POSITIONING AND GRABBING TECHNOLOGY OF INDUSTRIAL ROBOT BASED ON VISION

Yonghao Zhao*, Hailin Li
Guangzhou City Construction College, Guangdong Guangzhou, 510900, China.
Corresponding author: Yonghao Zhao
Email: zhaoyonghaobest1@yeah.net

***ABSTRACT**: In order to explore the role of machine vision technology in the positioning and grabbing of industrial robots, the application of machine vision image extraction technology in industrial robot positioning and grabbing was studied through experiments and analysis methods. The Gaussian filtering method was mainly used to smooth the image. On this basis, the edge extraction of the target object image was performed by Canny operator edge detection. The robot arm grab point was determined by binocular three-dimensional positioning coordinates, and then the PID (Proportion Integral Differential) position algorithm was used to analyze the dynamic positioning control of the robot. The results showed that the vision-based industrial robot positioning and grabbing technology can improve the recognition efficiency and grasping efficiency of industrial robots, which is superior to the traditional industrial robot positioning and grabbing function. In summary, the introduction of machine vision functions into industrial robot systems and for industrial production can improve the automation, intelligence and efficiency of industrial production, and has good industrial practical use value.*
***KEYWORDS**: Industrial Robot; Positioning and Grabbing; Vision; Edge Detection; PID*

1 INTRODUCTION

Industrial robots are machine devices that perform work tasks according to pre-programmed program codes, and have been widely used in the industrial field [1]. For example, it is applied in product line assembly, parts sorting, welding, loading and unloading, etc. Industrial robots can replace humans in some dangerous environments, repetitive high-tech scenes, reducing labor costs while improving the quality and efficiency of production. However, with the increasing complexity of the production environment, the requirements for the use of robots are also increasing. When positioning and grasping, in order to achieve the production tasks that are difficult and difficult to successfully complete the process, it is necessary to improve the industrial robot process system and design to add more automated processing functions [2]. The positioning and grabbing process requires the use of visually relevant techniques such as detection recognition, positioning, and dynamic tracking.

Vision-based industrial robot positioning and grabbing technology is a new functional technology application that combines machine vision technology with robot motion technology [3]. If make robots to replace human work well, the first thing to do is to let them be able to "see". Machine vision is the eye of an industrial robot. It processes and analyzes the image information acquired by the camera [4] to form the position of target object, and then leads the industrial robot to operate the corresponding tracking and grabbing actions. For robots, machine vision gives its sophisticated computing system and processing system [5], simulating the way of bio-visual imaging and information processing, allowing robot arm to perform more anthropomorphic and flexible operations [6] while simultaneously identifying, comparing, and processing scenes, to generate execution instructions, and then complete the action in one go. Machine vision is decisive for the flexibility and operability of robots.

In summary, this paper took the promotion of the grabbing efficiency of industrial robots as the direction of efforts, and increased the application function of industrial robots by using machine vision technology, which makes industrial robots have better flexibility and can further improve the working efficiency, application range and intelligent automation level of industrial robots.

2 RESEARCH STATUS

2.1 Extraction of vision system images

The machine vision system refers to extracting an image through machine vision, and then digitizing the image in a system processing unit, and determining the size, shape, and the like according to information such as pixel distribution and brightness, and then controlling the device operation at the site based on the result of the determination. In order to realize the machine eye function similar to the human eye function, it is necessary to optimize the image clarity and integrity. After all, the main information data is derived from the original image [7]. The filtering processing technology is to weaken the noise in the image obtained by camera, so that the image retains more key information, and the image quality is improved, and the validity and reliability of the image as a basis for subsequent technologies are guaranteed [8]. The Gaussian filtering method was used to eliminate interference in the image in this study. In addition, gray processing needs to be carried out for the image, that is, image enhancement, the color photo was converted into black and white grayscale display. This method can enhance the contrast ratio of the selected object, so that the target object has a stronger layering and highlighting. The edge-based image detection method mainly uses the difference value between the pixel value of the edge of the target object and the pixel value of the background image to extract.

Machine vision is widely used in industrial manufacturing. The characteristics of machine vision mainly include: firstly, high precision, which means that it can adapt to complicated processes and mass of fine processing steps, and there is no wear and tear on fragile components; secondly, flexibility, which means that it can capture a variety of different images and adapt to different functions. Thirdly, the cost is low, and as the price of computer processors decreases, the cost of machine vision systems also decreases. The fourth is high continuity. It can be used in place of people for continuous operation.

The application process of machine vision includes image acquisition, image processing, feature extraction, decision and control. Among them, image acquisition is to use the camera to acquire the image of the measured object [9], and convert the image and the intrinsic features into data that can be processed by the computer. Image processing is the processing of image enhancement, smoothing, edge sharpening, segmentation and other content. Feature extraction is a key feature that the processor recognizes and quantifies image, and then passes the data to the control program. Decisions and controls are based on data received by the processor. The application process is shown in Figure 1.

