A Design of Obstacle Avoidance and Grabbing Manipulator Based on Computer Vision for Warehouse Inspection Robot

Shaotong Zhang*

School of Mechanical Engineering and Automation. Northeastern University. Shenyang, Liaoning Province, 111000, China

* Corresponding Author Email: 20204865@stu.neu.edu.cn

Abstract. The automatic manipulator has several potential applications in industrial systems. It is critical to research machine route planning and obstacle avoidance in order to fulfill the rising technical requirements. This work provides a technique for realizing the aforementioned tasks based on an enhanced obstacle avoidance algorithm that can automatically skip barriers and complete target sorting using visual identification and improved obstacle avoidance algorithms. The whole system consists of three parts: image collection, pose estimation and target grasping. Firstly, computer vision technology is used to recognize and locate target objects. Secondly, the robot's kinematics model is created, and combined with the visual perception information, the target grabbing and sorting are completed. Finally, the environment is built with Gazebo, and the robot is simulated. Simulation results verify the feasibility and effectiveness of the method.

Keywords: Path planning; obstacle avoidance; mechanical arm; machine vision.

1. Introduction

Robotic arm is a kind of robot device which imitates the movement of human arm and completes all kinds of operations. The manipulator usually consists of a series of joints, drivers and actuators, which can be controlled by program or manual operation to achieve accurate positioning, grasping, handling and other operations. Robotic arms are widely used, such as material handling and processing on production lines in industrial automation, surgery and rehabilitation in health care, and operation and maintenance of drones and satellites in military and aerospace fields. With the continuous progress of robot technology, the application of robotic arm will be more and more extensive, and more industries and fields will use robotic arm to improve production efficiency and work efficiency. Industrial robot development offers a potential route for fostering company innovation on their own and achieving the innovation-driven development strategy of science and technology self-reliance [1].

Based on the ROS control system and GAZEBO. Currently, there are two types of path planning algorithms: conventional path planning algorithms, and intelligent optimization algorithms. Grid technique, visual graph method, artificial potential field method, and others are examples of conventional path planning methods. Intelligent algorithms mainly include ant colony algorithm [2], particle swarm optimization (PSO) [3], evolutionary optimization algorithms [4] and so on. The emergence of machine vision technology also meets the latest requirements of the development of robot arm automation, so it has made rapid progress in the field of robot application [5-7]. In the domestic aspect, a number of universities have studied the motion of the manipulator controlled by vision [8-10]. In order to capture occluded objects, Lei Zhiming of Guangxi University developed a technique for obtaining projection feature points of 3D objects using deep convolution neural networks based on multi-scale feature fusion [11]. In order to replace the subsequent ICP processing phase, Zheng Xinyu of Shenyang University of Technology employs the cyclic iteration optimization approach in conjunction with the target recognition and posture estimation algorithm based on deep learning [12]. In foreign countries, Vietnamese scholars have proposed a simple and efficient image processing algorithm, which helps to shorten the time of the sorting process [13]. Martin, Harold and other scholars have proposed to use vision to control prosthetic arms [14].

This paper simulates a possible working environment of the manipulator to complete the process of positioning, grasping and placing the grasping object to the specified position through machine vision. Firstly, computer vision technology is used to recognize and locate target objects. Secondly, the robot's kinematics model is created, and combined with the visual perception information, the target grabbing and sorting are completed. Finally, the environment is built with Gazebo, and the robot is simulated.

2. Methodolgy

2.1. The basic structure of the system

The structure of the manipulator grab system based on ROS designed in this paper is shown in Fig. 1. The whole system consists of three parts: image collection, pose estimation and target grasping, which realizes the grasping of grasping objects in the presence of occlusion.

