

# Optical Engineering

SPIEDigitalLibrary.org/oe

## Advanced geometric camera calibration for machine vision

Minh Vo  
Zhaoyang Wang  
Long Luu  
Jun Ma



# Advanced geometric camera calibration for machine vision

Minh Vo,<sup>a,b</sup> Zhaoyang Wang,<sup>a</sup> Long Luu,<sup>a,b</sup> and Jun Ma<sup>a</sup>

<sup>a</sup>The Catholic University of America, Department of Mechanical Engineering, Washington, DC 20064

<sup>b</sup>The Catholic University of America, Department of Electrical Engineering, Washington, DC 20064  
E-mail: wangz@cua.edu

**Abstract.** In many machine vision applications, a crucial step is to accurately determine the relation between the image of the object and its physical dimension by performing a calibration process. Over time, various calibration techniques have been developed. Nevertheless, the existing methods cannot satisfy the ever-increasing demands for higher accuracy performance. In this letter, an advanced geometric camera calibration technique which employs a frontal image concept and a hyper-precise control point detection scheme with digital image correlation is presented. Simulation and real experimental results have successfully demonstrated the superior of the proposed technique. © 2011 Society of Photo-Optical Instrumentation Engineers (SPIE). [DOI: 10.1117/1.3647521]

Subject terms: camera calibration; machine vision; frontal image; digital image correlation.

Paper 110686LRR received Jun. 17, 2011; revised manuscript received Sep. 13, 2011; accepted for publication Sep. 19, 2011; published online Nov. 3, 2011.

The recent decade has seen tremendous growth in the machine vision applications as they become the new driving forces in precision engineering. Many of these applications highly benefit from the improvements of camera calibration techniques, which primarily fall into two categories: photogrammetry calibration<sup>1</sup> and self-calibration.<sup>2</sup> Compared with the latter which does not use any specific calibration target, the former usually provides higher accuracy by using a precisely-constructed calibration target and is therefore the interest of this letter. A representative photogrammetry approach is Zhang's method, which uses a planar pattern as the calibration target and is well known for its flexibility, robustness, and low cost.<sup>1</sup> This method and its extensions implemented in the OpenCV library are considered as the conventional technique in this letter.

The literature in the relevant field addresses two major sources of error that affect the camera calibration results. The first one is the imperfection of the calibration target. Since the assumptions made for the conventional camera calibration are based on a perfect planar target with ideal patterns, the imprecision of the calibration target may lead to inaccurate results.<sup>3</sup> The second problem is the uncertainty in locating the control points directly from the geometries of the calibration patterns in the captured raw images which suffers from lens distortion as well as perspective distortion.<sup>4</sup>

This letter proposes an advanced geometric camera calibration approach to overcome the aforementioned two problems without loss of the original advantages of the conventional technique. The proposed technique uses a sophisticated lens distortion model that takes the radial, tangential, and prism distortion into account, and achieves a precise localization of the control points with a novel refinement process using a frontal image concept and an advanced digital image correlation (DIC) scheme; in addition, the imperfection of the calibration target can be compensated.

For the ideal pinhole model, assuming that the world coordinate is placed on the calibration target with its surface as the XY plane, the relation between the three-dimensional world coordinate of a calibration target point  $\mathbf{M} = [X_w, Y_w, Z_w]^T$  and its corresponding location  $\mathbf{m} = [u, v]^T$  in the image plane can be expressed as:

$$s \begin{Bmatrix} \mathbf{m} \\ 1 \end{Bmatrix} = \mathbf{A} [\mathbf{R} \ \mathbf{T}] \begin{Bmatrix} \mathbf{M} \\ 1 \end{Bmatrix}, \quad \mathbf{A} = \begin{bmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (1)$$

where  $s$  is a scale factor;  $\mathbf{A}$  is the intrinsic matrix, with  $\alpha$  and  $\beta$  the horizontal and vertical focal length in pixel unit,  $\gamma$  the skew factor, and  $(u_0, v_0)$  the coordinates of the principal point;  $\mathbf{R}$  and  $\mathbf{T}$  are the extrinsic parameters that denote the rotation and translation relating the world coordinate system to the camera coordinate system. Due to nonlinear optical distortion, the above equation is not sufficient for accurate camera calibration. In spite of the fact that some very complex models exist,<sup>5</sup> in practice they induce more instability rather than accuracy because of the high order distortion components. Here, the lens distortion is compensated by:

$$\begin{aligned} u' &= (1 + a_0 r^2 + a_1 r^4 + a_2 r^6)u + s_0 r^2 \\ &\quad + (p_0 + p_2 r^2)(r^2 + 2u^2), \\ v' &= (1 + a_0 r^2 + a_1 r^4 + a_2 r^6)v + s_1 r^2 \\ &\quad + (p_1 + p_3 r^2)(r^2 + 2v^2), \\ r^2 &= u^2 + v^2, \end{aligned} \quad (2)$$

where  $(a_0, a_1, a_2)$ ,  $(s_0, s_1)$ , and  $(p_0, p_1)$  represent the radial, prism, and tangential distortion coefficients, respectively,  $(u, v)$  denotes the distortion-free pixel location, and  $(u', v')$  is the corresponding distorted point. To avoid calculation instability during the optimization process, in Eq. (2),  $(u, v)$  and  $(u', v')$  are usually converted to their normalized terms  $(x_{cn}, y_{cn})$  and  $(x'_{cn}, y'_{cn})$  by applying the intrinsic matrix  $\mathbf{A}$ .

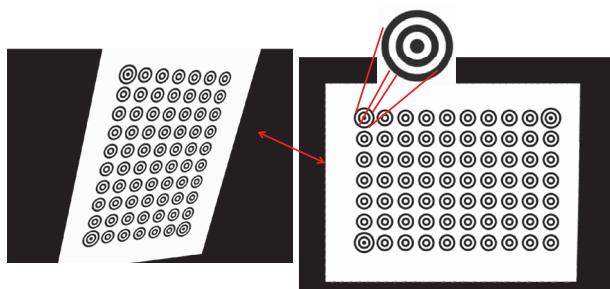
Given the locations of the control points in the world coordinate as  $\mathbf{M}_i$  and in the image plane as  $\mathbf{m}_{ij}$ , the calibration process involves a nonlinear optimization with the cost function defined as:

$$S = \sum_{i=1}^k \sum_{j=1}^l \|\mathbf{m}_{ij} - P(\mathbf{A}, \boldsymbol{\varphi}, \mathbf{r}, \mathbf{T}, \mathbf{M}_i)\|^2, \quad (3)$$

where  $k$  and  $l$  denote the number of control points in the calibration target and the number of images, respectively,  $\boldsymbol{\varphi} = (a_0, a_1, a_2, s_0, s_1, p_0, p_1)$ ,  $P$  denotes the projection of control points onto the image planes according to Eqs. (1) and (2), and  $\mathbf{r}$  is the Rodrigues vector presentation of  $\mathbf{R}$ .<sup>1,6</sup> The optimization is performed using the Levenberg–Marquardt (LM) algorithm.

The method described above relies on the intrinsic assumption that the control points on the calibration target are perfectly positioned. In reality, printing imprecision during calibration pattern fabrication is inevitable, especially for low-cost hardware systems. To deal with this issue, the proposed technique allows the world coordinates of the control points to be unknown, and they will be accurately determined together with the camera parameters. To ensure the uniqueness of the calibration target during optimization, a geometric constraint on three noncollinear control points, named markers, is applied.<sup>6</sup> The constraint enforces planarity of the markers by setting their Z world coordinates to zero, and requires the distance between any two of the three markers to be accurately measured to solve for the scale factor. The markers, normally made at the corners of the calibration panel, also help determine the orientation of the panel.

An accurate ellipse fitting technique can be employed to detect the centers of the calibration circles.<sup>7</sup> However, as the control points detected in the raw images suffer from lens and perspective distortion, their true locations cannot be accurately determined, and this will further lead to calibration errors. To cope with this issue, a two-step refinement with the frontal image concept, which is free from lens and perspective distortion, is employed. The refinement is conducted after the camera parameters are coarsely determined using the method described above. The raw images are successively undistorted by applying Eq. (2) and then reversely projected to the world coordinate system through using Eq. (1). The frontal image is generated from a direct scaling of the world coordinates. It is noteworthy that while the conventional subpixel edge and circle fitting method can be used to detect the control points in the frontal images, the method is not recommended because the gradient calculation involved in edge detection is sensitive to noise and the circle approximation does not yield high accuracy. On the other hand, as illustrated in Fig. 1, the transformation of each image from each location of the captured calibration board to the frontal image allows using the DIC concept to accurately locate the position of each control point by comparing the frontal image with the synthesized templates. Datta et al.<sup>4</sup> achieved subpixel-location detection of the control points by carrying out a quadratic fitting in the neighborhood regions based on their correlation coefficients, but such a peak-finding approach is less accurate than the iterative schemes.<sup>8</sup> Here, the proposed technique employs a high-accuracy cross-correlation algorithm, and the correlation coefficient function is:<sup>6,8</sup>