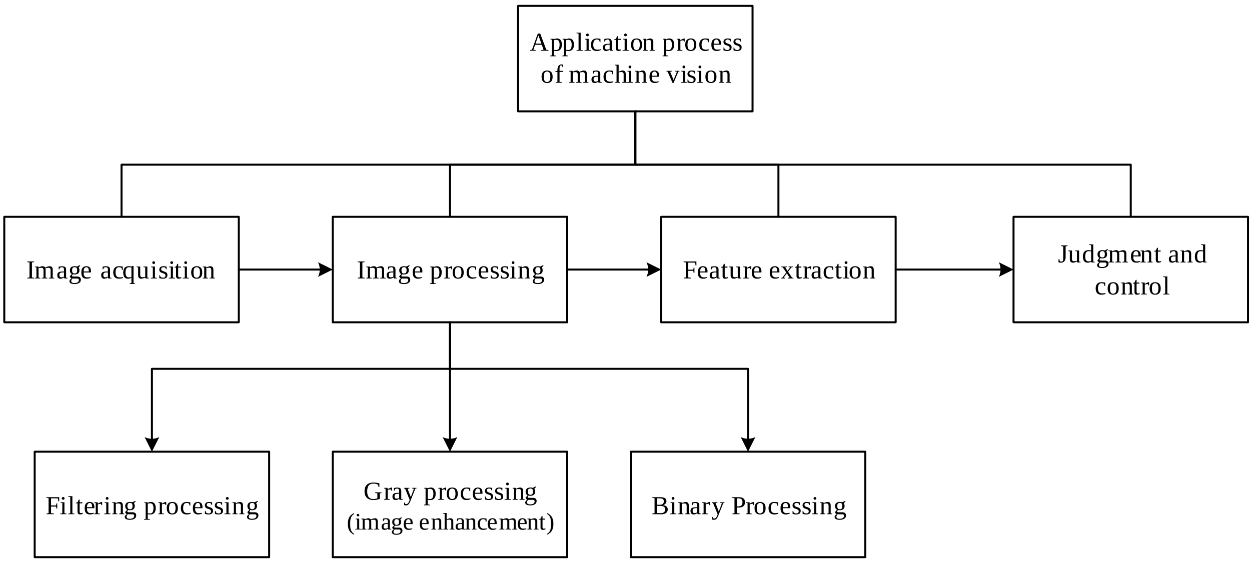

Figure. 1 Machine vision application process

Image preprocessing technology can reduce the noise in the image and maintain the image with clear quality [10]. Gaussian noise is one of the common noises. There are many ways to remove noise. The elimination effect of various methods is not the same. Gaussian filtering was used to denoise in this study. Gaussian filtering is a linear smoothing filter that is suitable for eliminating Gaussian noise and is widely used in the noise reduction process of image processing. In fact, Gaussian filtering is a process of weighting and averaging the entire image through a Gaussian function. The value of each pixel is obtained by weighted averaging of itself and other pixel values in the neighborhood [11]. Gaussian filtering is performed by using a Gaussian filter, which converts the noise signal in the image into a high-frequency signal, calculates the weight and performs denoising by linear convolution. The specific operation of Gaussian filtering is to scan each pixel in the image with a template and replace the value of the center pixel of the template with the weighted average gray value of the pixels in the neighborhood determined by the template. Gaussian filtering is a signal filter whose purpose is to smooth the signal. The filter is a mathematical model that is used to transform the image data into energy. The noise belongs to the high frequency part, and the Gaussian filter smoothes the noise to reduce the effect.

$$P(z) = \frac{1}{\sqrt{2\pi\sigma}} e^{-(z-\mu)^2/2\sigma^2} \quad (1)$$

Equation (1) is an expression of Gaussian noise, in the equation the parameters include: μ represents the expected value, σ represents the standard deviation, that is, the Gaussian radius, σ^2 represents the variance value. P (z) represents the probability.

$$g(x, y) = \frac{1}{2\pi} \exp\left(\frac{x^2 + y^2}{2\sigma^2}\right) \quad (2)$$

Equation (2) is an expression of image Gaussian filtering. Combined with the actual processing of the image, the Gaussian filtering method can remove the image interference noise on the basis of ensuring the original image feature as much as possible, and has a small image edge ambiguity.

