Accuracy Analysis: Influence of Camera Calibration on Vision Measurement

Sheng Gao†, Yu Song‡, Fei Wang‡ and Haiwei Yang‡

† Beijing Institute of Spacecraft System Engineering, Beijing, China
‡ The School of Electronic and Information Engineering, Xi'an Jiaotong University,
Xi'an, Shaanxi Province, China
Email: castgaosheng@163.com, songyuxjtu@stu.xjtu.edu.cn,
wfx@mail.xjtu.edu.cn, yanghw.2005@stu.xjtu.edu.cn

Abstract– In this paper we focus on the accuracy analysis problem of vision measurement system in the process of three dimensional pose estimation of cooperative target, systematically describe the camera model and a camera calibration method based on plane targets, and launch an in-depth research to seek for the relation between the camera calibration errors and pose measurement result. The camera calibration errors are usually ignored in practical applications and they directly affects the accuracy of vision measurement to a large extent. Our central idea is to make only one variable different at a time by controlling various variables. Then the effect of that single factor can be determined. In addition, a large number of synthetic and real data experiments have been developed. As a consequence, we establish the relationship between the intrinsic parameters and pose estimation errors effectively. This paper provides reference and basis to control errors during vision measuring subsequently.

1. Introduction

With the rapid development of computer hardware, software, image acquisition and processing techniques, the theories and techniques of computer vision have been widely used in medical image processing, robotics, character recognition, industrial inspection, geographical mapping and so on. The applications of computer vision in a variety of measurements are quantitative analyses with determined accuracy requirements.

Camera calibration is the first and critical step in vision measurement. It can be carried out by classical methods in which enough 3D points and their image correspondences are available as in [8]. Zhang [5] has proposed a flexible technique for camera calibration, which only requires the camera to observe a planar pattern shown at a few different orientations. Self-calibration is nonlinear and the correspondences between several images are required to get a unique solution. However self-calibration is lack of robustness [9]. Active vision calibration [10, 11] can overcome above shortcomings, but it undergoes with the known special motion of camera, which limits its range of application.

The traditional method by observing a calibration object with known geometry in 3D space has very good precision. In this paper we make use of Matlab Camera calibration toolbox [4] based on Zhang's calibration technique [5] to calibrate a camera and choose a widely used linear algorithm EPnP [6, 7] as the pose estimation algorithm.

Calibration methods usually cause errors to a certain degree. Thus a careful analysis of influence of camera calibration on pose measurement is necessary.

Behrooz Kamgar-Parsi [1] developed the mathematical tools to compute the average or expected error due to quantization. They derived the analytic expression for the probability density of error distribution of a function with an arbitrarily large number of independently quantized variables. Sunil Kopparapu [2] derived analytically the behavior of the camera calibration matrix under noisy conditions and further showed that the elements of the camera calibration matrix have a Gaussian distribution if the noise introduced into the measurement system is Gaussian. Jeffrey Rodriguez [3] derived the probability density function of the range estimation error and the expected value of the range error magnitude in terms of the various design parameters.

We launch an in-depth and meticulous research focusing on the influence of camera calibration on pose measurement. Theoretical analysis is carried out by controlling various variables and making only one variable different at a time, so that the effect of that single factor can be determined. In this paper, we construct error models for each parameter. Afterwards our approach is applied to both synthetic and real images, and then satisfying results are obtained. Theoretical analysis and experiments validate the performance of our approach.

Eventually the relationship between the intrinsic parameters and pose estimation error is established through theoretical derivation and experimental verification. As a consequence, this paper provides reference and basis for error controlling during vision measuring process subsequently to achieve more precise pose.

2. Camera Calibration

2.1. Camera Model

The pinhole camera model is used in this paper. In this model, a scene view is formed by projecting 3D points onto the image plane using a perspective projection

transformation. The model describes the mathematical relationship between the coordinates of a 3D point and its projection onto the image plane, which is given by

$$s\mathbf{p} = \mathbf{A}\cdot(\mathbf{R}\cdot\mathbf{P}_\mathbf{w} + \mathbf{t}), \text{ with } \mathbf{A} = \begin{bmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (1)$$

where s is the scale factor, **R** and **t** are extrinsic parameters, i.e. pose, representing rotation matrix and translation vector respectively, $\mathbf{P}_\mathbf{w} = (X_w, Y_w, Z_w)^{\mathrm{T}}$ are coordinates of a 3D point in world coordinate system, $\mathbf{p} = (u, v)^{\mathrm{T}}$ are coordinates of the projection point in pixels, and **A** is the matrix of intrinsic parameters, with (u_0, v_0) the principal point and (f_x, f_y) the equivalent focal lengths expressed in pixels.

