

## Research Article

# Table Tennis Capture System Based on Image Recognition and Modeling

Yajun Pang <sup>1,2</sup>

<sup>1</sup>College of Physical Education, Luoyang Institute of Science and Technology, 471023 Luoyang, Henan, China

<sup>2</sup>Henan Province Engineering Research Center of Industrial Intelligent Vision, 471023 Luoyang, Henan, China

Correspondence should be addressed to Yajun Pang; [yajun.pang@lit.edu.cn](mailto:yajun.pang@lit.edu.cn)

Received 31 December 2021; Revised 11 February 2022; Accepted 17 March 2022; Published 1 April 2022

Academic Editor: Zahid Mehmood

Copyright © 2022 Yajun Pang. 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.

This article takes table tennis as the research object, mainly extracts a large number of table tennis video trajectories, and combines the kinematic analysis method to reduce the noise of the extracted target to improve the accuracy of the table tennis trajectory and obtain the table tennis trajectory. The parameter group is used to simulate the trajectory of table tennis; based on the MATLAB environment to realize the table tennis trajectory simulation that provides the initial velocity and coordinates, complete the capture of the table tennis drop point. This article uses the two-dimensional information in the image to estimate the parameter values that form the three-dimensional information and then calculates the three-dimensional information based on the estimated parameter values. The unscented Kalman filter is used to estimate the trajectory parameters and rotation parameters of table tennis, and the algorithm for calculating and updating markers in real time is proposed, which reduces the influence of calculation errors on the estimation process and improves the accuracy of parameter estimation. First, establish a table tennis movement model and choose a suitable model to describe the translation and rotation movement of the table tennis. Then, based on the extended Kalman filter algorithm and the unscented Kalman filter algorithm, the ball flight trajectory estimation algorithm is established. Finally, an actual image acquisition system is built to collect trajectory information and surface information images from the actual table tennis movement. Processing is performed to obtain the positions of the center of the ping pong ball and the marked points on the corresponding sensors. Simulation and experimental results show that the proposed algorithm can effectively estimate the trajectory, linear velocity, and angular velocity parameters of a table tennis ball, and the selection of the initial value has little effect on the algorithm.

## 1. Introduction

In recent years, with the improvement of national living standards, people have become more and more interested in the improvement of physical fitness, and various sports have also developed vigorously. As China's national ball, table tennis has become an excellent way for people to increase their physical activity. In order to improve the accuracy of table tennis hits and promote the development of table tennis, a design for the table tennis capture system is designed in this article [1–10].

The concept of deep learning model needs to be traced back to 2006. Geoffrey Hinton used neural network to complete the dimensionality reduction of the data and published the results on "Science." Since then, the concept of

deep learning has been continuously extended to other fields and has been successfully used. For example, Yann LeCun, Yoshua Bengio, and Geoffrey Hinton, three leading figures in the field of deep learning, published a comprehensive research article titled "Deep Learning" in the journal Nature in 2015, in which they launched a detailed study of deep learning. In general, the main content of deep learning is that it mainly learns various feature expressions through a model composed of multiple cascaded network layers, and it also has the characteristics of multiple abstract levels. In addition, the back propagation algorithm is also adopted to guide the machine to self-learn by changing the internal variables and explore the deeper content contained in the data sample. In fact, this method of using backpropagation or hierarchical models to expand the corresponding learning has been used

in media, such as images, videos, text, and audio. For now, the most successful training network types are deep belief networks based on DBN algorithm, generative adversarial networks based on model optimization training, long-short-term memory to solve RNN, feedforward neural network, and so on. In addition, there are also many scholars who pay attention to the backpropagation algorithm of training the network. Therefore, more efficient algorithms are developed, including Adadelta, Adam, and RMSprop [11–15].

Convolutional neural networks (CNN) is the most common type in the ANN field. Because the neural network requires a lot of data for simulation training in the initial stage and requires the computer's own hardware equipment. It is difficult to obtain a network with relatively good performance through training. However, in recent years, with the continuous advancement of GPUs and corresponding labeled data, CNN has shown better and better results in dealing with image recognition or image classification problems. It is precisely because of this advantage of CNN that it is widely used in face recognition, object recognition, and other occasions. In recent years, the successful application of convolutional neural networks in image recognition has received widespread attention (Figure 1). Common image recognition methods can generally be divided into the following three types: decision theory recognition, syntactic pattern recognition, and fuzzy pattern recognition. Among them, a major feature of syntactic pattern recognition is the use of several structural features to form a single recognized object, which can accurately describe the characteristics of the image. Suppose a picture is composed of lines, curves, polylines, and the like according to specific conventions, then, it is often combined with the knowledge of statistical decision making in mathematical statistics to reconstruct the secondary space to achieve the purpose of image recognition. Commonly used methods include similar judgment method, similar analysis method, and function classification method. Among them, fuzzy pattern recognition is mainly through training to imitate the process of identifying things in the human brain, which can not only obtain a more accurate classification of objective things but also improve the simplicity of the recognition module. Therefore, it is also a method for the above two recognition methods [16–20].