The edge of the image is composed of some discontinuous gray values of the pixel. The edge of the image is the basic feature of the image. It is a collection of pixels that are not continuous with the grayscale of the neighborhood or have a sharp change in grayscale. It contains most of the key information of the image. The purpose of edge detection is to identify points in the digital image where the brightness changes significantly, and edge detection can effectively reduce the total amount of computation during image processing. Therefore, edge detection is a crucial part of image analysis processing and target recognition systems [12]. The edge detection technique generally detects the features by solving the first-order, second-order derivatives or gradients of the gray values of the pixels in the neighborhood, and analyzes whether the gray level of the image has a large variation to detect the edge of the image. The edge detection of the image is to weaken the background part and highlight the edge features of the target object [13]. There is a large gray value gradient amplitude at the edge of the image, so the maximum local amplitude of the gray gradient can be calculated to extract the edge points. Further, the contour of the target is obtained by tracking the straight edge of the pixel. The Canny edge detection operator is a multi-level edge detection algorithm. It is a multi-stage optimization operator with filtering, enhancement and detection. The implementation steps are as follows: first input the image, then perform Gaussian smoothing, then the gradient calculation, non-maximum suppression, finally double threshold detection, edge image. In order to satisfy the algorithm to identify the actual edges in the image as much as possible, and the identified edges should be as close as possible to the actual edges, Canny uses a variational method that closely approximates the first derivative of a Gaussian function for the possibility of image noise that may not exist to identify edges. Gaussian filtering smoothing is required in the Canny operator. The edge gray level change of the two-dimensional image can be represented by a gradient, and the first derivative also corresponds to the gradient operator. The two-dimensional image f(x, y), the gradient operator of the pixel at (x, y) can be represented by a vector, and the expression is expressed as equation (3):

$$\nabla f(x, y) = \left[f'_x \, f'_y\right] = \left[\frac{\partial f}{\partial f} \; \frac{\partial f}{\partial y}\right]^T \quad (3)$$

The gradient is used as the edge detection operator, the mode of the gradient is the edge intensity, and the direction angle of the gradient is the direction of the edge. Equation (4) is the direction angle of the gradient.

$$\theta(x, y) = \arctan\left(\frac{f'_x}{f'_y}\right) \quad (4)$$

After Gaussian filtering was used to smooth the image, noise interference was removed, image enhancement was realized, image and background values were separated, and then Canny edge detection operator was used to extract image edges and highlight image edge features. The Canny operator can be approximated by the first derivative of the Gaussian function and is the optimal solution of the product of the localization and the signal-to-noise ratio. This operator subtly converts the edge detection problem into the problem of finding the extreme point of the unit function. The five steps of the processing flow of the Canny edge detection algorithm are as follows;

First: Use a Gaussian filter to smooth the image and filter out noise.

Second: Calculate the gradient strength and direction of each pixel in the image.

Third: Apply non-maximum suppression to eliminate spurious responses from edge detection.

Fourth: Apply dual threshold detection to determine real and potential edges.

Fifth: Edge detection is finally completed by suppressing the weak edges of isolation.

The Canny operator edge detection method has obvious effect on image edge extraction, which makes the target object different from other parts of the color, and the extraction of weak edge information is relatively complete and obvious. The Canny operator is less affected by noise, and the detected edges are continuous and the edges are clear.

2.2 Target tracking position control

The three-dimensional positioning directly determines the accuracy in the grabbing. The binocular stereo vision can obtain the three-dimensional geometric coordinate information of the target object in the stereoscopic scene, which is equivalent to positioning the target object and providing the machine coordinate basis for the subsequent grabbing. Compared with two-dimensional image information, the three-dimensional positioning coordinate technology is suitable for large and complex production requirements in the industry, and realizes object grabbing of various types of shapes and sizes. The left and right cameras respectively collect image of the target object at different angles, and extract the feature according to the acquired image information and image library generation information. Then, the matching and recognition of the two are performed, and then the contour extraction is performed, and then the parallax calculation of the left and right camera images is performed, and the left and right double targets are determined, and finally the target positioning is achieved. And the left and right camera coordinates of the object were obtained. The binocular visual image plane forms a triangle with the target object image plane. On the basis of the known left and right camera image position information, the three-dimensional coordinate size of the target object can be obtained, as shown in Table 1 below:

Table 1 3D coordinates of the target object

No.	Left camera	Right camera	Three-dimensional coordinates
1	254,104	155,107	1.63,92.11,38.95
2	198,127	94,122	65.45,82.41,38.98
3	239,156	161,152	11.13,50.24,11.46
…	…	…	…

After obtaining the three-dimensional coordinates of the target object, the end effector of the robot performs a movement process from one point to another. The four-degree-of-freedom robot can control the position and speed requirements of the robot by controlling the joints of the robot arm. Joint motion control is accomplished in two steps, first with the completion of the joint motion servo command, transforming the motion of the position and posture of the end effector in the workspace into a time series represented by the joint variable or as a function of the joint variable as change of time. The servo control of the four joint motions is then completed, that is, the joint variable servo command generated in the first step is followed. The computer sends a control command, and the four joint axes can independently set parameters to exchange data with the motion controller and form a bridge of communication between the two. And the motion controller transmits the current motion position, velocity and acceleration information to the digital servo motor, then a control signal is generated in comparison with the actual position to initiate accurate position control. The structural parameters of the robot arm are shown in Table 2 below:

Table 2 Structural parameters of the robot arm

Parameters	Symbol
Motor frequency per revolution	P Value
Screw pitch	L(mm/r)
Ratio of transformation	n (When the motor is directly connected to the lead screw, n=l)
Target location	Pos(plusc)
Target speed	Vel(Pluse/ST)
Target acceleration speed	A.cc(Plusc/ST^2)
Control cycle ST	T(/μs)

The joint of the robot arm and the end position of the robot arm of industrial robot were taken as the starting point for positioning and grabbing. In the robot arm motion control, the inversion motion is performed according to the three-dimensional coordinate data, the target position of each joint is obtained, and the joint motion is controlled, so that the end effector of the robot obtains the trajectory of the motion route according to the joint, and the positioning and grabbing of the robot is ensured. The robot end point error and the rotation angle error of the arm joint are all declining, as shown in Figure 2.

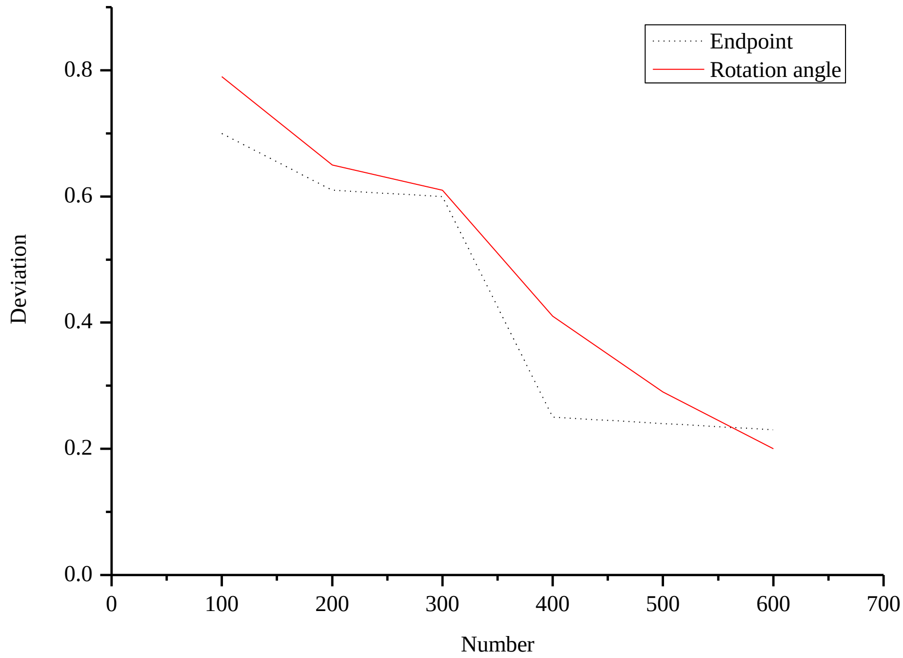

Figure. 2 End point and rotation angle error of robot arm

When the object is dynamically tracked, the PID algorithm can be used to realize dynamic tracking of the conveyor belt and dynamic monitoring of the target. The PID (Proportion Integral Differential) controller is the most widely used industrial controller. In the controller, the unit P is the basis of PID control; the unit I plays the role of eliminating the steady state error, but may increase the overshoot; unit D can increase the inertial response speed and weaken the overshoot. In the industrial automation production, according to the position information of the target, the robot arm trajectory is planned to realize dynamic and effective grabbing. After determining the grabbing position of the target, it is necessary to adjust the position between the robot arm and the conveyor belt target in real time. The conveyor belt is arranged in the XOY plane of the robot arm coordinate system, and the dynamic grabbing process only needs real-time tracking in the XY direction.

After the camera captures the image of the target object and then performs image pre-processing, since the target object is dynamically moving on the conveyor belt, the target has two positions, one is the current initial position Z1 of the target object, and the other is the intended grab position Z2 after the dynamism. The coordinates of the robot arm and the target object are calculated, and the X and Y components of the coordinate are used as the basis for controlling the grasping position Z2 of the industrial robot arm. By continuously adjusting the speed, position and other elements of the robot arm and the conveyor belt, the error between the robot arm and the conveyor belt is reduced, and according to the real-time situation, the difference between the coordinate Z2 of the intended grabbing position and the coordinate of the robot arm in the XOY plane is continuously reduced, so that The square is more matched, and when the position and speed of the robot arm and the target are agreed upon, the accurate grasping action is completed. The PID algorithm can be used to reduce this error.

$$m(t) = k_p\left[\varepsilon(t) + \frac{1}{T_i}\int_0^t \varepsilon(\tau)d\tau + T_d\frac{d\varepsilon(t)}{dt}\right] \quad (5)$$

In the PID time domain control expression equation (5) of PID control algorithm, m (t) refers to the control output, k_p refers to the proportional coefficient, T_i and T_d refer to the integral time constant and the differential time constant, respectively, and the addition in braces represents the proportional control term, the integral control term and the sum of the differential control terms.

