Research Article

Method of Analyzing and Managing Volleyball Action by Using Action Sensor of Mobile Device

Xu Sun,[1] Kai Zhao [iD],[2] Wei Jiang,[2] and Xinlong Jin[3]

[1]*Volleyball Teaching and Research Section, Shenyang Sport University, Shenyang, 110000 Liaoning, China*
[2]*China Volleyball College, Beijing Sport University, Beijing, 100089 Beijing, China*
[3]*Academic Affairs Office, Dalian University of Science and Technology, Dalian, 116000 Liaoning, China*

Correspondence should be addressed to Kai Zhao; paiqiu@bsu.edu.cn

With the development of electronic technology and sensor technology, more and more intelligent electronic devices integrate micro inertial sensors, which makes the research of human action recognition based on action sensing data have great application value. Data-based action recognition is a new research direction in the field of pattern recognition, which is essentially a process of action data acquisition, feature extraction, feature extraction, and recognition, the process of classification and recognition. Inertial motion information includes acceleration and angular velocity information, which is ubiquitous in daily life. Compared with motion recognition based on visual information, it can more directly reflect the meaning of action. This study mainly discusses the method of analyzing and managing volleyball action by using the action sensor of mobile device. Based on the motion recognition algorithm of support vector machine, the motion recognition process of support vector machine is constructed. When the data terminal and gateway of volleyball players are not in the same LAN, the classification algorithm classifies the samples to be tested through the characteristic data, which directly affects the recognition results. In this paper, the support vector machine algorithm is selected as the data classification algorithm, and the calculation of the classification process is reduced by designing an appropriate kernel function. For multiclass problems, the hierarchical structure of directed acyclic graph is optimized to improve the recognition rate. We need to bind motion sensors to human joints. In order to realize real-time recognition of human motion, mobile devices need to add windows to the motion capture data, that is, divide the data into a small sequence of specified length, and provide more application scenarios for the device. This method of embedding motion sensors into devices to read motion information is widely used, which provides a convenient data acquisition method for human motion pattern recognition based on motion information. The multiclassification support vector machine algorithm is used to train the classification algorithm model with action data. When the signal strength of the sensor is 90 t and the speed is 2.0 m/s and 0.5 m/s, the detection accuracy of the adaptive threshold is 93% and 95%, respectively. The results show that the SVM method based on hybrid kernel function can greatly improve the recognition accuracy of volleyball stroke, and the recognition time is short.

1. Introduction

In recent years, more and more researchers have begun to analyze the movements of the human body, and the field of research is also expanding. From medical rehabilitation engineering, motion sensing games to the production of movies and TV works, virtual reality, and professional sports analysis, human action analysis and recognition technologies are more and more widely used and have produced huge

value. Most of the mature research on the understanding of human behavior is based on the analysis of video or image sequences. The motion recognition method based on optical motion capture has obvious advantages. But it also has the most basic disadvantage: the acquisition of motion data is uncertain.

Using quantitative posture data as input data, the basic characteristics of motion can be reconstructed. In the aspect of action recognition, this method uses the support vector

machine algorithm to classify the longest common part as a kernel function to integrate, train, and compare the similarity of the time series of daily sports and realize common classification and recognition. Using the longest common subsequence as the kernel function of the support vector machine is completely different from the traditional time-based single point information. Using the time series information contained in sports, design a classification system for volleyball actions based on spatial methods.

Motion sensor technology and automatic fall detection systems have become reliable and low-cost solutions for falls. Yu et al. have developed a fall detection system based on Hidden Markov Model (HMM), which can use a single motion sensor to automatically detect falls for actual home monitoring scenarios. They proposed a new representation for the acceleration signal in HMM to avoid feature engineering and developed a sensor orientation calibration algorithm to solve the problem of sensor misalignment in actual scenes (misaligned sensor position and misaligned sensor orientation). The HMM classifier is trained to detect falls based on the acceleration signal data collected from the motion sensor. They collect data sets from experiments that simulate falls and normal activities. Their research process lacks theoretical foundation [1]. The smart phone sensor can measure the unique behavioral characteristics of the user when interacting with the smart phone according to the user's different habits, gestures, and angle preferences of touch operations. Shen et al. studied the reliability and applicability of active and continuous smart phone authentication using motion sensor behavior in various operating scenarios and systematically evaluated the uniqueness and durability of behavior. For each sample of sensor behavior, motion information sequences must be extracted and analyzed. These information sequences have statistics, frequency, and wavelet domain characteristics to provide accurate and fine-grained representations of user touch actions. Their research process lacks data [2]. Yurtman and Barshan proposed a novel noniterative direction estimation method based on the physical and geometric characteristics of acceleration, angular velocity, and magnetic field vector to estimate the direction of the motion sensor unit. They obtain the orientation of the sensor unit according to the rotation quaternion transformation between the sensor unit frames. He evaluated the proposed method by incorporating it into an activity recognition scheme for daily and sports activities, which requires accurate estimation of the orientation of the sensor unit in order to achieve the invariance of the orientation of the unit on the body. His research process lacks experimental data [3]. Signature recognition is to identify the owner of the signature, and verification is the process of finding the authenticity of the signature. Although both are important in the field of forensic science, verification is even more important for banks and credit card companies. Behera et al. proposed a method to analyze 3D signatures captured using leap motion sensors. They extended the original 2D function from the original signature to 3D and applied a well-known classifier for identification and verification. They used the leap sports interface to create a large data set containing more than 2,000 signatures registered by 100 volunteers. Their research method is not novel enough [4].

This research mainly uses the support vector machine method based on the hybrid kernel function to determine the movement recognition of the human body in the volleyball movement database. First, the moving images in the database are processed, and an action database suitable for the SVM algorithm is established. Then, PCA dimensionality reduction was performed before classification. Finally, the SVM classifier was used to identify the action type. In the research process of volleyball motion recognition, this research will focus on the SVM algorithm based on the hybrid kernel function and compare it with the experimental template matching method. The SVM algorithm based on the hybrid kernel function proves that the recognition of volleyball stroke is effective and fast.

