



Application of optical motion capture based on multimodal sensors in badminton player motion recognition system

Tiejun Zhang¹

Received: 17 October 2023 / Accepted: 17 November 2023 / Published online: 27 December 2023
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract

In the motion recognition system, the motion track of athletes can be obtained in real time and accurately by using the light motion capture technology, which provides strong support for the training and competition of athletes. The purpose of this study is to explore the application of light motion capture technology based on multi-mode sensor in badminton players' motion recognition system, so as to improve the athletes' training effect and competitive level. In this paper, multi-modal sensors combined with light motion capture technology are used to obtain real-time athletes' motion data through sensors installed on athletes, and the data is analyzed and processed by algorithms. The position of athletes is tracked by optical sensors to obtain their motion trajectory. The research results show that the light motion capture technology based on multi-modal sensors can accurately identify the movement posture and action of badminton players, real-time monitoring of athletes' movement status, and provide timely feedback and guidance. Through the application of sports recognition system, the training effect of athletes has been significantly improved, and the competitive level has been effectively improved.

Keywords Multi-mode sensor · Light motion capture technology · Badminton · An athlete · Motion recognition

1 Introduction

With the development of science and technology and the demand of sports training, the application of sports recognition system in badminton becomes more and more important. The traditional motion recognition system mainly relies on image analysis and sensor technology, but there are some problems such as low precision and large delay. As a new motion recognition technology, light motion capture technology has the advantages of high accuracy and strong real-time performance, so it has great potential in badminton players' motion recognition. Optical motion capture technology uses equipment such as high-speed cameras and infrared sensors to record and measure the position and motion state of moving objects in real time. By capturing the movement track

✉ Tiejun Zhang
zhangtiejun2020@163.com

¹ Henan University of Economics and Law, Zhengzhou 450046, China

and posture of athletes, accurate movement data and analysis results can be provided to provide comprehensive guidance and evaluation for the training and competition of athletes. However, at present, the application of light motion capture technology in badminton players' motion recognition system is relatively small. In today's Big data era, the sports data of athletes has also received the attention of researchers. Nowadays, in most sports events, there are professionals who collect sports data from athletes during exercise (Bohm et al. 2016). The sports characteristics, habits, and other available information of athletes can be obtained through statistical analysis of their sports data. After obtaining this information, it is possible to identify and fill in any gaps for each athlete, in order to tailor the training process accordingly and improve their competitive level. Badminton, as a fast-paced and technically demanding sport, has high requirements for athletes' swing movements (Phomsoupha and Laffaye 2015). However, currently there is relatively little research on the swing movement of badminton, and most of it remains in the academic research stage and has not been widely applied. Therefore, if there is a breakthrough in the recognition algorithm of badminton swing movements and it is put into practical application, it will be a great progress. By collecting, processing, and analyzing data on the swing movements of badminton players, personalized training suggestions and guidance can be provided for athletes (Chu and Li 2021). The characteristics, power distribution, and speed changes of athletes' swing movements can be obtained through motion data, which helps athletes understand the problems in their swing movements and make targeted improvements. By comparing and analyzing the swing movements of outstanding athletes, it can also help athletes learn and imitate their technical characteristics, further improving their technical level (Ting et al. 2015). The practical application of swing motion recognition algorithms can also help referees accurately determine whether an athlete's swing motion complies with the rules in competitions. By conducting real-time analysis of sports data, violations can be detected in a timely manner, ensuring the fairness and impartiality of the competition (Liu 2022).

In order to apply the badminton swing motion recognition algorithm to practical scenarios, this paper designs a badminton player motion recognition and tracking system based on multimodal sensors and deep learning technology (Ting et al. 2015). The system can collect real-time data on athletes' swing movements and transmit the data to smartphones or other smart terminals for processing and display. In the system, multiple sensors are worn on different parts of the athlete's body, such as the wrist and waist, to capture the athlete's movement information. These sensors can measure multi-dimensional data such as athlete's movement trajectory, force distribution, and speed changes. Motion data is transmitted to intelligent terminals through wireless transmission technology, and real-time recognition and analysis are carried out through deep learning algorithms (Seo et al. 2015). The system presents motion data to users in various intuitive ways. Users can intuitively understand their swing movement characteristics and improvement space through these visual data. The system can also provide personalized follow-up sports and training suggestions based on sports data, helping athletes to improve and enhance their technical skills in a targeted manner (Fang and Sun 2021). Through the system, athletes can timely understand their swing movements and train and adjust according to the suggestions provided by the system. Coaches and referees can also monitor athletes' swing movements through the system, providing more accurate guidance and judgment, ensuring the fairness and impartiality of the game (Rahmad et al. 2019). This will help athletes improve their technical level and promote the development of badminton.

