International Conference on Mechanical System Engineering and Technology Applications (ICoMSETA) 2024, pp. X – XX

Machine Vision Calibration for Industrial Robots Target Detection

Zhang XiaoYang[1,3],Muhammad Azmi Ayub[1], Fazlina Ahmat Ruslan[1], Sukarnur Che Abdullah[1], Shuzlina Abdul-Rahman[2]

[1] College of Engineering, UiTM Shah Alam,40450 Shah Alam Selangor, Malaysia

[2] School of Computing Sciences, College of Computing, Informatics and Media, UiTM Shah Alam,40450 Shah Alam Selangor, Malaysia

[3] Hebei Institute of mechanical and electrical technology, xingtai, China

[4] fazlina419@uitm.edu.my

Abstract. Industrial robots are increasingly used in various fields of industrial production. At present, the working mode of most industrial robots is fixed and single, and when the working environment or target position changes, industrial robots cannot accurately grasp and need to be repositioned. In order to improve the autonomy and intelligence of its work, machine vision is gradually applied to combine with robot technology to give robots the ability to perceive the environment.With the increase of industrial demand, rapid and accurate recognition and positioning of objects with arbitrary posture in complex space environment has become a new research direction.To solve this problem, this paper studies hand-eye calibration and pose recognition in 3D vision system detection based on depth camera, and realizes industrial robots to grasp objects based on pose results of vision system. In order to solve the problem of mapping between the camera coordinate system and the industrial robot coordinate system, this paper proposes a robot calibration method, which can quickly and accurately obtain the transformation matrix between the image coordinate system, the camera coordinate system and the robot coordinate system. The camera calibration process and the imaging system model are analyzed, and then image processing algorithms such as image segmentation, target classification, image matching, corner detection, target localization and tracking are used. Finally, the pose data of different positions of targets are obtained through experiments, and the error analysis of the data is carried out, which provides strong data support for accurate positioning and grasping of targets.

Keywords: Calibration, Industrial Robot, Machine Vision, Target Detection.

1. Introduction

As a key driving force for intelligent manufacturing and industrial development, industrial robots are expanding their application range. The use of industrial robots to replace manual handling, assembly, welding, and other repetitive heavy tasks can significantly reduce labor costs, enhance production efficiency, and meet product reliability requirements[1]. With the increasing requirements of complexity and flexibility of industrial robots and automated processing equipment in intelligent processing and manufacturing, the introduction of robot technology in the field of industrial production has become an inevitable trend. Therefore, robot technology has become the main symbol of measuring a country's scientific and technological innovation ability and high-end manufacturing level[2].

Industrial robots are particularly suitable for replacing humans working in harsh environments or engaging in boring repetitive work, such as processing, welding, spraying, assembly and handling, a large part of which is closely related to object grasping. The robotic arm is a multi-joint manipulator or multi-degree-of-freedom mechanical device, representing a crucial category within industrial robotics and holding significant relevance in the industrial manufacturing sector[3].

However, the traditional arm-type robot lacks the ability of environment perception and can't adjust the environment and target quickly in real time. This results in the lack of flexibility and robustness of traditional robots, which severely limits their application in dynamic scenarios. Machine vision technology gives industrial robots the ability to see like humans. In a complete machine vision system, the vision sensor collects the image of the objective object in the work scene, extracts and analyzes the important information through the intelligent processing unit to simulate the human brain, realizes the recognition and positioning of the target object, and even understands the work scene, so as to improve the adaptability of industrial robots to the changing external environment[4]. Enhance its perception and decision-making ability in complex industrial manufacturing environments.

In order to effectively to perform the target grasping, calibration of the robot system is essential. Robot calibration is the process of enhancing the accuracy of a robot manipulator through modification of the robot control software[5]. In general, robot calibration can be classified into two types, static and dynamic. Static calibration is an identification of those parameters influencing primarily the static positioning characteristic of the robot end-effector. Dynamic calibration is used to identify parameters influencing primarily the motion characteristics[6].

There are four steps involved in the calibration:

- Determination of a mathematical model that represents the robot geometry and its motion (Kinematic modeling)
- Measurement of position and orientation of the robot end effector in world coordinates (Pose measurement)
- Identification of the relationship between joint angle and end-point positions
- Modification of control commands to allow successful completion of a programmed task (Kinematic compensation)

In general, Figure 1 shows the dynamic or in-process robot calibration before the grasping task. The measurement of position and orientation of the robot end effector in world coordinates (Pose measurement) is important, for the grasping task can be accurately performed.

