Uncalibrated Workpiece Positioning Method for Peg-in-hole Assembly Using an Industrial Robot

Ming CONG[1], Fukang ZHU[1], Dong LIU[1*], Yu DU[2]
(1. *School of Mechanical Engineering, Dalian University of Technology, Dalian* 116024;
2. *Dalian Dahuazhongtian Technology Co., Ltd, Dalian* 116024)

Abstract: This paper proposes an uncalibrated workpiece positioning method for peg-in-hole assembly of a device using an industrial robot. Depth images are used to identify and locate the workpieces when a peg-in-hole assembly task is carried out by an industrial robot in a flexible production system. First, the depth image is thresholded according to the depth data of the workpiece surface so as to filter out the background interference. Second, a series of image processing and the feature recognition algorithms are executed to extract the outer contour features and locate the center point position. This image information, fed by the vision system, will drive the robot to achieve the positioning, approximately. Finally, the Hough circle detection algorithm is used to extract the features and the relevant parameters of the circular hole where the assembly would be done, on the color image, for accurate positioning. The experimental result shows that the positioning accuracy of this method is between 0.6-1.2 mm, in the used experimental system. The entire positioning process need not require complicated calibration, and the method is highly flexible. It is suitable for the automatic assembly tasks with multi-specification or in small batches, in a flexible production system.
Key words: Uncalibrated Workpiece Positioning, Industrial Robot, Visual Positioning, Peg-in-hole Assembly

1 Introduction

With the increase of requirements for intelligent and autonomous industrial robots in modern production, robot vision systems have appeared in more and more application scenarios. The traditional methods of teaching and pre-programming could accomplish some simple tasks such as handling and painting, for industrial robots [1]. But it is difficult to use them in flexible production and intelligent manufacturing. Thus, research on robot vision systems has been closely watched by experts and scholars at home and abroad [2-6]. In recent years, the emergence of three-dimensional (3D) data sensors like Kinect with low price and stable performance has led to extensive research on robotic stereo vision systems. Using the depth data in a 3D image, we can more flexibly segment the target object from the complex background, reducing the requirements of an ordinary 2D data sensor for sensing the light and the work area background.

The commonly used "template matching + visual compensation" positioning method [7] can meet the visual recognition and positioning requirements when using robots in a mass production assembly line, but it lacks flexibility and adaptability in the application of small batch and multi-specification, flexible production units. Subsequently, some have used binocular stereo vision [8], structured light [9], and deep learning [10] to guide robots to complete feature recognition and rapid positioning tasks. However, the assembly system will require complex and accurate calibration before use. The required technical level of the system operators is relatively high, and there is a lack of adaptability and flexibility in the actual production applications.

With this backdrop, the development of a robot vision positioning method that is suitable for the automated assembly in a flexible production system can reduce the development and manufacturing costs of special fixtures, and reduce the time consumption of

complex calibration process and machine halt due to working condition change, and improve the flexibility and adaptability of the robot. Such development is of great importance, to broaden the application field of industrial robots, enhance the intelligence level of industrial robots, and improve the production efficiency. In view of the necessity of recognition and positioning for peg-in-hole assembly using an industrial robot in a flexible production system, a novel uncalibrated visual positioning method, which can adapt to the changeable product specification working conditions is proposed in the present paper. It reduces the time consumption in a complex calibration process and process interruption due to working condition change in existing techniques.

2 Visual Positioning Principle

In a vision-based production process, it is very difficult to extract the target workpiece from the complex background of the RGB source image due to the effects of light and color [11]. The depth image is almost not affected by light and color. Therefore, the robustness and adaptability of this process are improved by using t depth images.

The implementation process of the proposed method is schematically presented in Fig.1. First, the depth information of the plane with the target feature is used to segment the global threshold. Meanwhile, the processed image can filter out most of the background interference. Then a connected domain search algorithm based on stroke [12] is used to find out the remaining all connected domain of the depth image, after threshold segmentation. After that, the area of smaller interference noise and the area that does not conform to the shape requirement are filtered out. After these two steps, the target workpiece can be extracted from the image. In order to prevent the use of incorrect target, with undesirable consequences, a pixel projection algorithm [13] is used to investigate whether the target is correctly extracted. Next, from the projection results we can obtain the outer contour's angular point coordinates of the target workpiece. After the rough positioning is completed, the workpiece will be basically in the center of the camera´s field of vision.

At this stage, the color image is in good condition and the number of pixels that occupies the target hole meets the accuracy requirements of the subsequent positioning. Then, Hough circle detection algorithm [14] is used to obtain the center and radius information of the target circular hole, which are used to drive the robot for secondary, accurate positioning.

