Research Article

Optimization of a Wireless Sensor-Based Tennis Motion Pattern Recognition System

Haoxuan Ruan [iD][1] **and Xinchen Zhang** [iD][2]

[1]*Institute of Sports and Health, Tianjin University of Sport, Tianjin 301617, China*
[2]*Department of Sociology, Yeungnam University, Gyeongsan, Gyeongsangbuk 38541, Republic of Korea*

Correspondence should be addressed to Haoxuan Ruan; ruanhaoxuan@tjus.edu.cn

With the rapid development of information technology in today's era, the application of the Internet, big data, and smart bracelet information technology in the field of sports has enhanced the intelligence of sports and plays an important role in promoting sports performance. This paper focuses on the application of wireless sensors in the field of tennis, using research methods such as literature research, video analysis, comparative research, and mathematical statistics, to explore and analyze the application of wireless sensors in the field of tennis big data, tennis robotics, and the implementation of tennis teaching and training, to provide a theoretical basis for promoting the application of wireless sensors in the field of tennis and also for the broader application of wireless sensors in sports to provide a theoretical reference. For the problem of multiple scales of motion targets in action videos, two video action recognition methods based on high- and low-level feature fusion are proposed, which are the video action recognition methods based on top-down feature fusion and the video action recognition methods based on bottom-up feature fusion. The multipowered mobile anchor nodes are allowed to move along a prescribed route and broadcast multiple power signals, and then, the location of the unknown node is estimated using a four-ball intersection weight center-of-mass algorithm. Simulations show experimentally that the algorithm reduces the average localization error and requires fewer anchor nodes.

1. Introduction

The wireless sensor network is one of the hot spots of rapid development in recent years; it combines the sensor field, wireless communication field, computer field, and a large number of other different fields of advanced technology and constantly developed into a new field of integrated technology. Many tiny low-power nodes constitute the wireless sensor network; tiny nodes can monitor complex external information in real time and transmit the monitoring results to the embedded system and after the system processing, by sending to the user terminal, so that these nodes can intelligently sense the outside world. However, these nodes can locate themselves in addition to sensing information such as temperature, humidity, and light intensity. Using this property, wireless sensor network technology quickly entered the wireless communication industry, giving rise to

many new technologies and applications that have attracted widespread attention worldwide [1]. The heavy use of sensors requires lower cost, better scalability, and more power savings than traditional technologies. Motion analysis allows one to learn the motion patterns of target objects and use them for analytical modeling. For example, in the field of medical rehabilitation, remote monitoring networks can be established for patients to enhance the monitoring of their behavior and thus provide timely feedback on medical data, while in the field of ergonomics it can also provide sufficiently accurate human posture data for research; in the field of sports, motion analytics can be used to simulate training, record athletes' movement data, and compare it with quasi-templates to generate corrective information for reference; in the entertainment industry, motion analysis technology is used in 3D graphics production to restore the movement of the target object, which can lead to lifelike character

modeling. In addition, distributed sensor architectures for motion capture can be installed on different mechanical devices, thus offering the possibility of achieving intelligent interaction [2].

Wireless sensor networks, as one of the important technologies for the new Internet of Things (IoT), have become the communication hub of society with their efficient, fast, and comprehensive features. The popularity of IoT has led to the rapid development of the wireless communication industry and the ubiquity of sensor networks. Compared to traditional technologies, the massive use of sensors demands low cost, good scalability, and more energy-efficient power consumption. WSN is usually a unified joint system consisting of communication, microelectronics, semiconductor, and embedded computer technologies. The ability of WSNs to reconfigure intelligently and dynamically allows them to collect and process the information sent by the nodes in large quantities and transmit it to the control center, which is the user terminal [3]. In this paper, around the theme of wireless sensors in the field of tennis, we use literature research, video analysis, comparative research, and mathematical statistics to explore and analyze the implementation of wireless sensors in tennis big data, tennis robotics, and tennis teaching and training approaches, to provide a theoretical basis for promoting the application of wireless sensors in the field of tennis and also to provide broader applications to provide theoretical references [4].