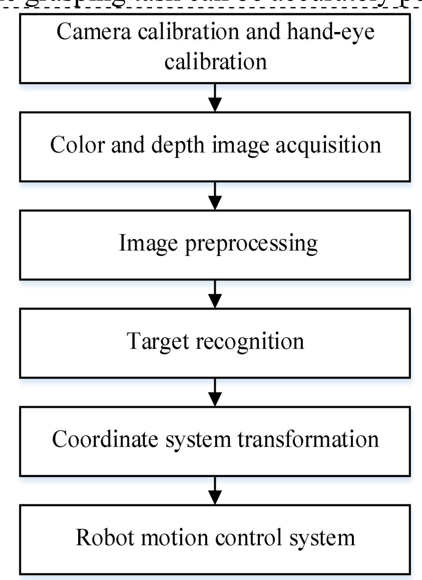


Figure 1. A structured sequence of robotic grasping task

This overall structured sequence of robotic grasping tasks can be summarized are as follows,

Camera calibration and hand-eye calibration: The initial step for the vision system involves completing both camera calibration and hand-eye calibration between the robot and the camera. This process enables the calculation of camera parameters and coordinate system conversion parameters.

Color and depth image acquisition: Depth information is obtained via a binocular vision system, while color image data is captured using an RGB camera. Following this, the depth information is correlated with the color image data at the pixel level.

Image preprocessing: Image preprocessing involves processing the acquired color image and depth image, including tasks such as image smoothing, enhancement, threshold segmentation, and edge extraction.

Target recognition: Following image processing, the color image is aligned with the depth image, allowing for the retrieval of pixel coordinates and depth values for the object's central point.

Coordinate system transformation: Through a series of coordinate transformations from the target object's coordinate system to the camera coordinate system, then to the end-effector coordinate system of the robot arm, and subsequently to the base coordinate system of the robot, the image data is converted into pose information in the robot's base coordinate system. This transformation results in the extraction of coordinates and pose details for the target object's grasp point, furnishing essential input parameters for the robot's motion planning.

Robot motion control system

Finally, a visual feedback control strategy is designed to achieve the grasping operation of the moving target on the conveyor belt through fixation approximation.

An experimental platform for 3D target grasping is built with binocular vision system. The accurate grasping of spatially arbitrary positional targets by industrial robots relies on the 6-DOF pose estimation of spatial targets by the 3D vision system. 3D point cloud data is generated by matched pairs of parallax image. In order to estimate the 6-DOF pose of target, a pose estimation method based on global structure feature constraints is proposed to quantify the global features which describe the target structure as constraints in the point cloud reconstruction process. Experiments are expected to verify that the robot can grasp arbitrary objects with corresponding posture and prove the feasibility of the method.

2. Model and Method

An experimental platform for 3D target grasping is built with a binocular vision system. The accurate grasping of spatially arbitrary positional targets by industrial robots relies on the 6-DOF pose estimation of spatial targets by the 3D vision system. 3D point cloud data is generated by matched pairs of parallax images. In order to estimate the 6-DOF pose of a target, a pose estimation method based on global structure feature constraints is proposed to quantify the global features which describe the target structure as constraints in the point cloud reconstruction process. Experiments are expected to verify that the robot can grasp arbitrary objects with the corresponding posture and prove the feasibility of the method[7&8].

For a robot system with vision, all information obtained by the camera is described in the camera coordinate system. In order to make the robot use the information obtained by the vision system, the first thing is to determine the relative relationship between the camera coordinate system and the robot, which is the research content of the robot hand-eye calibration[9&10].

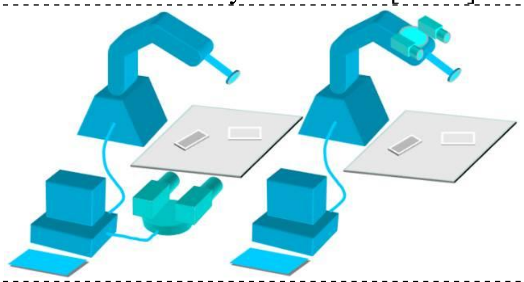


Figure 2. Eye-to-Hand and Eye-in-Hand Model

As shown in Figure 2, the installation of cameras in robot systems can be divided into two categories:

Eye-to-Hand: that is, the part where the camera is installed outside the arm, which is relatively fixed with the base of the robot (world coordinate system) and does not move with the movement of the robot arm.

