

Pose Estimation of 3C Components Based on Monocular 2D Camera

Qingzhi Mou, Xiansheng Yang, Han Zhang and Yunjiang Lou, *Senior Member, IEEE*

Abstract—In order to accurately and effectively estimate the pose of 3C product components in the manufacturing process and meet the needs of intelligent production, this paper proposes a complete 3C component pose estimation technology scheme based on the Perspective n-Points algorithm, and also analyzes and verifies the impact of errors in theoretical derivation and experiment. Showing in the framework, firstly, we compared common monocular camera pose estimation algorithms to choose the suitable one, identified and analyzed the factors affecting the accuracy of pose estimation, then the whole design of marker points is discussed and the corresponding image detection and localization algorithm is optimized. After completing the above theoretical analysis and realization, an accomplished technical work-flow is formed. Through 3D software simulation and actual experiment, the effectiveness and accuracy of the pose estimation scheme in monocular camera case are verified, and the theoretical basis for realizing the higher precision in multi-camera case is also provided.

Index Terms—error analysis, target design and optimization

I. INTRODUCTION

3C is the short name of computer, communication and consumer electronic products. With the fierce market competition and the improvement of product precision, the requirement of automatic 3C assembly becomes urgent. However, assembly action is different from planar applications such as sorting, packaging and defect detection, it needs to consider six-dimensional information (position and orientation) of components in space. Traditional 2D positioning scheme is no longer applicable. In order to provide accurate pose feedback information for intelligent assembly system, it is necessary to evaluate, analyze, validate and optimize the performance of pose estimation schemes, so as to provide a basis for later fusion to achieve high-precision pose estimation[1].

According to the measurement principle, pose estimation can be divided into vision and other physical sensors such as infrared, acoustic and laser. Zhang et al. have studied the object's pose measurement by using laser tracker[2], which is accurate, efficient but extreme high-cost. With the advantages of low cost, stable performance and easy to transplant, vision-based pose estimation has becomes a research hot-spot.

This work was supported partially by the NSFC-Shenzhen Robotics Basic Research Center Program (No. U1713202) and partially by the Shenzhen Science and Technology Program (No.JCYJ20180508152226630).

Mou, Yang, Zhang and Lou are with the School of Mechatronics Engineering and Automation, Harbin Institute of Technology Shenzhen, HIT Campus, University Town of Shenzhen, Xili, Nanshan, Shenzhen, China. (louyj@hit.edu.cn (Lou))

According to the type of visual system, visual measurement can be divided into monocular vision, binocular vision, multi-visual vision and projective depth vision shown as Fig. 1[3]. The monocular vision system is simple in structure, easy in calibration and low in cost but the geometric information of the object must be known to reproduce the depth. Binocular vision, i.e. Stereoscopic vision, can directly obtain geometric parameters but the stereo matching process is instable and time-consuming. The multi-vision system is so complex that it only used for 3D reconstruction in usual. The projective depth vision is easy to use but sensitive to noise and inaccurate[3-5]. Based on the controllability and weak texture characteristics of 3C product parts, as well as the requirement of high precision and efficiency of industrial production line, we adopt monocular vision measurement based on artificial marking features as the implementation scheme.



(a) Laser tracker (b) Monocular-vision (c) Multi-vision
 Fig. 1: Different types of pose estimation systems

Considering the monocular camera with artificial marker pose estimation method, the summary of predecessors' work is described as below. At the large workspace level, Brent E. Tweddle et al. in MIT discussed the design of pose estimation target in satellite system and proposed some non-linear optimization methods[6]; Matthias Faessler et al. proposed a new method of pose estimation system based on infrared LED, which has been applied to UAV[7]. At small workspace level, Yang et al. used LSD detection and K-means clustering to achieve the cooperative goal of manipulator grasping[8]; Liu et al. in SJTU used SIFT + EPnP method to estimate the pose of workpiece to assist assemble through VR[9]. In order to achieve human-guided robot assembly system, Wan et al. in AIST used AR marker with monocular camera to obtain the workpiece's pose and reproduced in robot operation[10].

