Basketball Posture Recognition Method in Physical Education Based on Machine Vision

Xiaobo Li[✉] and Zhiqiang Wang

Sports Center, Xian Eurasia University, Xian 710000, China
lixiaobo10034@126.com

Abstract. The traditional basketball posture recognition method has large error and the recognition effect decreases. A basketball posture recognition method based on machine vision is designed for physical education teaching. Collect basketball posture information; Extract the characteristics of basketball movement in physical education teaching; Basketball posture calibration based on machine vision; Then realize the recognition of basketball posture. By means of comparative experiments, it is verified that the attitude recognition effect of the new method is better and has great popularization value.

Keywords: Machine vision · Physical education · Basketball posture · Identification method

1 Introduction

In the 21st century, the innovation of science and technology has changed with each passing day. Cameras, video recorders and various video acquisition devices have been integrated into all aspects of our life. These devices are widely used in traffic management, medical experiments, shopping mall services and so on [1]. It can be said that our life is closely related to these technologies, and almost everyone can't live without these technology-related devices. Nowadays, the video data we can get has shown explosive growth, which brings some difficulties to our traditional video analysis. In the traditional methods, video analysis and other related processes through simple manpower are no longer applicable in this era. Now we often assist people to do corresponding video analysis by means of computer automatic video classification [2]. Its application scenarios involve video surveillance, human-computer interaction, virtual reality, video retrieval and so on [3]. People have done a lot of research on the content of automatic video recognition through computers, but now some video based recognition done by major companies often requires some special cameras, such as Kinect of Microsoft and somatosensory games developed by various game companies [4, 5]. However, there is little research on the videos taken by mobile phone cameras that are often used in our daily life. Indeed, compared with the video taken by the computer, the video taken by the mobile phone has some "jitter" problems. However, according to the psychology

of ordinary people, people are more willing to accept new products that meet the corresponding functions by transforming the objects they often use, rather than buying expensive equipment [6, 7]. Compared with the former, the equipment used in the latter has great universality and practicability.

In recent years, fitness has attracted more and more attention. Almost everyone pays attention to their own body, and everyone will do some fitness activities every day. Basketball is the most popular among fitness sports [8]. Basketball is a collective sport that throws the ball into the opponent's basket to score and prevents the opponent from getting the ball and scoring under the restriction of specific rules. Compared with other ball games, basketball has many technologies, various tactical forms, strong skills of players, and reflects the characteristics of individual combat and cooperation. In the basketball game, the basketball skill level of the players has an obvious impact on the whole team. If the basketball level of the players is insufficient, the weaknesses of the team will be exposed, and the defense and attack level will be greatly reduced, which is not conducive to the performance of the team in the basketball game. Therefore, it is very necessary to carry out scientific and reasonable basketball training for the players [9, 10]. In the traditional way of basketball training, coaches make training plans according to the training and competition of athletes. This method depends on the coaches' training theory and their own experience, and has a certain subjectivity. In addition, it is difficult to avoid wrong movements and possible injuries to athletes' muscles, soft tissues and bones through scientific observation in the training process, which will affect the normal progress of training and even shorten athletes' sports life. From the perspective of training quality evaluation, the evaluation work is operated manually. Coaches need to calculate the training performance of each athlete with reference to different test standards. This method also has some disadvantages [11]. Firstly, the test of athletes is carried out manually, which needs a lot of time for coaches, the process is complex and the accuracy is poor; Secondly, the test method has limitations. It is difficult to measure some important motion parameters such as acceleration and angular velocity directly, but it is impossible to measure the information such as muscle tension, sprint ability and body balance; Third, coaches lack scientific evaluation methods, and it is difficult to formulate corresponding decision-making schemes according to the test data. Therefore, if the sports parameters of athletes can be collected in real time and accurately, the sports posture of athletes can be analyzed and identified, and the training effect evaluation model can be constructed. Based on this, coaches can reasonably adjust the training scheme and scientifically evaluate the training quality, it is of great significance to improve the competitive ability of sports mobilization and the decision-making ability of coaches.