2. Volleyball Stroke

2.1. Internet of Things Technology. The Internet of Things technology can use smart terminals, communication base stations, data processing, and other display devices to optimize the collection, processing, and monitoring of information. The Internet of Things needs to meet special needs when applied, such as low power consumption, strong coverage, and low cost. In order to meet these needs, various industry organizations have formulated a series of communication standards. The current system based on the Internet of Things technology can meet the requirements of low cost and low power consumption, but there are still problems in security and capacity, so the demand for new Internet technology standards is becoming more and more urgent [5, 6]. The combination of wireless communication technology and wireless automatic identification technology is a network based on the computer Internet. On the basis of the Internet, sensors and controllers are particularly important. What we expect is that all devices or test objects on the network can communicate information without manual intervention [7].

2.2. Motion Feature Recognition. Motion detection is a general preprocessing process for computer vision applications, but motion detection is easily affected by the environment, for example, in scenes such as changes in light, dynamic background, camera movement, and shadows. It is very difficult to correctly detect moving targets. Currently, background subtraction, interframe difference, and optical flow have been widely used in motion detection [8, 9].

Median filtering is a type of nonlinear filtering that can effectively remove noise. The values around the calculated points are arranged in accordance with the gray value, and the gray value of the median value of the placed point is substituted by the gray value of the calculated point, thereby realizing the filtering function. This method can effectively remove noise and high-frequency deviation in Fourier space, so that the image is smoother and the resolution will not be reduced. If the pixel value of the point (x, y) is $f(x, y)$, and the pixel value of the corresponding point after median filtering is $g(x, y)$, the relevant formula is as follows [10]:

$$g(x, y) = \text{median}\{f(x - m, y - n), (m, n) \in Q\}. \quad (1)$$

Among them, Q is the size of the module. After feature representation, classifiers in machine learning are often used to recognize actions. Supervised learning is to build a model that represents the distribution between the classification label and the input features. The unsupervised learning mode is to directly learn training actions from the entire training data, that is, without inputting classification labels. The choice of the number of median filter banks depends on the actual situation. If the number is too small, the texture feature of the image cannot be fully expressed, and if the number is too large, the calculation will be complicated. Particularly during deep training, since deep neural networks can automatically learn image features, if the number of channels is set too much, the training speed will be greatly reduced, the memory footprint will be large, and the learning will be repeated between features, which will affect the overall robustness of the network [11, 12]. The parameters of the Gaussian function most suitable for pixel n can be obtained by a_n and k_n [13].

$$D(\alpha_n, k_n, z_n) = \frac{1}{2} \log \det \sum (\alpha_n, k_n) + \frac{1}{2} [z_n - \mu(\alpha_n, k_n)]^T$$
$$\cdot \sum (\alpha_n, k_n)^{-1} [z_n - \mu(\alpha_n, k_n)]. \quad (2)$$

Among them, z_n is the RGB value of pixel n. In actual application scenarios, the observation of the target is often multiangle and multidirectional, so these properties are very important for the effective extraction of texture features. They also directly determine the application field of the median filter, and compared to other edge filters, the median filter is closer to the continuously variable condition [14, 15].

2.3. Support Vector Machine.

Support vector machine (SVM) is a classic method in the field of pattern recognition and machine learning. It is a supervised learning model related to learning and training. It can be used for data analysis, pattern recognition classification, and regression. The support vector machine algorithm is based on the structural risk minimization of statistical learning methods and the VC dimension theory. Specifically, the support vector machine is constructed by eigenvalue dimensions, a high-dimensional space, in which a hyperplane is established so that the data categories are distributed on both sides of the hyperplane, and the hyperplane is optimized to make the separation between the separated data categories and the hyperplane. Maximize to get the optimal hyperplane. Motion characteristics are obtained by estimating continuous interframe changes. The local motion feature is based on a nonpredetermined target area motion analysis method, which directly searches for points of interest or regions from images or videos and describes their motion information [16]. The discrete form of the transformation function is

$$s_k = T(r_k) = \sum_{j=0}^{k} P_r(r_j) = \sum_{j=0}^{k} \frac{n_k}{n} k = 0, 1, 2, \cdots, L - 1, \quad (3)$$

where n is the sum of pixels in the image and n_k is the number of pixels whose gray level is n. Compared with the global shape feature, the different subshapes in the local shape feature are considered relatively independent, so the description of the shape is not as good as the global feature complete and accurate. However, because the local shape features are relatively less affected by the human body detection and tracking results, they can be used in complex scenes [17, 18]. Therefore, the heart in the initial state probability σ is defined as

$$(\sigma_{t+1}^i)^2 = (1 - \rho)(\sigma_u)^2 + \rho(x - \mu)^T (x_i - \mu_j). \quad (4)$$

It represents the probability that the first frame of this type of action is in the state ρ. Different initial models will produce different training results. The key issue of the application of the SVM model in action recognition is the determination of the hidden state. The traditional method is to divide each action sequence into N segments, and each segment corresponds to a state in the SVM model. In this way, we have added a penalty for samples that are out of bounds in the problem of maximizing the interval [19]. The trade-off between the two is controlled by the parameter C. The Lagrange equation [20, 21] at this time becomes

$$L(w, b, a) = \frac{1}{2} w^2 + C \sum_{n=1}^{N} \xi_n - \sum_{n=1}^{N} a_n \{ y_n (w^T x_n + b) - 1 \} - \sum_{n=1}^{N} \mu_n \xi_n. \quad (5)$$

Among them, a is the Lagrangian multiplier.

