Error Motion Tracking Method for Athletes Based on Multi Eye Machine Vision

Yanlan Huang[1](✉) and Chunshou Su[2]

[1] Guangxi College for Preschool Education, Nanning 530022, China
hyl5825@163.com
[2] Yulin Normal University, Yulin 537000, China

Abstract. Traditional methods are unable to perform three-dimensional detection of erroneous movements, resulting in insufficient accuracy in tracking athlete erroneous movements. Therefore, the tracking method for athlete erroneous movements based on multi eye machine vision is highlighted. Using multi eye machine vision technology to construct an athlete error motion tracking framework and obtain athlete error motion image timing. From the perspective of regional consistency and similarity, segment machine vision images of athlete's incorrect actions. Apply Canny operator to detect athlete's incorrect actions, obtain pixel values of edge images, and remove false edges. The design is based on a multi eye machine vision athlete error action recognition process, obtaining unknown vectors. With the support of a multi eye machine vision detection system, the absolute value of brightness difference between two frames of images is calculated, and the Hom Schunck algorithm is combined to track the optical flow field to achieve athlete error action tracking. From the experimental verification results, it can be seen that the tracking curve of this method for three types of erroneous actions is consistent with the actual curve, and the maximum tracking accuracy is 93%, which can accurately track athlete's erroneous actions.

Keywords: Multi Eye Machine Vision · Athlete Error Movement Tracking · Hom-Schunck Algorithm · Error Action Image

1 Introduction

A high-quality sports event is often inseparable from the tactical arrangement of the coach. Whether the real-time match information of athletes, such as speed and exact route, can be well obtained plays a decisive role in the coach's arrangement of personnel and tactics. Real time access to athletes' competition information is the core function of athletes' tracking system. With the continuous improvement of sports technology, in the process of sports technology research, the demonstration of correct movement has become a key topic in the teaching field. In the actual teaching process, due to the obvious difference between the students' cognitive and understanding level and their action ability, some students have more wrong actions, and they are slow to master correct sports actions. In this state, how to effectively correct wrong movements in sports has

L. Yun et al. (Eds.): ADHIP 2023, LNICST 550, pp. 150–164, 2024.
https://doi.org/10.1007/978-3-031-50552-2_10

become the main problem to be solved in this field. In recent years, with the promotion of computer image processing technology in China, computer vision feature analysis and image processing technology have been widely used in the analysis of human body structure, which can analyze various forms of human body in motion. In this context, the field of sports has also begun to introduce computer vision feature analysis technology to athletes' action recognition and correction, so as to improve the effectiveness and judgment of athletes' training.

In outdoor sports events, wearable GPS devices are often used to achieve real-time tracking of athletes. In indoor sports events, high-precision multi-sensor systems are generally used to obtain the location information of athletes. The above two methods will make the tracking system too complex and difficult to maintain. In view of the shortcomings of existing tracking methods, a tracking method that can be applied to mobile devices is proposed. Reference [1] proposed a tracking method based on the stacked short-term memory network. This method takes the original track of athletes' actions as input, uses the stacked short-term memory network to learn the feature representation of space-time window, and cascades the implicit spatial representation of athletes on the field through an additional full connection layer. The Softmax layer is used to estimate the probability of athletes' final actions. Each final action is associated with an expected score, and it is used to estimate the expected score to obtain the tracking results of athletes' wrong actions; Reference [2] proposed a tracking method based on LSTM neural network, which smoothed and denoised the original 3D skeleton data on the basis of 3D skeleton data to conform to the smooth rule of human joint movement. A fusion feature composed of static features and dynamic features is constructed to represent human actions, and a key frame extraction model is introduced to extract key frames in human action sequences to reduce the amount of computation. The human motion classification model of Bi-LSTM neural network based on LSTM neural network is established. Attention mechanism and Dropout are introduced to classify and recognize human motion. However, the two methods mentioned above can be well expanded and have strong applicability, which can meet the needs of somatosensory interactive applications, but the recognition accuracy is very low. Therefore, an error motion tracking method for athletes based on multi eye machine vision is proposed.

2 Tracking Image Processing Based on Multi Eye Machine Vision

The multi eye machine vision technology is used to build the athlete error motion tracking framework, and its structure is shown in Fig. 1.

