Automatic Recognition Method of Table Tennis Motion Trajectory Based on Deep Learning in Table Tennis Training

Liusong Huang[1(✉)], Fei Zhang[2], and Yanling Zhang[3]

[1] Software Engineering, Maanshan Teacher's College, Ma'anshan 243041, China
huangls2016@126.com
[2] Anhui University of Technology, Ma'anshan 243032, China
[3] Department of Computer and Art Design, Henan Vocational College of Light Industry, Zhengzhou 450001, China

Abstract. In table tennis training, in view of the problem of large errors when tracking fast moving targets, this study proposes an automatic identification method of table tennis motion trajectory based on deep learning. The multi-view image of the target object is collected by a multi-eye camera and a stereo image pair is formed. After stereo matching, the three-dimensional coordinate group of the target object is obtained by using the three-dimensional positioning principle of stereo vision. In the three-dimensional coordinate system, the mathematical model of table tennis motion is established. The initial position of the table tennis ball is detected by the Vibe algorithm, and the target area frame is marked, and the frame of the detected target area is used as the first frame of the KCF target tracking to track the table tennis ball. Based on this, a rotating table tennis trajectory recognition network is constructed based on LSTM. The experimental results show that the total trajectory error of this method is 39.00 mm, which can accurately identify the motion trajectory.

Keywords: Table tennis training · Deep learning · Trajectory recognition · Table tennis · Motion trajectory · Automatic recognition

1 Introduction

Table tennis is a sport with strong competitive, ornamental and interesting, it is very popular around the world and has a very strong mass base and enthusiasm in China. In table tennis international competitions, Table tennis players in China often achieve excellent results.

Although table tennis and tennis are both net-confrontation sports, table tennis is small in size and field, so it has higher requirements for the referee. When the tennis ball is rubbed, the referee can easily cause a misjudgment. At the same time, in the table tennis competition, because of the uncertain trajectory of the spinning ball, there

W. Fu and L. Yun (Eds.): ADHIP 2022, LNICST 468, pp. 191–204, 2023.
https://doi.org/10.1007/978-3-031-28787-9_15

are higher requirements for the skills and qualities of the athletes of both sides in the competition.

If the athlete can systematically observe and analyze the running trajectory of the spinning ball in the usual training, it will be of great help to the improvement of the athlete's skills. ITTF and some table tennis event organizers have repeatedly proposed to introduce "Eagle Eye" technology into table tennis competitions. "Eagle Eye" technology was widely used in tennis matches as early as 2003, and then in badminton and volleyball matches. "Eagle Eye" technology overcomes the blind spot and limit of human eyes. Eight or ten high-resolution fast black-and-white cameras are set around the court to quickly capture the movement data of the ball from different angles, and then the flight trajectory of the ball is calculated by the computer to simulate the landing position of the ball, so as to help the referee make a correct judgment.

For "Eagle Eye System" applied to the table tennis game, the main task is to obtain the real-time centroid position information of the table tennis ball and draw the three-dimensional centroid motion trajectory and rotation trajectory of the table tennis ball. In the field of ball games, coaches, ball players and game referees all want to improve their technical level and ability, but all of this is based on fully mastering the law of ball games, and all of this is not clear and accurate to get the law of ball games with human eyes.

In the field of table tennis robot research, it involves many aspects of knowledge learning, such as deep learning, robot kinematics, robot control, etc., which has very high research value and significance. Table tennis has the characteristics of fast moving speed, high-speed rotation, small size and small mass, which brings great research difficulty to the study of table tennis. Table tennis is about skills, with various tactics and increasingly perfect playing methods. Therefore, if a table tennis player or enthusiast wants to improve their skills, the awareness of hitting the ball and the ability to control the ball will be the most important necessary qualities. One is the need for more scientific training.

The traditional research method is mainly based on the color, outline and other characteristics of the table tennis ball, but the surrounding environment in the actual scene will bring great interference to the detection algorithm, resulting in the problem of low detection accuracy, and it needs to be set according to the color of the table tennis ball. The detection threshold is determined, resulting in the limitation of application scenarios. As a substitute for the human eye and human brain, computer stereo vision can obtain the various motion parameters of the ball in the process of movement through the high-speed camera shooting of the ball motion picture and the computer analysis of the shooting picture, so as to master the ball game movement rules. Most of the table tennis trajectory prediction methods are based on the establishment of a physical model, and then the parameters of the model are solved. When predicting the trajectory, the physical model is mainly based on the information of the current moment and the information of the previous moments to make the next position information. Prediction, and then use the predicted information to predict the next moment. When long-term prediction is required, errors are easy to accumulate and the prediction accuracy decreases.