In large workspace, especially for spacecraft aspect, the artificial features usually are complicated and can not be directly adapted to the assembly workpiece. However, in small workspace, the research is not enough to analyze the factors affecting the accuracy, nor does it make a deep study on target synthesis design, which results in large actual error of pose estimation and can only be applied to grasp and

assemble with low precision.

In this paper, a complete monocular vision pose estimation scheme suitable for 3C devices in small workspace is proposed and implemented, and its performance is analyzed and verified through simulation and experiment. The structure of this paper is as follows: In Section II, the common algorithms of pose estimation based on monocular vision are briefly introduced and compared, and the factors affecting the accuracy are discussed. In Section III, we discussed the marking target design and optimization including shape, size, etc. The corresponding detect and center point positioning algorithms scheme are also proposed and verified, and target perspective distortion is also considered. In Section IV, we designed the experimental process, using 3D software for simulation and actual platform experiments to verify the accuracy and effectiveness of the entire process in both theoretical and practical environments. Our contribution is proposed and verified a complete process of monocular camera for high precision pose estimation in small space, and to provide a basis for the next step of multi-2D camera fusion to achieve higher precision pose estimation.

II. ALGORITHM SELECTION & SENSITIVITY ANALYSIS

In this chapter, we first discuss and analyze the commonly used pose estimation methods, and simulate them simply to choose suitable one. Then the factors affecting the accuracy of pose estimation are considered and analyzed, the conclusion is drawn to guide the design of the system.

A. Pose Estimation Algorithm in Monocular Camera

1) Pose Estimation Based on Line Features

In the research of pose estimation, the methods which calculate pose by object's line features are called PnL methods, just like Pan et al. studied[11]. These methods have strong anti-interference due to line's characteristics. But when compared with the point feature, their intrinsic accuracy is poor in the magnitude level, so it is not suitable for high-precision pose estimation.

2) Pose Estimation Based on Point Features

The algorithm based on point feature can be divided into PosIT and PnP. PosIT uses affine projection model to iteratively calculates the pose without camera intrinsic parameters, which has limitations in accuracy. The PnP algorithm, which need use camera internal parameters, can be divided into iterative and non-iterative. Non-iterative algorithm uses algebraic algorithm to solve pose directly, which has less time cost but high instabilities. EPnP is recognized as a mature algorithm in this class[12]. Iterative algorithm expresses PnP as a constrained non-linear optimization problem to obtain the numerical optimization solution, which has high accuracy and stability but calculation consuming[12]. We select PnP/EPnP/Iterative-PnP for simulation, and consider the accuracy and time factors to complete the selection of the algorithm. The result is shown in Fig. 2.

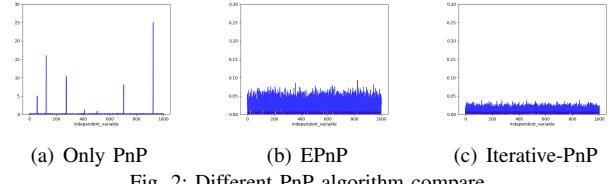


Fig. 2: Different PnP algorithm compare

It can be found that the iterative-PnP algorithm has the highest accuracy (Coordinate axis scales of (a) and (b)/(c) are different, unit both are mm), and 1000 iteration times $t < 0.3$ s, which meets requirements of industrial non-real-time pose estimation, so it is chosen as the experimental algorithm.

B. Sensitivity Analysis

Through the control variable and decouple method, the factors affecting the result of position and orientation are theoretically analyzed and simulated. Lu et al. have made a preliminary study on position aspect[13], and this paper has carried out a detailed analysis in considering more factors case. Zhu et al. discussed the factor influenced orientation[14] and we analyzed by simulation and proved them by experiments. It can be concluded that: in position aspect, the camera focal length calibration error is the primary factor affecting the depth measurement accuracy, and the pixel extraction accuracy is the primary factor affecting the plane measurement accuracy; in the aspect of orientation, the main factor affecting roll and pitch/yaw direction is the accuracy of pixel extraction, but the effect of the same error on roll direction is far less than that of pitch/yaw direction.

We give a brief analysis of the position accuracy as example.

