

Research Article

Binocular Vision-Based Recognition Method for Table Tennis Motion Trajectory

Chunfeng Lu and Xiyu Tang 

Zhejiang University of Science and Technology, Hangzhou, Zhejiang 310000, China

Correspondence should be addressed to Xiyu Tang; 102108@zust.edu.cn

Received 10 May 2022; Revised 28 June 2022; Accepted 13 July 2022; Published 10 August 2022

Academic Editor: Abid Yahya

Copyright © 2022 Chunfeng Lu and Xiyu Tang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The latest development in computer technology in sports has increased the popularity of ping pong among people. It calls for designing an intelligent table tennis robot using computer technology to cooperate with table tennis enthusiasts and professional table tennis players to practice. Much research work has already been done in this area. However, this study explores the use of binocular vision to precisely identify, locate, and anticipate the flight trajectory and landing position of table tennis balls. Major issues addressed in this study include identifying fast-moving ping pong balls, calibrating the camera, obtaining the camera's internal and external settings, and localizing the ping pong balls in three dimensions. A new target recognition method is proposed in combination with the actual needs of the combat ping pong robot. The method combines colour segmentation, background subtraction, and ellipse fitting, which can detect the tail range of flying ping pong balls and find the centre position of the balls. Based on the ellipse fitting analysis of the image, characteristics of the tail of the flying ping pong ball are studied. This study can aid in tracking the trajectories of high-speed flying objects, which is helpful for both aerospace and military industries.

1. Introduction

The table tennis robot has become an automated machine capable of competing with human sports through perception, prediction, decision making, and coordination of table tennis motion [1]. As a real-time, intelligent robot motion servo platform, it shows good system integration and technology and has a promising application. However, for a robot to play table tennis correctly and position with correct posture and speed, it needs to identify, track and estimate the trajectory of fast-moving ping pong balls, which involves several complex operations such as stereo vision and intelligent control. For this purpose, the development of vision systems is the first task in the development of table tennis robots. The perceptual part of the human visual system is the retina, which is equivalent to a three-dimensional world sampling system. After light hits the visible part of an object, it is projected onto the retina, producing a two-dimensional image of the object [2]. Based on this image, the human brain can understand the object from a three-dimensional perspective. By three-dimensional understanding, we imply

knowing the object's size, dimensions, form, motion direction, and speed. The vision system of the table tennis robot uses a camera instead of the retina and a computer instead of the human brain to simulate this human visual function to recognize and understand objects in a three-dimensional environment.

The study of modeling table tennis trajectory tracks and forecasts the table tennis ball's flight allowing the robot to calculate, where it will land and when to hit it. [3]. These robots can replicate human vision using a table tennis robot vision system. The table tennis ball movement data are then transferred to a PC for further processing. Vision systems are classified as monocular, binocular, or triangular based on the number of cameras utilized. Multilocular vision has higher recognition accuracy and a wider field of view. However, at present, most researchers employ binocular vision systems. Binocular vision views the same object from two angles, like human vision.

In this study binocular vision system has been employed to examine and model table tennis motion trajectory. Major tasks involve

- (i) Combining camera calibration to enhance the performance of the players
- (ii) Image models with target detection technology to find the centre of the moving table tennis ball in three dimensions

In this way, tracking the ping pong ball's flight trajectory and retrieving its motion information will become easier. Also, it will make the robot capable of playing table tennis.

The study is structured as follows. Related work in section 2 discusses the visual localization methods and table tennis trajectory prediction based on physical and machine learning models. In section 3, the methodology of this research work is explained. It comprises the design of an experimental binocular system and mathematical modeling. Results and conclusions are discussed in sections 4 and 5, respectively.

2. Related Work

In this section, the literature survey of different methods for visual localization and techniques utilized for determining the trajectory of a table tennis ball has been discussed in detail.

2.1. Visual Localization of Ping Pong Balls. Table tennis three-dimensional trajectory extraction is the basis of trajectory prediction research. Machine vision has aided researchers in extracting the flight object motion trajectory. The key points of the research are to meet the desired accuracy and real-time stability simultaneously.

Machine vision uses computers to measure and judge the objective environment in three dimensions. Human perception of the objective world is 70% through the visual system [4]; the human visual system collects image information through the retina; the visible part of the object is projected into the retina, whereas the brain analyzes the shape, size, spatial relative distance, colour texture, and motion characteristics of the observed object using two-dimensional imaging in the retina for three-dimensional processing in a short period. To function like the human eye imitates the human visual system through the camera for image acquisition and the computer for data analysis and processing. Machine vision systems for visual positioning are classified as monocular, binocular, or multivision [5].

