

# Pose Error Analysis Model Based on Binocular Vision for Rigid-Body

Niu Fenglian

School of Mechanical and Electrical Engineering, Ningbo Dahongying University,  
Ningbo 315175, P. R. China  
email: nfl1979@126.com

## Abstract

*In order to satisfy the orientation measuring requirements of rigid-body such as work piece, cutting tool in industry and medical instruments for invasive surgery, this paper presents a binocular vision detection technique based on spatial position information of markers to extract rigid-body pose information and analyzes the pose accuracy of rigid-body using the principal component analysis (PCA) and the least square method (LSM) when spatial position error of markers exist. The simulation experiment demonstrates the maximum angle error of orientation is about 0.59 degree when the position error of markers satisfy the Gaussian distribution with the mean is zero and the standard deviation is 0~3mm. The experimental results verify this method can robustly solve the orientation of rigid body using the position information of markers with position errors, and it provides a theoretical and experimental basis for orientation measurement of rigid body.*

**Keywords:** error analysis, orientation detection, binocular vision, camera calibration, the principal component analysis, the least square method

**Copyright © 2014 Institute of Advanced Engineering and Science. All rights reserved.**

## 1. Introduction

With the development of machine vision technology, orientation detection of work piece and medical instruments for invasive surgery is on the increase based on visual inspection in industrial technology and medical diagnosis [1-4]. Computer vision and photogrammetry have become a core issue, pose estimation as a research hot in the field of computer vision research has been carried on for many years [5]. The issue of study here is a dynamic pose estimation of the rigid body. For rigid-body pose estimation there are a large number of studies [6-7] including the method about the minimum sum of squares and singular value decomposition which is robust and fast calculation. Some researches have studied position error of the markers how to affect attitude solving, such as Woltring et al [8] proposed an estimation method of maximum error for solving a given tracking error of the probe when markers is symmetrically distributed on the probe, Morris and Donath et al [9] extended the Woltring's study and quantify the cumulative impact of multiple sources of error including dynamic deformation error algorithm targets and so on. However, no quantitative analysis about the influence of these methods position error how to influence the pose of rigid body. This article mainly focuses on pose estimation of rigid body based on binocular vision including calibration of intrinsic and extrinsic parameters of the camera, the solving about the orientation relation of two camera. According to the binocular camera parameters, image position information of markers in two image coordination systems are used to derive the spatial position of the markers in the world coordinate system. In pose solution of rigid body, the least squares method combined with principal component analysis are used to discuss how to use inaccurate location information of markers to analyze the orientation error and use simulation analysis to describe the entire research process.

## 2. Setting up of Vision Detection System

The pose detection system consists of two Toshiba industrial cameras, two PCI frame grabbers which type is Maxtor general and three infrared LED lights. The using of infrared LED is mainly on account of the ambient light affecting little to the acquired image in the experiment and deduce the segmentation difficulty of the ordinary image. Stereo vision system is used to

detect and track three IR markers and calculate the orientation and position of the rigid body relative to the world coordinate frame by merging virtual reconstruction system as Figure 1 shows.



Figure 1(a). Binocular vision system model

Figure 1(b). Virtual pose reconstruction system

### 2.1. Camera Calibration

To obtain spatial position of markers in the world coordinate system, the first procedure is to calibrate two cameras and obtain the relative orientation. Intrinsic parameters of the camera are used to acquire effective lens focal length, optical center, and lens distortion. Extrinsic calibration is needed to determine the orientation relation of the stereo cameras with respect to one another. Because the accuracy about camera calibration parameters directly affects the spatial location accuracy of markers, and there are some methods of the camera calibration to solve this problem. Taking into account the needs of system design, Zhang ZY's camera calibration algorithm [10] is adopted, which does not require the 3D spatial position information of markers and only needs to know 2D plane position information of markers in two camera coordinate systems, so it is very convenient to operate and has high robustness. Before calibration experiment, a high-quality printer is employed to print black and white checkerboard squares  $19 \times 21$  corner points on A0 paper, and paste the template on a flat wooden surface, where the length and width of each grid are equal and the side length is 20cm and pixel accuracy is 0.1 mm as Figure 3 shows. During the calibration, two cameras would be put in front of calibration template, translate distance and rotation angle don't exceed the scope of the camera view angle, the camera automatically detects the corner in the board using Susan corner detection algorithm<sup>[11]</sup>, Figure 2 is the left camera first image corner detected. In order to get the camera matrix and distortion parameters, we must get at least three or more different pose images.