2 Related work

The literature has designed an action recognition algorithm based on a two-layer classifier, aiming to accurately identify six common types of badminton swing movements (Ko et al. 2015). In order to achieve this goal, the literature chose to fix a single inertial sensor at the bottom of the badminton racket grip to collect swing motion data from 12 athletes. In order to achieve low-power and efficient data transmission and collection, low-power Bluetooth technology has been adopted. After data collection, the literature used a window cutting method that combines action windows and sliding windows to capture real-time data. Through this method, each swing action can be cut into multiple continuous windows to extract the feature of the action data after window truncation. In order to reduce the dimension of features and indirectly realize Feature selection, the literature uses principal component analysis and Linear discriminant analysis to reduce the dimension of feature data. The literature adopts a two-layer classifier based on support vector machine and adaptive lifting algorithm to identify six common swing movements in badminton (An 2018). These two classifiers collaborate with each other and can accurately classify each swing action into the corresponding category by training and learning feature data. Literature has developed a badminton assisted training application software based on the iOS platform (Kuo et al. 2020). This software utilizes low-power Bluetooth (BLE) technology to communicate with data acquisition devices, collect real-time swing motion data of athletes, and can instantly identify specific swing motion types. This application not only provides real-time data collection and motion recognition functions, but also includes registration and login, motion data statistical analysis, and social functions. Users can create personal accounts through the registration and login function, and can share and compare their swing movement data with other users. The application software can also conduct statistical analysis on the user's swing action data, provide Data and information visualization and report, and help users understand their training progress and improvement direction (Hassan 2017).

The literature annotates the swing action types corresponding to each video segment in the dataset for subsequent training and analysis. The literature has conducted research and analysis on swing movements, using computer vision and machine learning techniques to extract and classify features for each video segment, thereby achieving automatic recognition and classification of swing movements (Becker et al. 2017). The literature can also analyze video segments of different swing movements to understand the technical differences and improvement directions of athletes in different movements, in order to improve their badminton skills (O'Brien et al. 2021). The literature is based on the vision system and motion control system of badminton robots, which collect athletes' swing movements in real-time through cameras, and use computer vision and pattern recognition technology to identify and analyze the swing movements. The system first processes the collected video stream and extracts key features such as motion trajectory, speed, angle, etc. Then, through trained machine learning algorithms, the extracted features are compared and matched with pre-defined action models to achieve recognition and classification of swing movements. In addition to action recognition and classification functions, the system also has functions for action tracking and real-time feedback. The system can track athletes' swing movements in real-time, and provide real-time feedback and guidance to athletes through robot voice prompts or display screen instructions, helping them improve their movement skills and training effectiveness (Li 2020). The literature proposes non fixed length dense trajectory features, which extract athlete feature points from video frames and sample feature points at certain time intervals based on the athlete's movement speed and

direction, thus forming non fixed length trajectory features (Wen et al. 2012). This non fixed length trajectory feature can better reflect the motion characteristics of actions and improve the accuracy of action recognition.

3 Key technologies of badminton player motion recognition and tracking system

3.1 Sensor

In order to identify the lower limb movement intention of the human body, researchers use various sensors such as bioelectric signal sensors, visual sensors, and mechanical sensors to obtain the user's lower limb movement status and movement intention information. Among these sensors, light motion capture technology, as a non-contact motion recognition technology, has a unique advantage. By deploying equipment such as high-speed cameras and infrared sensors, it can accurately capture the movement trajectory and posture of athletes. More detailed and comprehensive sports data can be obtained through real-time monitoring and analysis of athletes' sports status. The application of light motion capture technology in the badminton players' motion recognition system can quickly and accurately capture the athletes' motion track and posture without direct contact with the players' bodies. This is very important for badminton players because players need to maintain flexible, agile movements and need to fully use their arm and body coordination. The use of light motion capture technology can accurately record and analyze the athlete's movement posture and movement, and provide more scientific and personalized training guidance for the athlete.

A bioelectrical signal sensor is a sensor used to collect bioelectrical signals generated during human movement. Bioelectric signals can directly reflect motion intention and have a high sampling frequency. The signal occurs earlier than the motion, meeting the real-time requirements of motion intention recognition. However, bioelectrical signal sensors also have some common drawbacks. Firstly, bioelectrical signals have a low signal-to-noise ratio and are susceptible to various subjective and objective factors, such as fatigue, hair, sweat, electrode displacement, and signal artifacts. Secondly, the process of wearing and using bioelectric signal sensors is relatively cumbersome. It is necessary to accurately install and adjust the position of the sensor, and there is a high requirement for the adhesion of the electrodes.

Visual sensors can obtain environmental and user limb motion information through cameras. It can obtain information such as motion trajectory, posture, and actions by capturing image or video data. Compared to other sensors, visual sensors have higher resolution and rich information content, which can provide more detailed and comprehensive motion information. But visual sensors also have some limitations. Firstly, visual sensors are susceptible to the influence of lighting conditions. Under different lighting conditions, the brightness, contrast, and color of images may change, which affects the stability and accuracy of image processing algorithms. Secondly, visual sensors are also affected when facing obstacles. If there are obstructions in the user's limbs or environment, it can lead to partial information loss in the image, thereby affecting the recognition and analysis of motion intention.

Mechanical sensors can be used to measure the interaction force between the human body and wearable robots, ground reaction force, and information such as angle,

velocity, and acceleration during the movement of human limbs. Mechanical sensors have some advantages. Mechanical sensors typically have lower prices and are easy to deploy on humans and robots. The signal of mechanical sensors is stable and reliable, and is not affected by complex motion environments. But there are also some issues with mechanical sensors. The signal of mechanical sensors may have a certain lag, resulting in a certain time difference between the measurement results and the actual situation. The measurement results of mechanical sensors may be affected by the installation position and require calibration and compensation. The measurement results of mechanical sensors may be affected by the accumulation of errors, resulting in a decrease in the accuracy of the measurement results.