LeNet is a more traditional convolutional neural network with a model depth of only five layers. Scholars such as Alex Krizhevsky put forward a series of views on this in 2012. They believe that although the current convolutional neural network technology can be competent for some data sets (such as MNIST and the like), the categories of these data sets are few. Therefore, they are more inclined to use other training data sets to train the model, such as ImageNet, to realize the trained convolutional neural network model, so that it has a stronger depth and breadth. With the further development of CNN, in order to improve the performance of neural networks, scholars often achieve this by increasing the number of layers of the model and the number of neurons in each layer [21–25]. However, in this method, in order to avoid the overfitting phenomenon of the model, many training samples are usually required, which will cause the model parameters to increase greatly. Therefore, in order

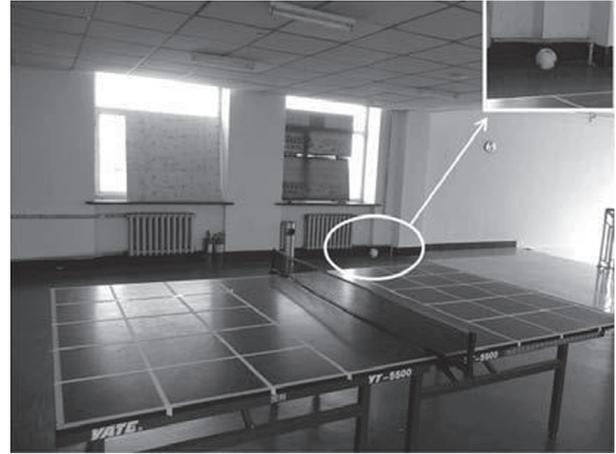


FIGURE 1: Image recognition.

to solve such problems, GoogLeNet uses the inception module. Its main principle is to control the cumbersome degree of calculation of model training, and it ingeniously overcomes the problem of gradient failure in reverse transfer. Computer scientist Andrew and scholar Karen of the University of Oxford in the United Kingdom have developed another DNN (deep neural Network) model: VGGNet. They used a convolution filter with a  $3 \times 3$  receptive field to produce the output of different models. After comparative analysis, the performance difference between convolutional neural network and network depth is studied systematically. Kaiming He et al. proposed the ResNet model in 2015 and won the ILSVRC championship that year. Kaiming He believed that the difficulty of training network parameters increases as the depth of the network increases. Once the model depth index exceeds a certain value, the problem of gradient explosion and gradient disappearance will occur, and the degree is positively correlated with the network depth, but the accuracy rate is reduced. Therefore, to solve this problem, they studied a new method, which includes adding connections between nonadjacent layers, skipping some layers in the network, changing the original stacking (plain network) form of the traditional network, join the identity mapping layer, and so on [26–30]. Existing convolutional neural network models constantly change the depth and width of the model, and at the same time, they continuously optimize the model architecture, so that its performance is significantly improved. In the image recognition task, as the current best-performing convolutional neural network, its classification performance on multiple image data sets (for example, the ImageNet data set) has been comparable to human capabilities. Therefore, on this basis, Hinton proposed a capsule network in October 2017. It uses neuron vectors to replace the single neuron node commonly used in traditional convolutional neural networks. Compared with traditional recognition methods, adding a deep learning algorithm module in the image processing field has the following advantages: the algorithm itself can complete the analysis of big data independently and extract features from it, without the need for manpower to complete this tedious task. Classification and recognition

work and feature extraction work can be coordinated to control the running consumables of the algorithm [31–33].

## 2. Model building

For rotating flying objects, the state of motion generally includes position, flight speed, and rotation speed. In this article, the three-dimensional position of the table tennis ball can be directly given by the external visual positioning system, so its motion state specifically refers to the flight speed and rotation speed, and the motion state space specifically refers to the high space composed of the flight speed state space and the rotation speed state space. Motion state estimation refers to the use of physical model-based methods or nonmodel purely mathematical methods to solve the motion state of a rotating flying object based on the observation information of a trajectory position time series; trajectory prediction is to bring the estimated initial motion state into the motion. The model predicts the movement state of the flying object over a period. The distance between the predicted value of the trajectory and the observed value of the actual trajectory is a direct standard to measure the accuracy of motion state estimation and trajectory prediction. It is not difficult to see that the key to estimating the state of the motion and predicting the trajectory lies in the accuracy of the rotating flight motion modeling. The changes of marking points on the surface of the table tennis ball as it rotates is shown in Figure 2.

As a typical example of rotating flying objects, table tennis balls are of small quality, fast in speed, and have high-speed rotation. The trajectory is easily affected by rotation and offset. Therefore, this article uses table tennis as a rotating flight motion model, motion state estimation, and trajectory prediction. When a rotating ping pong ball is flying in the air, it is mainly affected by the combined effects of gravity, air buoyancy, air resistance, and Magnus force. The air buoyancy is too small compared with the other three forces and can basically be ignored. According to aerodynamics, when a ping pong ball is flying in the air, it will rub against the air and generate air resistance. Its magnitude is proportional to the square of the flying speed of the ball, and the direction is opposite to the flying speed. Affected by the combination of flight speed and rotation speed, the relative speed of the surface of the ping pong ball is not consistent with that of the air. According to fluid mechanics, the pressure of air on both sides of the ping pong ball is not the same, resulting in the Magnus force. Ideally, the Magnus force is proportional to the outer product of flight speed and rotation speed. The comprehensive force of a rotating table tennis ball is nonlinearly related to the motion state, and the motion model is the double integral action of the comprehensive force in time, so the motion model has off-order nonlinear characteristics. This brings great challenges to the motion modeling, model parameter identification, motion state estimation, and trajectory prediction of rotating table tennis.

Some existing methods do not use the motion model based on force analysis but directly use pure mathematical methods such as polynomial fitting and local linear

regression BP neural network to fit the time series of the initial trajectory position of the pawn and the ball and the position of the ping pong ball hitting point and the mapping relationship between the state of motion. This type of method does not use the constraint information of the motion model and can only adapt to a small type of trajectory with basically no rotation speed and flying speed in a specific range, and the prediction accuracy cannot meet the servo requirements of the table tennis robot in the full-state space. Some have established a motion model based on force analysis but ignore the influence of rotation. Then, by constructing an extended Kalman filter to filter the real-time observations of the trajectory position time series, the least-square estimation of the motion state is obtained. Although the constraint information of the motion model is used when estimating the motion state, the motion model can only adapt to the situation of nonrotating or low-speed rotating table tennis, and the accuracy of motion state estimation and trajectory prediction will decrease as the speed of the table tennis increases. The best existing method proposes a motion model that fully considers the influence of rotation and derives the calculation method of the rotation speed according to the constraint relationship between the flight speed and the rotation speed in the motion model. Although this type of method effectively uses the constraint information of the motion model for the estimation of the rotation speed, a second-order polynomial to fit and derive the observation information of the trajectory position time series in three dimensions is adopted to obtain two adjacent table tennis balls. In fact, when the rotation speed is high, the motion trajectory is coupled with each other in three dimensions. It is inaccurate to regard them as independent motions that do not affect each other and to separately fit the trajectory, and iteratively predicting the trajectory through the discrete motion model will introduce a cutoff point. Therefore, this type of method cannot be effectively applied to table tennis balls with a high rotation speed. The coordinate system is shown in Figure 3.

