ISSN 2616-5775 Vol. 4, Issue 6: 72-77, DOI: 10.25236/AJCIS.2021.040612

Research on robot catching ball based on binocular vision positioning

Yingyue Xing[1], Zhijie Shao[2]

[1]School of Mechanical Engineering and Automation, Harbin Institute of Technology (Shenzhen), Shenzhen, Guangdong, China
[2]Dalian-Leicester Institute, Dalian University of Technology, Dalian, Liaoning, China

Abstract: *In this paper, the robot used to catch the ball is studied, and a faster and more accurate detection and recognition method for high-speed moving objects is proposed, which can automatically judge the position of the ball and output the position information through TCP communication protocol. Based on the binocular vision system, this method identifies the small ball by calibrating the camera, processing the gray-scale image, transforming the image coordinates to spatial coordinates and Hough transform, and then realizes the spatial position and distance detection of the small ball. In addition, the ball receiving system and the robot platform are studied, so that they can be connected in real time and stably. It ensures the real-time and accuracy of the vision system to track the ball and meets the design requirements of the ball-catching robot.*

Keywords: *Binocular vision, high-speed target positioning, robot ball receiving, communication protocol*

1. Introduction

Machine vision refers to the use of cameras, computers and other machines instead of human eyes to identify, track, measure and judge the target, and further graphics processing, so that it is more suitable for human eyes to observe or transmit to instruments for detection. Machine vision is mainly used to replace artificial vision and expand the function of artificial vision to complete the tasks of measurement and detection, recognition and classification, guidance and positioning efficiently, quickly and accurately [1, 2].

Moving target location is a technology that organically combines data association, image processing and other technologies, and can detect and track the target position in real time from the image sequence frame. In the case of high-speed target movement, how to locate and track the moving target is a hot topic in recent years [3].

At present, developed countries have developed corresponding machine vision software and hardware products for industrial field applications. In 2004, the University of reading in the UK and INRIA in France jointly completed the development of airport intelligent monitoring system AVITRAC [4], which can automatically track moving targets on the airport. A considerable number of scientific research institutes and universities in China are engaged in the research of moving target tracking and recognition. In 2013, Li Zhiqi of Harbin University of technology and others developed a small ball robot system with 7 degrees of freedom and 2 cameras, used the hitting strategy based on support vector regression to learn the motion trajectory of the racket when hitting the table tennis ball [5]. However, the stability and accuracy of small ball detection are still poor.

In the table tennis ball movement, the size of the table-tennis ball is small and the flight speed is fast. When designing a robot, it is necessary to solve the problems of reliable tracking and positioning, trajectory prediction and so on based on the characteristics of ball movement. Its vision system is similar to human's eyes, that is, binocular vision system, which requires the vision system to accurately measure the absolute position of the target in three-dimensional space on a moving platform. This problem involves the accurate visual positioning of high-speed moving objects, has a very broad application prospect, and the research is of great significance.

ISSN 2616-5775 Vol. 4, Issue 6: 72-77, DOI: 10.25236/AJCIS.2021.040612

2. Visual positioning and catch method

Binocular vision system can observe large-scale workspace, so as to realize visual detection of high-speed table tennis ball and accurately obtain the image coordinates of the target center. This requires the transformation from image coordinates to spatial coordinates, so as to locate table tennis in three-dimensional Cartesian space. After obtaining the position information, the data transmission between the manipulator and the computer is realized through TCP protocol, and then the robot can catch the ball accurately.

2.1. Binocular vision principle

The depth camera of binocular stereo vision does not actively project the light source to the outside, and completely depends on the taken pictures to calculate the depth. Its principle is as follows[6].

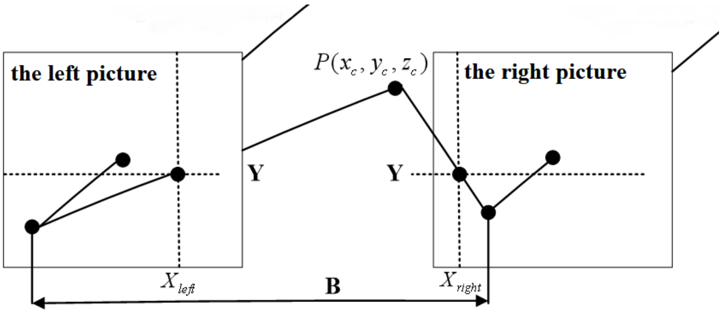

Figure 1: Principle of binocular stereo imaging

Where, the distance of the baselines (B) is the distance between the projection centers of the two cameras; The camera focal length is f.