(u, v) mentioned above are coordinates of the projection point in pixels, and (x, y) are its coordinates in millimeters. The relationship between them is given by

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} 1/dx & 0 & u_0 \\ 0 & 1/dy & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}, \quad (2)$$

where the physical sizes of every pixel in X and Y axis direction are dx and dy. $\mathbf{P}_\mathbf{w}$ are defined above, and $\mathbf{P}_\mathbf{c} = (X_c, Y_c, Z_c)^{\mathrm{T}}$ are the coordinates of a 3D point in camera coordinate system. The relationship between these two coordinate systems is given by

$$\mathbf{P}_\mathbf{c} = \mathbf{R}\cdot\mathbf{P}_\mathbf{w} + \mathbf{t}. \quad (3)$$

The projection equation of the pinhole camera model is given by

$$s\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \mathbf{P}_\mathbf{c}, \quad (4)$$

where the scale factor s is equal to Z_c. From (2), (3), (4), the projection from 3D world points to 2D image points is established as shown in (1).

2.2. Calibration Technique

We assume that the plane of the calibration target is on Z=0 in the world coordinate system. Let's denote the *i*th column of the rotation matrix **R** by $\mathbf{r}_i$. From (1), the projection equation in homogeneous form is given by

$$s\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{A}\cdot[\mathbf{r}_1 \quad \mathbf{r}_2 \quad \mathbf{r}_3 \quad \mathbf{t}] \begin{bmatrix} X_w \\ X_w \\ 0 \\ 1 \end{bmatrix}. \quad (5)$$

We still use $\mathbf{P}_\mathbf{w}$ to denote a point on the target plane but $\bar{\mathbf{P}}_\mathbf{w} = (X_w, Y_w, 1)^{\mathrm{T}}$. Therefore, a 3D point $\mathbf{P}_\mathbf{w}$ and its image points **p** are related by a homography **H**.

$$s\bar{\mathbf{p}} = \mathbf{H}\cdot\bar{\mathbf{P}}_\mathbf{w}, \text{ with } \mathbf{H} = \mathbf{A}\cdot[\mathbf{r}_1 \quad \mathbf{r}_2 \quad \mathbf{t}], \quad (6)$$

where $\bar{\mathbf{p}} = (u, v, 1)^{\mathrm{T}}$.

As is clear, the matrix **H** is defined up to a scale factor. Given an image of the model plane, a homography can be estimated. Let's denote it by $\mathbf{H} = [\mathbf{h}_1 \quad \mathbf{h}_2 \quad \mathbf{h}_3]$.

From (6), we have

$$[\mathbf{h}_1 \quad \mathbf{h}_2 \quad \mathbf{h}_3] = \lambda\mathbf{A}[\mathbf{r}_1 \quad \mathbf{r}_2 \quad \mathbf{t}], \quad (7)$$

where λ is an arbitrary scalar. According to the constraint that $\mathbf{r}_1$ and $\mathbf{r}_2$ are orthonormal, we have

$$\mathbf{h}_1^{\mathrm{T}}\mathbf{A}^{-\mathrm{T}}\mathbf{A}^{-1}\mathbf{h}_2 = 0, \quad (8)$$

$$\mathbf{h}_1^{\mathrm{T}}\mathbf{A}^{-\mathrm{T}}\mathbf{A}^{-1}\mathbf{h}_1 = \mathbf{h}_2^{\mathrm{T}}\mathbf{A}^{-\mathrm{T}}\mathbf{A}^{-1}\mathbf{h}_2, \quad (9)$$

These are the two basic constraints on the intrinsic parameters given one homography. Let

$$\mathbf{B} = \mathbf{A}^{-\mathrm{T}}\mathbf{A}^{-1} = \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{12} & B_{22} & B_{23} \\ B_{13} & B_{23} & B_{33} \end{bmatrix}. \quad (10)$$