2.1.1. Monocular Vision. Monocular vision extracts the target object's motion trajectory by image acquisition with one camera, but the image depth information cannot be extracted directly due to the single camera. The first proposed a single camera for ping pong ball localization, identifying the small ball, and the shadow of the small ball on the table and determining the 3D position coordinates of the ping pong ball through the geometric relationship between the camera, the light source, the ping pong ball, and the shadow of the small ball in space [6, 7]. Image coordinates of a table tennis ball and its shadow were recovered using a monocular camera system initially used to compute its spatial 3D coordinates. A single camera has also been used for ping pong ball localization, segmented the ball motion

from the background motion using the displacement frame difference method, and tracked the ball in 3D using the parameter calibration of a single CCD (charge-coupled device) camera [8]. It has determined the magnitude and direction of the ball's spinning speed in real-time by photographing the difference between the long and short axes of the elliptical trajectory formed by the markers on the ping pong ball during its rotational motion in the air [9].

2.1.2. Binocular Vision. Binocular vision has a broader detection range than monocular vision. Monocular vision can only utilize one camera and has strict environmental conditions. In binocular vision, the spatial object is mapped to the same name pixel point parallax in the image plane of the two cameras using geometric triangulation [10]. Also, the latest studies use binocular vision systems for table tennis trajectory extraction.

A wall-mounted binocular vision system has been designed using two coloured cameras with 50 fps and 320×240 resolution [11]. In contrast, the camera's field of view was about 50° from the horizontal plane. By using a simple colour segmentation scheme to detect balls in each image, the system could finally extract ping pong ball trajectories with a velocity range of 5–6 m/s. Similarly, a distributed parallel processing high-speed vision system based on smart cameras has also been designed [12]. The two high-speed cameras with a 250 fps acquisition rate build a binocular vision system that captures and processes images in parallel mode. They also proposed a grayscale image-based ball recognition and tracking algorithm. Thus, allowing the system to capture and process a grayscale image frame within 6 ms. The longer the binocular vision baseline, the higher the accuracy, but in practice, the table tennis robot does not have space to assemble a long-baseline binocular vision system. To solve this problem, a short baseline binocular vision system was developed with a baseline length of 0.18 m [13]. Two cameras with 640×480 were used for image acquisition. A projection matrix direct calibration and Gaussian fitting-based ball centre localization algorithm were used to limit the detection area of little tennis balls, improving the system's real-time performance.

For the problem of accuracy degradation caused by the nonsynchronous triggering of binocular vision, an onboard stereo vision system was designed to solve the problem of asynchronous observation between different cameras by considering the consistency of ping pong ball motion and achieve an accurate real-time estimation of ball trajectories, which can achieve the same effect as hardware synchronous triggering [14].

2.2. Table Tennis Trajectory Prediction. The current research for table tennis trajectory prediction predicts based on the classical physical motion model and machine learning model.

2.2.1. Prediction of Trajectories Using the Physical Model. Landing friction, gravity, Magnus force, air resistance, and ball rebound must be kept in mind in traditional table tennis

trajectory prediction models [14], which results in high model complexity and low robustness. An analytical model of the rebound between the table tennis ball and table/racket rubber was derived [15]. Where it is assumed that the kinetic energy of the tangent velocity is stored as potential energy because of the rubber's elasticity, and the impulse in the horizontal direction is proportional to that tangent velocity. A motion model proposed that it does not consider the rotation of a ping pong ball [16]. A nonlinear bounce model was developed that considers the spin of a ping pong ball based on the momentum theorem and the momentum moment theorem model [17]. It describes the collision process of the ball with the object table. The collision of a ping pong ball with a table was analyzed using an ultrahigh-speed camera and established a physical model of the self-spin and collision effects of the table tennis ball [18].

A motion model of a rotating table tennis ball was proposed [19]. Three second-order polynomials fitted the starting trajectory, and the ping pong ball's initial velocity comprised the flight and rotation velocity, which were calculated using polynomials. In the experiments, their proposed method demonstrated good prediction ability at a sampling time of 1 ms, but the accuracy of the flow velocity estimation was affected if the consecutive trajectory points were not dense enough or the sampling frequency was too low. Inserting those estimates into the complex equation calculation increased the distortion of the results. A motion model, in which the motion trajectories were first clustered using the K -means algorithm and then fitted to get the extended continuous motion model (ECMM) [20]. A novel strategy based on expectation-maximization was presented to anticipate motion trajectory better. ECMM's category is a potential variable that is represented as the difference between ballistic prediction and observation.

2.2.2. Machine Learning-Based Trajectory Prediction. In recent years, Machine learning methods have also been used to determine the trajectory of the table tennis ball. A Kalman filter-based table tennis ball trajectory prediction method was devised [21]. In the first step, the top, rear, left, and mixed spin dynamics of the ball were modeled, and then using the Kalman filter for curve optimization, flight trajectory was obtained, which was then used to simulate and predict the ball position. The misconceptions caused by blurred images, air resistance, and camera imaging distortion in table tennis' high-speed movements were addressed [22]. It was proposed that the adaptive measuring covariance discrete Kalman trajectories estimation. The algorithm tracks the target motion trajectory by dynamizing the measurement covariance, providing the ground for ping pong ball tracking prediction and hitting the arm. Experimental results show that the algorithm can effectively overcome the interference of measurement noise and data loss while achieving good tracking results when the image acquisition rate is greater than 70 frames/s, and the ping pong ball motion speed exceeds 5 m/s.

