Accuracy Assessment on Camera Calibration Method Not Considering Lens Distortion

Sheng-Wen Shih[‡†], Yi-Ping Hung[†] and Wei-Song Lin[‡]
[†]Institute of Information Science, Academia Sinica, Nankang, Taipei, Taiwan, 11529 R.O.C.
[‡]Institute of Electrical Engineering, National Taiwan University, Taipei, Taiwan, R.O.C.

Abstract

This work investigates the effect of neglecting lens distortion, and presents a theoretical analysis of the calibration accuracy. The derived error bound is a function of a few factors including the number of calibration points, the observation error of 2D image points, the radial lens distortion coefficient, the image size and resolution. This error bound provides a guide line for selecting both a proper camera calibration configuration and an appropriate camera model while satisfying the desired accuracy. Experimental results from both computer simulations and real experiments are included in this paper.

1: Introduction

The techniques for camera calibration can be classified into two categories: one that considers lens distortion [2] [4] [8] [9], and one that neglects lens distortion [3] [7]. A typical linear technique that does not consider lens distortion is the one estimating the perspective transformation matrix **H** [3] [7]. The estimated **H** can be used *directly* for forward and backward 3D–2D projection. If necessary, given the estimated **H**, the geometric camera parameters can be easily determined [3] [6].

Nonlinear optimization techniques are usually necessary for estimating the camera parameters when considering lens distortion [2] [8] [9]. Faig's method [2] is a good representative for those nonlinear methods. One disadvantage of this kind of method is that a good initial guess is required to start the nonlinear search. Recently, Weng showed some experimental results using a two-step method [9], which used the initial parameters estimated from a closed form linear solution based on a distortion-free camera model.

In general, considering lens distortion will not only complicate the camera calibration procedure, but also complicate the subsequent on-line processing (though not formidable) such as feature-point correspondence (in stereo) and camera re-calibration (in the case of having a moving camera). Notice that epipolar line is no longer a straight line if lens distortion is taken into account. Moreover, when lens distortion is small, if the noise in the 2D feature extraction is relatively large or the number of the calibration points is relatively small, the calibration results based on distortion camera model can be worse than those based on linear camera model. The question is then, "when should we consider lens distortion in camera calibration?" or "when does it worth all the troubles to consider lens distortion?" This work represents an effort toward the answer of the question.

2: Camera model

Let P be an object point in the 3D space, and $\mathbf{r}_O = (x_O, y_O, z_O)$ be its coordinates, in millimeters, with respect to a fixed object coordinate system (OCS). Let $\mathbf{s}_I = (u_I, v_I)$ denote the 2D image coordinates (in pixels), with respect to the computer image coordinate system (ICS), of the image point Q, where the origin of ICS is located at *the center of the frame memory coordinate* (e.g. the origin of the ICS is right at (256, 240) for a 480 by 512 image). The relationship between $\mathbf{r}_O$ and $\mathbf{s}_I$ can be expressed as [4]

$$(1-\kappa\varrho^2)(u_I-u_0)\delta_u = f\frac{x_O r_1 + y_O r_2 + z_O r_3 + t_1}{x_O r_7 + y_O r_8 + z_O r_9 + t_3} \qquad (1)$$

$$(1-\kappa\varrho^2)(v_I-v_0)\delta_v = f\frac{x_O r_4 + y_O r_5 + z_O r_6 + t_2}{x_O r_7 + y_O r_8 + z_O r_9 + t_3} \qquad (2)$$

where $\varrho = \sqrt{\delta_u^2(u_I-u_0)^2 + \delta_v^2(v_I-v_0)^2}$.

Using the above notations for camera parameters, t_1, t_2, t_3 and r_i, $i = 1 \ldots 9$, are parameters for coordinate transformation, f is the effective focal length, δ_u and δ_v are respectively the horizontal and vertical pixel spacing, u_0 and v_0 are the piercing point of the optical axis on the image plane, and κ is the coefficient of the lens distortion. Notice that, suppose there is no optical distortion (i.e., $\kappa = 0$), the relationship between $\mathbf{r}_O$ and $\mathbf{s}_I$ can be expressed as a linear transformation.

3: Accuracy assessment

This section introduces an approximate error bound for the linear calibration method [5] which does not consider lens distortion. The error bound is based on the following assumptions:

(a.1). The 3D positions of the calibration points are known exactly. In practice, it is easier to precisely locate the 3D position of a calibration point, comparing to its 2D image coordinates. Furthermore, the 3D position error can always be transformed to an equivalent 2D measuring error.