*2.1. The Table Tennis Capture System.* The table tennis capture system proposed is mainly composed of three parts:

- (i) Extraction of the trajectory coordinates of the table tennis. After analyzing the round shape, high-speed movement, and small size of the target object and comparing multiple image processing methods at the same time, we chose to use the circle detection method to identify the target object's trajectory coordinates. The circle detection of this system is realized by the Hough circle transformation algorithm. The Hough circles function is mainly used to obtain the center coordinates of the circle. In order to obtain more accurate trajectory coordinates, the function parameters set by this system are HoughCircles (src, circles, HOUGH\_GRADIENT, 1.5, 150, 130, 30, 1, 10).
- (ii) Background noise reduction of target trajectory. After obtaining the original coordinate samples of the target trajectory, the RANSAC algorithm is used to

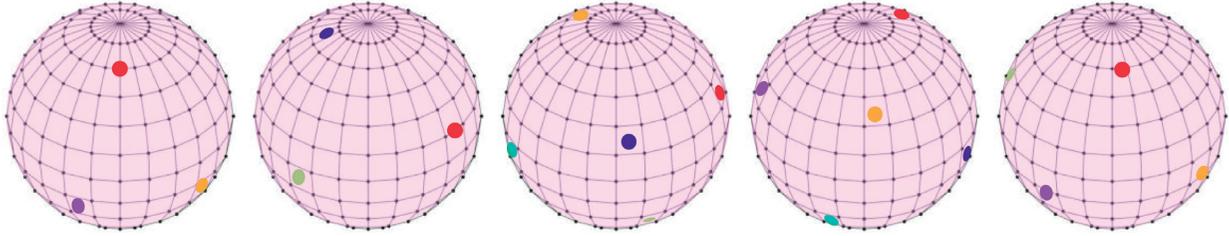


FIGURE 2: Changes of marking points on the surface of the table tennis ball as it rotates.

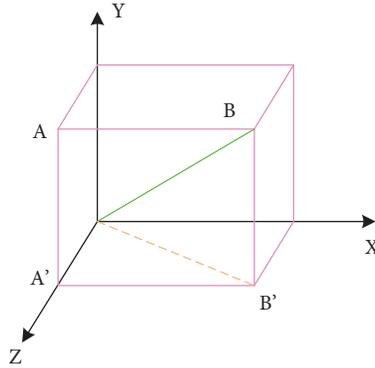


FIGURE 3: Coordinate system.

denoise the data. Because the trajectory of table tennis conforms to the laws of motion mechanics, a comprehensive trajectory model is established through the mechanical analysis of the trajectory of the table tennis, combined with the known trajectory samples. The RANSAC algorithm is used to reduce the noise of the obtained image trajectory coordinates. The above operations are performed several iterations to get more accurate trajectory data.

- (iii) Establishment of mathematical model of trajectory parameters.

2.2. Establishment of reference Coordinates. Through data retrieval, we get the size parameters of the standard table tennis case: the table is 2.74-m long, 1.525-m wide, 76-cm high from the ground, and 15.25-cm high from the net. Taking the upper left corner of the ping pong table as the origin, according to the camera position, we establish the following coordinate system (in centimeters): the coordinates of the four vertices of the ping pong table:

$$(0, 0, 0), (274, 0, 0), (0, 152.5, 0), (274, 152.5, 0). \quad (1)$$

Camera coordinates

$$(137, -2, 0). \quad (2)$$

Plane coordinates of the floor:

$$z = -76. \quad (3)$$

The plane coordinates of the table tennis case:

$$z = 0 (0 < x < 274, 0 < y < 152.5). \quad (4)$$

Straight line coordinates of the center line:

$$l = \begin{cases} x = 137, \\ z = 0. \end{cases} \quad (5)$$

2.2.1. The Real-Time Trajectory Coordinates of the Ping Pong Ball. We first convert the pixel coordinates obtained after capturing noise reduction into world coordinates. According to the formula, the process is as follows:

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{xd} & 0 & u_0 \\ 0 & \frac{1}{yd} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad (6)$$

$$\begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = L \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} l_1 & l_2 & l_3 & l_4 \\ l_5 & l_6 & l_7 & l_8 \\ l_9 & l_{10} & l_{11} & l_{12} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}.$$

Establish the conversion matrix from world coordinates to image coordinates:

$$P = \begin{bmatrix} P_{00} & P_{01} & P_{02} & P_{03} \\ P_{10} & P_{11} & P_{12} & P_{13} \\ P_{20} & P_{21} & P_{22} & P_{23} \end{bmatrix}. \quad (7)$$

Then,

$$\tilde{w} = P\tilde{X}. \quad (8)$$

According to the established coordinate system, the conversion matrix  $P$  is obtained by substituting relevant parameters. Through the conversion matrix, the pixel coordinate data imported into MATLAB is sequentially converted into world coordinates, and the real-time trajectory coordinates of the ping pong ball can be obtained. The flowing chart is shown in Figure 4.