A new trajectory prediction model was proposed [23], where a nonlinear filter based on fuzzy logic was employed

to eliminate the noise in the table tennis coordinate system. Then the least square method was used to calculate the initial flight speed and rotation speed based on the filtered ball position. Second, postbounce speed prediction was made using memory-based local modeling, and finally, a ball flight and bounce model was established to predict the subsequent trajectory of the ball. A deep conditional generative trajectory prediction model for the characteristics of table tennis trajectories belonging to time series data was presented [24].

A dual artificial neural network to mimic the table tennis ball was utilized [25], which divides the table tennis ball trajectory into two segments bounded by the landing point. Here, historical data were used to learn the patterns in them, with a final experimental error of 39.6 mm. Similarly, a deep conditional generation model based on the deep learning algorithm optimized method for ping pong ball trajectory prediction was proposed [26]. It first constructs a dataset of ping pong ball spatial location images with accurate labeling in various environments by utilizing the basis of traditional deep learning.

Meanwhile, a recurrent neural network-based ping pong ball trajectory prediction algorithm was proposed using a convolutional neural network as the location recognition algorithm, and the effectiveness of the method was verified by comparative analysis of multiple experiments.

While previous research work has been on motion detection and trajectory prediction of the table tennis ball using various complex methods. This study utilizes a tracing and tracking algorithm of the table tennis ball based on the binocular vision method that makes use of both hardware and software structures for this purpose. Key features for camera calibration are proposed, and trajectory prediction is made using the ellipse fitting technique.

3. Methodology

In this section, a prototype for designing a binocular vision system is proposed comprising both hardware and software structures. CV software identifies, track, trace, and predict table tennis ball trajectories. Mathematical modeling of target identification and extraction is done using the ellipse fitting method.

3.1. Design of the Binocular Vision Experimental Platform System. The system structure used in this project is based on the ontology of the parallel optical axis vision system. This structure requires the focal length of the left and right cameras and their internal parameters to be equivalent and the optical axis perpendicular to its imaging plane, the two, left and right cameras in the three-dimensional coordinate system, the X -axis and Y -axis coincide with parallel to each other.

As shown in Figure 1, the right camera will be moved in the opposite direction along the X -axis covering baseline distance denoted as b , where it can fully coincide with the left camera. In the polar plane, the optical centre of the left camera is determined by O_l the optical centre of the right

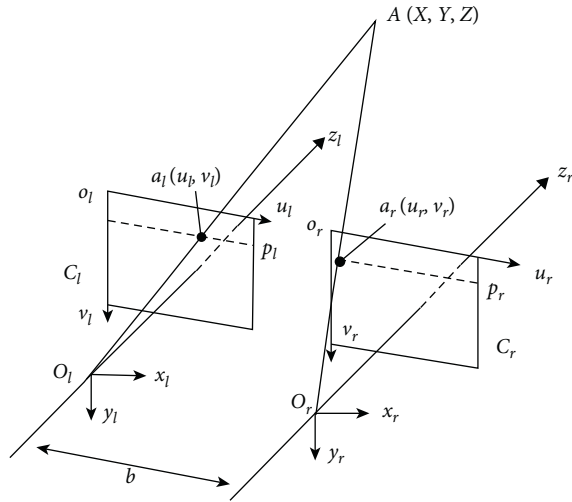


FIGURE 1: Schematic diagram of the ontology vision system based on a parallel optical axis.

camera by O_r . A point “A” is taken in three-dimensional space, which intersects the imaging planes C_l and C_r of the left and right cameras, respectively. The P_r and P_l intersection lines represent the conjugate polar pairs, the coordinate axis u_l represents the coordinate system of the imaging plane of the left camera and the coordinate axis u_r represents the coordinate system of the imaging plane of the right camera. Both u_l and u_r coincide and are parallel to P_l and P_r , respectively.

This structure is ideal, and the geometric relationship between the left and right cameras is the simplest. It is relatively convenient to solve the matching relationship between the points a_l and a_r projected the spatial point A onto the left and right imaging planes in this structure.

There are two key components to a computer: hardware and software. A binocular stereo vision system includes both hardware and software components. To create the ontology vision system of the humanoid robot, we first select the relevant hardware components and develop the software program structure. Figure 2 shows the specific structure diagram.

3.1.1. System Hardware Structure. Stereo binocular vision is divided into hardware and software structures, as shown in Figure 3. The hardware structure is subdivided into three major components. The necessary hardware tools required are discussed in this section. Two cameras are used to capture video data. They are connected to an image capture card through a camera controller.