It can be seen from Fig. 1 that the tracking target is determined according to the framework, and the end to end training is conducted through forward and reverse propagation to obtain the final tracking result [3]. Convolution operation is a sparse, parameter sharing, variable operation. In the convolution operation, a filter is used to scan the two-dimensional image as a whole (filter). Each pixel in the vector graph will be scanned on the same convolution kernel, so its convolution kernel is the weight. Weight assignment makes the deep convolution network only need to learn a group of parameters, thus greatly reducing the number of parameters.

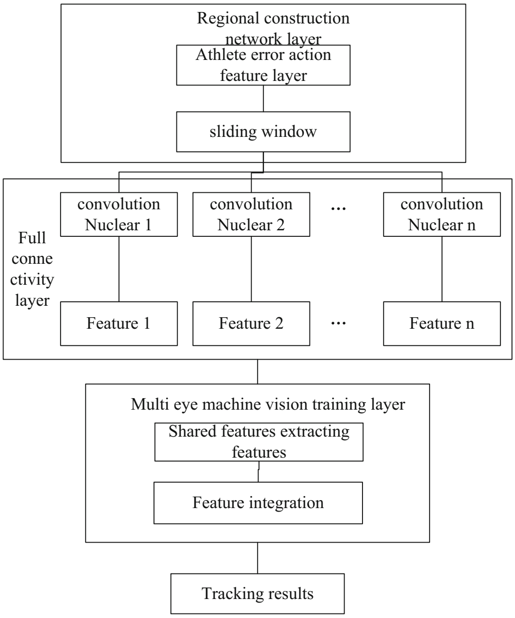


Fig. 1. Tracking framework of multi eye machine vision

2.1 Time Sequence Acquisition of Athletes' Wrong Action Images

Obtaining athletes' action images is the premise of extracting wrong action features. Due to the large visual error of action images obtained by traditional methods, it directly affects the accuracy of extracting wrong action features [4]. For this reason, the multi eye machine vision system is used to obtain athletes' action images. The components used to obtain images with the OV7670 camera are PL and PS, and the athletes' action images are displayed using the VGA interface according to the data integrity of line interruption and field interruption. The VGA timing is shown in Fig. 2.

In the multi eye machine vision system, the OV7670 camera is selected to obtain the movement image of athletes using the camera [5]. The OV7670 camera is a component of the COMOS camera and has the ability to acquire color images. The photosensitive array can achieve a maximum transmission rate of 640 * 680 and a maximum transmission rate of 30 frames/second. The camera has only one set of parallel data interfaces, marked as Y [7:0]. The pixel values of action images are read through the data interface, and the action images of football players are obtained in parallel.

2.2 Machine Vision Image Segmentation of Athletes' Wrong Actions

From the perspective of regional consistency and similarity, it can effectively prevent the integrity of region extraction caused by changes in image grayscale; Secondly, from the perspective of threshold segmentation, the time required for this method can be shortened [6]. Use the k-means clustering algorithm to converge the sample points to the position with the highest probability density, thereby obtaining regions with similar grayscale.

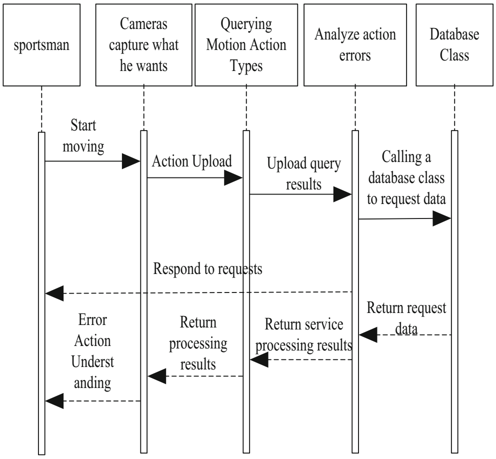


Fig. 2. Time sequence of athletes' action images

When applied to motion error machine vision image segmentation, the k-means clustering algorithm has advantages and characteristics such as simple, efficient, intuitive clustering results, applicability to large-scale datasets, and strong interpretability. This algorithm can quickly calculate the distance between sample points and the cluster center, allocate sample points to the nearest cluster center, and iteratively update the position of the cluster center, thereby segmenting erroneous action images into different clusters and intuitively distinguishing different types of errors.