Two cameras are set to view the same feature point $P(x_c, y_c, z_c)$ of space object at the same time, their image coordinates are $p_{left} = (X_{left}, Y_{left})$, $p_{right} = (X_{right}, Y_{right})$. From the geometric relationship, we can get that:

$$\begin{cases} X_{left} = f\frac{x_c}{z_c} \\ X_{right} = f\frac{(x_c - B)}{z_c} \\ Y = f\frac{y_c}{z_c} \end{cases} \tag{1}$$

Then the parallax is: $Disparity = X_{left} - X_{right}$. It can be calculated that the three-dimensional coordinates of feature point P in the camera coordinate system are:

$$\begin{cases} x_c = \frac{B \cdot X_{left}}{Disparity} \\ y_c = \frac{B \cdot Y}{Disparity} \\ z_c = \frac{B \cdot f}{Disparity} \end{cases} \tag{2}$$

Therefore, the three-dimensional coordinates of the point can be determined by finding the corresponding matching point on the camera image plane.

2.2. Identification and location

Hough transform is used to identify small balls. Put the circle $(x-a)^2 + (y-b)^2 = r^2$ into the three-dimensional coordinate system with a, B and R as the coordinate axis, and the corresponding point A (a, B, R) in the figure is the center of the circle. If the radius is assumed to be the same as the actually detected circle radius, the parametric coordinate system curve (circle) corresponding to several points will intersect at a point a, which is the center of the circle with the corresponding radius of R [7].

ISSN 2616-5775 Vol. 4, Issue 6: 72-77, DOI: 10.25236/AJCIS.2021.040612

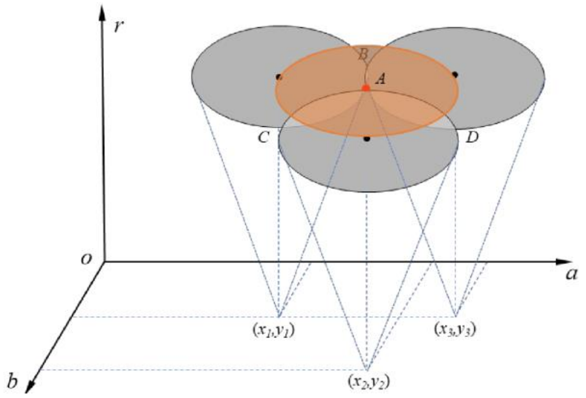

Figure 2: Principle of Hough transform to identify small ball

Define four coordinate systems, the world coordinate system $P_W(X_W, Y_W, Z_W)$, camera coordinate system $P_C(X_C, Y_C, Z_C)$, image coordinate system $P_I(X_I, Y_I)$, pixel coordinate system P (u, v), and the relationship between the four is shown in Figure 3.

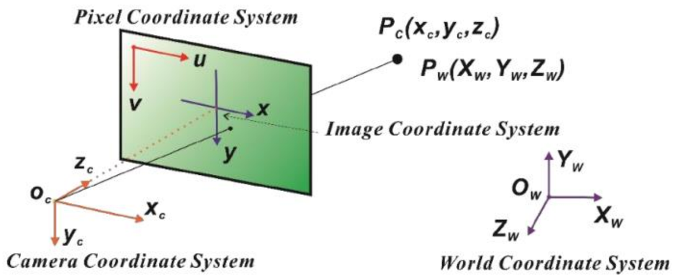

Figure 3: Coordinate transformation of visual system

After obtaining the image coordinates of the target point P in the left and right cameras, let the homogeneous coordinates on the normalized image plane of the left and right cameras calculated by the internal parameters[8] of the camera be $(x_1, y_1, 1)^T, (x_2, y_2, 1)^T$, the corresponding camera external parameters are: $(R_1, t_1)^T = ({m_{11}}^T, {m_{21}}^T, {m_{31}}^T), (R_2, t_2)^T = ({m_{12}}^T, {m_{22}}^T, {m_{32}}^T)$, substitute the following equations[9]:

$$\begin{cases} (m_{11} - x_1 \cdot m_{31}) \cdot X_w = 0 \\ (m_{21} - x_1 \cdot m_{31}) \cdot X_w = 0 \\ (m_{12} - x_2 \cdot m_{32}) \cdot X_w = 0 \\ (m_{22} - x_2 \cdot m_{32}) \cdot X_w = 0 \end{cases} \tag{3}$$

The optimal three-dimensional coordinates X_w of the ball in the world coordinate system can be obtained, realize the three-dimensional positioning of the ball.