Using scientific and technological means to detect the landing point of table tennis and reproduce the movement trajectory of table tennis to train the awareness of the landing point will effectively improve the ball control ability of table tennis players and contribute to the long-term development of table tennis professional training related fields. In this paper, an automatic identification method of table tennis motion trajectory is proposed based on deep learning, and the method is applied to table tennis training to help athletes better understand the running trajectory of table tennis.

2 Method Design

2.1 Establish the Mapping Relationship from Pixel Coordinates to Three-Dimensional Coordinates

Table tennis has the characteristics of fast speed and small size. In order to ensure that the trajectory of the ball can be captured in real time, a camera with a high frame rate must be used for shooting. Three high-speed industrial cameras are used to extract three-dimensional coordinates of the flight trajectory of table tennis. The "binocular vision system" can also be called "stereoscopic vision", which simulates the principle that the retinas of the two eyes are different when a person observes an object, and the brain fuses them to obtain a sense of depth. Two cameras with a fixed distance are used to obtain two images. Two-dimensional images with different angles, and then calculate the parallax between the corresponding points by the principle of triangulation, and obtain the information of the object in the three-dimensional space.

Let two cameras be in different perspectives, shoot the same target scene at the same time, obtain two two-dimensional plane images through the imaging principle of the camera, and calculate the coordinate deviation between the spatial points mapped to the left and right image pixels, thereby obtaining the spatial information of the three-dimensional scene. Firstly, multi-vision camera calibration is performed to obtain the mapping relationship between image pixel coordinates and world 3D coordinates.

The specific process of multi-vision 3D positioning is as follows: first, the multi-view image of the target object is collected by the multi-eye camera to form a stereo image pair, secondly, the multi-view image stereo matching of the target object is performed, and then the three-dimensional coordinate system of the target object is obtained by using the principle of stereo vision 3D positioning.

The imaging principle of the camera is similar to the ideal model of pinhole imaging. Since the field of view of the camera is very small, it can be regarded as a small hole through which the subject is projected onto the photosensitive element. However, the camera usually needs to be equipped with various convex or concave cameras, which will lead to imaging distortion and become a non-linear model, so the internal and external parameters of the camera need to be corrected. When the left and right cameras are placed completely horizontally, the left and right imaging models of binocular vision will only have deviations in the horizontal direction, but there will be no deviations in the vertical direction. Assuming that the camera coordinate system of the left camera in the stereo vision system coincides with the world coordinate system, and the left camera is set as the main camera, the relationship between the image physical coordinate system and the camera coordinate system is:

$$\gamma\begin{bmatrix} p \\ q \\ 1 \end{bmatrix} = \begin{bmatrix} w_1 & 0 & 0 \\ 0 & w_1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} \tag{1}$$

In formula (1), w_1 and w_2 represent the focal length of the camera in stereo vision; p and q represent the coordinates of the spatial target point mapped in the physical coordinate system of the stereo vision image; α, β and γ represent the three-dimensional coordinates. The transformation relationship between the geometric position of a point on the object in the three-dimensional space and the point mapped to the two-dimensional plane image depends on the camera imaging model parameters [1].

In general, camera parameters need to be obtained through experiments and corresponding calculations, and this process is camera calibration. Both the image coordinate system and the pixel coordinate system are coordinate systems on the imaging plane, but the unit of measurement is different from the origin. The unit of the image coordinate system is mm, the origin is the midpoint of the imaging plane, and the unit of the pixel coordinate system is pixel. Upper left corner of the image. The pixel coordinate system can be obtained by discretizing and offsetting the image coordinate system. Assuming that the camera coordinate system of the right camera has a rotation matrix relative to the camera coordinate system of the left camera, the expression of the three-dimensional coordinates can be obtained as:

$$\begin{cases} \alpha = \frac{\gamma w_1}{h_1} \\ \beta = \frac{\gamma w_2}{h_1} \\ \gamma = \frac{w_1(w_1 h_2 - w_2 h_3)}{w_1 h_1 - w_2 h_2} \end{cases} \tag{2}$$