2. Related Work

The development of video action recognition methods relies on the progress of fundamental research on video representation learning. Video representation can be divided into two aspects, manual feature representation and deep feature representation. The dense trajectory method (DT) was proposed in the literature [5] and applied to the video action recognition task. The basic idea of the dense trajectory method is to first use the optical flow field to obtain the trajectory in the video sequence, then extract motion descriptors HOF, HOG, MBH, and trajectory features along the trajectory, then encode the features using the Fisher Vector method, and finally train the SVM classifier based on the encoding results to give recognition results. An improved version of the dense trajectory method (IDT) is proposed in the literature [6]. IDT uses the SURF matching algorithm to match the key points of the optical flow between two frames before and after the video to attenuate the effect of camera motion on the video content and becomes the most effective method among traditional video motion recognition methods. Manual features mainly characterize low-level visual information, underrepresent high-level semantic information, and have the disadvantage of difficulty in handling large amounts of data and unsatisfactory recognition accuracy. To solve this problem, the literature [7] proposes the concept of intermediate-level features, which represent behavioral features through a set of action attributes learned from the training dataset, which is referred to as an intermediate concept in the paper. The literature [8] uses motion phrases and motion atoms to represent the features of

actions in videos. For high-level feature representation, the literature [9] uses an ordering function to model the evolution of motion over time. To better capture spatiotemporal information, literature [10] uses hidden Markov models to capture temporal information in videos and uses fixed dimensional vectors as descriptors of motion videos. The literature [11] uses a structural trajectory learning approach to extract relevant motion features.

The four methods based on ranging localization are angular arrival, timely arrival, time difference arrival, and received signal strength indication; AOA uses the angular relationship between two anchor nodes concerning the unknown node for localization, TOA and TDOA use the product of signal propagation time and propagation speed to calculate the distance, and trilateral localization or great likelihood estimation becomes the method to estimate the coordinates in the latter step. RSSI uses the received signal strength to measure the distance and then the base positioning method to achieve positioning. The main ones that are not based on ranging are the DV-HOP localization algorithm, APIT, center-of-mass localization, MDS-MAP, and amorphous localization: amorphous uses network connectivity as a basis for calculation. In indoor localization by WSN, the literature [12] can detect a single intruder through Wi-Fi devices with a high detection rate and small false-positive results; mobile anchor nodes can plan the path to achieve high coverage and are more flexible than static anchor nodes and do not depend on the topology of the network. The literature [13] proposes adaptive framework structures thus detecting variable speed objects in indoor environments. The authors conducted a series of experiments to learn empirically the effect of different speeds on localization accuracy and thus improve the accuracy of localization at different speeds. A novel indoor passive localization system in a wireless environment is proposed in the literature [14]. It provides low overhead and accurate and robust motion detection and gives tracking capability, using coordinates of different unknown nodes with the same anchor node to construct a new coordinate system to calculate the distance and then using trilateral localization for localization of nodes, which cleverly simplifies a large number of calculations using the coordinate method. In the literature [15], large-scale indoor passive localization and tracking are proposed. Although it has relatively high localization accuracy under multipath effect, the literature [2] better describes the localization classification model for passive localization, improves the quality of the dataset, and reduces the error caused by the multipath effect; mostly, the distance between the anchor node and unknown node is estimated by network connectivity, information passed between nodes, etc.; the accuracy is not very high, but it does not need to carry extra equipment so the cost is low and the power consumption is relatively low. The literature [16] proposes three passive indoor localization methods and discusses the effect of multiple targets on the results. Once the packet enters that grid, it is forwarded to the grid head node which also becomes the phantom source. If no node exists in the grid where the random location is located, the head node of the grid where the node that last cached the packet is located will become the phantom source.

3. Optimization of a Wireless Sensor-Based Tennis Motion Pattern Recognition System

3.1. Node Localization Algorithm for Wireless Sensor Networks. Wireless sensor network node localization algorithms can usually be divided into two categories: range-based localization algorithms and range-free localization algorithms. Range-based algorithms use geometric relationships to derive node unknowns by measuring the wireless signal angle or propagation time between the unknown node and the signal transmitting node. The measured information includes received signal strength, signal arrival time, signal arrival time difference, and signal arrival angle. These algorithms usually require the deployment of special components to obtain these variables and improve the localization accuracy by taking multiple measurements, resulting in incurring higher deployment costs. In contrast, range-free localization algorithms require only information about the anchor node and network connectivity and thus are cheaper to deploy and require no additional hardware support but have limited localization accuracy. The fingerprint localization algorithm belongs to the range-free localization algorithm, which requires several anchor nodes and reference nodes with fixed locations to be predeployed in the localization area. The anchor nodes continuously transmit wireless signals with rated power, and the signal RSS (Really Simple Syndication) of each anchor node is measured at each reference node location. The individual reference node locations and their measured RSS form a location fingerprint or fingerprint for short. The unknown node also measures the RSS of each anchor node and pattern matches it with the existing fingerprint to determine the node location. Fingerprint location algorithms not only are cheap to deploy but also have more accurate localization performance in complex and variable propagation environments, such as multipath and NLOS environments, and thus have been widely studied and applied in recent years.

