Fast Ball Detection Method for Ping-Pong Playing Robots

Jing Sun, Jiandong Tian, Yingkui Du
1. State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences
   Shenyang, 110016, China
2. Graduate University of the Chinese Academy of Sciences,
   Beijing, 100049, China

Yandong Tang
State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences
Shenyang, 110016, China
ytang@sia.cn

Abstract—Ping-pong ball detection plays an important role in the vision system of ping-pong playing robots. In the paper, we present a fast ball detection method for ping-pong playing robot. A reasonable searching region is proposed based on the ping-pong table and the prediction of the ball’s current position. Five key points are used to detect the ball in our method to meet the ping-pong ball’s high speed. Because we use points to the ping-pong ball instead of the ball’s region, the process of the detection can be finished during one scanning, which greatly increases the detecting speed. Experimental results verify the feasibility and the efficiency of our method.

Keywords-Image Processing; Ping-pong playing robot; Target Detection

I. INTRODUCTION

Ping Pong, interesting and simple, attracts all types of players-recreational to competitive. However, it is not a simple game for robots. The research of Ping-Pong playing robots is more than 20 years, which can be traced back to twenty years ago when Russell Andersson [1], a member of technical staff of the Robotics Systems Research Department at AT&T Bell Laboratories, constructed a system to play Ping-Pong using a real-time vision system. Since then, the researching on ping-pong playing robots has started in many countries. The development of the robots in some countries, such as Japan and USA, has reached a certain stage. However, China with great achievement in the ping-pong sport is still less of research on the subject of ping-pong playing robots.

For the ping-pong playing robots, vision is an unavoidable component and provides them with a detailed three dimensional knowledge of the environment. The task of the vision system is to estimate the ball position at each instant and to predict the future trajectory for the ball. Information obtained from the vision system of ping-pong playing robots is used to control the robot’s motion in real time. The development of vision systems for ping pong playing robots is tabulated in Table 1.

In the system, the ball detection in 2D camera images plays an important role, which provides the necessary data for further fast ball trajectory. Though the system developed by Andersson [1] locates the ball’s position accurately by four cameras, cameras’ calibrations in vision system are extremely difficult and involve complex mathematical calculations. Furthermore, to synchronizing the four cameras takes time. The playing robot in the Miyazaki Laboratory utilizes Quick-MAG which was a 3-D real-time motion analyzing system to recognize the 3-dimensional position through color tracking. However, the color may become trustless when the surrounding illumination changes [2].

| Developer(s)          | Vision element                  | Spatial resoln. | Time resoln. |
|-----------------------|----------------------------------|-----------------|--------------|
| Andersson, 1980 [1]   | Four cameras                     | 756*242         | 60 hertz     |
| Hashimoto, Asano, 1987 [3] | Binocular camera with four sensors | 2048*2048      | 100 hertz    |
| Fassler, Zurich, 1990 [4] | Two CCD cameras                  | 422*579         | 50 hertz     |
| Naghdly, Tran, 1993 [5] | Two cameras with frame grabber   | 1024*1024       | Around 60 hertz |
| Miyazaki, Kusano, 2002 [6] | Quick MAG III                   | 640*416         | 60 hertz     |

In tracking systems, locating the ball’s positions in images is the pre-acquisition for further analysis. Ball detection methods have been studied widely [7]. In Yow et al. [8], the ball is detected by template matching in each of the reference frames and then tracked between each pair of these reference frames. Tong et al. [9] employed indirect ball detection by eliminating non-ball regions using color and shape constraints. In Yamada et al. [10], white regions are taken as ball candidates after removing players and field lines. D’Orazio et al. [11] detected the ball using a modified Hough transform along with a neural classifier. The methods mentioned above either are too slow to meet the need of the ping-pong playing robot or require special precondition to process. To meet the special requirements of ping-pong ball detecting on accuracy, rapidity and precision, we propose a one-step scanning detecting method to recognize high-speed balls. The current method is for two external visual cameras in our system. The method includes three steps: Ball’s searching regions in the two images
gotten from the two cameras simultaneously are restricted to improve the accuracy of the detection; ping-pong balls are detected by five-point method rapidly; a recalculation step is adopted for the precision in locating the center of the ball.

The rest of the paper is organized as follows. Section 2 introduces the searching region calculating method; fast ball detecting method is illustrated in section 3, followed by experiment results in section 4. Section 5 is the conclusion.

II. BALL’S SEARCHING REGION

The detecting results of a ping-pong ball are easily influenced by the interference in the surrounding environment, such as the moving of persons and the swinging of arms. To simplify the background, it is necessary to restrict the searching region to a reasonable region in the captured image. Furthermore, restricting the searching region will shorten the processing time.

The ping-pong rules ensure that the ball is above the table most of the time in the rally. Enough information can be analysis the ball when it is above the table. Mapping into the captured image, most of the ball’s positions are at the middle parts of the image where the ping-pong table located. To limit the detection region in the image, we should firstly extract the table region.

The traditional thought to extract the table is to detect lines directly from the images, which bring some problems, such as false detection and time cost. We calculate it according to the relationship between the known scene point $P$ and its corresponding 2D image points $p$ instead.