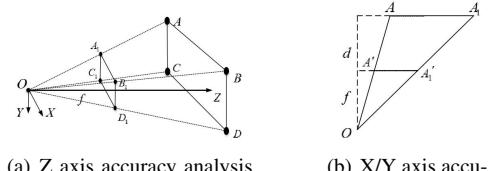


Fig. 3: PnP error analysis model

1) Considering Non-focal Length Calibration Error

The schematic diagram of Z axis accuracy is shown in Fig. 3 (a), The optical center of the camera is O . The feature points are $A \sim D$, and their imaging on a plane are $A_1 \sim D_1$. The distance between feature points is $l_0 = l_{AB} = l_{BC} = l_{CD} = l_{DA}$. The distance from the camera's optical center to the target is d , then ∂d is the accuracy of Z-direction position measurement. The number of pixels between the feature points is n , and the precision of the proposed feature point location is $\partial n = 0.2$ pixel. Pixel-unit resolution is $\delta = \mu d / f$.

According to the principle of camera imaging:

$$d = \frac{l_0 f}{n \mu} \quad (1)$$

Find partial derivatives from d to n , get:

$$\frac{\partial d}{\partial n} = -\frac{l_0 f}{\mu} \cdot \frac{1}{n^2} = -\frac{\mu d^2}{l_0 f} \quad (2)$$

As can be seen from the above formula, for specific cameras, given d and l_0 , the accuracy of Z-direction position measurement can be obtained.

The schematic diagram of X/Y axis accuracy is shown in Fig. 3 (b), target feature point A moves l_0 to A_1 and imaging point from A' to A'_1 . According to Formula (1)'s deformation, find partial derivatives from l_0 to n , get:

$$\frac{\partial l_0}{\partial n} = \frac{d\mu}{f} \quad (3)$$

For specific cameras, given d , the accuracy of X/Y-direction position measurement can be obtained.

2) Considering Focal Length Calibration Error

In the practical manufacturing experience of industrial cameras, there is a manufacturing error of about 0.25% in lens focal length, means that $f_{true} \neq f_{actual}$, which needs to be taken into consideration in high precision pose estimation in small workspace. According to the imaging principle, for the direction, there should be:

$$d_{true} = \frac{l_0 f_{true}}{n\mu} \quad (4)$$

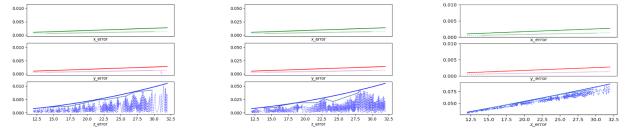
Then we have:

$$\Delta d = \frac{l_0}{n\mu} (f_{true} - f) = \frac{d_{true}}{f_{true}} (f_{true} - f) = d_{true} \cdot \frac{\Delta f}{f_{true}} \quad (5)$$

$$\Delta l_0 = \left(\frac{d_{true}}{f_{true}} - \frac{d_{true}}{f} \right) \mu \partial n = d_{true} \mu \partial n \cdot \frac{\Delta f}{f * f_{true}} \quad (6)$$

As (5) and (6) shows, the error of focal length calibration will significantly affect the accuracy of Z-axis direction estimation and the influence degree has a linear relationship with the actual work distance d_{true} . While in X/Y-axis part, the error of focal length calibration has no obvious influence on the accuracy of X/Y-direction estimation as $\mu \partial n \frac{\Delta f}{f * f_{true}}$ is so tiny that it can be negligible.

We set up a group of reference groups and verify the correctness of the theory through the simulation test of the control variable method. The default values for each variable parameter are as follows: $pixel_resolution = 0.2\ pixel$, $target_size = 8\ mm$, $d = 12 \sim 32\ mm$ (as abscissa axis), $f_{ideal} = 8\ mm$, $f_{true} = 8.02\ mm$. We changed the variables to simulate. The non-focal calibration error test results are shown as Fig. 4(a)(b) and focal calibration error test result is shown as Fig. 4(c). The red/green/blue color line means theoretical calculation curve and the light red/green/blue broken segment means maximum error fitting line (In particular, note that the theoretical calculation curve of Z direction in Fig. 4(c) is a simple superposition of two error factors after neglecting the minimum term).