The vision processing computer used in this project is configured with CPU: AMD Athlon (TM) II X2 245 Processor 2.90 GHz, memory 1.75 GB, and WINDOWS XP system. When choosing a camera, the following points must be kept in mind [27].

(1) *Frames per Second (FPS)*. The binocular vision system’s [28] key aims are speed and precision. Exposure time for a table tennis ball having a velocity of 5 m/s and a 40 mm diameter (international standard ball) is 40 mm/5 m/

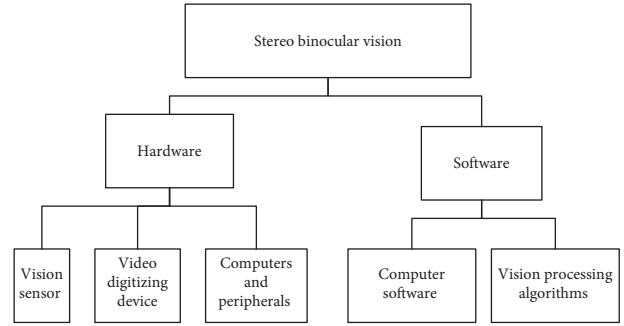


FIGURE 2: Stereo binocular structure (hardware and software).

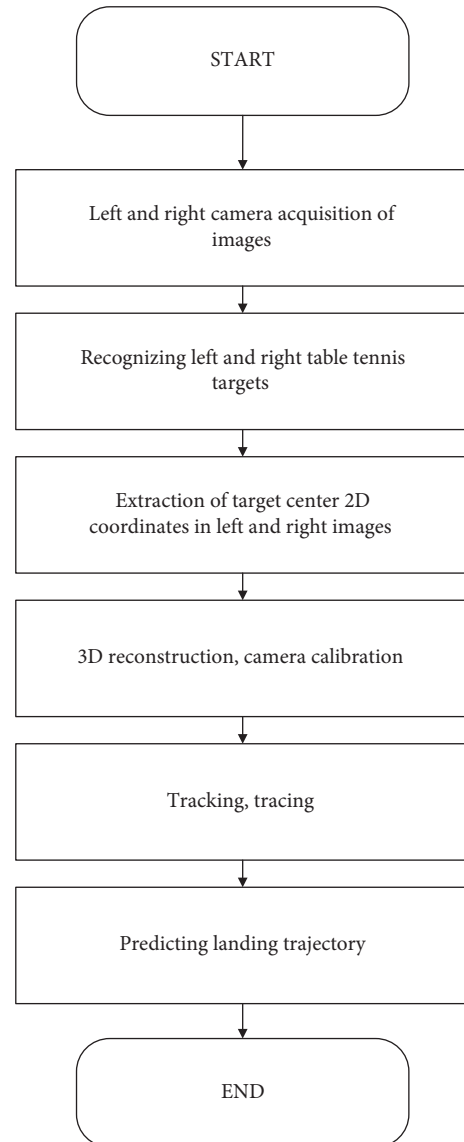


FIGURE 3: Vision software architecture.

$s = 8$ ms. The table tennis ball moves 40 mm within the exposure time, which will produce a “trailing shadow” on the 2D image, which will result in a big inaccuracy in the table tennis ball identification. The better the camera’s frame rate, the more continuous picture might well be made.

(2) *Colour*. The object's colour conveys detailed information; hence the colour camera is chosen.

(3) *Pixels*. Many pixels give better discrimination accuracy, but more data means longer transmission times, which results in lower frame rates. So, choose the biggest pixel value possible to satisfy the frame rate size.

(4) *Transmission Interface*. The size of the transmission rate has a greater impact on the real-time system, which directly impacts the camera's frame rate. It is necessary to have the optimal frame rate value for a faster transmission rate.

3.1.2. System Software Architecture. The system's software is written in Visual Studio 2008 using the OpenCV vision library, as shown in Figure 3. Using the target identification method, the left and right cameras detect the area and locate the centre of the captured photos with ping pong balls. The 3D position of the table tennis balls is reconstructed and estimated using the left and right projection pictures. To achieve fast-tracking, the ping pong ball's future position is estimated using its historical position information.

3.2. Table Tennis Target Identification. Target recognition is the basis of the whole binocular vision, providing raw data for the subsequent segments. Table tennis target recognition first requires detecting the target region of interest from the image, then retrieving the target region's two-dimensional picture coordinates, i.e., the table tennis ball's centre coordinates. The accuracy of the ball centre coordinates obtained from recognition has significant effects on target localization, tracking, and trajectory prediction accuracy. At first, the threshold range of the target's colour in the YCbCr colour space is determined offline, this colour space allows the computer to focus on moving the target in an image while keeping the high-resolution value. The target is then separated from the complete picture by the threshold segmentation technique, and then the centre point of the separated target, i.e., the ball centre coordinates, are obtained by the ellipse fitting technique. In this study, we analyze the trailing image characteristics of flying ping pong balls and the actual needs of a real-world ping pong robot. Integrate and complement colour segmentation, background subtraction, and ellipse fitting to detect the trailing range of flying spheres and find the centre of the sphere.