3 Image Processing and Feature Recognition

3.1 Segmentation of target workpiece area based on depth data

The work environment of a manufacturing system is very complex. In vision-based systems, the camera view contains background interference in addition to the target workpiece. A depth image can be used to overcome this problem. Because a round hole is in a workpiece is viewed in a rectangular plane, global threshold segmentation by using the depth information of the plane as threshold can easily segment the target round hole from the complex background. In this processing, there will be some errors and view offset due to accessibility limitations of the camera depth data. Hence, the plane of the target round hole is not a single value. Here, a range of depth data is used as the depth threshold. The original depth image and the threshold segmentation image are shown in Fig.1 (a) and (b), respectively.

$$\text{depth threshold} \in [700, 730] \quad (1)$$

3.2 Search and filter of connected domain based on arrangements

It can be seen from Fig.1 (b) that background interference is removed is extensively and the target workpiece is fully extracted after threshold segmentation. However, still there are some noise and interference objects, which are in the same depth of the target round hole in the depth image. Hence, we use an connected domain search algorithm to filter the

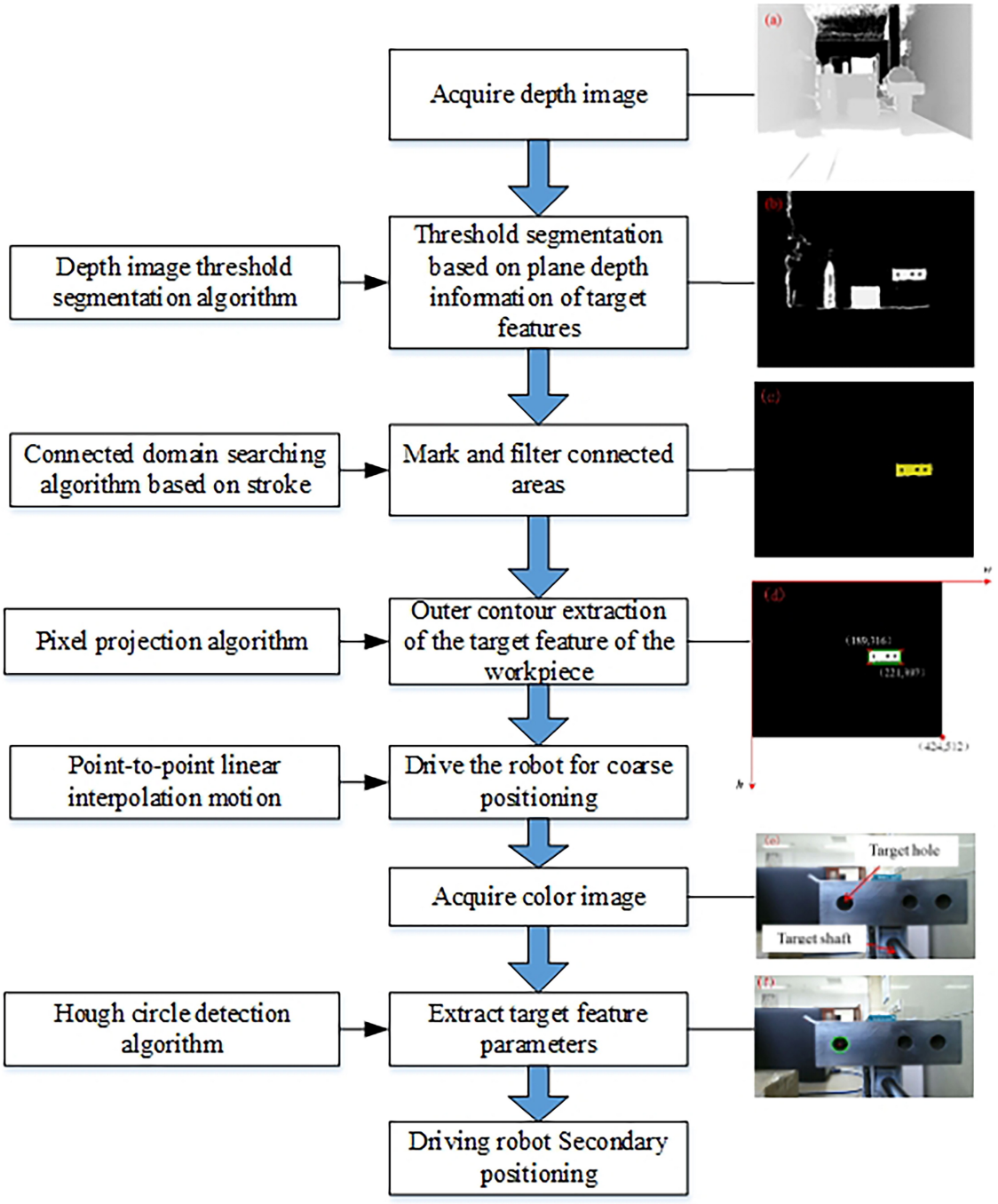

Fig. 1 The implementation process of the proposed method.

remaining interference pixels. The specific steps of the algorithm are as follows:

(1) Scan the image by line and record the starting point, the end point, and the line number of the successive pixel blocks in each line. Each successive pixel bar is called a block.

