Camera calibration approach using circle-square-combined target

Fuqiang Zhou (周富强)[1*], Yexin Wang (王晔昕)[1], Yi Cui (崔 毅)[1], and Haishu Tan (谭海曙)[2]

[1] *School of Instrumentation Science and Optoelectronics Engineering, Beihang University, Beijing 100191, China*

[2] *Department of Electronic Information Engineering, Foshan University, Foshan 528000, China*

[*] *Corresponding author: zfq@buaa.edu.cn*

Received April 15, 2011; accepted July 8, 2011; posted online September 22, 2011

Calibrating a small field camera is a challenging task because the traditional target with visible feature points that fit the limited space is difficult and costly to manufacture. We demonstrate a novel combined target used in camera calibration. The tangent points supplied by one circle located at the center of a square are used as invisible features, and the perspective projection invariance is proved. Both visible and invisible features extracted by the proposed feature extraction algorithm are used to solve the calibration. The target supplies a sufficient number of feature points to satisfy the requirements of calibration within a limited space. Experiments show that the approach can achieve high robustness and considerable accuracy. This approach has potential for computer vision applications particularly in small fields of view.

OCIS codes: 100.2000, 150.1488.

doi: 10.3788/COL201210.021003.

The demand for machine vision applications in small field of view (FOV), such as three-dimensional (3D) teeth reconstruction and machine parts detection and measurement etc., has increased in recent years. Consequently, significant attention has been devoted to this topic. Camera calibration is the primary task of these applications. It determines the mapping between the 3D world reference frame and the image plane based on the camera model[1,2]. In the calibration process, target plays a crucial role in determining the internal physical camera parameters (intrinsic parameters) as well as the 3D position and orientation of the camera relative to the world reference frame (extrinsic parameters).

According to the visual properties of the feature points, calibration target can be classified into two groups, namely, target using visible feature points (VFPs) and target using invisible feature points (IFPs). VFPs refer to all the feature points that can be observed and extracted from the images directly such as corners and intersections. On the contrary, IFP uses feature points that are hidden on the target such as circle centers, tangent points, and vanishing points which require feature extraction and further geometric calculation[3]. Generally, current calibration approaches use targets from one of the two types, two-dimensional (2D) square matrix and chessboard[2,4,5] targets. These two types of target are used widely in calibration because of the simplicity of their feature extraction and high calibration precision. On the other hand, due to its better manufactured accuracy and high robustness against random noise, the circle features, such as circle matrix, concentric circle, and their matrices, have always been adopted for IFP targets[6−13].

For a small FOV calibration task, 2D target is the best choice because a tiny 3D target would be difficult and costly to manufacuture and a one-dimensional (1D) target would require a large space[14]. However, the conventional 2D targets with square or circle matrices are difficult and costly to manufacture in a tiny size, and one square or circle cannot supply the sufficient number of points.

Figure 1 shows a novel circle-square-combined target. This target combines the two categories of targets mentioned above and uses square corners (points 1-4) as VFP and tangent points (points 5-12) from the corners to the circle as IFP. With this layout, the target is able to provide sufficient feature points within a limited space, which satisfies camera calibration task in small FOV.

The precondition of using this kind of tangent points in calibration process is that they should be invariant in perspective projection. Perspective projection is not a shape preserving transformation, hence, a circle is transformed into an ellipse[6]. The projection remains circular only when the object surface and the image plane are parallel[15].

Figure 2 shows the perspective projection of a circle. Let the world coordinate system Ω_1:$o - x_{\rm w}y_{\rm w}z_{\rm w}$ be centered at the camera optical center, and let $z_{\rm w}$ axis be perpendicular to the object plane π_1 where circle $\rm C_1$ is located. The homogeneous coordinates of a point on circle $\rm C_1$ are $\mathbf{x}_{\rm w}$= $(x_{\rm w},\ y_{\rm w},\ z_{\rm w},\ 1)^{\rm T}$, l_1 is the tangent line from arbitrary point Q_1 outside the circle on the plane, and P_1 is the tangent point. The equation of the circle in the world coordinate system is

$$\mathbf{x}_{\rm w}^{\rm T}\mathbf{C}_1\mathbf{x}_{\rm w} = 0, \tag{1}$$

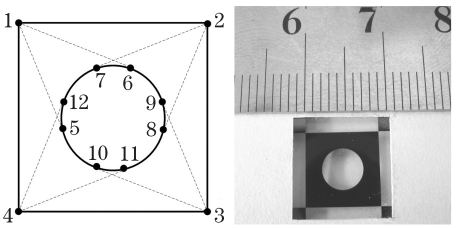

Fig. 1. Feature points in a novel combined target.

