

# Vision-Guided Recognition Method for Working State of Robot Arm based on Machine Learning

Peibo Li<sup>1</sup>, Peixing Li<sup>2</sup>

<sup>1</sup>School of Mechanical Engineering, Donghua University, Shanghai, 201620, China

<sup>2</sup>School of Mechanical and Automotive Engineering, Shanghai University of Engineering Science, Shanghai, 201620, China

## Abstract

*To realize the capture of the target in any position within the working range of the robot arm, a vision-guided target recognition and positioning control method based on machine learning was proposed to improve the accuracy of the working state recognition. The system grabs four different contours of workpieces (triangle, pentagon, round, square and place in the designated area as the task, and by using the MATLAB to process the image information, and all the connected domains are marked by the neighborhood area marking algorithm. Later, the logarithmic coordinates-Fourier transform template matching method is adopted to identify workpiece types and extract their centroid as the positioning reference coordinates. Combined with the 3-DOF robot arm, the standard D-H parametric method is used to establish the robot arm kinematics model, and through the inverse operation of the robot arm and according to the position of the workpiece coordinates, the joint angle of each robot arm can be obtained. Later, it is sent to the lower microcontroller Arduino through a serial port, and then the control instruction is completed by the Arduino to realize the capture and placement of the workpiece and complete the work state recognition task. The experimental results show that the working state recognition system can meet the design requirements.*

*Key words: machine learning; Visual guidance; Mechanical arm; recognition methods*

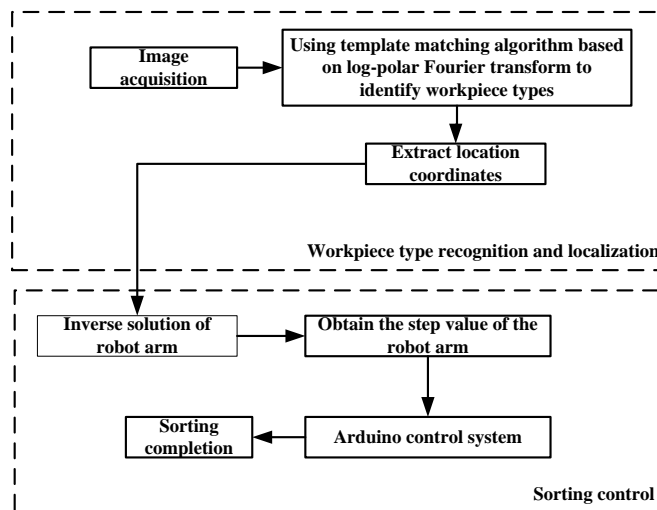
## I. Introduction

The working state recognition technology of robot arm has been applied to the operation of various factories. Traditional robot arm operation can complete complex point-to-point movement through accurate teaching operations, which significantly improved the working state recognition efficiency. However, this method requires the position of the grasping target to be relatively fixed, and when the target position deviation occurs, the robot arm will fail to grasp the target. To realize that the robot arm can automatically obtain the target in any position within its working range and adjust the position of the end-effector, and avoid the problem of grasping failure caused by position deviation, a target recognition and location method based on machine vision should be introduced. In this way, the system can automatically obtain the target picture through the camera and obtain the required target information through the algorithm analysis. Taking the 3-DOF robot arm as an example, and with four kinds of different shapes of workpiece work state recognition as a work state recognition task, the target recognition, centroid coordinates and positioning technology based on machine vision technology was studied. Later, combined with the kinematics model analysis of the 3-DOF robot arm, a method to realize the working state identification system was presented [1-3].

## II. Overall design of working state recognition system

The robot arm intelligent working state recognition system based on machine vision mainly includes workpiece recognition and robot arm working state recognition control. The main process of part of workpiece recognition is to collect image information from the camera and send it to Matlab, and then the type of workpiece is recognized and

the position coordinates are extracted by the recognition algorithm. As for the robot arm working state control part, by the inverse operation of the robot arm and according to the extracted workpiece coordinates, the joint angles of 3 connecting rods can be obtained, and they are converted into the step values of the robot arm. Then the way of serial port communication is used to send the step values to the Arduino microcontroller, and the activities of the robot arm are controlled by the Arduino, thus realizing the picking and placing operation of the workpiece and completing the working state recognition task. Figure 1 shows the system flow diagram[4].