2.4. Motion Sensor.

Generally, when selecting sensor positions, researchers often only consider the importance of different sensor positions to the action category. When a position is selected, the subsequent position selection will be selected according to the importance of the remaining position to the action category [22, 23]. If each position is not related to each other, then there is no problem with this method. However, in practice, the positions are related to each other, and they may contain the same information. If only the correlation between the sensor position and the action category is considered, the result will be the positions that cause serious overlap are selected at the same time, and the importance is relatively small, but the position and the selected position do not have any overlap are often ignored, so that some action categories that can only be distinguished by the ignored position cannot be well recognized [24]. When a single-turn optical fiber coil with a diameter of D rotates, two light waves are emitted clockwise (CW) and counterclockwise (CCW) from point P of the circular optical path, then

$$T_{CW} = \frac{\pi D}{V_{CW}} = \frac{\pi D}{V_f - \Omega D} = \frac{\pi D}{(c/n) - \Omega D}. \qquad (6)$$

Among them, c is the speed in a vacuum [25, 26].

3. Volleyball Stroke Management Experiment

3.1. Support Vector Machine (SVM) Action Recognition Algorithm Flow Design. The sensor node is the core of the system, and each node is fully integrated with the accelerometer and the inertial measurement unit of the magnetometer. The motion capture system connects all sensor nodes through wires and finally collects the motion capture data on the hub and realizes wired or wireless output through the hub based on the Internet of Things technology. The design process of the SVM algorithm is shown in Figure 1.

(1) The foreground part of the image is further segmented by the method based on image segmentation. Assign a pixel value to each divided block. In this way, the finally obtained image is composed of several pixel blocks, which have different pixel values for distinguishing each other

(2) After the first step, we will get the segmented image block. Refine each image block and extract bones. After all the image blocks are extracted, the results are added together to form the image T

(3) Perform vertical expansion on the obtained image T. The purpose of this expansion is to connect the obtained bone points to each other, because there may be gaps between the bone points obtained in the previous step, and they are not connected to each other

(4) Set a threshold, filter the image T, and filter out the smaller area (less than the threshold). Used to eliminate the interference of some noise points

3.2. Design of Foreground Detection Module. The foreground detection module takes the current frame as the input signal and then judges the value of each pixel to obtain a binary image of the same size as the current frame. If the pixel value is 1, then the pixel is the detected moving image pixel. If the pixel value is 0, it means that the pixel is a background pixel. Foreground detection is the basis of the entire support vector machine algorithm. In practical applications, the quality of the foreground detection module will directly affect the detection effect, and the foreground detection module processes pixel-level data, so the amount of calculation is very large. Generally, the relationship between algorithm complexity and foreground extraction effect is considered [27, 28].

3.3. Status Query Module Design. The cloud server functions as a WAN gateway in the Internet of Things technology. When the data terminal and the gateway of the volleyball swimmer are not in the same local area network, the cloud server needs to be used for communication and connection.

In addition, the security key comparison, the user's mode settings, and other configuration files are also stored in the cloud server.

3.4. Design of Volleyball Data Acquisition Module. To achieve the overall range of motion capture of the human body motion, more than 10 sensor nodes are required, and motion sensors need to be bound to the limb joints of the human body for motion capture. The sensors can detect the human body through the built-in angular rate gyroscope and accelerometer. Collect the motion information in the motion process. For the motion data collected by the motion sensor, the data acquisition module is required to process and collect the data. After the system is initialized, these sensor nodes will automatically build a sensor network using ZigBee technology to calculate the posture of each part of the human body independently and in real time. Then, under the scheduling of the sink node, the posture information of each part of the human body in the wireless sensor network is orderly sent to the sink node. Finally, the sink node uniformly uploads the data of each sensor node to the host computer, which calculates the final action form of the human body.

The function of the motion sensor data acquisition module is to process the motion data collected by the motion sensor equipment, including the following: (1) Analyze the obtained binary data to obtain the real data information of the motion. (2) Perform initial calculations on exercise data. (3) Convert the format of the motion data to a format that is easy to handle. (4) The movement data is handed over to the data transmission module, and the transmission module transmits the data to the terminal PC.

3.5. Design of Volleyball Data Storage Module. For the storage of sports data, the SaveDataFile method is implemented in the CDataManager class. The data storage format is completely in accordance with the data format received by the serial port and is saved in a txt file in binary form. The advantage of using this storage format is that the same data decoding method as when the serial port receives data can be used when the system software performs historical movement reproduction. According to the stored data in the form of a matrix, the effective data during the volleyball movement process are extracted, including the offset position information and the rotation angle information. Combining these two sets of data, the posture analysis algorithm is used to calculate the movement trajectory of the discrete data at the main joint points of the human body.

3.6. Trajectory Image Generation Module Design. The trajectory generation module takes the clump information, foreground mask information, and volleyball stroke data information as input information and then draws the corresponding motion trajectory information according to the position information of each clump and finally retains the motion trajectory information in the designated file. The trajectory generation module is the basis of the subsequent OPENCV program. If the module can give an effective moving object trajectory curve, it will greatly facilitate the
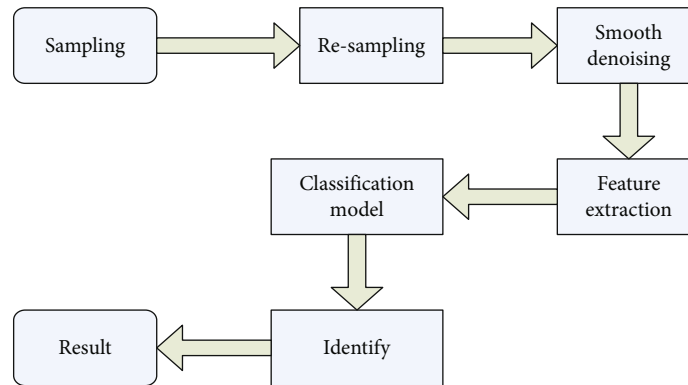
FIGURE 1: SVM algorithm design process.

convenience of the subsequent program to study the moving object. In real life, the algorithm itself is generally not considered complexity, usually choose the movement trend that can smoothly draw the trajectory curve, and prepare for the program to run.

