

# APPLICATION OF ROBOT SORTING TECHNOLOGY IN AUTOMATIC SORTING OF 3D VISION TARGET POSE MANIPULATOR

Sasa Zhao

Office of Quality Control and Performance Appraisal, Zibo Vocational Institute, Zibo, 255000, Shandong, China

Email: [zhaosasa\\_zbvi@163.com](mailto:zhaosasa_zbvi@163.com)

---

**Abstract:** A disordered sorting platform is designed to realize the application of robot sorting technology in the automatic sorting of 3D vision target positioning manipulator. First, the camera is used to detect the object to be sorted visually in order to obtain the pose information of the object. According to the environment and the specific information related to the Salsa object, the hardware of the sorting system is selected, and the process and scheme of the whole system are designed. Second, the calibration of the robot vision system composed of Kinect camera and ABB1200 robot is studied. The connection between the Kinect camera and the robot is determined as the installation mode of Eye-to-hand, and the Kinect camera itself is calibrated. Then, the Eye-to-hand system composed of the robot and the camera is calibrated, and the position relationship between the camera coordinate system and the robot base coordinate system is obtained. Subsequently, the pose estimation of randomly stacked objects is studied. The point cloud information collected by the camera is pre-processed, and the template is preliminarily matched with the target object by using the SHOT (Signature of Histogram of Orientation). And then the ICP (Iterative Closure Point) algorithm is used for further point cloud registration. Finally, the pose of the target object is estimated under the Eye-to-hand calibration results. The experimental platform of the whole unordered sorting system is built, and the multiple unordered sorting experiments of multiple objects are carried out, followed by the analysis of the experimental results. The error between the pose estimation results and the actual position is about 4mm, which verifies the feasibility of the system scheme and method and enables the robot to perform the unordered sorting task in centimetres, meeting the design requirements of the system. Therefore, the study provides a reference for future related research.

**Keywords:** Sorting technology; Vision system; Calibration; Position estimation.

---

## 1. Introduction

With the continuous development of China's economy, the automation production process based on the industrial robot is more common in parts assembly, cargo packaging and 3C industry [1]. More and more enterprises hand over the tasks that the previous factories complete manually to robots, which greatly saves the labour cost of the enterprise, improves production efficiency, and thus increases the economic benefits of the company [2]. Meanwhile, robots are also applied in the fields of automobile manufacturing and medical treatment, which ensures the workers' safety in a dangerous environment, and gradually changes the production mode, namely from the traditional workers as the leading force to the current industrial robot. Traditional robots mostly use teaching programming or offline programming to make robots do the point-to-point motion, and it is difficult to realize more complex motion trajectory, which cannot achieve the flexibility and automatic production of the pipeline [3,4]. Such traditional methods also have certain

requirements for the location of the grab objects. For example, when the shape or working environment of the grab object changes, the program needs to be rewritten, and the automation is not high. With the coming of the era of Chinese manufacturing in 2025, people attach great importance to intelligent technology. Nowadays, industrial robots are also developing towards intelligence. Mostly, the robot is combined with visual technology, enabling them to get the characteristics of these target objects, classify and identify them in the process of parts packaging, disassembly and transportation [5-7].

In the past decades, the research based on the 2D image is mature, and the vision technology based on the image is the most commonly used method in the target recognition, and it is also widely used in face recognition, medical detection and other fields [8]. It mainly applies the knowledge of pattern recognition and image processing to compare the common scenes of multiple images and process the image information, which includes the segmentation and feature extraction, to realize the target recognition and classification. Compared with 2D images, one-

dimensional information is added to 3D point cloud information of objects.

1) Deep information is added to 3D point cloud information to describe objects more comprehensively;

2) The feature extraction of point cloud information is not affected by scale, rotation and illumination;

3) 2D images based on vision technology can get 3DOF information of objects, while 3D vision technology can get 6DOF information. These advantages make the hot topic in the field of vision gradually enter the 3D vision.