**2.2.2. The Trajectory Model of Table Tennis in the Air.** According to the force of table tennis, through physical analysis, the trajectory equation of table tennis in the air can be obtained as follows, in the horizontal direction:

$$x = \begin{cases} \frac{m}{k} \ln \left( 1 + \frac{kv_{x0}t}{m} \right) + x_0, v_{x0} > 0, \\ -\frac{m}{k} \ln \left( 1 - \frac{kv_{x0}t}{m} \right) + x_0, v_{x0} < 0, \end{cases} \quad (9)$$

$$y = \begin{cases} \frac{m}{k} \ln \left( 1 + \frac{kv_{y0}t}{m} \right) + y_0, v_{y0} > 0, \\ -\frac{m}{k} \ln \left( 1 - \frac{kv_{y0}t}{m} \right) + y_0, v_{y0} < 0. \end{cases}$$

In the vertical direction:

$$z = \begin{cases} \frac{m}{k} \ln \left| \cos \left( \sqrt{\frac{kg}{m}} t \right) + v_{z0} \sqrt{\frac{kg}{m}} \sin \left( \sqrt{\frac{kg}{m}} t \right) \right| + z_0, v_{z0} > 0, \\ -\frac{m}{2k} \ln \frac{\left( 1 + e^{2t\sqrt{kg/m}} \right)^2}{4e^{2t\sqrt{kg/m}}} + z_0 + v_{z0}t, v_{z0} < 0. \end{cases} \quad (10)$$

The components of the initial velocity of the table tennis on the  $X$ ,  $Y$ , and  $Z$  coordinate axes.

**2.2.3. Fitting of Table Tennis Trajectory Equation.** In MATLAB, according to the above table tennis trajectory equation in the air, the least square method is used to fit the table tennis trajectory. Along the three directions of  $X$ ,  $Y$ ,  $Z$ , respectively, according to the trajectory equation (formula) to fit, the initial velocity component of the ping pong ball on the coordinate axis is regarded as an unknown number, and the trajectory of the ping pong ball in this direction can be obtained. Combining the predicted trajectories in three

directions, the predicted trajectory of the table tennis ball in space can be obtained.

**2.3. Judgment of the Rebound of the Collision between the Table Tennis Ball and the Table.** If the table tennis collides with the ball case during the movement, the fitting prediction will be made according to the new trajectory after the rebound. Currently, the speed of the rebound is a new variable, that is, if the observed data of the table tennis satisfies

$$z = 0 (0 < x < 274, 0 < y < 152.5). \quad (11)$$

Then remove the previous data and perform fitting based on the new data. Prediction of the placement of table tennis: the predicted placement of the table tennis is the intersection of the predicted trajectory and the ground plane. In the above-mentioned coordinate system, that is, the value of  $x$  and  $y$  when  $z = -76$ .

**2.4. Model Checking and Prediction Accuracy.** This article has conducted many experiments with the above model, and the experimental results are as follows:

- (i) Through experiments, we found that this system runs faster, has real-time performance, and has certain predictive capabilities; it can better capture, extract, and transform data, so it has better performance. Good recording and real-time monitoring functions.
- (ii) In terms of data accuracy, the prediction results of this system and the actual landing error are within a reasonable range. By analyzing the errors in all valid experiments, we can think that the confidence interval for the prediction error within  $\pm 2$  cm is 0.99. Therefore, this system has strong practicability.

After analysis, the table tennis capture system can be put into practical use, but there is still a certain error. The cause of the error is table tennis has the characteristics of small targets and fast movement speed. Under the shooting of high-speed cameras, some target samples are distorted and causes the loss of some trajectory coordinates; it is difficult to completely filter the background noise of table tennis image recognition; due to environmental factors, before using the table tennis capture system, the mathematical model parameters should be adjusted accordingly; otherwise, it will be A large error occurred. The samples are compared in Figure 5.

### 3. Simulation

The above content has completed the estimation of the motion trajectory parameters and the estimation of the ball rotation parameters for the table tennis motion model and the camera working model. In order to verify whether the algorithm is reasonable and can meet the actual table tennis motion parameter estimation requirements, this section builds the image. The acquisition system platform uses two cameras to separately collect table tennis motion

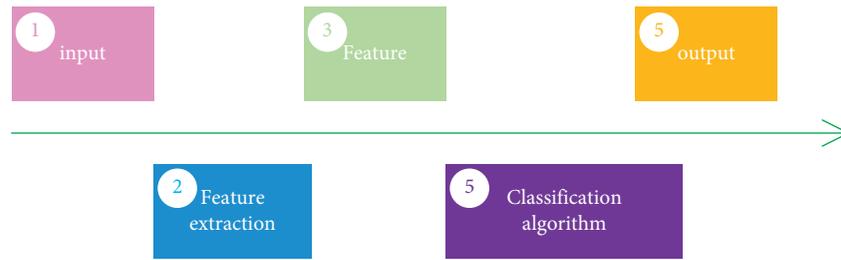


FIGURE 4: Flowing chart.

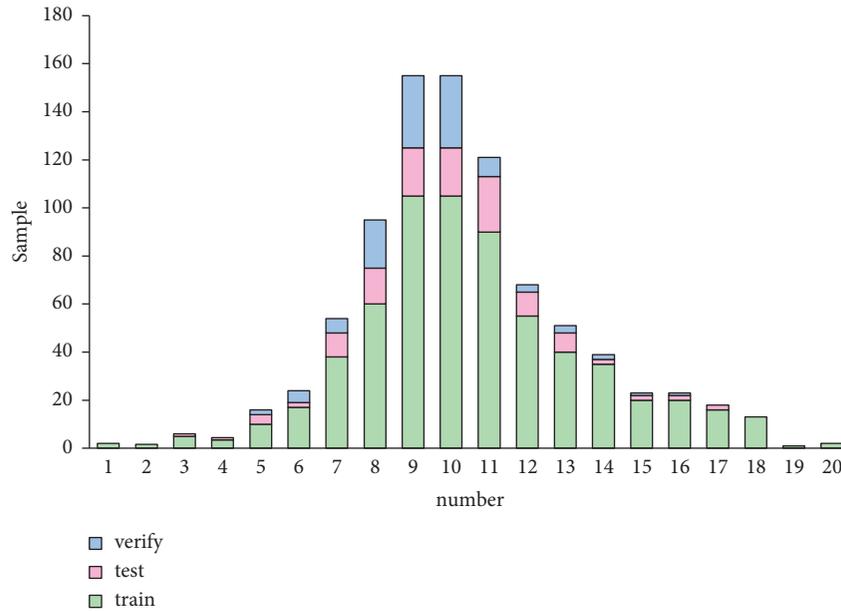


FIGURE 5: Sample comparison.

information images and rotation information images and performs offline processing and analysis on the two sets of images obtained, to obtain the actual measurement value matrix that can be used in the algorithm, and adjust the simulation algorithm Parameters are processed and analyzed to verify the correctness of the algorithm.