(2) Starting from number 1, if the current block is not connected with either block in the last row, give the current block a new label. If it is connected with either block in the last row, give the current block the smallest label of the last line, and write the block in the last row into the equivalent set

(3) The blocks in the equivalent set are connected. For the convenience of labeling the last pixel, we should determine the connected equivalent set in order to form an equivalent sequence. Starting from number 1, and giving each equivalent sequence a label, several equivalent sequences can be acquired.

(4) An equivalent sequence is searched according to the label of the starting block in each equivalent sequence, and a new identical mark is given to all the blocks in the sequence.

(5) Mark each equivalent sequence corresponding to the image and mark them with different pixel colors, so that all the connected areas in the image can be clearly seen.

After threshold segmentation, a connected region searching algorithm is used on the depth image to obtain a marked effect diagram, where the color of the working region is yellow and other regions are distinguished by different colors. Next, according to the ratio between the width and the height (ROI_width/ROI_height) and the area (ROI_area) of the circumscribed rectangle of each connected region, we can set filtering conditions to completely separate the workpiece features from the image. The final result is shown in Fig.1(c).

$$2 < \text{ROI\_width} / \text{ROI\_height} < 3 \quad (2)$$

$$\text{ROI\_area} > 45 \quad (3)$$

3.3 Depth image pixel projection of target workpiece

The next step performs pixel projection processing on the feature of the workpiece to ensure that the workpiece is correctly segmented, to prevent the robot from malfunctioning during coarse positioning. When the workpiece is correctly segmented, the corner coordinates and the center point coordinates of the workpiece in the image coordinate system can be obtained from the pixel projection results, and we also can draw the circumscribed image rectangle of the workpiece.

In the real world, the appearance of an object can be determined through its three views. In a planar image, the shape of a regular geometric object can also be determined by the pixel projection results in both horizontal and vertical directions. The detailed process is shown in Fig. 2. A pixel projection algorithm is applied to the depth image after searching the connected domain. As shown in Fig.1(d), the values of the corner points of the circumscribed rectangle of the workpiece image in the depth image coordinate system can be obtained from the pixel projection results of the workpiece depth image. After this calculation, we can get the approximate image coordinates of the center point of the workpiece (205, 356.5).

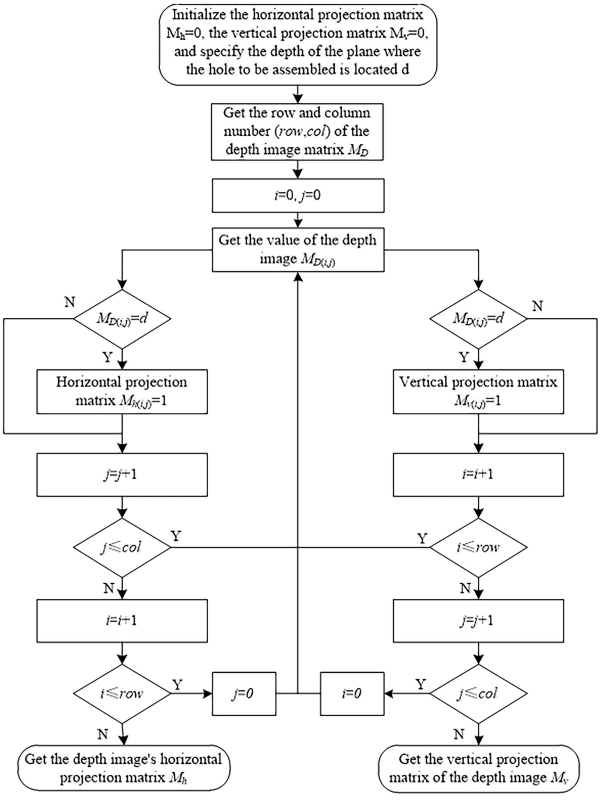

Fig. 2 Pixel Projection Algorithm.