In formula (2), h_1, h_2 and h_3 are the intrinsic parameters of the left and right cameras. For binocular camera calibration, it is necessary to determine the internal and external parameters of the camera itself. The internal parameters are related to the attribute structure of the camera itself, including focal length, the coordinates of the center point of the imaging plane, the physical size of the pixel, and the distortion parameters. The internal parameters are generally fixed. The external parameters describe the positional relationship between the camera and the world coordinate system, including the third-order orthogonal rotation matrix and the three-dimensional translation vector. Three cameras are used for image acquisition, so two-to-two interactions can form up to seven sets of stereo vision pairs, and up to seven spatial three-dimensional coordinate values of target points can be obtained at each moment [2]. The average value of the stereo vision coordinate values of each group is calculated, and finally the three-dimensional coordinate value of the target point can be obtained.

2.2 Mathematical Modeling of Table Tennis

Since the flight motion of the no-spin table tennis ball can be regarded as a free fall in the vertical direction and a uniform linear motion in the horizontal direction, the ideal state of the no-spin table tennis motion is a flat throwing motion, and its trajectory is a parabola. Its kinematic trajectory is shown in Fig. 1.

Under standard conditions, gravity and air buoyancy are generally constant; air resistance is generally proportional to the square of the flight speed of a table tennis ball, and the direction is opposite to the flight speed [3].

In the space three-dimensional coordinate system o-x-y-z, the trajectory of the horizontal throwing motion is only on the x-o-z plane, and the coordinate origin is established at the vertex of the trajectory. At this time, the trajectory equation is related to the

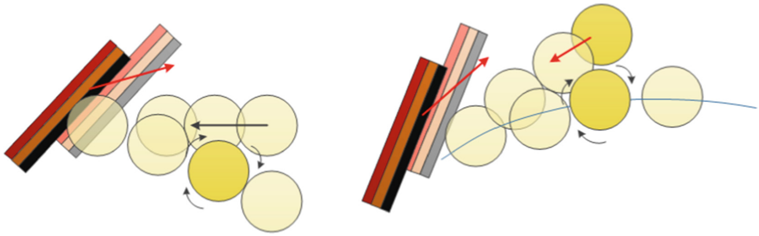

Fig. 1. Kinematic trajectory of table tennis

speed of the horizontal throwing. At present, the commonly used sampling strategies are divided into proportional sampling method and proportional modified symmetric sampling method. Compared with the former, the latter can solve the problem of non-local effects of sampling [4]. Therefore, the sampling strategy for generating the sigma point set in this paper is the proportional correction symmetry method. Among them, the first point is the mean value of the current state, and the other points use symmetrical sampling. Assuming that the coordinate position of the ping-pong ball at the initial moment is (x_0, y_0, z_0), the ping-pong ball moves at the initial velocity u_0, and the angle between u_0 and the xoy surface is ϑ_1, and the angle between the projection of u_0 on the xoy surface and the x-axis is ϑ_2, then the mathematical model of table tennis without rotation at any time is:

$$\begin{cases} x = x_0 + u_0 \cos\vartheta_1 \cos\vartheta_2 \\ y = y_0 + u_0 \cos\vartheta_1 \sin\vartheta_2 \\ z = u_0 \sin\vartheta_1 \end{cases} \tag{3}$$

In formula (3), (x, y, z) represents the position coordinates of the no-spin table tennis ball at any time. In table tennis training, the movement of table tennis is no longer a simple flat throwing movement. Athletes tend to serve spinning balls, and due to the viscous nature of air, spinning table tennis balls are not only affected by gravity, buoyancy, and air resistance, but also by Magnus force. According to hydrodynamics, the pressure on the mountain of fluid with different flow rates is different, so there will be a pressure difference on both sides of the table tennis ball, that is, Magnus force. It can be seen that the Magnus force is not only caused by the rotation speed, but caused by the combined action of the rotation speed and the flight speed.

Compared with the non-spin table tennis ball, the motion model of the spinning table tennis ball is a time-varying nonlinear model [5]. If we ignore the influence of friction on the rotation of the table tennis ball in the air, the attenuation of its rotation speed is very small, so this paper can regard the rotation speed as a constant value.

