An Eye-to-Hand Monocular Vision Robot Visual Servo System

Yajun Cheng [a], Xiaohong Ren [b, *], Shi Fan [c], Zhengguang Zheng [d] and Shaotang Cai [e]

School of Sichuan University of Science & Engineering, Zigong 643000, China.

[a]364063542@qq.com,[b]549859443i@qq.com,[c]121945157@qq.com, [d]2277606934@qq.com, [e]18855214@qq.com

Abstract

Industrial robots are widely used in our w. Most industrial robots use the teaching device to perform fixed motions on robots to reach specified positions and perform repetitive tasks. For industrial robots to grab target objects (such as mechanical parts, beverage bottles, etc.), if the target is offset by a short distance from the designated position, the industrial robot will not capture it, causing the industrial robot to fail to complete the work. The impact of the class situation is even more pronounced when it comes to picking up finer parts. In response to this problem, the eye-to-hand monocular vision servo system structure is proposed to enable the manipulator to have a small distance even if the target object is offset, and the feedback compensation of the visual servo system also enables the industrial robot to complete the task of capturing the target. As long as the distance from the target object does not exceed the camera's field of view, the camera uses the object image in the field of view for template matching to obtain the target's center coordinates. This coordinate is compared with the origin coordinates determined at the time of calibration, and a difference is generated. This difference represents the distance of the target on the X-axis and Y-axis. This offset distance parameter is fed back to the industrial robot. The industrial robot completes the grabbing of the offset target by compensating for the corrected motion trajectory.

Keywords

Robot; Visual positioning; Visual servoing.

1. Experimental Platform

This experiment uses ABB's Model 1410 industrial robot, which has six degrees of freedom and repeatability of positioning accuracy of 0.05mm. The camera uses a Cognex industrial camera CAM-CIC-5000R-14-G with 500W pixels, a resolution of 2592*1944, and a frame rate of 14fps. Industrial camera SSV 0358 F1.4, focal length 16mm. The overall structure is shown in Figure 1.

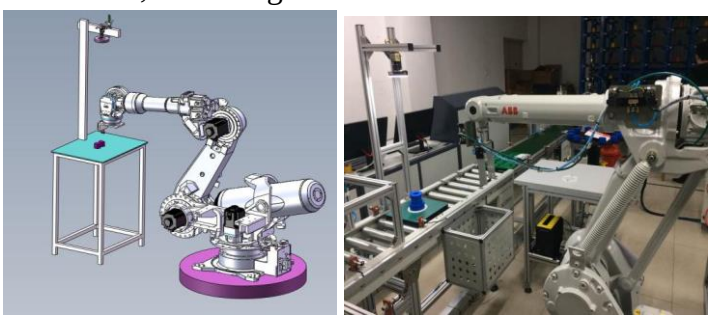


Fig. 1. Experimental platform

The camera is mounted to the outside of the robot, the distance from the camera to the target carrying plate is 695mm, the height of the target object is 85mm, and the diameter is 71mm. The visual field of the camera corresponding to the distance of the target carrying plate is approximately 249.2mm*186.1mm, That is to say, the target has an offset within the range where the standard point is centered, the vision system can identify the target, and make corresponding compensation, and transmit it to the industrial robot to make corresponding control.

2. Conversion of Coordinate System

The visual servo system involves three coordinate systems, a camera coordinate system, a target object coordinate system, and a manipulator coordinate system. The origin of the image coordinates of the coordinate system set by the camera is (1254.25, 1456.44).The transformation between coordinate systems is shown in Figure 2:

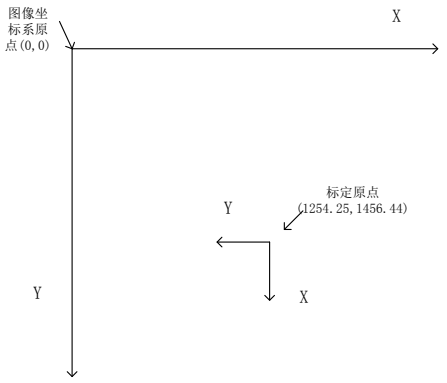


Fig. 2. Coordinate transformation

The visual servo has three coordinate systems, one is the coordinate system where the target object is located, one is the coordinate system that the camera plane occupies, and the last is the coordinate system that the image plane occupies. The relationship between their three coordinate systems is the transformation between the camera coordinates and the coordinates of the target object [1]:

$$\begin{pmatrix} x_c \\ y_c \\ z_c \\ 1 \end{pmatrix} = \begin{pmatrix} R & T \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_w \\ y_w \\ z_w \\ 1 \end{pmatrix} \tag{1}$$