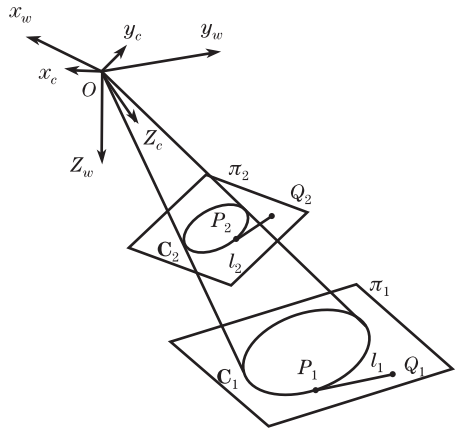

Fig. 2. Perspective projection of a circle.

where $\mathbf{C}_1$ is the parameter matrix of the circle. Let the camera coordinate system Ω_2: $O - x_\mathrm{c}y_\mathrm{c}z_\mathrm{c}$ be centered at the camera optical point, and let z_c axis be perpendicular to the image plane π_2. Note that the perspective projection of the circle C_1 is ellipse C_2. The homogenous coordinates of a point on ellipse C_2 are $\mathbf{x}_\mathrm{c}$= (x_c, y_c, z_c, 1)$^\mathrm{T}$. The transformation from world to camera coordinate system is described as

$$\mathbf{x}_\mathrm{c} = \mathbf{A}\mathbf{x}_\mathrm{w}, \tag{2}$$

where $\mathbf{A}$ is the tranformation matrix from Ω_1 to Ω_2. The tangent line l_1 after perspective projection is line l_2 on π_2 (perpendicular projection is excluded in the discussion). Note that P_1 and Q_1 in $\mathbf{\Omega}_1$ are transformed into P_2 and Q_2 in $\mathbf{\Omega}_2$, with their coordinates marked as $\mathbf{x}_{\mathrm{w}P1}$, $\mathbf{x}_{\mathrm{w}Q1}$, $\mathbf{x}_{\mathrm{c}P2}$, and $\mathbf{x}_{\mathrm{c}Q2}$. The relations of the four points are

$$\mathbf{x}_{\mathrm{c}P2} = \mathbf{A}\cdot\mathbf{x}_{\mathrm{w}P1} \quad \text{and} \quad \mathbf{x}_{\mathrm{c}Q2} = \mathbf{A}\cdot\mathbf{x}_{\mathrm{w}Q1}, \tag{3}$$

where matrix $\mathbf{A}$ is non-singular and consists of standard orthogonal bases. From Eqs. (1) and (2), we have

$$\mathbf{x}_\mathrm{c}^\mathrm{T}\left(\mathbf{A}^{-\mathrm{T}}\mathbf{C}_1\mathbf{A}^{-1}\right)\mathbf{x}_\mathrm{c} = 0. \tag{4}$$

Note that $\mathbf{C}_2 = \mathbf{A}^{-\mathrm{T}}\mathbf{C}_1\mathbf{A}^{-1}$ is the paramter matrix of ellipse C_2 which is transformed from C_1. The point P_1 is transformed into P_2 on C_2. Note that P_2 and Q_2 are on line l_2. Assuming l' is the tangent line of C_2 on point P_2 in Ω_2 coordinate system, then it can be expressed as

$$l' = \mathbf{C}_2\cdot\mathbf{x}_{\mathrm{cp}2} = \mathbf{A}^{-\mathrm{T}}\mathbf{C}_1\mathbf{A}^{-1}\cdot\mathbf{x}_{\mathrm{cp}2}. \tag{5}$$

