A Camera Calibration Method Based on Nonlinear Model and Improved Planar Pattern

Zhang Yanqing Wang Zhiyan

School of Computer Science and Engineering, South China University of Technology,
Guangzhou, Guangdong, P.R.China, 510641

Abstract

In this paper, the topic on camera calibration is discussed. Firstly, lens distortion in nonlinear camera model is described in details. And then, according to perspective projection principle and effect of lens distortion, an improved planar pattern used for calibration is given. At last, a practical calibration method based on nonlinear model and improved planar pattern is put forward. In testing the proposed methods, satisfactory results are achieved in real data experiments.

Keywords: camera calibration, camera modeling, nonlinear distortion and planar pattern.

1. Introduction

Camera calibration is the process of determining the intrinsic and extrinsic parameters of a camera. It is a curial problem for further metric scene measurement in computer vision. Much work has been conducted in this area. Camera calibration is divided into two steps. The first step, called camera modeling, deals with the mathematical approximation of the physical and optical behavior of the sensor by using a set of parameters. The second step deals with the use of direct or iterative methods to estimate the values of these parameters. Most of the existing methods require either special motion of the camera or precise placement of the calibration target [1], which causes difficulties in practical use.

To overcome these limitations, Li and Chen [2,3] investigated the use of a single view for automatic recalibration of active vision systems in their recent study. Zhang [1] study the use of a planar pattern for automatic calibration of the intrinsic and extrinsic camera parameters. In this paper, an improved planar pattern is employed. The planar pattern contains a series of equilateral octagon. Therefore it is easy to make the planar pattern. Furthermore, no knowledge is required about the motion of the calibration target. In this paper, a nonlinear camera model and a refinement method are used to calibrate the camera. A real data experiments was conducted to verify the proposed method and satisfactory results were achieved. Experimental results show that this method has high potential for practical applications.

This paper is structured as follows. Section 2 deals with camera modeling and how the camera model is gradually obtained by a sequence of geometrical transformations. Section 3 describes the improved planar pattern used for camera calibration. Then, a practical calibration method based on nonlinear model and improved planar pattern is put forward in section 4. Finally, a testing experiment is carried out, and some results are obtained.

2. Camera Modeling

Camera model is a mathematical formulation that approximates the behavior of any physical device by using a set of mathematical equations. Camera modeling is based on approximating the internal geometry along with the position and orientation of the camera in the scene. The simplest are based on linear transformations without modeling the lens distortion. However, there are also some applications where greater precision is required, so nonlinear models which accurately model the lens are useful.

Camera modeling is usually broken down into 4 steps, as is hereafter detailed (see also Figure 1) [4].

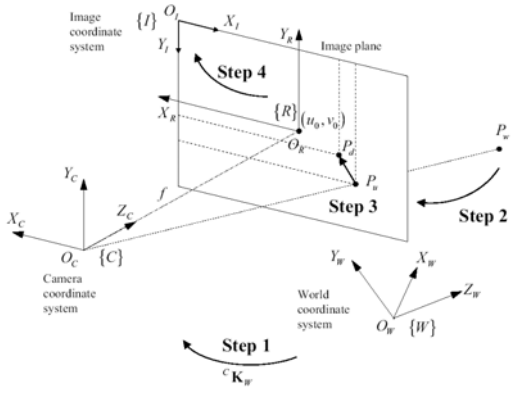

Fig. 1: The geometric relation between a 3D object point and its 2D image projection.

The first step consists of relating a point ${}^{W}p_w$ from the world coordinate system to the camera coordinate system, obtaining ${}^{C}p_w$. This transformation is performed by using a rotation matrix and a translation vector.

Next, it is necessary to carry out the projection of point ${}^{C}p_w$ on the image plane obtaining point ${}^{C}p_u$ by using a projective transformation.

The third step models the lens distortion, based on a disparity with the real projection. Then, point ${}^{C}p_u$ is transformed to the real projection of ${}^{C}p_d$ (which should coincide with the points captured by the camera).