Note that **B** is a symmetric matrix defined by a 6D vector $\mathbf{b} = (B_{11}, B_{12}, B_{22}, B_{13}, B_{23}, B_{33})^{\mathrm{T}}$. Let the *i*th column vector of **H** be $\mathbf{h}_i = (h_{i1}, h_{i2}, h_{i3})^{\mathrm{T}}$. Then, we have

$$\mathbf{h}_i^{\mathrm{T}}\mathbf{B}\mathbf{h}_j = \mathbf{v}_{ij}^{\mathrm{T}}\mathbf{b}, \quad (11)$$

with

$$\mathbf{v}_{ij}^{\mathrm{T}} = [h_{i1}h_{j1}, h_{i1}h_{j2} + h_{i2}h_{j1}, h_{i2}h_{j2}, h_{i3}h_{j1} + h_{i1}h_{j3}, h_{i3}h_{j2} + h_{i2}h_{j3}, h_{i3}h_{j3}]^{\mathrm{T}}. \quad (12)$$

Therefore, the two fundamental constraints (8) and (9) from a given homography can be rewritten as two homogeneous equations as follows.

$$\begin{bmatrix} \mathbf{v}_{12}^{\mathrm{T}} \\ (\mathbf{v}_{11} - \mathbf{v}_{22})^{\mathrm{T}} \end{bmatrix}\mathbf{b} = \mathbf{0}, \quad (13)$$

If n images of the model plane are observed, by stacking n such equations as (13), we have

$$\mathbf{V}\mathbf{b} = \mathbf{0}, \quad (14)$$

where **V** is a $2n \times 6$ matrix. If $n \geqslant 3$, we will have in general a unique solution **b** defined up to a scale factor.

The solution to (11) is the eigenvector of $\mathbf{V}^{\mathrm{T}}\mathbf{V}$ associated with the smallest eigenvalue. Since **b** is known, we can compute the intrinsic parameters from matrix **B** by the following equations.

$$\begin{aligned} v_0 &= (B_{12}B_{13} - B_{11}B_{23})/(B_{11}B_{22} - B_{12}^2) \\ \lambda &= B_{33} - [B_{13}^2 + v_0(B_{12}B_{13} - B_{11}B_{23})]/B_{11} \\ f_x &= \sqrt{\lambda/B_{11}} \\ f_y &= \sqrt{\lambda B_{11}/(B_{11}B_{22} - B_{12}^2)} \\ u_0 &= cv_0/f_x - B_{13}f_x^2/\lambda \end{aligned} \quad (15)$$

3. Accuracy Analysis of Intrinsic Parameters

According to (1) and (3), we obtain $s\mathbf{p} = \mathbf{A}\cdot\mathbf{P}_\mathbf{c}$, which indicates that the intrinsic parameters affect the accuracy of the 3D points in camera coordinate system, when the 2D image points **p** are acquired precisely. Then we have

$$X_c = \frac{u-u_0}{f_x} Z_c \tag{16}$$

$$Y_c = \frac{v-v_0}{f_y} Z_c . \tag{17}$$

The Taylor expansions of (16) at true values of f_x and u_0 respectively are

$$X_c = \bar{X}_c + \frac{(-1)(u-u_0)Z_c}{\bar{f}_x^2}(f_x - \bar{f}_x) + \frac{(-1)^2(u-u_0)Z_c}{\bar{f}_x^3}(f_x - \bar{f}_x)^2 + \cdots + \frac{(-1)^n(u-u_0)Z_c}{\bar{f}_x^{n+1}}(f_x - \bar{f}_x)^n + R_n(f_x) \tag{18}$$

$$X_c = \bar{X}_c + \frac{Z_c}{\bar{f}_x}(u_0 - \bar{u}_0), \tag{19}$$

where variables with overline are true values of the corresponding ones, and R_n is the remainder of the Taylor formula.

The calibration result f_x is far greater than the error $(f_x - \bar{f}_x)$. Therefore we can ignore those high order terms in the Taylor expansion (18). It can be obtained that, the error of $\mathbf{P_c}$ has inverse relationship with the error of the equivalent focal lengths approximately, when the principal point coordinates are constant values. From (19), the error of $\mathbf{P_c}$ is linear with respect to the error of principal point strictly, when the equivalent focal lengths are constant.

It also applies to (17) to be expanded at f_y and v_0 respectively.