Eye-in-Hand: that is, the camera is installed on the mechanical arm and moves with the movement of the mechanical arm[11&13].

The solutions are slightly different, but the basic principles are similar.

2.1. *Eye-to-Hand*

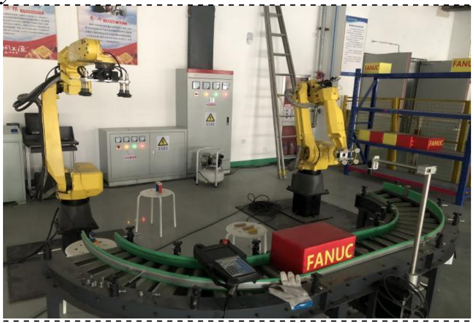


Figure 3. Eye-to-Hand

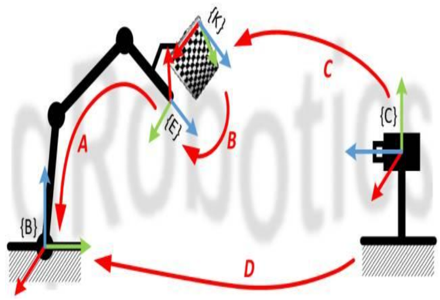


Figure 4. Eye-to-Hand system model

As shown in Figure 3 and Figure 4, the Kinect camera sits on a table next to the arm, and the camera is relatively fixed to the base of the arm as long as no one touches it.

We get coordinate systems like these:
{B}： Robot base coordinate system
{E}： Coordinate system of robot terminal connecting rod (connecting rod fixedly with calibration plate)
{K}： Calibration plate coordinate system
{C}： Camera coordinate system
And coordinate transformation between them:
{A}： The position and pose of the end of the robot in the coordinate system of the base of the manipulator, which is actually the problem of the forward kinematics of the robot;
{B}： The position and pose of the calibration plate in the robot terminal coordinate system are unknown because the calibration plate is installed randomly;
{C}： The pose of the camera in the calibration plate coordinate system, which is actually solving the extrinsic parameter of the camera;
{D}： The orientation of the camera in robot base coordinates, that's what we want to figure out.

$$D = A * B * C \tag{1}$$

Therefore, as long as we can calculate the B transformation, the pose D of the camera in the robot coordinate system can be obtained naturally[14].

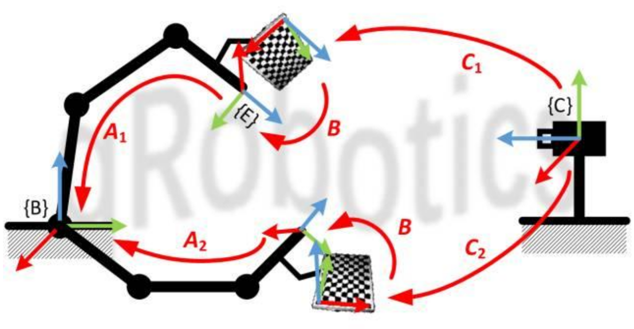


Figure 5. Eye-to-Hand system model

As shown in Figure 5, we let the robot move two positions to ensure that the camera can see the calibration plate in both positions, so as follows:

$$A_1 * B_1 * C_1 = A_2 * B_2 * C_2 \tag{2}$$

Because the calibration plate and the connecting rod at the end of the manipulator are fixed, the B transformation is also fixed and invariable.
Change it a little:

$$\left(A_2^{-1} * A_1\right) * B = B * \left(C_2 * C_1^{-1}\right) \tag{3}$$

This is a typical AX = XB problem, where X is, by definition, a 4x4 homogeneous transformation matrix:

$$X=\begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \tag{4}$$

So, with A, B, and C, we can directly solve for the desired hand-eye calibration result D [15].

