

# Development of a Real-time Detection System for Shaped Objects with Binocular Vision of a Mobile Robot

<sup>1</sup>Guo-Shing Huang and <sup>2</sup>Wen-Lang Zhan

## Abstract

The purpose of this paper is to develop a real-time detection system for trapezoid-shaped and ellipse-shaped objects, which is capable of obtaining the 3D position coordinates of objects when it is applied to the binocular machine vision of an order taking robot. The rectangular tray to be detected by the system may appear to be in trapezoid shape on the camera screen due to different ways of placement, but the round plate may appear to be in ellipse shape. This has led to the conclusion that the machine vision of robots must be capable of detecting trapezoid and ellipse shapes. The image coordinates can be used in coordination with 3D reconstruction principle to obtain the 3D coordinates of the two access points on the left and right of the object, so that the pair of mechanical arms of an order taking robot can more precisely handle trays and round plates. Based on actual detection results, the success rate, accuracy, and reliability of such identification can all meet the system requirement.

**Keywords:** meal service robot, dual robotic arms, binocular vision system, mobile robot

## 1. Introduction

The purpose of this paper is to develop a robot which can identify geometric objects and obtain coordinates of pixels from such geometric objects.

The clear binary image of an object is necessary for detection of geometric object. The literatures involved in the following discussion will contribute to this study. In terms of enhanced image identity, a so-called digital image interpolation method can be used to calculate new pixel data in accordance with data of surrounding pixels [5]. In literature [6] the stepwise refinement method has been used for problem solving, and such method did contribute to image pre-processing, image identification, and image enhancement. Literature [7] is about detection of feature points which contain the

---

\*Corresponding Author: Guo-Shing Huang  
(E-mail: hgs@ncut.edu.tw)

Department of Electronic Engineering, National Chin-Yi University of Technology, 57, Sec.2, Zhongshan Rd., Taiping Dist., Taichung 41170, Taiwan, ROC.

features of the entire image, and such method has been widely utilized in the field of machine vision. The algorithm in dissertation [8] performs better than traditional algorithm in terms of detection. Accuracy of the signal edge has been enhanced, and image details have been retained. In [9] an adaptive edge detection algorithm has been proposed in accordance with multiform and multi-scale structure elements, which provides effective noise suppression and enhanced detection accuracy. For measuring the distance between the camera and the rectangular object, a detection algorithm for the distance of video image of rectangular figure was proposed in [10], which has successfully overcome the deficiency of traditional method. There are a total of eight sections in this article. In addition to the introduction in Section I, Section II is about the introduction of hardware and system appearance and architecture, while the image detection method is introduced in Section III, Section IV and Section V. Section VI is about the description of actual testing procedure, while the experimental results and discussions can be found in Section VII. The overall conclusion will be made in Section VIII.

## 2. Hardware Architecture

The purpose of this article is to investigate a set of binocular machine vision system which can be installed on the platform of intelligent robots. The Logitech QuickCam E3500 Webcam has been used in this system, and the binocular camera platform of this vision system is shown in Figure 1. The step motor controlled by 8051 chip is used by the circuit board, and the data is transmitted from the computer to 8051 chip through RS-232 in order to control the binocular camera platform of this vision system to move upward and downward by 30 $\mu$ m as shown in Figure 2.

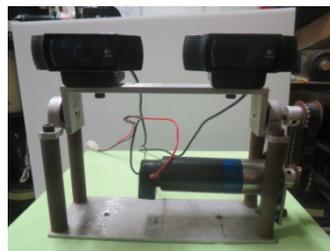


Figure 1: Camera platform



Figure 2: Order taking robot

The image resolution used in this article is 320×240 RGB color space with the refresh rate of 30 frames per second. The circuit diagram of data transmission from PC to control board is shown in Figure 4, while the architecture of binocular vision is shown in Figure 3.