The trajectory characteristics of the eight kinds of spinning balls compared to the no-spinning balls are shown in Table 1.

When the rotation speed and the flight speed have a large angle and the rotation speed is high, the movement trajectory of the table tennis ball will have a large deviation, that is, the flight trajectory of the rotating table tennis ball is an arc. In the actual movement process of table tennis, it is often not a simple rotation around the coordinate axis, and its

Table 1. Characteristics of the trajectory of the spinning ball

Project	Trajectory trend	Flight arc	Flight distance
Left hand	Yaw to the right	Consistent	Same
Swing ball	Yaw left	Consistent	Same
Topspin	Decline	Little	Relatively close
Backspin	Rise	Big	Relatively far
Left backspin	Turn right, rise	Big	Relatively far
Topspin	Turn right, descend	Little	Relatively close
Right backspin	Turn left, rise	Big	Relatively far
Topspin	Turn left, descend	Little	Relatively close

rotation direction and rotation axis are arbitrary. Therefore, we need to study the rotation around any axis.

The magnitude of air resistance is proportional to the square of the flight speed, and the proportionality factor is determined by the air resistance coefficient, air density and the cross-sectional area of the ping pong ball. The direction of air resistance is opposite to the direction of flight speed. For the rotation of the table tennis ball around any axis, this paper adopts the method of coinciding the rotation axis and the coordinate axis, and then performing the rotation transformation. When the table tennis ball has both flight speed and rotation speed and the directions of flight speed and rotation speed are not parallel, the relative motion between one side of the table tennis ball and the air becomes larger due to the superposition of the rotation speed and the flight speed, and the relative motion between the other side and the air becomes larger. Relative motion is reduced due to the mutual cancellation of rotational speed and flight speed.

Translate the object so that the rotation axis passes through the coordinate origin, rotate the object around the x axis, and make the rotation axis rotate to the xoz plane, the rotation angle is φ_1; rotate the object around the y axis, so that the rotation axis coincides with the z axis, and the rotation angle is φ_2; therefore, the rotation transformation matrix to rotate the angle η around any axis is:

$$B = A_1(\varphi_1)A_2(\varphi_2)A_3(\eta)A_1^{-1}(\varphi_1)A_2^{-1}(\varphi_2)A_3^{-1}(\eta) \tag{4}$$

In formula (4), B represents the rotation transformation matrix that rotates the angle η; A_1, A_2 and A_3 represent the rotation transformation matrix of the point around the x-axis, y-axis and z-axis. The formula for calculating the rotation transformation matrix of any point around the x-axis is given below:

$$A_1(\varphi_1) = \begin{bmatrix} x & y & z & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\varphi_1 & \sin\varphi_1 & 0 \\ 0 & -\sin\varphi_1 & \cos\varphi_1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{5}$$

Similarly, the rotation transformation matrix of any point around the Y-axis and the z-axis can be obtained. Given the initial motion state, the continuous motion model can directly calculate the motion state of table tennis at any time without iteration, which not only effectively eliminates the iteration error and cut-off error, but also directly describes the trajectory position time series and the initial motion state. It can effectively use the time series information of the trajectory position of consecutive multiple frames, and can conveniently obtain the gradient of the trajectory position time series relative to the initial motion state [6]. Thus, the mathematical model of the movement trajectory of the table tennis ball is obtained.

2.3 Motion Target Detection in Table Tennis Training

A complete moving target detection mainly includes three parts: preprocessing, detection algorithm and motion area analysis and processing. The specific table tennis detection flow chart is shown in Fig. 2.