The experimental results show that the method can effectively overcome trailing shadow, identify, and track the high-speed flying spheres more accurately, and has good anti-interference and real-time performance.

3.3. Target Extraction of Ping Pong Balls. Due to the length of exposure, the camera used in the acquisition of high-speed movement of the ping pong ball will produce the phenomenon of "motion blur," also known as trailing shadows. Table tennis target extraction is based on segmenting out the flying table tennis body and extracting the image coordinates of its centre point. The ping pong ball is a circular sphere, and the binarized image obtained after image segmentation,

which is the motion area of the ping pong ball target, is approximately elliptical, so this study takes the least squares ellipse fitting algorithm to extract the centre point coordinates of the target.

The basic idea of ellipse fitting is to find an ellipse that can be as close to these data points as possible per the given set of data points on the plane. In other words, it is to fit a dataset on the image plane using the elliptic equation as a model. The equation is used to find the closest dataset. The parameters obtained are then inserted into the elliptic equation. Finally, the centre point of the ellipse is extracted.

3.4. Mathematical Modeling. The least square ellipsoid fitting method minimizes the total error rates. Using the ellipse fitting algorithm, the mathematical modeling is done in the following steps:

Step 1. Find a set of parameters.

Step 2. Minimize the distance between data points and ellipse.

Here, the distance metric uses algebraic distance. Making an integer multiple of the solution is considered the same elliptic representation. To avoid the generation of zero solutions, a restriction is made on the parameters with the constraint $A + C = 1$. The extracted target contour points are subjected to least squares to obtain the coefficients of the equation.

$$f(x, y) = \sum_{i=1}^N (Ax_i^2 + Bx_iy_i + Cy_i^2 + Dx_i + Ey_i + F)^2, \quad (1)$$

that is, find the minimum value of the function $f(x, y)$ in (1). By the extreme value principle, when $(\partial f/\partial B = \partial f/\partial C = \partial f/\partial D = \partial f/\partial E = \partial f/\partial F = 0)$, the $f(x, y)$ value is minimum.

This results in a linear system of equations and solving this equation with the constraints leads to the values of the five parameters ($A, B, C, D,$ and E).

Then, according to (2), the centre coordinate position (x_0, y_0) of the moving target can be found.

$$\begin{aligned} x_0 &= \frac{BE - 2CD}{4CD - B^2}, \\ y_0 &= \frac{BD - 2AE}{4AC - B^2}. \end{aligned} \quad (2)$$

3.4.1. Identifying the Ping Pong Ball Body in a Trailing Shadow. The tennis ball's exposure time is 8 ms, and the ping pong ball moves 40 mm within the exposure time, resulting in a "trailing shadow" phenomenon on the 2D image. Therefore, we solve this problem by identifying the ping pong ball with high-speed movement.

The combination of the above two methods is used to obtain the binarized contour corresponding to the trailing image of the ping pong ball body, which is approximated by

an ellipse. So, the ellipse fitting method can obtain the trailing contour representation.

Figure 4 shows the coordinate diagram of ellipse fitting and sphere centre position, where the solid line represents the fitted ellipse, (x_c, y_c) is the centre coordinate, “ b ” is the long axis, “ a ” is the short axis (can be regarded as the radius of the sphere), the long axis of the ellipse (can be regarded as the flight direction of the sphere), and these parameters are obtained by ellipse fitting.

$$\begin{cases} x_p = x_c + (b - a)\cos \theta \\ y_p = y_c + (b - a)\sin \theta \end{cases} \quad (3)$$

The long axis of the ellipse in each frame represents the ball’s instantaneous flight direction, while the short axis’ length is approximately equal to the ball’s radius. Derivation of the 2D centre of the ping pong ball is done after fitting ellipses. The dashed circle in Figure 4 can be regarded as the position of the sphere and (x_c, y_c) is the coordinates of the centre of the sphere, which can be calculated using equation (3), i.e., the centre of the ellipse (x_c, y_c) is obtained by translating the difference between the lengths of the long and short axes ($b-a$) along the direction of the long axis (θ angle).

4. Experimental Results and Analysis

In this section, the experiment dataset has been explained, followed by model training and target recognition, trajectory extraction and error analysis, and prediction of bounce coefficients of table tennis trajectories.

4.1. Dataset. Table tennis balls have the characteristics of fast speed and small size. To ensure that the trajectory of small balls can be captured in real-time, a camera with a high frame rate must be used for shooting. In this experiment, a high-speed black-and-white industrial camera with specific colour components of YCbCr space, a high resolution of 1280×1024 , and a good frame rate of up to 210 fps (model MV-CA013-21UM) have been used for image acquisition. Because of the simple experimental setup with few surrounding distractions, a large dataset was not required, and 1000 table tennis motion images were collected as the dataset for the experiment. Its specific parameters for ideal calibration are discussed in Section 3.1.1.