The probability density value of similar regions is calculated using the formula:

$$\rho(a) = \frac{1}{kD} \sum_{x_i \in s_D} f\left(\frac{a}{D}\right) \tag{1}$$

In formula (1), a represents the sampling point; s_D represents the area with a fixed bandwidth of D; $f(\cdot)$ represents a symmetric kernel function; k represents the number of calculations [7–9]. After determining the initial point, use, k-means clustering algorithm to perform iterative processing according to the following steps:

Step 1: Use the density clustering method to classify the athletes' error movement tracking data, and divide the normal data and error data.In the athlete's error action tracking data set, error data contains some feature quantities that must be preprocessed [10]. The preprocessed training data set can be regarded as a feature matrix, in which the data in the same row is consistent with the features of multi eye machine vision. For the clustering problem in multidimensional vector space, the density of attributes is used to cluster each attribute, and the form of adjacent regions of eigenvalues can be expressed as:

$$q(z) = \{z_1, z_2 \in V_m | \ dist(z_1, z_2) \leq \vartheta\} \tag{2}$$

In formula (2), z_1 and z_2 represent two eigenvalue objects; V_m represents a set of m eigenvalues; ϑ represents the core object in the cluster.

Step 2: Using density clustering technology to conduct unsupervised machine learning method, it can determine the density of its distribution by taking the size and concentration of adjacent areas as indicators without setting the number of clusters in advance, so as to find clusters with irregular shapes [11–13]. The calculation formula of neighborhood radius and neighborhood density can be expressed as:

$$l = \frac{d_{z_2} - d_{z_1}}{\max\left(d_{z_2}, d_{z_1}\right)} \tag{3}$$

$$\rho = \frac{1}{n}\sum_{k=1}^{n} l_k \tag{4}$$

In the above formula, d_{z_1} and d_{z_2} respectively represent the distance between the two samples z_1 and z_2 and other points in the same category; n represents the number of samples. This method can mark dense scattered points as one type, and scattered scattered points as another type to distinguish between normal data and error data [14, 15].

Step 3: Continuously iterate and set the convergence threshold λ. The constraint conditions for convergence values can be expressed as:

$$\|a_{n+1} - a_n\| < \lambda \tag{5}$$

When the calculation result of formula (5) is satisfied, it means that the convergence density reaches the maximum value, otherwise, it cannot. This clustering method can avoid the problem of region segmentation errors.

2.3 Tracking Image False Edge Processing

The edge of an image is the grayscale value space boundary in which the pixel neighborhood has a stepping property. Extracting a set of pixels with significant grayscale changes from the image is the result of image edge detection. The Canny operator is a first-order derivative Gaussian method that accurately extracts image edges from an image. By analyzing image gradients, the edges of each pixel in the image can be obtained. The main reasons why Canny operator can accurately detect edges when extracting image edges are as follows. Firstly, the Canny operator adopts multi-stage operations, including Gaussian filtering, gradient calculation, and non maximum suppression, which can effectively suppress noise, locate edges, and refine edges. Secondly, the Canny operator uses Gaussian filtering to smooth the image, reducing noise interference and making subsequent edge detection more stable and reliable. Then, by calculating the gradient size and direction of each pixel, the Canny operator can locate the edge regions in the image, as the pixel values at the edges change greatly, and the gradient values also increase accordingly. Next, the Canny operator uses non maximum suppression to refine edges, preserving local maxima in the gradient direction, and further improving the accuracy of edge detection. Finally, the Canny operator uses dual threshold processing to filter edges, identifying pixels with gradient values greater than the high threshold as strong edges, while pixels between the low and high thresholds are identified as weak edges.

By connecting strong edges with adjacent weak edges, the Canny operator can obtain the final accurate edge result. Therefore, the Canny operator can extract accurate edge information from images through multiple stages of processing and parameter settings. Canny operator is applied to detect athletes' wrong actions, and the pixel value of edge image obtained can be expressed as:

$$\beta'(a, b) = \Delta\beta(a, b) \otimes \Delta c(a, b) \tag{6}$$

In formula (6), $\Delta\beta(a, b)$ represents the loss compensation pixel value; $\Delta c(a, b)$ represents the pulse response compensation value. In order to solve the problem of false edges, a weighted steering filter is constructed by combining the steering filter and edge sensing weight, which can strengthen the low-frequency components in the image and prevent false edges from appearing in the image. The calculation formula of edge perception weight is:

$$\omega(i) = \frac{1}{k}\sum_{i=1}^{k}\frac{S^2(i)}{S^2(I)} \tag{7}$$

In formula (7), S represents the variance of pixel point i as the guide image I within the pixel neighborhood. The pixel value of the image smoothing position is different from that of the edge position, where the edge perception weight of the smoothing position is less than 1 and the edge position is greater than 1. In the image, the two nearest edges are taken as true edges, which are very different from the average value of each pixel. For other pixels, if the pixel value jumps too large, it is called false edge. The false edge is determined according to the dispersion of each pixel in the image. The smallest of the two pixels is selected as the true edge, and the remaining pixels are false edges.