**Fig. 1** The conversion from raw images to frontal images enables the correlation with the ring templates.

$$C = \sum_{i=1}^N [af(x_i, y_i) + b - g(x'_i, y'_i)]^2, \quad (4)$$

where  $a$  is a scale factor,  $b$  is an intensity offset, and  $f(x_i, y_i)$  and  $g(x'_i, y'_i)$  denote the intensity values at the  $i$ -th pixel in the template and frontal images, respectively. The template pattern is a square subimage of  $N$  pixels with its center as the center of the circular target point. Denoting the subimage center as  $(x_0, y_0)$  and its shift amount between two subimages as  $(u, v)$ , the correlation shape function considering both translation and rotation is

$$x'_i = x_i + u + u_x(x_i - x_0) + u_y(y_i - y_0), \quad (5)$$

$$y'_i = y_i + v + v_x(x_i - x_0) + v_y(y_i - y_0),$$

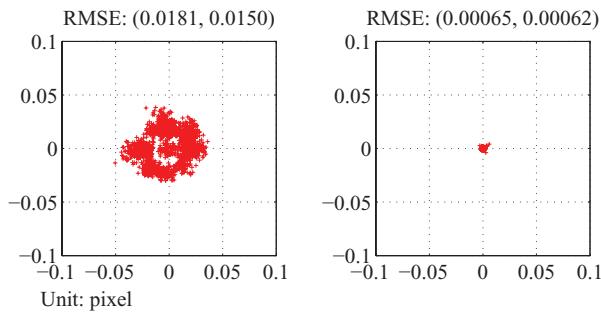
where  $u_x, u_y, v_x$ , and  $v_y$  are coefficients of the shape function. To determine the eight unknowns  $(u, v, u_x, u_y, v_x, v_y, a, b)$ , the Newton–Raphson algorithm is employed to minimize  $C$  in Eq. (4). With the detected  $(u, v)$ , the location of each control point in the frontal image can be directly determined. Then, these points in the frontal image are reversely projected back to the image plane to achieve hyperaccurate localization of the control points.

The rationale of the proposed method is that although the imprecise control points detected in the raw images lead to inaccurate camera calibration parameters, they offer good initial information about the relation between the camera and each scene which can be further processed to achieve very accurate localization of the control points. This information helps detect the camera intrinsic and extrinsic parameters as well as the world coordinates of the control points with high accuracies. The procedure of the proposed technique is summarized as follows:

1. Detect the control points in the raw images using the edge detection and ellipse fitting method.
2. Optimize the camera parameters and world coordinates of the control points using the LM algorithm. This involves a similar procedure as the conventional technique using Eqs. (1)–(3).
3. Obtain the frontal images, and use the calculated positions of the control points in the frontal images as initial guess to refine the positions using the DIC method.
4. Reversely project the detected control points in the frontal images back to the raw images.
5. Re-optimize the camera parameters together with the world coordinates of the control points.

To demonstrate the validity of the proposed technique, computer simulation along with real experiment have been conducted. The results suggest that no less than 40 control points and 15 positions are required to ensure a reliable calibration. Both simulation and experiment presented below use a flat panel with  $10 \times 7$  ring patterns whose grid distance is 25.4 mm (as illustrated in Fig. 1), and 20 images at different positions are used in the calibration.

In the simulation, the images are synthesized with camera parameters obtained from a real calibration where the radial, tangential, and prism lens distortion are considered. In addition, Gaussian noise with a standard deviation of 0.2% of the 25.4 mm distance is added to the position of each ring pattern



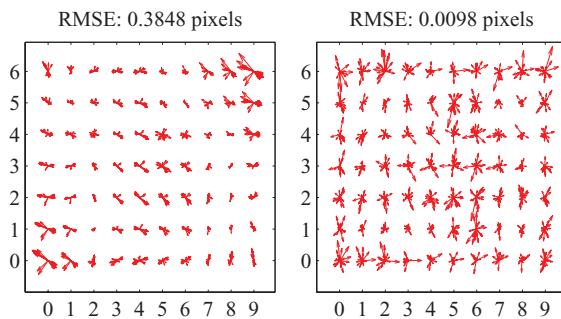
**Fig. 2** Localization errors of the control points obtained by Heikkila's method and the proposed method.