## 2 Design of Basketball Posture Recognition Method in Physical Education Teaching Based on Machine Vision

2.1 Collect Basketball Posture Information

At present, there are two main recognition methods of human posture recognition, namely, recognition technology based on image analysis and recognition technology based on inertial sensor. The recognition technology based on image analysis mainly

recognizes human posture by collecting video, image and other information. Therefore, it is necessary to place cameras and other monitoring equipment in the detection environment in advance for data acquisition. The application of image analysis technology in human posture recognition is early and mature, In addition, the technology uses multiple cameras to detect human action posture from multiple angles, and uses neural network algorithm to train and classify image and video data. Although this method can accurately recognize other people's daily actions, it contains a large amount of data, It is difficult to realize real-time monitoring [12, 13]. There are still many deficiencies in the recognition technology based on image analysis, which requires high precision of the equipment, and the equipment is bulky and inconvenient to carry. Video acquisition is easy to produce dead corners, some places are not easy to be observed, the monitoring range is obviously limited, and the large amount of image acquisition data is easy to lead to insufficient storage and can not achieve the purpose of real-time monitoring. The recognition technology based on inertial sensor makes up for the shortcomings of image recognition technology. The development of science and technology drives the improvement of sensor technology. Sensor equipment has become the best method to obtain human posture information with the characteristics of small volume, high precision, flexibility and easy to wear, low environmental requirements, high sensitivity, low energy consumption and good real-time performance, It is widely used in various fields, such as competitive sports, rehabilitation therapy, somatosensory games and so on. Multiple inertial sensor devices are used together to form a body area network, which has been widely used. Therefore, on this basis, this paper designs the data acquisition module by using the way of sensor.

In the data acquisition stage, the attitude information of human body completing different actions is collected by sensors. The sensor node formed by the combination of multiple sensor devices can convert the action information in the process of action completion into electrical signals for uploading, so as to meet the requirements of subsequent logic operation, data storage and communication. According to the requirements of practical application, it is difficult for a single sensor module to meet the work requirements. The information required in human posture recognition is complex and diverse, including physical and physiological information such as acceleration, angular velocity or heart rate. The analysis and processing work needs to be completed inside the node, so the design of the node needs to include multiple sensor modules, It can be used together to meet the work requirements of the system. Generally, a sensor node includes four modules, as shown in Fig. 1 below.

As shown in Fig. 1, it is mainly composed of processor module, power supply, sensor module and communication module. The processor module controls the normal operation of each functional module of the sensor node and carries out the relevant processing of each signal; The sensor module realizes the function of detecting object motion information and realizes the transformation from motion information to electrical signal; The communication module is responsible for signal transmission and transmitting node data to other devices by wireless means; The power supply provides energy for the normal operation of the whole sensor. At present, mobile devices such as mobile phones have also begun to integrate sensor modules, which have the function of wireless communication. They will replace sensor nodes to wear them on key parts of the human body for
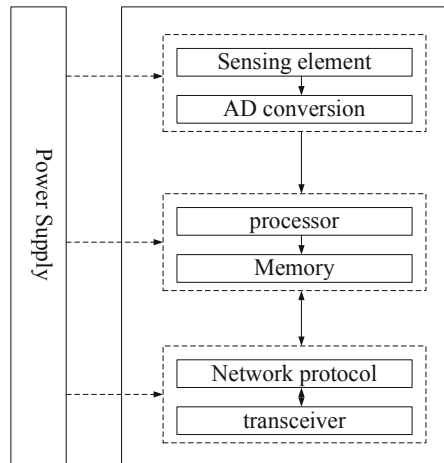
Fig. 1. Sensor node structure

signal acquisition. Compared with sensor nodes, the wearing position of mobile devices is not fixed, which will have an impact on the recognition results of the system. When detecting motion information through sensors, The equipment can be placed in a fixed position to avoid this effect. At present, there are many human posture recognition platforms. Researchers need to build a posture detection platform according to their own research needs.