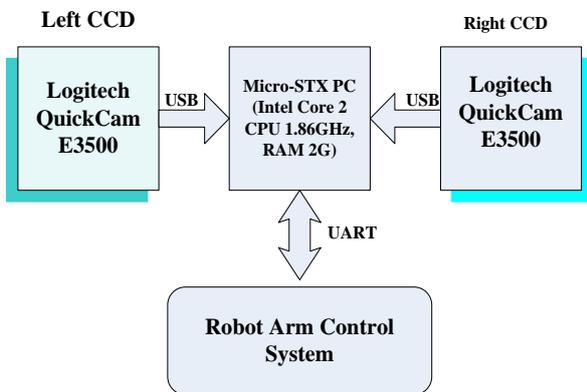


Figure 3: Binocular vision architecture

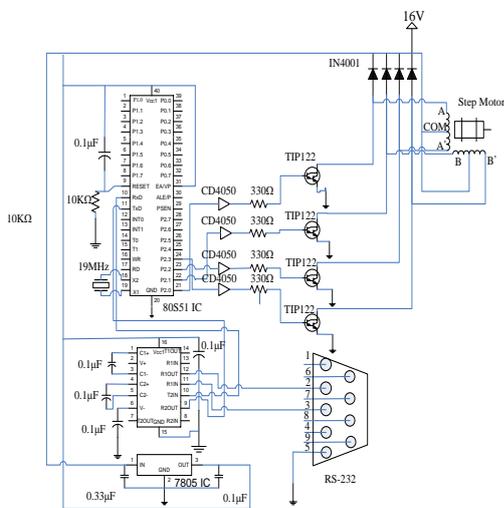


Figure 4: Circuit diagram of data transmission from PC to control unit

### 3. Detect Straight Lines

#### 3.1 Hough Transform

The data of edge points can be obtained by edge detection, and all points in binary image can be projected to the parameter space by using Hough Transform in order to obtain the possible shape in the parameter space [4].

#### 3.2 Line Detection Via Hough Transform

After the image has gone through the processing of edge detection, every edge point will be converted from  $(x, y)$  space to parameter space as  $(\gamma, \theta)$  based on the relation as shown in (1):

$$\gamma = i \cos \theta + j \sin \theta \quad (1)$$

As shown in Figure 5, the  $(i_1, j_1)$ ,  $(i_2, j_2)$  and  $(i_3, j_3)$  are substituted in (1). The slope is represented by red line, and the shortest distance between the slope line and origin is  $\theta$ , which will vary within the range of  $(-90^\circ, 90^\circ)$  along with changing slope of (1). If  $\gamma$  is expressed in absolute value, it represents the shortest distance between the slope of each point and  $(0, 0)$ . The voting on edge points along the straight line will be conducted in Hough Transform as shown in Figure 6. The coordinate of one point  $(i, j)$  is captured in the image for calculating corresponding  $\gamma$  value in the given range of  $\theta$  by using (1) before it is accumulated to the 2D matrix of  $(\gamma, \theta)$ , and such accumulated  $(\gamma, \theta)$  is the result of Hough Transform. The number of intersections of curve is set to be the vote value, and the threshold will be set to determine whether it is a straight line or not.

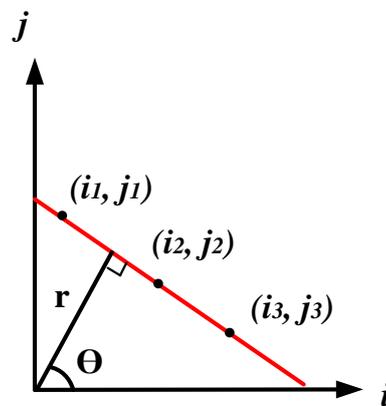


Figure 5: Hough transform

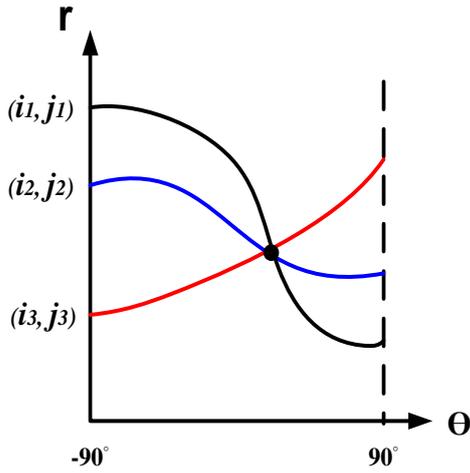


Figure 6: The intersection of two curves is the which is 2. vote value, which is 2.