3.4 Detecting the hole feature of the target workpiece

For the recognition of hole feature on the target workpiece, the HoughCircles( ) function provided in the Open Source Computer Vision Library (OpenCV) is implemented [15]. Because this algorithm has a high demand on the quality of the im-

age, only when the image of the detected round hole in the image is close to a perfect circle and occupies a certain number of pixels, the accuracy of the detection could be guaranteed. Therefore, the camera has to be adjusted to a proper position through coarse positioning in the early stage, to obtain a high-quality image for examination, as shown in Fig. 1(e). The detection result is shown in Fig. 1(f). The circle center and radius parameter information of the target circular hole in the image coordinate system is: {(607,851), 63.3877}.

4 Robot Motion Control

In the coarse positioning process, the target workpiece is separated from the depth source map, to obtain the coordinate values of its outer contour corner point in the depth image coordinate system. Next, according to relationship of the coordinate conversion between the image coordinate system and the robot coordinate system, these coordinate values are converted into the control signal in the robot coordinate system, to drive the robot to the correct position. As shown in Figure 3, the horizontal and vertical viewing angles of the Kinect v2.0 camera depth data sensor are 70° and 60°, respectively, and the depth image has a total width of 512 pixels and a total height of 424 pixels. The width of the target workpiece in the image coordinate system is w_{oi} (unit: pixels), and the actual distance between the target workpiece and the camera is d mm. Then we can determine that the true view width of depth sensor of the camera is

$$b_1 = 2d \cdot \tan(70°/2) \tag{4}$$

According to the proportional relationship, we can also obtain the true width of the target workpiece as

$$w_{or} = b_1 \cdot w_{oi}/512 \tag{5}$$

Also, a true distance corresponding to a width pixel of the target workpiece in the depth image can be obtained as w_{or}/w_{oi}. In the same way, we can get the real distance corresponding to a height pixel of the target workpiece in the depth image as h_{or}/h_{oi}, which determine the robot control increment as

$$\triangle y = -\triangle w_d \cdot w_{or}/w_{oi}, \triangle z = \triangle h_d \cdot h_{or}/h_{oi} \tag{6}$$

After coarse positioning, the end face of the target axis coincides with the plane of the target hole, and the image quality of the target hole in the color image is good, which meets the accuracy requirements of the secondary positioning. As shown in Fig.1(f), the coordinate value of the end face center of the target axis in the color image coordinate system is (977, 1257), and the coordinate value of the center point of the target hole is (607, 851). If we want the image coordinates of the two center points to coincide, we need to execute the camera pixel increments

$$\triangle w_c = 406, \ \triangle h_c = 370 \tag{7}$$

Next, according to the relationship of the coordinate conversion between the color image coordinate system and the robot coordinate system, we convert these coordinate values into the control signal in the robot coordinate system, to drive the robot to the correct position.

5 Experiments and Verification

5.1 Experimental setup

The overall experimental platform is composed of an upper software control system and a lower hardware execution system, as shown in Figure 4. Control software development is carried out in the operating environment Windows 8.1 x64, compiled in Microsoft Visual Studio 2010/MFC. Moreover, OpenCV 2.4.9 Library and Kinect SDK v2.0 provide image processing program and the related class library support, and robot offline programming interface library provides the related class library support for the robot control program. The hardware system is composed of a PC (Core i3/4G/500G), a robot control cabinet, a teach pendant, an SR10C six-axis industrial robot, a Kinect v2.0 camera, and a workpiece. The camera communicates with the host computer (PC) through the USB 3.0 interface, and the robot control cabinet and the PC communicate through the network port using the TCP/IP protocol.