Finally, the last step consists of carrying out another coordinate system transformation in order to change from the metric coordinate system of the camera to the image coordinate system of the computer in pixels, obtaining ${}^{I}p_d$.

2.1. Camera Position and Orientation

Changing the world coordinate system to the camera coordinate system is carried out by this transformation which is modeled using a translation vector and a rotation matrix, as shown in equation (1).

$$\begin{pmatrix} {}^{C}X_w \\ {}^{C}Y_w \\ {}^{C}Z_w \end{pmatrix} = {}^{C}R_w \begin{pmatrix} {}^{W}X_w \\ {}^{W}Y_w \\ {}^{W}Z_w \end{pmatrix} + {}^{C}T_w \tag{1}$$

Then, given a point ${}^{W}p_w$ related to the world coordinate system, and applying equation (1), the point ${}^{C}p_w$ in relation to the camera coordinate system is obtained.

2.2. Perspective Projection

Consider that any optical sensor can be modeled as a pinhole camera [4]. That is, the image plane is located at a distance f from the optical center O_C , and is parallel to the plane defined by the coordinate axis X_C and Y_C . Moreover, given an object point (${}^{C}p_w$) related to the camera coordinate system, if it is projected through the focal point (O_C), the optical ray intercepts the image plane at the 2D image point (${}^{C}p_u$). This relation is shown in equation (2).

$${}^{C}X_u = f\frac{{}^{C}X_w}{{}^{C}Z_w} \qquad {}^{C}Y_u = f\frac{{}^{C}Y_w}{{}^{C}Z_w} \tag{2}$$

2.3. Lens Distortion

Equations (3) transform the undistorted point ${}^{C}p_u$ to the distorted point ${}^{C}p_d$, where δ_x and δ_y represent the distortion involved.

$${}^{C}X_u = {}^{C}X_d + \delta_x \qquad {}^{C}Y_u = {}^{C}Y_d + \delta_y \tag{3}$$

The displacement given by the radial distortion dr can be modeled by equations (4), which consider only k_1 the first term of the radial distortion series. It has been proven that the first term of this series is sufficient to model the radial distortion in most of the applications [5].

$$\begin{aligned} \delta_{xr} &= k_1 {}^{C}X_d ({}^{C}X_d^2 + {}^{C}Y_d^2) \\ \delta_{yr} &= k_1 {}^{C}Y_d ({}^{C}X_d^2 + {}^{C}Y_d^2) \end{aligned} \tag{4}$$

The decentering distortion is due to the fact that the optical center of the lens is not correctly aligned with the center of the camera [6]. This type of distortion introduces a radial and tangential distortion [7], which can be described by the following equations,

$$\begin{aligned} \delta_{xd} &= p_1(3{}^{C}X_d^2 + {}^{C}Y_d^2) + 2p_2 {}^{C}X_d {}^{C}Y_d \\ \delta_{yd} &= 2p_1 {}^{C}X_d {}^{C}Y_d + p_2({}^{C}X_d^2 + 3{}^{C}Y_d^2) \end{aligned} \tag{5}$$

The thin prism distortion arises from imperfection in lens design and manufacturing as well as camera assembly. This type of distortion can be modeled by adding a thin prism to the optic system, causing radial and tangential distortions [6]. This distortion is modeled by,

$$\delta_{xp} = s_1({}^{C}X_d^2 + {}^{C}Y_d^2) \quad \delta_{yp} = s_2({}^{C}X_d^2 + {}^{C}Y_d^2) \tag{6}$$

The total distortion will be the sum of these three distortions. Usually, tangential distortion can be reduced by adopting high quality camera and lens, which has been by actual experiments. In addition, for most application, the accuracy obtained by those nonlinear models that only consider radial distortion is sufficient. At the same time, with the increasing of model complexity, the calibration algorithm will become less stability, and new error will be introduced. Therefore, radial distortion will be paid more attentions in this paper, that is,

$$\delta_x = \delta_{xr} \qquad \delta_y = \delta_{yr} \tag{7}$$

2.4. Computer Image Frame

This final step deals with expressing the ${}^{C}p_d$ point with respect to the computer image plane in pixels {I}. This transformation can be carried out by equation (8) as follows [5],

$$
\begin{aligned}
{}^{I}X_d &= -s_x d'^{-1}_x {}^{C}X_d + u_0 \\
{}^{I}Y_d &= -d_y^{-1} {}^{C}Y_d + v_0
\end{aligned} \tag{8}
$$