(a) 0.2 pixel resolution (b) 1 pixel resolution (c) With focus error
Fig. 4: Effect of Pixel Extraction/Focal Error on Pose Estimation Accuracy

As shown in the Fig. 4, without considering the focal length calibration error, the pose estimation error caused by the extraction error of feature points is linear in the X/Y direction and quadratic in the Z direction. After considering the focal length calibration error, the estimation error in X/Y direction is not significantly affected, while the estimation error in Z direction can be simply regarded as a superposition of linear and quadratic errors dominated by linear terms. These simulation results validate our conclusions and can be used to guide the subsequent target design.

In orientation aspect, according Zhu's analysis[14], we analyzed the influencing factors when the target surface is parallel to the camera plane and obtained the analysis that the main factor affecting roll and pitch/yaw direction is the accuracy of pixel extraction, but the effect of the same error on roll direction is far less than that of pitch/yaw direction. Results are shown in the TABLE I.

TABLE I: Orientation sensitivity analysis

Orient axis	Influence factor	Source error	Orient error
yaw\pitch	pixel extraction error ($\Delta x/\Delta y$)	0.2pixel	0.165
roll	pixel extraction error ($\Delta x/\Delta y$)	0.2pixel	0.051
yaw\pitch	focus calibration error (%)	0.25%	1.387e-4
roll	focus calibration error (%)	0.25%	6.131e-7

III. TARGET DESIGN & DETECTION

High precision pose estimation by PnP algorithm requires excellent target design and accurate feature point center extraction. In this chapter, feature points' design is discussed and analyzed, and the detection and central extraction algorithm are experimented and optimized.

A. Shape of Feature Points

1) Feature Points Based on One-dimensional Projection

Square to vertex feature point is shown in Fig. 5(a), which has good precision and is not affected by inherent perspective projection distortion, but its encoding ability is weak and greatly affected by camera lens distortion. It is mostly used in the form of checkerboard to calibrate the camera[15].

2) Feature Points Based on Two-dimensional Projection

Two-dimensional feature points are normally classified as circle/ring, rectangle and cross. For these feature points, they will be all affected by perspective distortion in small distance. Circle/ring can be compensated by mathematical method and rectangle/cross can be compensated by diagonal intersection. Therefore, in this part, the comprehensive performance of non-distortion situation is considered first and distortion situation will be analyzed below. The circle/ring

has affine invariance and concise features, high positioning accuracy[16]; for rectangle, high accuracy can be achieved when determining the center through the intersection of diagonals, but pseudo-rectangle detections cause insufficient robustness[17]; for cross, high accuracy can be achieved when the feature points area ratio is high (*over*50%), but it's coding capacity is insufficient and the research of localization algorithm is not enough[18]. The two-dimensional feature points is shown in Fig. 5(b)(c)(d).

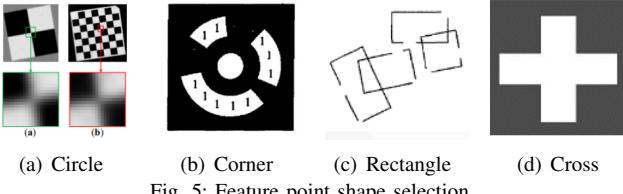


Fig. 5: Feature point shape selection

According above review, considering positioning accuracy, recognition stability, coding ability and anti-occlusion performance, we consider concentric circles as feature points for recognition.

B. Feature Point Center Detection

After selecting the shape, it is need to detect the feature points and extract their centers. We divide it into three parts: region extraction, sub-pixel interpolation and precise center fitting. The algorithm flow is shown as Algorithm 1.