2. 2. *Eye-in-Hand*

Figure 6. Eye-in-Hand

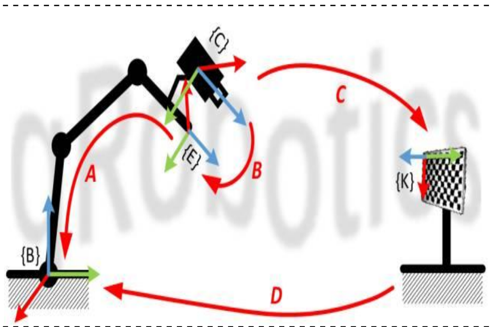


Figure 7. Eye-in-Hand system model

As shown in the Figure 6 and Figure 7, the Realsense camera is fixed at the end of the manipulator and moves with the movement of the manipulator.
We get a bunch of coordinate systems similar to the previous one and coordinate transformation relationships (slightly different) as shown in Figure 8.

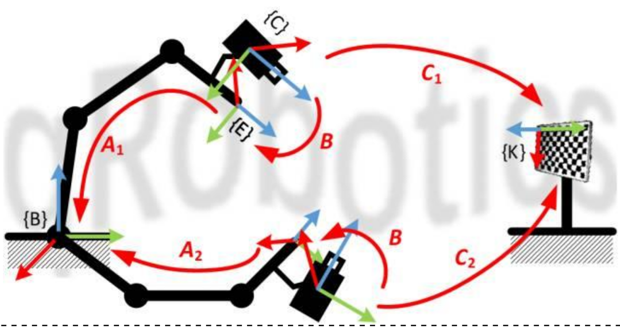


Figure 8. Eye-in-Hand system model

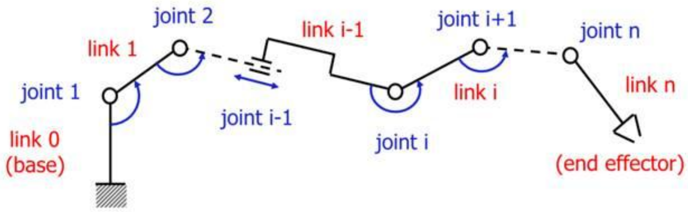
{A} : The position and pose of the end of the robot in the coordinate system of the base of the manipulator, which is actually the problem of the forward kinematics of the robot;
{B} : The pose of the camera in the coordinate system at the end of the robot, the transformation is fixed. As long as we know the transformation, we can calculate the actual position of the camera at any time, so this is what we want to find;
{C} : The pose of the camera in the calibration plate coordinate system, which is actually solving the extrinsic parameter of the camera;
{D} : The camera pose in the robot base coordinate system does not exist in the actual use of the calibration board, so we do not care about this transformation relationship.
This is similar to the previous one, which directly moves the manipulator in two positions to ensure that the calibration plate can be seen from both positions. Then, the spatial transformation loop is constructed:

$$A_1 * B * C_1^{-1} = A_2 * B * C_2^{-1} \tag{5}$$

$$\Downarrow$$

$$\left(A_2^{-1} * A_1\right) * B = B * \left(C_2^{-1} * C_1\right) \tag{6}$$

So it's an A*X=X*B problem again.
Let's take it one at a time:
Transformation A: It is the position and pose of the end of the robot arm in the coordinate system of the robot base. In fact, it is the most basic forward kinematics solution in robotics.



$${}^{0}_{n}T = {}^{0}_{1}T\,{}^{1}_{2}T\,{}^{2}_{3}T\,{}^{3}_{4}T \cdots\cdots {}^{n-2}_{n-1}T\,{}^{n-1}_{n}T \tag{7}$$

This is use DH matrix and other methods to calculate the forward kinematics of the robot. If it is in ROS, it can be obtained directly via KDL, TF, and etc. [16].In short: As long as we know the angles of each joint of the robot in its current state, we can calculate the A transformation[17].
Transformation C: it is the pose of the camera in the calibration plate coordinate system. In fact, it is the extrinsic parameter of the camera.

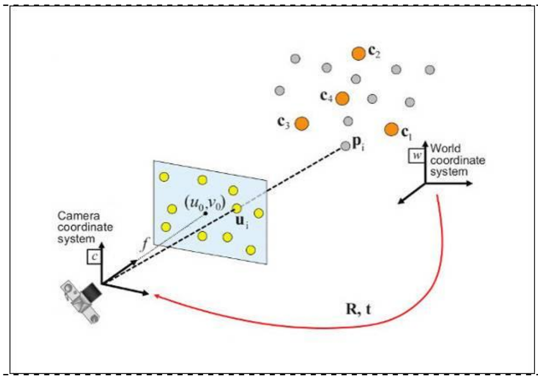


Figure 9. The keyhole model

Just to explain, for a typical camera, we use the keyhole model to model it as shown in Figure 9. Therefore, in the camera coordinate system, 3D points (X_c, Y_c, Z_c, 1) in space and corresponding 2D point (X_c Y_c) in the image exactly meet the following relationship:

$$\lambda \begin{bmatrix} x_c \\ y_c \\ 1 \end{bmatrix} = \begin{bmatrix} f_c & 0 & x_{c0} \\ 0 & fy & y_{c0} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} \tag{8}$$

