An Application of Vision Technology on Intelligent Sorting System by Delta Robot

Shiyu Wang
Shenyang Institute of Computing Technology
University of Chinese Academy of Sciences
Beijing, P.R. China
wangshiyu0410@163.com

Hu Lin
Shenyang Institute of Computing Technology
Chinese Academy of Sciences
Shenyang, P.R. China
linhu@sict.ac.cn

Rongli Gai*
School of Information Engineering
Dalian University
Dalian, P.R. China
gairli@sict.ac.cn

Yilan Sun
Shenyang Golding NC & Intelligence Tech Co., Ltd
Chinese Academy of Sciences
Shenyang, P.R. China
sunyilan@sict.ac.cn

***Abstract*—Combine the visual module with the Delta parallel robot to set up the sorting experiment platform based on visual technology. Matching results are sensitive to environmental interference. In order to eliminate the influence such as illumination and noise on matching process, the traditional grayscale features are replaced by edge features. The advantages of using edge features are that the edge contains complete information about the image and there is less data to be deal with. The input images are analyzed under multi-scale filtering in the Canny algorithm frame to extract accurate edge features, then build template sets according to the possible rotation angles of the objects and achieve the matching process between the targets and the templates. Experimental results conducted that the visual system can identify and locate the objects accurately and this experiment platform is able to collect the experimental objects with high success rate.**

Keywords- Delta parallel robot; edge features; Multi-scale filtering; Canny operator

I. INTRODUCTION

Sorting is an important part of most industrial assembly lines in the fields of electronics, automobiles, food and pharmaceuticals. Traditional robot sorting process generally adopt “teach-in” method or off-line programming to control the movement of the robot, it is difficult to adapt to the complex working environment in the present industrial production. In recent years thanks to the country to develop the robot industry, robot vision technology has been able to flourish. Combined with the visual technology, robots have higher intelligent degree and better adaptability to the environmental.

Dalta parallel robot [1] is invented by Dr. Clavel, Swiss federal Institute of Technology in Lausanne, in 1985. After decades of development, the current Delta series robots presents the characteristics of high speed, high precision, flexibility, it is very suitable for high-speed sorting operation with small workpiece [2]. In this paper, we have designed and built an intelligent sorting experiment platform, which consists mainly of Delta parallel robot and Cognex In-Sight series industrial camera, based on the related visual technology. Visual system guides the robot end-effector for subsequent operations according to the position and category information obtained by testing, orientating and classifying the sorting objects. In terms of the visual system, this paper proposes an improved matching algorithm based on multi-scale Canny operator to avoid the interference caused by environmental factors such as illumination variation and noise. Computer simulation results show that the improved matching algorithm significantly reduces the effects of illumination and noise, so that no additional light source is needed. This experimental platform combine the advantages of robot vision and Delta parallel robot, it improves the efficiency and the flexibility of the sorting system. Therefore, the Delta robot sorting system combined with visual system has a broad prospect of application and research value.

II. THE OVERALL DESIGN OF THE SORTING EXPERIMENT PLATFORM

Fig. 1 shows the complete experimental platform. The sorting experiment platform consists of several parts, such as Delta parallel robot, industrial camera, conveyor belt and sorting objects. The structure character of Delta parallel robot is simple and compact, and three drive mechanisms of driving arms are distributed on the fixed platform uniformly by the angle of 120°. Driven arms and moving platform are both made of lightweight material, so the end-effector can obtains high speed and acceleration which make the Delta parallel robot has good kinematic and dynamic characteristics [3]. The parallelogram structure of driven arms eliminate the rotational degree of the moving platform and makes Delta robot achieve translational motion of three degrees of freedom (DOF) in the motion space. Furthermore, adding a telescoping shaft that mainly used to modify the angle of the experimental subjects connecting the center of

978-1-5090-6704-6/17/$31.00 ©2017 IEEE

the fixed platform and moving platform can achieve the motion form composed of translational motion and rotating motion. Industrial camera is fixed above the conveyor belt. When the objects come into the view area of the camera, the computer obtain the classification information, coordinate information and the rotation angle with certain processing algorithm, and later it will send these information to the robot controller. Based on the information acquired from the visual system, the robot control system will drive the end-effector to track and pick up objects in the appropriate area.