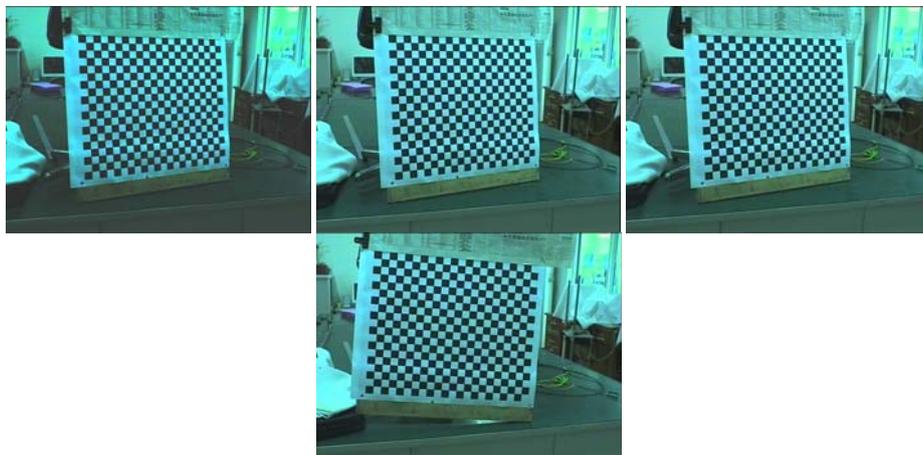


Figure 2. Acquired images on the left camera

The calibration process can be described as follows, firstly calibration template is placed in front of binocular camera, and acquire image points by detecting chessboard corner so that image corner point and correspondence corners on-board are built. which should be explained that the first corner of the lower left corner on the calibration plate is used as a calibration plate coordinate system origin point, while the image coordinate system origin point is in the upper left corner of the image. In order to solve the intrinsic and extrinsic parameter of camera, three or more different pose images are needed, so that the camera matrix and distortion parameters proposed by Zhang Zhengyu are obtained. When the four corner points and the corresponding image of the punctuation board corners corresponding relations are established, Internal and external parameters of the equation are solved. In the calibration, the relative positions of the cameras are fixed, right and left camera images are available, and so the internal and external parameters of two cameras are obtained at the same time.

For the development of calibration software, we make use of the Intel's OpenCV library to implement the calibration for our research. The image resolution is 768 pixel×576 pixel. The model plane contains a pattern of 8x11 checker boards, and the size of every square is 20cmx20cm. It is printed with a high-quality printer. In theory, three images positioned in three different directions can satisfy the requirement, in order to get better solution of intrinsic and extrinsic parameter, 6 images are used to calibrate the left and right cameras. Figure 3 shows the images and its comers acquired by the left camera.

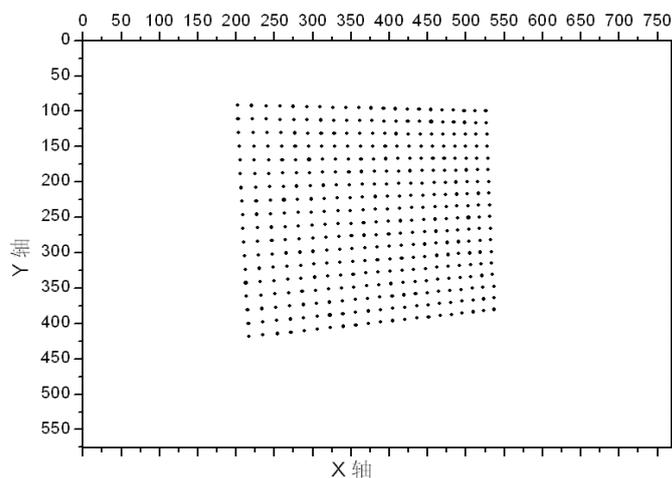


Figure 3. Corner point extraction of calibration image

Using Zhang's nonlinear optimization technique based on the maximum likelihood criterion; we can get the two camera parameters. Table 1 describes the left and right parameter.