### 3.3 Detection of Straight Line Intersection

The intersection of two non-parallel Hough straight lines can be calculated by the simultaneous equations as shown in (2):

$$\begin{aligned} \gamma_l &= i \cos \theta_l + j \sin \theta_l \\ \gamma_k &= i \cos \theta_k + j \sin \theta_k \end{aligned} \quad (2)$$

## 4. Ellipse Detection

The stochastic ellipse detection method has been used in this study to detect the ellipse shape, which is faster than the traditional ellipse detection method [11].

### 4.1 Determination of Ellipse Center

An ellipse can be represented by the equation below:

$$u(x - x_c)^2 + v(x - x_c)(y - y_c)^2 + w(y - y_c)^2 = 1 \quad (3)$$

where  $(x_c, y_c)$  represents the ellipse center, and there are three other variables  $u, v,$  and  $w$  meeting the conditions of  $u > 0, w > 0$  and  $4uw - v^2 > 0$ . If the rotational angle of the ellipse is set to be  $\theta$  and the lengths of two axes are set to be  $a$  and  $b$ , respectively, then the five variations in (3)  $(x_c, y_c, u, v, w)$  can be converted to  $(x_c, y_c, \alpha, \beta, \theta)$  while meeting (4), (5) and (6).

$$\theta = \frac{\tan^{-1} \frac{v}{u-w}}{2} \quad (4)$$

$$\alpha = \sqrt{\frac{1}{u \cos^2 \theta + v \sin \theta \cos \theta + w \sin^2 \theta}} \quad (5)$$

$$\beta = \sqrt{\frac{1}{w \cos^2 \theta - v \sin \theta \cos \theta + u \sin^2 \theta}} \quad (6)$$

Assuming  $P_s, t = (x_s, t, y_s, t), s, t = 1, 2, 3, 4, s \neq t$ .

Four edge points will be randomly generated, and two edge points will be selected. If these two edge points are set to be  $P_s$  and  $P_t$ , and assuming the tangent slope of  $P_s$  is not parallel to the tangent slope of  $P_t$ , the two tangents passing through  $P_s$  and  $P_t$  will intersect at  $D_{st} = (d_x, d_y)$ . In addition, the  $E_{st} = (e_{xst}, e_{yst})$  is set to be the midpoint of section  $\overline{P_s P_t}$  as shown in (7), (8), (9), and (10).

$$d_{x_{st}} = \frac{x_s + x_t}{2} \quad (7)$$

$$d_{y_{st}} = \frac{y_s + y_t}{2} \quad (8)$$

$$e_{x_{st}} = \frac{y_s - y_t - s_s x_s + s_t x_t}{s_t - s_s} \quad (9)$$

$$e_{y_{st}} = \frac{s_s s_t (x_t - x_s) - y_t s_s + y_s s_t}{s_t - s_s} \quad (10)$$

There is a very important geometric property between  $D_{st}$  and  $E_{st}$ , and the straight line connecting  $D_{st}$  and  $E_{st}$  will pass through the ellipse center as shown in Figure 7.