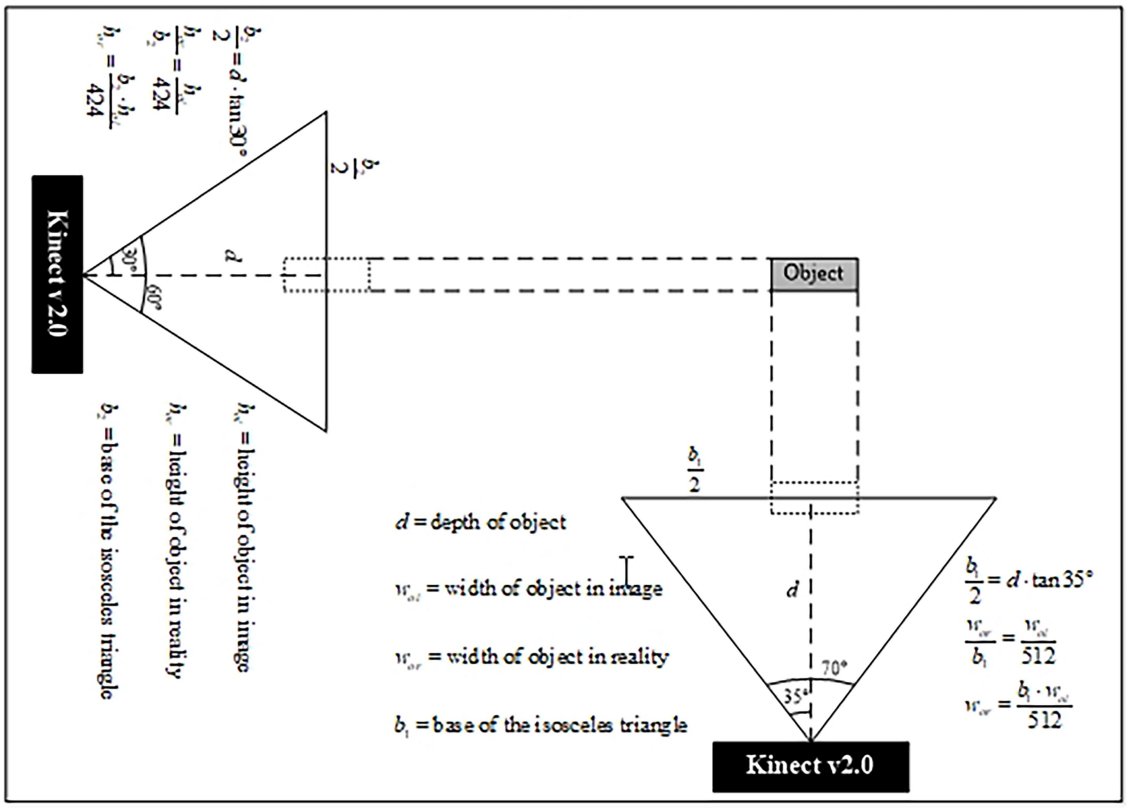

Fig. 3 The conversion relationship between the depth image coordinates and the real coordinates.

The visual system is built using the eye-in-hand format. The computational time of the proposed algorithm is about 1s. As shown in Fig.4, the camera is mounted on the end of the robot through an adapter rack, and the target shaft workpiece is also fixed at the end of the robot through the adapter. The target hole workpiece is placed on the experimental bench. The following steps are carried out:

(1) Obtaining a depth image of a field of view of the target workpiece, and extracting the coordinates of the center point of the target workpiece by using an image processing algorithm designed according to the depth information of the plane of the target hole;

(2) According to the coordinate conversion relationship between the depth image coordinate system and the robot coordinate system, obtain the spatial position information of the target workpiece in the robot coordinate system, and generate a control command to guide the robot to complete the depth image field of view alignment;

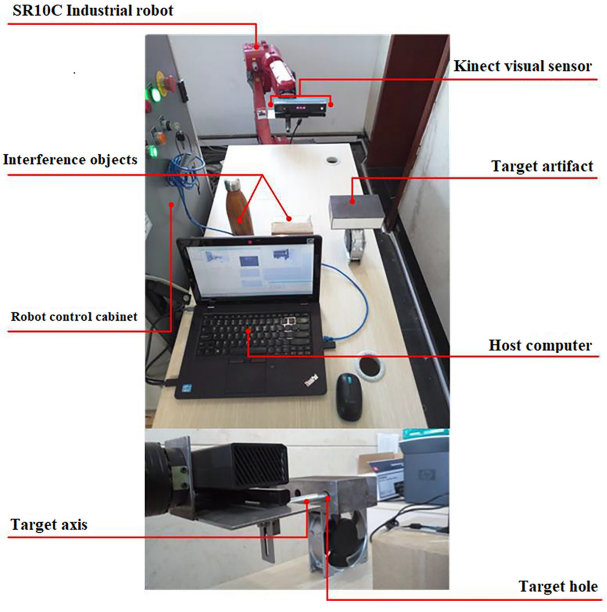

Fig. 4 The experimental platform.

(3) Guide the robot to approach the workpiece, based on the plane depth information of the target hole, so that the end surface of the target axis coincides with the plane of the target hole;

(4) Obtain the color image of the target workpiece, and use the hole feature detection algorithm to get the parameter information of the center and the radius of the target hole, so as to guide the robot to complete the final positioning process.

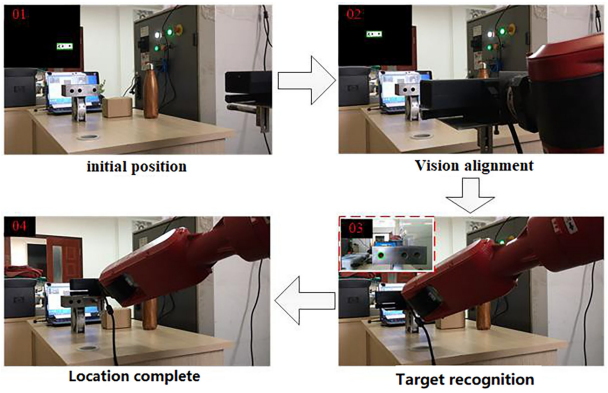

Fig. 5 The visual positioning experiment.