Algorithm 1 Feature Point Center Detection Process

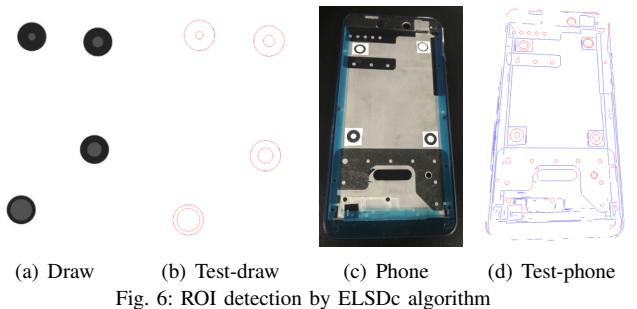
```

Region of Interest(ROI) Detection (ELSDc)
if ROI Number = Feature Points Number then
    for each ROI in set{ROIs} do
        Edge Sub-pixel Interpolation (Zernike Moment)
        Fitting Ellipse Center (Least Square)
    end for
    Use Rings' Area to Determine Point Order
    Perspective Projection Distortion Compensation
    Input Information to Pose Estimation
end if
```

1) Feature Points ROI-area Detection

To obtain the point center, it is necessary to roughly recognize the target feature points area first. ELSDc detection algorithm is an ellipse and line detection algorithm, which can obtain reasonable accuracy of ellipse and line segment detection results in a relatively short time. The advantages of this algorithm are fast detection speed, high accuracy and robustness. It is suitable for fast feature extraction and detection[19]. Fig. 6(a)(b) shows the image and test results we used as the positioning accuracy test and Fig. 7(c)(d) prove it's stability in the actual scene.

The deviations between the circle center coordinate detected by ELSDc and the actual value are shown in the TABLE II(GN means gaussian noise), which shows because



(a) Draw (b) Test-draw (c) Phone (d) Test-phone

Fig. 6: ROI detection by ELSDc algorithm

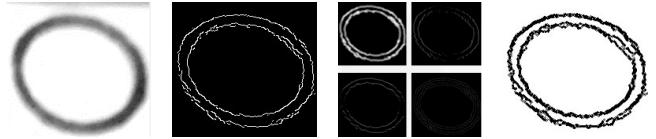
region growing and linking algorithm makes error beyond our requirement.

TABLE II: Centroid coordinate extraction error by ELSDc

Truth	GN0	GN1	GN2	GN4	GN8
1057,513	4.9667	4.9667	4.9639	4.9639	4.9696
2153,601	4.9497	4.9497	4.9497	4.9526	4.9611
2097,2041	4.9582	4.9639	4.9582	4.9639	4.9667
881,3409	4.9441	4.9526	4.9583	4.9469	4.9554

2) Sub-pixel Interpolation Based on Zernike Moment

Sub-pixels are techniques that subdivide pixel units to improve image resolution. After obtaining the approximate target region, sub-pixel extraction of the edge of the target using sub-pixel positioning based on Zernike moment is beneficial to improve the accuracy of the late center fitting [20]. The process is shown in Fig. 7.



(a) ROI circle (b) After canny (c) Zernike convolution (d) Sub-pixel result

Fig. 7: Sub-pixel positioning process

3) Fitting Ellipse Center

After obtaining the sub-pixel coordinates of the target image edge, the final high-precision center coordinates are obtained by using the least square ellipse fitting algorithm[21]. The algorithm minimizes the distance error through specific constraints. The least square method is used to solve the problem, and the optimal fitting center of the ellipse is obtained as shown in Fig. 8. In order to obtain higher precision, it is actually to fit the concentric circles separately.



(a) After sub-pixel (b) Fitting ellipse
Fig. 8: Ellipse fitting process

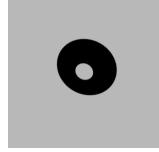
We test the accuracy after implementing all the processes, the result is shown in TABLE III. As can see that the accuracy has been greatly improved(Only if the target is not tilted).