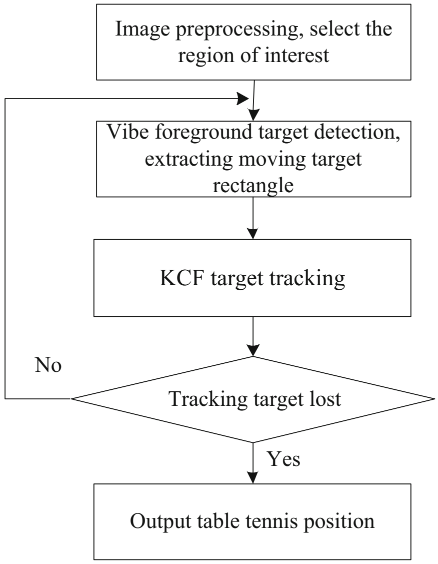

Fig. 2. Table tennis detection flow chart

In this paper, the color information of the image is used to overcome the problem that the result of the difference cannot accurately obtain the moving target area. The image is directly subjected to the difference operation in the color space and the image is binarized by the color threshold, so that the color of the image is preserved information. KCF combines the cyclic matrix in the concept of correlation filtering to solve the problem of a large number of missing negative samples in the target tracking process. The introduction of circulant matrix can increase the number of negative samples, and at the same time can avoid complex matrix inversion operations, effectively reducing the computational load of the classifier.

The Vibe target detection algorithm includes three steps: the construction and initialization of the background model, the detection of foreground moving targets, and the update of the background model set. The image is binarized on the three components of R, G, and B, respectively, and the image is divided by the color threshold, that is, the pixels belonging to different color ranges in the image are distinguished by the set color threshold. For subsequent video frames, the Vibe algorithm uses the two-dimensional Euclidean distance to classify and judge the pixels, and realizes by calculating the Euclidean distance between the pixel value of the current pixel position and the pixel value of the corresponding pixel position in the background model set, namely:

$$H(g_1, g_2) = \sum_{j=1}^{3} (g_1 - g_2)^2 \tag{6}$$

In formula (6), g_1 represents the pixel value of the same pixel point in the background model set; g_2 represents the pixel value of the pixel point of the current input video frame; j represents the three channels of R, G, and B in the color image; H represents the pixel value. Euclidean distance.

Since the area occupied by the ping-pong ball in the image is very small, and the distance between the moving balls in two adjacent frames is small and has continuity and correlation, it will waste a lot of time to process the entire image in each frame [7]. A sub-sampling factor ε is introduced. After the background point in the background model set completes a detection, there is a probability of $\frac{1}{\varepsilon}$ to perform its own update, and there is a probability of $\frac{1}{\varepsilon}$ to update its neighborhood samples [8]. When the number of foreground detections of a certain pixel reaches a critical number, it is immediately set as a background point, and has a probability of $\frac{1}{\varepsilon}$ to update its own background model set. The principle of the dynamic window is to set a fixed-size window. When a moving target is detected in a certain frame of the image sequence, the search range of the next frame can be narrowed to the dynamic window, which can not only reduce the calculation time but also It can eliminate external noise and improve the efficiency of motion trajectory tracking.

Use the spatial propagation between pixels to update the neighborhood to ensure that the update of the background model is gradually diffused outward. When replacing the sample value, a sample value is randomly selected to update to ensure that the sample value has a smooth life cycle. The probability that the sample value remains in the minimal time is:

$$W(\tau) = e^{-\ln\left(\frac{c}{c-1}\right)d\tau} \tag{7}$$

In formula (7), $W(\tau)$ represents the probability that the sample is retained in the minimum time τ; c represents the probability of being updated; e is a natural constant.

The initial position of the table tennis ball is detected by the Vibe algorithm, and the target area frame is marked, and the frame where the detected target area is located is used as the first frame of the KCF target tracking, and the tracking task of the table tennis ball is performed. For the image under the left camera, first determine whether it is the initial frame, if it is the initial frame, first determine the position of the target area, the KCF tracking algorithm obtains from the original image, extracts the features of the moving target, and performs a cyclic shift on it to Training samples to complete the training of the classifier. If the target area is detected in the next frame, continue the tracking process and update the target sample set of KCF. If the target is not detected in the next frame, that is, the target is occluded or lost, the current tracking process is ended. The Vibe algorithm is used again to detect the position of the ping-pong ball, and the frame where the position of the ping-pong ball is detected is taken as the first frame of target tracking. If it is not the initial frame, first determine the position of the candidate region, extract the features of the candidate region and calculate the similarity with the target through the kernel function, and finally select the candidate region with the largest similarity as the final output target, and use it as the tracking target of the next frame [9]. And so on, until the conditions that trigger the system to stop running. In the process of detection and tracking, the coordinates of the center of mass of the tracking frame are always output as the basis for subsequent trajectory recognition.