Figure 1. Sorting experiment platform.

III. TARGET MATCHING

Visual module is the important embodiment of the intelligent upgrade in sorting system and reliable target matching is an important task in robot vision. The accuracy of the sorting system depends on the matching accuracy. Hence, the research emphasis of this article is concentrated on the process of the image matching.

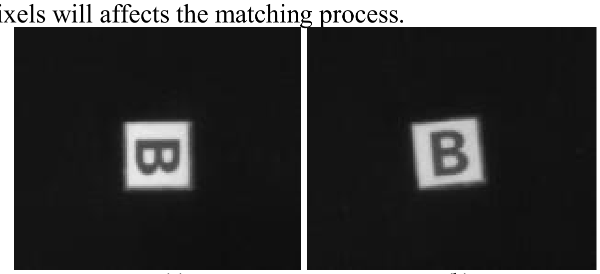
The simplest method to locate a known target in an image is to search the perfect copy in pixel level, which requires to control the pose of the targets and illumination condition. At the same time, it also has the requirement on the targets without scale and rotation transformation, it is difficult to achieve in reality. The SIFT- scale-invariant feature transform proposed by Lowe [4] in 2004 is the classic way to solve this kind of problem, then Bay [5] proposed SIFT's accelerated vision SURF. Both of them need to establish the scale space and they are commonly used in areas such as 3D reconstruction and visual tracking. The industrial assembly line based on the visual system is mostly 2D flat application, the working distance of industrial camera is fixed at certain height. In 2D flat applications, the image information collected is only affected by rotation and illumination, the scale of the image doesn't change. The sorting experimental platform designed in this article has put forward the solution to the inevitable phenomenon about rotating and illumination variation during the matching process [6].

A. *Template matching*

The idea of template matching is really simple [7]. Compare the grey value between the known template and a same size area in the original image [8]. There is a given template whose dimension is $r \times c$ and an image I. Assume that the image is pinned to the offset (x_0, y_0) , when the template can match the target completely, there will be equation (1)

$$E(x) = \sum_{i=1}^{r} \sum_{j=1}^{c} \left(M_{i,j} - I_{x_0+i, y_0+j} \right)^2 = 0 \qquad (1)$$

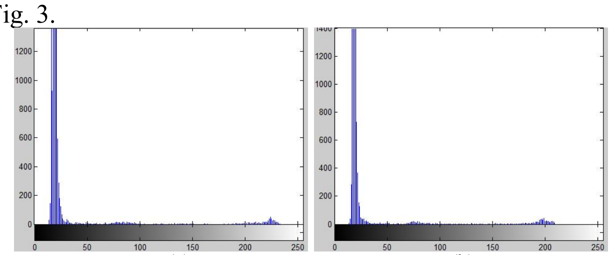
Equation (1) includes the information about position and gray value, where E represents the measurement error of the matching. The minimum value of E indicates the result of template matching to some extent. The sum term in (1) is sensitive to illumination variation and rotation. Illumination variation manly affects the variable that represents gray values in (1) and the target rotation affects the position information in the matching process. Fig. 2 shows the images of same target taken in different light conditions, and the illumination in Fig. 2(a) is more abundant than that in Fig. 2(b). So we can observe that the white block in Fig. 2(a) is brighter than that in Fig. 2(b). As for the pixels in the background, it is difficult to distinguish the difference of the gray value between Fig. 2(a) and Fig. 2(b). However, the illumination variation does change the gray value of the pixels in the two images and the varying gray value of the pixels will affects the matching process.



(a) (b)

Figure 2. Images of different lighting conditions

The gray distribution of the image shown in Fig. 2 is not obvious and the effect of illumination on the gray distribution is more intuitional with gray histogram shown in Fig. 3.