Repeatability is an important indicator for measuring sensor performance. Repeatability refers to the ability of a sensor to obtain similar results multiple times under the same measurement conditions. In this article, the repeatability of the pneumatic muscle motion sensor (PMMG) signal during human linear walking is verified by analyzing it. The results are shown in Fig. 1.

Through the light motion capture technology, the athlete's movement track, posture and action information can be obtained and recorded in the form of time series. Through the analysis of these data, the gait cycle characteristics of athletes can be obtained. During the experiment, the light motion capture technology can use equipment such as infrared cameras and infrared transmitters to accurately capture the muscle activity of the athlete during each gait cycle. Pneumomusculargram (PMMG) data can be obtained by using pneumatic muscle motion sensors and using light motion capture technology to measure muscle activity during each gait cycle. These data can reflect the cyclical characteristics of muscles. By looking at the PMMG data in Fig. 1, the shape and amplitude of the PMMG signal curve for each gait cycle are roughly similar. This shows that the light motion capture technology can stably measure the muscle activity of athletes, and the signal curve of each gait cycle has a good periodicity. The PMMG signals obtained from several consecutive experiments also showed similar characteristics, indicating that the sensor has good repeatability and can steadily record and capture the muscle activity of athletes.

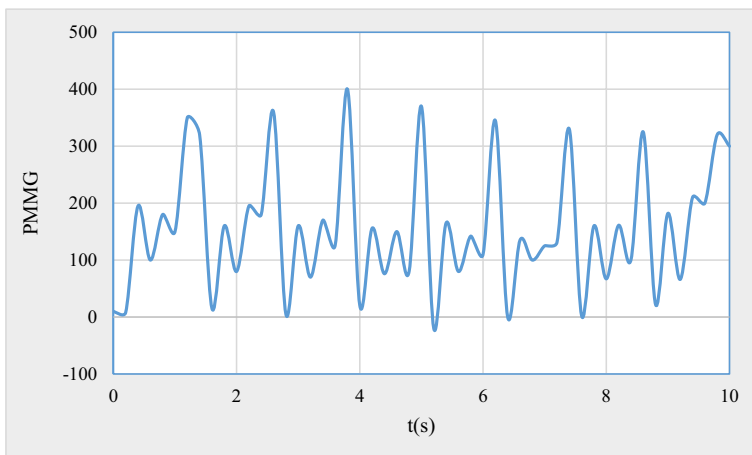


Fig. 1 PMMG signal of the left thigh muscle during uniform straight walking

In order to quantitatively evaluate the repeatability of PMMG sensors, past gait analysis studies were referred to and variance ratios were used for evaluation. The variance ratio (VR) is a numerical value that varies from 0 to 1, and the closer it is to 0, the better the repeatability of the sensor signal.

The formula for calculating the variance ratio is shown in formula (1):

$$VR = \frac{\sum_{i=1}^m \sum_{j=1}^n (P_{ij} - \bar{P}_i)^2 / m(n - 1)}{\sum_{i=1}^m \sum_{j=1}^n (P_{ij} - \bar{P})^2 / (mn - 1)} \tag{1}$$

In order to calculate the average PMMG signal value within each gait cycle (Formula 2), it is necessary to sum the PMMG signal values within each gait cycle and divide them by the number of gait cycles:

$$\bar{P}_i = \frac{1}{n} \sum_{j=1}^n P_{ij} \tag{2}$$

In order to calculate the global average value of the PMMG signal (Formula 3), it is necessary to sum the PMMG signal values within all gait cycles and divide them by the total number of gait cycles:

$$\bar{P} = \frac{1}{m} \sum_{i=1}^m \bar{P}_i \tag{3}$$

During the training process, weighted joint position loss (Loss) is used to measure the error between the generated pose and the actual pose. This Loss function is composed of three linear combinations, namely, key node error (Loss_core), body node error (Loss_body) and hand node error (Loss_hand):

$$Loss = \alpha_1 Loss_{core} + \alpha_2 Loss_{body} + \alpha_3 Loss_{hand} \tag{4}$$

$$Loss_{core} = P_{core} - P_{core}^{GT} \tag{5}$$

$$Loss_{body} = P_{body} - P_{body}^{GT} \tag{6}$$

$$Loss_{hand} = P_{hand} - P_{hand}^{GT} \tag{7}$$

According to Table 1 of the experimental results, when using Kinect as input data, the joint error is lower, indicating that Kinect data can provide more accurate output results compared to sparse IMU data. When using complete Kinect data as input conditions and comparing the results of different numbers of IMU inputs, it can be found that increasing IMU data can reduce the estimation error of three types of nodes. When the input data is complete Kinect and IMU data, that is, when the input data dimension is the largest, the estimation error of each type of node is the lowest.

These results indicate that using richer and more accurate sensor data in pose estimation tasks can improve the accuracy of pose estimation. Kinect data provides more joint position information, making the generated pose closer to the real pose. Adding IMU data can provide more pose information, thereby reducing the estimation error of nodes.