2.4 Establishment of Automatic Recognition Model of Table Tennis Motion Trajectory Based on Deep Learning

LSTM is a special recurrent neural network, which draws on the long-short-term and forgetting characteristics of human neural memory, and introduces unit state into the neural network, which can solve the long-term dependency that traditional RNN cannot learn and process due to the problem of gradient disappearance. The memory unit of LSTM can consider the connection between distant data, and the flight trajectory of a ping-pong ball is a close connection between long distances. Therefore, this paper attempts to apply LSTM to the trajectory recognition task of table tennis.

This chapter builds an LSTM-based rotating table tennis trajectory recognition network, and trains and optimizes the network. Taking the rotation of the ping-pong ball into consideration, the approximate rotation type of the ping-pong ball is judged by its flight trajectory, which improves the accuracy of the prediction point while ensuring that the detection speed meets the requirements of the vision system. Since the motion state space of table tennis is continuous, the more categories are theoretically classified, the better the adaptability of the extended continuous motion model to table tennis in different motion states [10]. However, the time complexity and space complexity of the algorithm are proportional to the number of categories of the extended continuous motion model. Therefore, considering the accuracy and efficiency of the model, we divide the training data into 8 categories according to the different characteristics of the flight speed decay curve with time. The table tennis trajectory recognition network needs to receive the three-dimensional table tennis position information at the current moment, and output the three-dimensional coordinates of the table tennis ball at the next moment. The basic idea of LSTM is to maintain a state vector internally. During the process of sequence input, the model continuously updates the state vector, and outputs a vector based on the current input and current state. The current moment signal input is combined to output the output value of the LSTM at the current moment.

The overall network structure consists of three parts, namely the input layer, the hidden layer, and the output layer. Input the three-dimensional coordinate information of the ping-pong ball at time t to the network in chronological order. In the LSTM unit, it is necessary to calculate the output of the current state according to the state of the cell and the input, and update the cell information and pass it to the next hidden layer unit, and finally pass the output $t + 1$. The three-dimensional coordinate information of the moment. The output of the LSTM network can be expressed as:

$$Y_t = \theta\left[\varsigma(k_{t-1}, s_t) + \upsilon\right] \tag{8}$$

In formula (8), Y_t represents the output at time t; θ represents the activation function; k_{t-1} represents the output at the previous time; s_t represents the input at the current time; ς represents the update weight; υ represents the bias term. In the research requirements of table tennis trajectory prediction in this paper, it is necessary to predict the long trajectory of table tennis in the future according to the previous piece of table tennis trajectory data. Input at time t + 2, and this cycle can complete the prediction of long-term table tennis trajectory information. Given the same initial motion state, the trajectory prediction based on the continuous motion model only depends on the prediction time and is not related to the motion state of the previous and subsequent frames, that is, the error of the trajectory prediction value of each frame is independent and identically distributed. Then the joint probability density function of the error distribution of the trajectory prediction value of consecutive multiple frames is equal to the product of the probability density function of the error distribution of each frame.

In this paper, the Gaussian mixture model is used to describe the likelihood between the trajectory prediction and the trajectory observation, and the three-dimensional coordinate position of the ping-pong ball with various steps is used as the network input for experiments. In the process of network training, the network evaluates and optimizes the output predicted value according to the known table tennis trajectory information, and calculates the average three-dimensional space distance between the predicted value and the actual value based on the accuracy of the predicted coordinates. The calculation formula is as follows Eq. (9) is shown.

$$L = \sqrt{(F_1 - R_1)^2 + (F_2 - R_2)^2 + (F_3 - R_3)^2} \tag{9}$$

In formula (9), (F_1, F_2, F_3) represents the predicted coordinate value; (R_1, R_2, R_3) represents the actual coordinate value; L represents the three-dimensional space distance between them. The decay curve of flight speed with time of each category has a similar variation law. In the initial stage, the speed in the horizontal direction of the rotating table tennis ball is relatively large, and air resistance plays a major role, so the flight speed is attenuated at this stage. The approximate rotation type and flight direction of the ball are determined by calculating the flight speed in the three coordinate directions by obtaining the first five position information during the flight of the table tennis ball.

As the flight speed decays, the effect of air resistance becomes smaller and smaller until it is close to the effect of gravity. At this time, the size of the flight speed remains basically unchanged (the horizontal component continues to decrease, the vertical component increases, and the total remained largely unchanged). Then gravity plays a major role, and the flight speed starts to increase. Thus, the design of the automatic identification method of table tennis motion trajectory based on deep learning in table tennis training is completed.