3.7. Data Windowing. In order to realize the real-time recognition of human movements, it is necessary to window the motion capture data, that is, to divide the data into a short sequence of specified length. At the same time, in order to enhance the ability to describe actions, there is a 50% overlap between adjacent windows. The length of the window is related to the frequency of the human body's behavior. If the length of the window is shorter, the cycle of the human body motion included is also less, and the recognition effect is worse. If the length of the window is longer, the cycle of the human body motion included is also more, the more stable the recognition effect. But the length of the window should not be too large; otherwise, it will cause system delay. Considering that the sampling rate of the motion sensor used in this study is 100 Hz, this study sets the length of the window to 230.

3.8. Denoising Processing of Volleyball Image. The key to median filtering is to replace the local average value with the local median value. In the $N \times N$ $(N = 2n + 1)$ mask window centered on the pixel (x, y) of the gray image F, the gray values of the $N \times N$ pixels are arranged in descending order. The gray value of the middle position is defined as a, a is called the median value, and then $f(x, y) = a$. The median filter will affect the impact function. When half of the window is larger than the width of the alluvial function, the impact function tends to disappear, and the upper part of the triangular function is flattened, so edge blur can be avoided and discrete impact noise can be reduced. In this study, the median filtering method was mainly used to denoise the volleyball image.

3.9. SVM Parameter Training. Using the multiclass SVM algorithm, use the action data to train the classification algorithm model. Randomly select 650 groups from the 1300 groups of sample data as the training set, and the remaining 650 groups as the test set. The final training process of the

classification algorithm is to input the training action data in the parameters of the volleyball, and the SVM classifier model for action recognition can be obtained. For different parameters, the recognition results and time are different.

3.10. Network Simulation Program Realization. In the process of network simulation, when the density of network nodes is small, the positioning calculation based on cooperation mode is selected. The simulation program is composed of two parts: one is a user interface program written with VisualC++ development tools; the other is a simulation example program for volleyball wireless monitoring sensor network technology written with OMNeT++ discrete event simulation software. The latter is the core part of the simulation program, which is composed of application layer simulation sample program and MAC/route simulation sample program, which are embedded in the user interface program, and the two simulation sample programs can be started in the user interface program.

3.11. Interactive Communication Design between the Client and the Server. The server publishes related programming interfaces through UDDI, and the client queries these interfaces through UDDI and programs these interfaces to realize the secure communication between the client and the server.

3.12. System Test. The accelerometer data used for training needs to be tagged, and the experimenter needs to enter the type of volleyball style used. Four common volleyball styles can be selected: breaststroke, freestyle, backstroke, and volleyball. Add nickname information to prevent data confusion during the experiment. In addition, the experiment recorder needs to manually record the time point when each experimenter swims and turns and the number of strokes. Click the start button before volleyball, and click the end button after going back and forth through the pool. Then, the acceleration data can be exported to the mobile phone SD card in the txt format through the export button. The txt data format contains the system's time stamp and the *xyz* three-axis data components of the acceleration sensor. The naming format is id+stroke style+pool distance +sensor data source category.

TABLE 1: SVM parameter identification results.

Method	GS	GA	PSO
Parameter	(0.743, 0. 188)	(0.673, 0.981)	(0.665, 0.100)
Training sample accuracy	90.33 (542/600)	91.5% (549/600)	92.2% (553/600)
Test sample accuracy	92.0 (552/600)	96% (576/600)	93.8% (563/600)
Support vector number	305	179	289
Training time	26	20 s	24 s

4. Analysis of Volleyball Stroke Management

4.1. Analysis of Parameter Recognition Results. SVM parameter identification results are shown in Table 1. Compare the combined training results of three parameters (C, g). Looking at the test accuracy of the three classifiers as a whole, except for the genetic algorithm, the other two have an accuracy close to 94%. Further analysis of the impact of different parameters on the accuracy rate found that the smaller the penalty factor C, the greater the number of support vectors, the larger the C, the less the number of corresponding support vectors. This is in line with the definition of C. The smaller the C, the smaller the emphasis on noise points. During training, more noise points are treated as data samples. However, when the parameter g is 0.1, the test accuracy is higher, and the genetic algorithm has the highest accuracy in the training samples, which indicates that the kernel function affects the generalization ability of the classifier, and the effect of processing unknown samples is better. In summary, the parameter combination obtained by the grid search algorithm corresponds to the classifier with higher accuracy and better performance. The actual test time of the three methods is within 0.5 seconds. For a well-trained interactive system, it meets the real-time requirements. The training process is performed offline and does not affect real-time recognition. There is also a positive correlation between training time and the number of support vectors, indicating that more support vectors require a lot of training time. It can be seen from Table 1 that even if only the relative position features between the pair of bone nodes are used, the average recognition rate of this study is the highest. This shows that the SVM based on the two-layer AP can select a more representative posture. Taking the selected posture as the hidden variable of SVM, this method can achieve good results for action recognition. After adding bone angle features, the average recognition rate of this study can reach 96%. It can be seen that for the UTKinect database, the combination of special recognition features can enrich the feature information and improve the recognition rate.

4.2. Area Window Size on the Detection Effect. After SVM detects the peak, it needs to detect the window of the area where the peak is located and eliminate the influence of some spike interference on the detection effect. The setting of different area windows will affect the monitoring effect of the algorithm. This research will find the best through experiments. The size of the area window makes the detection accuracy the highest. The experiment extracts the complete data of 4 users from the volleyball data and measures

TABLE 2: Influence of area window size on detection effect.

Numbers	Time window	Check accuracy
1	2	93%
2	4	95%
3	6	97%
4	8	98%

the accuracy of the measurement in different area windows through the above algorithm. The effect of the area window size on the detection effect is shown in Table 2. In the data collection, when each subject is recording the acceleration data during volleyball while wearing the equipment, the experimenter in charge of recording will observe with his eyes and record the number of strokes. By comparing the actual recorded data with the measured data, the accuracy of the algorithm can be obtained. Observe the average accuracy of the experiment by setting different sizes of time windows in the algorithm. The ideal recognition accuracy is the dynamic threshold detection method and the regional peak detection method proposed in this paper. As the acceleration curve is difficult to avoid, there are some spike interference signals that make the curve difficult to be completely smooth. Therefore, some of the peak signals may be misjudged as arm stroke behavior, resulting in experimental detection results generally greater than actual results. The dynamic threshold detection method will miss the detection situation, resulting in the detection result is generally less than the actual result. Based on the results of volleyball style detection, the measured value of the regional peak detection method is the closest to the actual situation, which can show that the method proposed in this study is feasible.