Table 1 Results of multimodal ablation experiment

Serial number	Kinect input	IMU input	Critical node error (cm ²)	Body node error (cm ²)	Hand node error (cm ²)
1	Complete	None	0.004087	0.006122	0.006443
2	None	The whole body	0.013258	0.014977	0.012996
3	Complete	lower limbs	0.004122	0.005821	0.006303
4	Complete	Upper limbs	0.003801	0.005896	0.005816
5	Complete	The whole body	0.002846	0.004031	0.003210

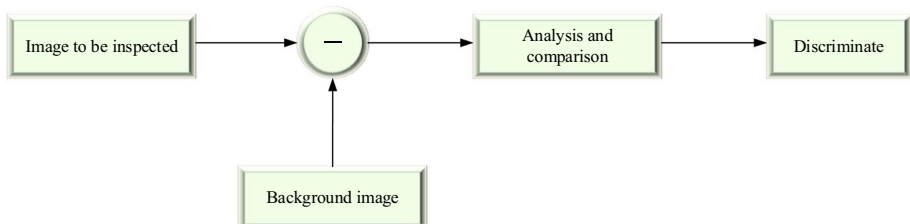
According to the results of the multi-stage ablation experiment in Table 2, when using Kinect as input data, the output results of multi-stage prediction have lower positional errors. Compared to using only IMU input data, key node errors, body node errors, and hand node errors are reduced. This indicates that Kinect data provides more accurate joint position information and can significantly improve the accuracy of pose estimation. When Kinect is combined with IMU as input data, the output results of multi-stage prediction have similar effects as using Kinect only as input data. The position error of multi-stage output results is lower than that of single stage prediction results, and the error of three types of nodes is reduced. This indicates that combining IMU data with Kinect data can further improve the accuracy of pose estimation.

3.2 Deep learning action tracking model

In the badminton player motion recognition system, the light motion capture technology can realize motion tracking by background difference method. Figure 2 shows the basic principle of background difference method. First obtain a background image, which can be obtained by collecting a static background image at the beginning of the experiment. Then the differential image is generated by the difference operation between the continuous

Table 2 Results of multi-stage ablation experiments

Experiment number	Input	Structure	Critical node error (cm ²)	Body node error (cm ²)	Hand node error (cm ²)
1	Kinect	Single stage	0.006493	0.005056	0.005337
2	Kinect	Multistage	0.003355	0.004056	0.005168
3	Kinect + IMU	Single stage	0.005228	0.003296	0.002594
4	Kinect + IMU	Multistage	0.002451	0.002422	0.002390

**Fig. 2** Background difference principle

frame image and the background image. Each pixel value in a differential image represents the degree of difference between the current frame and the background frame of that pixel. By setting a threshold, the foreground pixel in the differential image, that is, the moving part, is extracted. Using light motion capture technology, high-speed cameras are used to capture the movements of badminton players. These successive frame images can be compared to the background image to identify the position and posture of the athlete. Through the background difference method, the changes occurring in the frame image are detected and treated as moving foreground pixels. These foreground pixels can be further used to analyze and identify the movements of athletes.

In the background difference method, the RGB color information of an image is used to distinguish foreground objects from background information. By calculating the difference of each pixel in the image to be detected on the RGB channel (ΔI_r , ΔI_g , ΔI_b), the total information difference value of the pixel between the image to be detected and the background image can be obtained (ΔI).

The difference of RGB color information is calculated by formula (8):

$$\begin{cases} \Delta I_r = |I_r - I_{rback}| \\ \Delta I_g = |I_g - I_{gback}| \\ \Delta I_b = |I_b - I_{bback}| \end{cases} \quad (8)$$

By substituting the difference of RGB color information into formula (9), the total information difference value (ΔI) can be obtained:

$$\Delta I = \sqrt{\Delta I_r^2 + \Delta I_g^2 + \Delta I_b^2} \quad (9)$$

Finally, the value of ΔI is substituted into the formula (10) to determine whether the pixel is a foreground target. The specific judgment conditions of formula (10) can be set according to actual needs and image characteristics.

$$P_i = \begin{cases} 1, & \Delta I_i \geq T_0 \\ 0, & \Delta I_i < T_0 \end{cases} \quad i = 1, 2, \dots, n \quad (10)$$

By using RGB color information for background subtraction processing, compared to using only grayscale information, foreground targets and background information can be more accurately and reliably distinguished. However, the background difference method may still be affected by factors such as changes in lighting and background during its application. Therefore, it is necessary to comprehensively consider the actual situation and select appropriate thresholds and judgment conditions for foreground extraction.

4 Design of action recognition and tracking system for badminton players

4.1 System requirement analysis

The application of light motion capture technology based on multi-mode sensor in badminton players' motion recognition system provides accurate and reliable real-time data analysis for badminton training. The system is connected to the badminton robot, and software

development is carried out on the Windows 10 system in C++ language to ensure compatibility and stability between the software and the robot. Through the light motion capture technology, the system can accurately capture, identify and analyze the athlete's swing. The system realizes the accurate capture of athletes' swing motion by light motion capture technology. When the athlete is swinging, the system can continuously collect the athlete's movement track at high speed, and analyze the athlete's movement type and quality in real time. According to the different movement type and quality, the system can give the corresponding score evaluation. This provides players with real-time feedback and evaluation to help them improve and improve their swing technique.

The application of light motion capture technology in this system enables the players' swing movements to be accurately captured and analyzed in real time. By accurately extracting the details and features of the athletes' swing movements, the system can provide detailed technical guidance and training suggestions for the athletes. Players can understand the accuracy and quality of their movements by observing the scores and feedback provided by the system, and adjust and improve their swing technique according to the guidance and suggestions of the system.