From Eqs. (4) and (5), we obtain

$$\begin{aligned}\mathbf{x}_{\mathrm{c}Q2}^\mathrm{T}\cdot l' &= (\mathbf{A}\cdot\mathbf{x}_{\mathrm{w}Q1})^\mathrm{T}\cdot(\mathbf{A}^{-\mathrm{T}}\mathbf{C}_1\mathbf{A}^{-1}\cdot\mathbf{x}_{\mathrm{cp}2})\\ &= \mathbf{x}_{\mathrm{w}Q1}^\mathrm{T}\cdot\mathbf{C}_1\cdot\mathbf{x}_{\mathrm{w}P1} = \mathbf{x}_{\mathrm{w}Q1}^\mathrm{T}\cdot l_1 = 0.\end{aligned} \tag{6}$$

It shows that point Q_2 is also on line l', indicating that l' and l_2 are the same line. Therefore, these points are justified inviarant during perspective projection.

The calibration steps for the small FOV proposed in this letter are illustrated in Fig. 3. The rectangle marked with dotted lines is the illustration of the feature extraction algorithm of a combined target. The specific steps are as follows.

1) The image coordinates of the corner points are extracted by applying the sub-pixel Harris method and saved according to the sequence shown in Fig. 1. The sequences of the corner points are derived easily by contrasting their image coordinates.

2) Adaptive threshold method is used to obtain binary images. Subsequently, the ellipse contour is selected by filtering unclosed counters and using a given length threshold.

3) Using least square method to fit the ellipse contour obtained from step 2), the ellipse parameters are determined, including image coordinates of the ellipse center, lengths of the major and minor axes, and angle of the major axis.

4) Let the ellipse coordinate system $O_\mathrm{e} - x_\mathrm{e}y_\mathrm{e}$ be centered at the ellipse center and let its x_e axis be the ellipse major axis. Solve the coordinates of tangent points using the simultaneous equations combined by ellipse equation and its tangent chord equation in O_e- $x_\mathrm{e}y_\mathrm{e}$. Transform the points back into the image coordinate system and save them according to the sequence shown in Fig. 1.

Figure 4 introduces the pinhole camera model used in the calibration. The world and camera image coordinate systems are $O_\mathrm{w} - x_\mathrm{w}y_\mathrm{w}z_\mathrm{w}$, $O_\mathrm{c} - x_\mathrm{c}y_\mathrm{c}z_\mathrm{c}$, and $O_\mathrm{u} - x_\mathrm{u}y_\mathrm{u}$ respectively. Define $O_\mathrm{c}x_\mathrm{c}//O_\mathrm{u}x_\mathrm{u}$, $O_\mathrm{c}y_\mathrm{c}//O_\mathrm{u}y_\mathrm{u}$, and $O_\mathrm{c}z_\mathrm{c}$ as perpendicular to the image plane. Point $\mathbf{P}$, with homogeneous coordinates $\widetilde{\mathbf{P}}_\mathrm{w}$= ($x_\mathrm{w}$, y_w, z_w, 1)$^\mathrm{T}$ in the 3D world coordinate system is transformed into $\widetilde{\mathbf{p}}_\mathrm{u}$= ($x_\mathrm{u}$, y_u, 1)$^\mathrm{T}$ in the image coordinate system. The camera model can be described as

$$s\widetilde{\mathbf{p}}_\mathrm{u} = \mathbf{A}\left[\begin{array}{cc}\mathbf{R} & \mathbf{T}\end{array}\right]\widetilde{\mathbf{P}}_\mathrm{w} = \mathbf{H}\widetilde{\mathbf{P}}_\mathrm{w}, \quad \mathbf{A} = \left[\begin{array}{ccc} f_x & 0 & u_0\\ 0 & f_y & v_0\\ 0 & 0 & 1\end{array}\right], \tag{7}$$

where s is a scale factor, 3×3 $\mathbf{R}$ and 3×1 $\mathbf{T}$ are the rotation matrix and translation vector, respectively, $\mathbf{A}$ is the intrinsic parameters matrix, where f_x and f_y represent the focal length in terms of pixel in the x and y directions, respectively, (u_0, v_0) are the image coordinates of the camera principal point, and 3×4 $\mathbf{H}$ is the mapping

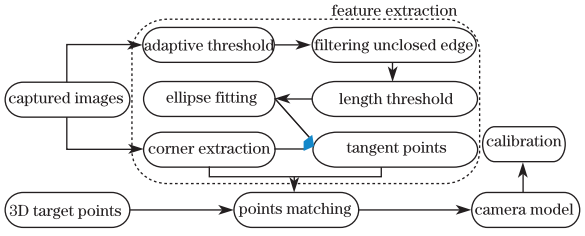

Fig. 3. Calibration using combined target.