(a) (b)

Figure 3. The corresponding greyscale distribution in Fig 2

It can be seen that most gray components belong to the black background and the pixels are mainly distributed in two intervals by analyzing the histogram. The distribution of the pixels on the bright interval can reflect the effects of the illumination variation more clearly. However, there is also variation in the distribution of pixels on the dark interval which cannot be seen clearly in Fig.2.

B. Edge feature extraction

Some theories assume that images are preprocessed by an edge detector before matching is done [9]. In this paper, computer image processing technology is used to carry on binary processing for template image and target image in order to eliminate the effects of varying grey value on the matching process. Firstly, obtain the edge feature by edge detection operator and then make the response to the edge expressed in white pixels and set the other pixels to zero. Therefore, no matter how the light environment changes, it doesn't affect the matching process. On the other hand, it also requires the accuracy of edge feature extraction.

Edge is the location where the grey value changes sharply. Canny [10] proposed the three criteria of edge detection: good detection, good localization and only one response to a single edge. What is the first thing need to do is to smooth the image and reduce the noise. Marr and Hildreth proposed two criteria in literature [11] about how to select the smoothing filter: in order to avoid the influence of extra frequency on image function, filter should be smooth and should be limited bandwidth in the frequency domain. Another is that the response to the filter should be from the adjacent points in the image. According to the uncertainty principle, the band-limited function in frequency domain is infinite in spatial domain, which will not satisfy the second criteria mentioned above. The Fourier transform of Gaussian function is still Gaussian also, and the only parameter of Gaussian is the standard deviation σ, which is proportional to the size of filtering neighborhood. The effect of pixels can be ignored when the distance is more than 3 σ from the center. These make the two criteria for selecting filter satisfied simultaneously.

Canny operator always omits feature synthesis in the general implementation and adopts the single filter scale for noise suppression and edge detection. Although the complete structure of the edge will be maintained when using a small scale filtering parameter, the filter is sensitive to noise. On the contrary, the large scale filtering parameter will cause the loss of the edge structure while inhibiting the noise. Compared with the accumulated method to generate edge in feature synthesis, the improved Canny algorithm based on scale multiplication, which is proposed by Pan Dafu with his colleagues in literature [12], has the simpler idea and the less calculation.

The good performance of edge detection requires to maximize signal-to-noise ratio. Let the impulse response of the filter be $f\left(x\right)$ and assume the filter has a finite response bounded by [-T, T]. The edge itself is denoted by $G\left(x\right)$. Assume the edge is centered at $x=0$. Then the response of the filter to edge at its center is given by the convolution integral:

$$H_G=\int_{-T}^{+T} G\left(-x\right) f\left(x\right) dx \tag{2}$$

The response to the noise $n_0^{\ 2}$ only will be

$$E\left(H_n\right)=n_0^{\ 2}\int_{-T}^{+T} f^2\left(x\right) dx \tag{3}$$

Where $n_0^{\ 2}$ is the mean-squared noise amplitude per unit length. Hence, the mathematical form of the edge detection can be expressed as the SNR at the edge

$$SNR=\frac{\left|\int_{-T}^{+T} G\left(-x\right) f\left(x\right) dx\right|}{n_0\sqrt{\int_{-T}^{+T} f^2\left(x\right) dx}} \tag{4}$$

Defining the operator $f_w\left(x\right)=f\left(x/w\right)$ which is the result of stretching f by scale factor of w. So the response to the noisy edge is given by

$$H^w=H_G^w+H_n^w \tag{5}$$

In order to guarantee the precision of the edge feature extraction, keep the edge information complete and reduce the effects of noise on edge response as far as possible, use the larger scale factor of w_1 and the smaller scale factor of w_2 for edge detection. The corresponding responses of the step edge $G(x)$ is H^{w_1} and H^{w_2} respectively. The mathematical form of multi-scale edge detection is obtained by bringing different scale parameters into (4)

$$SNR_w=\frac{\sqrt{\left|H_G^{w_1}\cdot H_G^{w_2}\right|}}{\sqrt{\left|E\left[H_n^{w_1}\cdot H_n^{w_2}\right]\right|}} \tag{6}$$