According to the above requirements, the badminton player's motion recognition and tracking system should have the following functions:

- (1) Badminton swing Motion capture function: Using the vision system of the badminton robot, the video segment of players performing the swing can be captured. The system should have high frame rate and high-resolution video capture capabilities to ensure accurate capture of athlete movements.
- (2) Badminton swing motion recognition function: Identify the captured swing motion and accurately determine the type of motion. The recognition algorithm can use machine learning or deep learning methods to identify different types of swing movements through training models.
- (3) Badminton swing action analysis function: The system can calculate the similarity between the analyzed action and the standard action of the type to which it belongs. Measure the similarity between two actions by comparing parameters such as key-frames, trajectories, and speeds between them. Then, based on similarity, a scoring formula is used to calculate the score of the action.
- (4) Result display function: Display the type and score of the action on the display screen of the badminton robot vision system. Athletes can see their movement types and scores in real-time for self-evaluation and improvement.
- (5) Overall requirement: The system should be stable and reliable, with clear and clear interfaces between various functional modules to ensure normal and accurate operation of the system. The system should have a good user interface design, making the operation simple and intuitive, and convenient for athletes and coaches to use.

4.2 Overall framework and functional design of the system

As shown in Fig. 3, the badminton player motion recognition and tracking system mainly consists of five parts. The hardware part, namely the visual system of the badminton robot.

In order to realize the application of light motion capture technology in the badminton player's motion recognition system, this paper designs each module of the software and applies it to the vision system of the badminton robot. Among them, the camera SDK and OpenCV 4.1 image processing C++ library are used to write a wobble motion capture program. The

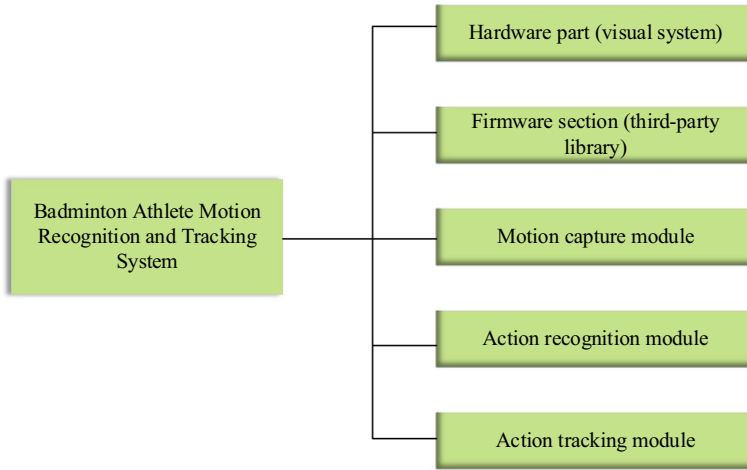


Fig. 3 Overall framework diagram of badminton player motion recognition and tracking system

program can obtain real-time video stream from the camera and process each frame image to extract the key frame, that is, the important image of the badminton swing action. On the basis of the wobble motion capture program, this research also uses OpenCV 4.1 image processing function library and VLFeat data analysis C++ function library to write the wobble motion recognition and analysis program. The program is able to extract features from captured key-frames and utilize recognition algorithms to accurately identify and analyze the type of action. The application of these function libraries makes the realization of image processing, feature extraction and recognition algorithms more concise and efficient. Through the light motion capture technology and the corresponding program, the system can obtain the key frame image of the athlete’s swing action in real time, and analyze and recognize the action type accurately through the feature extraction and recognition algorithm. In this way, the athlete’s swing can be accurately captured and analyzed, providing a reliable basis for subsequent data processing and feedback.

4.3 Design of system action recognition and tracking module

Regarding the diversity of datasets for the same hitting action in the sample, the correction formula for improving the model recognition rate in this article is as follows:

$$\bar{\pi}_i = \sum_{m=1}^M \frac{\alpha_1^{(m)}(i)\beta_1^{(m)}(i)}{P(O^{(m)}|\lambda)} \tag{11}$$

$$\bar{a}_y = \frac{\sum_{m=1}^M P_m \sum_{t=1}^{T_m-1} \alpha_t^{(m)}(i)a_{ij}b_j(\mathbf{o}_{t+1}^{(m)})\beta_{t+1}^{(m)}(j)}{\sum_{m=1}^M P_m \sum_{t=1}^{T_m} \alpha_t^{(m)}(i)\beta_t^{(m)}(j)} \tag{12}$$

$$\overline{b_j(k)} = \frac{\sum_{m=1}^M P_m \sum_{t=1}^{T_m-1} \alpha_t^{(m)}(i) \beta_t^{(m)}(j) I(o_t, v_k)}{\sum_{m=1}^M P_m \sum_{t=1}^{T_m} \alpha_t^{(m)}(i) \beta_t^{(m)}(j)} \tag{13}$$

Table 3 provides the average recognition rates for different action models, using three methods: single training, mean training, and frequency weighted training. These methods are used in training HMM parameter models to improve the recognition accuracy of different actions.

Through comparison, it can be found that both mean training and frequency weighted training can improve the recognition rate of the model compared to single training. Mean training reduces the differences between samples and improves the model’s generalization ability through data layer fusion. Frequency weighted training more accurately reflects the contribution of different samples to the model by considering the occurrence frequency of samples, further improving the recognition rate.