Assuming the two cameras are accurately calibrated, the projection of a world point $P$ through the camera model is

$$p = MP = K(R|t)P$$

where, $K$ is the matrix of intrinsic parameters, $(R|t)$ is the matrix of extrinsic parameters.

We establish the world coordinate taking the point of intersection of the center line and the line close to the cameras as the origin, as shown in Fig. 1. Because the four corners of the table are fixed in relation to the intersection, they are measurable in our world coordinate. When the distance between the corner of the table and the origin of our the world coordinate is given, the related points in the left and right images of the four corners of the table can be calculated by function (1).

Figure 1. The origin of our established world coordinate

Because two points determine a straight line, we get the two side lines in both images. Fig.2. shows the example of the table lines in the left image gotten by our method. The searching region is restricted to the part between the two lines in each image.

Figure 2. The results of extracting the table region by our method. The upper: the original image; the below: The results of extracting the table region.

The searching region is decided by the intersection part of the prediction of current position and table region. The prediction algorithm refers to [12].

III. FAST BALL DETECTING METHOD.

The vision system must operate at high speed to acquire enough data for trajectory analysis. To meet the high speed of the ping pong ball, all the time resolutions in current systems are more than 60 hertz, that is, the data cycle should at least be less than 20mm. The operations for one data cycle basically includes: the caught of the images, ball
detection in both images simultaneously, the acquisition of the centre of the ping pong ball, the calculation of position real ball, and the prediction of the future trajectory for the ball. To save the detection time, we use five points to represent the ball and detect it in one one-step scanning.

A. Five Key Points in the method

Because we have reduced the detecting area in the image, the background is relatively stable. The method is processed on the difference of adjacent frames, as shown in Fig.3.

![Figure 3. the difference of adjacent frames](image)

The five key points in the method are used to describe a ping-pong ball in the difference image, as shown in Fig.4.

![Figure 4. Five key points are used to describe a ping-pong ball](image)

The five points are distributed at the center and around the boundary of the ball in the images, so the round feature of the ball can be sure by the five key points. The five points are chosen because they can represent a ball in the researching region of the binary image gotten by our method where there are few disturbances and because they are easy to get during one image scanning.

Supposing that \( I(i, j) \) is the binary image gotten by the difference of adjacent frames and \( m \times n \) is the size of the image, we simply set the threshold \( T \) as follows,

\[
T = \frac{1}{2} \times \frac{1}{m \times n} \sum_{i=1}^{m} \sum_{j=1}^{n} I(i, j)
\]  

(2)

When a pixel value is higher than the threshold, we take it as a possible upper left boundary point of the ball and then judge if the other four points comply with the rules one by one. If all the values of the five points are higher than the threshold, the region surrounded by the five points is the ball in the image. The current possible position of the ball, which is calculated according to the prediction algorithm of the future trajectory\([7]\), decides the size of the ball in the image. The current size of the ball is calculated according to the predicted position. The distance between the five points is decided accordingly.

B. The center of the detected ball

The five-key-point method can only find the possible center \((u, v)\) of the ball. Because the possible region is detected, we can use more complex method to locate the accurate center of the ball on small regions. The method of elliptic object detection based on the C-V model and a shape restraint Main title\([8]\) are applied to locate the center.

Based on the shape feature, an elliptic shape restraint is imposed on the zero level set of a Lipschitz function in C-V model. The C-V model with an elliptic shape restraint can be written as follows:

\[
\inf\{E[c_1, c_2, \phi] | u_0 \} = \\
\alpha \int_{\Omega} \left( (u_0 - c_1)^2 H(\phi) + (1 - \alpha) \int_{\Omega} (u_0 - c_2)^2 (1 - H(\phi)) \right) \]  

(3)

Subject to:

\[
\phi = 1 - \frac{((x - x_0) \cos \theta + (y - y_0) \sin \theta)^2}{a^2} + \frac{(- (x - x_0) \sin \theta + (y - y_0) \cos \theta)^2}{b^2} \]  

(4)

Where \( \alpha > 0 \) are fixed parameters and \( H(\cdot) \) is the Heaviside function. Here, \( x_0, y_0, \theta, a, b \) are the parameters of the ellipse: \( \phi = 0 \).

After the round shape of the ball is detected by the method, the center of the ball is located accurately.

C. The flow of the method

The proposed fast one-step scanning detecting algorithm mainly involves the following steps:

- To calibrate two cameras and to calculate the researching region according to the table region and the prediction. Let \( M \) be the fundamental matrixes; \( \Omega' \) the calculated region in the image gotten from left camera.
• Input: Two adjacent frames gotten from the left camera, \( I_n^l \) and \( I_{n+1}^l \); Two gotten from the right camera, \( I_n^r \) and \( I_{n+1}^r \).

• Find the ball possible region in the left image \( I^l \) gotten by \( I_{n+1}^l - I_n^l \) according to five key points.

• To locate the center of the detected ball.