(x_c , y_c , z_c) and (x_w , y_w , z_w) represent the coordinates of the target point in the camera coordinate system and the world coordinate system respectively, R is a 3*3 rotation matrix, and T is a translation vector.

Conversion between image coordinate system and camera coordinate system:

$$z_c \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_c \\ y_c \\ z_c \\ 1 \end{pmatrix} \tag{2}$$

(x, y) is the image coordinates and is represented by the actual size, f is the focal length of the camera.

Conversion between actual size and pixel size in the image coordinate system:

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} dx & 0 & -u_0 dx \\ 0 & dy & -v_0 dy \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} \tag{3}$$

dx,dy represent the physical size of a pixel on the x-axis and y-axis, respectively, and（u0 , v0）are the center coordinates of the image.

Conversion between industrial robotic coordinate systems

Changes from mechanical axis 0 to mechanical axis 1:

$$ {}^{0}_{1}T = \begin{bmatrix} c\theta_0 & -s\theta_0 & 0 & a_0 \\ 0 & 0 & 1 & d_0 \\ -s\theta_0 & -c\theta_0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{4} $$

Changes from mechanical axis 1 to mechanical axis 2:

a_0=150mm,d_0=475mm.

$$ {}^{1}_{2}T = \begin{bmatrix} c\theta_1 & -s\theta_1 & 0 & a_1 \\ s\theta_1 & c\theta_1 & 0 & 0 \\ 0 & 0 & 1 & d_1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{5} $$

The change matrix from mechanical axis 2 to mechanical axis 3:

a_1=600mm,d_1=0mm.

$$ {}^{2}_{3}T = \begin{bmatrix} c\theta_2 & -s\theta_2 & 0 & a_2 \\ 0 & 0 & 1 & d_2 \\ -s\theta_2 & -c\theta_2 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{6} $$

The change matrix from mechanical axis 3 to mechanical axis 4:

$$ {}^{3}_{4}T = \begin{bmatrix} c\theta_3 & -s\theta_3 & 0 & a_3 \\ 0 & 0 & -1 & -d_3 \\ s\theta_3 & c\theta_3 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{7} $$

The change matrix from mechanical axis 4 to mechanical axis 5:

$$ {}^{4}_{5}T = \begin{bmatrix} c\theta_4 & -s\theta_4 & 0 & a_4 \\ 0 & 0 & 1 & d_4 \\ -s\theta_4 & -c\theta_4 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{8} $$

The change matrix of the mechanical axis 5 to the end effector 6:

$$ {}^{5}_{6}T = \begin{bmatrix} c\theta_5 & -s\theta_5 & 0 & a_5 \\ s\theta_5 & c\theta_5 & 0 & 0 \\ 0 & 0 & 1 & d_5 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{9} $$

The end effector is the clip, the length of the clip plus the extension is 236.2mm, and the working range of the clip is: 44.3mm to 84.3mm.

The relative position of the end effector relative to the 0-axis reference point is [2]:

$$ {}^{0}_{6}T = {}^{0}_{1}T(\theta_0)\,{}^{1}_{2}T(\theta_1)\,{}^{2}_{3}T(\theta_2)\,{}^{3}_{4}T(\theta_3)\,{}^{4}_{5}T(\theta_4)\,{}^{5}_{6}T(\theta_5) \tag{10} $$