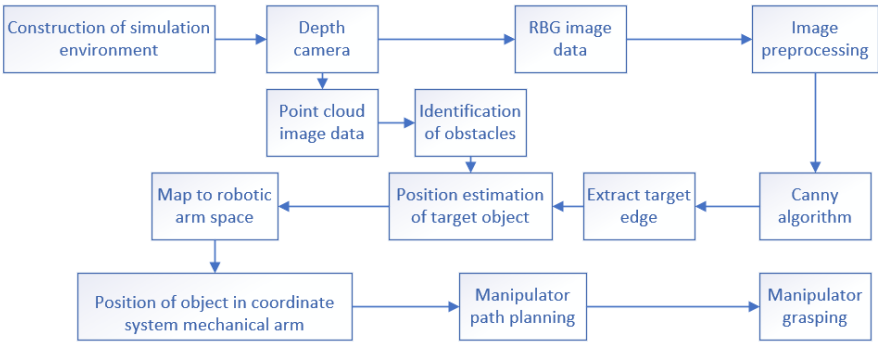


Fig.1 the basic structure of the system

The manipulator used in this system is a six-degree-of-freedom UR5 manipulator. The gripper adapted to the manipulator adopts two-finger mechanical gripper of RobotiQ. In order to collect relevant point cloud data and depth data, the Kinect V2 depth camera is used in the experiment. Firstly, the edge of the target object is extracted through image processing technology, and the position of the object is estimated using the point cloud information gathered by the depth sensor. Subsequently, the relative position of the target object is calculated depending on the robot's location and posture data. Finally, combining the above information, the robot arm's trajectory is planned, and the target item gripping job is done.

2.2. Target location estimation

2.2.1 Image filtering

The camera records the target picture, which includes not only the target data but also the background and noise data. This noisy information will degrade the image quality, resulting in inadequate or incorrect processing results. As a result, filtering off the image's noise information is required to lessen the influence on contour detection. There are two types of image filtering: linear filtering and nonlinear filtering. Linear filtering encompasses square frame filtering, mean filtering, Gaussian filtering, and other techniques. Nonlinear filtering techniques include median filtering, bilateral filtering, and others. Gaussian filtering is used to match the canny method used in this experiment.

Gaussian filtering algorithm is also a class of linear spatial filtering methods, it belongs to low-pass filter, which can suppress the Gaussian noise in the image. Gaussian filtering uses a weighted average of the whole image, and each pixel's value is derived by averaging its own value and the values of surrounding pixels. The weighted average of every pixel neighborhood, which is what the Gaussian function produces, is closer to the value of the central pixel. The algorithm emphasizes the weight of the center point, so it has a better smoothing effect than the mean filter.

2.2.2 Binarization of image

After binarization, the grayscale image is converted to a black and white image. The image's pixel values are merely 0 and 1. Binary image processing can reduce the amount of picture data and the complexity of future processing greatly. When the background is relatively simple and there are only a few objects in the image, the object can be easily separated from the background, which is convenient for object edge feature extraction and object recognition. Image binarization processing is to distinguish the target object information from the background information by setting an appropriate threshold and comparing the small relationship between the pixel value and the threshold of each point in the grayscale image.

The setting of threshold has a great influence on the effect of image binarization. The threshold selected by artificial measurement often does not adapt to the changes of the current environment and affects the execution efficiency of the visual system. In order to make the image processing process more adaptive, the adaptive threshold algorithm should be used to experiment the binarization of the image. In this paper, OTSU algorithm is used to obtain the threshold automatically. For threshold selection in picture segmentation, the OTSU algorithm is regarded as the best algorithm. it is straightforward to calculate and is not impacted by image brightness and contrast. It separates the picture into two halves based on the grayscale qualities of the image: background and foreground. Because variance is a measure of grayscale uniformity, the difference between an image's foreground and background is proportional to how dissimilar the two are from one another. As a result, segmentation that optimizes the variation across classes has the lowest likelihood of misclassification.

2.2.3 Canny algorithm

The edge of the image is the part of the image where the local brightness varies significantly, and this area's gray profile can be thought of as a step, going from one gray value in a very little buffer area to another gray value with a significant gray difference. The majority of the image's information is concentrated at its edges. The recognition and interpretation of the complete image scene relies on the detection and extraction of the picture edge, which is also a property on which image segmentation depends. Edge detection is primarily concerned with measuring, detecting, and locating grayscale changes in images. JohnF. Canny created the Canny edge detection technique in 1986. Through multi-stage algorithms, it can identify a wide variety of edge information in a picture. The traditional Canny detection algorithm is divided into the following four steps.

The first step is to use Gaussian filter to process the original image. Because the edge of the image is easily disturbed by noise, in order to avoid detecting the wrong edge information, it is usually necessary to filter the image to remove the noise in order to get a more accurate edge. In the actual processing process, Gaussian filtering is usually used to remove the noise in the image.

The gradient is calculated in the second step. The gradient's direction is perpendicular to the edge. The operator in charge of edge detection has returned. The gradient always runs perpendicular to the edge, and the closest values are generally horizontal, vertical, diagonal, and so on.