In data acquisition, multiple sensor nodes are generally used to collect relevant information. In addition to the structure of the node itself, the effective and complete transmission of data is another main problem. At present, according to different data transmission media, data transmission forms are mainly wired and wireless. Wired transmission mode is more stable and reliable, but it is not widely used because of its complex installation and wiring and many restrictions on motion detection; Wireless transmission mode has many advantages in the field of human posture recognition. The commonly used wireless communication technologies include wireless radio frequency identification, Bluetooth, Zig Bee, wireless ultra wideband, etc. wireless transmission mode can reduce the impact of sensors on normal activities, so most systems use this form of data transmission. When designing wireless transmission protocol, we need to consider network architecture, radio technology, communication protocol and energy control. Among the common network topologies, star topology and mesh topology are widely used in practical applications. In the body area network structure, it mainly realizes the aggregation of the data information collected by the nodes attached to the human body, so as to transmit the data to the upper computer for calculation. In the application of body area network, the star topology requires multiple nodes to be directly connected with the receiving node. Its network communication structure is relatively simple and easy to implement, so it is often used. Compared with the star topology, the mesh topology is more complex, but it can use multiple ways to reduce the path loss caused by diffraction, and the data transmission only exists between adjacent nodes, which can keep the node

small transmission energy. In the network protocol setting, we need to set a reasonable network structure according to our own research needs.

2.2 Extraction of Basketball Action Characteristics in Physical Education Teaching

Feature is the abstract representation of a set of data, and the representation of data features can have many forms. At the same time, feature can be equivalent to attribute value, so feature extraction is the process of transforming the original data set into an attribute set that can be used to express these data characteristics. When the amount of input data set is too large and some of the data is redundant, it is necessary to use the feature extraction algorithm to extract the feature attribute set of input data to form the feature vector. As a representative feature set, it simplifies the data set to a certain extent, improves the operation efficiency of the algorithm and reduces the complexity. Generally, the combination of feature extraction and feature selection is used to construct the appropriate feature vector. The selection of the appropriate feature set is very key to the recognition of human posture. After extracting the data set representing each action through data division, analyze the data of unit action to extract the attribute set of reaction characteristics. This process is the feature extraction process [14]. The main methods of feature extraction are to analyze the characteristics of different data and extract the corresponding features. At present, the more common features are time-domain features, frequency-domain features and time-frequency features. Time domain features: there are many time domain features, which are often used in research: mean, variance, extreme value, standard deviation, covariance, correlation coefficient, root mean square, etc. It is also called signal statistical feature. Its calculation method is simple and low complexity. It is often used in the field of human pose recognition. Frequency domain characteristics: in order to obtain the frequency domain characteristics of the signal, the time domain signal is usually transformed into the frequency domain signal, which is generally realized by Fourier transform. The frequency domain signal represents the components of the signal at different frequencies, which is different from the time domain characteristics, and reflects the signal characteristics on the other hand. The commonly used frequency domain features include frequency domain line, spectral energy, etc., and the Fourier transform coefficient $L + L$ and its absolute value are generally used as frequency domain features for reference. Time frequency feature: the time-frequency feature reflects the characteristics of the signal in time domain and frequency domain. This feature is mainly used for feature extraction of unstable signals. In the process of time-frequency feature extraction, wavelet analysis is mainly used to realize the extraction, and the local transformation of frequency and time can be used to realize multi-scale thinning of the signal.

Feature selection is a variable selection method, also known as attribute selection or variable subset selection. It is a process used to select relevant attribute subsets in order to build a classification model. The primary reason why feature selection is proposed is that in the feature set obtained by feature extraction, not all attributes are relevant and useful, and some attribute selection may be redundant. The introduction of irrelevant attributes not only has no effect on the construction of the model, but also makes the constructed model more complex due to the redundancy and irrelevance of data. Therefore,

reasonable feature screening is extremely necessary. There is a big difference between feature selection and feature extraction. The purpose of feature extraction is to extract feature vectors from the original data, and feature selection is to select appropriate feature sets from these feature vectors. There are three main purposes of feature selection: (1) simplify the model and reduce the computational complexity; (2) Shorten training time; (3) Strengthen the promotion to avoid the problem of over fitting. The commonly used feature selection algorithms are generally obtained by combining the evaluation function with sequential forward / backward search, decision tree, best first search and genetic algorithm. Among them, the evaluation algorithm is a function that can reflect the advantages and disadvantages of the selected feature subset, which can be used to solve the correlation between feature and classification, classifier error rate and so on.