Where: (u_0, v_0) are the components of the principal point in pixels; s_x is the image scale factor; $d'_x = d_x \frac{N_{cx}}{N_{fx}}$; d_x is the center to center distance between adjacent sensor elements in the X direction; d_y is the center to center distance between adjacent sensor elements in the Y direction; N_{cx} is the number of sensor elements in the X direction; and N_{fx} is the number of pixels in an image row as sampled by computer.

3. Planar Pattern for Camera Calibration

The effect of radial distortion is the radial displacement of image point in distortion, not the tangential displacement of it. The main reason for this distortion is flawed radial curvature of lens. There are two trends of radial distortion: one is the distortion of image point deviating from the center, so that the rectangle frame is saddle-shaped; the other is the distortion of image point gathering towards the center, so that the rectangle frame is drum-shaped [8], as shown in figure 2.

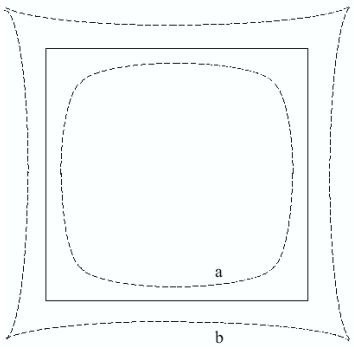

Fig. 2: Radial distortion effect (a: negative, b: positive)

The effect of tangential distortion is the tangential displacement of image point in distortion, not the radial displacement. The main reason for this distortion is decentering distortion and thin prism distortion. The trend of tangential distortion of the image is shown in figure 3.

Perspective projection has one important feature, that is, the straight line in three-dimensional space still has the quality of being straight after it turned into two-dimensional picture by perspective projection. Among all the images that are non-linearly distorted, the radial straight line from light center stays straight both before and after the distortion, which shows all non-radial straight lines will be distorted. The distortion is obvious in the mid point of the straight line. Therefore, during the course of camera calibration, the mid point of radial straight line not starting from light center should be paid more attention, and can be used in the improvement of traditional planar calibration pattern based on this. Improved planar pattern is shown in figure 4.

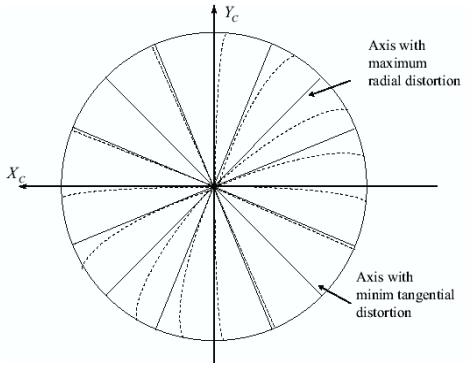

Fig. 3: Tangential distortion effect

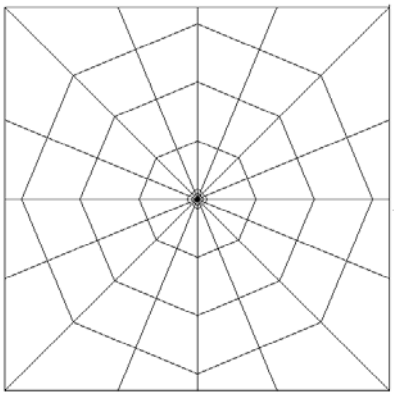

Fig. 4: Improved planar pattern

Figure 5 is the magnifying picture of improved pattern center; it is mainly used as target center that can keep the pattern center and the optical center of camera on a same line.