in the  $x$  and  $y$  directions, and the images are also blurred and Gaussian noise with  $\sigma = 2.0$  is added. Figure 2 shows the errors in locating the positions of the control points against their true values. Using Heikkila's technique<sup>7</sup> to detect the control points directly in the raw images, the effect of perspective distortion where the ellipse center is not the actual circle center is obvious. In contrast, with the frontal image and DIC concepts, the distortion effect is removed, and over 20 times localization improvement in terms of the root-mean-square error (RMSE) in both  $x$  and  $y$  directions is achieved. The residual of the calibration, named reprojection error and defined as the RMSE between the projection of the control points to the image planes and their measured locations, is 0.00086 pixels. Table 1 shows the results of the intrinsic parameters. It can be seen that the camera parameters can be

**Table 1** Intrinsic calibration parameters.

	True	Retrieved	Error
$\alpha$	5776 <sup>a</sup>	5775.9906 <sup>a</sup>	0.00016%
$\beta$	5776 <sup>a</sup>	5775.9910 <sup>a</sup>	0.00015%
$\gamma$	0.2767 <sup>a</sup>	0.2772 <sup>a</sup>	-0.18%
$u_0$	1067 <sup>a</sup>	1067.0262 <sup>a</sup>	0.0025%
$v_0$	751 <sup>a</sup>	750.9649 <sup>a</sup>	0.0047%
$a_0$	0.179	0.1789	0.012%
$a_1$	-0.334	-0.3339	0.33%
$a_2$	28.110	28.0892	0.074%
$p_0$	-0.00371	-0.0037067	0.19%
$p_1$	-0.00229	-0.0022943	-0.14%
$s_0$	0.00631	0.0063065	0.16%
$s_1$	0.00627	0.0063018	-0.36%

<sup>a</sup>Unit: pixel.



**Fig. 3** Reprojection errors at the control points obtained by the conventional method and the proposed method. The vector scales are different in the figures for clear illustration.

accurately retrieved once the control points are detected with high precision.

For the real experiment, Fig. 3 shows the displacement error vectors between the detected and projected positions at each control point of the calibration target. With the conventional method and the proposed one, the reprojection errors are 0.3848 and 0.0098 pixels, respectively. Because of the insufficiency of the lens distortion model and the imperfection of the calibration target, the reprojection error is noticeably polarized with the conventional method in spite of its control refinement. With the proposed method, the error behaves more isotropic, indicating that the lens distortion and the imprecision of the calibration target have been successfully compensated.

#### Acknowledgments

This work was supported by the Army Research Office (ARO) under Grant No. W911NF-10-1-0502 and the National Science Foundation (NSF) under Grant No. CMMI-0825806.

#### References

1. Z. Zhang, "A flexible new technique for camera calibration," *IEEE Trans. Pattern Anal. Mach. Intell.* **22**(11), 1330–1334 (2000).
2. M. Pollefeys, R. Koch, and L. van Gool, "Self-calibration and metric reconstruction in spite of varying and unknown intrinsic camera parameters," *Int. J. Comput. Vision* **32**(1), 7–25 (1999).
3. J. M. Lavest, M. Viala, and M. Dhome, "Do we really need an accurate calibration pattern to achieve a reliable camera calibration?," in *Proceedings of European Conference on Computer Vision*, Freiburg, Germany, p. 158 (1998).
4. A. Datta, J. S. Kim, and T. Kanade, "Accurate camera calibration using iterative refinement of control points," in *Proceedings of IEEE Conference on Computer Vision Workshops* (2009).
5. D. Douxchamps and K. Chihara, "High-accuracy and robust localization of large control markers for geometric camera calibration," *IEEE Trans. Pattern Anal. Mach. Intell.* **31**(2), 376–383 (2009).
6. M. A. Sutton, J. J. Orteu, and H. W. Schreier, *Image Correlation for Shape, Motion, and Deformation Measurements*, Springer, New York (2009).
7. J. Heikkila, "Moment and curvature preserving technique for accurate ellipse boundary detection," in *Proceedings of Fourteenth International Conference on Pattern Recognition* (1998).
8. B. Pan, K. Qian, H. Xie, and A. Asundi, "Two-dimensional digital image correlation for in-plane displacement and strain measurement: a review," *Meas. Sci. Technol.* **20**, 062001 (2009).