Mark edges at local maxima in the response of the filtering operator $f\left(x\right)$, the first derivative of the response will be zero at these points. Let H_n be the response of the filter to noise only and H_G be the response to edge. We have assumed that edges are centered at $x=0$, so there should be a local maximum in the response at $x=0$ in the absence of noise. For the noisy signals, if there is a local maximum in the total response at the point $x=x_0$, then we have

$$H_n^{'}\left(x_0\right)+H_G^{'}\left(x_0\right)=0 \tag{7}$$

The response of multi-scale multiplication to the noisy edges can be expressed as (8)

$$H^{w_1{'}}\left(x_0\right)\cdot H^{w_2}\left(x_0\right)+H^{w_1}\left(x_0\right)\cdot H^{w_2{'}}\left(x_0\right)=0 \tag{8}$$

The Taylor expansion of $H_G^{'}\left(x_0\right)$ about the origin gives

$$H_G^{'}\left(x_0\right)=H_G^{'}\left(0\right)+H_G^{''}\left(0\right)x_0+O\left(x_0^{\ 2}\right) \tag{9}$$

Since we have considered the edge is centered at $x=0$, the first term in the expansion can be ignored because the response of the filter in the absence of noise has a local maximum at the origin. The displacement x_0 of the actual maximum is assumed to be small so we can also ignore quadratic and higher terms. And equations (7) and (9) give

$$H_n^{'}\left(x_0\right)+H_G^{''}\left(0\right)x_0=0 \tag{10}$$

According to the conclusion of (10), the Taylor expansion of (8) gives

$$x_0 = \frac{-H_n^{w_1'}(0)\cdot H^{w_2}(0) - H_n^{w_2'}(0)\cdot H^{w_1}(0)}{H_G^{w_1''}(0)\cdot H^{w_2}(0) + H_G^{w_2''}(0)\cdot H^{w_1}(0)} \quad (11)$$

Equations (6) and (11) are the mathematical forms for the edge detection and localization by multi-scale multiplication. We seek to maximize (6) and minimize (11). To simplify the analysis for edge feature extraction, define the reciprocal of (11) as the mathematical form of the edge localization. For the present the criterion of edge feature extraction is converted to maximize the product of (6) and the reciprocal of (11).

In one dimension we have analyzed the characters of the multi-scale multiplication in edge feature extraction. The image processing in two dimension can be view as the combination of one-dimensional process in two directions. Apply the method mentioned in preceding paragraphs to actual images to obtain edge features. Compare with the traditional Canny operator, the operator based on multi-scale multiplication has better performance on edge detection and better accuracy on edge localization. Fig. 4 shows the results of the traditional Canny operator and the operator based on multi-scale multiplication respectively.

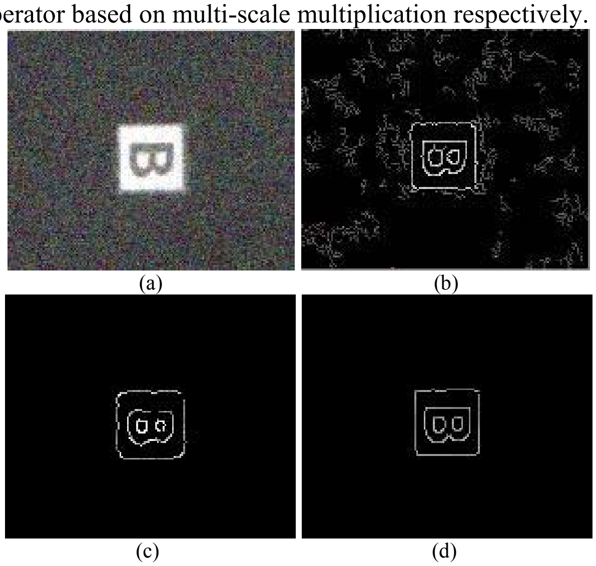


Figure 4. Edge feature extraction (a) noisy image (b) the result of smaller scale parameter (c) the result of larger scale parameter (d) the result of multi-scale multiplication