4. Calibration method based on nonlinear camera model and improved planar pattern

The calibrating method depends on the model used to approximate the behavior of the camera. The

linear models, the method in this paper uses a two-stage technique. As a first stage, a linear approximation with the aim of obtaining an initial guess is carried out and then a further iterative algorithm is used to optimize the parameters. In this section, calibrating method is explained by detailing the equations and the algorithm used to calibrate the camera parameters.

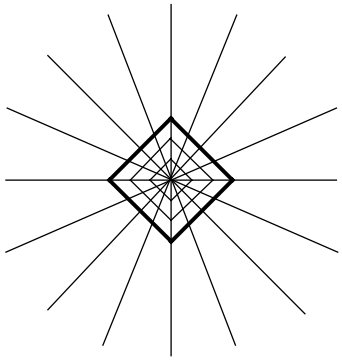

Fig. 5: The magnifying picture of improved pattern center

By combining equations (1), (2), (3), (4), and (7), the equations (9) are obtained.

$$
\begin{aligned}
{}^{C}X_d + {}^{C}X_d k_1 r^2 &= f\frac{r_{11}{}^{W}X_w + r_{12}{}^{W}Y_w + r_{13}{}^{W}Z_w + t_x}{r_{31}{}^{W}X_w + r_{32}{}^{W}Y_w + r_{33}{}^{W}Z_w + t_z} \\
{}^{C}Y_d + {}^{C}Y_d k_1 r^2 &= f\frac{r_{21}{}^{W}X_w + r_{22}{}^{W}Y_w + r_{23}{}^{W}Z_w + t_y}{r_{31}{}^{W}X_w + r_{32}{}^{W}Y_w + r_{33}{}^{W}Z_w + t_z} \\
r &= \sqrt{{}^{C}X_d{}^{2} + {}^{C}Y_d{}^{2}}
\end{aligned}
\tag{9}
$$

Once ${}^{C}X'_d$ and ${}^{C}Y'_d$ are obtained in metric coordinates by using equation (8), they can be expressed in pixels (${}^{I}X_d$ and ${}^{I}Y_d$) and the following equations are obtained.

$$
{}^{C}X'_{di} = -({}^{I}X_{di} - u_0)d'_x \quad {}^{C}Y'_{di} = -({}^{I}Y_{di} - v_0)d_y \tag{10}
$$

Where,

$$
{}^{C}X'_{di} = {}^{C}X_{di}s_x \qquad {}^{C}Y'_{di} = {}^{C}Y_{di} \tag{11}
$$

It is necessary to find a relationship between the image point P_d (in metric coordinates) with respect to the object point P_w. It can be observed that the segment $\overline{O_R P_d}$ is parallel to the segment $\overline{P_{oz}P_w}$, here, O_R is the center of image, P_{oz} is the origin of camera coordinate system. Considering this constraint, the following relationship is established.

$$
\overline{O_R P_d} \,//\, \overline{P_{oz}P_w} \Rightarrow \overline{O_R P_d} \times \overline{P_{oz}P_w} = 0 \tag{12}
$$

By using equation (12), the following equations are obtained.

$$
({}^{C}X_d, {}^{C}Y_d) \times ({}^{C}X_w, {}^{C}Y_w) = 0 \tag{13}
$$

$$
{}^{C}X_d\,{}^{C}Y_w - {}^{C}Y_d\,{}^{C}X_w = 0 \tag{14}
$$

Equation (14) can be arranged expressing the object point P_w with respect to the world coordinate system, instead of expressing it with respect to the camera coordinate system.

$$
\begin{aligned}
{}^{C}X_d(r_{21}{}^{W}X_w + r_{22}{}^{W}Y_w + r_{23}{}^{W}Z_w + t_y) = \\
{}^{C}Y_d(r_{11}{}^{W}X_w + r_{12}{}^{W}Y_w + r_{13}{}^{W}Z_w + t_x)
\end{aligned}
\tag{15}
$$