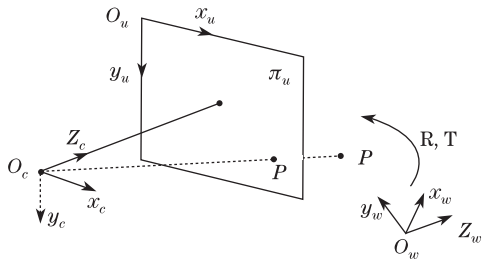

Fig. 4. Pinhole camera model.

matrix from the 3D world to the 2D image coordinate system. According to Eq. (6), at least six points solve the problem because every point provides two independent equations.

The calibration approach using a combined target is described below according to the camera model and the flow chart in Fig. 3.

1) The 2D combined target is composed of a square and a circle, which are homocentric with inverse colors, and the diameter of the circle is shorter than the length of the square.

2) Let $O_{\rm w} - x_{\rm w}y_{\rm w}z_{\rm w}$ represent the world coordinate system centered at one corner of the target, and plane $x_{\rm w}o_{\rm w}y_{\rm w}$ represent the target plane. The world coordinates of the 12 feature points can be calculated easily.

3) With the square and circle in the FOV, move the combined target at least twice and capture an image by each move.

4) The image coordinates of all the feature points in each image can be extracted with the algorithm presented previously.

5) With both the world and image coordinates of the feature points, the mapping matrix of the projection can be calculated. An initial guess of the camera parameters is acquired based on the orthogonal property of matrix $\mathbf{R}$[2]. Subsequently, the camera parameters are optimized using the nonlinear Levenberg-Marquardt algorithm with

$$\min \sum_i \sum_j ||\widehat{\mathbf{X}}^u_{ij} - \widehat{\mathbf{X}}^u_{ij}(f_x, f_y, u_0, v_0)||, \qquad (8)$$

where $\mathbf{X}^u_{\rm ij}$ represents the extracted image coordinates of the ith point on the jth image and $\widehat{\mathbf{X}}^u_{\rm ij}$ is calculated by the calibrated camera parameters based on the model. However, because lens distortion is not discussed in the model, this calibration procedure works only for lens with little distortion or for images corrected by distortion separated calibration method.

In order to simulate the ellipse on the combined target, 30 equal step-points on the circle edge were used. The 30 points and 4 corner points in 3D world were projected onto 10 virtual images using the parameters shown in Table 1 (intrinsic parameters) and Table 2 (extrinsic parameters). Thirty projected points on each virtual image were used to fit the ellipse and calculate the image coordinates of tangent points.

The tangent points from the corners to the ellipse are IFP, thus, in order to simulate the noise effects in ellipse extraction and its influence on computing tangent points, Gaussian noise (location deviation) with 0 mean value and standard deviations from 0.1 to 1.0 with increasing pace 0.1 (pixel) was added to the points. After ellipse extraction as well as fitting and calculation of tangent points, the calibration process was implemented. Fifty independent experiments were conducted to avoid randomness. The average results are shown in Fig. 5. To evaluate the calibration results, the root mean square (RMS) re-projection and normalized calibration errors (NCE)[16] are defined as

$$\mathrm{RMS} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left[\left(x_{(\mathrm{u}i)} - \hat{x}_{(ui)}\right)^2 + \left(y_{\mathrm{u}i} - \hat{y}_{\mathrm{u}i}\right)^2\right]}, \qquad (9)$$

$$\mathrm{NCE} = \frac{1}{n}\sum_{i=1}^{n}\left[\frac{\left(\hat{x}_{\mathrm{c}i} - x_{\mathrm{c}i}\right)^2 + \left(\hat{y}_{\mathrm{c}i} - y_{\mathrm{c}i}\right)^2}{\hat{z}^2_{\mathrm{c}i}\left(f_x^{-2} + f_y^{-2}\right)/12}\right]^{1/2}, \qquad (10)$$