To further refine the edge, the next step is to suppress the amplitude and direction of the gradient by non-Maxima. After determining the gradient's amplitude and direction, traverse the pixels in the picture and delete all non-edge points. In the concrete implementation, it is important to examine each pixel individually to see if it has the highest value and the same gradient as the pixels around it, if it is the maximum value, the current point is determined as the edge point; if it is not the maximum value, then determine that the current point is not part of the edge point.

The fourth stage is using multiple thresholds to improve the image edge. When non-maximum suppression is finished, the image's strong edge is already present in the edge image that is now being acquired. The edge image might potentially contain some virtual edges, though. These fictitious edges could be produced by real images or by noise. It is necessary to eliminate the noise-induced virtual edges. As a result, two thresholds—high threshold and low threshold—are established. The attribute of the edge is determined by comparing the gradient amplitude of the current edge pixel to the two thresholds. If the gradient amplitude is greater than or equal to the high threshold, the current edge

pixel is labeled as a strong edge and retained. If the current edge pixel's gradient amplitude is between an excessive and a minimal threshold, the pixel is retained as a virtual edge. If the current edge pixel's gradient amplitude is less than or equal to the low threshold, it is suppressed and suppressed. The acquired virtual edges must be processed further. The edge is considered as an edge if it is linked to a strong edge. If there is no link to the powerful edge, the edge becomes weak and muted.

2.3. Manipulator motion control

2.3.1 Forward kinematics analysis

The forward kinematics of the manipulator is used to solve the position and pose result of the manipulator's end tool actuator relative to the manipulator's base using the connecting rod parameters and joint rotation angle parameters without taking force into account. The manipulator connecting rod coordinate system is established by the standard Dmurh method, as shown in Fig. 2. The parameters such as joint coordinate system, rotation angle around z-axis θ_i, length of extended z-axis d_i, rotation angle around x-axis α_i and length of extended x-axis a_i are marked in the figure.

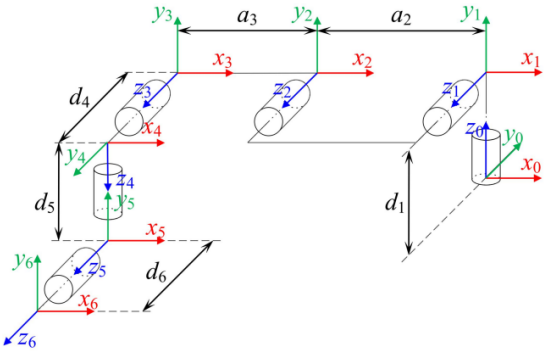


Fig.2 Simplified model of UR5 manipulator and linkage coordinate system [11]

The standard D-H parameter table is made from the UR5 manipulator connecting rod coordinate system, as shown in Table 1.

Table 1. Standard D-H parameter table of UR5 robot

i	d_i	a_i	α_i	ϑ_i
1	d_1=89.459	0	$\pi/2$	ϑ_1
2	0	a_2=-425	0	ϑ_2
3	0	a_3=-392.25	0	ϑ_3
4	d_4 =109.15	0	$\pi/2$	ϑ_4
5	d_5 =94.65	0	$-\pi/2$	ϑ_5
6	d_6 = 82.3	0	0	ϑ_6

After obtaining the D-H parameter table, the transformation matrix between two adjacent coordinate systems can be obtained, as shown in follow equation. ${}^{a}_{b}T$ is the pose transformation relationship between joint a and joint b; c and s are abbreviations for cos and sin, respectively.

$$ {}^{0}_{1}T = \begin{bmatrix} c1 & 0 & s1 & 0 \\ s1 & 0 & -c1 & 0 \\ 0 & 1 & 0 & d1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{1} $$

$$ {}^{0}_{1}T = \begin{bmatrix} c1 & 0 & s1 & 0 \\ s1 & 0 & -c1 & 0 \\ 0 & 1 & 0 & d1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{2} $$

$$ {}^{1}_{2}T = \begin{bmatrix} c2 & -s2 & 0 & a2c2 \\ s2 & c2 & 0 & a2s2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3} $$

$$ {}^{2}_{3}T = \begin{bmatrix} c3 & -s3 & 0 & a3c2 \\ s3 & c3 & 0 & a3s2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{4} $$

$$ {}^{3}_{4}T = \begin{bmatrix} c4 & 0 & s4 & 0 \\ s4 & 0 & -c4 & 0 \\ 0 & 1 & 0 & d4 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{5} $$

$$ {}^{4}_{5}T = \begin{bmatrix} c5 & 0 & -s5 & 0 \\ s5 & 0 & c5 & 0 \\ 0 & -1 & 0 & d5 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{6} $$

$$ {}^{5}_{6}T = \begin{bmatrix} c6 & -s6 & 0 & 0 \\ s6 & c6 & 0 & 0 \\ 0 & 0 & 1 & d6 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{7} $$