Such algorithms use network-wide connectivity information to make location decisions. One of the best-known algorithms is DV-hop. This algorithm has distance vector routing at its core, where each anchor node broadcasts a beacon message containing its location coordinates. The initial value of the number of hops in the beacon is 1, and 1 is added for each node passed. When beacons from multiple anchor nodes are transmitted in the network, each node on the transmission path records the minimum number of hops for each anchor node. Due to the diversity of action modes covered in the set, the energy base of each action varies, and even the magnitude difference between different performers under the same type of action is huge, so it is unrealistic to use a constant value as a threshold to complete the interception of all actions. Therefore, it is necessary to propose a threshold determination scheme with self-adaptive capability. In an isotropic sensing network, the single-hop physical distance of the signal is approximately the same in all directions. Unknown nodes estimate the distance to each anchor node based on the number of hops. However, in complex networks, the presence of interference and other factors lead to large differences in the single-hop distances

in each direction, making it difficult to achieve precise positioning, as in Figure 1 bit wireless sensor network node localization process.

Fingerprint localization is a localization algorithm that has gained more attention among the range-free localization algorithms. A certain number of anchor nodes are deployed in the localization area with a fixed location and known coordinates with the signal transmitting function. The sensor nodes measure the wireless signal strength RSS of each anchor node. The measured HSS value and the position coordinates of that node are called the signal fingerprint of that position. The fingerprint localization approach does not derive the node location based on RSS and distance equations but rather fuses RSS with the anchor node approximation algorithm to derive the sensor node location. The fingerprint localization algorithm requires a fingerprint database in the localization space, i.e., the location coordinates of each point in the space are linked to the RSS information of different anchor nodes at that location. The fingerprint localization process is to convert the RSS information received by the unknown node into location information based on the fingerprint and location relationship information in the fingerprint database. The process of converting RSS into a target location is known as fingerprint matching and fingerprint localization. Fingerprint localization can also be described as a multiple hypothesis testing problem, where the best hypothesis (location of the target) is deduced based on the preobtained observations (i.e., fingerprints). The fingerprint localization process can also be considered a decision process, where the decision target is the unknown node location based on the information available (fingerprint database) and the RSS measured by the unknown node. The fingerprint localization algorithm requires two phases: an offline measurement phase and an online localization phase:

$$P(\theta \mid k) = \frac{\int \alpha \cdot (\theta - \mu)/\sigma d\theta}{kr}. \tag{1}$$