With the continuous development of sensor technology, Microsoft launches the depth sensor Kinect a few years ago, which can obtain the three-dimensional data of the object surface and is convenient to use. Kinect is favoured by researchers in the fields of three-dimensional reconstruction, gesture recognition, human bone recognition and so on. Kinect camera can obtain the point cloud information of all the object surfaces and recognize and classify the target object in the process of vision processing. Besides, ASUS's Xtion and Intel's Realsense SR300 have the same functions and are also widely used in the field of scientific research [9].

The above information reflects that the acquisition and processing of 3D information of the target object is an indispensable technology in the development of robot intelligence [10]. The recognition and position estimation of the target object to be captured by the robot are studied. The Kinect camera is used to build a vision system to obtain the point cloud data, and the related algorithms are used to complete the recognition and segmentation of the target object to estimate the position of the target object itself, which has great significance to the development of robot intelligence.

## 2. Software and Hardware

### 2.1 Overall Design Scheme

The traditional production line adopts off-line programming for robot motion and has certain requirements for the placement of objects to be grasped, resulting in low production efficiency. The basic design scheme is as follows: (1) The teaching blocks are taken as the target objects to be captured in the experiment; (2) the Kinect camera and ABB1200 robot are used to build the whole experimental platform; (3) the 3D vision algorithm is applied to process the target objects to enable the robot to recognize the target objects; (4) the sorting task of the randomly stacked target objects in centimetre is completed. The design can improve the production efficiency. The flow chart of the overall scheme is shown in Figure 1:

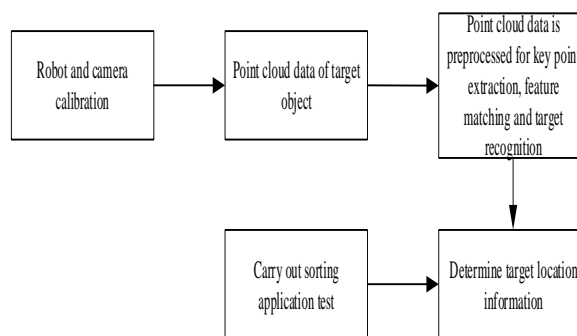


Figure 1: Overall design scheme

To realize the functions as expected, the hardware and software requirements for performing the function of automatic sorting of 3D vision target positioning manipulator should be met according to the flow chart in Figure 1.

#### (1) Hardware

The hardware mainly includes the ABB1200 robot, Kinect camera and gripper. The robot modelled IRB1200-5 / 0.9 is selected, and the maximum payload is 5kg respectively. It is widely used in small object assembly, and food, beverage and object transportation. The principle is shown in Figure 2 [11]. The speed of Kinect camera transmitting colour images and deep images can reach 30fps. It can capture the colour image and deep images in sight simultaneously. Based on the relationship between the deep map and the depth of spatial points, point cloud data can be obtained. The way to get the depth of Kinect camera is: the infrared emitter emits a laser in a specific area; the light shines on the surface of the object and then reflects back to the infrared camera to identify the current depth. The end effector of the robot system is SCHUNK's electric 2-finger translational machine modelled 0306101EGN100 made in Germany. The internal multi-tooth guide rail can provide a larger grasping force and torque load. And its minimum grasping force is 170N and the maximum is 720N, which can meet the needs of most users. The finger part at the lower end of the manipulator can be customized according to the actual situation [12].