3 Identification and Tracking of Athletes' Wrong Movements

3.1 Athlete Error Action Recognition Based on Multi Eye Machine Vision

The iterative algorithm can quickly calculate the local maximum probability of the target to adapt to the multi-target deformation. Once the target is occluded, the number of multi-target maximum points will become more. At this time, the segmented multi-target tracking and positioning results will be biased, and the coordinate information will be lost. Segmented multi-target tracking and location method can track the target well for the problem of target occlusion, but the selection of location information is relatively strict.

Compared to other factors, multi eye machine vision systems need to focus on occlusion issues when applied. Occlusion issues can lead to information loss, target deformation, and target loss. Therefore, when using multi eye machine vision systems for multi object tracking, the influence of other factors should be ignored and occlusion issues should be emphasized.

In the whole tracking process, the feature information needs to be fused according to Bhattacharyya coefficient, and the target model is established for the current frame through Dempster-Shafer evidential reasoning. After iterative processing, candidate targets are captured, and similarity judgment is made according to the selected targets. If the similarity is greater than the set threshold, it means that the multiple targets are not blocked, and the multi target tracking is continued using the multi eye machine vision system; On the contrary, if the similarity is less than the set threshold, it means that multiple targets are occluded. At this time, the target template should not be updated, and the Dempster-Shafer evidential reasoning method should be directly switched to track. The method of comparing similarity with threshold can ensure good tracking effect when multiple targets are not occluded. Once the target is occluded, the method can enhance the effectiveness of tracking.

The process of athlete error action recognition based on multi eye machine vision is shown in Fig. 3.

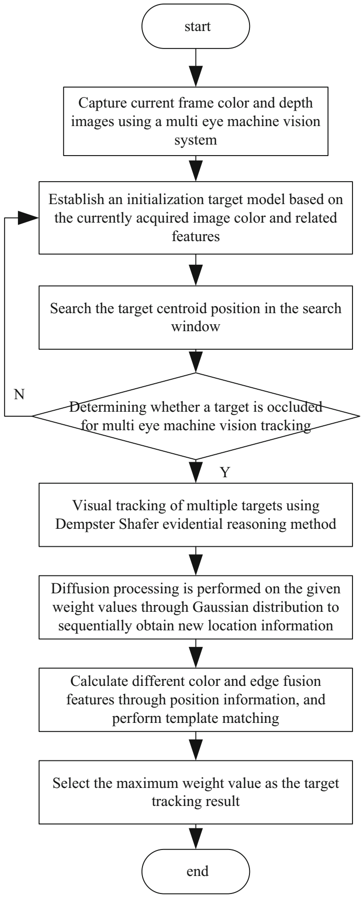


Fig. 3. Athlete's Error Action Recognition Process Based on Multi eye Machine Vision

It can be seen from Fig. 3 that the camera orientation is calculated by the coordinate information obtained from the multi object tracking and positioning based on multi eye machine vision and the corresponding points as well as the internal parameters of the multi eye machine vision system, the tracking target position is estimated by Dempster-Shafer evidential reasoning method, and the virtual control points are used to represent the points of the multi object position. The camera perspective problem is transformed into the control point problem under the camera coordinate system, and the coordinate points under the camera coordinate system are represented by the marked coordinate points to build the imaging model. The imaging model is corrected by using the model parameters of camera pinhole imaging, and the image coordinates are obtained. The camera internal parameters can be obtained by using the least square method, thus the unknown vector can be obtained. Identify the target points marked on the ground. If the number of points can not be completely recognized, the target position information can be obtained by identifying several of the points, so as to achieve multi-target visual positioning.

3.2 Error Motion Tracking Based on Multi Eye Machine Vision

Select appropriate instruments to establish a fast, accurate and effective imaging data entry and analysis system. When designing computer aided devices, it is necessary to make specific analysis on camera, special image processing system, lighting device and other components. The structure of the applied machine vision inspection system is shown in Fig. 4.