So, the transformation matrix from coordinate system 0 to coordinate system 6 is as follows.

$$ {}^{0}_{6}T = {}^{0}_{1}T \cdot {}^{1}_{2}T \cdot {}^{2}_{3}T \cdot {}^{3}_{4}T \cdot {}^{4}_{5}T \cdot {}^{5}_{6}T = \begin{bmatrix} n_x & o_x & a_x & p_x \\ n_y & o_y & a_y & p_y \\ n_z & o_z & a_z & p_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{8} $$

The parameters in the Eq. (8) are described as follows.

$$ \begin{cases} n_x = c_1c_{234}c_5c_6 + s_1s_5c_6 - c_1s_{234}s_6 \\ n_y = s_1c_{234}c_5c_6 - c_1s_5c_6 - s_1s_{234}s_6 \\ n_z = s_{234}c_5c_6 + c_{234}s_6 \\ o_x = -c_1c_{234}c_5s_6 - s_1s_5s_6 - c_1s_{234}c_6 \\ o_y = c_1s_5s_6 - s_1c_{234}c_5s_6 - s_1s_{234}c_6 \\ a_z = c_{234}c_6 - s_{234}c_5s_6 \\ a_x = s_1c_5 - c_1c_{234}s_5 \\ a_y = -s_1c_{234}s_5 - c_1c_5 \\ o_z = -s_{234}s_5 \\ p_x = s_1d_{234} + c_1c_2a_2 + c_1c_{23}a_3 + c_1s_{234}d_5 + s_1c_5d_6 - c_1s_{234}s_5d_6 \\ p_y = -c_1d_{234} + s_1c_2a_2 + s_1c_{23}a_3 + s_1s_{234}d_5 + c_1c_5d_6 - s_1c_{234}s_5d_6 \\ p_z = d_1 + s_2a_2 + s_{23}a_3 - c_{234}d_5 - s_{234}s_5d_6 \end{cases} \tag{9} $$

2.3.2 Inverse kinematics analysis

Inverse kinematics is the rotation angle of each joint corresponding to each joint of the manipulator according to the target position and posture of the end tool effector of the manipulator. According to the joint order, based on the Pieper criterion, the analytical solution of the manipulator is as follows.

$$\begin{cases} \theta_1 = atan2(p_y, p_x) - atan2\left(-d_x, \pm\sqrt{p_x^2 + p_y^2 - d_4^2}\right) \\ \theta_2 = atan2(B, A) - atan2\left(C, \pm\sqrt{A^2 + B^2 - C^2}\right) \\ \theta_3 = \theta_{23} - \theta_2 \\ \theta_4 = \theta_{234} - \theta_{23} \\ \theta_5 = atan2(s_5, s_1 a_x - c_1 a_y) \\ \theta_6 = atan2\left(\frac{-s_1 o_x + c_1 o_y}{s_5}, \frac{s_1 n_x - c_1 n_y}{s_5}\right) \\ B_1 = c_1 p_x + s_1 p_y - d_5 s_{234} \\ B_2 = p_z + d_1 - d_5 c_{234} \\ A = -2B_2 a_2 \\ B = 2B_1 a_2 \\ C = B_1^2 + B_2^2 + a_2^2 - a_3^2 \\ \theta_{23} = atan2\left(\frac{B_2 - a_2 s_2}{a_3}, \frac{B_1 - a_2 c_2}{a_3}\right) \\ \theta_{234} = atan2\left(-\frac{a_z}{s_5}, -\frac{c_1 a_x + s_1 s_y}{s_5}\right) \end{cases} \quad (10)$$

3. Experimental results and analysis

3.1. The introduction of the simulation platfor

In the previous chapter, the author discusses the related technologies used in this robot system. In order to verify the technical feasibility, a warehouse simulation environment is set up in GAZEBO to carry out simulation experiments, as shown in Fig. 3.



Fig.3 the simulation environment

The specific simulation experiment flow is shown in the figures below.