(a.2). The only source of measurement noise is the error in estimating the image coordinates of the calibration points, i.e., the 2D observation noise (in pixels). In both horizontal and vertical directions, we assume the 2D observation noise have the identically independent Gaussian distribution with zero mean and the variance, σ^2.

(a.3). The depth components of both calibration and test points, z_C, can be approximated by a constant (i.e. the depth of field is small relative to the object distance). This assumption holds in most computer vision applications, since the depth of field for a practical camera is usually limited to a small range comparing to the object distance.

To evaluate the accuracy of the camera calibration for 3D vision application, it is necessary to define an error measure. The error measure adopted in this paper is the *2D prediction error*, i.e, the image distance (in pixels) between the Q and $\hat{Q}$, where Q is the true 2D image coordinates of the test point P, and $\hat{Q}$ is the predicated image coordinates of the 3D test point P using the estimated parameters[5].

The *expectation* of the average square 2D prediction error due to measurement noise has been shown to be (see [5])

$$\epsilon_n^2 = \frac{11\ \sigma^2}{N_{calib}} \qquad \text{(in pixels)} \qquad (3)$$

where N_{calib} is the number of calibration points and σ^2 is the variance of the 2D observation noise.

Also, the 2D error bound due to the modeling error (the negligence of lens distortion) has been derived in [5]

$$\epsilon_M^2 = \kappa^2 \frac{R^6}{36\delta_a^2} \qquad \text{(in pixels)} \qquad (4)$$

where δ_a is the average pixel spacing in all directions (refer to [5]), and R (in millimeters) is half the diagonal size of the image sensing area.

Assume that the interaction between measurement noise and modeling error is negligible. Then, we have the approximate total mean square 2D prediction error by combining (3) and (4):

$$\epsilon_{Bound}^2 \approx \epsilon_M^2 + \epsilon_n^2 . \qquad (5)$$

Notice that the second term, ϵ_n^2, of equation (5) is an expectation value, which means that the violation of the approximate upper bound, ϵ_{Bound}, is possible.

4: Experiments

The first experiment tested the error bound by computer simulations. Totally, ten thousand trials were simulated with randomly selected camera parameters f, u_0, v_0, κ, and 2D noise σ. The calibration points and test points were randomly generated, too. More details are described in [5]. For each random trial, the computed 2D prediction error is normalized by its theoretic bound. Fig. 1 shows the histogram of the normalized error which shows that, in most trials the 2D prediction error is close to and less than the theoretic bound. Still, there are some points which exceed the theoretic bound. This is partially because of the violation of the assumption (a.3) in section 3. Another reason is that ϵ_n^2 is an expectation value, not an upper bound.

The second experiment tested the bound by a real experiment. We took 21 images of a moving calibration plate having 25 calibration points on it, which was mounted on a translation stage. One image was taken each time the translation stage was moved toward the camera by 25 millimeters. A typical 480 × 512 image is shown in Fig. 3. Thus we have 21 × 25 = 525 pairs of 2D-3D coordinates of points. The image coordinates of