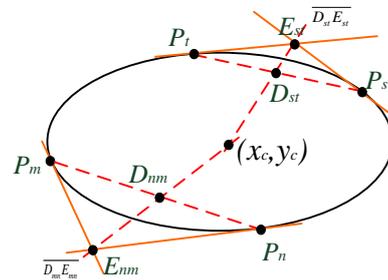


Figure 7: The ellipse center

The simultaneous equations of (11) and (12) have been used to obtain the intersection as the coordinate of ellipse center.

$$(t_{x_{st}} - m_{x_{st}})y - (t_{y_{st}} - m_{y_{st}})x = t_{x_{st}} m_{x_{st}} - m_{x_{st}} t_{y_{st}} \quad (11)$$

$$(d_{x_{mn}} - e_{x_{mn}})y - (d_{y_{mn}} - e_{y_{mn}})x = d_{x_{mn}} e_{x_{mn}} - d_{x_{mn}} e_{y_{mn}} \quad (12)$$

### 4.2 Determination of the Remaining Three Variables

If we assume the ellipse center to be  $(x_c, y_c)$  with the origin shifted to the ellipse center, the ellipse equation (3) can be simplified as:

$$dx^2 + exy + fy^2 = 1 \tag{13}$$

These three variables can be solved by using three out of the selected four edge points. The following linear system as shown in (14) can be obtained by substituting the edge point  $(x'_i, y'_i)$  into previous equation:

$$\begin{bmatrix} (x'_i)^2 & x'_i y'_i & (y'_i)^2 \\ (x'_j)^2 & x'_j y'_j & (y'_j)^2 \\ (x'_k)^2 & x'_k y'_k & (y'_k)^2 \end{bmatrix} \begin{bmatrix} d \\ e \\ f \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \tag{14}$$

It must be noted that  $(x'_i, y'_i)$  is the coordinate of edge point  $(x_i, y_i)$  shifted by  $(x_c, y_c)$ . From the perspective of combination, there are a total of four groups of  $(d, e, f)$  solutions.

### 4.3 Determination of Candidate Ellipse

The four groups of  $(d, e, f)$  solutions will be further investigated to see if they meet  $d > 0, f > 0$  and  $4df - e^2 > 0$ . If so, the corresponding ellipse is a legitimate possible ellipse. If the distance between edge point and actual ellipse is shorter than the threshold value, this possible ellipse is recognized as a candidate ellipse. The calculation formula of the distance between edge point and ellipse is shown as (15):

$$\left| d(x_i - x_c)^2 + e(x_i - x_c)(y_i - y_c) + f(y_i - y_c)^2 - 1 \right| \tag{15}$$

## 5. 3D Vision

The 3D vision is about the research on how to use more than two images with different angles to obtain the 3D structure and depth of field of target object in the scene [12][13][14][15].

## 6. Experimental Procedure

The colors and shapes of objects will be required for the detection and determination of more than 2 kinds of objects. This system is capable of detecting the colors of two kinds of objects, while the shapes detected are rectangular and ellipse.

The system flow chart of this study is shown in Figure 8.

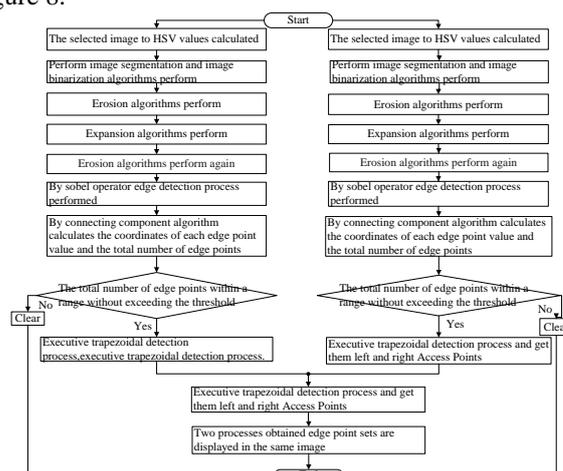


Figure 8: System flow chart

In this system the HSV conversion [1] is used before the image cutting to retain the image color so that the system will only display the color of target object. The binary algorithm will be then implemented to simplify the image and to facilitate the determination of object feature.

The small noise in the image which may affect the determination of object feature can be eliminated by the implementation of an erosion operation [3]. And the split pattern can be put back together into a shape after dilation operation [3]. The oversized object after dilation operation can be restored to the normal size after another implementation of an erosion operation. The purpose of these three process flows is to put the split patterns in the image back together to form a geometric shape while eliminating a lot of unnecessary noises.