4.2. Model Training and Target Recognition. The experimental platform for model training and target recognition is i7-7700K CPU, 16 GB memory, TITAN X graphics card, and Ubuntu 16.04 operating system. After the network is compiled, the number of training iterations is set to 10000, the learning rate is 0.00001, the batch is 64, the subdivision is 16, and the decay is 0.00001.16, and the decay is 0.0005.

4.3. Trajectory Extraction and Error Analysis. This experiment extracts the ping pong ball trajectory. At first multivision calibration is carried out, and the specific calibration parameters are identified as discussed in the system hardware structure.

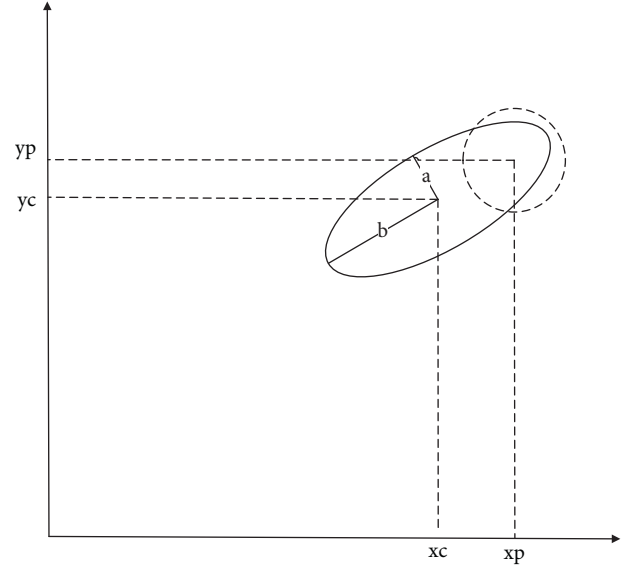


FIGURE 4: The ellipse fitting and coordinates of sphere centre position.

TABLE 1: The real-time coordinate values of trajectory points.

	X (mm)	Y (mm)	Z (mm)
1	86.8	66.2	94.6
2	107.7	82.4	77.1
3	124.5	95.6	60.1
4	140.9	108.5	41
5	156.9	121.2	19.9
6	172.6	133.6	0.4
7	182.2	140.5	13.4
8	194	148.9	31.2
9	205.6	157.4	46.1
10	217.1	165.7	58.2

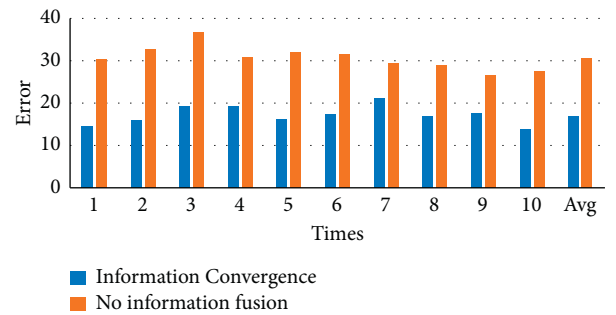


FIGURE 5: The comparison table showing the average error of trajectory extraction.

Secondly, multiview image target recognition is conducted, and then the centre point of the target frame is extracted. The matching pixel point coordinates are inserted into the multivision 3D positioning formula to get the 3D coordinate values. This process is done using a multicamera information fusion strategy.

Table 1 provides 10 sets of 3D coordinate values of table tennis ball trajectory points in real-time, whose reference

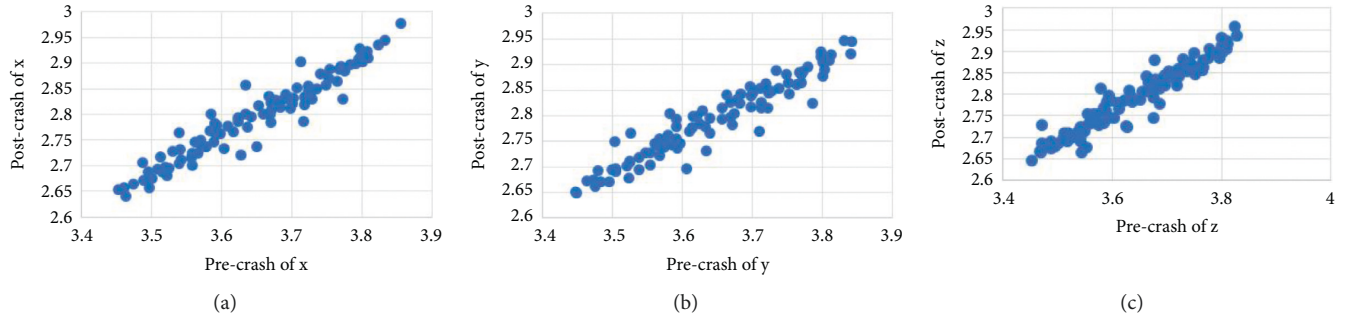
FIGURE 6: The rebound coefficient. (a) x -axis. (b) y -axis. (c) z -axis.