Table 1. The Intrinsic Parameter of the Left and Right Camera

	Left camera(mm)		Right camera(mm)	
Focal length	1002.406	1034.403	996.238	1028.667
Principal point	358.465	321.970	376.992	315.856

## 2.2. Relative Orientation Solution of Two Cameras

When intrinsic and extrinsic parameters of matrix are obtained, the following method is to obtain the relationship of position and orientation about two cameras. The orientation relation for two cameras can be derives as follows:

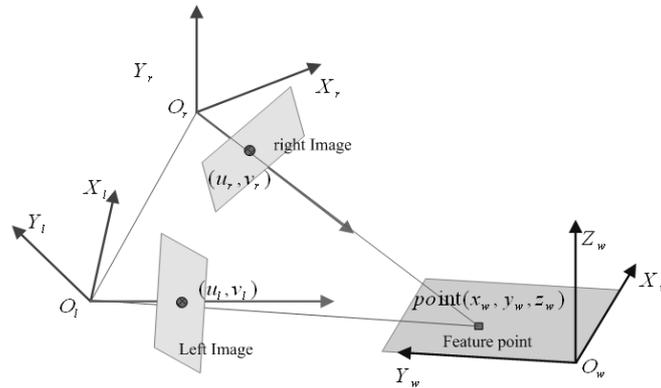


Figure 4. Location of feature point based on binocular vision

Supposing there is a feature point  $P$  in space as shown in Figure 4, the relation about the feature point in left and world coordinate system can be expressed as:

$${}^L P = {}^L L_w \bullet {}^W P \tag{2}$$

Where  ${}^L P$  is the spacial coordinate between the left camera coordinate system,  ${}^W P$  is the spacial coordinate in the world coordinate system,  ${}^L L_w$  is the relative orientaiton about the world coordinate system and left camera coordinate system.

Similarly, the relation about the feature point between right and world coordinate system can be expressed as:

$${}^R P = {}^R L_w \bullet {}^W P \tag{3}$$

By (2) and (3) the equation can be gotten,

$${}^L L_T^{-1} {}^L P = {}^R L_T^{-1} {}^R P$$

That is,

$${}^L P = {}^L L_T {}^R L_T^{-1} {}^R P$$

Therefore, the rotation and translation matrix about the cameras can be expressed as:

$${}^R L_T = {}^R L_w {}^L L_w^{-1} \tag{4}$$

Table 2. Orientation matrix about two cameras for four calibration experiments

0.89578	0.014684	-0.44426	551.51	0.89698	0.013141	-0.44187	549.48
-0.0059524	0.99976	0.021043	3.5124	-0.0021422	0.99968	0.025379	-3.7665
0.44446	-0.016205	0.89565	151.8	0.44206	-0.021819	0.89672	149.72
0	0	0	1	0	0	0	1
0.89849	0.017266	-0.43866	544.3	0.89594	0.017502	-0.44384	551.13
-0.0087504	0.99973	0.021427	2.6137	-0.0057056	0.99959	0.027899	-5.9424
0.43891	-0.015413	0.8984	146.4	0.44415	-0.022463	0.89567	150.26
0	0	0	1	0	0	0	1

Considering the orientation relation between two cameras didn't changed during the camera calibration process, thus  ${}^R_L T$  should be constant matrix, but experiments results showed that the orientation matrix is not entirely consistent by four calibration experiments. In theory, the two cameras relative orientation remained unchanged after completion of the package, Table 2 is the orientation matrix for four calibration experiments.

From the calibration matrix we could conclude that calibration model of binocular vision is limited by the experimental. The limitation was caused by the detection error of corner point and the vibration of the camera, which affects the small changes of calibration parameter. To improve calibration accuracy, the least-squares method can optimize calibration matrix to obtain the optimal parameters, the least-squares can be expressed as:

$$\begin{cases} \min \sum_{k=1}^K \left\| {}^R_L t_k - t_{optimal} \right\|^2 \\ \min \sum_{k=1}^{K=4} \left\| R_{optimal} - {}^R_L R_k \right\|^2 \end{cases} \quad (5)$$

The rectified orientation matrix is,

$$T_{optimal} = \begin{bmatrix} 0.8968 & 0.015648 & -0.44216 & 549.11 \\ -0.0056376 & 0.99969 & 0.023937 & -0.8957 \\ 0.44239 & -0.018975 & 0.89661 & 149.54 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

### 3. Marker Tracking and Location

In order to realize real time and robust tracking of markers, each camera should be processed at least more than 25HZ and image resolution should be less than one pixel. Although using stereo vision contributes to high and robust tracking, but these requirements need a computational cost. To overcome this problem, firstly we predict the region of interest. The procedure of tracking is implemented as follows:

- a) Grab the frame
- b) Threshold the Frame into binary image
- c) Filtering and labeling
- d) Find the center of marker

Connected component labeling algorithm is used to extract the blob region and canny edge detection is employed to get contour line and extract blob area. For every blob we can give two thresholds about area to discard the blobs with less than one pixel and more than fifty pixels. If the number of every image is more than three, we use the roundness of the blob to get the best three blobs. After that, geometric center as criteria is used to discover every blob's center, Figure 5 shows two images of markers acquired by the left camera and right camera. When knowing the relative orientation matrix about two cameras and image point of markers in the two image coordinates, the next question is how to extract the spatial point in the camera coordinate system from the image coordinate system.



Figure 5. Marker image in the left and right cameras

In previous studies, the commonly triangulation method is used to calculate 3D coordinate of markers according to the corresponding left image points coordinates  $(u_l, v_l)$  and right image points coordinates  $(u_r, v_r)$ . For simplified calculation in the solution process, two camera coordinate systems normalized into the same coordinate system is necessary. The intersection spacial point of two straight lines in the same coordinate system can be obtained. The image point of left camera coordinate system in the right coordinate system is as follows:

$$V_{L-R} = {}^R_L T \cdot (A_L^{-1} \cdot [u_l \quad v_l \quad 1]^T) \quad (6)$$

Where,  $A_L^{-1} \cdot [u_l \quad v_l \quad 1]^T$  represents image point of markers in the left image coordinate transforming into the left camera coordinate,  $V_{L-R}$  indicates that image points in the left camera coordinate are transforming into the right camera coordinate system. So the whole problem is converted into solving intersecting point of the two straight lines in the same coordinate system. Taking into account the actual measurement, camera calibration error and extraction errors of the image points will result in no intersection about two straight spacial lines, the public vertical line of two spacial line are solved to define the intersecting point which is defined by the the center of public vertical line.

#### 4. Solving the Pose of Rigid-body

The above discusses how to solve the problem of spacial location of marker, next we'll discuss how to use marker spatial location information to solve rigid-body position and orientation changing. It is not robust using the direct linear solution method to solve because markers position errors would influent the pose accuracy. In practical applications, it is necessary to find a more robust pose optimization algorithm for the solution using optical tracking marker location information. As for the location information with errors, the optimization method of principal component analysis will reduce the impact on position and orientation.

For the problem of pose estimation, markers can be elaborated as follows: For a given three-dimensional data points  ${}^1y_i$  and  ${}^2x_i$ , where  $i = 1, 2, \dots, m$ , they respectively represent coordinate values of one data point in two coordinate systems, the rotation matrix and translation matrix expressing transformation relation from  ${}^2y_i$  to  ${}^1x_i$ , respectively is  ${}^1R$  and  $t$  and therefore:

$${}^2y_i = {}^1R \cdot {}^1x_i + t \quad i = 1, 2, \dots, m \quad (1)$$

For this equation, in previous studies, there are several methods for solving such as linear least square method, singular value analysis, quaternion, etc. All these methods share a common solution, i.e. the solution is divided into two steps, the first step is to use a given two pairs of points series to solve rotating matrix, and then solve the translation matrix. Some of orientation estimation algorithms are used such as least-squares method and principal component analysis [6]. Here, least squares optimization method combined with principal component analysis are employed to solve the rigid-body pose, the whole solution process can be expressed as the following equation:

$$\mathcal{E}(R, t) = \sum_{k=1}^K w_k \|y_k - R x_k - t\|^2 \quad (7)$$

Where,  $k$  is the number of markers,  $y_k$  is coordinate value of No. k marker in the global coordinate system (right camera coordinate system),  $x_k$  is coordinate value of No. k marker in the local coordinate system (rigid-body coordinate system),  $w_k$  is the weighting factor for the

rigid-body, all the weight coefficients are set to 1. Error matrix can be written as:

$$\varepsilon(R, t') = \sum_{k=1}^K \left\| \tilde{\mathbf{y}}_k - R\tilde{\mathbf{x}}_k - t' \right\|^2 \quad (8)$$

Where,  $\tilde{\mathbf{x}}_k = \mathbf{x}_k - \mathbf{x}$ ,  $\tilde{\mathbf{y}}_k = \mathbf{y}_k - \mathbf{y}$ ,  $\mathbf{y}$  and  $\mathbf{x}$  is respectively center point of all markers in the global coordinate system and local coordinate system.