4.3. Different Dimensionality Reduction Methods on the Performance of Action Recognition. In this study, the selected part of the action sequence and the collected action sequence in the action database are used as the research data set, and the angle characteristic sequence of each action sequence is extracted, and each sequence is further extracted by windowing each sequence. The time-frequency characteristics of the data. Part of the collected human movement data is used as training data, the movement data selected from the CMU movement database and the rest of the collected data are used as test data. Three different feature reduction methods are used to reduce the dimensionality of the test data. It can be seen from the figure that in the dimensionality reduction results obtained by linear discriminant analysis, the same types of action data are almost clustered

together, but the action data between different types are too close and the degree of discrimination is too small. In the dimensionality reduction results obtained by principal component analysis, all the motion data are intertwined in a mess, which is difficult to distinguish with the naked eye. In contrast, in the dimensionality reduction results obtained by the generalized discriminant analysis, the same type of action data has good aggregation, and different types of actions have a large degree of distinction, which lays a good foundation for the classification of the next step.

In order to further quantitatively study the impact of different dimensionality reduction methods on the performance of action recognition, this study uses the collected action data as the training set, and the data selected from the CMU action database as the verification set. The hierarchical SVM classifier is used to compare the three different dimensionality reduction methods compared. The verification results obtained by different dimensionality reduction methods are shown in Figure 2. The action recognition rate obtained by generalized discriminant analysis dimensionality reduction is the highest, followed by linear discriminant analysis, and principal component analysis is the lowest. The worst results obtained by using principal component analysis to reduce dimensionality can be explained as follows: principal component analysis is a linear dimensionality reduction method, and it is also an unsupervised dimensionality reduction method. The effect of sample labels is not considered in the process of dimensionality reduction. Important information is not extracted. In contrast, linear discriminant analysis, as a supervised dimensionality reduction method, has better feature extraction capabilities. Due to the introduction of the kernel function, generalized discriminant analysis can map nonlinear motion characteristics to high-dimensional space and then perform linear discriminant analysis, so generalized discriminant analysis has a stronger ability to extract information from human movements.

4.4. Feasibility Analysis of SVM. According to the best effect of this feature, this study set the number of joints to 6 and the segmentation window to 60. The SVM descriptor has the fastest recognition time. This is because the descriptor has only a few simple operations in the process of extracting time-frequency features, and the dimensionality reduction process can be obtained only by multiplying with the projection matrix. In addition, the descriptor applied generalized discriminant analysis to effectively extract the important information in the action features, so a good recognition result was obtained. Through comparison, it can be found that the method proposed in this study is superior to other methods in recognition rate and recognition time. The two descriptors proposed based on the topic model, LDA and LDA+SVM, have the best recognition accuracy, which can be explained as follows: the topic model-based method refines the difference of different types of actions to the difference of the posture frame. Sampling makes action words with smaller discrimination in similar types of actions have greater weight, which makes action descriptors of similar action types have greater differences. The feature recognition

result of SVM for volleyball stroke is shown in Figure 3. Since the types of actions in this study are few and simple, it uses fewer action words and action topics to achieve the optimal recognition effect and has a shorter recognition time. Compared with the two types of action descriptors proposed in this study, both HOD and SMIJ have longer recognition time. Since the HOD descriptor needs to count the sum of the displacement of each joint point in different directions in different dimensions, the calculation amount is large and noise is easily introduced, so the effect of the descriptor is relatively poor. The SMIJ descriptor needs to select the first 6 joints with the largest amount of information from all joints in each time window. The selection process is more complicated and consumes a lot of time. The experimental results show that although the accuracy of the SVM method based on the hybrid kernel function in the action recognition result is only slightly higher than that of the template matching method, the time it takes to recognize the action is much lower than the template matching method. This is because when using the template matching method, to identify the input action category, it is necessary to traverse the entire template library, which greatly increases the running time of the program. When using the SVM method of the mixed kernel function, due to the principal component analysis method, only need to extract the features of the picture, no need to traverse and analyze each image, which greatly shortens the recognition time, so it proves that the SVM method based on the hybrid kernel function is feasible in action recognition, and it is a kind of higher efficiency. Method. The template matching method and the SVM method based on the hybrid kernel function are used to identify the action category of the person in the image. The results show that for nondatabase images, the calculation speed of the two methods is basically unchanged, and the recognition rate is reduced. This is mainly because the types of actions in the database are limited, they are all actions under the same background, and the direction of application is the same.