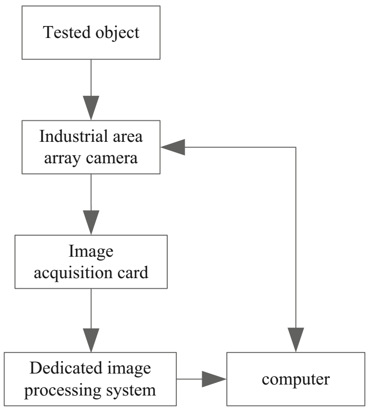


Fig. 4. Schematic diagram of application of multi eye machine vision inspection system

On the basis of athletes' wrong actions, the tracking process of wrong technical actions is designed by calculating athletes' wrong action descriptors. Multi eye machine vision technology is based on the characteristics of optical flow. The optical flow information set is established by the time displacement of each pixel, which requires high accuracy of optical flow. In order to transform the optical flow vector of moving video into a vector field, and then form a movable spatial distribution relationship, especially the optical flow field needs to be analyzed. Because optical flow can only reflect the motion information in the tennis player's foreground image when tracking the image, the background in the tracking image will affect the calculation of optical flow. Therefore, the background must be cleaned up first. Not only the uniformity of background color should be considered, but also the global foreground image of athletes should be obtained after the region growth algorithm is processed based on the Gaussian mixture model.

The inter frame difference method is used to eliminate the background color centered on athletes. By analyzing the gray difference of the corresponding pixels in the image, and then using double thresholds to select the gray difference in the image, the frame difference method is obtained by performing the difference operation on the adjacent frames in the image. When tracking athletes' wrong actions, there will be significant brightness difference between frames. Based on this, the absolute value of brightness difference between two frames can be expressed by the following formula:

$$\Delta G(a, b) = |G(a, b, t) - G(a, b, t - 1)| \tag{8}$$

In formula (8), $G(a, b, t)$ represents the grayscale value of image pixels at time t. According to the definition of the camera, the user can manually adjust the image sequence of the error action, and then estimate the length of the light field. First, according to the change degree of the camera's flash and intensity, the brightness of the error action image is tracked in real time, which will cause errors in the optical flow calculation results. Therefore, image difference is used to distinguish brightness and eliminate the influence caused by brightness change. Secondly, by analyzing the theory of biological vision system, it can be seen that machine vision cells are very sensitive to the edge movement of objects. In the direction and speed, different optical flows are formed due to different images, which reflects how the human vision system affects the changes of optical flow. On the basis of the difference image, the Hom-Schunck algorithm is used to estimate how athletes track the Hom-Schunck optical flow field, as follows:

$$\begin{cases} A_i = H_i - H_{i-1} \\ O_i = E(A_i) \end{cases} \tag{9}$$

In formula (9), In formula (9), A_i represents the differential image of tracking error action H_i and H_{i-1}; E represents the estimated expression of the Horn Chunck algorithm; O_i represents the optical flow field. The athlete's position in the adjusted error action image is related to the relative displacement of the body, which exists in the corresponding image area. For different postures of tennis players, the spatial distribution of optical flow field is different.

In the machine vision system, the binary image of a pixel in the image can be obtained by comparing the pixels between the correct action and the wrong action of the athlete, which can be expressed as:

$$I(a, b) = \begin{cases} 1 & \Delta G(a, b) < \xi \\ 0 & \Delta G(a, b) \geq \xi \end{cases} \tag{10}$$

In formula (10), ξ is the threshold for distinguishing between correct and incorrect actions by athletes. When the calculation result in the above formula is 0, it indicates that the difference of one pixel gray level in the two images is the difference data represented by the correct action of the athlete; When the calculation result in the above formula is 1, it indicates that the difference of one pixel gray level in the two images is the difference data represented by the athlete's wrong action. Between two adjacent frames, due to the environment, lighting and other factors, it is inevitable that there will be a pixel area with a calculation result of 1. At this time, it is necessary to determine whether this area is the area where the athlete's wrong action is located. In the machine vision system, the region screening method is used to select the pixel connected area with the operation value of 1 to judge whether this area is the wrong action area of athletes, so as to effectively screen the wrong action area of athletes and eliminate other interference.