Figure 2 shows the basic process of fingerprint localization. In the offline measurement phase, firstly, a certain number of reference nodes are laid out in the current localization environment and the location coordinates of all reference points are recorded. Usually, the reference nodes are laid out in a grid-like manner, and the reference nodes can be either physical or virtual nodes. Then, the RSS values of each anchor node are measured and collected in some way at all reference nodes, called raw observation data, or samples. Due to the inevitable signal interference in the localization area, the RSS measurements are subject to errors and certain methods are needed to preprocess the samples. The preprocessed RSS data and the coordinates of the reference node establish a correspondence to form a fingerprint database. In the online localization phase, the target node measures the RSS value of each anchor node at its location and sends it to the backend localization service. The localization algorithm matches this RSS value with all samples of the fingerprint database according to the set algorithm and finds one or more reference nodes with the highest matching
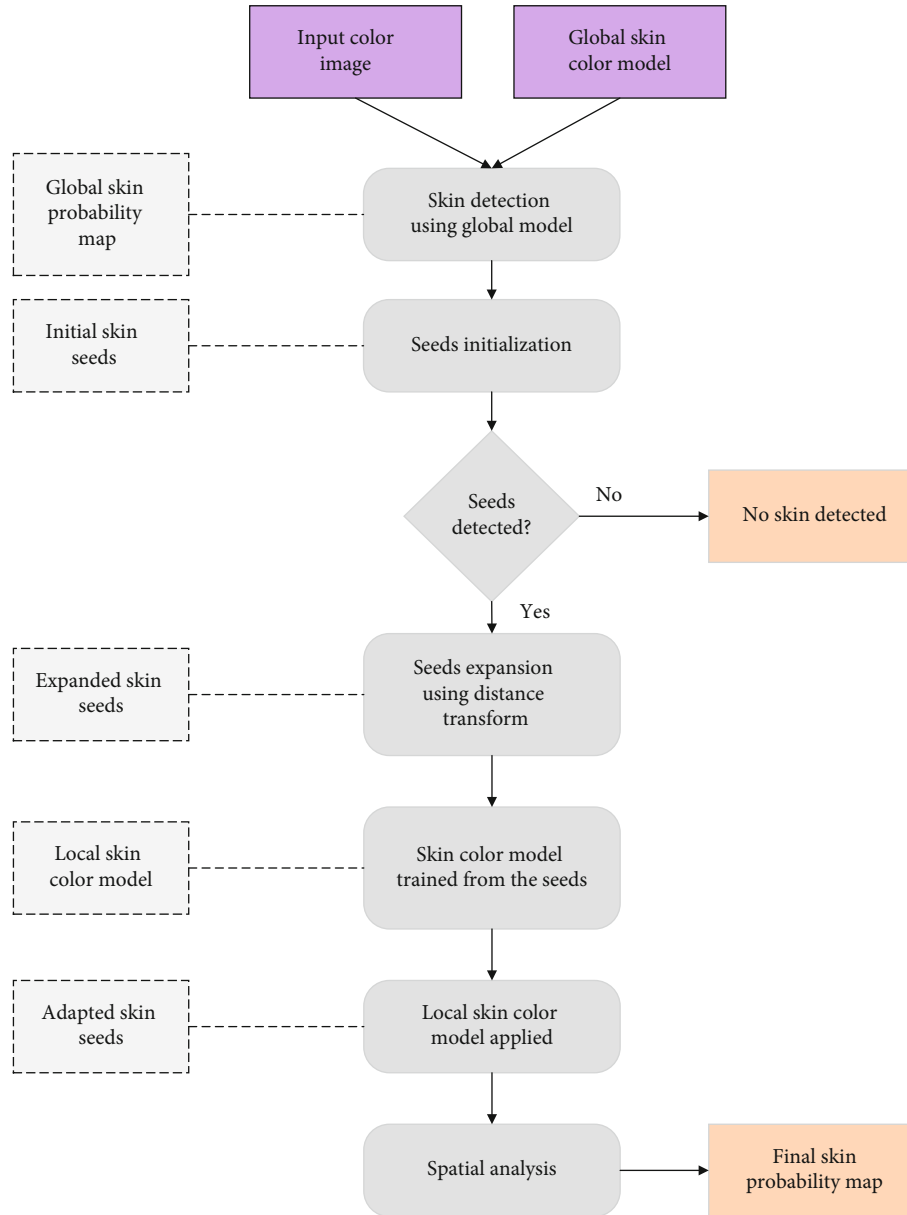
FIGURE 1: Flowchart of wireless sensor network node localization.

degree. Finally, these reference point location coordinates are converted to the location corresponding to the target node according to the characteristic algorithm, i.e., the location estimate of the target node.

In a fixed localization environment, RSS samples usually obey some probability distribution. This is usually described using a joint probability distribution and assuming that the RSS of each anchor node measured by the reference node is independent of each other and does not interact with each other, using the product of the edge distributions of the RSS as the joint distribution. A common data form is the basis for sharing research results. This paper gives a common inertial device standard, motion recording scheme, and data storage form and establishes a simple error calibration scheme for MEMS devices in motion capture application scenarios and a data cleaning