Operating equation (15) and arranging the terms,

$$
\begin{aligned}
{}^{C}X_d = {}^{C}Y_d{}^{W}X_w\frac{r_{11}}{t_y} + {}^{C}Y_d{}^{W}Y_w\frac{r_{12}}{t_y} + {}^{C}Y_d{}^{W}Z_w\frac{r_{13}}{t_y} + {}^{C}Y_d\frac{t_x}{t_y} \\
-{}^{C}X_d{}^{W}X_w\frac{r_{21}}{t_y} - {}^{C}X_d{}^{W}Y_w\frac{r_{22}}{t_y} + {}^{C}X_d{}^{W}Z_w\frac{r_{23}}{t_y}
\end{aligned}
\tag{16}
$$

In order to compute equation (16) for the n points obtained from equations (10), it is necessary to combine equation (16) with the equations (11), obtaining

$$
\begin{aligned}
{}^{C}X'_{di} = {}^{C}Y'_{di}{}^{W}X_{wi}\frac{s_x r_{11}}{t_y} + {}^{C}Y'_{di}{}^{W}Y_{wi}\frac{s_x r_{12}}{t_y} + {}^{C}Y'_{di}{}^{W}Z_{wi}\frac{s_x r_{13}}{t_y} + {}^{C}Y'_{di}\frac{s_x t_x}{t_y} \\
-{}^{C}X'_{di}{}^{W}X_{wi}\frac{r_{21}}{t_y} - {}^{C}X'_{di}{}^{W}Y_{wi}\frac{r_{22}}{t_y} + {}^{C}X'_{di}{}^{W}Z_{wi}\frac{r_{23}}{t_y}
\end{aligned}
\tag{17}
$$

At this point, a system with n equations and 7 unknowns is obtained, which can be expressed in the following form,

$$
\begin{pmatrix}
{}^{C}Y'_{di}{}^{W}X_{wi} \\
{}^{C}Y'_{di}{}^{W}Y_{wi} \\
{}^{C}Y'_{di}{}^{W}Z_{wi} \\
{}^{C}Y'_{di} \\
-{}^{C}X'_{di}{}^{W}X_{wi} \\
-{}^{C}X'_{di}{}^{W}Y_{wi} \\
-{}^{C}X'_{di}{}^{W}Z_{wi}
\end{pmatrix}^{T}
\begin{pmatrix}
t_y^{-1}s_x r_{11} \\
t_y^{-1}s_x r_{12} \\
t_y^{-1}s_x r_{13} \\
t_y^{-1}s_x t_x \\
t_y^{-1}s_x r_{21} \\
t_y^{-1}s_x r_{22} \\
t_y^{-1}s_x r_{23}
\end{pmatrix}
= {}^{C}X'_{di}
\tag{18}
$$

In order to simplify the notation, the 7 unknowns components of the vector renamed.

$$
\begin{aligned}
a_1 &= t_y^{-1}s_x r_{11} \qquad a_5 = t_y^{-1}r_{21} \\
a_2 &= t_y^{-1}s_x r_{12} \qquad a_6 = t_y^{-1}s_x r_{22} \\
a_3 &= t_y^{-1}s_x r_{13} \qquad a_7 = t_y^{-1}s_x r_{23} \\
a_4 &= t_y^{-1}s_x t_x
\end{aligned}
\tag{19}
$$

Note that the a_i components can be easily computed by using a least-squares technique. Therefore, the point of interest is to extract the calibrating parameters of the camera from these a_i components. First t_y can be obtained by using equations (19) in the following manner,

$$t_y = \frac{\|r_2\|}{\|a_{5,6,7}\|} \tag{20}$$

and equation (20) is simplified because the norm of the vector r_2 is equal to the unity, obtaining the parameter t_y.

$$|t_y| = \frac{1}{\sqrt{a_5^{\,2} + a_6^{\,2} + a_7^{\,2}}} \tag{21}$$

However, equation (21) is insufficient since it does not provide the sign of the t_y component. In order to determine this sign, a point $({}^{I}X_d, {}^{I}Y_d)$ located at the periphery of the image, far from the center, is taken from the set of test points (its corresponding 3D point is also kept). It is then supposed that the t_y sign is positive, and the following equations are computed.