Schematic diagram of the image acquisition system. The system uses an external vision processor to connect two cameras and a PC upper computer. The two cameras are arranged in a nonparallel arrangement. The vision processor is CVS-1456. The two cameras are connected to the visual processor CVS1456 via IEEE1394 cables, respectively, and the visual processor is connected to the PC upper computer with a twisted-pair network cable. The field of view of camera 1 is the entire table tennis table, so that it can collect the entire trajectory information of table tennis, and the field of view of camera 2 is the second half of the trajectory of table tennis, mainly extracting the surface information of the image, which is set by the image acquisition program. The image acquisition time and image acquisition frequency of the two cameras, the images acquired by the two cameras are saved to the PC through the CVS. The fitting curve graph is shown in Figure 6.

The core of the image acquisition system is the vision processor CVS-1456, which is a compact vision processing system. The vision processor is a simple to use and distributed real-time image processing system. The real-time image processing system is embedded in the CVS. For real-time high-speed image acquisition and processing, it can acquire, process, and display images from IEEE1394 cameras that comply with the IIDC1394 digital camera specifications. A variety of digital input/output ports are also provided in CVS for communication with external devices. The CVS-1456 is connected to the PC via Ethernet, which can display measurement results and status information, and it can also perform system and network settings on the vision processor. The visual processor CVS-1456 has 2000×2000 image resolution, 1623 Mbps processing speed, and 256 MB flash memory, which can store a large amount of image data to meet image processing needs.

There are 3 IEEE1394a interfaces on the CVS-1456 panel to connect to the camera. There are currently more than 80 IEEE1394 cameras to choose from, and up to three IEEE1394 cameras can be connected at the same time. The bandwidth provided by the IEEE 1394 bus is a fixed amount, which is shared by the ports connected to the camera. The bandwidth provided by each port depends on the amount of

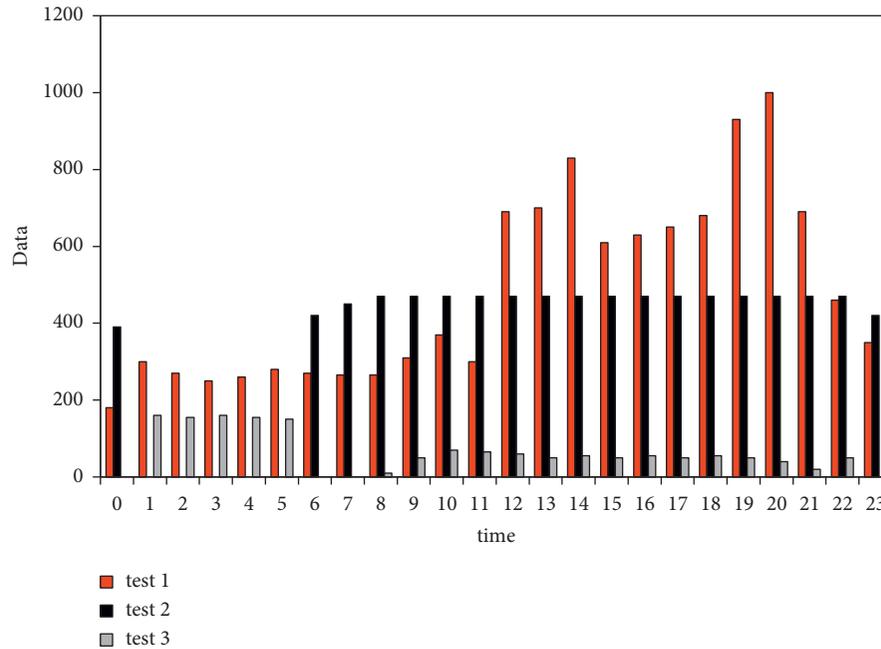


FIGURE 6: Fitting curve graph.

bandwidth required by each camera. High frame rate and large-size images require higher data transmission. Rate requires more bandwidth.

The image data transmission interface, communication module and image processing module are integrated in the CVS1456. There are multiple input and output terminals, which can output or receive trigger signals to control the work of the cameras connected to the processor and realize the operation of multiple cameras, synchronous control. There are two trigger signal ports on the panel. The independent signal input port can be connected to external devices, such as sensors or start/stop buttons, and the TTL signal output port is used to send signals to trigger external devices, such as cameras or lighting equipment. Therefore, the use of the visual processor eliminates the need for a separate processing unit, which speeds up the image acquisition speed and saves time and cost.

Among them, camera 1 is used to collect position information of the ping pong ball, and the lower resolution can meet the requirements, and camera 2 is mainly used to collect the information of the marked points on the surface of the ping pong ball, so a higher resolution is selected. The two cameras are connected to the vision processor, the sampling frequency is 100 Hz, and the acquisition is triggered synchronously by the software, and a set of sphere center position images and a set of marked point information images are obtained, respectively. The two cameras have the following characteristics:

- (1) All adopt CMOS sensor, which has higher sensitivity and shorter exposure time compared with CCD sensor, which is beneficial to realize high-speed image acquisition.