Proportional P control is one of the simplest control methods, which makes the output of the controller proportional to the input error signal. When there is only proportional control, the system output has a steady-state error. Both the integral I control and the differential D control are the control of the deviation. In the I control, the output of the controller is proportional to the integral of the input error signal. In order to eliminate the steady-state error, the "integral term" must be introduced. The relative error of the integral term depends on the integral of time, and the integral term will increase with time. It pushes the controller's output up so that the steady-state error is further reduced until it equals zero, leaving the system with no steady-state error after entering steady state. The D control makes the output of the controller proportional to the differential of the input error signal (the rate of change of the error). It can predict the trend of error variation, adjust the mutation in advance, solve the lag problem, and improve the dynamic response speed. The P+D controller can make the control effect of the suppression error equal to zero or even a negative value in advance, thereby avoiding the serious overshoot of the controlled quantity and improving the dynamic characteristics of the system during the adjustment process.

During the dynamic grabbing process of the robot arm, the robot arm and the target object will change in the XOY plane of the conveyor belt. The dynamic response change is made according to equation (5). The system response error during the dynamic grabbing process is as shown in Figure 3, indicating that the previous error margin is concentrated. As time increases, the displacement error of the robot arm and the target object tends to zero, which realizes the unification of the two positions and facilitates subsequent grabbing. PID position control technology is used for dynamic tracking and dynamic monitoring of target objects, which can adjust parameters to achieve dynamic response of moving targets, and improve dynamic picking efficiency in industrial production and manufacturing.

3 RESULTS AND DISCUSSION

As can be observed from Figure 2, the end error of the arm and the rotation angle of the joint are both declining, and the end point error directly affects the result of the grab. The decrease of the error indicates that the visual-based industrial robotic robot arm has higher grabbing accuracy.

It can be observed from Figure 3 that in the dynamic grabbing process, at the early stage of positioning and tracking of industrial robots and target objects, the displacement error between the robot and the target object is large, and the error amplitude fluctuates greatly. As time increases, the actual distance between the two is reducing, and the displacement error gradually approaches 0, which means that the position of the robot and the target object is unified, and then the positioning and grabbing is completed, indicating that the industrial robot based on PID position control under machine vision can achieve accurate positioning and tracking and grabbing in dynamic.

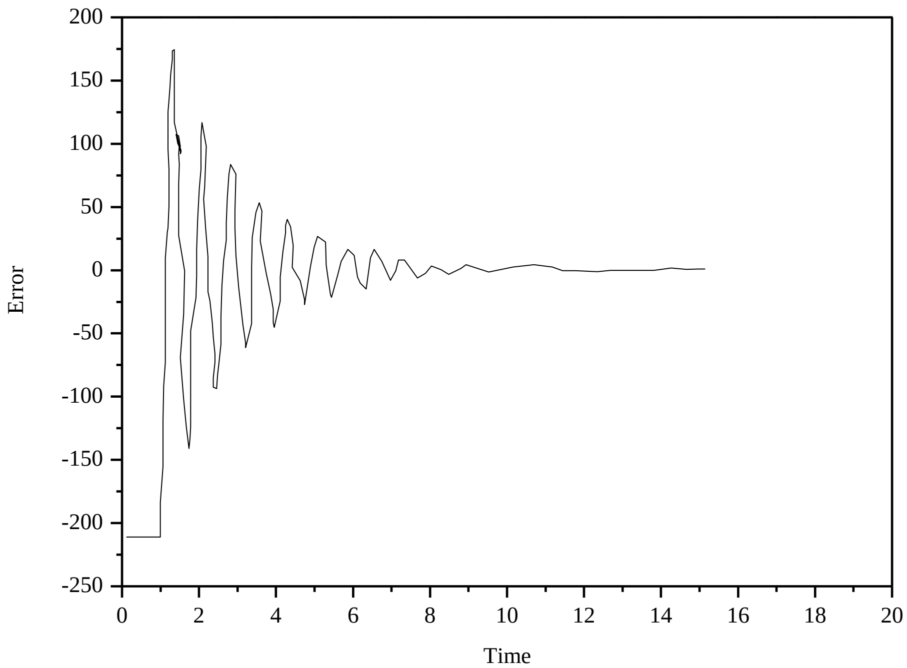

Figure. 3 Response error of dynamic grabbing process system

In Figure 4 is an error diagram after adjusting the differential time constant parameter in Figure 4 twice. It can be observed from the figure that after a small amplitude adjustment, the position error tends to approach 0 faster. The application of PID position algorithm control reduces the range of the fluctuation range of the error value and reduces the time taken for the fluctuation, which can tend to be stable more quickly and enhance the system response performance. It is shown that the vision-based industrial robot can reduce the coincidence time with the target object during dynamic positioning and grabbing, which is beneficial to the positioning and grabbing.