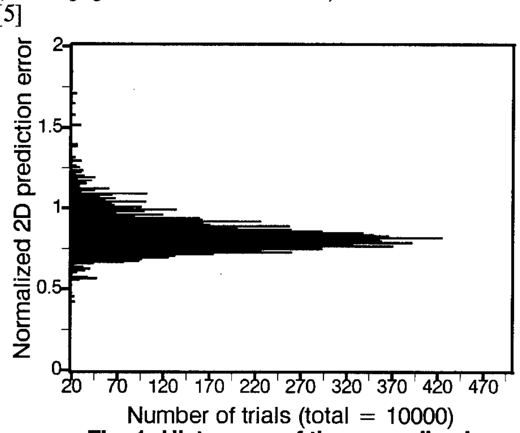

Fig. 1. Histogram of the normalized 2D prediction error

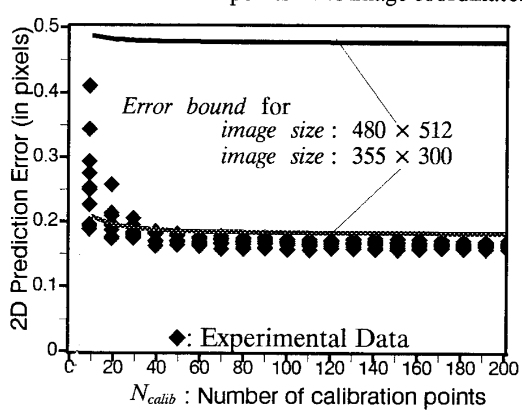

Fig. 2. Real experiment

the center for each circle was estimated, with an error of about 0.1 pixel. We randomly chose N_{calib} points (N_{calib} = 10, 20, 30, ... or 200) from the 525 2D-3D pairs to calibrate the camera and used all remaining points to test the calibrated parameters. The above random trials were repeated ten times to obtain ten sets of the 2D prediction error. Fig. 2 shows the ten sets of data and two predicted bounds based on two different effective image sizes (here κ = 0.00035 millimeter^{-2}, which corresponds to roughly 2 to 3 pixels of distortion near the four image corners). Since all the calibration and test points are distributed in the central part of the image, whose size is roughly of 355 by 300 pixels (see Fig. 3), the bound calculated with this image size is much closer to the experimental results. To use every pixels in the 480×512 image, the error bound will be approximately three times of the experimental results.

5: Conclusions

For 3D applications, e.g., stereo vision, it is of great importance to determine the accuracy of 3D position estimation. Knowing the 2D prediction error, the 3D position error can be derived as in [1]. Thus, the error bound can be used as a criterion to decide whether the linear camera model is sufficient or not, for a specific application. In the following, a general guide line is provided for using this error bound:

1). Determine the acceptable 2D prediction error, denoted as ϵ_{spec}.

2). Calculate the approximate error bound, ϵ_{Bound}, by equation (5) according to the parameters of the equipments to be used.

3). If $\epsilon_{spec} > \epsilon_{Bound}$ then it is good enough to use the linear camera model.

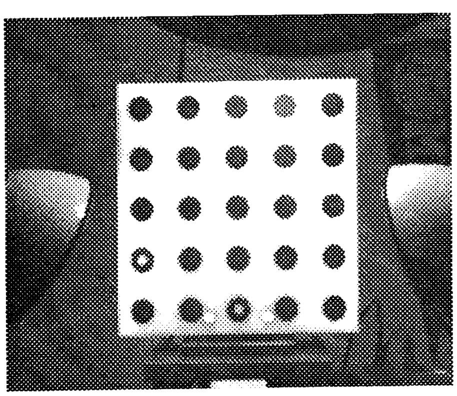

Fig. 3. A typical image of the calibration plate containing 25 calibration points used in the real experiment.

4). If $\epsilon_{spec} < \epsilon_{Bound}$ then try to reduce ϵ_n in equation (5) as much as possible, by making the feature extraction more accurate (reduce σ) and increase the number of calibration points.

5). If ϵ_{Bound} still can not meets the requirement after the reduction of ϵ_n in step 4), then try to reduce the effective size of the image to an acceptable level (see equation (4)).

6). If the efforts in step 4) and 5) fail to reduce ϵ_{Bound} such that $\epsilon_{spec} > \epsilon_{Bound}$, then a nonlinear camera model should be considered in the camera calibration procedure as in [2] [4] [8] [9].

6: References

[1] S.D. Blostein, T.S. Huang, "Error Analysis in Stereo Determination of 3-D Point Positions," IEEE PAMI, Vol. 9, No. 6, Nov., 1987, pp. 752–765.

[2] W. Faig, "Calibration of Close-Range Photogrammetry Systems: Mathematical Formulation," Photogrammetric Engineering and Remote Sensing, Vol. 41, No. 12, 1975, pp. 1479–1486.

[3] Y.P. Hung, "Three Dimensional Surface Reconstruction Using a Moving Camera A Model-Based Probabilistic Approach," Ph.D dissertation, Division of Engineering, Brown University, 1990.

[4] S.W. Shih, Y.P. Hung, W.S. Lin, "An Efficient and Accurate Camera Calibration Technique for 3D computer Vision," Proc. SPIE Conf. on Optics, Illumination, and Image Sensing for Machine Vision VI, November 1991, pp. 133–145.

[5] S.W. Shih, Y.P. Hung, W.S. Lin, "Accuracy Assessment on Camera Calibration Method not Considering Lens Distortion," Technical Report TR-92-001, Institute of Information Science, Academia Sinica, Taipei.

[6] T.M. Strat, "Recovering the Camera Parameters from a Transformation Matrix," DARPA Image Understanding Workshop, 1984.

[7] I. Sutherland, "Three-Dimensional Data Input by Tablet," Proceedings of the IEEE, Vol. 62, No. 4, 1974, pp. 453–461.

[8] 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. of Robotics and Automation, Vol. RA-3 No. 4, 1987, pp. 323–344.

[9] J. Weng, P. Cohen, M. Herniou, "Calibration of Stereo Cameras Using a Non-linear Distortion Model," Proc. 10th Inter. Conf. on Pattern Recognition, 1990, pp. 246–253.