$$=$$

$$\begin{bmatrix} c_0[c_{12}(c_3c_4c_5-s_3s_5)-s_{12}s_4c_5]+s_0(s_3c_4c_5+c_3s_5) & c_0[c_{12}(-c_3c_4s_5-s_3c_5)+s_{12}s_4s_5]+s_0(c_3c_5-s_3c_4c_5) & -c_0(c_{12}c_3s_4+s_{12}c_4)-s_0s_3s_4 & c_0[a_1c_1+a_2c_{12}-d_3s_{12}]-d_1s_0 \\ s_0[c_{12}(c_3c_4c_5-s_3s_5)-s_{12}s_4c_5]-c_0(s_3c_4c_5+c_3s_5) & s_0[c_{12}(-c_3c_4s_5-s_3c_5)+s_{12}s_4s_5]-c_0(c_3c_5-s_3c_4c_5) & -s_0(c_{12}c_3s_4+s_{12}c_4)+c_0s_3s_4 & s_0[a_1c_1+a_2c_{12}-d_3s_{12}]+d_1c_0 \\ -s_{12}(c_3c_4c_5-s_3s_5)-c_{12}s_4c_5 & -s_{12}(-c_3c_4s_5-s_3c_5)+c_{12}s_4s_5 & s_{12}c_3s_4-c_{12}c_4 & -a_1s_{12}-a_1s_1-d_3c_{12} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The coordinate system established for each axis, and the rotation angle as the standard of movement.

Axis 1 is the chassis responsible for the robot's left and right movements, as if the left rotation was positive.

Axis 2 is responsible for the overall back and forth movement of the robot and moves forward to the positive direction.

The axis 3 is responsible for the up and down movement of the robot, and the downward movement is the positive direction.

The axis 4 is responsible for the left and right rotation of the connecting rod, and the right rotation is the positive direction.

The shaft 5 is responsible for the up and down movement of the end effector and moves downwards in the positive direction.

The shaft 6 is responsible for the left and right rotation of the end effector, and is rotated to the right in the positive direction.

The state of the robot shown in the figure is a state in which all six axes are 0 degrees, and the reference state is shown in Figure 3 .

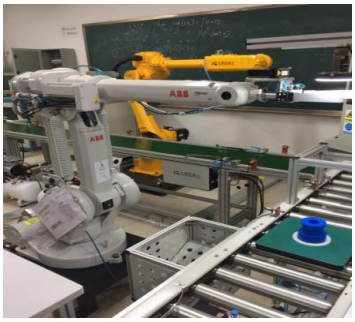


Fig. 3. The initial state

With the end-effector as the coordinate system, 6 axes are linked for cooperation. The X-axis controls the end-effector movement back and forth, forward movement is positive. The Y axis controls the left and right movement of the end effector and the leftward movement is the positive direction. The Z axis controls the end effector to move up and down, and the upward movement is positive.

Camera calibration and teaching of industrial robots: The standard position of the robot is shown in Figure 4, the teach pendant reads the end clip center Cartesian coordinates for $(x, y, z) = (1193.39, -37.29, 640.50)$

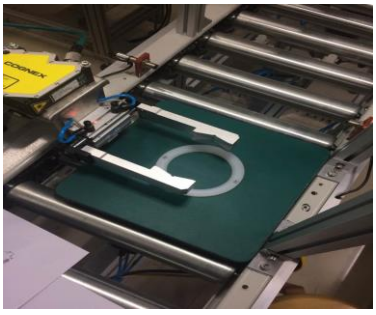


Fig. 4. Standard position of manipulator

The calibration results are as follows:

Use Matlab's Calibration Toolbox to calibrate the 10mm calibration plate placed on the carrier plate. The two blank cuboids in the figure represent the X-axis and Y-axis of the image coordinate axis, respectively. Irrespective of the distortion factor and cutoff distortion, the calibration results are shown in Figure 4 [3]:

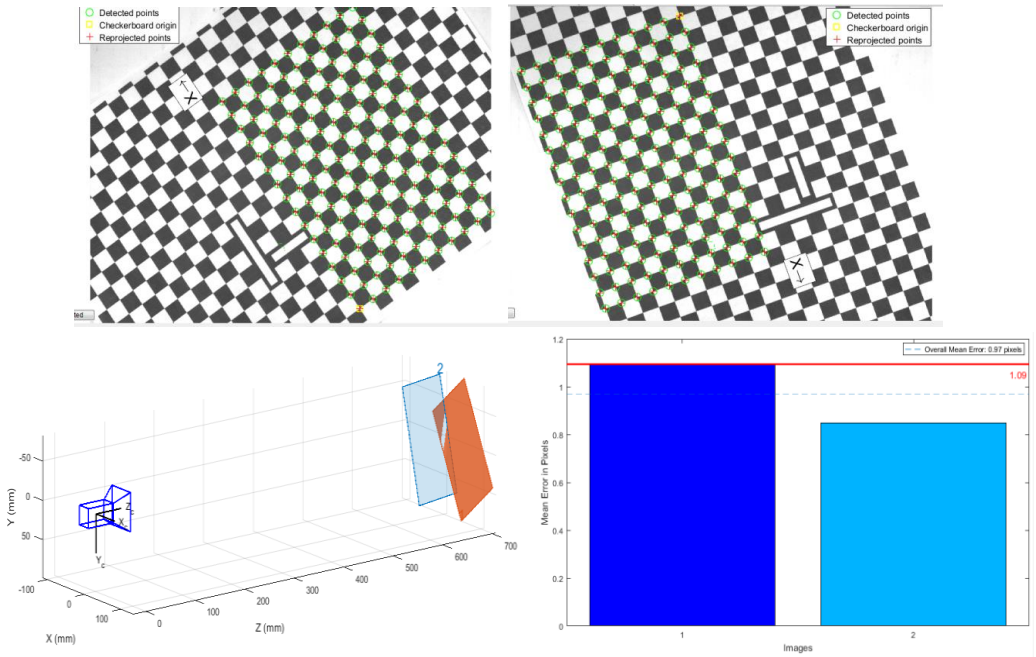


Fig. 5. Camera calibration

Internal parameter matrix:

$$K=\begin{pmatrix}\alpha_u & s & u_0\\ 0 & \alpha_v & v_0\\ 0 & 0 & 1\end{pmatrix}=\begin{pmatrix}7276.4 & 0 & 1221.5\\ 0 & 7272.0 & 841.4\\ 0 & 0 & 1\end{pmatrix}$$

$\alpha_u=\frac{f}{dx}$ $\alpha_v=\frac{f}{dy}$ (The unit here is pixel/)

Outer parameter matrix:

Rotation component $R=\begin{pmatrix}-0.0380 & -0.0476 & -2.1793\\ -0.0580 & -0.0112 & 1.2282\end{pmatrix}$

Translation component $T=\begin{pmatrix}50.3161 & 97.6980 & 697.8742\\ -17.5461 & -75.7725 & 705.3290\end{pmatrix}$

Radial distortion: [-0.3191 5.3922]

From the three-dimensional figure, we can see that the camera-to-calibration plate distance calculated by calibration is close to 700mm, which is very close to the actual measured camera-to-carriage plate distance of 695mm. It indicates that the calibration result is relatively accurate, and the mean squared error of the calibrated pixel is 0.97 pixels.

With the image obtained by the camera, one pixel size represents the actual distance of the physical space:dx=0.085mm/pix, dy=0.084mm/pix.

In order to make the end effector reach any position and posture in three-dimensional space, the robot needs at least six degrees of freedom, including three position degrees of freedom and three degrees of azimuth degrees of freedom.

The two links of an industrial robot are represented by two abstract geometric parameters. they are the common normal (at the same time perpendicular to the z axis of axis i-1 and axis i) distance a_{i-1}, and the angles between Z_{i-1} and Z_i in the plane perpendicular to the common normal are shown in Figure 6.

Table 1. Mechanical arm parameter

linki	ai-1(mm)	α_{i-1}(°)	Di-1(mm)	θ_{i-1}(°)
1	150	-90	475	0
2	600	0	0	-90
3	120	-90	0	0
4	0	90	720	0
5	0	-90	85	180
6	0	0	200	0