The transformation matrix $\begin{bmatrix} f_c & 0 & x_{c0} \\ 0 & fy & y_{c0} \\ 0 & 0 & 1 \end{bmatrix}$, which is called the Intrinsic parameter matrix, contains the camera focal length and other parameters only related to the internal structure of the camera[18].

Then, we usually describe objects in world coordinates, that is, (X_w, Y_w, Z_w, 1), not in camera coordinates. So, we need to first transform the object coordinates from the world coordinates to the camera coordinates.

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = \begin{bmatrix} r11 & r12 & r13 & t_x \\ r21 & r22 & r23 & t_y \\ r31 & r32 & r33 & t_z \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \tag{9}$$

Thus, we have the complete camera model:

$$\lambda \begin{bmatrix} x_c \\ y_c \\ 1 \end{bmatrix} = \begin{bmatrix} f_c & 0 & x_{c0} \\ 0 & fy & y_{c0} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r11 & r12 & r13 & t_x \\ r21 & r22 & r23 & t_y \\ r31 & r32 & r33 & t_z \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \tag{10}$$

λ:Degree of freedom (for the depth);

$\begin{bmatrix} x_c \\ y_c \\ 1 \end{bmatrix}$:Point in captured image;

$\begin{bmatrix} f_c & 0 & x_{c0} \\ 0 & fy & y_{c0} \\ 0 & 0 & 1 \end{bmatrix}$:Intrinsic parameter matrix(focal length f,center of the image c0);

$\begin{bmatrix} r11 & r12 & r13 & t_x \\ r21 & r22 & r23 & t_y \\ r31 & r32 & r33 & t_z \end{bmatrix}$:Extrinsic parameter matrix(Rotational component r,translate component t);

$\begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$:Point in maker coordinate.

With the camera model, what we need to do is to determine the internal parameter matrix first, and then obtain the extrinsic parameter matrix according to the calibration plate pictures taken at different locations: namely, the transformation relation C of the camera in the world coordinate system[19&20].

2.3. *Experimental procedures*

We will use the 3D vision system to complete the pose data of a cuboid object relative to the basic coordinates of the industrial robot on the current experimental platform. The dimensions of this rectangular object are 54mm long, 37mm wide and 5mm high as shown Figure 10. The computer algorithm for image acquisition, processing and feature recognition was developed under the software Win10: Python + OpenCV3.4.3 + OpenGL.

Figure 10. Experimental platform

The camera hand-eye calibration as shown in Figure 11 is carried out to obtain the internal and external reference matrices of the camera. Then, in the developed 3D computer vision calibration algorithm, the fixed object 3D coordinate recognition task of the camera is completed.

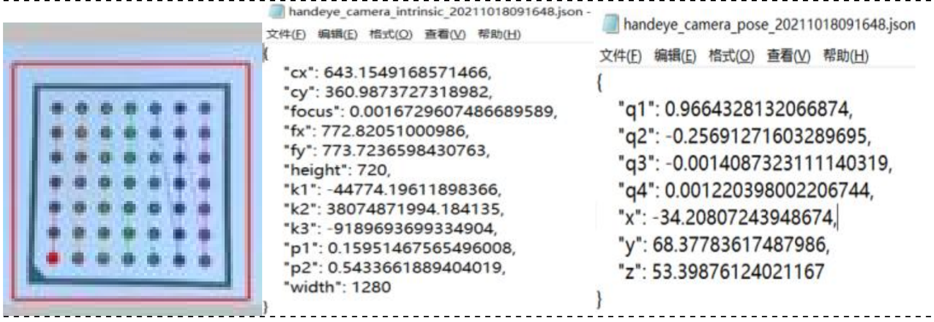


Figure 11. Hand-eye calibration

The 2D image is acquired and depth image is computed as shown in Figure 12 and Figure 13.