TABLE 2: Rebound coefficient.

	K_x	K_y	K_z	B_x	B_z
Average	0.78	0.63	0.78	0.14	0.18

coordinate system overlaps with the physical coordinate system of one camera.

In this study, ping pong balls are fixed at 9 different spatial locations for error analysis, and the actual distance and positioning distance between each fixed ball and its neighboring balls are measured by a laser distance meter and multivision system, respectively. Figure 5 shows the error analysis graph.

Figure 5 shows the error comparison between the adjacent positions of ping pong balls in the first trajectory extraction. The average error is 16.65 mm in the first test, when the information fusion strategy is used, while the average error is 30.74 mm when no information fusion strategy is used. After the first test, the placement of the balls was adjusted, and the test was carried out 10 times.

4.4. Prediction of Bounce Coefficients of Table Tennis Trajectories. The experiment takes the velocity before the bounce as the horizontal coordinate and the velocity after the bounce as the vertical coordinate and then uses the least squares fitting to determine the bounce coefficients in the horizontal and vertical directions, respectively. The fitted bounce velocity curves are shown in Figure 6.

Figure 6 shows the bounce velocity curves for each velocity component fitted by the least square method, whereas the bounce coefficients can be specified based on the slope and intercept of each fitted curve. The detailed coefficients are given in Table 2.

5. Conclusion

This study develops a table tennis ball identification and tracking system based on binocular vision. The system can accurately identify the speed of the ping pong ball, combine the camera imaging model, and realize the 3D positioning of the ping pong ball body. The trajectory tracking of the ping pong ball according to the ball motion identification and positioning information based on theoretical study and algorithm implementation of the position prediction model

has been done. The extensive use of the ellipse fitting method has aided in achieving fast-flying ping pong ball ghost image recognition tracking and calculating its position. This method effectively overcomes ghosting, and the calculation cost is low. The strong anti-interference ability for the actual type of combat will help to achieve continuous, real-time fast-flying table tennis position identification. In addition, this method is not only applicable to orange ping pong balls but also can be applied to other colour ghost images for rapid target identification and object tracking.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] Y. YaoJun-Cai and L. LiuGui-Zhong, "Objective assessment method of image quality based on visual perception of image content," *Acta Physica Sinica*, vol. 67, no. 10, Article ID 108702, 2018.
- [2] M. W. Burke, *Image Acquisition: Handbook of machine vision engineering*, Vol. 1, Springer Science & Business Media, Berlin, Germany, 2012.
- [3] X. Lin, J. Wang, and C. Lin, "Research on 3D reconstruction in binocular stereo vision based on feature point matching method," in *Proceedings of the 2020 IEEE 3rd International Conference on Information Systems And Computer Aided Education (ICISCAE)*, pp. 551–556, IEEE, Dalian, China, 2020, September.
- [4] S. Vallerand, M. Kanbara, and N. Yokoya, "Binocular vision-based augmented reality system with an increased registration depth using dynamic correction of feature positions," in *Proceedings of the IEEE Virtual Reality, 2003*, pp. 271–272, IEEE, Los Angeles, CA, USA, 2003, March.
- [5] K. Okada, M. Inaba, and H. Inoue, "Integration of real-time binocular stereo vision and whole body information for dynamic walking navigation of humanoid robot," in *Proceedings of the IEEE International Conference On Multisensor Fusion And Integration For Intelligent Systems*, pp. 131–136, IEEE, Tokyo, Japan, 2003, August.
- [6] Y. Fang, I. Masaki, and B. Horn, "Depth-based target segmentation for intelligent vehicles: fusion of radar and