- (2) It has three trigger modes: free mode, software starts, and hardware trigger, which can realize continuous acquisition and trigger acquisition.
- (3) The data transmission interface is a standard IEEE1394 interface, which can realize plug and play.
- (4) Support multiresolution image acquisition, support custom image acquisition window, speedup image acquisition speed.

In the simulation system, the position and posture of the two cameras are known, but in practice, only the position of the camera is fixed when working. When building the system, you need to adjust the position and posture of the camera according to the actual table and ball trajectory. To meet the requirements, these pose parameters are the basis for estimating the trajectory parameters and rotation parameters of the table tennis ball. However, if these pose parameters are obtained by direct measurement methods, there are large errors. Therefore, according to the camera imaging rules, some feature points with known coordinate values in the space are used to correspond to the coordinates of the image points on the surface of the camera sensor. The corresponding relationship between the two can calculate the position and attitude parameters of the camera. This method is more accurate and more convenient than direct measurement. It is called camera calibration.

Camera calibration is mainly divided into three categories: traditional calibration method, camera self-calibration method, and camera calibration method based on active vision. The traditional calibration method uses a reference object or a calibration object with known geometric information and position information, selects the feature points on the reference object or calibration object, and calculates the pose parameters of the camera according to

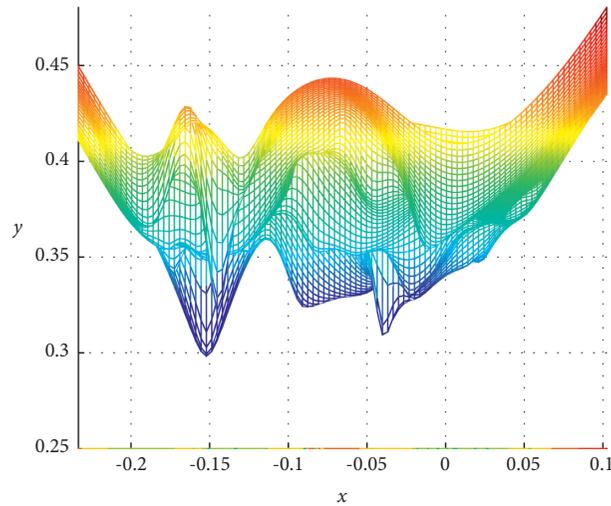


FIGURE 7: Predicted result.

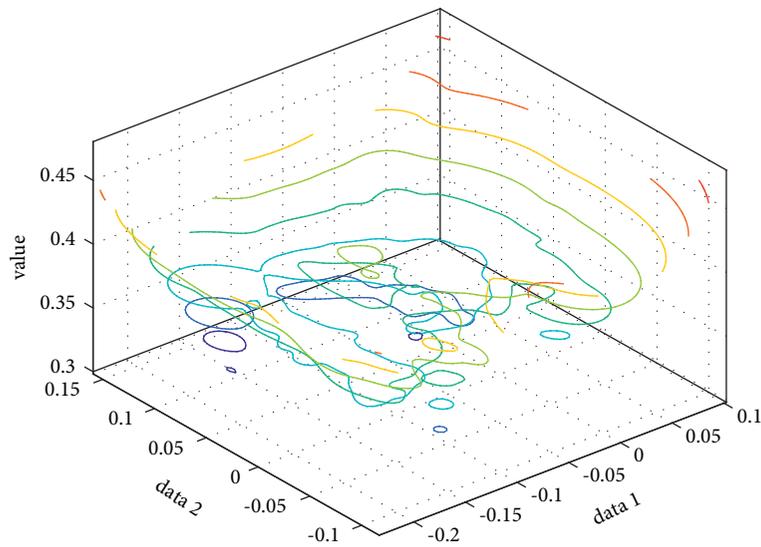


FIGURE 8: Contour.

the corresponding relationship between the feature points and the image points. The operation is simple, and the accuracy is easy to guarantee. The predicted result is shown in Figure 7.

The camera self-calibration method does not need the three-dimensional information of the points in the space and uses the camera's motion constraint relationship or scene constraint relationship to calculate the camera's pose parameters. This method is flexible in application and is mainly used for camera online calibration and real-time calibration.

The calibration method based on active vision is to control the camera to perform simple linear motion with known motion information and to linearly solve its internal parameters based on the image information obtained by the moving camera, without reference objects or calibration objects, and is characterized by simple algorithms and robustness. It is highly robust, but this method requires a high-

precision vision system platform, the system is complex, and the cost is high.

The system established in this article mainly performs online image acquisition, offline data processing, and analysis. The parameter estimation method adopted makes it unnecessary to calibrate the two cameras. Therefore, the two cameras can be calibrated separately, and because there is no need to perform online parameter estimation, in the process of image acquisition, there is no requirement for real-time camera calibration. We only need to adjust the appropriate camera position and posture to ensure the viewing angle range of each camera. Therefore, comparing the advantages and disadvantages of various methods and the characteristics of the image acquisition system we have built, we choose the traditional calibration method with the aid of calibration objects.

Through the image acquisition system, two sets of pictures are obtained, one is the panoramic image of the

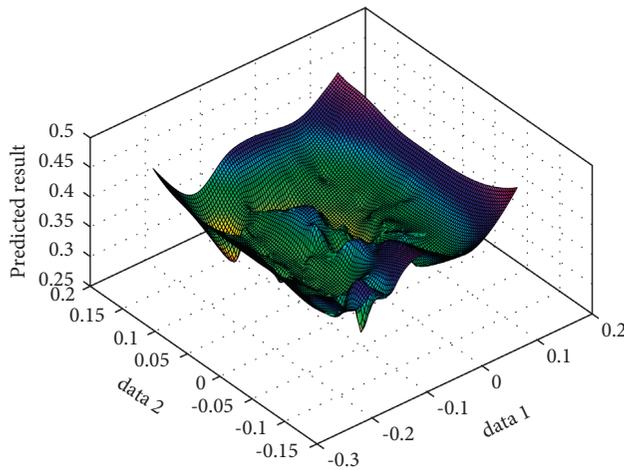


FIGURE 9: Predicted result.