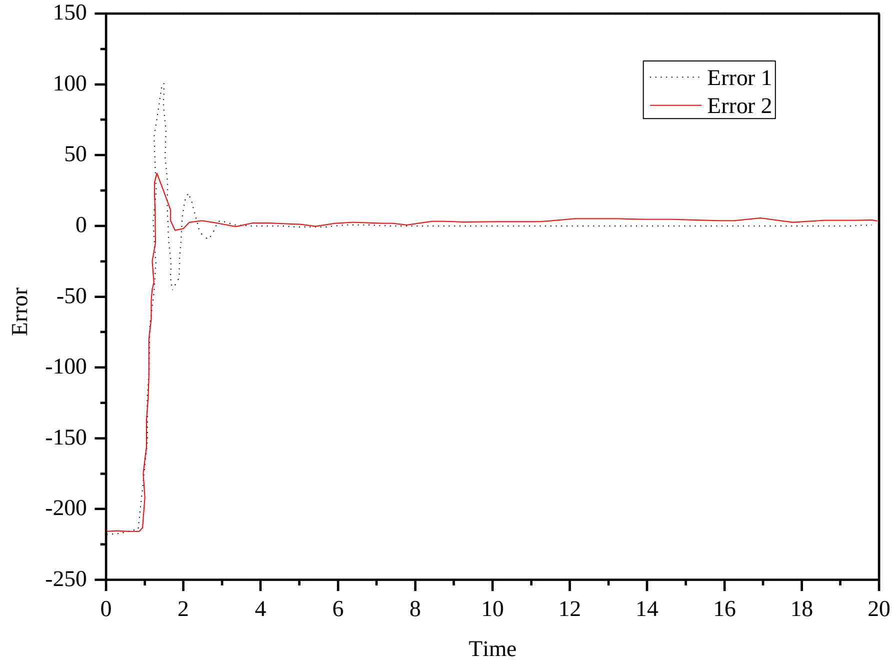

Figure. 4 Position error after adjusting parameters in the dynamic grabbing process

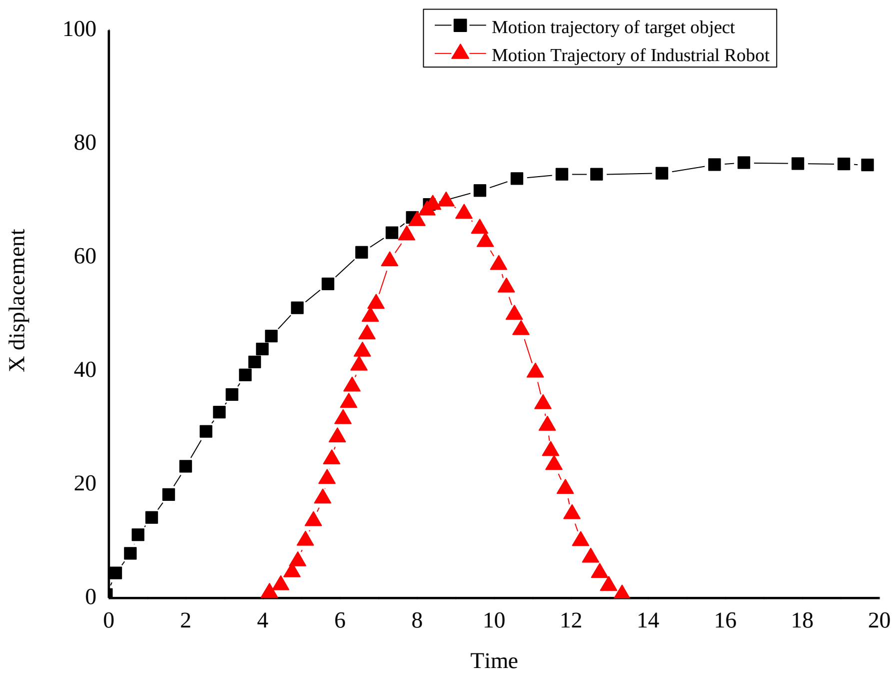

Figure. 5 PID tracking and grabbing of dynamic targets

As shown in Figure 5, in the PID dynamic target tracking, the black square curve and the red triangle curve table represent the motion trajectory curve of the target object and the motion trajectory curve of the robot. The two intersect at a time position of 0.9s, indicating that the robot arm of the industrial robot meets the target object at this time, which is also the time point for performing the positioning and grabbing. When the motion state of the target object changes with time, the distance between the robot arm and the target object in the position direction and speed is unified by continuously adjusting the distance between the robot arm and the target object, and then the grabbing action is performed. It shows that the dynamic position tracking of PID based on machine vision has a good effect.