*Fig. 1 Block diagram of working state recognition system*

### III. Workpiece recognition and positioning design

#### 3.1 Mark connected areas

Recognizing multiple objects in the image and extracting coordinates is the purpose of image preprocessing. The images collected by the camera are conducted with grayscale and binarization processing by Matlab, but there are still many small noise points. To better extract target features, the image after binarization processing should be reprocessed through morphological filtering operations such as corrosion filling, so as to reduce small noise points in the image, obtain the binary image after morphological filtering, and finally analyze the target object in the image. However, the specific analysis and recognition of the target are different from the grayscale and binarization of the image. The latter can process the image as a whole, while the former needs to divide all the objects in the image into individuals for analysis. Therefore, each target in the binary image after morphological filtering needs to be divided into an independent connected area, and then these connected regions are marked. Based on this, the number of targets in the image can be detected, and the parameters such as target types and coordinates of each target can be extracted[5].

The system uses the four-neighborhood marker algorithm introduced in Literature to calculate all the connected areas in the image, and then marks the connected areas with continuous natural numbers in the order of top to bottom and left to right and creates the minimum encasing rectangle, making preparations for the subsequent recognition of workpiece types and extraction of centroid coordinates[6-8].

#### 3.2 Workpiece type recognition

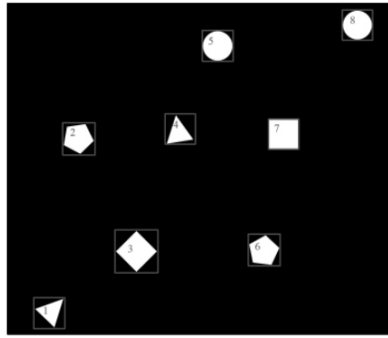
##### 3.2.1 Image cutting

The position of the workpiece on the workbench is arbitrary. It is necessary to cut the image whose connected area is

previously marked into N(number of workpiece) small pictures according to the minimum outer rectangle of each target, and then analyze the content of each picture, respectively.

### 3.2.2 Workpiece type recognition algorithm

Given the rotation changes of the workpiece during placement, this system adopts the template matching algorithm of Log-PolarTransformationFourier Transformation to identify the type of the workpiece. After the log polar transformation of the recognized image, the rotation change of the image can be converted into the translation change, and then the rotation invariance of the image can be realized by combining with the modulus invariance of the Fourier translation transformation.



**Fig. 2 Marked image of connected area**

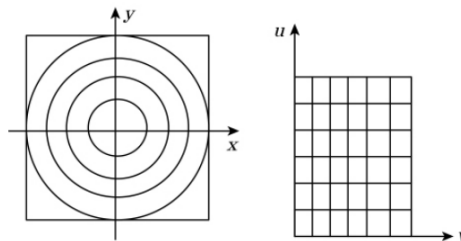
#### Log-polar transformation

Log-polar transformation of the image aims to convert the Cartesian coordinate system to the logarithmic coordinate system, and its corresponding transformation formula is as follows

$$\begin{cases} u = k \cdot \ln \rho = k \cdot \ln(\sqrt{(x - x_0)^2 + (y - y_0)^2}) \\ v = m \cdot \theta = m \cdot \arctan[(y - y_0) / (x - x_0)] \end{cases} \quad (1)$$

Where  $(x, y)$  and  $(u, v)$  are the coordinates of the image before and after transformation, respectively;  $(x_0, y_0)$  is the coordinate transformation center;  $k$  and  $m$  are the amplification parameters introduced to enlarge the transformed image.

The corresponding relation of Cartesian - log polar coordinate system is shown in Figure 3.



**Fig. 3 Cartesian - log polar coordinate system**

According to Formula(1), when the original image rotates a radian of  $\phi$ , the following formula can be obtained:

$$v' = m \cdot (\theta + \phi) = v + m\phi \quad (2)$$

At this moment,  $v'$  represents the abscendant of the rotated image in log-polar coordinates. Thus, a rotation change in Cartesian coordinate system can be transformed into a translation change in log-polar coordinate system. Fig. 4 shows the translation transformation of the same kind of workpiece at different placement angles.