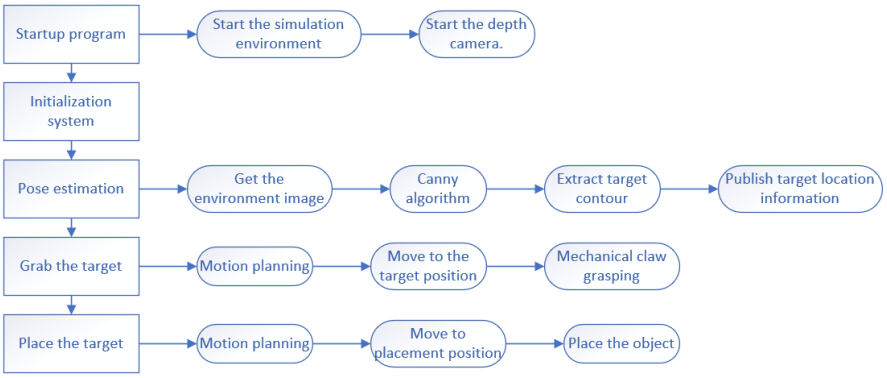
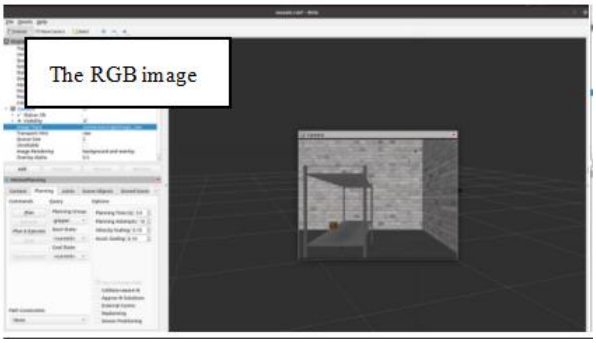


Fig.4 Simulation experiment flow

The experiment takes the cube to be grasped as the experimental object. At the beginning of the experiment, start the simulation environment and related running programs. Then the camera publishes RGB topic information and depth information, and adopts relevant algorithms to estimate the position and pose of the grasped object. After calculating the position in the coordinate system of camera, the object will be converted to the mechanical arm base coordinate system. Then the manipulator uses inverse kinematics to plan the grasping path and complete the operation of grasping and placement.

3.2. Simulation results and analysis

At the beginning of the simulation experiment, start the simulation program. The first step is image collection. The camera captures the RGB and depth images of the object, and generates the point cloud data of the environment. The RGB and depth images of the are visually displayed, and the point cloud data are generated in Rviz, as shown in Fig. 5 and Fig. 6. The system uses canny algorithm to extract the contour and centroid of the object to be grasped.



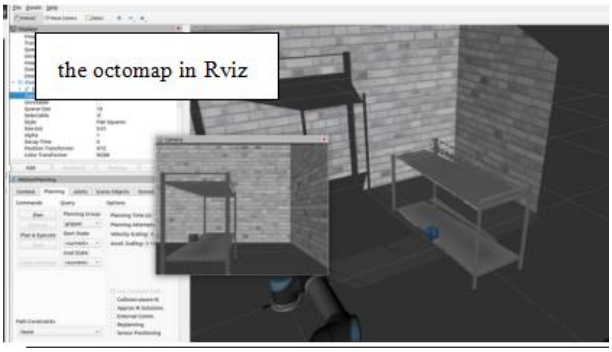
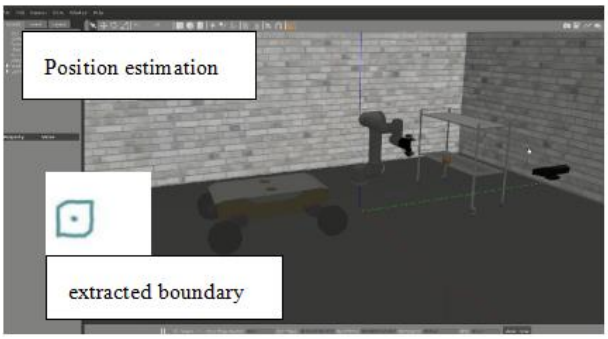


Fig.5 the RGB image **Fig.6** the octomap in Rviz

The second step is position estimation. The system estimates the position of the target object through a series of algorithms. After the centroid position is calculated, it is converted into the manipulator coordinate system and the information is returned to the manipulator. The experimental effect is shown in the following figure.