• Find the ball possible region in the right image \( I^r \) gotten by \( I_{n+1}^r - I_n^r \). Because we have located the center of the ball in \( I^l \) accurately, the searching area in \( I^r \) can be reduced to the region centered around the point \((u_r, v_r)\).

\[
(u_r, v_r) = M^* [u_j, v_j, 1]^T.
\]

The ball is detected in \( I^r \) by five key points and the center is calculated by the method in the former steps.

IV. EXPERIMENT RESULTS

We test our method practically. The time resolution in our systems is 100Hz. Because we restrict the searching region according to the tale and the predicted position of the ball, the man with white T-shirt in the background and the swinging arm in the left do not affect the detection results. Fig.3 shows the experiment results on several frames of an image sequence by our method. The left column is the original images captured by the left camera; the right column is the outputs. The red stars in the right column represent the detected centers of the ball. Because only five key points are used in our method, the detecting process is fast enough to meet the need of Ping-Pong playing robots.

V. CONCLUSION

Though the study on the vision system of Ping-Pong playing robots has been over years, it is still a challenge topic due to its high-speed and complicated background. In the paper, we present a fast one-step scanning method to detect the ping-pong ball to meet the special needs of the vision system. Our method has the following advantages:

1) Though the ball region detected by the five-key-point method may be smaller than the real region or be part of the region, it restricts the detected region to be round-like;

2) Since only five points are used in the method, the process can be finished in one image scanning;

3) Our method has the anti-noise ability to a certain extent. Four points chosen by us are located at the edge of the ball, and the other one is at the center. Discrete noises and small pieces will be removed in the detecting in such layout.

The experiments verify its efficiency and accuracy, but there are still some problems unsolved. The method fails when the ball extracted from the binary image is of the incomplete round shape. The future work focuses on the searching region restricting method and thresholds selection.

Figure 5. The results detection of the ping-pong ball practically. Left column: the original images captured by the left camera; the right column: the detecting results represented by the red stars.

ACKNOWLEDGMENT

The work is carried out at State Key Laboratory, Shenyang Institute of Automation, Chinese Academy of Science, and is was supported in part by the Natural Science Foundation of China (Grant Number: 60871078 and 60835004).
REFERENCES

[1] A. L. Russell, “A Robot Ping-Pong Player: Experiment in Real-Time Intelligent Control”, AT&T Bell Laboratories, MIT Press, 1988.

[2] G.D. Hager, W.-C. Chang, and A.S. Morse, “Robot hand-eye coordination based on stereo vision,” IEEE Control Syst. Mag., vol. 15, pp.30-39, Jan. 1995.

[3] H. Hashimoto, F. Ozaki, K. Asano and K. Osuka, “Development of a ping pong robot system using 7 degrees of freedom direct drive arm,” Industrial applications of Robotics and machine vision (IECON), pp. 608-615, 1987.

[4] H. Fassler, H.A. Vasteras, and J.W.Zurich, “A robot ping pong player: optimized mechanics, high performance 3D vision, and intelligent sensor control,” Robotersystem, Berlin: Spring-verlag, pp.161-170, 1990.

[5] F. Naghdy, J Wyatt, and S. tran, “A transputer-based architecture for control of a robot ping pong player,” Parallel computing and Transputers, pp. 311-317, 1993.

[6] F. Miyazaki, M. Takeuchi, M. Matsushima, T. Kusano and T.Hashimoto, “Realization of the table tennis task based on virtual targets,” Proc. International Conference on Robotics and Automation Vol 4, pp. 3844-3849, May 2002.

[7] B. Peng, Y.C. Hong, S.S. Du and W. Wei, “An Approach to Hit Point Prediction for Ping-Pong Robot”. Journal of Jiangnan University (Natural Science Edition), Vol. 6, pp.433-437, Aug. 2007.

[8] J. Ren; M. Xu, J. Orwell , and G.A.Jones, “Real-Time Modeling of 3-D Soccer Ball Trajectories From Multiple Fixed Cameras”, Circuits and Systems for Video Technology, IEEE Transactions on. Vol. 18, pp.350-362, March 2008.

[9] D.Yow, B. L. Yeo, M.Yeung, and B. Liu, “Analysis and presentation of soccer highlights from digital video,” in Proc. 2nd Asian Conf. Comput.Vis., Singapore, pp. 499-503,1995.

[10] X. F. Tong, H. Q. Lu, and Q. S. Liu, “An effective and fast soccerball detection and tracking method,” in Proc. IEEE Int. Conf. Pattern Recognit., Cambridge, U.K., vol. IV, pp. 795-798, 2004.

[11] A. Yamada, Y. Shirai, and J. Miura, “Tracking players and a ball in video image sequence and estimating camera parameters for 3-D interpretation of soccer games,” in Proc. IEEE Int. Conf. Pattern Recognit.,Québec, QC, Canada, vol. I, pp. 303-306, 2002.

[12] Y. Tang, X.Li, A. Freyberg, and G. Goch, "Automatic Segmentation of the Papilla in a Fundus Image Based on the C-V Model and a Shape Restraint," Pattern Recognition, 18th International Conference on Pattern Recognition (ICPR'06), vol. 1, pp. 183-186, 2006.