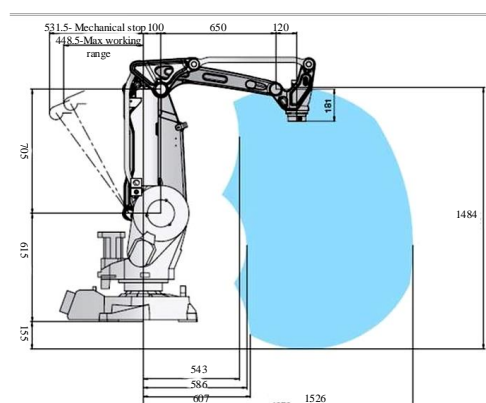


Figure 2: IRB100-5/0.9 rotor

(2) Software

With the rapid development of industrial automation, the pressure of market competition is also increasing. Enterprises need to reduce production costs and improve product quality to obtain more profits. However, it is impossible to spend time on industrial robot programming to test or operate at the beginning of producing a new product, because the industrial robot programming needs to stop all the processes of the products being produced, and then program the parts of the new product that need to be operated. The RobotStudio software can simulate the off-line programming of the robot. This process can be carried out, as well as the actual work of the robot, which improves work efficiency. In RobotStudio, many functions can be realized. For example, it can import various formats of CAD model data, and designers can program the current system with higher accuracy. It can also simulate the actual robot system and the working system of the industrial robot before the formal work, verifying the feasibility and safety of the scheme, and then performing the production, which provides a guarantee for the whole production process.

2.2 Robot and Camera Calibration

(1) Classification of the robot vision system

The camera and the end of the manipulator constitute the machine vision system. According to the fixed position of the camera, the hand-eye system is often divided into two types: Eye-in-hand and Eye-to-hand [13]. As shown in Figure 3, the camera of the Eye-in-hand system on the left is installed at the end of the manipulator. When the manipulator is working, the camera will move with the movement of the manipulator. The camera of the Eye-to-hand system on the right is installed at a fixed position outside the manipulator. When the robot is working, the camera will not move with the manipulator.

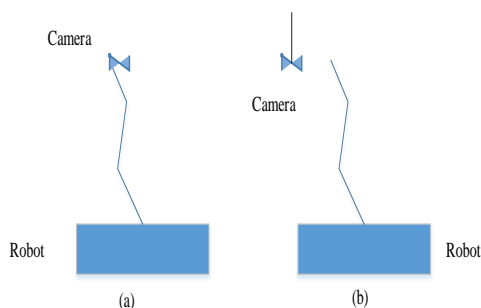


Figure 3: Hand-eye system (a: Eye-in-hand; b: Eye-to-hand)

Eye-in-hand is used more industrial robots. When the manipulator is closer to the target, the camera is closer to the target accordingly.

In this case, the absolute error of the camera measurement will be reduced and the accuracy will be improved. But in practical application, because the view field of the camera in the Eye-in-hand system changes with the movement of the manipulator, it cannot guarantee that the target is always in sight. And sometimes there may lead to target loss. The Eye-to-hand system is widely used in humanoid robots and mobile robots with the manipulator. The manipulator can operate the target according to the visual measurement, and it may also block the target [14]. Because the calibration of the Eye-to-hand system is relatively easy to achieve and it is used for the visual inspection in the whole sorting system.

(2) Calibration principle of robot and camera

Calibrating the Eye-to-hand system is to gain the relationship between the camera coordinate system and the robot coordinate system. This process is as follows: make the robot end in different poses; obtain the parameters at external and the target of the camera, and record the position and attitude information of the robot end in different poses; solve the external parameters of the camera relative to the robot coordinate system. In the process, the target is fixed at the end of the manipulator, and the camera is installed at a fixed position outside the robot. When the target moves with the robot, the pose of the robot end needs to be recorded. The robot can move to any position of the two, as shown in Figure 4.

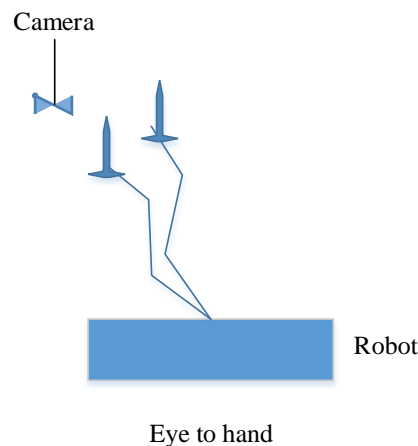


Figure 4: Motion diagram

The calibration experiment of the vision system is mainly divided into two parts, one is to calibrate the Kinect camera and get the internal and external parameters of the camera, the other is to calibrate the hand-eye system composed of the camera and the robot to get the external parameters of the camera relative to the robot coordinate system [15]. The Eye-to-hand camera installation method is used to build the vision system, that is, the Kinect camera is installed in a fixed position outside the manipulator, and the calibration plate is fixed at the end of the manipulator.