According to the kernel density estimation and grid histogram of the error motion image, the optical flow histogram is collected as the motion descriptor of the athlete during the movement. For a given optical vector of the optical flow field coordinates, calculate the amplitude $\eta(o)$ and direction angle $\theta(o)$ of the athlete's erroneous action, with the formula:

$$\begin{cases} \eta(o) = \sqrt{\sigma_x^2(o) + \sigma_y^2(o)} \\ \theta(o) = \arctan \frac{\sigma_x(o)}{\sigma_y(o)} \end{cases} \tag{11}$$

In formula (11), $\sigma_x(o)$ and $\sigma_y(o)$ represent the horizontal and vertical components of the optical vector at the given optical flow field coordinate o. In conclusion, on the basis of multi eye machine vision technology, error action features are extracted, and error action descriptors are determined by tracking and adjusting error actions, so that error action tracking is realized.

4 Experiment

4.1 Establishment of Experimental Platform

In order to make full use of the convenience of mobile devices and the power of computer GPU floating point computing capability, the system adopts the C/S architecture mode, which is a typical distributed architecture. The system structure is shown in Fig. 5.

It can be seen from Fig. 5 that it mainly includes three parts, namely, the client, the network transport layer and the server.

The client refers to the mobile phone, which is mainly an athlete's long-term tracking system APP. First, you need to customize an IDL file, which is an essential work with the Thrif framework. First, define a structure to encapsulate the image frame of the video sequence, and then define two services. One is the init service to initialize the tracking algorithm of the server, and transfer the first frame (template image) and the initial position of the tracking target to the tracking algorithm for operation. The other is the tracking service. It is used to transfer the subsequent image frames (search images) to the tracking algorithm for target position prediction. Next, you can use the code generation engine of the Thrift framework to generate Java code, and write the corresponding

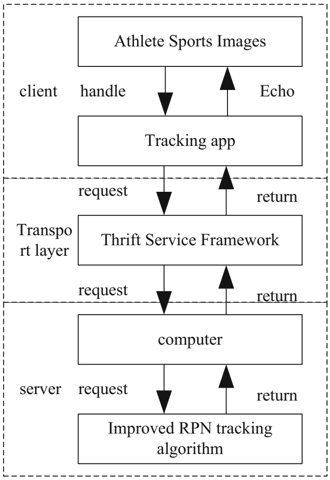


Fig. 5. Experimental Platform

operation interface and logic processing to build an Android client. It should be noted that the init service and tracking service are implemented on the server, and the client only borrows the injection interface. In addition, since the use of Thrift framework needs to consider the problem of network transmission delay, the video captured at the client adopts a caching strategy, which converts the video into a frame image, and then transmits it to the server for processing, and finally displays the results on the Android screen. Although this will cause the results displayed on the screen to be slightly slower than the actual shooting, it can effectively avoid the impact of network delay.

The network transport layer is implemented using TCP transport protocol and has been encapsulated by the Thrift framework.

The server of the server refers to the computer. Through the IDL file defined in the figure, the corresponding Python code can be generated. Then the improved SiamRPN algorithm can be written into the corresponding inti service and tracking service.

4.2 Experimental Evaluation Criteria

In order to quantitatively evaluate the athlete's error action tracking method based on multi eye machine vision, the recall rate index and accuracy rate index are introduced to determine the recognition ability of each action in the error action. The calculation method for recall index R and accuracy index P is:

$$\begin{cases} R = \frac{m_1}{m_1+m_2} \times 100\% \\ P = \frac{m_1}{m_1+m_3} \times 100\% \end{cases} \tag{12}$$

In formula (12), m_1, m_2, and m_3 respectively represent the number of correctly identified erroneous actions, the number of unrecognized erroneous actions, and the number of incorrectly identified erroneous actions.

4.3 Analysis of Experimental Data

Taking the high jumper as an example, the correct high jump action is shown in Fig. 6.

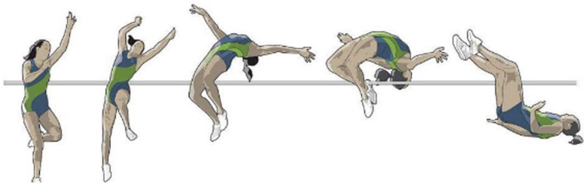


Fig. 6. Example of correct high jump

The correct action tracking curve and error action tracking curve corresponding to Fig. 6 are shown in Fig. 7.