#### Discrete Fourier Transform

In the time domain, the modulus of the Fourier transform has a translation invariance. When carrying out image processing, the signal given by the image is two-dimensional discrete, so the two-dimensional discrete Fourier transform is adopted. It is assumed that the two-dimensional discrete Fourier transform  $F(u, v)$  of the function  $f(x, y)$  of the picture with a size of  $M \times N$  is as follows:

$$F(u, v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi(ux/M + vy/N)} \quad (3)$$

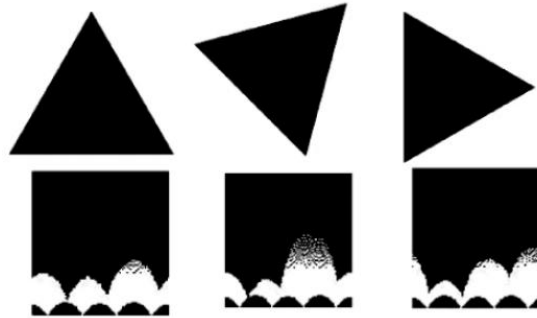


Fig. 4 Translation transformation in log-polar coordinates

After image translation,  $f(x - x_0, y - y_0)$  is transformed into

$$F'(u, v) = F(u, v) \cdot e^{-j2\pi(ux_0 + vy_0)} \quad (4)$$

The modulus before and after translation is obtained:

$$|\text{fft } 2[f(x, y)]| = |\text{fft } 2[(x - x_0, y - y_0)]| \quad (5)$$

From Formulas (1) to (5), the rotation invariance of the image can be seen.

As the four kinds of workpieces selected by the system are symmetric, and there isn't any other design on the surface, so besides the circular workpiece only takes 0° state diagram, the other three kinds of workpieces choose seven different rotating images (0°, 15°, 30°, 45°, 60°, 75°, 90°) to constitute a template library containing 22 images.

To sum up, the recognition part of workpiece types in this system can be divided into the following processes:

- (1) Create a template library.
- (2) Gray scale, binarization and other operations are first conducted to the whole image, and then each independent workpiece is marked with the connected area and cut into the corresponding several small images.
- (3) The image to be recognized and the image in the template library are transformed into log-polar coordinates and the correlation coefficient is obtained, and the workpiece with the largest correlation coefficient is matched with it. The recognition results of triangle, pentagon, circle and square are represented by numbers 1 ~ 4.
- (4) Repeat steps (2) ~ (3) for the next recognition.

### 3.3 Workpiece location positioning

#### 3.3.1 Extraction of centroid coordinates

After completing the workpiece recognition, it is necessary to locate the target in the image. Firstly, the centroid

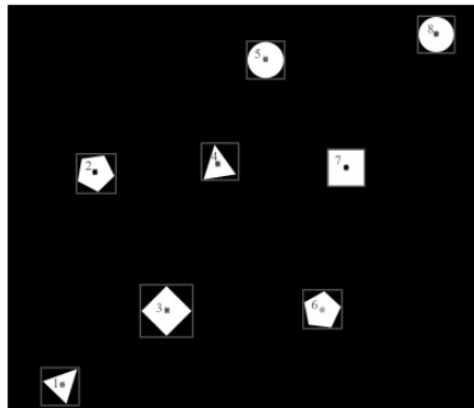
coordinates of the workpiece on the image are obtained. In the binary image, the target area is white, which is represented by 1, while the non-target area is black, which is represented by 0, as shown in Formula (6). Formula (7) is the centroid calculation formula of the target image.

$$f(x, y) = \begin{cases} 1 & \text{target area} \\ 0 & \text{Non target area} \end{cases} \quad (6)$$

$$x_0 = \frac{\sum xf(x, y)}{\sum f(x, y)}, y_0 = \frac{\sum yf(x, y)}{\sum f(x, y)} \quad (7)$$

Where  $x_0, y_0$  is the centroid coordinate of the target workpiece;  $f(x, y)$  is the pixel value of the point  $(x, y)$  on the binary image;  $xf(x, y), yf(x, y)$  are the pixel values of  $x$  and  $y$  coordinates of the point  $(x, y)$  in the binary image, respectively .

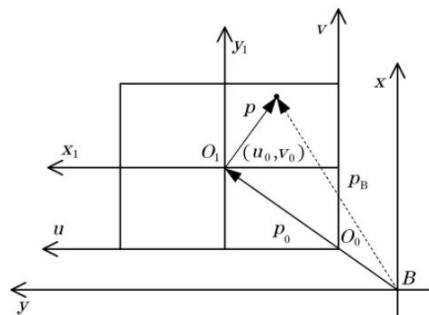
According to Formulas (6) and (7), the centroid coordinates of each connected area in the binary image can be obtained, that is, the pixel coordinates of the target object in the imaging coordinate system. Fig. 5 shows the pixel coordinate points of the workpiece obtained.