The system computing load can be greatly reduced by implementing edge detection [2]. The connected component algorithm [2] is used to calculate the total number of edge points of every complete figure in the image with connectivity. The sets of edge points with too large or too small difference from detected object will be eliminated, so that this step can eliminate a huge amount of unnecessary computing load of the system and thus speed up the object detection. After the completion of trapezoid detection and ellipse detection, the remaining sets of edge points will be displayed on the same screen together with the left access point and right access point of the detected object.

### 6.1 Quadrilateral Detection

The quadrilaterals detected by this system include trapezoid, rectangular, and parallelogram. And the area of certain quadrilaterals will be used for the setting of threshold value in order to reduce the number of misjudgments due to irregular shapes.

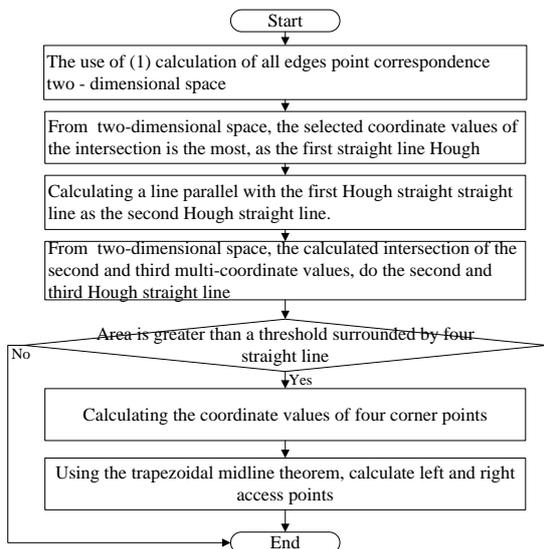


Figure 9: Flow chart of trapezoid detection

In Figure 8, where the procedure of trapezoid detection is shown in Figure 9. The algorithm of this trapezoid detection can only be implemented after the image is processed by edge detection. And then the connected component algorithm [2] will be implemented to substitute the edge point coordinates into (1) and to generate a 2D space. After calculation, if the edge point coordinates on the image plane are identical to those in the corresponding 2D space, the vote value will be increased because they are in the same straight line on image plane. The points with the highest number of votes will form the first Hough straight line. To ensure the quadrilateral is a trapezoid, the second Hough straight line must be parallel to the first Hough straight line. After all 4 Hough straight lines have been confirmed, the enclosed area should be calculated to see if it is larger than threshold value. Otherwise, it will not be regarded as a quadrilateral because the feature of general quadrilateral is the larger area.

### 6.2 Ellipse Detection

This system is also capable of detecting circles and ellipses. If the length of short axis of an ellipse is the same as the long axis, it is a circle.

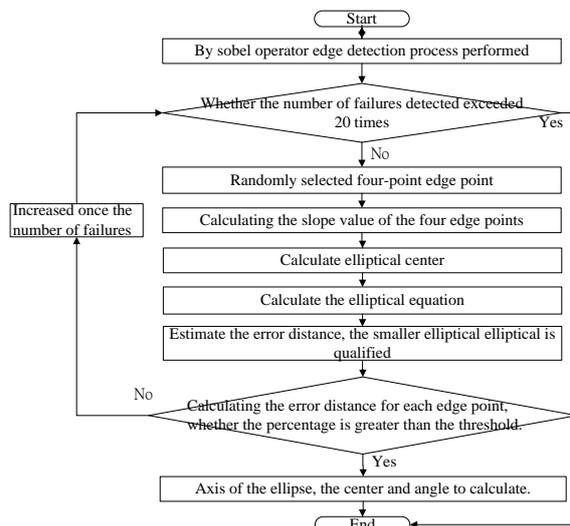


Figure 10: Flow chart of ellipse detection

In Figure 8, where the process flow of ellipse detection is shown in Figure 10. After image edge detection is completed, the trapezoid detection can be implemented. The connected component algorithm [2] can be used to obtain a binary image with connectivity, the pixel coordinate of every edge point, and the total edge points of that particular image. After knowing the locations of pixels, the coordinates of 4 edge points will be randomly selected, and the slopes of coordinates of these 4 points will be calculated. The intersection of these calculated slopes can be calculated by the simultaneous formula of (11) and (12) to obtain 3 sets of ellipse centers. (14) is the elliptic equation based on ellipse center. It should be noted that (14) must be modified into (16) to avoid the resulting figures to be too large for the system to calculate.