3 Experimental Test

3.1 Experiment Preparation

In this experiment, three high-speed black-and-white industrial cameras are used for image acquisition. The camera model is HIKVISON/MV-CA013-21UM, the pixel is 1.3 million pixels, and global exposure is adopted, and the maximum frame number can reach 210 frames. The camera is installed on the straight line where the table tennis net is located, 1 m from the lower edge of the table, and 1.5 m above the plane where the table is located, that is, the upper side of the table. Adjust the shooting angle and focal length of the camera to ensure that the ball can be captured. The complete area of the right table top of the table. Table tennis uses standard game balls with a diameter of 40 mm and a weight of 2.7 g. The table specifications are 2.74 m long, 1.525 m wide and 0.76m high. The experimental simulation platform is Intel®Core™ i7-7700K CPU@4.20 GHz×8 processor, 16 GB memory, TITAN X (Pascal)/PCle/SSE2 graphics card, Ubuntu16.04 operating system.

The 500 ping-pong trajectories collected and the 1,700 ping-pong ball trajectories acquired by the network are used, of which 1,900 are used as training data sets and 300 are used as test data sets. The experiments set the network with stride 10 to train for 1000 epochs. According to the training step size, input the trajectory data of the first 15 frames, and use the errors at 30, 40, 50, and 60 frames to analyze the accumulation of prediction errors with the increase of the number of frames. The prediction speed of 60 frames reaches 85.17 ms, which meets the real-time requirements.

3.2 Results and Analysis

Input the initial 10 trajectory point information of each curve in the test set into the trained LSTM model, and output the three-dimensional coordinate value of the trajectory point in the future accordingly. Calculate the error between the predicted value of the trajectory point and the actual coordinate value.

In order to avoid too single experimental results, the table tennis trajectory automatic recognition method based on deep learning proposed in this paper was used as the experimental group, and the trajectory recognition method based on mean-Shift algorithm and S_Kohonen neural network was selected as the control group. Test the recognition effect of different methods, and compare the average error and total error of different recognition methods on each coordinate axis component. The results are shown in Tables 2, 3, 4 and 5.

In the measurement results of the x-axis component, the average error of the automatic recognition method based on deep learning is 21.05 mm, which is 50.56 mm and 63.14 mm lower than the trajectory recognition method based on the Mean-Shift algorithm and the S_Kohonen neural network.

In the y axis component measurement results, the average error of the automatic recognition method based on deep learning is 21.84 mm, which is 54.98 mm and 62.89 mm lower than the trajectory recognition method based on Mean-Shift algorithm and S_Kohonen neural network.

Table 2. Errors of x axis components (mm)

Testing frequency	Recognition method based on deep learning	Recognition method based on Mean-Shift algorithm	Recognition method based on S_Kohonen neural network
1	21.44	68.42	78.42
2	22.88	72.84	79.84
3	21.65	69.68	82.68
4	21.56	65.20	86.26
5	22.23	78.53	85.03
6	20.32	72.96	85.35
7	20.65	76.39	89.69
8	19.59	74.65	87.22
9	20.92	69.28	84.54
10	19.25	68.14	82.90

Table 3. Errors of y coordinate components (mm)

Testing frequency	Recognition method based on deep learning	Recognition method based on Mean-Shift algorithm	Recognition method based on S_Kohonen neural network
1	18.44	72.03	82.49
2	19.87	73.87	79.88
3	20.68	75.94	84.67
4	25.26	76.68	85.35
5	23.52	79.35	86.23
6	24.95	82.26	88.05
7	19.33	75.53	89.74
8	20.64	78.85	82.51
9	22.50	74.22	85.12
10	23.25	79.48	83.28

In the z-axis component measurement results, the average error of the automatic recognition method based on deep learning is 21.03 mm, which is 55.80 mm and 62.68 mm lower than the trajectory recognition method based on Mean-Shift algorithm and S_Kohonen neural network.