$$\begin{aligned} r_{11} &= a_1 t_y / s_x \quad & r_{21} &= a_5 t_y \\ r_{12} &= a_2 t_y / s_x \quad & r_{22} &= a_6 t_y \\ r_{13} &= a_3 t_y / s_x \quad & r_{23} &= a_7 t_y \\ t_x &= a_4 t_y \end{aligned} \tag{22}$$

By using the corresponding 3D point ${}^{W}X_w, {}^{W}Y_w, {}^{W}Z_w$, the linear projection of this 3D point on the image plane (without considering lens distortion) can be computed by using equations (23).

$$\begin{aligned} {}^{C}X_u &= r_{11}\,{}^{W}X_w + r_{12}\,{}^{W}Y_w + r_{13}\,{}^{W}Z_w + t_x \\ {}^{C}Y_u &= r_{21}\,{}^{W}X_w + r_{22}\,{}^{W}Y_w + r_{23}\,{}^{W}Z_w + t_y \end{aligned} \tag{23}$$

At this point, the t_y sign can be verified. If both components of the point $({}^{C}X_u, {}^{C}Y_u)$ have a sign equal to the components of the point $({}^{I}X_d, {}^{I}Y_d)$, it means that the t_y sign was correctly chosen as positive. Otherwise, it has to be considered negative.

The second parameter to be extracted is the scale factor (as). Note that by arranging equations (19), the following equation is obtained,

$$s_x = \frac{\|a_{1,2,3}\| t_y}{\|r_1\|} \tag{24}$$

where it is known that the norm of r_1 is the unity and the scale factor is always positive. Then, s_x is obtained by using equation (25).

$$s_x = \sqrt{a_1^{\,2} + a_2^{\,2} + a_3^{\,2}}\,|t_y| \tag{25}$$

Furthermore, the 2D points, with respect to the camera coordinate system $({}^{C}X_d, {}^{C}Y_d)$, can be computed from the same point with respect to the image coordinate system, that is $({}^{I}X_d, {}^{I}Y_d)$, by using equations (11). Moreover, by using equations (22), the r_1 and r_2 vectors of the rotation matrix ${}^{C}R_W$, and the first element of the translation vector ${}^{C}T_W$, i.e. t_x, can be calculated. Finally, the third orientation vector (r_3) can be computed by a cross product between r_1 and r_2 because of the property of orthogonality, (note also that the determinant of any rotation matrix is the unity, i.e. $|{}^{C}R_W| = 1$). At this point, the first three steps of the method are completed, see Figure 6.

However, the following parameters are still unknown: the focal distance (f), the radial lens distortion coefficient (k_1), and the translation of the camera with respect to the Z axis (t_z). In order to compute these three parameters, a linear approximation is first used without considering the k_1 parameter. The linear approximation is shown in equation (26), which was obtained from equations (9).

$$\begin{aligned} &\left(r_{21}\,{}^{W}X_{wi} + r_{22}\,{}^{W}Y_{wi} + r_{23}\,{}^{W}Z_{wi} + t_y - {}^{C}Y_d\right)\begin{pmatrix} f \\ t_z \end{pmatrix} = \\ &\qquad \left(r_{31}\,{}^{W}X_{wi} + r_{32}\,{}^{W}Y_{wi} + r_{33}\,{}^{W}Z_{wi}\right){}^{C}Y_d \end{aligned} \tag{26}$$

Equation (26) has now been applied to the whole set of test points, obtaining a system of n equations and two unknowns. The linear approximation of both unknowns, f and t_z, is obtained by using a pseudo-inverse. However, in order to calculate a better approximation including the k_1 parameter, it is necessary to iterate equations (9) by using an optimization method considering the linear method with k_1=0 as an initial solution.

Finally, all the parameters are optimized iteratively with the aim of obtaining an accurate solution. The entire process is explained in Figure 6.