The hitting speed is estimated using angular velocity data collected by a six axis sensor at the end of the racket. Perform peak detection on the data of each stroke, where the peak time represents the stroke time. The instantaneous angular velocity at the stroke time is shown in formula (14), and the linear velocity $v(t_0)$ at the swing time is shown in formula (15):

$$\omega(t_0) = \sqrt{\omega_x^2(t_0) + \omega_y^2(t_0) + \omega_z^2(t_0)} \tag{14}$$

$$v(t_0) = \omega(t_0) \times l \tag{15}$$

This paper collects data from 10 rounds of singles badminton matches, and conducts Real-time computing calculation and statistics on scores and errors of individual players. Actual statistics and measurements were conducted on the scores and error evaluations of various hitting movements. Table 4 shows the program calculation results and actual measurement results of the scores and errors of ten hitting movements of both players in ten matches.

For video frames that fail to detect the human body, the rectangular box of the human body in the previous frame is expanded to obtain it. The specific expansion formula is as follows:

Table 3 Recognition rates of three HMM training algorithms

Action model	Single training	Mean training	Frequency Weighted Training
Serve	0.8738	0.9026	0.9408
Forehand rubbing ball	0.7654	0.9215	0.9319
Forehand putt	0.8313	0.9487	0.9362
Backhand push	0.9179	0.8894	0.9229
Backhand pick	0.7383	0.8760	0.9700
Forehand high and far	0.7941	0.8909	0.9242

Table 4 Scores and Error Results of Ten Stroke Actions

Action	Score (%)	Mistake (%)
Serve	0.00	0.00
Forehand rubbing ball	7.38	6.38
Forehand putt	8.31	0.00
Backhand push	0.00	14.28
Backhand pick	0.00	0.00
Forehand high and far	0.00	0.00

$$p_{top}^1 = \begin{cases} 0, & p_{top}^0 < 10 \\ p_{top}^0 - 10, & p_{top}^0 \geq 10 \end{cases} \tag{16}$$

$$p_{down}^1 = \begin{cases} p_{down}^0 + 10, & p_{top}^0 < x_{max} - 10 \\ x_{max}, & p_{top}^0 \geq x_{max} - 10 \end{cases} \tag{17}$$

$$p_{left}^1 = \begin{cases} 0, & p_{left}^0 < 10 \\ p_{left}^0 - 10, & p_{left}^0 \geq 10 \end{cases} \tag{18}$$

$$p_{right}^1 = \begin{cases} p_{right}^0 + 10, & p_{top}^0 < y_{max} - 10 \\ y_{max}, & p_{top}^0 \geq y_{max} - 10 \end{cases} \tag{19}$$

When an error occurs during the detection process, this article uses the same method to obtain the human rectangle in that frame. By comparing the center positions of the rectangular boxes in the two frames before and after, if the distance between the two rectangular boxes is greater than 50 pixels, it is judged as a detection error. For the detection of multiple human rectangular boxes, this article adopts the method of detecting feature points in all rectangular boxes for processing. By detecting the feature points of the human body, the position and posture of the human body can be more accurately determined.

The indicators used to represent the features of non fixed length trajectories include the pixel coordinates of the initial and end points of each trajectory, the average and variance of the position coordinates of all feature points of the trajectory, the pixel length of the trajectory, and the velocity and acceleration of the motion of feature points on the trajectory.

For trajectory $\{p_1, p_2, \dots, p_L\}$ with a length of L , its pixel length can be calculated using the following formula (20):

$$d = \sum_{i=1}^L \sqrt{p_{i+1}^2 - p_i^2} \tag{20}$$

According to the definition of velocity, the average value of the coordinate difference between adjacent frames in the trajectory is defined as the motion velocity of the feature point, while the difference in velocity between adjacent frames is defined as the acceleration of the feature point's motion. This can describe and analyze the speed and acceleration changes of the swing action through the movement of feature points on the trajectory.

For adjacent feature points p_i and p_{i+1} , their motion velocity can be calculated using the following formula (21).

$$v_x = \frac{\sum_{i=1}^{L-1} \Delta p_{xi}}{L-1}, v_y = \frac{\sum_{i=1}^{L-1} \Delta p_{yi}}{L-1} \quad (21)$$

$$a_x = \frac{\sum_{i=1}^{L-1} (\Delta p_{x(i+1)} - \Delta p_{xi})}{L-1}, a_y = \frac{\sum_{i=1}^{L-1} (\Delta p_{y(i+1)} - \Delta p_{yi})}{L-1} \quad (22)$$

By utilizing the shape features of non fixed length trajectories, the recognition rate of badminton swing movements can be significantly improved. Compared to fixed length trajectory features, non fixed length trajectory features have higher recognition rates in the swing action dataset under the same parameter settings. Grid 5 provides recognition results for different combinations of feature descriptors, and compares the performance differences between fixed length trajectory features and non fixed length trajectory features. Through comparison, it can be seen that the recognition rate of non fixed length trajectory features is significantly higher than that of fixed length trajectory features.