For the table tennis motion trajectory and real trajectory identified by the model, the total trajectory error of the automatic recognition method based on deep learning is 39.00 mm, which is 112.84 mm and 126.59 mm lower than the trajectory recognition

Table 4. Errors of z coordinate components (mm)

Testing frequency	Recognition method based on deep learning	Recognition method based on Mean-Shift algorithm	Recognition method based on S_Kohonen neural network
1	19.64	75.47	82.45
2	18.47	78.81	81.68
3	19.85	75.56	85.32
4	22.68	79.63	86.50
5	24.35	76.35	82.28
6	23.26	74.22	79.64
7	22.53	77.51	85.32
8	21.94	78.04	84.27
9	19.31	77.42	86.41
10	18.27	75.28	83.25

Table 5. Total error of table tennis trajectory (mm)

Testing frequency	Recognition method based on deep learning	Recognition method based on Mean-Shift algorithm	Recognition method based on S_Kohonen neural network
1	35.56	142.08	162.26
2	36.47	145.44	160.44
3	35.84	157.87	158.81
4	42.61	152.55	159.56
5	41.28	161.63	167.37
6	35.53	149.23	168.58
7	38.23	154.52	172.20
8	39.62	138.25	174.54
9	41.54	162.12	175.88
10	43.28	154.74	156.26

method based on the Mean-Shift algorithm and the S_Kohonen neural network. It can be seen from the above research results that the average error and total trajectory error of the table tennis motion trajectory automatic identification method based on deep learning are the smallest compared with the two comparison methods, so this method can be accurately used for training. Table tennis trajectory recognition in

4 Conclusion

The research on the space trajectory of flying objects has great research significance. In this paper, table tennis is taken as the research object, and the research on automatic identification of the motion trajectory is carried out. The method can reduce the average error of the components in each direction of the coordinate axis and the total error of the trajectory, and realize the accurate identification of the movement trajectory of table tennis.

In the training process of table tennis, it is inevitable that there will be some edge balls, tennis touches, etc. This system only considers some common situations. When there are special situations, it is often impossible to accurately identify and judge, thus affecting the applicability of the method. In the follow-up, the special trajectory in table tennis training can be studied to improve the practicability of the automatic trajectory identification method.

Fund Project. Natural Science Research Project of Education Department of Anhui Province "Man-machine comparative verification of table tennis based on multiple feature fusion" (KJ2021A1289).

References

1. Ma, R., Du, Z., Xiao, Y., et al.: Three dimensional trajectory reconstruction of table tennis. Comput. Syst. Appl. **30**(1), 250–255 (2021)
2. Li, J., Li, C., Xie, Y., et al.: Research on the influence of the characteristics of the racquet base plate on the table tennis motion track based on CFD method. Sports Sci. Res. **24**(2), 27–29 (2020)
3. Wang, D., Ruan, C., Yang, Y.: The service trajectory detection method of pneumatic table tennis server based on binocular vision. Autom. Instrum. **3**, 164–167 (2020)
4. Chang, Q., Liu, R.: Intelligent recognition method of bounce trajectory of table tennis under different dynamics. J. Chifeng Univ. (Nat. Sci. Ed.) **36**(04), 54–56 (2020)
5. Ji, Y.: Reverse calculation of rotation using trajectory of table tennis based on genetic algorithm. J. NanJing Sports Inst. **20**(2), 53–60 (2021)
6. Ma, R., Du, Z., Xiao, Y., et al.: Three dimensional trajectory reconstruction of table tennis. Comput. Syst. Appl. **30**(01), 250–255 (2021)
7. Zhang, X., Wang, X.: Continuous correction of moving path error of 3D obstacle space scene target. Comput. Simul. **37**(5), 420–424 (2020)
8. Wang, D., Ruan, C., Yang, Y.: The service trajectory detection method of pneumatic table tennis server based on binocular VI. Autom. Instrum. **2020**(03), 164–167 (2020)
9. Li, B., Jin, P., Wu, Z.: Design of ping-pong recognition based on S_Kohonen neural network. J. Huazhong Univ. Sci. Technol. (Nat. Sci. Ed.) **48**(3), 52–56 (2020)
10. Kong, X., Shi, J., Gao, L.: Detection and trajectory tracking of spherical objects motion. Comput. Digit. Eng. **48**(7), 1606–1610, 1681 (2020)
11. Zhang, X., Wang, W., Tang, Y., et al.: Real-time table tennis motion pattern recognition based on ESP32. Comput. Syst. Appl. **31**(6), 117–124 (2022)