According to equation (8), error matrix can be expressed as:

$$\varepsilon(R, t') = \sum_{k=1}^K (\tilde{\mathbf{y}}_k - R\tilde{\mathbf{x}}_k - t')(\tilde{\mathbf{y}}_k - R\tilde{\mathbf{x}}_k - t')$$

And it can be simplified as:

$$\varepsilon(R) = \sum_{k=1}^K (\tilde{\mathbf{y}}_k^T \tilde{\mathbf{y}}_k + \tilde{\mathbf{x}}_k^T \tilde{\mathbf{x}}_k - 2\tilde{\mathbf{y}}_k^T R\tilde{\mathbf{x}}_k) \quad (9)$$

Minimization of function  $\varepsilon(R)$  is equivalent to maximization of function  $f(R)$  which can be expressed as:

$$f(R) = \sum_{k=1}^K \tilde{\mathbf{y}}_k^T R\tilde{\mathbf{x}}_k \quad (10)$$

The above equation can be solved by singular value decomposition and be written as:

$$f(R) = \text{Trace}\left(\sum_{k=1}^K \tilde{\mathbf{y}}_k^T R\tilde{\mathbf{x}}_k\right) = \text{Trace}(RH) \quad (11)$$

Where,  $H = \sum_{k=1}^K \tilde{\mathbf{x}}_k \tilde{\mathbf{y}}_k^T$ , if  $H$  can be broken into  $AA^T$ , in which  $\text{Trace}(RH)$  is maximum value of the function. using the singular value decomposition  $H$  can be expressed as,  $H = U\Gamma V^T$ . where,  $U$  and  $V$  are for orthogonal matrix.  $\Gamma$  is a non-negative diagonal matrix expressed as:

$$\begin{aligned} \Gamma &= \begin{bmatrix} \gamma_1 & 0 & 0 \\ 0 & \gamma_2 & 0 \\ 0 & 0 & \gamma_3 \end{bmatrix} \\ &= \begin{bmatrix} \sqrt{\gamma_1} & 0 & 0 \\ 0 & \sqrt{\gamma_2} & 0 \\ 0 & 0 & \sqrt{\gamma_3} \end{bmatrix} \begin{bmatrix} \sqrt{\gamma_1} & 0 & 0 \\ 0 & \sqrt{\gamma_2} & 0 \\ 0 & 0 & \sqrt{\gamma_3} \end{bmatrix} \\ &= CC^T \end{aligned}$$

Supposing  $X = VU^T$ , then:

$$\begin{aligned} XH &= VU^T U\Gamma V^T = V\Gamma V^T \\ \text{So } A &= VC, \quad XH = AA^T. \end{aligned}$$

Where,  $XH$  is a positive symmetric matrix, when  $X = VU^T$ ,  $\varepsilon(R, t)$  is the smallest, so the solution has been a rotation matrix  $R$ , as for  $VU^T$  the orthogonal matrix, so  $R$  can be expressed as:

$$R = V \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \det(VU^T) \end{bmatrix} U^T \quad (12)$$

So the translation matrix is  $t = \mathbf{y} - R\mathbf{x}$

### 5. Pose Error Analysis

For pose estimation, it is the key problem about how to determine the pose estimation error because of spacial position error of markers. Here we can express the orientation error as a micro-change for the real position and orientation. Given a real rotation matrix  $R$  and translation matrix  $T$ , the micro-variable rotation  $\Delta R$  and translation matrix  $\Delta T$ , the relation between real matrix and micro-variable matrix can be described as:

$$\begin{aligned} R_{err} &\approx \Delta R R \\ T_{err} &\approx \Delta T + T \end{aligned}$$

Because  $R = VDU^T$ ,  $U, V$  is the matrix of the SVD solution. In order to determine  $\Delta R$  and  $\Delta T$ , we must know marker position error how to affect matrix  $H$ , the error matrix  $\Delta V$  and  $\Delta U$  is respectively applied to the matrix  $U, V$ , the change of matrix  $U, V$  can expressed as:

$$\begin{aligned} V &\rightarrow V + \Delta V \\ U &\rightarrow U + \Delta U \end{aligned}$$