## 2.3 Point Cloud Data Processing and Recognition of Target Objects

The main formats of point cloud data are PCD, PLY, STL, OBJ. The point cloud data collected by Kinect camera is PCD format. Generally, the point cloud data collected by 3D acquisition devices like radar and laser scanning are unevenly distributed, and the number of data is large. These data are different from the traditional solid grid data, and there is no geometric topology information between the points. The topological relationship between these points is established to reach the fast search between the neighbourhood areas. In addition, point cloud data is mainly used to record the information of the object surface. It is generally based on the surface normal and curvature of point cloud data to describe the features of points and data.

### (1) Point cloud data pre-processing

Point cloud data collection is generally affected by SUCH factors as noise in the environment, nature of target object surface, camera accuracy and researcher operation, which will cause the data far away from the target object group point [16]. Before processing the point cloud data, the data should be filtered first to reduce the error brought by these factors in the experiment. The object to be identified is the teaching block, which is widely used. The side length of the square block is 5cm, and all of them are placed on a plane. After the camera collects the point cloud data in the whole scene, the background needs to be modelled and removed to extract the information of the target object. The more common RANSAC (RANDOM SAMPLE CONSENSUS) algorithm is selected, that is, a random sampling consistency algorithm is used to segment the carrier plane. In short, the algorithm is an iterative method to estimate the parameters of the mathematical model. The flow chart of the algorithm is shown in Figure 5:

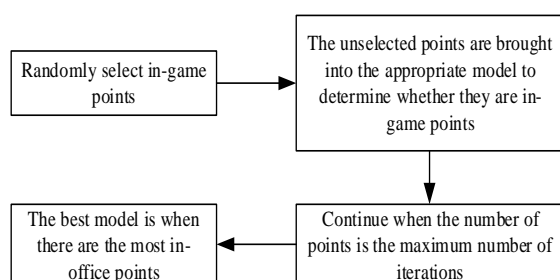


Figure 5: The flow chart of the algorithm

RANSAC method is used to extract the point cloud plane. Assume that the plane model equation is  $ax+by+cz+d=0$ , and the purpose of using RANSAC is to find the four parameters  $a, b, c,$  and  $d$  in the model. The specific fitting process is as follows:

- 1) Three non-collinear points  $\{p_1, p_2, p_3\}$  in the point cloud scene are randomly selected;
- 2) A plane can be obtained from the three points above;

- 3) For any point  $P$  in the point cloud, the distance from the point to the plane is calculated. If the distance is less than the set threshold, then  $p$  is the focal point;

- 4) Traverse the points in the whole scene, calculate the distance from the point to the plane, select the interior points, and count the number of interior points;

- 5) Repeat the above steps to the maximum number of iterations, and select the plane model with the most interior points as the optimal fitting of the background plane;

- 6) The set of all points from the point to the optimal plane is calculated as the point cloud data of the plane in the scene;

The load plane model is obtained by using the random sampling consistent algorithm, and then the plane is removed by using the direct filter, which performs on the  $z$ -axis and removes the points outside the  $z$ -axis. The point cloud data of the target object in the scene has a certain distance from the point cloud data of the background of the load plane. The random sampling consistent algorithm is used to fit the load plane, on which the load plane of the background is removed. Meanwhile, the invalid points in the origin cloud are deleted, and only the point cloud data of the object to be sorted is left in the scene, which is convenient for the following segmentation and matching.

Because the target object used is a cube block, there are three situations in which the camera can see the object, that is, when the object is placed horizontally, one face can be seen, and when the object is placed obliquely, two or three faces can be seen. The template is created in three cases using the same method.