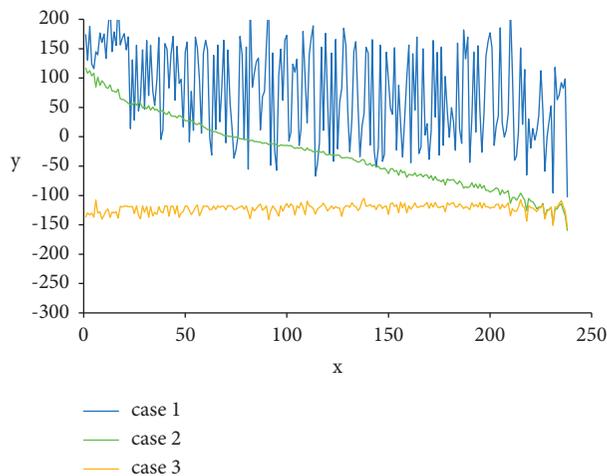


FIGURE 10: Case comparison.

trajectory information of table tennis collected by camera 1, and the other is the surface information of the ping pong ball taken by camera 2 about the position of the marked point. The image of the ping pong ball on the sensor is a circle. To obtain the position of the center of the ping pong ball more accurately, the circle that the ball is imaged on the sensor must be separated and then the center of the sphere must be calculated. Therefore, for the panoramic information collected by camera 1, we need to obtain the image coordinates of the center of the sphere through image processing. Because the size of the ping pong ball is small, its imaging on the sensor is also small, and the marking point information on the surface is not easy to separate from the image. Therefore, in offline image processing, this process can be simplified as directly from the image. Extract the position of the marker from the image and read its position information. The contour is shown in Figure 8.

The edge is the most basic feature of the object in the image, and it contains important information about the detected target. Especially in complex background, only the edge of the object can be easily distinguished from the background, and irrelevant information can be eliminated.

Reduce the amount of data, which is the most basic and most versatile application in machine vision, image processing, and other fields. Edges are areas that reflect significant changes in attributes. These changes include brightness, texture, color, geometric attributes, and other aspects of the image. The image processing concentration is expressed as a sharp change in the image gray level. The image gray level function gradient can be used. The Canny operator is one of the most widely used algorithms in the direction of edge detection. It uses Gaussian function to smooth the image and calculates the point with the local maximum gradient amplitude by first-order differentiation. The predicted result is shown in Figure 9, and different cases are compared in Figure 10.

#### 4. Conclusion

Aiming at problems such as table tennis trajectory prediction and rotation parameter estimation, this article adopts a parameter estimation method that is different from the traditional state estimation method, that is, no 3D reconstruction is required, and the required 3D information is estimated only based on the 2D image information captured by the camera. The three-dimensional information is calculated according to the estimated parameter values. This article also designs the calculation method of the three-dimensional coordinate value of the mark point according to the spatial geometric relationship and the constraint relationship of each mark point and reduces the influence of calculation error on the estimation of the rotation parameter through the real-time update method of the mark point position.

The research content and results of this article are summarized as follows:

- (1) The overall simulation environment is constructed, including the mathematical model of table tennis and the visual acquisition system.
- (2) Based on the extended Kalman filter and the unscented Kalman filter, the table tennis trajectory parameter estimation method is derived, and the estimation effect is compared, and the influence of the selection of each parameter on the estimation result is analyzed and compared. Based on the result of trajectory prediction, an algorithm for mark point calculation and real-time update is proposed. Also, the estimation method of table tennis rotation parameters is derived based on the unscented Kalman filter. Combining the trajectory prediction part and the rotation parameter estimation part, a complete motion estimation algorithm is constituted.
- (3) An experimental verification platform is established, and the experimental results verify the effectiveness of the proposed algorithm.

#### Data Availability

\*The data set can be accessed upon request.

## Conflicts of Interest

The author declares no conflicts of interest.

## Acknowledgments

This work was supported by the project source: Science and Technology Projects of Henan Science and Technology, project name: Assessment of Taichi Action Based on Vision Transformer, Department No. 222102320016.