**Fig. 5 Pixel position coordinate points of the workpiece**

### 3.3.2 Coordinate transformation

The centroid coordinates obtained can not be directly used as the positioning coordinates of the workpiece, and they need to be transformed from the pixel coordinate system to the base coordinate system of the robot arm. As shown in Fig. 6,  $\{O_0\}$  is the pixel coordinate system,  $\{O_1\}$  is the physical coordinate system, and  $\{B\}$  is the base coordinate system.



**Fig. 6 Coordinate System**

$(u_0, v_0) = (640, 512)$  is the center point of the whole image.  $x_1$ 、 $y_1$  represent the actual physical position of the workpiece under  $\{O_1\}$ , and  $dx_1$ 、 $dy_1$  represents the unit pixel length of the coordinate axis direction of  $x_1$  and  $y_1$ , respectively, so as to obtain the existing relationship of each pixel in the coordinate system of  $\{O_0\}$  and  $\{O_1\}$ :

$$\begin{cases} u = x_1 / dx_1 + u_0 \\ v = y_1 / dy_1 + v_0 \end{cases} \quad (8)$$

The representation of each pixel under  $\{O_1\}$  is obtained:

$$\begin{cases} x_1 = (u - u_0) \cdot dx_1 \\ y_1 = (v - v_0) \cdot dy_1 \end{cases} \quad (9)$$

According to Formula (9), the actual physical position of the workpiece on the image can be obtained. However, in order to be used as the positioning coordinates of the end of the robot arm to finally grasp, it needs to be transformed into the representation in the base coordinate system. As for the 3-DOF robot arm of the system, the end-effector only needs to describe its position. Through Formula (10), the base coordinate system  $\{B\}$  and  $\{O_1\}$  can have the same direction, so that there is only a translation relationship between the two coordinate systems.

$$\begin{cases} x_1 = (v - v_0) \cdot dy_1 \\ y_1 = (u - u_0) \cdot dx_1 \end{cases} \quad (10)$$

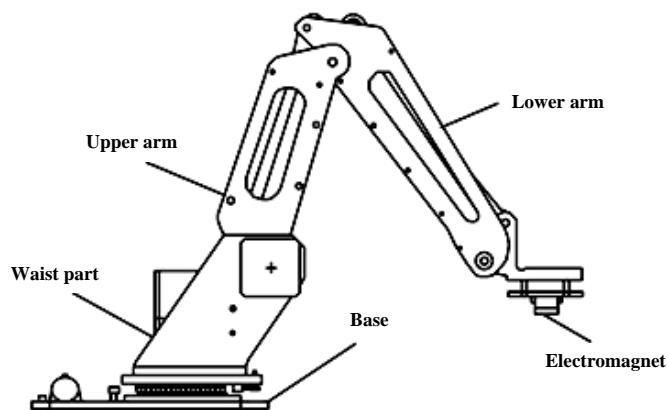
It is assumed that the position vector  $\mathbf{p}_0 = [x_0 \ y_0 \ z_0]^T$  is the position of  $\{O_1\}$  relative to  $\{B\}$ , and  $\mathbf{p} = [x \ y \ z]^T$  is a point in  $\{O_1\}$ , then the position  $\mathbf{p}_B = [p_x \ p_y \ p_z]^T$  of point p relative to the coordinate system  $\{B\}$  can be obtained by adding the vectors, namely:

$$\mathbf{p}_B = \mathbf{p} + \mathbf{p}_0 \quad (11)$$

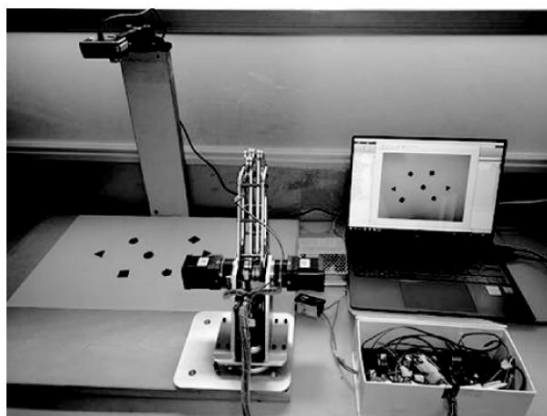
The pixel coordinates of each workpiece calculated according to Formulas (6) and (7) are substituted into Formulas (9) - (11) to obtain the final positioning coordinates of the workpiece in the base coordinates. Then the horizontal ordinate values are placed in the vectors  $\mathbf{p}_x, \mathbf{p}_y$ , respectively. As the robot arm needs to identify the target object in three-dimensional space coordinates during the capturing process, so a set of vectors  $\mathbf{p}_z$  need to be set to represent the height of the workpiece in three dimensional space.