C. Template sets

At present, we've got accurate edge feature information, which can reduce the noise interference during the matching process. As for the changes in grey value caused by the illumination variation, turn the feature images of templates and targets into binary images. There are only two kinds of pixels in binary image where 0 represents black pixels and 1 represents white pixels. The computer will store the binary image as a binary matrix in the memory. The actual grey scale of the image ranges from 0 to 255, and binarization process can be given by

$$dst(x,y) = \begin{cases} 0 & if\ src(x,y) < thresh \\ 255 & otherwise \end{cases} \quad (12)$$

The *src* in (12) represents the feature image while the *dst* represents the final binary image. Binary images avoid the effect of grey value on matching process. Fig. 5 shows the matching process of binary images.

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

(a) (b)

Figure 5. Matching process

The above matching process needs to make the rotation angles of templates consistent with that in target images. Different kinds of cube blocks are scattered on the conveyor during the sorting process and the rotation phenomenon is inevitable. The robot controller need to acquire the rotation information in order to ensure the accuracy of the matching process. The visual system will select the coordinates (x_1, y_1) with the minimum x and the coordinates (x_2, y_2) with the maximum y once a rectangle target is detected in the visual coordinate as shows in Fig. 6.

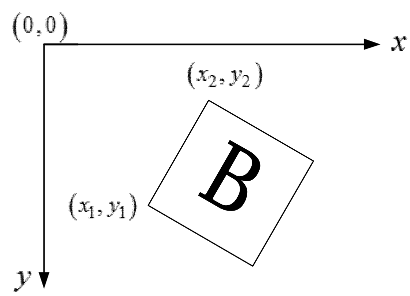


Figure 6. The rotation angle of the objects

The pattern feature of the object block is asymmetric, so there will be several rotation angles possibly. The rotation angles θ will be given by

$$\theta = \arctan\left(\frac{|x_2 - x_1|}{|y_2 - y_1|}\right) + \frac{\pi}{2} i \quad (i = 0,1,2,3) \quad (13)$$

Then rotate the template

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (14)$$

The coordinates (x, y) in (14) represents the position information of the input template while the coordinates (u, v) represents the rotated template. Save the possible rotation angles to build the template sets as shows in Fig. 7.

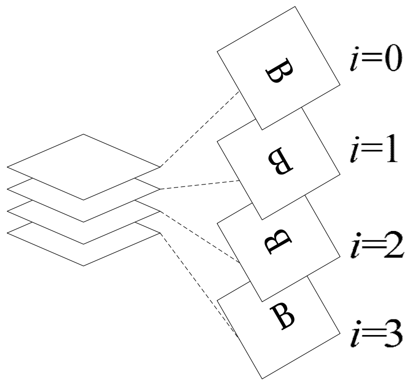
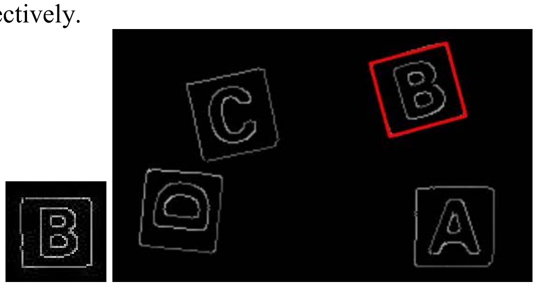
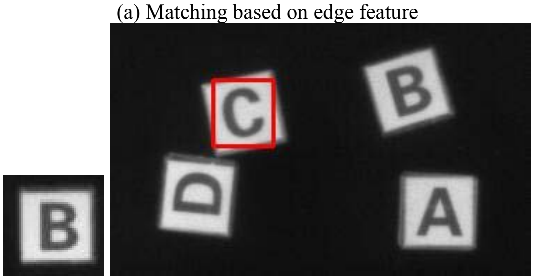


Figure 7. Template sets

Visual system will use template sets to match the images collected by the industrial camera. Once the match is successful, the type information, coordinate information and rotation angle of the experiment cube blocks will be recorded and sent to the robot control system for subsequent sorting operation and angle compensation [13]. For the situation that edges of the objects to be identified are irregular, the visual system can obtain the minimum enclosing rectangle of objects before using the above method. Fig. 8 shows the results of the template matching based on edge feature and the gray-scale correlation respectively.