4.5. Sample Size and Time Interval on Recognition Results. The influence of sample size and time interval on the recognition result is shown in Figure 4. The recognition accuracy rate of volleyball stroke is represented by a blue curve and a red curve with a star. The number of samples increases from 1 to 5, and the recognition accuracy increases faster, while from 5 to 50, it increases more and more slowly, until it reaches the highest value at 50. The sampling interval is in seconds, starting from 0.1 s and ending at 0.5 s, respectively, 0.1, 0.2, 0.3, 0.4, and 0.5 seconds. From the experimental results, the recognition accuracy of the DTW algorithm will indeed increase as the number of template samples increases, and the recognition accuracy of the corresponding DTW+SVM hierarchical recognition method will also increase as the number of template samples increases. However, when the number of samples reaches 50, the recognition accuracy reaches the maximum, which means that only 50 high-quality samples need to be found from 200 samples as template samples. The internal voting method is adopted, and the number of votes is selected from more
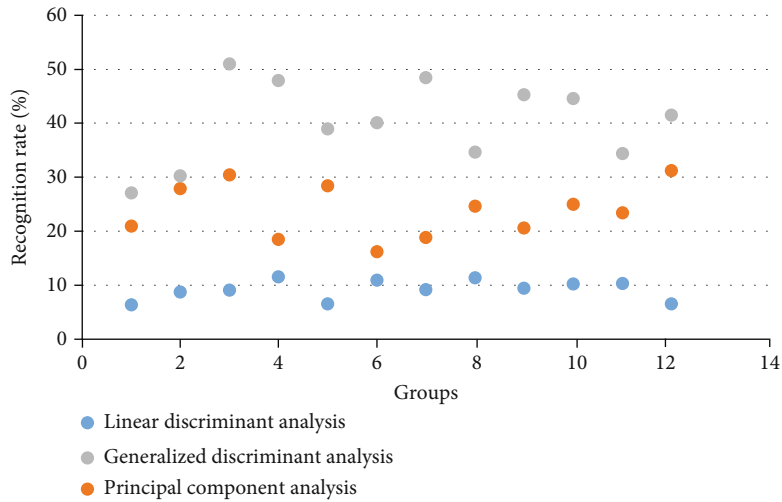
FIGURE 2: Recognition results of verification data obtained by different dimensionality reduction methods.
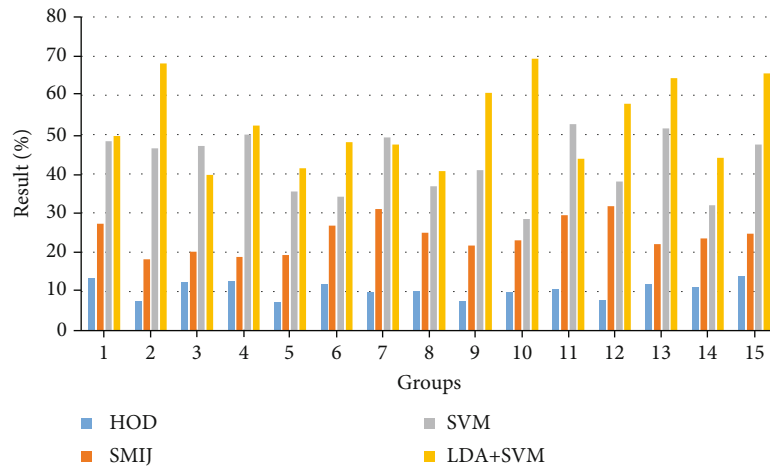

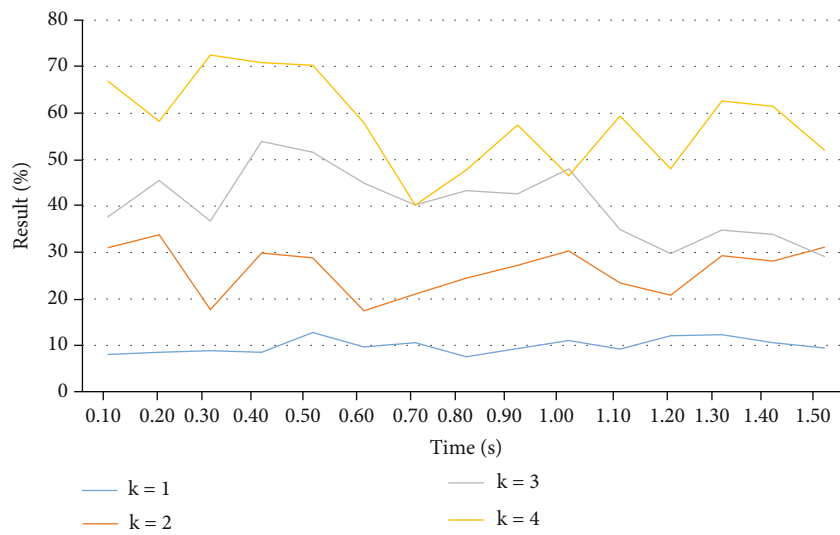FIGURE 3: SVM feature recognition results for volleyball stroke.

FIGURE 4: Influence of sample size and time interval on identification results.

to less. Arrange, take the first 50 samples. It can be seen from Figure 4 that the recognition time is closely related to the number of samples, which means that the more samples, the longer the recognition time. The reason is that the number of samples increases, and the number of template matches for the DTW algorithm to recognize simple arm movements will also be corresponding; the identification time will also increase. From Figure 4, it can be seen that the sample size range is between 10 and 50. It is more appropriate to combine the sample size between 10 and 30. If you need to improve the recognition accuracy, increase the number of samples, and if you need to shorten, recognition consumes time and reduces the number of samples. It can be seen from Figure 4 that both the red curve and the blue curve will decrease as the sampling interval increases, that is, the recognition accuracy of simple arm movements and the recognition accuracy of arm movements will decrease as the sampling interval increases; the decreasing range is getting bigger and bigger. When the sampling interval is 0.5 seconds, the recognition accuracy of arm movements is only about 60%. The time spent on recognition decreases as the sampling interval increases. It can be seen that it is more appropriate to set the sampling interval to less than 0.8 seconds. If you need to improve the recognition accuracy, reduce the sampling interval, and if you need to reduce the identification consumption time, increase the sampling interval. Considering the number of candidate segments, the number of candidate segments k performs length statistics on all simple arm motion templates and sorts them from high to low according to the number of occurrences. The first k lengths are selected as the length of the candidate segments, so the larger the value of k, the more the number of candidate segments, the better the segmentation effect, the higher the recognition accuracy, and the smaller the k value, the less the number of candidate segments, the worse the segmentation effect, and the lower the recognition accuracy. And when k is from 1 to 2, the degree of improvement in recognition accuracy is small, and the degree of improvement from 3 to 4 is very small, so the number of suitable candidate segments k is selected between 3 and 4; if you need to improve the recognition accuracy, it can be appropriate. Modifying the value of the number of candidate segments k, increasing k can slightly improve the recognition accuracy, but it will also increase the recognition time. The larger the value, the more the number of candidate segments, the better the segmentation effect, the higher the recognition accuracy, and the longer the recognition time. Conversely, the smaller the k value, the worse the segmentation effect, the lower the recognition accuracy, and the more expensive the recognition, the shorter the time.