We defined two matrices  $A$  and  $B$  in order to express this transformation:

$$\Delta V \approx AV, \quad \Delta U \approx BU$$

Given that these parameters  $V, U, D$  are from markers and related information to determine the error, and  $R_{err}$  can be expressed as:

$$\begin{aligned} R_{err} &\approx \Delta R R \\ &\approx (\mathbf{V} + \Delta \mathbf{V}) \mathbf{D} (\mathbf{U} + \Delta \mathbf{U})^T \\ &\approx (I + A) R (I - B) \end{aligned}$$

Taking into account the matrix exponential can be carried out using Taylor expansion approximation:

$$e^A \cong 1 + A + \frac{1}{2}A^2 + \frac{1}{6}A^3 + \dots,$$

The first-order approximation can be written as:

$$\begin{aligned} R_{err} &\approx e^A R e^{-B} \\ &= e^{(I+A)} R e^{(I-B)} \\ &= e^A R e^{-B} I = \underbrace{e^A R e^{-B} R^T}_{\Delta R} R \end{aligned} \quad (13)$$

Therefore, the rotation error  $\Delta R$  can be expressed as"

$$\Delta R = e^A R e^{-B} R^T$$

This  $\Delta R$  effectively describe the impact about spatial position error of markers to orientaiton errors for rigid-body.

## 6. Simulation analysis

This research makes use of matlab software to simulate position error of markers how to affect orientation of rigid-body, and employees the least-squares method to fit the entire data. Gaussian noise is added to the marker point in the local coordinate system, and the variance  $\sigma$  is used to determine the error characteristics of markers. The range of variance  $\sigma$  is 0-3mm and the mean is zero. Assumption  $R_{err}$ ,  $T_{err}$  is the expected error matrix of local coordinates according to the above statement, average error matrix  $E$  in the tracking coordinate system can be expressed as:

$$E = \frac{1}{K} \sum_{k=1}^K \|y_k - R_{err} x_k - T_{err}\|$$

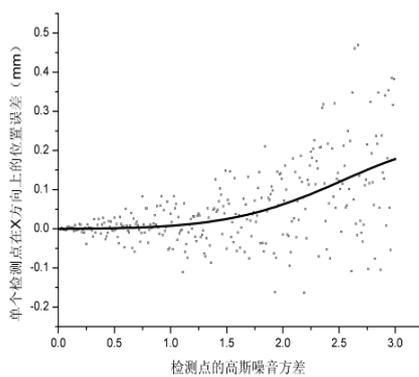


Figure 6(a). Position error of single marker in X direction

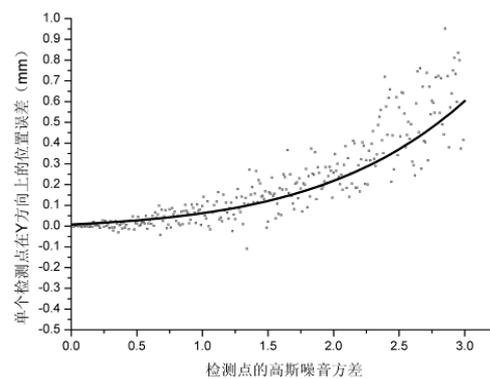


Figure 6(b). Position error of single marker in Y direction

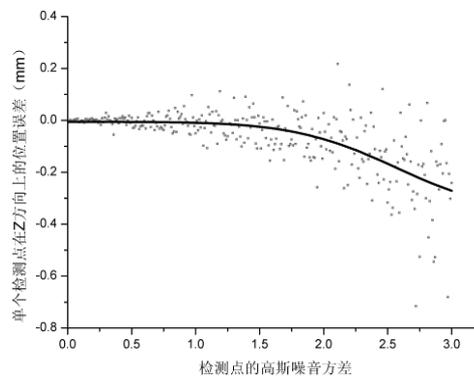


Figure 6(c): Position error of single marker in Z direction

When the range of the variance is 0-3mm, we simulate 1000 times and use the average error to analyze and predict the relationship between pose error and gaussian noise. Simulated