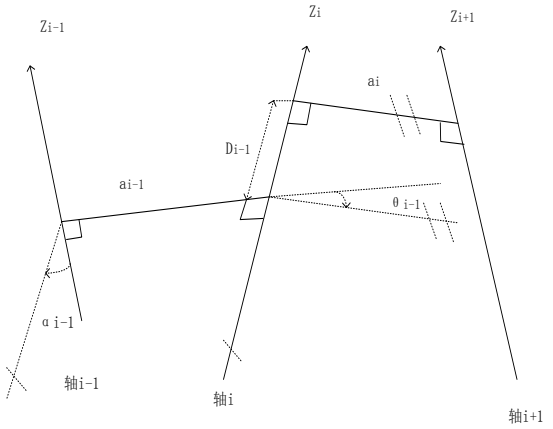


Fig. 6. The transformation between two links

The connection relationship between three adjacent links can also be abstracted into two other quantities. One is the common normal a_{i-1} of axis i and axis i-1, and the common normal a_i of axis i and axis i+1. Distance D_{i-1} on the Z_i axis. The second is the angle θ_i between the common normal line a_{i-1} and the common normal a_i on a plane perpendicular to the Z_i axis. In summary, the four variables describing the relative relationship between the various links of the industrial robot are link length, connecting rod twist angle, two-link distance D_i, and two-link angle θ_{i-1} [4].

α_{i-1} indicates the angle of rotation around the X-axis, θ_{i-1} indicating the angle of rotation about the Z-axis.a_{i-1} indicates the distance between two Z-axes, and D_{i-1} indicates the distance between two X-axes.

If there is a simultaneous rotation of θ_{i-1} and α_{i-1}, First rotate the Z axis θ_{i-1}, then revolve around the X axis α_{i-1}, The overall robot coordinate system structure is shown in Figure 7.

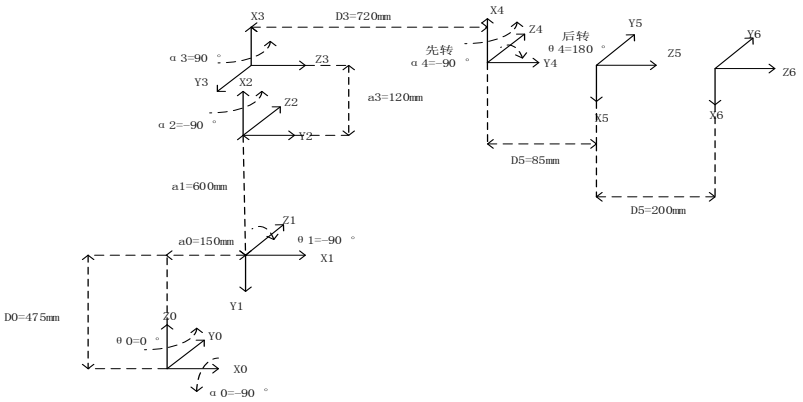


Fig. 7. Robot coordinate transformation

3. Image Processing

For different parts, select different features to achieve the correct and rapid identification of the target parts.

Target artifacts, using a blue cylinder for grab simulation. As shown in Figure 8:

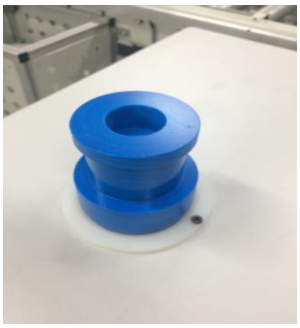


Fig. 8. The target artifact

For the workpiece shown in the figure, the selected feature method is to identify the circle, the white circle on the tray is the feature of constant size, as the main reference feature of the target coordinate position, by extracting the center coordinates of the white circle and initial calibration The coordinates of the white circles of the set standard reference pictures are compared, and a coordinate difference is obtained. This coordinate difference is fed back to the industrial robot, so that the compensation can be performed so that the industrial robot can still give a deviation. Accurately grab the target [5].

As shown in the figure, the standard reference diagram as the target part is shown in Figure 9.

Fig. 9. The standard image

The current image is compared with the reference image, and the similarity between the two features in the feature area reaches a certain ratio to determine whether the parts in the two images are the same and whether they need to be captured [6]. In order to avoid noise interference in the visual field of the

image, the limit and The method of narrowing the feature area to avoid unnecessary noise interference with the target recognition result is as follows:

The feature extraction area is mainly concentrated in the white circle of the tray. For other interference objects outside the white circle, the system automatically ignores it and does not recognize it. Unless the features other than the white circle and the feature height within the white circle are known, this time will cause A misjudgment of the system will make the system confused about who is the target area to measure[7]. Under normal circumstances, this situation will not occur, unless the disturbing material outside the white circle meets the following two characteristics, one is a circle, and the other is that the inner and outer diameters of the circle are basically the same as the inner and outer diameters of the white circle. For the interference in the white circle, for example, there are other interferences on the circle, and the circle is partially blocked, resulting in the lack of completeness of the circle. In this case, the system will give a weight, as long as the occlusion part is not larger than a proportional value, The system automatically fits the white circle. Generally, it does not affect the detection result. This weight can be manually set according to the specific situation [8].