4.6. Analysis of Median Filter Denoising Results. First, divide the pixels into three groups, namely, P1, P2, and P3. Then, sort the three groups separately to find the maximum, middle, and minimum of the three groups. Finally, take the smallest value among the three sets of maximum values, the middle value among the three sets of intermediate values, and the largest value among the three sets of minimum values, and then compare the intermediate values. This
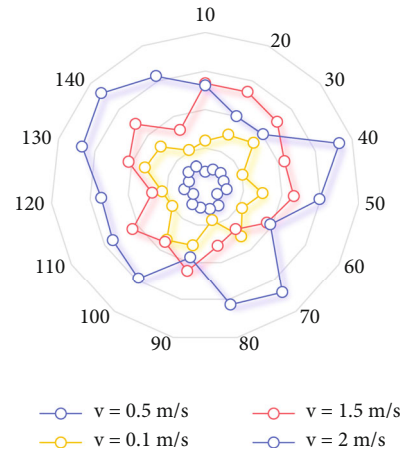

FIGURE 5: Median filter denoising results.

intermediate value is the required pixel value. This is the content of the quick sort module 2, a total of 3 rows, 3 columns, and 6 sorts; each sorting only needs to delay 3 clock cycles to get the result. Human body movements are characterized by high complexity and style diversity, and it is difficult to avoid the introduction of noise when using wearable sensors to collect movement data. Therefore, the time-frequency features extracted from movement data have a nonlinear relationship with different types of movements. In the actual design, only one value sorting module is needed, and the design can be completed by including them in the top file, respectively. The median filter denoising result is shown in Figure 5. It can be seen from Figure 5 that using the pipeline method for design, after the first group of 9 pixels are quickly arranged, after a delay of three clock cycles, the final median value is the value of data 2. The median value obtained in the second and third groups is also delayed by three clock cycles before getting the result, which shows that the design meets the requirements. It can be seen from Figure 5 that the gray image has more noise, and after median filtering, most of the noise can be filtered out. When the signal strength is 90 T, at 2.0 m/s and 0.5 m/s speed, the detection accuracy of the adaptive threshold reaches 93% and 95%, respectively. Since the adaptive threshold dynamically adjusts the detection threshold according to the signal conditions, it maintains a high detection rate under the low signal condition of 9 T, and the accuracy of 3.0 m/s and 0.5 m/s motion detection reaches 84% and 92%.

5. Conclusion

In this study, firstly, the analysis and management method of volleyball action using a mobile device's action sensor is proposed, and a two-layer AP gesture selection SVM action recognition algorithm is proposed, and the effectiveness of the algorithm is simulated and analyzed. The second is to make use of resources. Based on Visual Studio programming technology, a motion recognition software system platform based on MFC is designed and developed, which can recognize the trained motion captured by Kinect in real time. In the aspect of feature selection of action recognition, a SVM action

recognition method based on two-layer AP gesture selection is proposed. This method can quickly, accurately, and stably select multiple gestures that can represent this kind of action and accurately recognize them through support vector machine. It avoids the shortcomings of the traditional clustering initialization algorithm which is complex or unstable.

According to different application scenarios and information acquisition methods, human action recognition is roughly divided into two categories: image-based and inertial sensor-based. With the development of electronic technology, information shows that the national surveillance cameras have reached 30 million, recording a lot of video information every day. In this study, under the basic framework of traditional video surveillance, the access detection module, motion tracking module, motion recognition module, intelligent SMS notification module, cheek control module, and analysis result display module are added, which basically realizes the intelligent video surveillance function in a simple environment. Because the development of an ordinary video surveillance system has been quite mature, in each module, action recognition is the most difficult, and the algorithm is also the most complex. This paper will briefly introduce the related research of video surveillance, focusing on the related research of action recognition module. Adapt to more changing application scenarios. By embedding the motion sensor into the small wearable device, it is not affected by the external environment, such as light, background, range of motion, and obstacles. There is no need for complex additional equipment, such as cameras; the user's personal real environment does not affect the recognition effect, so it has more application scenarios. More direct access to motion data. Inertial sensors can directly obtain the acceleration and angular velocity information in the process of motion. Through filtering the measured data, high-precision measurement can be achieved, which greatly improves the accuracy of feature extraction in the process of recognition quality.

Through the effective processing of acceleration information, we can extract the corresponding action information and infer the intention of the action executor. On the basis of consulting a large number of fields of human motion capture and behavior recognition, the behavior referred to in pattern recognition is what we call action. Most human activities can be divided into a single combination of actions, so the recognition of behavior can be summed up as the recognition of a series of actions. Data-based action recognition is a new research field in the field of pattern recognition. The process can be simplified as the following steps: firstly, the motion sensor obtains the acceleration and angular velocity data generated during the motion and transmits the measured data to the mobile device through wireless transmission. Secondly, the mobile device processes the data and extracts the characteristic value of the action. Finally, the action is classified and recognized according to the characteristic value through the pattern recognition algorithm.

Data Availability

No data were used to support this study.

Conflicts of Interest

The author states that this article has no conflict of interest.

References

[1] S. Yu, H. Chen, and R. A. Brown, "Hidden Markov model-based fall detection with motion sensor orientation calibration: a case for real-life home monitoring," *IEEE Journal of Biomedical and Health Informatics*, vol. 22, no. 6, pp. 1847–1853, 2018.

[2] C. Shen, Y. Li, Y. Chen, X. Guan, and R. A. Maxion, "Performance analysis of multi-motion sensor behavior for active smartphone authentication," *IEEE Transactions on Information Forensics & Security*, vol. 13, no. 1, pp. 48–62, 2018.

[3] A. Yurtman and B. Barshan, "Novel noniterative orientation estimation for wearable motion sensor units acquiring accelerometer, gyroscope, and magnetometer measurements," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 6, pp. 3206–3215, 2020.

[4] S. K. Behera, D. P. Dogra, and P. P. Roy, "Analysis of 3D signatures recorded using leap motion sensor," *Multimedia Tools and Applications*, vol. 77, no. 11, pp. 14029–14054, 2018.