2.3. Manipulator interface communication protocol

The data transmission between the manipulator and the computer requires TCP protocol[10]. The manipulator is the customer service end and the controller is the service end. Through a specific programming port, the client establishes a TCP/IP connection with the controller, writes a script program on the manipulator and sends it directly to the robot controller. After the program runs, the coordinates of the ball are obtained, and then transmitted to the manipulator through TCP protocol, and the manipulator is controlled to move and catch the ball.

The communication protocol is shown in Figure 4. Among them, instruction confirmation (cmdcfm): after receiving and processing the instruction, the underlying device assigns this byte to feed back the instruction execution.

Head	DeviceAdd	Cmd	cmdcfm	Len_1	Len_2	data	Check
1	1	1	1	1	1	n	1
0x5A	0xFF	/	/	/		/	/

Figure 4: Communication Protocol

ISSN 2616-5775 Vol. 4, Issue 6: 72-77, DOI: 10.25236/AJCIS.2021.040612

3. Experiment

3.1. Basic process

In this experiment, the binocular camera principle is used to identify and locate the small ball. After the camera calibration, a series of processing such as grayscale and spatial coordinate transformation are carried out for the obtained image, then the small ball is identified by Hough transform for the obtained image, and finally the position of the small ball is determined by the derived distance calculation formula. The basic flow diagram is shown in Figure 5.

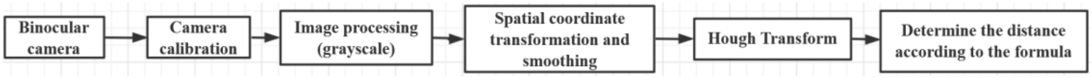

Figure 5: Basic flow chart of binocular recognition ball experiment

3.2. Experimental process

After printing the calibration board, take photos from different angles, as shown in Figure 6. Use matlab to calibrate and export the calibration data.

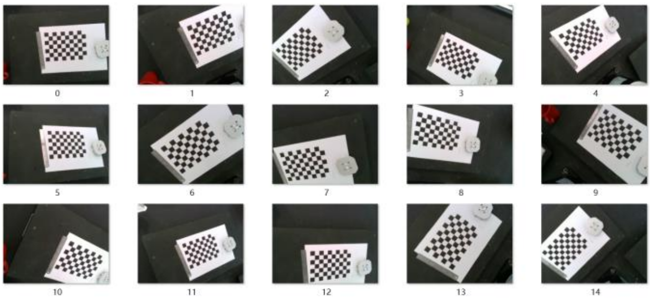

Figure 6: Photo of calibration plate

The maximum value method is used for image graying. The color intensity of three channels of color image is compared to obtain the maximum value, and then assigned to the graying parameter, i.e. $f_{(i,j)} = max(R_{(i,j)}, G_{(i,j)}, B_{(i,j)})$. The comparison between the original drawing and the effect drawing is shown in Fig.7-8.

Figure 7: Original picture

Figure 8: Processed picture

Use Hough transform to identify the sphere, and the results are shown in Fig. 9-10.

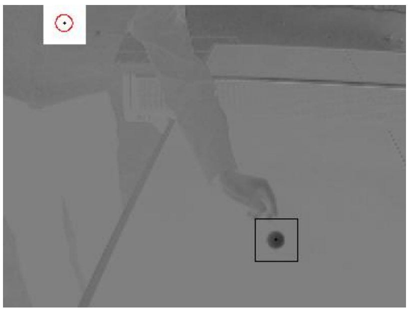

Figure 9: Result of sphere recognition

Figure 10: Result of sphere recognition

ISSN 2616-5775 Vol. 4, Issue 6: 72-77, DOI: 10.25236/AJCIS.2021.040612

Calculate the distance,

$$Z = \frac{fT}{X_1 - X_r} \tag{4}$$

Where f is the focal length, T is the distance between two cameras and the denominator $X_1 - X_r$ is the pixel difference between the two pictures. After correction, if the pixel coordinates of the spherical center in the first picture are (U, V1) and the pixel coordinates in the second picture are (U, V2), the pixel difference is | V1-V2|.The pictures before and after calculating the distance are shown in Figure 10-11.

Figure 11: Picture before processing

Figure 12: Picture after processing

4. Result analysis

When the camera works in high-speed acquisition mode, the image resolution is 656×490 pixels and the camera observation area is 1400×1400mm. The specific data comparison results are shown in Table 1.