## References

- [1] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, Las Vegas, NV, USA, 2016.
- [2] K. Muelling, A. Boularias, B. Mohler, B. Schölkopf, and J. Peters, "Learning strategies in table tennis using inverse reinforcement learning," *Biological Cybernetics*, vol. 108, no. 5, pp. 603–619, 2014.
- [3] S. S. Tabrizi, S. Pashazadeh, and V. Javani, "Comparative study of table tennis forehand strokes classification using deep learning and SVM," *IEEE Sensors Journal*, vol. 20, no. 22, pp. 13552–13561, 2020.
- [4] C. Jordan, P. Renaud, L. Mascarilla, and T. Benoit, "Table Tennis ball kinematic parameters estimation from non-intrusive single-view videos," in *Proceedings of the 021 International Conference on Content-Based Multimedia Indexing (CBMI)*, pp. 1–6, Lille, France, 2021.
- [5] J. Kim, M. Ra, H. Lee, J. Kim, and W.-Y. Kim, "Precise 3D baseball pitching trajectory estimation using multiple unsynchronized cameras," *IEEE Access*, vol. 7, pp. 166463–166475, 2019.
- [6] H.-I. Lin, Z. Yu, and Y.-C. Huang, "Ball tracking and trajectory prediction for table-tennis robots," *Sensors*, vol. 20, no. 2, pp. 333–356, 2020.
- [7] O. Koç, G. Maeda, and J. Peters, "Online optimal trajectory generation for robot table tennis," *Robotics and Autonomous Systems*, vol. 105, pp. 121–137, 2018.
- [8] C. He, X. Zhang, Y. Gui, Y. Liu, and W. Zhang, "Mathematical modeling and simulation of table tennis trajectory based on digital video image processing," *Image Processing based on Partial Differential Equations*, vol. 2021, p. 11, Article ID 7045445, 2021.
- [9] A. G. Melo, M. F. Pinto, A. L. M. Marcato, I. Z. Biundini, and N. M. S. Rocha, "Low-cost trajectory-based ball detection for impact indication and recording," *Journal of Control, Automation and Electrical Systems*, vol. 32, no. 2, pp. 367–377, 2021.
- [10] 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.
- [11] A. M. Fathollahi-Fard and M. Hajiaghaei-Keshteli, "A stochastic multi-objective model for a closed-loop supply chain with environmental considerations," *Applied Soft Computing*, vol. 69, pp. 232–249, 2018.
- [12] B. Lindemann, T. Müller, H. Vietz, N. Jazdi, and M. Weyrich, "A survey on long short-term memory networks for time series prediction," *Procedia CIRP*, vol. 99, no. 2, pp. 650–655, 2021.
- [13] X. Wang, T. Xin, H. Wang, L. Zhu, and D. Cui, "A generative adversarial network based learning approach to the autonomous decision making of high-speed trains," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 3, pp. 2399–2412, 2022.
- [14] W. Deng, H. Liu, J. Xu, H. Zhao, and Y. Song, "An improved quantum-inspired differential evolution algorithm for deep Belief network," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 10, pp. 7319–7327, 2020.
- [15] M. Johnson, M. Schuster, Q. V. Le et al., "Google's multi-lingual neural machine translation system: enabling zero-shot translation," *Transactions of the Association for Computational Linguistics*, vol. 5, no. 1, pp. 339–351, 2017.
- [16] Y. He and G. Ding, "Deep transfer learning for image emotion analysis: reducing marginal and joint distribution discrepancies together," *Neural Processing Letters*, vol. 51, no. 3, pp. 2077–2086, 2020.
- [17] X. Wang, X. Yu, L. Guo, F. Liu, and L. Xu, "Student performance prediction with short-term sequential campus behaviors," *Information*, vol. 11, no. 4, p. 101, 2020.
- [18] Q. Guo, Z. Zhu, Q. Lu, D. Zhang, and W. Wu, "A dynamic emotional session generation model based on Seq2Seq and a dictionary-based attention mechanism," *Applied Sciences*, vol. 10, no. 6, pp. 1–10, 2020.
- [19] P. Li, M. Wang, and J. Wang, "Named entity translation method based on machine translation lexicon," *Neural Computing & Applications*, vol. 33, no. 9, pp. 3977–3985, 2021.
- [20] J. Frazer, P. Notin, M. Dias et al., "Disease variant prediction with deep generative models of evolutionary data," *Nature*, vol. 599, no. 7883, pp. 91–95, 2021.
- [21] T. Moon, T. I. Ahn, and J. E. Son, "Long short-term memory for a model-free estimation of macronutrient ion concentrations of root-zone in closed-loop soilless cultures," *Plant Methods*, vol. 15, no. 1, pp. 59–12, 2019.
- [22] N. Pourdamghani and K. Knight, "Neighbors helping the poor: improving low-resource machine translation using related languages," *Machine Translation*, vol. 33, no. 3, pp. 239–258, 2019.
- [23] L. Bote-Curiel, S. Muñoz-Romero, A. Gerrero-Curieses, and J. L. Rojo-Álvarez, "Deep learning and big data in healthcare: a double review for critical beginners," *Applied Sciences*, vol. 9, no. 11, pp. 1–11, 2019.
- [24] J. Zhang and T. Matsumoto, "Corpus augmentation for neural machine translation with Chinese-Japanese parallel corpora," *Applied Sciences*, vol. 9, no. 10, pp. 1–12, 2019.
- [25] Y. Chen, Y. Ma, X. Mao, and Q. Li, "Multi-task learning for abstractive and extractive summarization," *Data Science and Engineering*, vol. 4, no. 1, pp. 14–23, 2019.
- [26] P. Zhou and Z. Jiang, "Self-organizing map neural network (SOM) downscaling method to simulate daily precipitation in the Yangtze and Huaihe River Basin," *Climatic and Environmental Research*, vol. 21, no. 5, pp. 512–524, 2016.
- [27] X. Xiao, "Analysis on the employment psychological problems and adjustment of retired athletes in the process of career transformation," *Modern Vocational Education*, vol. 5, no. 12, pp. 216–217, 2018.
- [28] S. Sahoo and M. K. Jha, "Pattern recognition in lithology classification: modeling using neural networks, self-organizing maps and genetic algorithms," *Hydrogeology Journal*, vol. 25, no. 2, pp. 311–330, 2016.
- [29] L. Chen and W. Wang, "Analysis of technical features in basketball video based on deep learning algorithm," *Signal Processing: Image Communication*, vol. 83, Article ID 115786, 2020.
- [30] J. Pang, "Research on the evaluation model of sports training adaptation based on self-organizing neural network," *Journal*

- of Nanjing Institute of Physical Education*, vol. 16, no. 1, pp. 74–77, 2017.
- [31] G. Querzola, C. Lovati, C. Mariani, and L. Pantoni, “A semi-quantitative sport-specific assessment of recurrent traumatic brain injury: the TraQ questionnaire and its application in American football,” *Neurological Sciences*, vol. 40, no. 9, pp. 1909–1915, 2019.
- [32] J. Wang, X. Luo, and H. Yan, “Correlation analysis between injuries and functional movement screening for athletes of the National Shooting Team,” *Journal of Capital Institute of Physical Education*, vol. 5, no. 4, pp. 352–355, 2016.
- [33] G. Ma, “Research on the design of juvenile football players’ sports injury prediction model,” *Automation Technology and Application*, vol. 277, no. 7, pp. 141–144, 2018.