TABLE III: Test result by whole process

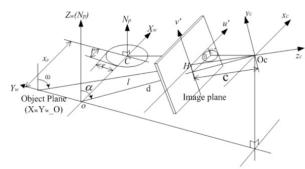
Truth	GN0	GN1	GN2	GN4	GN8
1057,513	0.0025	0.0188	0.0167	0.0065	0.0080
2153,601	0.0104	0.0171	0.0144	0.0142	0.0131
2097,2041	0.0080	0.0168	0.0085	0.0045	0.0150
881,3409	0.0392	0.0384	0.0374	0.0319	0.0318

4) Distortion Characteristics of Two-Dimensional Graphics in Perspective Projection Transform

Because of the inherent characteristics of perspective projection transformation, two-dimensional graphics which are not parallel to the image plane in space will be distorted when projected to the image plane, resulting in the spatial geometric center and the image pixel center no longer coincide, as shown in Fig. 9(a). Dong et al. proposed a method to determine the true projection position of a concentric circular target based on perspective transformation and analytic geometry principle, according to the known radius relationship of the concentric circular target[22]. The perspective projection model is shown in Fig. 9(b). The deduced compensation formula is shown as (7), where $(u_{b1}, v_{b1}, r_1), (u_{b2}, v_{b2}, r_2)$ is the fitting center and true radius of the big circle/small circle in the concentric circle. The simulation results are shown in the TABLE IV. It can be seen that after considering the compensation of the circle center deviation of the perspective projection, the fitting result can be really improved and meet our needs.



(a) Distortion



(b) 2d distortion compensation model

Fig. 9: Distortion phenomenon and compensation model

$$u_c = K_1 \cdot u_{b2} - K_2 \cdot u_{b1} \text{ with } K_1 = \frac{r_1^2}{r_1^2 - r_2^2}, K_2 = \frac{r_2^2}{r_1^2 - r_2^2} \quad (7)$$

$$v_c = K_1 \cdot v_{b2} - K_2 \cdot v_{b1}$$

TABLE IV: Distortion Compensation Effect Test

Distortion compensate	Target size(mm)	Camera distance(mm)	Target tilt angle(deg)	Fitting error(pixel)
False	5/10	160	5°	0.1811
True	5/10	160	5°	0.0075
False	5/10	160	10°	0.2723
True	5/10	160	10°	0.0141

IV. SIMULATION & EXPERIMENT

A. Sketch

The process is divided into two parts: simulation and experiment. The simulation was carried out using Blender software, and experiment was carried out through the platform built. In the simulation, we are fully know the accurate six-dimensional information before and after the object moves, which allows us to better verify the algorithm. In the experiment, due to the calibration error, the moving six-dimensional

information cannot be obtained completely accurately, but in robotic system, the difference between the Euclidean distance of the object movement and the Orientation angle before and after the movement can still be considered as accurate value, we mainly consider the estimation accuracy of the distance and orientation angle difference when the test object moves as the performance standard of the algorithm. The overall verification process design is shown as Algorithm 2.

Algorithm 2 Process Design

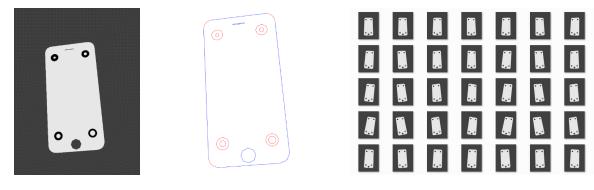
```

Definition:
Initial point position  $P_{init} = init\ point's\ pose$ 
Set different poses from the initial point set{ $P_i$ }
 $p_i = \{x, y, z\}$ ,  $o_i = \{r, p, y\}$ ,  $P_i = \{p_i, o_i\}$ ,  $d_i = |p_i - p_{init}|$ 
In simulation:
for point in set{ $P_i$ } do
    Perform pose estimation{ $p_{i\_est}, o_{i\_est}$ } and compare with
    simulated image reference truth{ $p_{i\_true}, o_{i\_true}$ }
end for
In experiment:
for point in subset{ $P_i$ } do
    Perform pose estimation { $d_{i\_est}, o_{i\_est}$ } and compare with
    robot reference truth{ $d_{i\_true}, o_{i\_true}$ }
end for

```

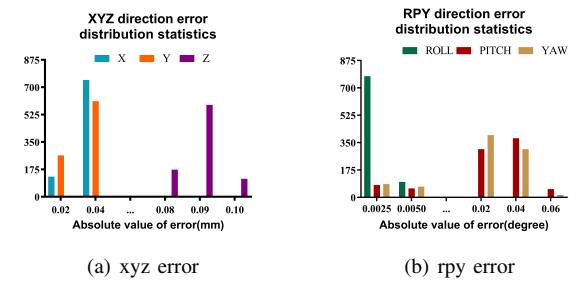
B. Simulation Test

In simulation part, we rendered 875 photos(rotate around each axis from -5° to 5° , at intervals of 2.5° at the original point and points on each coordinate axis and $\pm 10mm$) as Fig. 10, and estimated their positions and orientation to verify non-focal length calibration error situation(focal length calibration error also can be verified but not shown here). The accuracy of the process was observed by comparing information in the 3D software with the estimated values. The results are shown in Fig. 11.