Table 1: Comparison of position measurement errors (unit: mm)

No.	X_1	Y_1	Z_Z	X_2	Y_2	Z_2	Error
1	437.5	1828	494	438.7	1826.8	492.8	2.4658
2	437.5	1528	494	438.4	1528.2	492.4	1.8805
3	-162.5	2053	494	-162.8	2054.3	494.5	3.6463
4	-162.5	1678	494	-162.5	1677.4	492.7	3.3671
5	-12.5	1678	494	-11.4	1676.05	493.7	2.7274
6	-12.5	1678	494	-11.1	1676.925	492.975	2.6180
7	-462.5	1828	494	-460.6	1825.65	493.55	3.3554
8	-462.5	1528	494	-461.5	1525.3	494.6	2.9891

Where, (X_1, Y_1, Z_1) are the actually measured three-dimensional coordinates of the object, (X_2, Y_2, Z_2) are the three-dimensional coordinates calculated by this method, and Error is the error value of the two, both in mm.

It can be seen from the above table that theoretically, the average value of three-dimensional measurement error caused by the error of one pixel in the picture coordinates is 2.8812 mm, which has fully met the requirements of table tennis position measurement and trajectory prediction.

The above experiments show that this method can accurately identify the position and distance of the small ball and realize the requirements of small ball positioning and recognition. In addition to the above recognition process of the small ball flight trajectory by the vision system, the time period is about 6.5ms, which meets the real-time requirements of the small ball target recognition algorithm, i.e. > 60fps.

5. Achievements and Prospects

5.1. Achievements

In this paper, a small ball recognition and positioning method based on binocular vision is proposed to extract the position information of moving targets quickly and accurately.

Firstly, the camera model is established, the transformation relationship between each coordinate system in the camera model is analyzed, and Zhang Zhengyou camera calibration method is used for calibration. Secondly, after obtaining the real image of the small ball, the image is grayed. By analyzing

ISSN 2616-5775 Vol. 4, Issue 6: 72-77, DOI: 10.25236/AJCIS.2021.040612

the image feature detection technology of Hough transform, the small ball matching and positioning are realized.

Experiments show that compared with the existing regional target search method, this method can accurately and robustly extract the ball center, and meet the requirements of the rapidity and accuracy of the vision system when the ball robot receives the ball. In addition, this experiment also studies the communication protocol of robot ball receiving. Compared with the original robot ball receiving device, the docking compatibility between the vision system and the manipulator platform is stronger and more stable.

5.2. Expectation

The ball receiving method can be further tested and debugged, and other methods can be added or selected in the process of image recognition and processing, so as to optimize the process and speed up the identification of small balls.

In addition, the analysis method of this study can be used for reference to analyze the characteristics of different ball sizes, surface characteristics and application environment. It can be adjusted in the steps of image preprocessing and matching. Finally, its effectiveness is verified by experiments.

References

[1] Lu Rongsheng, Shi Yanqiong, Hu Haibing. Overview of robot vision 3D imaging technology [J]. Progress in laser and optoelectronics, 2020, 57 (4)

[2] Huang Zhuo. Research on real-time detection and recognition algorithm of UAV moving target [D]. Harbin Institute of technology

[3] Hou Zhiqiang, Han Chongzhao. Overview of visual tracking technology. Journal of automation, 2006, 32 (4): 603-617

[4] Blauensteiner,P. and Kampel,M. Visual surveillance of an airports apron-anoverview of the AVITRACK project [R]. Digital imaging in media and education, annual workshop of AAPR, 2004.

[5] Wang Q, Zhang K, Wang D. The trajectory prediction and analysis of spinning ball for a table tennis robot application [C]. //IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems. IEEE, 2014:496-501.

[6] Luo Dan, Liao Zhixian. Binocular stereo vision ranging based on OpenCV [J]. Popular technology, 2011 (4): 52-71

[7] Cui Jiwen, Tan Jiubin. Fast circle contour detection technology based on constrained sampling Hough transform [J]. Journal of Harbin University of technology, 2005, 37 (10): 1394-1396

[8] Qiu Maolin, Ma Songde, Li Yi. Overview of camera calibration in computer vision. Journal of automation, 2000, 26 (1): 43-55

[9] Zhang Zhengtao, Xu de. Research on high-speed vision system based on intelligent camera and its target tracking algorithm. Robot, May, 2009, 31 (3): 229-234

[10] Xie Xiren. Computer network (Seventh Edition) [M]. Beijing: Electronic Industry Press, 2013.238-240