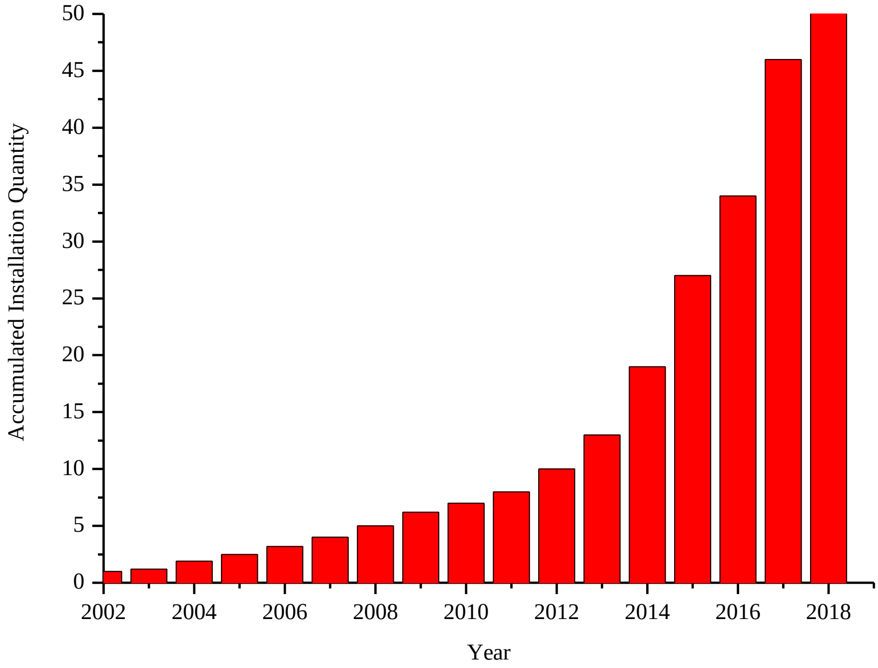

Figure. 6 Accumulated installation capacity of industrial robots in China in recent years (10,000 units)

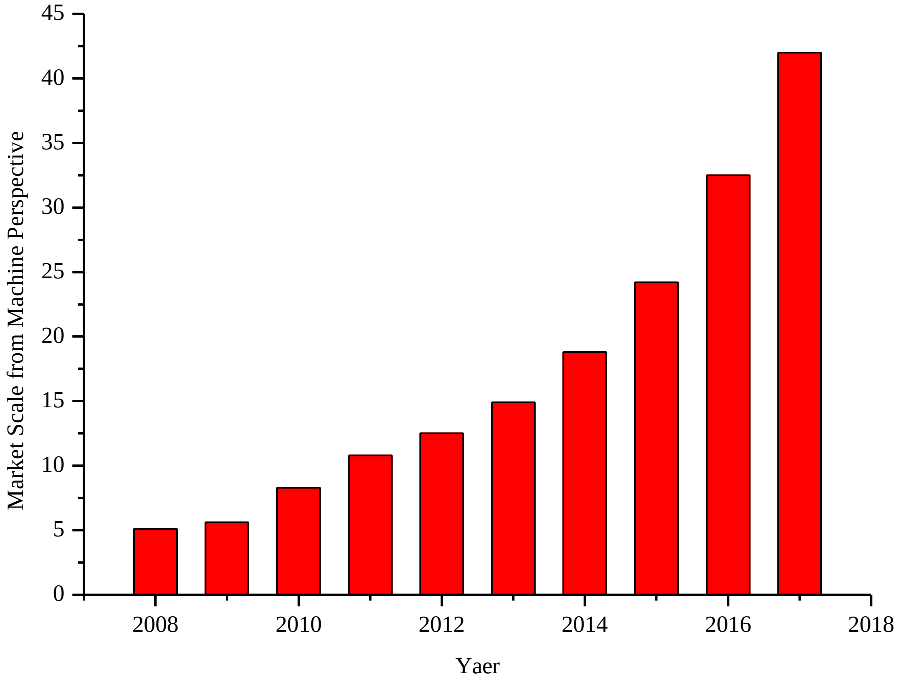

Figure. 7 China's machine vision industry market scale (100 million CNY)

Figure 6 shows the cumulative installation capacity of industrial robots in China in 2002-2018. Since 2003, the market share of industrial robots in China has risen linearly and is expected to maintain rapid growth, indicating that the use of industrial robots in China will increase.

As shown in Figure 7, the scale of China's machine vision market has shown an upward trend in recent years. In 2014, it was about 1.88 billion CNY. In 2015-2017, the industry entered a stage of rapid development, and in 2017 it reached 4.2 billion CNY. From 2016 to 2020, the growth rate of China's machine vision market is expected to remain above 20%, and the market space is growing.