where $(x_{\mathrm{u}i}, y_{\mathrm{u}i})$ represent the extracted image coordinates of the ith point and $(\hat{x}_{\mathrm{u}i}, \hat{y}_{\mathrm{u}i})$ are calculated by the calibrated camera parameters based on the model. In Eq. (10), $(x_{\mathrm{c}i}, y_{\mathrm{c}i}, z_{\mathrm{c}i})$ represent the true coordinates of the ith point in the camera coordinate system and $(\hat{x}_{\mathrm{c}i}, \hat{y}_{\mathrm{c}i}, \hat{z}_{\mathrm{c}i})$are evaluated as the intersection of the back-projection line of the sensed point and the plane $z = z_{\mathrm{c}i}$.

In Fig. 5(a), when the noise deviation for the points on the ellipse is less than a pixel, the relative error of camera intrisic parameters f_x, and f_y satisfies the calibration accuracy significantly. The RMS and NCE errors are shown in Figs. (b) and (c). Both demonstrate that this approach has high accuracy. When the ellipse point location error rises to one pixel, the RMS error of the proposed approach is less than 0.16 pixel, whereas the NCE error is less than 0.4 pixel. Adoption by the ellipse extraction and fitting of a considerable number of contour points which decrease the noise influence

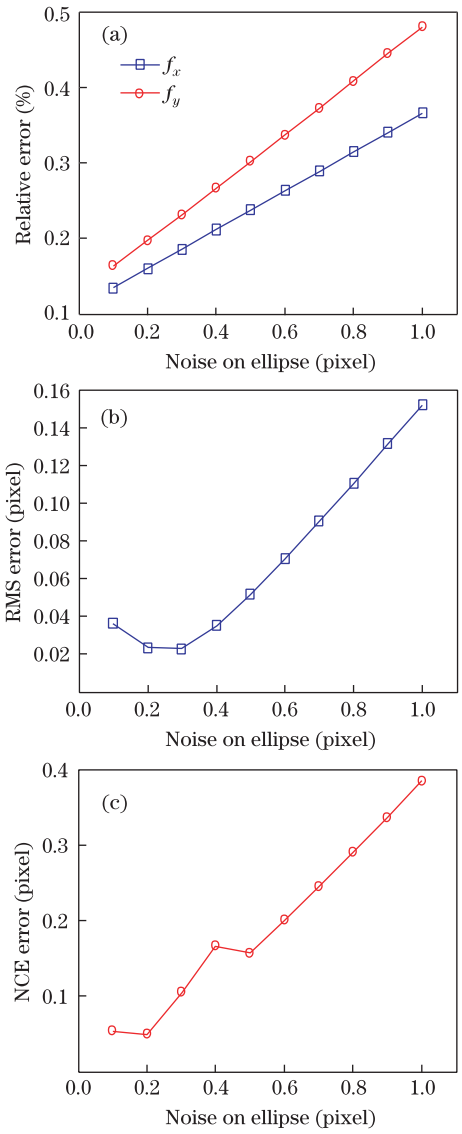

Fig. 5. Calibration results of different noise levels on ellipse points. (a) Relative error; (b) RMS error; (c) NCE error.

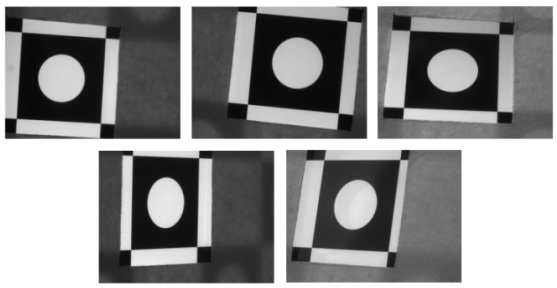

Fig. 6. Five images captured by camera in small FOV.

account for these results. Moreover, these guarantee the precision of the invisible tangent points and high accuracy of the calibration results.