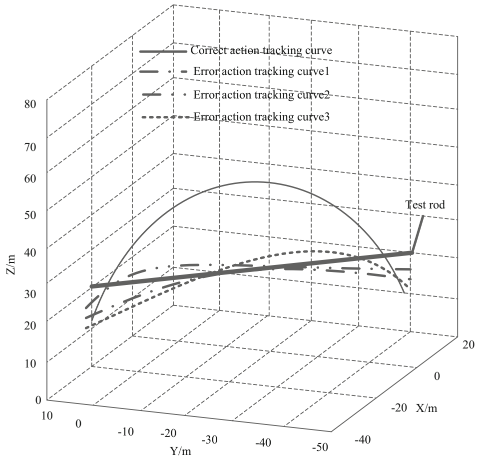
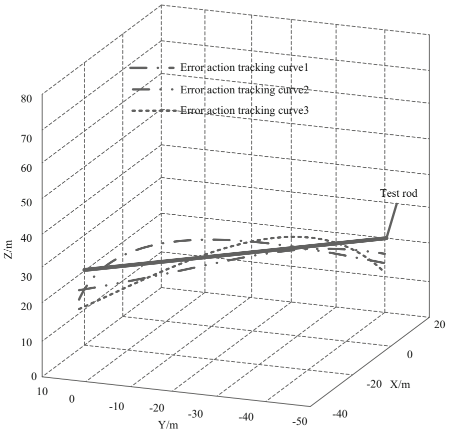


Fig. 7. Correct and error action tracking curve

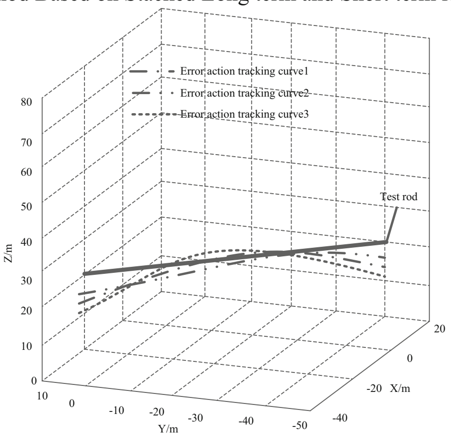
It can be seen from Fig. 7 that the correct action tracking curve shows that the athlete successfully skips the test bar, while the wrong action tracking curve shows that the athlete fails to skip the test bar. The first type will hit the bar during the descent, the second type will hit the bar directly because of the short approach distance, and the third type will hit the bar because of the poor posture during the ascent.

4.4 Test Result

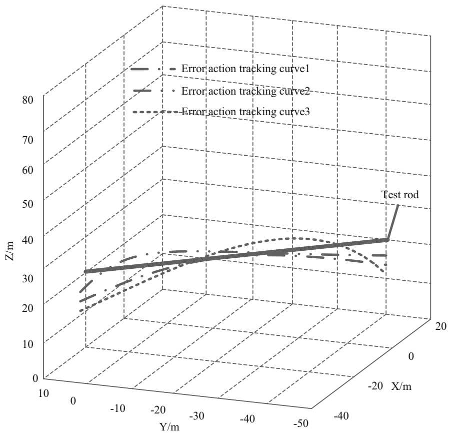
Use the tracking method based on the stacked short-term memory network, the tracking method based on LSTM neural network and the tracking method based on multi eye machine vision to compare and analyze whether the error action tracking curve is consistent with the curve connected to the experimental data, as shown in Fig. 8.



(a) Tracking Method Based on Stacked Long term and Short term Memory Network



(b) Tracking Method Based on LSTM Neural Network



(c) Tracking method based on multi eye machine vision

Fig. 8. Error action tracking curve analysis of different methods

As shown in Fig. 8, using the tracking method based on the stacked short-term memory network and the tracking method based on the LSTM neural network, the three error action tracking curves are inconsistent with the actual curves, while using the tracking method based on multi eye machine vision, the three error action tracking curves are consistent with the actual curves.

In order to further verify the high tracking accuracy of the research method, the tracking effects of the three methods are compared again, as shown in Fig. 9.

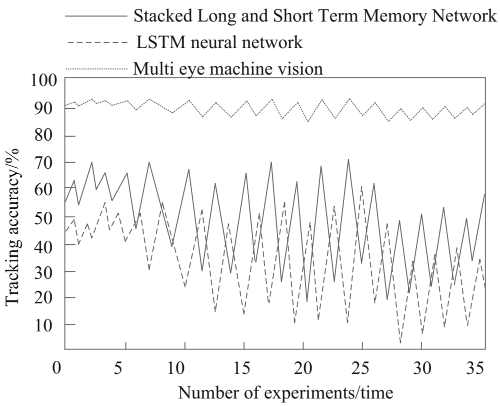


Fig. 9. Comparison and analysis of tracking effects of different methods

It can be seen from Fig. 9 that the maximum tracking accuracy of the tracking method based on the stacked short-term memory network is 71%, the maximum tracking accuracy of the tracking method based on LSTM neural network is 66%, and the maximum tracking accuracy of the tracking method based on multi eye machine vision is 93%.