4. Experiments

4.1. Simulations

In the simulation experiment, firstly we produce synthetic 3D to 2D correspondences in a 640 × 480 image acquired using a virtual calibrated camera with an equivalent focal length of 800 and a principal point at (320, 240). Then we generated the 3D word points which are uniformly distributed in a cube. Given a rotation and translation, we obtain 2D image points by a perspective projection. Next we calculate pose using the pose estimation algorithm.

The error of the rotation matrix **R** is represented by the corresponding angle error **q**, which is defined by

$$E_q = \| \mathbf{q}' - \mathbf{q} \|, \tag{20}$$

and the error of the translation vector **t** is defined by

$$E_t = \| \mathbf{t}' - \mathbf{t} \|, \tag{21}$$

where $\mathbf{t}'$, $\mathbf{q}'$ are calculated translation vector and Euler angle respectively, and $\mathbf{t}$, $\mathbf{q}$ are corresponding true values.

Deviations distributed in the interval [−10, 10] with a step of 0.2 are generated. Then they are added to the equivalent focal lengths (f_x, f_y), and the principal point (u_0, v_0). We calculate the camera pose using the intrinsic parameters attached with deviations and compute errors of **R** (represented by Euler angle **q**) and **t**.

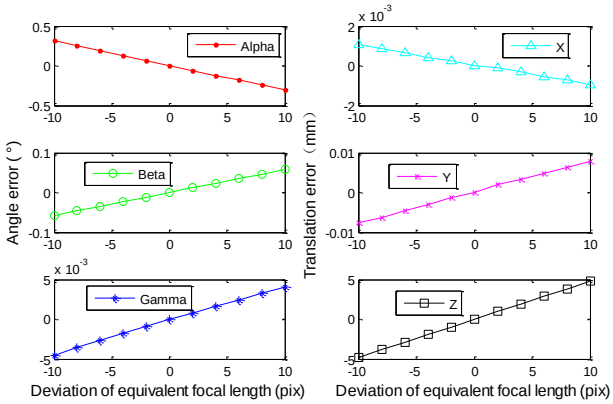


Fig. 1. Pose errors caused by deviations of the equivalent focal length

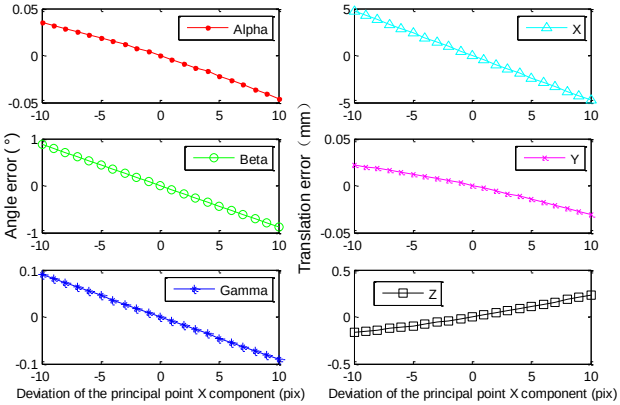


Fig. 2. Pose errors caused by deviations of the principal point X component u_0

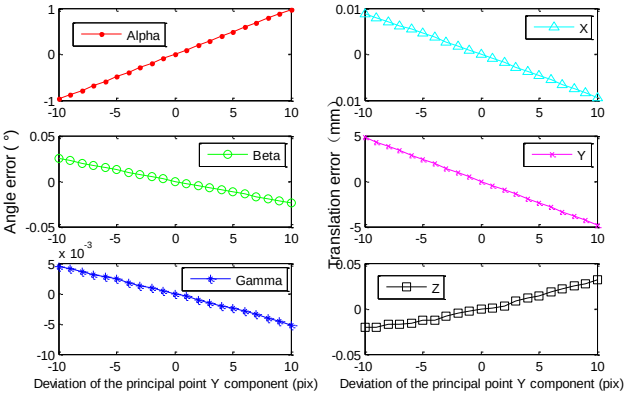


Fig. 3. Pose errors caused by deviations of the principal point Y component v_0

As shown in figures above, errors of all pose components are linear with the deviations. Moreover the focal length has a great effect on α and z, the principal

point in X axis effects β and x, and the principal point in Y axis effects α and y.

4.2. Real Data

In the real data experiments, images are taken by a Basler sca1600-14gc industrial camera with the resolution of 1624×1234 and a Computar 8mm lens.