Figure 12. 2D-Image

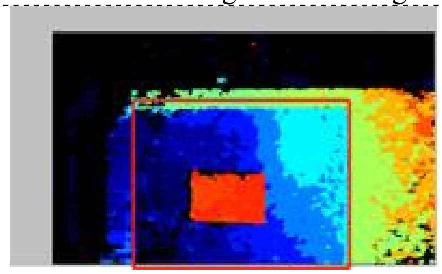


Figure 13. Depth-Image

For image processing, the color extraction is carried out to obtain the gray image of the object as shown in Figure14.

Figure 14. Color extraction

Then feature recognition and data analysis are carried out. Figure 15 shows the pose data of the object relative to the camera coordinate system that can be calculated through the internal reference matrix and directed object positioning. Hence, through the external parameter matrix, the pose data of the object relative to the base coordinate system are computed.

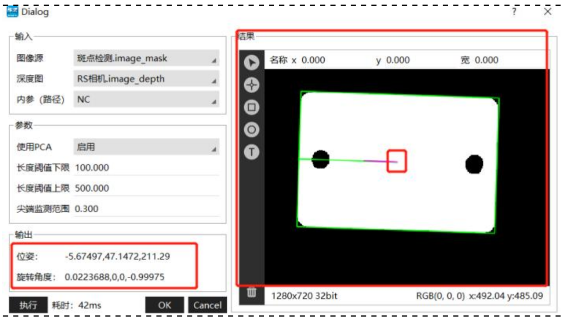


Figure 15. The camera pose

Through the external parameter matrix, the pose data of the object relative to the base coordinate system are calculated.

Figure 16. Object pose data

We fixed the camera directly above the object as shown in Figure 17, 220mm away from the shooting plane.Moving the target, and obtained three groups of data, namely three different positions of object:A, B, and C;three different positions of camera:P, Q and R.

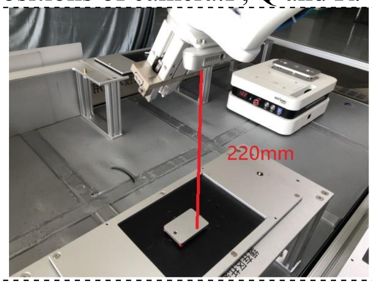
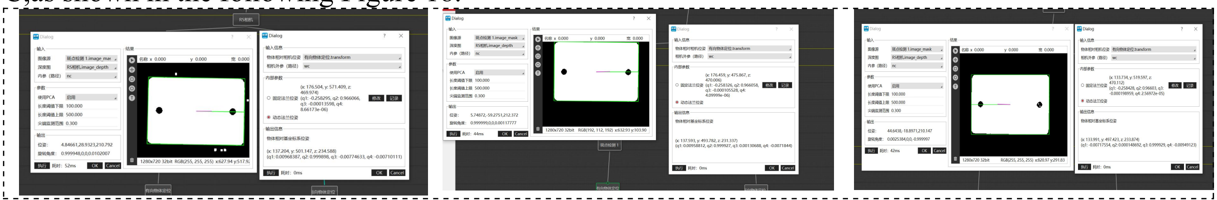


Figure 17. The camera height

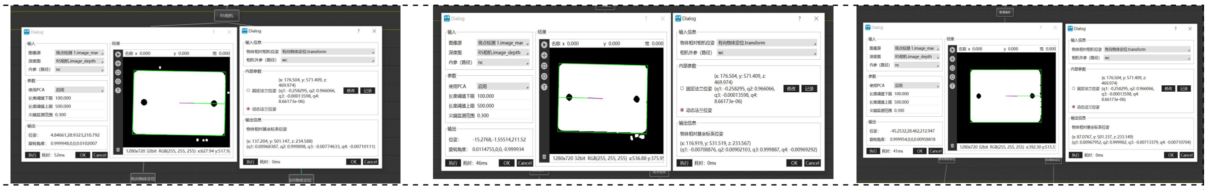
We put the object on the platform, fix the camera, move the object, get three sets of data P-A, P-B, P-C,as shown in the following Figure 18.



P-A P-B P-C

Figure 18. Three different locations of object and fix location of camera

We put the object on the platform, move the camera, adjust the camera position, and get three sets of data P-A, Q-A, R-A, as shown in the following Figure 19.