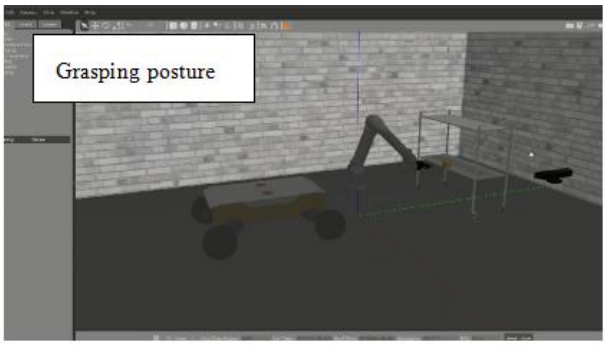


Fig.7 get target location information **Fig.8** adjust the grasping posture

The third step is to grab the object and the fourth step is to place the object in a specified position. The manipulator moves to the target position, and grasps the target object. The robotic arm can analyze the point cloud data, plan the grasping route, and complete the grasping and placement actions. The experimental effect is shown in the following figure.



Fig.9 grab and place the target

According to the aforementioned simulated testing findings, the design of the warehouse inspection robot to circumvent obstacles grasping manipulator according to automated vision created in this study has some practicality.

4. Conclusions

As the warehouse inspection process requires a lot of manpower, this paper studies the warehouse inspection manipulator based on vision. The whole system consists of three parts: image acquisition, pose estimation and target capture. First of all, the system uses the depth camera to identify and locate the grasped object. Then the manipulator grasps and places the object according to the acquired data. Finally, the simulation experiment is carried out based on the ROS control system and Gazebo simulation environment. From the experimental results, the system can better complete the task of target detection and acquisition, which verifies the feasibility of the method.

References

[1] Wang Lei, Xiao Qian, Deng Fangfang. Research on the Impact of Artificial Intelligence on Innovation in China's Manufacturing Industry: Evidence from the Application of Industrial Robots [J/OL]. Journal of Finance and Economics:1-14.

[2] Li, Xue, Lei Wang, and L. Wang. "Application of improved ant colony optimization in mobile robot trajectory planning." Mathematical Biosciences and Engineering 17.6 (2020): 6756-6774.

[3] Zhou Wei Research on trajectory planning of robotic arms based on particle swarm optimization algorithm [D]. Nanjing University of Posts and Telecommunications,2022.

[4] Malyshev, Dmitry, et al. "Synthesis of Trajectory Planning Algorithms Using Evolutionary Optimization Algorithms." Advances in Optimization and Applications: 13th International Conference, OPTIMA 2022, Petrovac, Montenegro, September 26–30, 2022, Revised Selected Papers. Cham: Springer Nature Switzerland, 2023.

[5] Ficuciello, F., et al. "Vision-based grasp learning of an anthropomorphic hand-arm system in a synergy-based control framework." Science robotics 4.26 (2019): eaao4900.

[6] Åstrand, Björn, and Albert-Jan Baerveldt. "An agricultural mobile robot with vision-based perception for mechanical weed control." Autonomous robots 13 (2002): 21-35.

[7] Lee, Hai-Wu. "The study of mechanical arm and intelligent robot." IEEE Access 8 (2020): 119624-119634.

[8] Zheng, Zhaohui, et al. "Industrial part localization and grasping using a robotic arm guided by 2D monocular vision." Industrial Robot: An International Journal (2018)

[9] Chen, Joy Iong-Zong, and Jen-Ting Chang. "Applying a 6-axis mechanical arm combine with computer vision to the research of object recognition in plane inspection." Journal of Artificial Intelligence 2.02 (2020): 77-99.

[10] Chiang, Mao-Hsiung, and Hao-Ting Lin. "Development of a 3D parallel mechanism robot arm with three vertical-axial pneumatic actuators combined with a stereo vision system." Sensors 11.12 (2011): 11476-11494

[11] Lei Zhiming. Research on pose estimation based on RGB image robotic arm grasping [D] Guangxi University, 2021 one thousand one hundred and seventy-nine

[12] Zheng Xinyu Research on Visual Based Manipulator Grasping [D]. Shenyang University of Technology,2022.DOI:10.27323/d.cnki.gsgyc.2022.000047.

[13] Cong, Vo Duy, et al. "Design and development of robot arm system for classification and sorting using machine vision." FME Transactions 50.1 (2022): 181-181.

[14] Martin, Harold, et al. "A novel approach of prosthetic arm control using computer vision, biosignals, and motion capture." 2014 IEEE Symposium on Computational Intelligence in Robotic Rehabilitation and Assistive Technologies (CIR2AT). IEEE, 2014.SHI Biao, LI Yu Xia, YU Xhua, YAN Wang. Short-term load forecasting based on modified particle swarm optimizer and fuzzy neural network model. Systems Engineering-Theory and Practice, 2010, 30(1): 158-160.