5.2 Experimental results

First, we conduct 10 times the uncalibrated visual positioning experiment on the target hole (diameter: Φ 10) of the workpiece to be assembled, as shown in Fig.4. The results are given in Table 1.

Due to the influence of the accuracy of the robot's own motion, the visual sensor accuracy and the complex external conditions, the positioning result will contain errors. For the convenience of analysis and observation, the position of the center of the target axis hole after the positioning process is described as in Fig.6, where red + represents the actual position of the target axis center, and black + represents the actual position distribution of the target axis center after the 10 positioning experiments.

In order to further analyze the size and distribution behavior of the positioning error, the positioning error of the 10 experiments in the directions of y and z, radial to the plane of the target hole is counted as shown in Fig.7. It can be seen from Fig.7 that the positioning error in the y direction is basically stable between 0.6 and 1.05 mm, and the positioning error in the z direction is basically stable between 0.6 and 1.1 mm.

In order to verify the versatility of the method

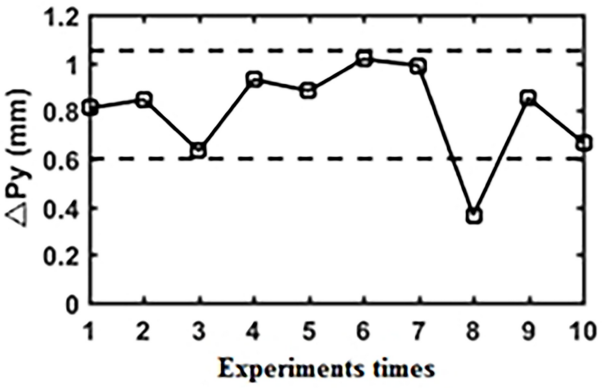

Fig. 6 The center points distribution of the target peg and hole.

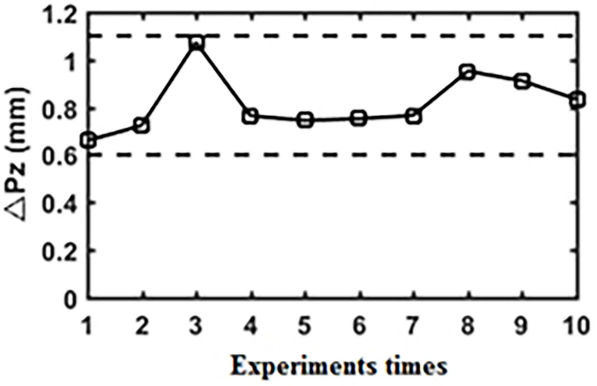

Fig. 7 The positioning error distribution.

and to study the effect of the aperture size on the visual positioning error, a visual positioning experiment is performed on a workpiece with a set of concentric holes with diameters of Φ30 and Φ60, as shown in Fig.8. Under the same conditions, 10 visual positioning experiments are performed on the circular holes to be assembled with diameters of Φ30 and Φ60, and the position feedback data is collected to determine the error distribution behavior. The experimental results are given in Table 2.

In order to more intuitively represent the position distribution of the center of the shaft and the hole after each positioning, the data collected in Table 2 is plotted as a scatter plot, where, the red "+" represents the actual position of the target hole center, the blue "+" represents the experimental result of visual positioning of the Φ30 target hole, and the

rose red " +" represents the target hole positioning result with a diameter of Φ60. It can be seen from Fig 9 that the position of the center of the circle obtained by identifying the hole with a diameter of 60 mm is more accurate than the position of the center of the circle with a diameter of 30 mm.

Table 1 The experimental results of visual positioning.

Experiment number	Target hole site		Aperture size	Target axis position	
	P_y (mm)	P_z (mm)		P'_y (mm)	P'_z (mm)
1				198.837	510.091
2				198.872	508.708
3				197.394	508.358
4				198.956	508.667
5				198.909	508.686
6	(198.026,509.431)		Φ10	197.011	510.183
7				197.038	510.197
8				198.391	510.383
9				198.879	510.342
10				198.693	508.597

Table 2 The experimental results of visual positioning.