P-A Q-A R-A

Figure 19. Three different locations of camera and fix location of object

3. Results and Discussion

In this experiment, two groups of experiments were carried out to realize the detection and recognition of static objects by 3D camera. The first group of experiments is to fix the camera in one position, while placing the object in three different positions: A, B and C for target detection. The second group of experiments is to fix the object in one position and move the camera to three different positions: P, Q and R for target detection. All other parameters of the robotic visual servoing system remained constant during the image acquisition. The algorithms for both image processing and feature recognition were executed when the object and camera were moved and placed under the same conditions.

An average of six readings was taken for both sets of experiments. The results for the pose data of the object relative to the camera coordinate are shown in Table 1. The pose data of the object relative to the base coordinate system were shown in Table 2.

As shown in Table 1: The object relative to the camera coordinate.
The experimental results show that no matter the fixed camera, moving object, or fixed object, moving camera, in the camera coordinate system, the system can accurately detect the values of X, Y, Z coordinates of the object in the camera coordinate system, as well as the rotation angles of each axis θ, α, β.

By measuring the target object on the platform actual position (relative to the zero) of the camera coordinate system, and comparing the experimental data, the error range is less than 2mm, indicating that the internal parameter calibration of the camera is basically correct, and the correct conversion of the target object to the coordinate system of the photographed image can be realized.

As shown in Table 2: The pose data of the object relative to the base coordinate system.
No matter it is fixed camera, moving object, or fixed object, moving camera, the results of these two groups of experiments show that the system can accurately obtain the values of X, Y and Z coordinates of objects in the basic coordinate system of industrial robots through the calibration of the external reference matrix, as well as the rotation angles of each axis θ, α, β.

When the camera is fixed and the object is moved, through the measurement of the coordinate X, Y and Z values of the object in the base coordinate system of the industrial robot and the measurement of the attitude data of each axis, and comparing the experimental calculation data, it shows that the calculation of the external parameter matrix of the camera can realize the correct conversion of the camera coordinate system to the base coordinate system of the industrial robot, and the measurement error is within 1mm.

When the object is fixed and the camera is moved, the correct conversion of the camera coordinate system to the base coordinate system of the industrial robot can be realized by measuring the coordinate X, Y and Z values of the object in the base coordinate system of the industrial robot and measuring the attitude data of each axis, comparing the experimental calculation data and calculating the external parameter matrix of the camera. The measurement error is within 8mm. For the measurement of the target in the same position, it can be seen from the different data obtained that the measurement error is obviously larger, which may cause certain interference to the robot to grasp the target.

Therefore, when detecting static objects, the camera can be fixed in one position for target detection. If the camera needs to move along with the end axis and take pictures at the same time, the error will have a large deviation. In the later research, when the camera moves to take pictures, the error should be corrected to ensure the correct detection of the target.

The overall experimental results show that the average error is larger when the camera is moving than when the camera is fixed. The position and orientation of the object are inconsistent with respect to the robot coordinate system. The cumulative error when the camera is fixed is less than the cumulative error when the camera is moving. Generally speaking, the errors of robot visual servo system mainly come from two sources. These errors are caused by the robot's inherent mechanism and vision system. Examples of errors caused by the inherent mechanism of a robot are gear backlash, friction, and wear of the mechanism. Errors due to the visual system can come from the presence of dust and dirt in the inspection area, uneven light distribution, and random speckles produced by the camera.

Table 1. The pose data of the object relative to the camera coordinate

Location Camera	Location Object	Position Camera			Orientation Camera		
		X	Y	Z	θ	α	β
P	A	4.8466	28.932	210.792	0	0	1.169
P	B	-15.276	1.555	211.520	0	0	178.685
P	C	-45.253	28.462	212.947	0	0	1.099
P	A	4.8466	28.932	210.792	0	0	1.169
Q	A	5.7487	-59.275	212.372	0	0	0.135
R	A	44.643	-18.897	210.147	0	0	-179.709

Table 2. The pose data of the object relative to the base coordinate system

Location Camera	Location Object	Position object			Orientation object		
		X	Y	Z	θ	α	β
P	A	137.204	501.147	234.588	178.897	-0.822	0.880
P	B	116.919	531.519	233.567	178.861	-0.812	-178.977
P	C	87.0767	501.337	233.149	178.896	-0.823	0.809
P	A	137.204	501.147	234.588	178.897	-0.822	0.880
Q	A	137.593	493.782	231.337	178.900	-0.821	-0.157
R	A	133.991	497.923	233.874	-178.913	0.822	-179.991