### (1) Key point extraction based on ISS3D algorithm

When the point cloud data are matching, some representative points are usually used to represent the information of the whole model. The reduction of the number of points also accelerates the processing speed of the program. Feature point extraction is to use these representative feature points to complete the matching. The feature points are applied to match the objects in disorder with a few obstacles, gaining a good recognition effect [17]. In 2D image processing, key point extraction algorithms include SIFT (scale invariant feature transform), FAST (features from accelerated segment test), and ORB (oriented fast and rotated brief). In 3D, many key point extraction algorithms come from 2D algorithms, such as SIFT3D, HARRIS 3D, ISS3D. The literature shows that the ISS3D (intrinsic shape signatures) algorithm is more stable for key point extraction, faster and accurate in object surface description.

Based on the above, the ISS3D algorithm is selected to extract the key points of the template

point cloud data and the target object, according to the three models in (2).

(2) Feature matching based on SHOT

After the feature of the SHOT is explained in detail and the key points of the template and the target point cloud are extracted, the SHOT feature is used for the matching experiment. KD tree is adopted to calculate the nearest feature in the template point cloud feature set, gaining a matching set. In the matching process, since the Geometric Consistency Grouping will produce one or more matching results, the Global Hypotheses Verification is applied to get an optimal matching result for subsequent fine registration [18, 19]. The template is created by taking a single object offline and then removing the carrier plane. The initial pose of the template is the actual pose of the object in the robot base coordinate system when the camera collects the template data and then is combined with the hand-eye calibration results.

(3) Target recognition based on ICP algorithm

The descriptors of each model in the cloud database are matched. Therefore, the kd-tree search structure is used to perform the nearest neighbour search, retrieve the Euclidean distance between descriptors, and select and set the maximum distance value as the threshold. After feature point matching, it belongs to coarse registration, and then fine registration can be carried out by using the classic ICP. The registration is equivalent to the recognition of the current target [20-22]. This method is a matching method of point set relative to the point set and calculates the optimal coordinate transformation iteratively through the least square method, that is, rotate matrix R and translate vector t to minimize the error and calculate the pose change of the object relative to the template in the scene.

### 3. Visual System Calibration and Target Sorting Experiment Results

#### 3.1 Calibration Experiment of the Vision System

Kinect camera is used to take 10 checkerboard images under different postures as calibration images under Windows 7 system. The final calibration results are shown in Figure 6.

Figure 6 indicates the reprojection error of the target image. The error of each image is less than 0.15 pixels, and the overall average error is 0.132 pixels, showing a good effect.

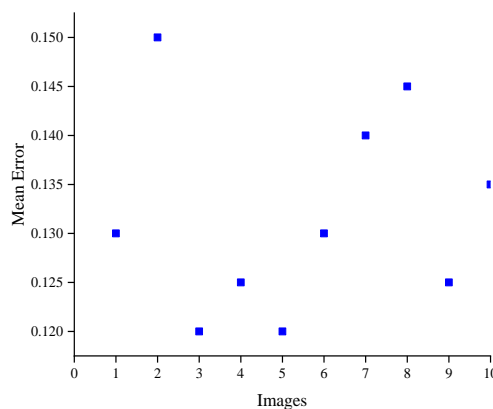
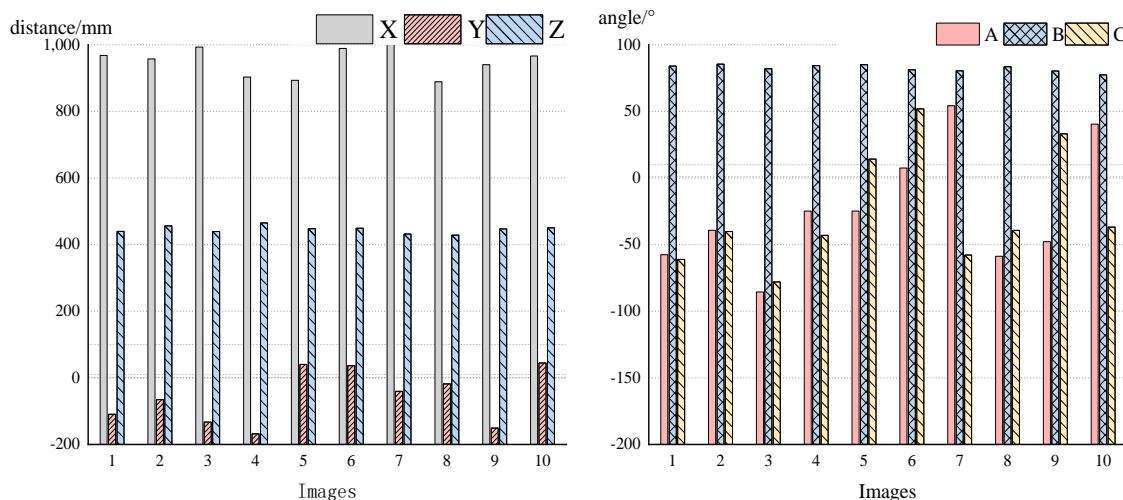