(a) Matching based on edge feature



(b) Matching based on gray-scale correlation

Figure 8. Matching results

It is easy to obtain wrong matching results for template matching based on gray-scale correlation while the process based on edge feature gets the better matching results as shown in figure 8. Although extra data may be introduced to the processor in process of building the template sets, the binary images also exclude a large amount of unnecessary data.

D. User coordinate system

The robot control system needs to convert the position information of the visual coordinates into the coordinates in the robot operating area, which need us to establish the corresponding relationship between visual module and the robot end-effector. The corresponding relationship between the visual system and the end-effector of the robot is shown in Fig. 9.

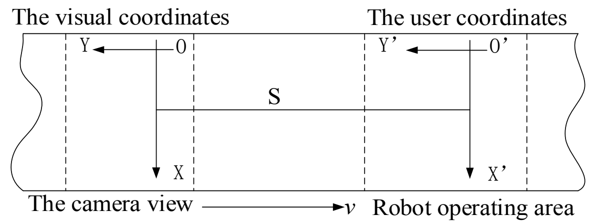


Figure 9. The camera view and robot operating area

In order to guarantee all the objects on the conveyor can be identified by visual module, the speed of the conveyor v and the camera sampling frequency f should be satisfy the relationship

$$S \times f - v > 0 \tag{15}$$

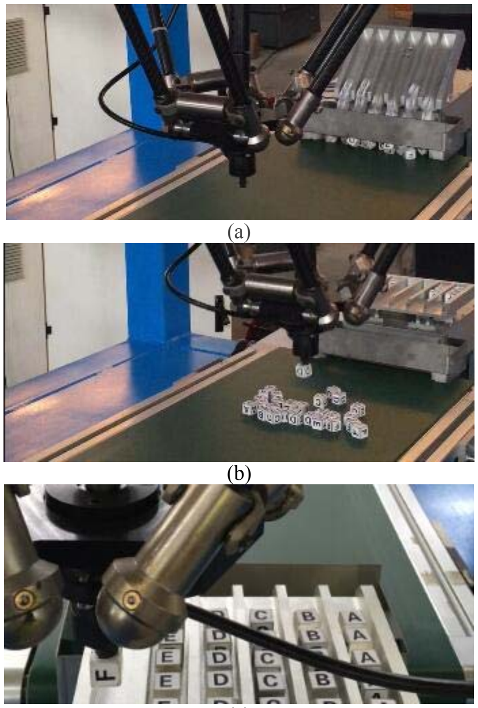
The encoder is a device for recording the conveyor. Firstly, Select a reference location on the conveyor and record the corresponding encoder value a. Then start the conveyor to keep the reference location moving a distance and record the value b of the encoder at this time. Let S represents the difference between a and b. Finally, we operate the robot to establish the user coordinates. It is important that the origin in user coordinates and the coordinate axis should be consistent with that in visual coordinates. At this point, the corresponding relationship between the visual module and the robot end-effector has been set up

$$\begin{cases} y' = y - S + vt \\ x' = x \\ z = -H \end{cases} \tag{16}$$

Visual technology is integrated into the robot sorting system by establishing the corresponding relationship between the visual module and the robot end-effector.

IV. THE ANALYSIS OF EXPERIMENTAL RESULT

Fig. 10 shows some screenshots of the video that recorded the whole sorting process. Different kinds of cube blocks are scattered on the conveyor randomly by opening the cylinder. At the same time, the robot controller trigger the camera at intervals to capture images and the visual system will examine whether there are objects in the view area of the camera. Then the visual system matches the pre-established templates with the collected images using the methods described above. After obtaining the relevant information, the robot end-effector will pick up the objects according to the acquired information and position them in the appropriate area. The changing lighting conditions are specially introduced to the experimental process to test and verify the accuracy of the matching algorithm proposed in this paper. Fig. 10(c) shows the experimental result that robot sorting system can correctly identify and grasp experimental objects. And the matching algorithm proposed in this paper can effectively link the process of identify and locate together and strengthen the practicability of matching algorithms.