In order to obtain clear results, a series of deviations distributed in the interval [0, 20] with a step of 5 are generated in pixel. Then they are added to the equivalent focal lengths (f_x, f_y), the X component of the principal point u_0 and the Y component of the principal point v_0 respectively. We project the 3D points on the image by the calculated camera pose.

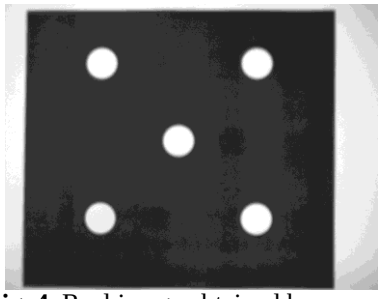
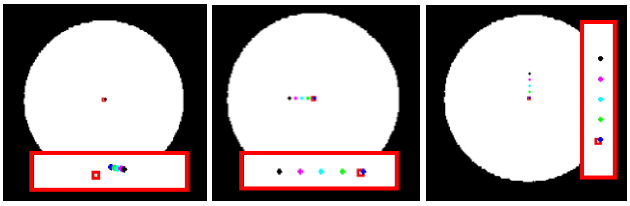


Fig. 4. Real image obtained by camera



(a) (b) (c)

Fig. 5. Experiment results (We take the central point in Figure 4 as example. Figure (a), (b) and (c) are the results when (f_x, f_y), u_0 and v_0 are added with deviation. The images in the red rectangles are the enlarged views corresponding to the central part of circles.)

As shown in figure 5 clearly, the red square is the center of a circle. The blue point is projection of the 3D point without deviation, the green one with 5 pixels, the cyan one with 10 pixels, the magenta one with 15 pixels and the black one with 20 pixels. The translation vector in X and Y axis are mainly affected by the X and Y component of the principal point respectively. The figures above show that the results of real data experiments agree with those of simulations and the theoretical analysis.

5. Conclusion

In this paper we launch an in-depth research to seek for the relation between the camera calibration errors and pose measurement result. Theoretical analysis is carried out by controlling various variables to make one variable different at a time to determine the effect of a single parameter. We construct error models for every intrinsic parameter. Afterwards our approach is applied to a large number of synthetic and real data experiments, and satisfying results are obtained. The relationship between the influencing factors and pose estimation error is established through theoretical derivation and experimental verification. As a consequence, this paper provides reference and basis for error controlling during vision measuring process subsequently to achieve more precise pose.

Acknowledgments

This work was supported by Supported by the Programme of Introducing Talents of Discipline to Universities (B13043), the National Natural Science Foundation of China (61273366) and the National High Technology Research and Development Program of China (2013AA014601).

References

[1] Kamgar-Parsi B. Evaluation of quantization error in computer vision [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1989, 11(9): 929-940.

[2] Kopparapu S K, Corke P. The effect of measurement noise on intrinsic camera calibration parameters [C]. Proceedings of IEEE International Conference on Robotics and Automation, 1999: 1281-1286.

[3] Rodriguez J J, Aggarwal J K. Quantization error in stereo imaging [C]. Computer Vision and Pattern Recognition, 1988. Proceedings CVPR'88, Computer Society Conference on. IEEE, 1988: 153-158.

[4] Camera calibration toolbox for matlab [CP/OL]. http://www.vision.caltech.edu.

[5] Zhang Z Y. A flexible new technique for camera calibration [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2000, 22(11), 1330~1334.

[6] Lepetit V, Noguer M F, Fua P. EPnP: An Accurate O(n) Solution to the PnP Problem [J]. International Journal of Computer Vision, 2009, 155–166.

[7] Noguer M F, Lepetit V, Fua P. Accurate Non-Iterative O(n) Solution to the PnP Problem [C]. Proceedings of International Conference on Computer Vision (ICCV2007), 2007:1-8.

[8] Faugeras O. Three-dimensional computer vision: a geometric viewpoint [M]. MIT press, 1993.

[9] Lei C, Wu F Z. Kruppa equations and camera self-calibration [J]. ACTA AUTOMATICA SINICA. 2001, 27(5): 621~630.

[10] Du F L, Brady M. Self-calibration of the intrinsic parameters of cameras for active vision systems[C]. In: Proceedings of Computer Vision and Pattern Recognition. 1993, 477~482.

[11] Ma S D. A self-calibration technique for active vision systems [J]. IEEE Transaction on Robotics and Automation, 1996, 12 (1): 114~120.