- binocular stereo,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 3, no. 3, pp. 196–202, 2002.
- [7] C. F. Olson, H. Abi-Rached, M. Ye, and J. P. Hendrich, “Wide-baseline stereo vision for Mars rovers,” vol. 2, pp. 1302–1307, in *Proceedings of the 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003)*(Cat. No. 03CH37453), vol. 2, pp. 1302–1307, IEEE, Las Vegas NV USA, 2003 October.
 - [8] C. Tong, “Three-dimensional reconstruction of the dribble track of soccer robot based on heterogeneous binocular vision,” *Journal of Ambient Intelligence and Humanized Computing*, vol. 11, no. 12, pp. 6361–6372, 2020.
 - [9] K. Zhang, Z. Cao, J. Liu, Z. Fang, and M. Tan, “Real-time visual measurement with opponent hitting behavior for table tennis robot,” *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 4, pp. 811–820, 2018.
 - [10] K. Lu, X. Wang, Z. Wang, and L. Wang, “Binocular stereo vision based on openCV,” in *Proceedings of the IET International Conference on Smart and Sustainable City (ICSSC 2011)*, Shanghai, July 2011.
 - [11] L. Zhang, F. Cai, J. Wang et al., “The SGM algorithm based on census transform for binocular stereo vision,” in *Proceedings of the 2022 International Conference on Machine Learning And Knowledge Engineering (MLKE)*, pp. 50–54, IEEE, Guilin China, 2022 February.
 - [12] D. Jiang, Z. Zheng, G. Li et al., “Gesture recognition based on binocular vision,” *Cluster Computing*, vol. 22, no. S6, pp. 13261–13271, 2019.
 - [13] B. Bhowmick, S. Bhadra, and A. Sinharay, “Stereo vision based pedestrians detection and distance measurement for automotive application,” in *Proceedings Of The 2011 Second International Conference on Intelligent Systems, Modelling And Simulation*, pp. 25–29, IEEE, Phnom Penh Cambodia, 2011, January.
 - [14] M. Kytö, M. Nuutinen, and P. Oittinen, “Method for measuring stereo camera depth accuracy based on stereoscopic vision,” in *Three-Dimensional Imaging, Interaction, and Measurement* vol. 7864, , pp. 168–176, SPIE, 2011, January.
 - [15] G. Di Leo, C. Liguori, and A. Paolillo, “Covariance propagation for the uncertainty estimation in stereo vision,” *IEEE Transactions on Instrumentation and Measurement*, vol. 60, no. 5, pp. 1664–1673, 2011.
 - [16] Y. M. Mustafah, R. Noor, H. Hasbi, and A. W. Azma, “Stereo vision images processing for real-time object distance and size measurements,” in *Proceedings Of The 2012 International Conference on Computer and Communication Engineering (ICCCE)*, pp. 659–663, IEEE, Kuala Lumpur, Malaysia, 2012, July.
 - [17] X. B. Lai, H. S. Wang, and Y. H. Xu, “A real-time range finding system with binocular stereo vision,” *International Journal of Advanced Robotic Systems*, vol. 9, no. 1, p. 26, 2012.
 - [18] H. Tsung-Shiang, Tsung-Shiang, and T. C. Wang, “An improvement stereo vision images processing for object distance measurement,” *International Journal of Automation and Smart Technology*, vol. 5, no. 2, pp. 85–90, 2015.
 - [19] M. Rezaei, M. Terauchi, and R. Klette, “Robust vehicle detection and distance estimation under challenging lighting conditions,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 5, pp. 2723–2743, 2015.
 - [20] C. Sánchez-Ferreira, J. Y. Mori, M. C. Q. FariasFarias, and C. H. Llanos, “A real-time stereo vision system for distance measurement and underwater image restoration,” *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, vol. 38, no. 7, pp. 2039–2049, 2016.
 - [21] C. Li, L. Li, G. Wu, N. Xu, H. Gao, and G. Tang, “A visual 3D modelling system for the robot autonomous obstacle negotiation in substation,” in *Proceedings Of The 2016 4th International Conference on Applied Robotics For the Power Industry (CARPI)*, pp. 1–4, IEEE, Jinan, China, 2016, October.
 - [22] C. Lu, “Kalman tracking algorithm of ping-pong robot based on fuzzy real-time image,” *Journal of Intelligent and Fuzzy Systems*, vol. 38, no. 4, pp. 3585–3594, 2020.
 - [23] H. Su, Z. Fang, D. Xu, and M. Tan, “Trajectory prediction of spinning ball based on fuzzy filtering and local modeling for robotic ping-pong player,” *IEEE Transactions on Instrumentation and Measurement*, vol. 62, no. 11, pp. 2890–2900, 2013.
 - [24] S. Gomez-Gonzalez, S. Prokudin, B. Scholkopf, and J. Peters, “Real time trajectory prediction using deep conditional generative models,” *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 970–976, 2020.
 - [25] H. I. Lin, Z. Yu, and Y. C. Huang, “Ball tracking and trajectory prediction for table-tennis robots,” *Sensors*, vol. 20, no. 2, p. 333, 2020.
 - [26] L. Chaoming, “Prediction and analysis of sphere motion trajectory based on deep learning algorithm optimization,” *Journal of Intelligent and Fuzzy Systems*, vol. 37, no. 5, pp. 6275–6285, 2019.
 - [27] G. L. Cai, “A Method for Prediction the Trajectory of Table Tennis in Multirotation State Based on Binocular Vision,” *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 8274202, 10 pages, 2022.
 - [28] Y. Lei, J. Chen, N. Wang, H. Ma, Y. Zhang, and F. Hou, “miR-1285 rgccganr,” *Critical Reviews in Eukaryotic Gene Expression*, vol. 32, no. 3, pp. 83–93, 2022.