(a) Rendered (b) Processed (c) Test picture set

Fig. 10: Images of simulation process



(a) xyz error (b) rpy error

Fig. 11: Results of simulation

C. Experiment Test

In experiment, we use 20 million resolution camera(Basler acA5472) with 8mm lens(ML-U0817SR-18C) and UR5 robot to finish the actual verification process to verify focal length calibration error situation. Firstly, we carried out static experiments to verify its accuracy in fixed environment and confirmed that its static accuracy meets the requirements. Then in X/Y/Z direction, we choose points 1 ~ 5mm away from respective axis to the origin and take photos. The distance between the feedback current point and the origin point position in UR5 robot is compared with the distance get by the our method, the accuracy of the pose estimation process in each axis is obtained. Then we rotate the RPY angle $-5^\circ \sim 5^\circ$ at the origin respectively, and compare in the same way of position to get the measurement accuracy of the orientation. (Note that the difference between the Z and the X/Y axis is due to the influence of pixel extraction error (including depth of field factor) as Fig. 4(a)(b) on pose estimation. The error caused by camera calibration error factor is difficult to reflect when Euclidean distance is used as the evaluation criterion because of its tiny increment).

The picture of experiment are shown as Fig. 12.

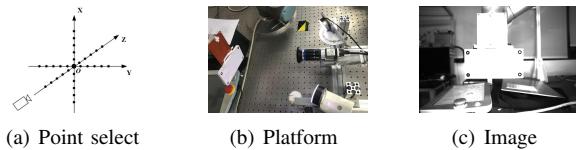


Fig. 12: Images of experiment process

The experiment results are shown in Fig. 13.

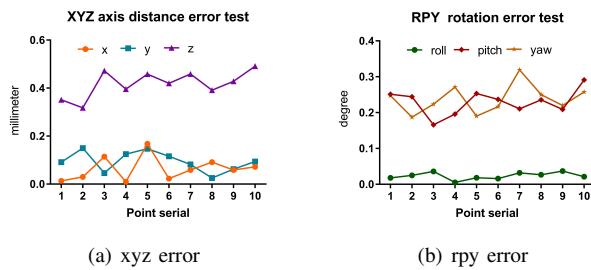


Fig. 13: Results of experiment

The experimental results show that our process has high precision under controllable working conditions.

V. CONCLUSIONS

This paper presented a complete 3C component pose estimation technical solution based on the perspective n-point algorithm. The process firstly identified and analyzed the factors affecting the accuracy of pose estimation, then the parameter of feature points are discussed, corresponding image recognition and localization algorithms are optimized. Finally through 3D software simulation and the actual experiment we verified the effectiveness and accuracy of this scheme in monocular camera case, it also provided the theoretical basis to achieve higher precision in multi-camera data fusion as our future work.