probe has three marker points, the curve relationship between position error of single marker and variable  $\sigma$  is shown in Figure 2. As expected, when the variance  $\sigma$  increases, position error is also on the increase. Error distribution of the marker is different for a given point in different directions when error variance  $\sigma$  is changing which respectively is 0~1.8mm in the x axis, 0 ~ 0.61mm in the y axis and 0.27 ~ 0mm in the z axis. In order to describe better  $\Delta R$  changing, quaternion method is used to represent pose error as in the above studies  $R_{error} = \Delta RR$ , therefore  $\Delta R = RR_{error}^{-1}$ , so the corresponding quaternion can be expressed as follows,  $\Delta q = qq_{error}^{-1}$ , which is angle error as well as including rotation angle error of the rotation axis. Figure 7 shows the relationship between angle error of rigid-body and the detection position errors of three marker points, the fluctuation range of the error is also growing, when the location error of the point mark satisfied with Gaussian noise variance are changed from 0~3 mm, its pose changes is in the angle range -0.59 - 0.41 degree.

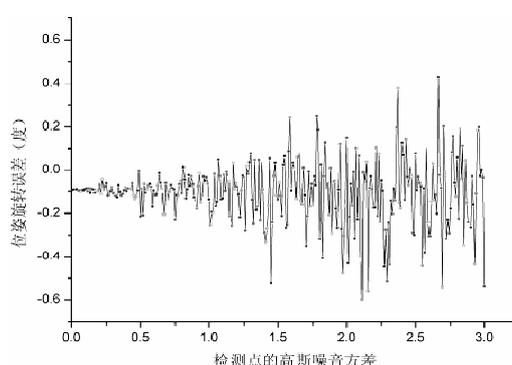


Figure 7. Relationship between position variance and pose changes

## 7. Conclusion

In this paper, the position and pose relation about two cameras are derived by camera calibration experiments using the intrinsic and extrinsic parameters of the two cameras. Due to the detecting error of image corner, the orientation relationship obtained by each image group is not consistent, the optimized orientation parameters are acquired by optimizing the obtained orientation matrix of multi group images. and use the same coordinate system to solve the problem of spacial points. The least squares method combined with principal component analysis method are employed to analyze work-piece orientation and set up the mathematical model of orientation error propagation. Simulation analysis established the relation about the position error of markers and pose error of rigid-body and verified that it is feasible and effective, it provides a theoretical analysis and experimental basis about the pose detection of rigid body such worke piece and medical instrument.

## Acknowledgements

This project is supported by Natural Science Foundation of Ningbo City (2013A610048)

## References

- [1] Agustin Navarro, Edgar Villarraga, Joan Aranda. Relative Pose Estimation of Surgical Tools in Assisted Minimally Invasive Surgery. *Pattern Recognition and Image Analysis*. 2007; 4478: 428-435.
- [2] Vinas FC, Zamorano L, et al. Application accuracy study of a semipermanent fiducial system for frameless stereotaxis. *Computer Aided Surgery*. 1997; 2(5): 257-263.
- [3] Fabrice C, Lavalley S. *Experimental Protocol for Accuracy Evaluation of 6-d Localizers for Computer-Integrated Surgery: Application to Four Optical Localizers*. Lecture Notes In Computer Science. 1998; 1496: 277-284.
- [4] Micha Hersch, Aude Billard, Sven Bergmann. Iterative Estimation of Rigid-Body Transformations.

- Journal of Mathematical Imaging and Vision*. 2011; 43(1): 1-9.
- [5] Yi Xinhua, Wang Mingjun, Cheng Xiaomin. Deformation sensing of colonoscope on FBG sensor net. *Telkomnika Indonesia Journal of Electrical Engineering*. 2012; 10(8): 2253-2260.
- [6] Arun K, Huang T, Blostein S. Least-squares fitting of two 3-D Point Sets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1987; 9(5): 698–700.
- [7] SG Kong, J Heo, BR Abidi. Recent advances in visual and infrared face recognition review. *Computer Vision and Image Understanding*. 2005; 97(1): 103 {135.
- [8] Woltring H, Huiskes R, De Lange A, et al. Finite centroid and helical axis estimation from noisy landmark measurements in the study of human joint kinematics. *Journal of Biomechanics*. 1985; 18(5): 379–389.
- [9] Morris T, Donath M. Using a maximum error statistic to evaluate measurement errors in 3d position and orientation tracking systems. Presence: Teleoperators
- [10] Zhang ZY. A flexible new technique for camera calibration. Technical Report MSR-TR-98-71, 1998.
- [11] Jie Chen. The Comparison and Application of Corner Detection Algorithms. *Journal of multimedia*. 2009; 4(6): 435-441.