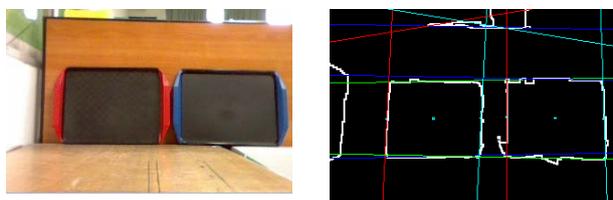
$$\begin{bmatrix} (x'_i)^2 / 100 & x'_i y'_i / 100 & (y'_i)^2 / 100 \\ (x'_j)^2 / 100 & x'_j y'_j / 100 & (y'_j)^2 / 100 \\ (x'_k)^2 / 100 & x'_k y'_k / 100 & (y'_k)^2 / 100 \end{bmatrix} \begin{bmatrix} d \\ e \\ f \end{bmatrix} = \begin{bmatrix} 1/100 \\ 1/100 \\ 1/100 \end{bmatrix} \quad (16)$$

(15) can be used for error estimation of distance between the estimated elliptic equation and the actual image in this study. A threshold value must be set up here, and when the error is less than the threshold value, it can be confirmed to be an ellipse, with the pixel as the distance unit. As shown in the figure, the number of allowed failures is 20. If the number of allowed failures is too high, the system operation will be rather slow. If the number of allowed failures is too low, the system may not be able to detect the object. Therefore, 20 is the most proper number.

## 7. Experimental Results

The image coordinates of the detected object will be converted into 3D coordinates.

### 7.1 Tray Detection



(a) (b)

**Figure 11: Situations of detection of two trays.** (a) Situation of detection by the camera on the left; (b) Image processing of camera on the left

Situations of tray detection are shown in Figures 11 (a)(b). The tray with red handles is called Tray A, and the tray with blue handles is called Tray B. The 3D coordinates of Tray A and Tray B are as shown in Table 1 and Table 2, respectively:

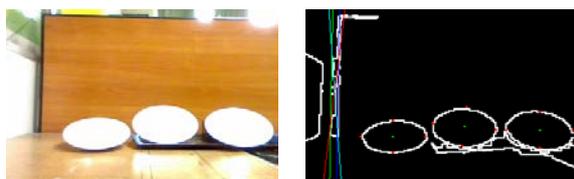
**Table 1: The 3D coordinates of Tray A**

Unit (mm)	Left access point (estimated distance)	Right access point (estimated distance)	Left access point (actual distance)	Right access point (actual distance)
X	81	405	85	445
Y	59	70	-30	-30
Z	1127	1100	1200	1200

**Table 2: The 3D coordinates of Tray B**

Unit (mm)	Left access point (estimated distance)	Right access point (estimated distance)	Left access point (actual distance)	Right access point (actual distance)
X	-312	19	-342	18
Y	65	67	-30	-30
Z	1140	1184	1200	1200

### 7.2 Saucer Detectio



(a) (b)

**Figure 12: Situations of detection of three saucers.** (a) Situation of detection by camera on the left; (b) Image processing of camera on the left

The situations of detection of three saucers are shown in Figures 12 (a) and (b). The saucer on the left is called saucer C, the saucer in the middle is called saucer D, and the saucer on the right is called saucer E. Their actual 3D coordinates are as shown in Table 3, and the estimated coordinates in this study are shown in Table 4:

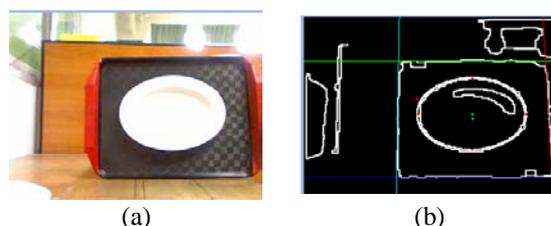
**Table 3: Actual 3D coordinates of three saucers**

Unit (mm)		X	Y	Z
Saucer C	Left access point	-212	75	870
	Right access point	-30	75	870
Saucer D	Left access point	-20	60	880
	Right access point	177	60	880
Saucer E	Left access point	190	50	870
	Right access point	373	50	870

**Table 4: Estimated 3D coordinates of three saucers**

Unit (mm)		X	Y	Z
Saucer C	Left access point	-208	136	806
	Right access point	-44	152	899
Saucer D	Left access point	-22	110	882
	Right access point	151	122	850
Saucer E	Left access point	182	127	866
	Right access point	342	118	806

### 7.3 Tray and Saucer Detectio



(a) (b)

**Figure 13: Situation of tray overlapping with saucer.** (a) Situation of detection by camera on the left; (b) Image processing of camera on the left

The situation of tray overlapping saucer is shown in Figures 13 (a) and (b). The actual and estimated 3D coordinates of tray and saucer are shown in Table 5 and Table 6, respectively.

**Table 5: Actual 3D coordinates of tray overlapping with saucer**

Unit (mm)	Saucer		Tray	
	Left access point	Right access point	Left access point	Right access point
X	-37	213	-87	273
Y	-40	-40	-30	-30
Z	670	670	680	680

**Table 6: Estimated 3D coordinates of tray overlapping with saucer**

Unit (mm)	Saucer		Tray	
	Left access point	Right access point	Left access point	Right access point
X	-34	200	-80	263
Y	12	12	23	24
Z	649	632	654	668

As shown in Table 5 and Table 6, the statistic error is rather significant. This experiment is for proving the capability of this system to detect multiple saucers and trays, where the 3D coordinates of multiple saucers can be obtained in coordination with binocular vision geometry. The longer the distance between the camera and the object, the greater the error in 3D coordinates. A proper distance between the camera and the object will lead to higher accuracy of 3D coordinates.

As shown in Table 6, with shorter distance between the camera and the object, the errors in the actual and estimated 3D coordinates have been significantly improved. The system used in this study can still detect the tray overlapping with the saucer. This is an indication of the capability of this system for detecting ellipses and quadrilaterals.

## 8. Conclusion

A tray may appear in the shapes of trapezoid, square, parallelogram, circle, and ellipse in every position in the image, and the probability of appearing as trapezoid and ellipse is the highest so that there must be an algorithm to be developed for solving this problem.

The image processing software used in this paper is compiled by Borland C++ Builder 6 (BCB6), and is capable of detecting trapezoid objects and elliptic objects of most angles. Even if the saucer is overlapping with the tray, this system is still capable of detecting these two objects so as to meet the requirements of service robots. After obtaining the image coordinates, the 3D coordinates of that particular object can be obtained by using 3D reconstruction theory. This coordinate information can then be provided to the control system of mechanical arms in order to accomplish the action of gripping the target object. If the saucers and trays are specially designed, the system capability of identifying trays and saucers can be greatly improved.

## Acknowledgment

This work was supported by the National Science Council of Taiwan, R.O.C. under grant NSC 101- 2221-E-167-020 -MY2.