method for the low automation of the data acquisition process. The RSS vector measured by the unknown node is set, the probability of getting this vector at each reference node is obtained, and the reference node with the highest probability is selected as the estimated location. Probabilistic algorithms are mainly based on Bayesian theory, or Bayesian combined with clustering algorithms, to calculate the location estimate of the unknown node on the posterior probability of the unknown node. Plain Bayes, hidden Bayes, Bayesian networks, and maximum likelihood estimation are also widely used methods. The process of node localization based on RSS fingerprinting is usually divided into two phases: an offline measurement phase and an online localization phase. In offline measurement, the RSS data of the anchor node is measured at multiple reference nodes to build a fingerprint database.
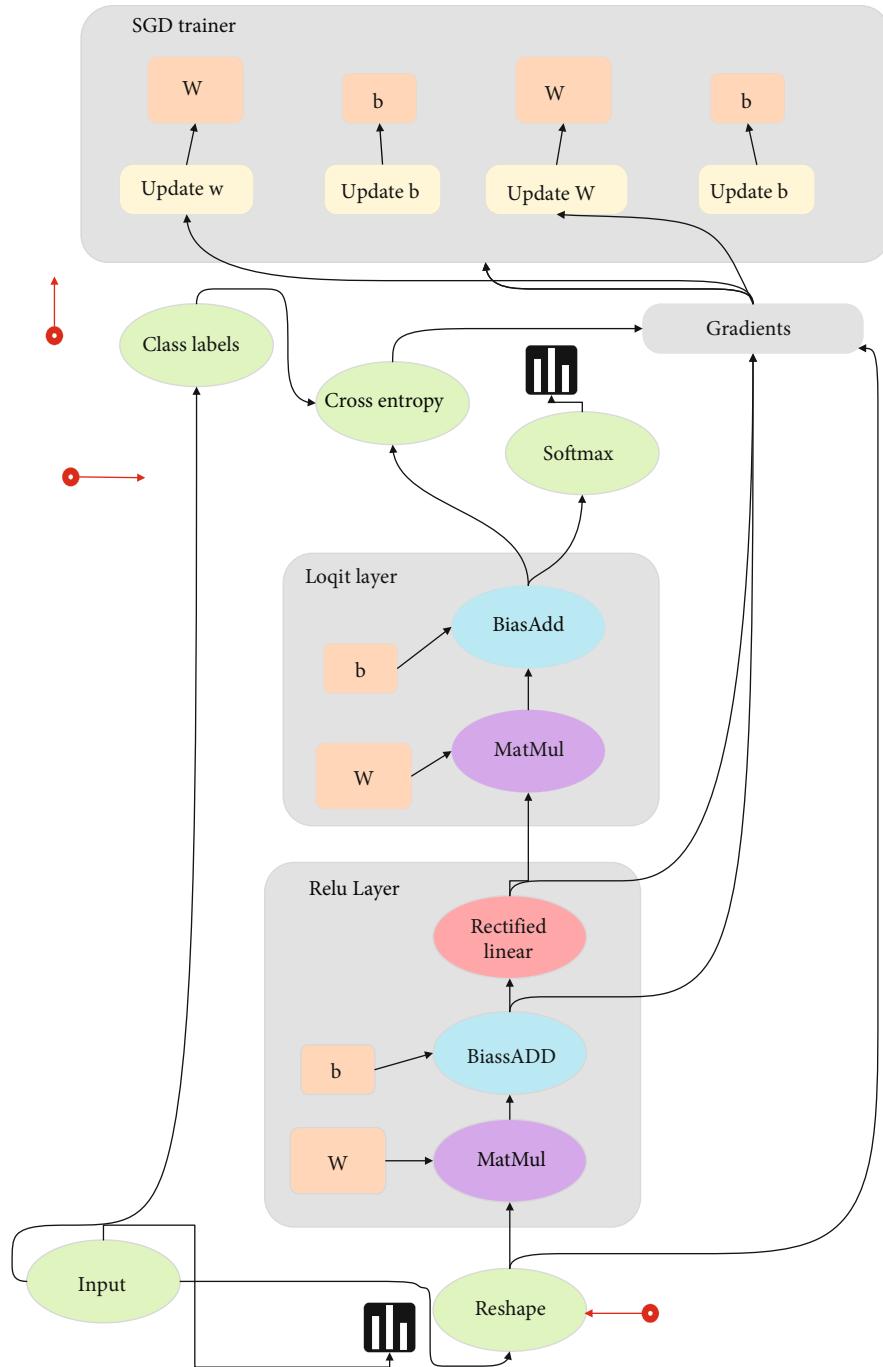
FIGURE 2: Fingerprint location process.

Since environmental noise and obstacles interfere with the wireless signal propagation, it is also necessary to remove the noise in the fingerprint database using statistics, filtering, and fitting; in the online measurement phase, the location of the unknown node is estimated by matching the RSS data collected from the unknown node with the fingerprint database. Therefore, the research of fingerprint localization algorithms mainly includes two aspects: enhancing fingerprint data accuracy and improving localization accuracy.