Experiments number	Target hole site		Aperture size	Target axis position	
	P_y (mm)	P_z (mm)		P'_y (mm)	P'_z (mm)
1			176.100		584.156
2				174.805	582.536
3				176.028	582.733
4				174.830	582.637
5				174.497	582.662
6			Φ30	174.671	582.523
7				176.132	582.596
8				176.191	582.549
9				174.521	583.956
10			174.612	582.511	
11	(175.368,583.293)		174.892	583.649	
12			175.736	582.716	
13			175.711	583.663	
14			175.951	583.663	
15				174.777	582.955
16			Φ60	174.913	583.771
17				175.870	582.829
18				174.841	583.846
19				174.971	583.875
20				175.798	582.772

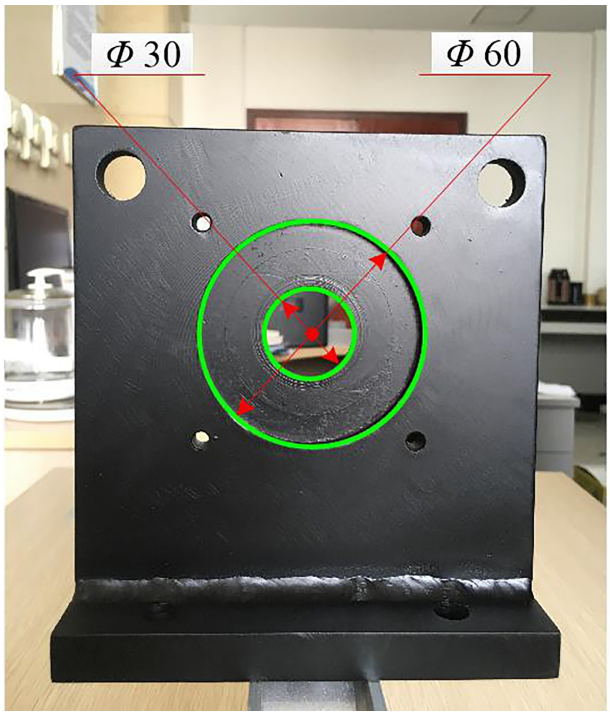

Fig. 8 The workpiece with concentric circles.

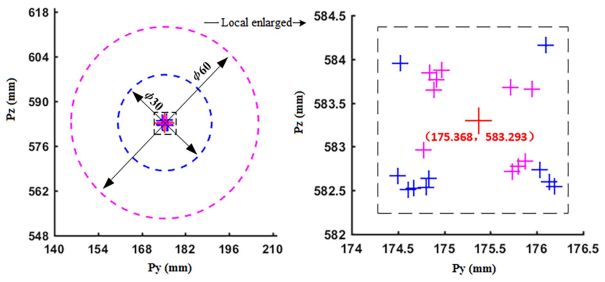

Fig. 9 The center point distribution of the target peg and hole.

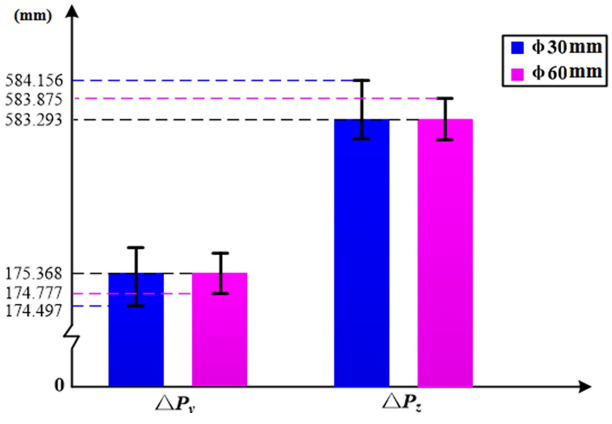

Fig. 10 Error band distribution.

Comparing the distribution of the error zone in Fig. 10, for the visual positioning error distribution of the circular hole with a diameter of 30 mm, the *y*-direction positioning error is less than 0.90 mm, and the *z*-direction positioning error is less than 0.87 mm. In the visual positioning error distribution of a circular hole with a diameter of 60 mm, the positioning error in the *y* direction is less than 0.60 mm, and the positioning error in the *z* direction is less than 0.59 mm. We can make the following conclusions from the results of the conducted experiments with the specifications Φ10, Φ30 and Φ60 t of the circular hole visual positioning: With the increase of the diameter of the circular hole, the positioning error of the proposed visual positioning method is smaller and tends to be stable, which provides a good reference for error compensation and actual production applications. Due to the resolution limitation of the used vision sensor, there are limitations to the small holes positioning.