5 Conclusion

With the continuous development and progress of society, people's attention to multi eye machine vision is increasing, and the research efforts of various institutions in this area have also been strengthened. Many research results have also been applied to our lives, and are widely used in entertainment, production, military and other aspects. This paper proposes a method of tracking athletes' wrong actions based on multi eye machine vision. After image recognition and analysis, the recognition and tracking of athletes' wrong actions are completed. The experimental results show that the research method can effectively track the wrong actions of athletes, and the identification parameters of foul actions are large, which has certain practical significance.

The outlook for future research is that with the continuous development of multi eye machine vision technology, more application scenarios will emerge. On the one hand, multi eye machine vision can play an important role in the field of sports training, helping coaches and athletes more accurately identify and correct incorrect movements, and improving training effectiveness. On the other hand, multi eye machine vision can also be applied in the medical field, helping rehabilitation training and treatment by monitoring and analyzing the patient's movement process. In addition, in the field of intelligent security, multi eye machine vision can provide more comprehensive and accurate monitoring and recognition functions, improving security and early warning capabilities. With further improvements in algorithms and hardware, multi eye machine vision technology will become more intelligent, efficient, and able to adapt to more complex environments and task requirements. In short, the research and application prospects of multi eye machine vision in the future are broad and will have a positive impact on various fields.

References

1. Niu, C., Lu, D., Zheng, Y.: Evaluation method of basketball player micro motions based on stackable long and short term memory network. J. Hunan Univ. Sci. Technol. Nat. Sci. Edn. **37**(02), 95–103 (2022)
2. Yang, S., Yang, J., Li, Z., et al.: Human action recognition based on LSTM neural network. J. Graph. **42**(02), 174–181 (2021)
3. Tao, L.: Application of data mining in the analysis of martial arts athlete competition skills and tactics. J. Healthc. Eng. **36**(3), 557–563 (2021)
4. Zhang, Y., Wang, K., Jiang, J., et al.: Research on intraoperative organ motion tracking method based on fusion of inertial and electromagnetic navigation. IEEE Access **51**(19), 4881–4889 (2021)
5. Wang, Y.: Real-time collection method of athletes' abnormal training data based on machine learning. Mob. Inf. Syst. **12**(3), 1–11 (2021)
6. Gu, K., Li, Y., You, X., et al.: Location tracking scanning method based on multi-focus in confocal coordinate measurement system. Precis. Eng. **32**(9), 170–177 (2021)
7. Liu, Y., Dong, H., Wang, L.: Trampoline motion decomposition method based on deep learning image recognition. Sci. Program. **114**(9), 1–8 (2021)
8. Guo, Y., Liu, Z., Luo, H., et al.: Multi-person multi-camera tracking for live stream videos based on improved motion model and matching cascade. Neurocomputing **492**(12), 561–571 (2022)
9. Wu, Z.: Human motion tracking algorithm based on image segmentation algorithm and Kinect depth information. Math. Probl. Eng. **41**(8), 64–69 (2021)
10. Zhao, J., Zhang, D.: Simulation of human motion information capture in time-space domain based on virtual reality. Comput. Simul. **38**(08), 391–395 (2021)
11. Pan, Y.L.: Run-up track tracking method of 110-meter hurdle athletes based on meanshift algorithm. J. Changchun Univ. **32**(10), 15–19 (2022)
12. Yang, J.: A sensor-based player tracking method. Microcomput. Appl. **39**(03), 12–16 (2023)
13. Wang, Y., Fang, W.C., Ma, J.W.: Posture segmentation based comprehensive assessment of actions. J. Signal Process. **38**(02), 300–308 (2022)
14. Ren, Y., Luo, J.T., Liang, X.P.: Algorithm for detecting occluded basketball players based on adaptive keypoint heatmap. J. Comput.-Aided Des. Comput. Graph. **33**(09), 1450–1456 (2021)
15. Luo, S., Qin, L.R.: Detection of basketball video events and key roles based on attention model. Comput. Appl. Softw. **38**(01), 186–191 (2021)