Figure 6: Key point extraction results of three types of templates

When the camera is calibrated to capture the target image, record the position and pose of the corresponding manipulator end effector relative to the robot base coordinate system, as shown in Figure 7.



(a)

(b)

Figure 7: Information of robot end corresponding to the target image (A: position information; B: angle information)

### 3.2 Random Sorting Experiment of Target Objects

The pose of each small block to be grasped in the robot base coordinate system is statistically sorted out. X, Y and Z represent the position of the centre of mass of the point cloud on the surface of the object in

the robot coordinate system, and A, B and C represent the rotation angle of the target around the x, y and Z axes [23,24].

The comparison between the position and angle information of the target object to be sorted and the information of the actual grab is shown in Figure 8.

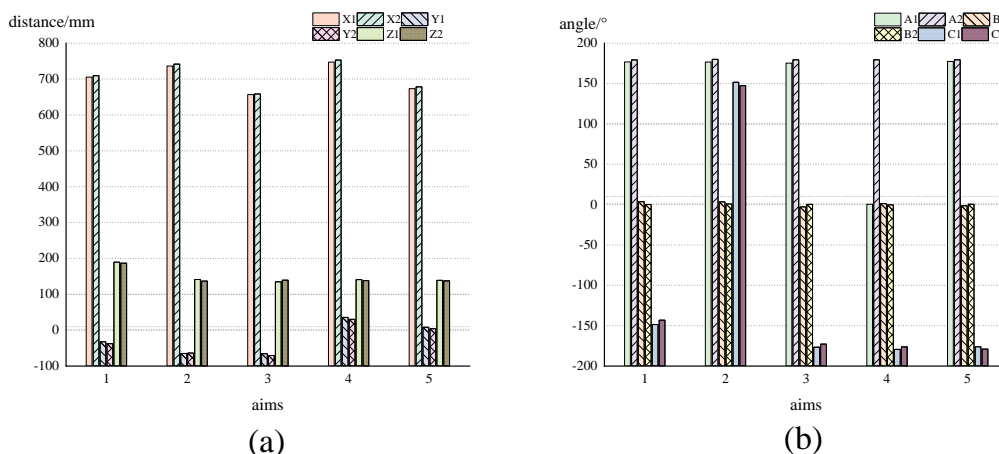


Figure 8: Comparison between the information to be sorted and the actual information to be grasped (A: position comparison; B: angle comparison)

Figure 8 shows the actual pose information of the object in the scene in the robot base coordinate system can be read from the robot teaching device, and the robot can be manually operated to move its end to the designated position for recording. The rotation angle of the target around any of the x, y, and z axes is around 180° [25-27]. Therefore, the error between the pose of the object obtained by the program matching and the actual pose of the object in the scene is shown in the figure:

Table 1 Errors

Position error/mm		Angle error /°	
X	4.12	A	3.02
Y	3.56	B	2.86
Z	3.87	C	3.97

Table 1 shows that through the statistics and analysis of the pose error results of the target object, the average error on the x, y and z axes is 4.12mm, 3.56mm and 3.87mm respectively, and the overall average position error is 3.85mm. The angle error around the x, y and z axes is 3.02°, 2.86° and 3.97° respectively, and the overall average angle error is 3.28°. Through the experiments on multiple target objects and the analysis of experimental data, the target recognition matching algorithm and the robot sorting system designed based on ICP can complete the sorting of randomly stacked target objects and the sorting task in centimetres [28,29].