4. Conclusions

Traditional industrial robots lack the ability to sense and judge the environment, so they can only perform presetting repeated actions in fixed scenes, and the feasibility of the actions still needs human judgment. In order to improve the automation performance of the robot, calibration of the robotic system for the adaptive movement in the dynamic complex environment was realized. It is essential to obtain information about the external environment and to process and judge the robot's grasping accuracy so as to provide the basis for robot trajectory planning. In this paper, the dynamic calibration of the visual servoing robotic system using the Eye in hand was performed. That is, the camera is installed on the mechanical arm and moves with the movement of the mechanical arm. The result shows that the errors are bigger when the camera is moving with the robot arm. However, Eye in hand system will give greater flexibility if the object is moving.

References

[1] Shen T, Radmard S, Chan A, et al. Optimized vision-based robot motion planning from multiple demonstrations [J]. Autonomous Robots, 2018, 42(6): 1117-32.

[2] Pérez L, Rodríguez Í, Rodríguez N, et al. Robot Guidance Using Machine Vision Techniques in Industrial Environments: A Comparative Review [J]. Sensors, 2016, 16(3): 335.

[3] Zhang Xi, Cheng DongQin, Li Wei.Construction of large field and high precision visual calibration of virtual stereoscopic target [J]. Optics and Precision Engineering, 2017,25(4):891-9.

[4] Gong Z L, Gu Y H, Zhu T T, Ren B. A method for setting costmap adaptive inflation radius based on robot operating system [J]. Science Technology and Engineering, 2021;21(09):3662-8.

[5] Jiang Tao, Cheng XiaoSheng. Binocular vision Calibration Method based on homography matrix [J]. Journal of the optical, 2018,38(3): 353-61.

[6] Pang C T. Research on Robot Tracking and Grasping Technology based on Visual servo [D]: Jiangsu University; 2021.

[7] Ma S C. Research on Robot Grasping System Based on Visual Inspection [D]: Tianjin University of Technology; 2020.

[8] Herrero H, Pacheco R, Alberdi N, et al. Skills for vision-based applications in robotics application to aeronautics assembly pilot station[C] IEEE EUROCON 2015 - International Conference on Computer as a Tool. IEEE, 2015: 1-6.

[9] Zhang T, Ye J Y, Liu X G. Weighted Hand-eye Calibration Algorithm for Robot Belt Grinding [J]. Journal of mechanical engineering, 2018,54(17):142-8.

[10] Wang J T. Research on automatic sorting system of industrial robot based on hand-eye vision [J]. Industrial technology innovation, 2014,1(2):216-20.

[11] Xu Peng. Research on Visual Servo Control Method of Industrial Robot [D]. Changchun University of Technology, 2017.

[12] Lin C C , Gonzalez P , Cheng M Y , et al. Vision based object grasping of industrial manipulator[C]. International Conference on Advanced Robotics & Intelligent Systems. IEEE,Taipei, 2016.

[13] Song Y Y. Research on Industrial Robot Visual Feature Extraction Technology Based on Deep Learning [D]. Harbin. Harbin Institute of Technology, 2017.

[14] Minaeian S, Liu J, Son Y J. Vision-Based Target Detection and Localization via a Team of Cooperative UAV and UGVs[J]. IEEE Transactions on Systems Man & Cybernetics Systems, 2016, 46(7): 1005-16.

[15] Song Y Y. Research on Industrial Robot Visual Feature Extraction Technology Based on Deep Learning [D]. Harbin. Harbin Institute of Technology, 2017.

[16] Borji A, Cheng M M, Jiang H, et al. Salient Object Detection: A Benchmark[J]. IEEE Transactionson Image Processing, 2015, 24(12): 5706-22.

[17] Xia J, Qian K, Ma X D, et al. Fast pose detection for planar grasping robot based on cascaded convolutional neural network Robot[J]., 2018,40(6):794-802.

[18] Zhao Z, Huang P, Chen L. Visual tracking and grasping of moving objects and its application to an industrial robot[C]. IEEE International Conference on Real-time Computing and Robotics (RCAR). IEEE, Okinawa, 2017.

[19] Ji X Q, Wang J C, Zhao J D, et al. Intelligent assembly method of space borne equipment based on robot and vision Guidance,Journal of mechanical engineering, 2018,54(23):63-72.

[20] He Z Q. Design of Industrial Robot Sorting System based on Machine Vision [D]. Harbin Institute of Technology; 2016.