Since the concept of machine vision was put forward, many large companies, research institutes and universities have made unremitting efforts in this field. At present, they have made great progress in many fields. Many products of machine vision have been widely used in our society and life, such as human-computer interaction, virtual reality, video retrieval, video surveillance and so on. Most of the major companies do the corresponding research on machine vision to meet their own needs, or to explore the way and clear the obstacles for the future of the company. Therefore, each company has its own focus, and each company also has its own advantages and disadvantages. In various applications of machine vision, gesture recognition based on video is a popular and indispensable direction in the research direction of human-computer interaction. Gestures are closely related to our lives. Almost everyone has to make different gestures every day, such as waving, clapping, boxing and so on. Among the gestures people make, some gestures are complex and some gestures are simple. But every gesture is basically inseparable from the beginning, process and end of the gesture. For the research on the application of gesture recognition based on video, we can roughly make these classifications: one-dimensional gesture recognition, two-dimensional gesture recognition and three-dimensional gesture recognition. The specific application scenarios of gesture recognition of these three classifications are very different. One dimensional gesture recognition is mainly used to recognize some static hand types. It is relatively simple. The gestures he wants to recognize are almost unchanged. For example, you need to recognize a static V hand type, a static fist hand type, and a static number 0 hand type; Compared with one-dimensional gesture recognition, two-dimensional gesture recognition still has no relevant depth information. Its process is mainly detected and recognized by the changing characteristics of a gesture action in the two-dimensional plane. Although this kind of recognition extracts more features than one-dimensional gesture recognition, it is difficult to recognize because it not only recognizes simple hand shape, but a series of continuous actions without corresponding depth information. Because modern people often carry some devices, such as mobile phones and cameras, their cameras have no depth information, but these devices are widely used now, and almost everyone is using these devices. Therefore, two-dimensional gesture recognition still has corresponding research prospects and application scenarios. Compared with two-dimensional gesture recognition, three-dimensional gesture recognition has a layer of depth information, so its recognition efficiency is more accurate than two-dimensional gesture recognition, but the calculation of three-dimensional gesture recognition is much

larger than that of two-dimensional gesture recognition. Therefore, three-dimensional gesture recognition is generally applied to some devices equipped with depth cameras, It can not be used for mobile phones and other equipment, such as home game consoles, some security locks of security companies, places with large personnel mobility in hotel amusement parks, etc.

In this paper, feature extraction is divided into two categories, one is global analysis method, the other is local analysis method. The overall analysis method usually extracts the foreground of the video by means of background subtraction method, optical flow method and difference method to obtain the corresponding binary image, then obtains the overall region of interest through the binary image, and then processes these regions of interest accordingly. The next step is to obtain the contour and edge information of these regions of interest. However, in these information, we may often contain a lot of unnecessary and redundant information, which we call noise. How to filter these noises and obtain the corresponding accurate eigenvalues is often the focus of feature extraction. People have done a lot of research in this area. Extracting video eigenvalues can distinguish the corresponding eigenvalues. This method mainly obtains the contours through the background subtraction method. These contour information often contains corresponding energy information. The second step of this method is to analyze the energy of these contours, in which the analysis mainly analyzes the corresponding motion potential energy, and kinetic energy, Then the corresponding video contour energy map is obtained by analysis, and the motion history map of the contour is obtained by using the motion image. Because the process of obtaining eigenvalues by this method is single perspective and the perspective is often unchanged, it is easy to obtain some features of invariant moments.

2.3   Calibrating Basketball Posture Based on Machine Vision

After the basketball feature extraction above, this paper will further calibrate the basketball posture based on machine vision. In general, in the extracted feature set, only the largest and second largest features are taken as the corresponding real feature values, while for other smaller features, they are deleted on the current interface, so as to ensure that there are two external rectangular features of the image of each frame, Prevent the situation that more than two features are drawn in a connected area. This prevents "system noise". By removing these noises, we can further improve the accuracy of our recognition. The principle of machine vision is shown in Fig. 2 below.

In Fig. 2, P represents the light source, X_1 and X_R represent the left and right cameras, d represents the distance between the left and right cameras, L and R represent the object to be measured, B represents the left and right boundary distance of the object to be measured, D represents the distance between the camera and the object to be measured, and Z represents the distance between the light source and the object to be measured. As shown in Fig. 2, the original data signals collected by the machine vision equipment often have noise signals generated by the influence of the external environment and themselves. These signals are inaccurate. Therefore, it is very necessary to denoise the original data signals. There are many methods to denoise data information, and the signal denoising method implemented in software design is generally called digital filtering. This technology mainly includes classical filter and modern filter. Among them, the
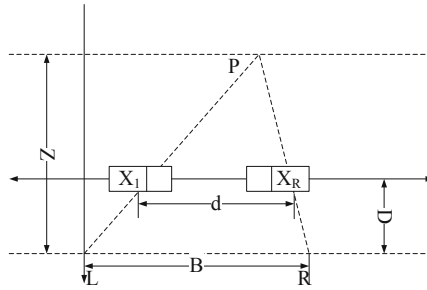
Fig. 2. Principle of machine vision