REFERENCES

- [1] Y. Ding, J. Xiao and Kenji Fukasawa, "2D/3D localization and pose estimation of harness cables using a configurable structure representation for robot operations," *U.S. Patent 9 478 035 B2*, Oct 25 2016.
- [2] Z. Zhang, M. Li and Y. Tian, "Research on Real-Time Measurement Method for Objects Space Attitude," *Machine Design and Research*, Vol. 24, No. 5, pp. 83-87, Oct 2008.
- [3] Pérez Luis, et al., "Robot Guidance Using Machine Vision Techniques in Industrial Environments: A Comparative Review," *Sensors*, Vol. 16, No. 3, pp. 335-, Mar 2016.
- [4] Tobias Nöll, Alain Pagani and Didier Stricker, "Markerless Camera Pose Estimation - An Overview," *Visualization of Large and Unstructured Data Sets - Applications in Geospatial Planning, Modeling and Engineering*, Vol. 19, pp. 45-54, 2011.
- [5] A. Tellaeche and I. Maurtua, "6DOF pose estimation of objects for robotic manipulation. A review of different options," in *Proceedings of the 2014 IEEE Emerging Technology and Factory Automation (ETFA)*, Barcelona, 2014, pp. 1-8.
- [6] Brent E. Tweddle and Alvar Saenz-Otero, "Relative Computer Vision-Based Navigation for Small Inspection Spacecraft," *Journal of Guidance, Control, and Dynamics*, Vol. 38, No. 5, pp. 969-978, May 2015.
- [7] M. Faessler, E. Mueggl, K. Schwabe and D. Scaramuzza, "A monocular pose estimation system based on infrared LEDs," in *Proceedings of the 2014 IEEE International Conference on Robotics and Automation (ICRA)*, Hong Kong, 2014, pp. 907-913.
- [8] P. Yang, "Application Research on Monocular Visual Grab Technology of Supply Robot," *M.S. dissertation*, Dept. Computer. Science., Southwest University of Science and Technology Univ., Mianyang, SC, 2018.
- [9] R. Liu, et al., "Research on Combinational Pose Estimation of Base Part and Assembly Part for Augmented Reality Aided Assembly," *Machine Design and Research*, Vol. 34, No. 6, pp. 119-125+137, Dec 2018.
- [10] W. Wan, F. Lu, Z. Wu and Kensuke Harada, "Teaching robots to do object assembly using multi-modal 3D vision," *Neurocomputing*, Vol. 259, pp. 85-93, Oct 2017.
- [11] S. Pan, et al., "Simulation of Visually-guided Autonomous Landing System of Unmanned Helicopter," *Aerospace Control*, Vol. 26, No. 2, pp. 63-67, 2008.
- [12] Sumant Sharma and Simone D'Amico, "Comparative assessment of techniques for initial pose estimation using monocular vision," *Acta Astronautica*, Vol. 123, pp. 435-445, June/July 2016.
- [13] Y. Lv, et al., "Research on Mono-vision Pose Measurement for Space Cooperative Target," *M.S. dissertation*, Dept. Mechanical. Eng., Chinese Academy of Sciences., Changchun, JL, 2018.
- [14] F. Zhu, et al., "Accuracy Analysis of Camera Orientation Calibration Based on P4P Method," *Acta optica sinica*, Vol. 38, No. 11, pp. 241-249, Nov 2018.
- [15] T. Yang, et al., "Sub-Pixel Chessboard Corner Localization for Camera Calibration and Pose Estimation," *Applied Sciences*, 2118, Nov 2018.
- [16] B. Liang, et al., "Sub-pixel location of center of target based on Zernike moment," in *Proceedings of the 2010 Sixth International Symposium on Precision Engineering Measurements and Instrumentation*, 75443A, Dec 2010.
- [17] K. Chen, "Research on Random Hough Transform and Neural Network Method for Rectangular Detection," *M.S. dissertation*, Dept. Computer. Science., Xiangtan Univ., Xiangtan, HN, 2018.
- [18] J. Shi, et al., "Method of Straight Edge Detection with Sub-pixel Localization Based on SUSAN and Hough Transform," *Opto-Electronic Engineering*, Vol. 35, No. 6, pp. 89-94, Jun 2008.
- [19] V. Patraucean, et al., "Joint A Contrario Ellipse and Line Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 39, No. 4, pp. 788-802, 2017.
- [20] B. Wei and Z. Zhao, "A sub-pixel edge detection algorithm based on Zernike moments," *The Imaging Science Journal*, Vol. 61, No. 6, pp. 436-446, 2013.
- [21] A. Fitzgibbon, M. Pilu and R. B. Fisher, "Direct least square fitting of ellipses," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 21, No. 5, pp. 476-480, May 1999.
- [22] H. Dong, et al., "Correction of Circular Center Deviation in Perspective Projection," in *Proceedings of SPIE - The International Society for Optical Engineering*, 8499, Oct 2012.