In the actual experiment shown in Fig. 6, five images of the combined target were captured by a camera in small FOV with a resolution of 800 × 600. The size of one pixel on CCD is 4.4 × 4.4 (μm). The length of the square is 9 mm and the diameter of the circle is 5 mm. After the feature extraction and calibration procedure mentioned above have been completed, the calibration results are derived and shown in Table 3.

The RMS and NCE errors are low, indicating that calibration using combined target in small FOV is capable of good accuracy. Compared with the commonly used chessboard target, the combined target has advantage of robustness against noise and manufacturing cost. The calibration approach also uses both the VFP and IFP, which supply a sufficient number of feature points to satisfy the calibration in a limited space.

Table 1. Accurate Camera Intrinsic Parameters Set for Simulation

Parameter	f_x	f_y	u_0	v_0
Value (pixel)	1 855	1 855	400	300

Table 2. Accurate Camera Extrinsic Parameters Set for Simulation

Image No.	R			T		
	r_1	r_2	r_3	t_1	t_2	t_3
2	–1.71	–2.16	0.73	–10	–45	477
3	2.07	2.21	–0.27	–34	–43	495
4	1.89	2.31	0.44	–34	–62	476
5	1.85	1.82	–0.49	–40	–46	527
6	–1.40	–2.46	0.78	–1	–50	490
7	1.76	2.01	0.17	–30	–58	518
8	–1.96	–2.30	–0.80	–29	–47	491
9	2.06	2.23	0.35	–33	–45	483
10	–1.81	–2.33	0.78	–7	–39	523

Table 3. Calibration Results Using Combined Target

Parameter	f_x	f_y	u_0	v_0	RMS	NCE
Value (pixel)	4 234.92	4 234.12	399.5	299.5	0.110 4	0.223 2

In conclusion, a camera calibration approach with a new combined target is developed. The perspective projection invariance of the tangent points generated by the target is proved. Both visible and invisible feature points are extracted from the captured images to calibrate the intrinsic and extrinsic parameters of the camera. The experimental results, together with simulation and experiment data, show that the proposed approach is effective and robust. Compared with classical approaches generally used, the proposed approach provides a way to calibrate a camera in small FOV with high accuracy.

This work was supported by the National Natural Science Foundation of China (No. 60972086) and the Research Fund for the Doctoral Program of Higher Education of China (No. 20101102110033).

References

1. J. Salvi, X. Armangue, and J. Batlle, Pattern Recogn. **35,** 1617 (2002).
2. Z. Zhang, IEEE Trans. Pattern Anal. Mach. Intell. **22,** 1330 (2000).
3. Z. Chen and P. Shi, Chin. Opt. Lett. **3,** 69 (2005).
4. M. Fiala and C. Shu, Mach. Vis. Appl. **19,** 209 (2008).
5. L. Zhang, J. Zhang, and Y. Hu, in *Proceedings of International Conference on Information and Automation* 352 (2009).
6. J. Heikkila and O. SilvCn, in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognitions* 1106 (1997).
7. W. Ma, Y. Zheng, and Y. Liu, Opt. Eng. **5,** 053602 (2009).
8. F. Zhou and W. Zhang, Proc. SPIE **7129,** 71290C (2008).
9. J. Kim, P. Gurdjos, and I. Kweon, IEEE Trans. Pattern Anal. Mach. Intell. **4,** 637 (2005).
10. Z. Yang, F. Chen, J. Zhao, and H. Zhao, in *Proceedings of the Third IEEE Conference on Industrial Electronics and Applications* 2222 (2008).
11. X. Meng and Z. Hu, Pattern Recogn. **36,** 1155 (2003).
12. C. Sun, X. Zhang, and Y. Qu, Chin. Opt. Lett. **3,** 585 (2005).
13. H. Cui, W. Liao, X. Cheng, N. Dai, and T. Yuan, Chin. Opt. Lett. **8,** 33 (2010).
14. I. Miyagawa, H. Arai, and H. Koike, IEEE Trans. Image Process. **6,** 1528 (2010).
15. H. Richard and Z. Andrew, *Multiple View Geometry in Computer Vision* (Cambridge University Press, Cambridge, 2000).
16. J. Weng, P. Cohen, and M. Herniou, IEEE Trans. Pattern Anal. Mach. Intell. **10,** 965 (1992).