6 Conclusion

This paper aimed at the identification and positioning of workpiece in the process of shaft assembly of flexible production cell using industrial robots. A calibration-free visual positioning method for industrial robots based on depth images was proposed. Compared with the traditional robot visual positioning method, the method proposed in this paper solved the problem of identification and positioning of workpieces in the axle hole assembly process using industrial robots in a flexible production system, and breaking through the limitations and limitations of vision for visual positioning with fixed cameras. The two positioning processes made full use of the advantages of depth image and color image in the corresponding stages, avoiding the interference of environmental noise on the image processing process. This provided better flexibility and adaptability in practical applications. This method can reduce the time consumed for re-calibration and teaching of the robot vision system due to the change of

assembly conditions in multi-specification and small-batch axle hole assembly tasks.

Due to the limitation of the development level of the robot vision system, the error caused by pure visual positioning was unable meet the needs of direct assembly of precision shaft holes. The subsequent assembly process can only be completed by multi-sensor information fusion such as force sensor and flexible control of the special fixture at the end of the robot, by continuously optimizing the target recognition and localization algorithm and compensating for the error to improve the positioning accuracy and system stability. This direction is suitable for future work.

References

[1] Tang, Q. (2017). Coordinate transformation of automated spraying system based on point cloud and image matching. *Computer Integrated Manufacturing Systems*, 23(11), pp.2392-2398.

[2] Vijayan, A T, and Ashok, S. (2017). Integrating visual guidance and feedback for an industrial robot. *International Conference on Control & Robotics Engineering*.

[3] Perez, L. (2016). Robot guidance using machine vision techniques in industrial environments: A comparative review. *Sensors*, 16(3), pp.335

[4] Zhai, J M. (2014). Positioning and Grasping System Design of Industrial Robot Based on Visual Guidance. *Machine Design and Research*, 30(05), pp.45-49

[5] Wu, X R., Ling, X Y., and Liu, J X. (2019). Location Recognition Algorithm for Vision-Based Industrial Sorting Robot via Deep Learning. International Journal of Pattern Recognition and Artificial Intelligence, 33(7), pp.1-18

[6] Timotei I E., Zsolt M., and Géza. (2016). Robot visual and virtual control technology in industrial environment. *International Symposium on Small-scale Intelligent Manufacturing Systems IEEE*, pp.71-75.

[7] Pena, M., Lopez, I., and Osorio, R. (2006). Invariant Object Recognition Robot Vision System for Assembly. *Electronics, Robotics & Automotive Mechanics Conference. IEEE*, pp.30-36.

[8] Chang, W C. (2017). Robotic assembly of smartphone back shells with eye-in-hand visual servoing. *Robotics and Computer-Integrated Manufacturing*.

[9] XIE, Z X. (2016). Industrial Robot Positioning System Based on the Guidance of the Structured-Light Vision, *Acta Optica Sinica*, 36(10), pp.400-407

[10] WU, X R. (2016). Fast Visual Identification and Location Algorithm for Industrial Sorting Robots Based on Deep Learning. *Robot*, 38(06), pp.711-719

[11] Wang, W L. (2017). Intelligent robot object detection algorithm based on spatial pyramid and integrated features. *Computer Integrated Manufacturing Systems*, 23(11), pp.2382-2391

[12] Zhang, Y J. (2012). *Image Processing*, Tsinghua University.

[13] LIU, HB. (2010). *Digital Image Processing Using Visual C++*, China Machine

[14] Kaehler, A., & Bradski, G. (2016). *Learning OpenCV 3: Computer Vision in C++ with the OpenCV Library*, O'Reilly Media, Inc

[15] MAO, X Y. (2015). *Introductory Programming of OpenCV 3*, Publishing House of Electronics Industry

Authors' Biographies

Ming CONG, received his Ph.D. from Shanghai Jiao Tong University in 2003. He is currently a professor and PhD supervisor in Dalian University of Technology, PR China. His main research interests include robotics and automation, intelligent control, and biomimetic robotics.

Email: congm@dlut.edu.cn.

Fukang ZHU, received his Master's degree from Dalian University of Technology, PR China, in 2018. Now he is an engineer at Suzhou Huichuan Technology Co., Ltd. His main research interests include industrial robotics and machine vision.

Email: zhufukang@mail.dlut.edu.cn

Dong LIU, received his PhD from Dalian University of Technology (DUT), PR China, in 2014. He is currently an Associate Professor and Master supervisor in DUT. His main research interests include intelligent robotics and systems, cognitive control, and motion planning.
Email: liud@dlut.edu.cn

Yu DU, received her PhD from the University of British Columbia, Vancouver, Canada, in 2018. She is currently the Chief Executive Officer of Dalian Dahuazhongtian Technology Co. Ltd. Her main research interests include robotics and automation, and intelligent control.
Email: dutduyu@gmail.com