[5] M. Hirata, R. Watanabe, Y. Koyano et al., "Using a motion sensor-equipped smartphone to facilitate CT-guided puncture," *Cardiovascular & Interventional Radiology*, vol. 40, no. 4, pp. 609–615, 2017.

[6] J. K. Min, E. J. Cheon, J. M. Kim, S. C. Lee, and K. Choi, "Comparison of the 6-DOF motion sensor and stain gauge data for ice load estimation on IBRV ARAON," *Journal of the Society of Naval Architects of Korea*, vol. 53, no. 6, pp. 529–535, 2016.

[7] I. I. Duma and S. Giurgiu, "Circadian activity and nest use of Dryomys nitedula as revealed by infrared motion sensor cameras," *Folia Zoologica*, vol. 61, no. 1, pp. 49–53, 2018.

[8] L. B. Huang, W. Xu, C. Zhao et al., "Multifunctional water drop energy harvesting and human motion sensor based on flexible dual-mode nanogenerator incorporated with polymer nanotubes," *ACS Applied Materials And Interfaces*, vol. 12, no. 21, pp. 24030–24038, 2020.

[9] T. A. Donaldson, "Inline calibration of motion sensor," *Journal of Chinese Pharmaceutical ences*, vol. 48, no. 23, pp. 2022–2025, 2016.

[10] G. Marin, F. Dominio, and P. Zanuttigh, "Hand gesture recognition with jointly calibrated leap motion and depth sensor," *Multimedia Tools and Applications*, vol. 75, no. 22, pp. 14991–15015, 2016.

[11] H. Jin, Q. Chen, Z. Chen, Y. Hu, and J. Zhang, "Multi-Leap-Motion sensor based demonstration for robotic refine tabletop object manipulation task," *Caai Transactions on Intelligence Technology*, vol. 1, no. 1, pp. 104–113, 2016.

[12] Z. Wei, W. Zhang, and W. Liu, "Attitude theory and experimental research of micro-aircraft based on MEMS sensor," *Piezoelectrics & Acoustooptics*, vol. 40, no. 4, pp. 516–520, 2018.

[13] D. Laurijssen, S. Truijen, W. Saeys, W. Daems, and J. Steckel, "An ultrasonic six degrees-of-freedom pose estimation sensor," *IEEE Sensors Journal*, vol. 17, no. 1, pp. 151–159, 2017.

[14] D. Korpi, J. Tamminen, M. Turunen et al., "Full-duplex mobile device - pushing the limits," *IEEE Communications Magazine*, vol. 54, no. 9, pp. 80–87, 2016.

[15] L. Zhou, "Mobile device-to-device video distribution: theory and application," *ACM Transactions on Multimedia Computing Communications and Applications*, vol. 12, no. 3, pp. 1–23, 2016.

[16] E. Noei, M. D. Syer, Y. Zou, A. E. Hassan, and I. Keivanloo, "A study of the relation of mobile device attributes with the user-perceived quality of Android apps," *Empirical Software Engineering*, vol. 22, no. 6, pp. 3088–3116, 2017.

[17] K. F. Chiang and H. H. Wang, "Nurses' experiences of using a smart mobile device application to assist home care for patients with chronic disease: a qualitative study," *Journal of Clinical Nursing*, vol. 25, no. 13-14, pp. 2008–2017, 2016.

[18] X. Du and K. Yang, "A map-assisted WiFi AP placement algorithm enabling mobile device's indoor positioning," *IEEE Systems Journal*, vol. 11, no. 3, pp. 1467–1475, 2017.

[19] R. Gotor, P. Ashokkumar, M. Hecht, K. Keil, and K. Rurack, "Optical pH sensor covering the range from pH 0-14 compatible with mobile-device readout and based on a set of rationally designed indicator dyes," *Analytical Chemistry*, vol. 89, no. 16, pp. 8437–8444, 2017.

[20] M. I. Gofman and S. Mitra, "Multimodal biometrics for enhanced mobile device security," *Communications of the ACM*, vol. 59, no. 4, pp. 58–65, 2016.

[21] C. Hauptmann, A. Wegener, H. Poppe, M. Williams, G. Popelka, and P. A. Tass, "Validation of a mobile device for acoustic coordinated reset neuromodulation tinnitus therapy," *Journal of the American Academy of Audiology*, vol. 27, no. 9, pp. 720–731, 2016.

[22] R. D. Keegan, M. C. Oliver, T. J. Stanfill et al., "Use of a mobile device simulation as a preclass active learning exercise," *Journal of Nursing Education*, vol. 55, no. 1, pp. 56–59, 2016.

[23] W. Sun, F. Brännström, and E. G. Ström, "Network synchronization for mobile device-to-device systems," *IEEE Transactions on Communications*, vol. 65, no. 3, pp. 1193–1206, 2017.

[24] H. X. Zhang, F. F. Sheng, P. Y. Xu, and Y. Tang, "Visualizing user characteristics based on mobile device log data," *Journal of Software*, vol. 27, no. 5, pp. 1174–1187, 2016.

[25] S. J. Lofgreen, K. Ashack, K. A. Burton, and R. P. Dellavalle, "Mobile device use in dermatologic patient care," *Current Dermatology Reports*, vol. 5, no. 2, pp. 77–82, 2016.

[26] H. Chen and W. Li, "Mobile device user's privacy security assurance behavior: a technology threat avoidance perspective," *Information & Computer Security*, vol. 25, no. 3, pp. 330–344, 2017.

[27] M. Zhou, Y. Long, W. Zhang et al., "Adaptive genetic algorithm-aided neural network with channel state information tensor decomposition for indoor localization," *IEEE Transactions on Evolutionary Computation*, p. 1, 2021.

[28] Y. L. Zhou, Y. Li, M. J. Tahir, X. Geng, Y. Wang, and W. He, "Integrated statistical test of signal distributions and access point contributions for Wi-Fi indoor localization," *IEEE Transactions on Vehicular Technology.*, vol. 70, no. 5, pp. 5057–5070, 2021.