### 4. Conclusion

The workpiece recognition and sorting in the application field of an industrial robot is always a hot research topic. The disordered sorting technology of industrial robots based on target recognition is studied. The hardware experimental platform is built by using a Microsoft Kinect camera, robot and XiongKe claw. The program is written on the windows 7 operating system combined with PCL point cloud library and RobotStudio software, and the results and the experiment verify the feasibility of the method. According to the research content, the following research work is done: (1) read a large number of domestic and foreign literature, establish the research background and significance, analyse the research status of some domestic and foreign companies applying robot sorting technology, summarize and compare the research status by target object detection technology, and the whole content is displayed. (2) According to the overall design of the sorting system, the experimental hardware is selected and the Kinect camera and robot are used to build the robot vision system and the camera parameters are obtained. The robot vision system is calibrated by using the camera parameters and the robot pose information recorded in the calibration picture, and the position relationship between the camera and the robot is obtained. (3) The Kinect camera is used to collect the scene information of the target object and pre-process the obtained point cloud data. Furthermore, the template of the target object is established to

remove the background of the target object and the template, and the key points of the point cloud data are extracted. The extracted key points are described by SHOT features, and the preliminary registration is carried out. Finally, the ICP (Iterative Closure Point) method is used for fine registration. Experiments are designed to verify the feasibility of the method. (4) The scheme of the whole sorting system is verified by experiments. According to the actual information of the target object, the hand claw and other equipment are programmed and debugged in advance, and multiple target objects in the scene are sorted. Finally, the overall average position error and average angle error are 3.92mm and 3.14 ° respectively, which verifies the feasibility of this method and the disordered sorting system.

### Acknowledgment

This work was supported by Sciences Planning Project of ministry of Education (Subject No. 16YJA880050).