former first assumes that the useful signal and noise signal in the signal are located in different frequency bands, and the noise signal can be removed simply by placing the signal in the linear system. The common high-pass and low-pass filters are constructed based on the principle of different frequency band distribution of the signal. However, the classical filter has defects, which is no longer applicable to the original signal data with overlapping frequency bands of noise and signal. For this type of signal, the modern filter is generally used. Different from the classical filter principle, the modern filter classifies the useful signal and noise into random signals, and then deduces and estimates the characteristics of the useful signal or noise by using its corresponding statistical characteristics such as autocorrelation function and self power spectrum. The common modern filter algorithms mainly include Kalman filter, Wiener filter and so on.

Normalization, also known as standardization, is a basic step in data information mining and plays a role in simplifying calculation. In the process of human pose recognition, normalization method is also used to process the data. Because in the body area network system, the placement positions of nodes are different, and in the process of movement, nodes will shake with the movement of human body. The change of node position information will cause a certain deviation in the collected data, which may reduce the accuracy of human posture recognition. Data normalization can eliminate the dimensional influence between different data and solve the problem of data comparability. The normalization processing means is to scale the data to a certain range, and transform its expression, so that different data can be comprehensively compared and evaluated. Two methods of data normalization processing will be introduced in detail below. First, linear function transformation. This method maps the original data to the closed interval between 0 and 1 through linear transformation, and can complete the equal scaling function of the original data signal. The specific effects are as follows.

$$X_{norm} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \tag{1}$$

In formula (1), X_{norm} is the result of linear function conversion; X is the original data; $X_{\max}$ is the maximum value of sample data; $X_{\min}$ is the minimum value of sample data. On this basis, the 0-means normalization method is used to process the original data into a normal distribution set with mean value of 0 and variance of 1, as shown below.

$$y = \frac{x - \mu}{\delta} \tag{2}$$

In formula (2), x data set to be processed; y is the result obtained after calculation; μ is the mean value of the original data; δ is the variance of the original data. In order to solve the attitude accurately and reduce the noise interference of the sensor, the angular velocity, acceleration and magnetic field strength are fused in the process of calculating the node attitude, and the quaternion method is used to represent the space attitude, and then the attitude is calibrated. Quaternions are composed of a real number and three imaginary units. The definition formula is as follows:

$$Q = w + x_i + y_j + z_k \tag{3}$$

In formula (3), w, x, y and 4 are real numbers; i. j and k are three imaginary units; Q is a unit quaternion, which is normalized as follows:

$$Q_{norm} = \frac{Q}{\sqrt{w^2 + x^2 + y^2 + z^2}} \tag{4}$$

In formula (4), Q_{norm} represents the normalized quaternion. In this paper, it is applied to describe the rotation of rigid bodies in three-dimensional space, as follows:

$$\hat{Q} = 0.5 \cdot Q \cdot p \tag{5}$$

$$p = 0 + w_{xi} + w_{yj} + w_{zk} \tag{6}$$

In formula (5–6), $\hat{Q}$ is the derivative of quaternion to time; p is the angular velocity. Through visual conversion, the unit quaternion used to describe the conversion of rigid body from one attitude to the next can be obtained. The update formula is as follows:

$$Q_{k+1} = Q_k + \Delta t \cdot \hat{Q}_k \tag{7}$$

$$\hat{Q}_k = 0.5 \cdot Q_k \cdot p_k \tag{8}$$

In formula (7–8), at this time, k is a non negative integer; Q_{k+1} is the unit quaternion of basketball posture at $k+1$ times; Q_k is the unit quaternion of basketball posture at k times; Δt is the time interval between two samples; $\hat{Q}_k$ is the derivative of quaternion to time at time k; p_k is the quaternion at time k. Calibrate the quaternion as follows:

$$x_{k+1} = \phi_k x_k + w_k \tag{9}$$

$$z_k = H_k x_k + w_k \tag{10}$$

In formula (9–10), x_{k+1} is the basketball posture calibration state at time $k+1$; ϕ_k is the state transition matrix; x_k is the basketball posture calibration state at time k; w_k is the noise vector; z_k is the secondary measured value; H_k is the relationship between basketball teaching posture and visual measurement in the ideal state. Through this calibration state, basketball teaching posture can be accurately identified.
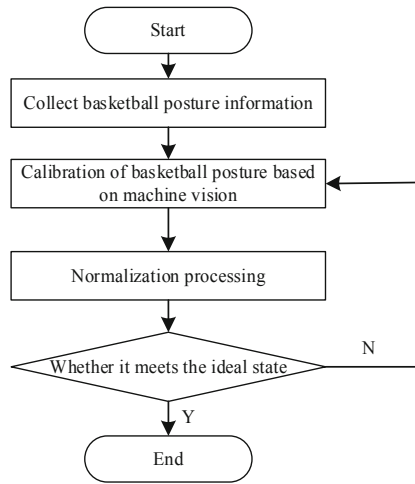
Fig. 3. Basketball posture recognition process

2.4 Recognition of Basketball Posture

The implementation process of basketball posture recognition method based on machine vision is shown in Fig. 3.

In order to realize the recognition of basketball posture, the machine vision equipment is fixed on the back of the basketball player's hand, and the acceleration and angular velocity data of the player's hand in the process of jump shot are collected. The jump shot process is divided into four stages, the shooting posture of each stage is analyzed, and the sound feedback reminder is used to help the athletes correct the shooting posture. However, the jump shot posture of athletes is reflected by the movement posture of their arms and legs. This method only analyzes the posture of hands in the shooting process, and lacks the consideration of the posture of other parts of the body. By collecting the data of athletes' heart rate, oxygen consumption and acceleration, this paper analyzes the physiological characteristics of basketball players in the process of sports, compares the physiological performance of athletes in different forms of competition and training, and proves that wearable devices are very helpful to the quantification of basketball, but there is no specific research on the quantification of basketball; The basketball posture recognition system is constructed by using acceleration machine vision, the lower limb data of basketball players are collected, and the recognition of 8 kinds of actions in basketball is completed. In this paper, only the acceleration data is used as a reference, the characteristics of the constructed data set are single, and the average recognition accuracy is less than 70%. There are many researches on basketball based on machine vision, but they mainly focus on the movement information of some limbs in basketball, such as the hand posture in the shooting process, the lower limb state in the movement process or the physiological characteristics of basketball players. There is little research on the comprehensive analysis of basketball players' upper and lower limb movements. Basketball movement is a complex movement completed by the upper and lower limbs. The recognition of basic basketball movements plays an important role in improving

the skills of basketball players. Therefore, this paper mainly studies the classification of basic movements of upper and lower limbs in basketball, and preliminarily realizes the recognition of basic basketball movements.

3   Experiment and Analysis

In order to verify whether the method designed in this paper has practical effect, this paper carries out experiments on the above methods. The experimental process and results are as follows.

3.1   Experimental Process

The PC of this training and testing program is configured with Intel Core 2 2.70 GHz processor, 2 G memory, Windows XP operating system, and Visual Studio 2005 and OpenCV programming environment. The original resolution of the camera is 1024 × 768, the resolution of the collected video image is adjusted to 512 × 384 by using the interpolation algorithm. The experiment starts at the angle of 0° and ends at the angle of 360°. The sensor node is continuously rotated and the data is sampled every 45°. The experiment is divided into two groups. In order to compare the recognition effect of the method designed in this paper, the first group adopts the traditional attitude recognition method; The second group of the experiment uses the method designed in this paper to recognize the running posture of basketball.

3.2   Experimental Results and Discussion

In the above experimental environment, two methods are used to recognize the basketball action posture in the collected image, and compared with the standard action. The smaller the error between the standard action and the standard action, the better the recognition effect of the corresponding method. The experimental results are shown in Table 1.