#### IV. Control part of robot arm

The structure of the 3-DOF robot arm is shown in Figure 7. The waist parthas a rotating motion on the horizontal plane, and the upper arm and lower arm have a pitch motion perpendicular to the horizontal rotation plane of the base. An electromagnet is installed at the end of the lower arm, and the grasping and placing operation of the workpiece is realized by controlling the power gain and loss of the electromagnet. Figure 8 shows the system construction diagram.



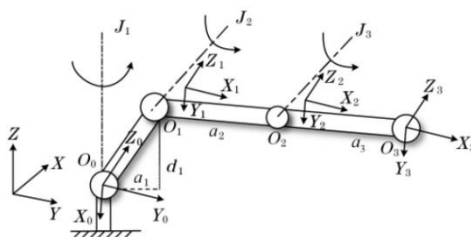
*Fig. 7 Structure chart of the robot arm*



*Fig. 8 System construction diagram*

To make the end of the robot arm move to the top of the workpiece accurately, it is necessary to control the rotation angle of each joint. Therefore, the system needs to substitute the solved positioning coordinates into the inverse operation of the robot arm to obtain the rotation angle of each mechanical connecting rod. The link coordinate system is established for the robot arm of this system, as shown in Fig. 9. Then, according to the connecting rod parameters shown in Table 1, the standard D-H method is adopted to establish the kinematics model of the 3-DOF robot arm.

The forward solution of the robot arm is obtained according to the transformation rules between the coordinate systems of adjacent two connecting rods shown in Table 1 and Formula(12), and the transformation matrix between the end of the robot arm and the base coordinate system can be obtained, as shown in Formula (13).



*Fig. 9 Coordinate system of connecting rod of robot arm*

**Table 1 D-H parameter list of each joint**

Connecting rod	$\theta_i$ Joint angle	/mm Offset $d_i$ /mm	/mm Length connecting rod $a_i$ /mm	Torsion angle $\alpha_i$
1	$\theta_1$	$d_1 = 102$	$a_1 = 62$	$\pi / 2$
2	$\theta_2$	0	$a_2 = 155$	0
3	$\theta_3$	0	$a_3 = 180$	$-\pi / 2$

$${}^i T_i = \begin{bmatrix} \cos \theta_i & -\sin \theta_i \cos \alpha_i & \sin \theta_i \sin \alpha_i & a_i \cos \theta_i \\ \sin \theta_i & \cos \theta_i \cos \alpha_i & -\cos \theta_i \sin \alpha_i & a_i \sin \theta_i \\ 0 & \sin \alpha_i & \cos \alpha_i & d_i \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (12)$$

$${}^0 T_3 = {}^0 T_1 T_2 T_3 = \begin{bmatrix} \mathbf{n}_x & \mathbf{o}_x & \mathbf{a}_x & p_x \\ \mathbf{n}_y & \mathbf{o}_y & \mathbf{a}_y & p_y \\ \mathbf{n}_z & \mathbf{o}_z & \mathbf{a}_z & p_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (13)$$

Where  $\mathbf{n}_x$ 、 $\mathbf{n}_y$ 、 $\mathbf{n}_z$ 、 $\mathbf{o}_x$ 、 $\mathbf{o}_y$ 、 $\mathbf{o}_z$ 、 $\mathbf{a}_x$ 、 $\mathbf{a}_y$ 、 $\mathbf{a}_z$  are the direction vectors of the end coordinate system axis of the robot arm in coordinate system B, respectively;  $p_x$ 、 $p_y$ 、 $p_z$  are the coordinates of the end position of the robot arm in the base coordinate system, respectively.