(a)

(b)

(c)

Figure 10. Experiment results

V. CONCLUSIONS

This paper designed and built a set of robot sorting experimental platform. Combining the advantages of robot vision and Delta parallel robot improves the efficiency and the flexibility of the sorting system. For most 2D flat application, for example, the sorting application mentioned in this article, template matching is a good choice. But the gray-scale correlation based template matching algorithm is easily influenced by lighting and rotation. Transform the above problems into the variation of the grey value and position information in the template matching algorithm. Firstly, we extracted image edge feature with Canny operator based on multi-scale multiplication and obtained edge feature images. Then the binarization of the feature images excluded the changes in gray value caused by the illumination variation. Aiming at the rotation of the experiment objects, we set up the template sets by acquiring the possible rotation angle of the objects for overcoming the influence of the variation of position information on the matching process. The introduction of the template sets seems to increase the computation of the computer, but the result of the binarization also greatly reduces the data that the computer actually needs to process during the operation. Finally, the experimental results also prove that the robot sorting system with the matching algorithm mentioned in this paper can correctly identify and locate the experiment objects.

ACKNOWLEDGEMENT

This work was supported by Natural Science Foundation of China: NO.61602074. Rongli Gai is the corresponding author.

REFERENCES

[1] Vischer, Peter, Clavel, et al. "Kinematic calibration of the parallel Delta robot." Robotica, 2001, vol.16, no.2, pp.207-218.

[2] Sumin Yin, Changjun Zheng, Qixiang Xu, et al. "Control System Design of Sysmac Automatic Platform." Machine Tool & Hydraulics, 2017, vol.45, no.3, pp.1-5

[3] Xuming Dong, Zhibin Li. "Application of the Computer Vision in Kinematic Calibration of Parallel Robot." Process Automation Instrumentation, 2016, vol. 37, no.5, pp.16-19.

[4] Lowe D. G. "Distinctive Image Features from Scale-Invariant Keypoints." International Journal of Computer Vision, 2004, vol.60, no.2, pp. 91-110.

[5] Bay H, Tuytelaars T, Gool L V. "SURF: Speeded Up Robust Features." Computer Vision & Image Understanding, 2006, vol.110, no.3, pp. 404-417.

[6] Ngo, H.Q.T, Nguyen, et al. "Design and implementation of high performance motion controller for 2-D delta robot." Seventh International Conference on Information Science and Technology. IEEE, 2017, pp.129-134.

[7] Zhaohui Guo, Stanley Osher. "Template Matching via l_1 Minimization and Its Application to Hyperspectral Data." Inverse Problems & Imaging, 2017, vol.5, no.1, pp.19-35.

[8] Milan S, Roger B, Vaclav H. "Image Processing, Analysis and Machine Vision." Journal of Electronic Imaging, 2014, vol.19, no.82, pp. 685-686.

[9] D. C. Marr, and T. Poggio, "A computational theory of human stereo vision." Proc. Roy. Soc. London.,1979, vol.B 204, pp.301-328

[10] Canny J,. "A computational approach to edge detection." Pattern Analysis & Machine Intelligence IEEE Transactions on, 1986, PAMI- 1986, vol.8, no.6, pp. 679-698.

[11] Marr D. Hildreth E. Theory of Edge Detection[C]//Proceedings of Royal Soc.. London, UK: [s. n.], 1980: 187-217.

[12] Dafu Pan, Bo Wang, Zhiqiang Zhou. "Improved Canny Edge Detection Algorithm Based on Scale Multiplication." Computer Engineering, 2009, vol.35, no.3, pp. 222-224.

[13] Hepeng Ni, Ya'nan Liu, Chengrui Zhang, et al. "Sorting System Algorithms Based on Machine Vision for Delta Robot." Robot, 2016, vol.38, no.1, pp.49-55.