### References

- [1] Yang K, Yang W., & Wang C. (2018) Inverse dynamic analysis and position error evaluation of the heavy-duty industrial robot with elastic joints: an efficient approach based on lie group. *Nonlinear Dynamics*, 93(4), 1-18.
- [2] Yao B., Zhou Z., Wang L, et al. (2018) Sensorless and adaptive admittance control of industrial robot in physical human robot interaction. *Robotics & Computer Integrated Manufacturing*, 51(jun.), 158-168.
- [3] Zhang G. H., Li K., Zhao B, et al. (2019) Clinical effects of robot-assisted total knee arthroplasty compared with traditional surgery: a meta-analysis. *Zhongguo gu shang China journal of orthopaedics and traumatology*, 32(9), 846-852.
- [4] Liu Z., Liu N., Wang H, et al. (2020) A new type of industrial robot trajectory generation component based on motion modularity technology. *Journal of Robotics*, 2020(4) 1-11.
- [5] Jakubiec B. (2020). Virtual environment as a tool for analysing the operation of an industrial robot. *Przeglad Electrotechnics*, 96(2), 98-101.
- [6] Wang Z. (2020) Design and implementation of polishing and polishing flexible workstation based on industrial robot. *International Core Journal of Engineering*, 6(4), 193-197
- [7] Brunot M., Janot A., Young P, et al. (2018) An improved instrumental variable method for industrial robot model identification. *Control Engineering Practice*, 74(MAY), 107-117.
- [8] Suszyński M., Rogalewicz M. (2020) Selection of an industrial robot for assembly jobs using multi-criteria decision making methods. *Management and Production Engineering Review*, 11(1), 62-72.
- [9] Thu T L T., Quoc K D., Thanh L P. (2020) Calibration of industrial robot kinematics based on results of interpolating error by shape function. *Journal of Engineering and Applied Sciences*, 15(6), 1451-1461.
- [10] Polivanov A Y., Ivanov Y. V., Kholin D. V. (2020) The method of the computer vision system coordinate transformation for an industrial robot for a laser welding operation. *Mekhatronika Avtomatizatsiya Upravlenie*, 21(3), 166-173.
- [11] Gao H., Lu S., Wang T. (2020) Motion path planning of 6-dof industrial robot based on fuzzy control algorithm. *Journal of Intelligent and Fuzzy Systems*, 38(3), 1-10.
- [12] Cerva M., Liberto M., Gurreri L, et al. (2017) Coupling Cfd With A One-Dimensional Model To Predict The Performance Of Reverse Electrodialysis Stacks. *Journal of Membrane Science*, 5(41), 595-610.
- [13] Kuts V., Cherezova N., Sarkans M. (2020) Digital twin: industrial robot kinematic model integration to the virtual reality environment. *Journal of Machine Engineering*, 20(2), 53-64.
- [14] Chemogonov P., Makarov A., Potseluko V, et al. (2020) Design and research of an automated laboratory stand based on a prototype of the "hand" type industrial robot. *IZVESTIA VOLGOGRAD STATE TECHNICAL UNIVERSITY*, 3(238), 93-96.
- [15] Yan Z., Ouyang B., Li D, et al. (2020) Network intelligence empowered industrial robot control in the f-ran environment. *IEEE Wireless Communications*, 27(2), 58-64.
- [16] Xiao N., Zhang X. (2020) A target positioning method for industrial robot based on multiple visual sensors. *Traitement du Signal*, 37(3), 469-475.
- [17] Krinitsyn N S., Babaev A S., Stolov E V, et al. (2020) Precision in single-coordinate positioning of an industrial robot manipulator. *Russian Engineering Research*, 40(1), 83-85.
- [18] Truc L N., Nghia N T., Thanh N T, et al. (2020) Effect of actuator torque degradation on behavior of a 6-dof industrial robot. *Universal Journal of Mechanical Engineering*, 8(2), 114-128.
- [19] Kim Y., Park J., Na K, et al, (2020) Phase-based time domain averaging (ptda) for fault detection of a gearbox in an industrial robot using vibration signals. *Mechanical systems and signal processing*, 138(Apr.), 106544.1-106544.19.
- [20] Chen S C., Huang C C. (2019) Application of robotic arm for medical surgery positioning aid. *Journal of China University of Science and Technology*, 1(77), 1-21.
- [21] Jin H., Fan W., Chen H, et al. (2020) Anti-corrosion wood automatic sorting robot system based on near-infrared imaging technology. *Journal of Mechanical Science and Technology*, 34(7), 3049-3055.

- [22] Grambo P., Mullick T., Furukawa T, et al. (2019) Automatic sorting-and-holding for stacking heterogeneous packages in logistic hubs - science direct. IFAC-PapersOnLine, 52(10), 109-114.
- [23] Kirschenstein M., Krasuski K., Go A. (2020) Methods of precise aircraft positioning in the gps system with an application of the troposphere correction. Scientific Journal of Silesian University of Technology. Series Transport, 109(4), 73-84.
- [24] Lee D., Kim S., Kim P, et al. (2018) Automatic sea squirt sorting algorithm based on the hsv colour model and weight estimation. Journal of Intelligent and Fuzzy Systems, 35(2), 1-8.
- [25] Aung N P., Wai M M., M Htay L. (2018) Model of metal sorting conveyor system using siemens s7-1200 plc. International Journal of Scientific & Technology Research, 7(8), 45-49.
- [26] Jiang G., Luo M., Bai K. (2019) Optical positioning technology of an assisted puncture robot based on binocular vision. International Journal of Imaging Systems & Technology, 29(2), 180-190.
- [27] Meng O. K., Pauline O., Soon L. E, et al. (2018) Robotic arm system with computer vision for colour object sorting. International Journal of Engineering & Technology, 7(4), 50-56.
- [28] Afanasiev M. Y., Fedosov Y. V., Krylova A, et al. (2020) Application of machine vision in the tasks of automatic positioning of modular equipment tool. Izvesti vysih uebnyh zavedenij Priborostroenie, 63(9), 830-839.
- [29] Wataru H., Hajime S. (2019) A study on an optimal coil positioning in a wireless power transfer system with a function of human tracking. Journal of the Japan Society of Applied Electromagnetics and Mechanics, 27(3), 302-307.