## References

- [1]. Y. C. Luo, "Extracting the Foreground Subject in the HSV Color Space and Its Application to Human Activity Recognition System," National Chiao Tung University, Hsinchu, Taiwan, July 2007.
- [2]. Y. Z. Lin, "Binocular Vision Based Human Head Tracking," National Taiwan University of Science and Technology, Taipei, Taiwan, June 2008.
- [3]. Z. C. Huang, "Design and Implementation of Monocular Vision-Based Ball-Batting Robots," National Chiao Tung University, Hsinchu, Taiwan, September 2010.
- [4]. R. O. Duda and R. E. Hart, "Use of the Hough Transform to Detect Lines and Curves in Pictures," *Comm. ACM*, Vol. 15, No. 1, pp. 11 - 15, January 1972.
- [5]. L. Rui and L. Qiong, "Image Sharpening Algorithm Based on a Variety of Interpolation Methods," 2012 International Conference on Image Analysis and Signal Processing (IASP), pp.1 - 4, November 9-11, 2012.
- [6]. L. Yaping, Z. Jinfang, X. Fanjiang and S. Xv, "The Recognition and Enhancement of Traffic Sign for the Computer-Generated Image," 2011 11th International Conference on Control, Automation and Systems (ICCAS), pp.750 - 753, October 26-29, 2011.
- [7]. D. Jiang and J. Yi, "Comparison and Study of Classic Feature Point Detection Algorithm," 2012 International Conference on Computer Science & Service System (CSSS), pp. 2307 - 2309, August 11-13, 2012.
- [8]. D. Xia, F. Dong, G. Yao and Yong Liu, "An Edge-Detection Algorithm Based on Improvement Average Filter for Noise Image," World Automation Congress (WAC), pp. 31 - 34, June 24-28, 2012.
- [9]. H. Huang, H. Wang, F. Guo and J. Zhang, "A Gray-Scale Image Edge Detection Algorithm Based on Mathematical Morphology," 2011 Third International Conference on Measuring Technology and Mechatronics Automation (ICMTMA), Vol. 1, pp. 62-65, January 6-7, 2011.

- [10]. G. Wei, L. Xiaobo, Y. Li, C. Wujun and L. Yang, "Detection Algorithm of Video Image Distance Based on Rectangular Pattern," 2012 5th International Congress on Image and Signal Processing (CISP), pp. 856 - 860, October 16-18, 2012.
- [11]. K. L. Chung and Y. H. Huang, "A Pruning-and-voting Strategy to Speed up the Detection for Lines, Circles, and Ellipse," Journal of Information Science and Engineering, 24(2), pp. 503-520, 2008.
- [12]. N. K. William and A. C. Bovik, "FOVEA: A Foveated Vergent Active Stereo Vision System for Dynamic Three-Dimensional Scene Recovery," IEEE Transactions on Robotics and Automation, Vol. 14, No.5, October 1998.
- [13]. R. Klette and A. Koschan, "Computer Vision, Three-Dimensional Data from Images," Springer, Singapore, 1998.
- [14]. G. S. Huang and Y. L. Shen, "Development of a Real-time Two-Eye Vision System," Proceedings of 2009 CACS Automatic Control Conference, SaA04, 00370, pp.1-6, Taipei, Taiwan, November 2009.
- [15]. Y. T. Ho, "Three-Dimensional Target Trajectory Estimation and Interception by Visual Servo Technology," National Cheng Kung University, Tainan, Taiwan, July 2002.



Wen-Lang Zhang was born in Kaohsiung, Taiwan. He received the B.S. degree in electrical engineering from Kao Yuan University of Technology, Kaohsiung, Taiwan, in 2009. He is currently a graduate student in the Department of Electronic Engineering of National Chin-Yi University of Technology. His research interests include image processing, and software development.



Guo-Shing Huang (M'02 -SM' 06) was born in Tainan, Taiwan, R. O. C. in 1957. He received the B. S. and the M. S. degree in electrical engineering and automatic control engineering from the Feng Chia University in Jun. 1980 and Jan. 1983, respectively, and earned the Ph. D. in Department of Electrical Engineering, National Cheng Kung University from Sep. 1994 to Dec. 1998. He is currently a Distinguished Professor in the Department of Electronic Engineering of National Chin-Yi University of Technology. His research interests include control application, integrated GPS/INS electronic navigation, fuzzy/robust control, intelligent robotic location navigation and control.