Then, the inverse solution of the robot arm is obtained, and both ends of Formula (13) are left multiplied by the unknown connecting rod inverse transformation simultaneously to separate the joint variables. Thus,  ${}^0 T_1^{-1} \cdot {}^0 T_3 = {}^0 T_2 \cdot {}^0 T_3$ , namely:

$$\begin{bmatrix} c_1 & s_1 & 0 & -a_1 \\ 0 & 0 & 1 & -d_1 \\ s_1 & -c_1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \mathbf{n}_x & \mathbf{o}_x & \mathbf{a}_x & p_x \\ \mathbf{n}_y & \mathbf{o}_y & \mathbf{a}_y & p_y \\ \mathbf{n}_z & \mathbf{o}_z & \mathbf{a}_z & p_z \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} & b_{13} & a_2 c_2 + a_3 c_2 c_3 - a_3 s_2 s_3 \\ b_{21} & b_{22} & b_{23} & a_2 s_2 + a_3 c_2 s_3 + a_3 c_3 s_2 \\ b_{31} & b_{32} & b_{33} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (14)$$

Let the elements on both sides of the Formula (1,4)、(2,4)、(3,4) be equal to each other, and then the following formula can be obtained:



$$\begin{cases} p_x s_1 - p_y c_1 = 0 \\ p_x c_1 + p_y s_1 - a_1 = a_2 c_2 + a_3 c_2 c_3 - a_3 s_2 s_3 \\ p_z - d_1 = a_2 s_2 + a_3 c_2 s_3 + a_3 c_3 s_2 \end{cases} \quad (15)$$

Then:

$$\begin{cases} \theta_1 = \arctan(p_y / p_x) \\ \theta_2 = \arctan(A / C) + \arctan(\pm \frac{B}{\sqrt{C^2 + A^2 - B^2}}) \\ \theta_3 = -(\pi - \frac{\arctan(A \sin \theta_2 + C \cos \theta_2)}{a_2 + C \sin \theta_2 - A \cos \theta_2}) \end{cases} \quad (16)$$

$$s_i = \sin \theta_i, c_i = \cos \theta_i, A = p_x c_1 + p_y s_1 - a_1, B = (a_3^2 - a_2^2 - A^2 - C^2) / (2a_2), C = -(-p_z + d_1)$$

Where  $s_i = \sin \theta_i, c_i = \cos \theta_i, A = p_x c_1 + p_y s_1 - a_1, B = (a_3^2 - a_2^2 - A^2 - C^2) / (2a_2), C = -(-p_z + d_1)$ .

At this point, the rotation angle of each mechanical connecting rod can be obtained only by knowing the situation that the end of the robot rod is about to move to the target position[14], that is, knowing the positioning coordinates and of the workpiece  $p_x$ 、 $p_y$  and  $p_z$ .

## V. Conclusion

As the motion planning of the general robot arm is only applicable to a specific environment, or it can be used in a complex environment but it needs a very long calculation time, and it is unable to immediately interact with the environment and lacks detection methods for the determination of the location of environmental objects, so an intelligent working state recognition system based on machine vision was designed for the robot arm. The workpiece placed in the workspace of the robot arm arbitrarily can better realize the recognition and positioning function of the workpiece, thus improving the situation that the position of the workpiece could not be grasped accurately due to the change of position.

## Reference

- [1] Ren Yabing. Planning Research of Robot Arm Motion Based on Machine Learning [D]. University of Electronic Science and Technology, 2019.
- [2] Xie Yinbo, Wei Tianqi, Tian Yuan, etc. Research on intelligent mechanical vehicle based on synchronous remote control of human arm [J]. Experimental technology and management (12): 85-89.
- [3] Zhou Xinyu. A Research on Robot Demonstration Learning Method Based on Track Simulation Learning and Task Planning [D]. Harbin University of Technology, 2019.
- [4] Yao Qicai, Wang Di, Liao Maosheng. Design of intelligent grasping system of robot arm based on machine vision [J]. Measurement and test technology, 2020, v.47;No. 341 (10): 37-42.
- [5] Liu Yiming. Mobile arm motion control and target capture based on machine vision [D]. Dalian Maritime University, 2020.
- [6] Pan Lei, Qian Wei, Zhang Zhiyan, etc. Kinematic analysis and MAT simulation of Four-Degree-of-Freedom robot arm [J]. Mechanical Science and Technology, 2013 (03): 115-119.
- [7] Meng Shaonan, Liang Yanbing, Shiheng. Motion planning of 6-DOF robot arm based on ROS platform [J]. Journal of Shanghai Jiaotong University (S1 issue): 94-97.
- [8] Wang Wenxi, Wang Xiaolan. Research on Motion Control Strategy of Flexible Manipulator [J]. Electrical Transmission Automation, 2006, 028 (005): 6-9.