From Table 5, it can be seen that for the badminton swing action dataset, the recognition rate of non fixed length trajectory features is significantly better than that of fixed length trajectory features under different combinations of descriptors. For a single descriptor, the recognition rate of non fixed length trajectory features is significantly higher than that of fixed length trajectory features. This indicates that non fixed length trajectory features can better express the details and changes of swing movements. For combination descriptors, the recognition rate of non fixed length trajectory features is also significantly higher than that of fixed length trajectory features in the case of combination descriptors. Non fixed length trajectory features are more effective in identifying badminton swing movements, and can better capture the details and changes of the movements. This indicates that non fixed length trajectory features have better applicability in intra class action recognition problems, and have obvious advantages in identifying swing movements.

Figure 4 shows the flow of SR-SA network with light motion capture technology. The network structure adopts the stacking recurrent neural network (RNN) and the feature learning method of the self-focusing sensor. Deep learning is enabled by multi-layer RNNs and bidirectional gated loop units (GRUs), combined with multi-head self-focusing mechanisms. This design enables the network to better extract the interaction information between the time series and the sensor, thus improving the effect of feature learning. Through the continuous stacking of RNN layers and bidirectional GRU layers, the network

Table 5 Comparison of action recognition rates between dynamic trajectory and non fixed length trajectory when represented by different descriptors

Descriptor	Fixed-length trajectory (%)	Non-fixed trajectory (%)
Traj	43.72	57.16
HOG	61.90	61.33
HOF	55.91	59.51
MBH	60.37	59.42
Traj + HOF	59.93	60.96
HOG + MBH	62.87	62.04
Traj + HOF + MBH	62.43	62.70
Traj + HOG + HOF	64.05	67.16
Combined	64.77	67.49

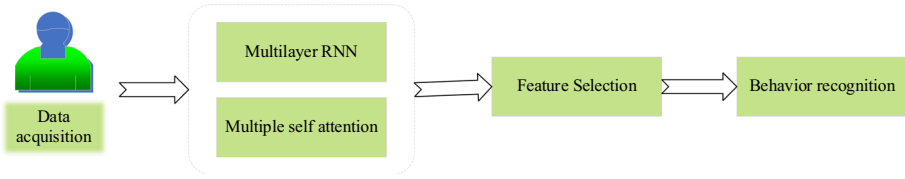


Fig. 4 Stacked Loop Self Attention Model Network

can gradually deepen its learning and extract richer feature information. The proportional dot product model is used as the attention scoring function, which can effectively normalize the attention weight and improve the stability and convergence of the model. Through the attention mechanism, the network can automatically focus on key feature points, thus improving the accuracy and stability of motion recognition.

Figure 4 shows a stacked loop self attention model network consisting of an input layer, a hidden layer, and an output layer. Among them, the hidden layer contains a bidirectional GRU layer. In the hidden layer, the data between layer i and layer $i + 1$ can be calculated using formula (23). This formula is used to update hidden layer data:

$$\mathbf{X}_{i+1}^r = \mathbf{X}_i^r * \mathbf{W}_{i,i+1}^r + \mathbf{b}_i^r \tag{23}$$

The calculation method of the bidirectional GRU layer can be carried out according to formulas (24)–(28). These formulas describe the calculation process of the bidirectional GRU layer, where W , W' , and b represent the corresponding weights and biases:

$$r_t = \text{sigmoid}\left(T\left(\mathbf{X}_{(i-1)j}^r, \mathbf{X}_{(i)j-1}^r\right)\right) \tag{24}$$

$$z_t = \text{sigmoid}\left(T\left(\mathbf{X}_{(i-1)j}^r, \mathbf{X}_{(i)j-1}^r\right)\right) \tag{25}$$

$$\tilde{h}_t = \tanh\left(T\left(\mathbf{X}_{(i-1)j}^r, \mathbf{X}_{(i)j-1}^r\right)\right) \tag{26}$$

$$\vec{h}_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \tag{27}$$

$$\overleftarrow{h}_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \tag{28}$$

Through these calculations, the stacked loop self attention model network can learn the temporal and dependency relationships in the input sequence, and generate corresponding outputs. The use of a bidirectional GRU layer enables the network to simultaneously consider past and future information, better capturing contextual information in input sequences.

4.4 Analysis of system experiment results

According to Fig. 5, when running the designed application on an Android phone, the CPU memory usage is monitored and recorded.

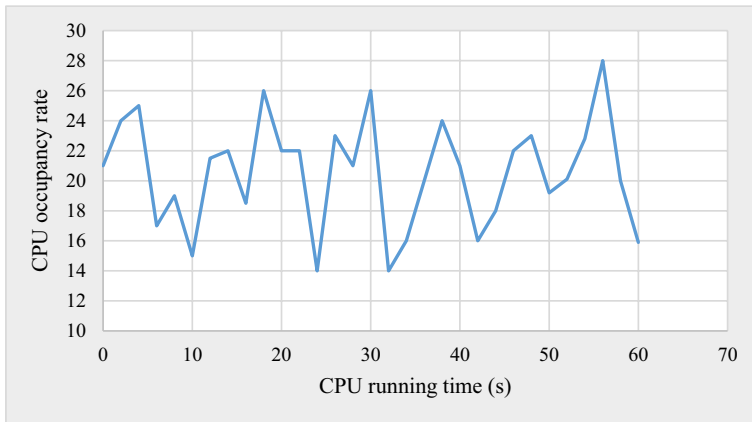


Fig. 5 CPU usage of mobile phones during runtime of behavior recognition application

According to the measurement results, the designed application performs well on Android phones, with low CPU and memory usage, making it suitable for implementing behavior recognition methods based on the SR-SA model. This means that applications can effectively recognize behavior on mainstream mobile devices without having a significant impact on user experience and device performance.