5. Experiment results

This chapter presents the experimental results obtained from the implementation of the method described in the previous sections. The calibration steps are the following:

(1) Place the camera to be calibrated in a location where the entire working area is viewed by the given camera.

(2) Set a world coordinate system in a fixed position with respect to the camera.

(3) Place a set of 3D object points spread throughout the working area. In experiment, the mid point of each edge of octagon is selected as object point. In order to obtain more accurate results, at least twenty-four object points together with their 2D image

projections are well distributed throughout the image plane.

(4) Capture an image of the working area and measure the 2D projections of each one of the object points.

(5) Solve the correspondence between the 3D object points with their 2D image projections

(6) Calibrate the camera.

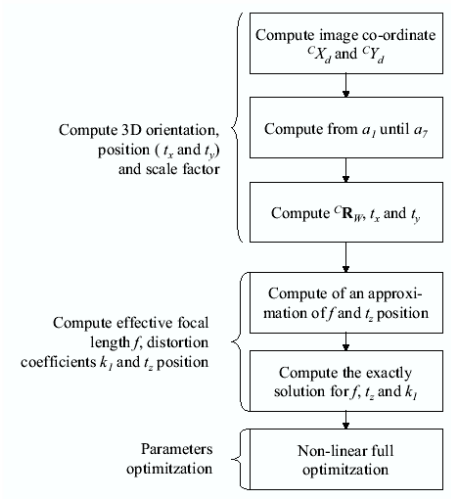

Fig. 6: Flowchart of the calibration method

Camera calibration results are given in table 1.

Table 1. Camera calibration experiment results

R11=-0.173510	R12=-0.983689	R13=0.047409
R21=-0.796698	R22=0.111903	R23=-0.593929
R31=0.578937	R32=-0.140818	R33=-0.803117
Tx=40.123047	Ty=-56.816054	Tz=974.907012
f=1200.34	K1=4.76497e-8	

By comparing computational value of 3D object points with actual value, the accuracy evaluation of experiment results is present as follows,

Emin=0.110096; Emax=1.480197;

Sigma=0.69183.

Here, Emin denotes the minimal value of calibration error, Emax denotes the maximal value of calibration error, sigma is even value of mean root square error.

6. Conclusion

An improved planar pattern for camera calibration is presented. In addition to the normal parameters in a pinhole model, nonlinear lens distortion can also be taken into account by the calibration method in this paper. Experiments on real data were conducted to verify the validity of proposed method. The satisfactory results obtained demonstrated that this method provides a correct and practical solution for camera calibration. This method offers the advantage of ease of use in practical applications.

7. Acknowledgment

This paper is supported by project B6-109-497 of Provincial Nature Science Foundation of Guangdong, China.

8. References

[1] Beiwei Zhang and Fu-Chao Wu, “Planar pattern for automatic camera calibration,” Opt. Eng. 42(6), June, 2003,pp. 1542–1549.

[2] Y. F. Li and S. Chen, ‘‘Automatic recalibration of an active structured light vision system,’’ IEEE Trans. Robot. Automat. in press.

[3] S. Chen and Y. F. Li, ‘‘Self-recalibration of a colour-encoded light system for automated 3-D measurements,’’ Meas. Sci. Technol. in press.

[4] Xavier Armangue Quintana, Modelling Stereoscopic Vision Systems for Robotic Applications, Ph.D dissertation, University de Girona, July 2003, pp.7-41.

[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 International Journal on Robotics and Automation, vol. RA-3, no. 4, August 1987, pp 323-444.

[6] J. Weng, P. Cohen and M. Herniou, “Camera Calibration with Distortion Models and Accuracy Evaluation.” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 14, no. 10, October 1992, pp 965-980.

[7] C. C. Slama, C. Thenrer and S. W. Henriksen. Manual of photogrammetry. American Society of Photogrammetry, Falls Church, VA, 4 edition, 1980.

[8] Jiang Dazhi, Research and Application on 3D Reconstruction of Computer Vision, Ph.D dissertation, Nanjing University of Aeronautics and Astronautics, PRC, October 2001, pp.56-61.