The application of machine vision technology to industrial robots is necessary to optimize the function of robots. The market-oriented industrial robot positioning and grabbing technology has a large market demand and maintains a rapid growth trend.

It has good practical effects in industrial production, low cost of use and high production efficiency, which further meets the intelligent manufacturing needs of industrial automation production lines.

4 CONCLUSION

This study tries to solve the limitation of robot positioning and grabbing, and proposed a vision-based robot positioning and grabbing method. By combining the machine vision system image processing technology with the industrial robot control technology, and pre-processing the acquired image to optimize the quality, the three-dimensional positioning of the target object was then realized. The Canny detection was then used to obtain good edge features of the target object, and then the PID position control was used to acquire the motion tracking of the target object. The results showed that the vision-based robot system had fast and accurate image recognition, and the positioning was smart and precise, and it had high industrial application value. This technology is an important symbol of industrial modernization and automation development, and is also a technical support for industrial enterprises to improve production efficiency. This technology not only expands the range of applications of industrial robots, but also enhances the adaptability and working ability of robots in the industrial field. The scale of machine vision industry and the cumulative installation capacity of industrial robots in China are growing rapidly. In actual production, visual-based industrial robots are needed to complete the positioning and grabbing work. Faced with the increasingly complex production requirements in the intelligent manufacturing process of industrial robots, further expanding the integration of various functions can be considered. In the subsequent optimization design, it is feasible to start from the robot multi-arms and study the coordination of multiple arms to achieve better production efficiency of industrial robots.

5 REFERENCES

[1] Tae-jae Lee, Chul-hong Kim, Cho D I D. A Monocular Vision Sensor-Based Efficient SLAM Method for Indoor Service Robots. IEEE Transactions on Industrial Electronics, 2018, 66(1): 318-328.

[2] Wang X, L. He, T. Zhao. Mobile Robot for SLAM Research Based on Lidar and Binocular Vision Fusion. Chinese Journal of Sensors & Actuators, 2018, 31(3): 394-399.

[3]. Li L, Liu, Y, Jiang T, et al. Adaptive Trajectory Tracking of Nonholonomic Mobile Robots Using Vision-Based Position and Velocity Estimation. IEEE Trans Cybern, 2018, 48(2): 571-582.

[4] Fang, Ge, Wang, Xiaomei, Wang, Kui, et al. Vision-based Online Learning Kinematic Control for Soft Robots using Local Gaussian Process Regression. IEEE Robotics and Automation Letters, 2019, 4(2): 1194-1201.

[5] Colmenero-Martinez J T, Blanco-Roldán Gregorio L, Bayano-Tejero Sergio, et al. An automatic trunk-detection system for intensive olive harvesting with trunk shaker. Biosystems Engineering, 2018, 172: 92-101.

[6]. Eisenbruch M. The cloak of impunity in Cambodia I: cultural foundations. International Journal of Human Rights, 2018: 1-17.

[7] Zhang Z, Dequidt, Jeremie, Duriez, Christian. Vision-Based Sensing of External Forces Acting on Soft Robots Using Finite Element Method. IEEE Robotics & Automation Letters, 2018, 3(3): 1529-1536.

[8]. Calli B, Caarls, Wouter, Wisse, Martijn, et al. Active Vision via Extremum Seeking for Robots in Unstructured Environments: Applications in Object Recognition and Manipulation. IEEE Transactions on Automation Science & Engineering, 2018, (99): 1-13.

[9] Kudryavtsev A V, Chikhaoui, Mohamed Taha, Liadov, Aleksandr, et al. Eye-in-Hand Visual Servoing of Concentric Tube Robots. IEEE Robotics & Automation Letters, 2018, 3(3): 2315-2321.

[10] Omisore O M, Han, Shi Peng, Ren, Ling Xue, et al. Towards Characterization and Adaptive Compensation of Backlash in a Novel Robotic Catheter System for Cardiovascular Interventions. IEEE Transactions on Biomedical Circuits & Systems, 2018, (99): 1-15.

[11]. Mingmei Cheng, Yigong Zhang, Yingna Su, et al. Curb Detection for Road and Sidewalk Detection. IEEE Transactions on Vehicular Technology, 2018, 67(11): 10330-10342.

[12] Fan Z, Tang H, Liu Y H. Odometry-Vision-Based Ground Vehicle Motion Estimation With SE(2)-Constrained SE(3) Poses. IEEE Transactions on Cybernetics, 2018, (99): 1-12.

[13] Ikeuchi, Katsushi, Ma, Zhaoyuan, Yan, Zengqiang, et al. Describing Upper-Body Motions Based on Labanotation for Learning-from-Observation Robots. International Journal of Computer Vision, 2018, 1-15.