Table 1. Experimental results

Rotation angle	Recognition error of traditional attitude recognition methods	The recognition error of the attitude recognition method designed in this paper
0	0.25	0.00
45	0.29	0.03
90	0.32	0.04
135	0.68	0.08
180	0.87	0.12
225	1.38	0.36

(*continued*)

Table 1. (*continued*)

Rotation angle	Recognition error of traditional attitude recognition methods	The recognition error of the attitude recognition method designed in this paper
270	2.72	0.44
315	2.95	0.48
360	3.03	0.51

As shown in Table 1, the traditional attitude recognition method has large recognition error and poor recognition effect; The attitude recognition method designed in this paper has small recognition error and good recognition effect, which is in line with the purpose of this paper.

4   Conclusion

Basketball training is the process of improving competitive ability. It refers to the educational process specially organized to continuously improve and maintain the competitive level of athletes under the guidance of coaches and the participation of athletes. The content of basketball training mainly includes five aspects: Athletes' physical training, technical training, tactical training, psychological training and intelligent training. In the competition, the athletes' technical movement is unstable and do not play well, and they can't keep up with the changes of their opponents and miss people in defense. On the one hand, it shows that the basic skills are not solid, and it also shows that the trainers and coaches don't pay attention to the basic skills practice in peacetime, so it will lead to the athletes' repeated mistakes and omissions due to the basic skills in the competition. Therefore, we should strengthen the basic technical training of athletes. In this paper, the recognition method of basketball action posture in physical education teaching based on machine vision is designed. By collecting basketball action posture information, the basketball action characteristics in physical education teaching are extracted, and the basketball action posture is calibrated by machine vision method, in order to improve the recognition effect, reduce the recognition error, and provide guiding suggestions for the standardization of basketball.

References

1. Wang, S.M., Liao, C.L., Ni, Y.Q.: A machine vision system based on driving recorder for automatic inspection of rail curvature. IEEE Sens. J. **99**, 1–1 (2020)
2. Zhao, C., Ding, H., Cao, G., et al.: A new method for detecting compensation hole parameters of automobile brake master cylinder based on machine vision. J. Adv. Trans. **3**, 1–14 (2021)
3. Sergiyenko, O., Tyrsa, V.: 3D optical machine vision sensors with intelligent data management for robotic swarm navigation improvement. IEEE Sens. J. (99), 1–1 (2020)
4. Xue, B., Wu, Z.: Key technologies of steel plate surface defect detection system based on artificial intelligence machine vision. Wirel. Commun. Mob. Comput. **2021**(6), 1–12 (2021)

5. Qy, A., Jin, F.A., Jian, T.B., et al.: Development of an automatic monitoring system for rice light-trap pests based on machine vision. J. Integr. Agric. **19**(10), 2500–2513 (2020)

6. Liu, S., Wang, S., Liu, X., et al.: Fuzzy detection aided real-time and robust visual tracking under complex environments. IEEE Trans. Fuzzy Syst. **29**(1), 90–102 (2021)

7. Zhang, Y., Hong, G.S., Ye, D., et al.: Powder-bed fusion process monitoring by machine vision with hybrid convolutional neural networks. IEEE Trans. Industr. Inf. **16**(9), 5769–5779 (2020)

8. Chen, Y., Shu, Y., Li, X., et al.: Research on detection algorithm of lithium battery surface defects based on embedded machine vision. J. Intell. Fuzzy Syst. **5**, 1–9 (2021)

9. Ali, M., Atia, M.R.: A lead through approach for programming a welding arm robot using machine vision. Robotica, 1–11 (2021)

10. Chen, J.H., Zhang, J.J., Gao, R.J., et al.: Research on modified algorithms of cylindrical external thread profile based on machine vision. Measure. Sci. Rev. **20**(1), 15–21 (2020)

11. Liu, S., Liu, D., Muhammad, K., Ding, W.: Effective template update mechanism in visual tracking with background clutter. Neurocomputing **458**, 615–625 (2021)

12. Li, M., Jia, J., Lu, X., et al.: A method of surface defect detection of irregular industrial products based on machine vision. Wirel. Commun. Mob. Comput. **1**, 1–10 (2021)

13. Liu, W.: Simulation of human body local feature points recognition based on machine learning. Comput. Simul. **38**(6), 387–390,395 (2021)

14. Wang, S., Liu, X., Liu, S., et al. Human short-long term cognitive memory mechanism for visual monitoring in IoT-assisted smart cities. IEEE Internet of Things J. **9**(10), 7128–7139 (2021). https://doi.org/10.1109/JIOT.2021.3077600