It is very important to test the application software after its initial completion. The purpose of testing is to verify whether the software meets the previously established requirements and can operate safely and stably as expected. Through testing, the stability, reliability, and practicality of the software were evaluated to ensure that it can achieve the expected results in practical use.

5 Conclusion

With the development of technology and the increasing attention paid to sports, the training and performance analysis of athletes has become an important task. In badminton, accurately capturing and identifying players' movements can provide valuable information for training and skill improvement. However, the traditional motion recognition methods are limited by sensor types and data acquisition methods, which cannot meet the requirements of accuracy and real-time performance. In order to solve this problem, this paper presents an application of light motion capture technology based on multi-mode sensor in badminton player motion recognition system. The technology utilizes the camera SDK and OpenCV image processing library to extract and analyze the movement of athletes by capturing and processing key frame images in real-time video streams. The VLFeat data analysis library and self-focusing sensor feature learning method are also applied to improve the accuracy of feature extraction and motion recognition. Through comparative experiments, it is proved that the application of light motion capture technology based on multi-mode sensor in badminton players' motion recognition system can effectively capture and recognize athletes' movements, and ensure real-time and accuracy. Compared with traditional methods, this technique has higher accuracy and reliability. Light motion capture technology based on multi-mode sensor provides an efficient, accurate and real-time solution for badminton player motion recognition system, and has a broad application prospect. Future research could further explore and improve this

technology to meet the needs of more sports and provide better support for athletes' training and skill development.

Author contributions Tiejun Zhang has contributed to the paper's analysis, discussion, writing, and revision.

Funding The authors have not disclosed any funding.

Data availability The data will be available upon request.

Declarations

Conflict of interest The authors declare that they have no competing interests.

Ethical approval Not applicable.

References

- An, F.P.: Human action recognition algorithm based on adaptive initialization of deep learning model parameters and support vector machine. *IEEE Access* **6**, 59405–59421 (2018)
- Becker, R.M., Keefe, R.F., Anderson, N.M.: Use of real-time GNSS-RF data to characterize the swing movements of forestry equipment. *Forests* **8**(2), 44 (2017)
- Bohm, P., Scharhag, J., Meyer, T.: Data from a nationwide registry on sports-related sudden cardiac deaths in Germany. *Eur. J. Prev. Cardiol.* **23**(6), 649–656 (2016)
- Chu, Z., Li, M.: Image recognition of badminton swing motion based on single inertial sensor. *J. Sensors* **2021**, 1–12 (2021)
- Fang, L., Sun, M.: Motion recognition technology of badminton players in sports video images. *Futur. Gener. Comput. Syst.* **124**, 381–389 (2021)
- Hassan, I.H.I.: The effect of core stability training on dynamic balance and smash stroke performance in badminton players. *Int. J. Sports Sci. Phys. Educ.* **2**(3), 44–52 (2017)
- Ko, B., Hong, J., Nam, J.Y.: Human action recognition in still images using action poselets and a two-layer classification model. *J. vis. Lang. Comput.* **28**, 163–175 (2015)
- Kuo, K.P., Tsai, H.H., Lin, C.Y., Wu, W.T.: Verification and evaluation of a visual reaction system for badminton training. *Sensors* **20**(23), 6808 (2020)
- Li, C.: Badminton motion capture with visual image detection of picking robotics. *Int. J. Adv. Rob. Syst.* **17**(6), 1729881420969072 (2020)
- Liu, L.: Arm movement recognition of badminton players in the third hit based on visual search. *Int. J. Biometrics* **14**(3–4), 239–252 (2022)
- O'Brien, B., Juhas, B., Bieńkiewicz, M., Buloup, F., Bringoux, L., Bourdin, C.: Sonification of golf putting gesture reduces swing movement variability in novices. *Res. q. Exerc. Sport* **92**(3), 301–310 (2021)
- Phomsoupha, M., Laffaye, G.: The science of badminton: game characteristics, anthropometry, physiology, visual fitness and biomechanics. *Sports Med.* **45**, 473–495 (2015)
- Rahmad, N.A., Sufri, N.A.J., Muzamil, N.H., As'ari M A.: Badminton player detection using faster region convolutional neural network. *Indonesian J. Electr. Eng. Comput. Sci.* **14**(3), 1330–1335 (2019)
- Seo, J., Starbuck, R., Han, S., Lee, S., Armstrong, T.J.: Motion data-driven biomechanical analysis during construction tasks on sites. *J. Comput. Civ. Eng.* **29**(4), B4014005 (2015)
- Ting, H.Y., Sim, K.S., Abas, F.S.: Kinect-based badminton movement recognition and analysis system. *Int. J. Comput. Sci. Sport* **14**(2), 25–41 (2015)
- Wen, X., Kleinman, K., Gillman, M.W., Rifas-Shiman, S.L., Taveras, E.M.: Childhood body mass index trajectories: modeling, characterizing, pairwise correlations and socio-demographic predictors of trajectory characteristics. *BMC Med. Res. Methodol.* **12**, 1–13 (2012)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.