In summary, the whole system has high anti-interference and stability in terms of detection in the entire visual field of the camera, no matter whether it is inside or outside the white circle.

For different situations of these types of interference, the renderings are as follows:

The interior of the white circle is partially blocked. As shown in the figure, there is a green circle in the figure indicating correct recognition. The recognition effect is shown in the figure.

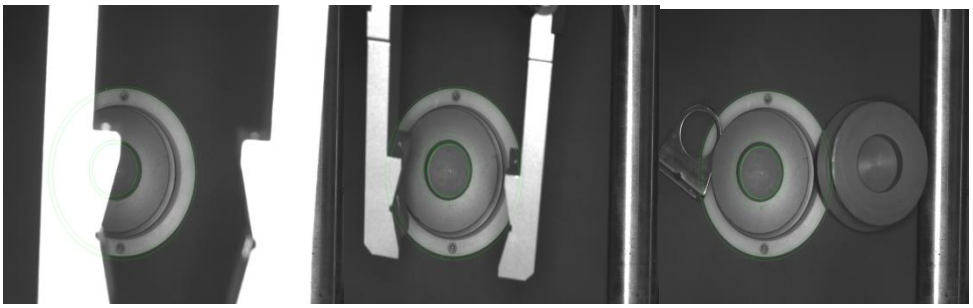


Fig. 11. Interference experiment 1

It can be seen from the figure that the target object is partially occluded, and the vision system can still complete the recognition of the target object through the fitting function. Although the recognition effect of the target object being occluded will be lower than that when it is not occluded, it will reduce The experimental results show that this reduced deviation is very small.

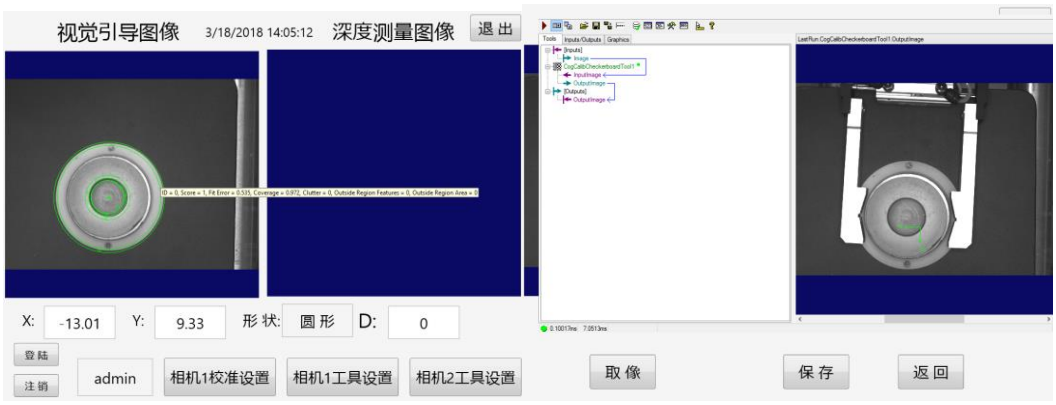
Outside the white circle, there are other interferences, such as circular objects. The experimental results are as follows:

Fig. 12. Interference experiment 2

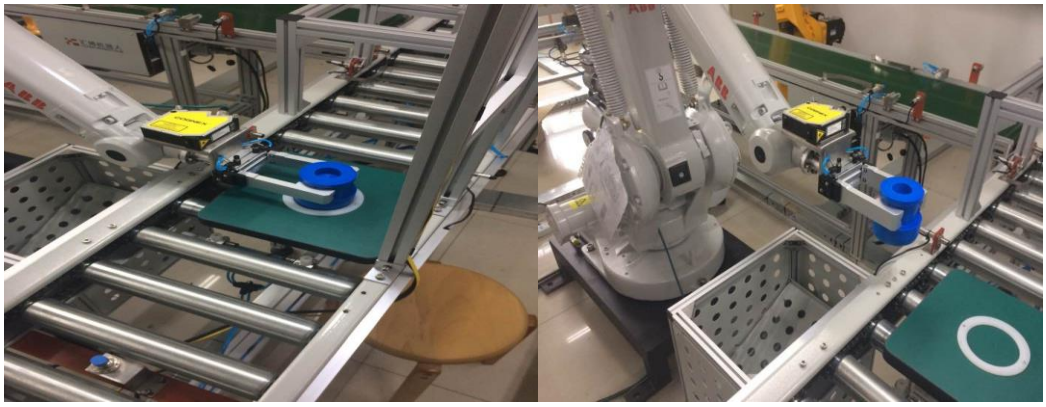
It can be seen from the figure that the interfering objects outside the white circles do not affect the experimental results. The vision system automatically ignores these interferences without the need for denoising.

4. The Experimental Results

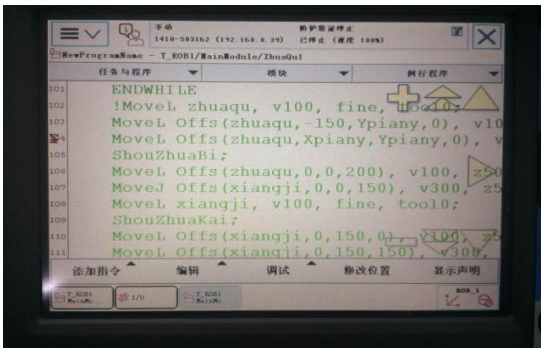
The target workpiece deviates from the standard position. The effect of the industrial robot crawling task experiment is shown in Figure 13.



a) calculation deviation b) compensation deviation



b) Clipping d) Exercise



e) Robot compensation program

Fig 13. Grab task experiment

The image processing interface created on the conex desigen software is shown in Figure 12a). From the experimental results, the deviation of the experimental workpiece from the set standard position on the x-axis y-axis is (-13.01mm, 9.33mm) Corresponds to the teach pendant (Xpiany, Ypiany), the deviation of the workpiece in the camera field of view, through the visual processing can give the robot a more satisfactory compensation, to achieve our desired results.

5. Conclusion

In this paper, the vision system is introduced into industrial robots, and the workpieces offset from the standard position can be accurately captured. The experimental results show that the image processing strategy adopted in this paper is also robust under the interference of various complicated conditions. Can correctly identify the target, with better reference value and practical value.

Acknowledgements

Natural Science Foundation. Intelligent signal and information processing sichuan youth science and technology innovation research team（2015TD0022）; The first batch of key r&d projects in sichuan province in 2017.（2017GZ0068）

The first author:Yajun Cheng

*represent Corresponding author:Xiaohong Ren

References

[1] PeterCorke, K. Ke, R. Liu. Robotics, machine vision and control: Basic MATLAB algorithm [M](Electronic Industry Press, China 2016).
[2] Y.C. Fang. Research on Robot Vision Servo[J], Journal of Intelligent Systems, vol. Vol. 3 (2008) No. 2, p. 109-114.
[3] Z.W. Zhu. Research on the visual servo control system of joint robot[D]. Zhengzhou University, 2009.
[4] Chaumette F. Visual servo control-part II: advanced approaches[J]. IEEE Robotics Automat Mag, Vol. 14 (2007) No. 1, p. 109-118.
[5] Horaud R, Dornaika F, Espiau B. Visually guided object grasping[J]. IEEE Transactions on Robotics & Automation, Vol. 14 (1998) No. 4, p. 525-532.
[6] L.K. Qiu. Research on visual servo control based on manipulator and its application[D]. University of Science and Technology of China, 2006
[7] John C. Russ. Digital Image Processing [M]. Publishing House of Electronics Industry, 2014.
[8] Z.X. Cai. Fundamentals of Robotics[M]. Mechanical Industry Press, 2015.