3.2. Wireless Sensor-Based Algorithm for Tennis Motion Pattern Recognition. The fundamental research in the field of tennis motion analysis can be divided into two directions: namely, motion analysis based on the pose layer and analysis based on the action primitive layer, the essential difference being whether the extraction of data meaning is more focused on positional or velocity information. We can know that tennis sports actions from two perspectives. One way of thinking is to consider it as a continuous-time sequence, i.e., the body joints complete a spatial displacement, then the

velocity information of the point movement can be a complete response to the movement. The other idea is to consider the serving action as a segment of motion with wrist force and posture change, then we can achieve the recognition of a segment of motion by keeping continuous detection of body posture. The two ideas focus on different motion information; the first idea is more concerned about the absolute motion of space differential information; if you use the video capture scheme, you need to extract the spatiotemporal motion trajectory of the moving target and then only through the position information interest inverse calculation of the speed information, resulting in the calculation accuracy being seriously limited by the number of frames shot and a large amount of calculation. The inertial motion sensor can be worn to directly capture the velocity information of the moving object, and the video capture does not have the advantage in this scheme. The second idea is more concerned about the location of the target point information; using the video program is roughly the data processing process: first from a single frame image to extract the relative position of the target feature points and then compared with the standard template to determine the former human pose, and for the inertial sensor program, the need to use inertial navigation integration algorithm from the device output to measure the location of the target point and posture information, so the integration of inertial data. The accuracy of the operation determines the feasibility of the scheme, which is also the core focus of almost all inertial guidance research.

The wireless sensor network is a combination of four components which are sensor nodes, aggregation nodes, mobile communication network, and task management desk. The sensor nodes are mainly placed in the monitoring area and are responsible for the collection of the required information, such as temperature and humidity. There are a small number of anchor nodes carrying self-locating hardware and a large number of unknown nodes whose locations are not known in advance. The main role of the aggregation nodes is to gather the information propagated from the nodes in the monitoring area and then deliver it to the higher level, similar to the role of a gateway. The mobile communication network is mainly responsible for carrying the transmission of information. Usually, the reference nodes are laid out in a grid-like pattern, and the reference nodes can be physical nodes or virtual nodes. Then, the RSS value of each anchor node is measured and collected in some way at all reference nodes, which is called raw observation data or called sample. The task management desk is mainly responsible for processing the collected information for use in higher-level applications.

From a mathematical point of view, an important issue that must be considered in algorithm selection is the trade-off between bias and variance. Classification models with high bias have a high error rate in prediction, while models with high variance will perform erratically across different datasets. Bias and variance are defined in statistics as follows: bias describes the difference between the predicted value and the true value as shown in

$$I^\alpha f(\eta) = m(z)(i\omega)^\alpha f(\omega) = |\omega|^\alpha e^{i2\pi\alpha} f(\omega). \tag{2}$$

Variance describes the instability of the model predictions themselves as shown in

$$Rf(x) = \int\int g(t)dt = \left(\frac{1+\gamma}{n}\right) \cdot \sum (x-1)f(t). \tag{3}$$

Ideally, with an infinitely large sample size of training data and a model algorithm that tends to be perfect, we could obtain models with small bias and variance, but in real engineering problems, this ideal situation does not often exist. Learning algorithms with low bias values tend to be more "flexible" and respond to the higher complexity of the model, thus being able to fit the data very accurately. The feature space is divided into two, with positive and negative classes on each side of the plane, and the specific classification decision function is as follows:

$$R^m f(x) = \int\int g(t)dt = \left(\frac{m+\gamma}{n}\right) \cdot \sum (x^m - t^{m-1})f(t). \tag{4}$$

For linearly differentiable problems, the sample points in T that are closest to (w, b) are called support vectors, and they are mathematically characterized in such a way that equation (5) holds

$$L_{R-D}^n f(x) = \frac{1}{\eta(\alpha)}\int\int (x-\eta)^\alpha g(\eta)d\eta. \tag{5}$$

Since the coverage of node M_i contains the intersection A, we have

$$Mi = \begin{vmatrix} A \cdot \sin u \cos x & I_n \cdot \cos \gamma \\ I_m \cdot \sin \gamma & -\sin y \end{vmatrix}. \tag{6}$$

In the process of covering the void repair, the void is not split if the NNICI generated by all nodes in the set of the mobile nodes and the void boundary nodes that make up the covering void is not more than 2, based on the guarantee that the void inferior arc of the driving node is completely covered, with

$$\overline{v_i} = \frac{\delta x}{\delta t}\left(\frac{n!}{r!(n-r)!}x^\gamma + \mu\right). \tag{7}$$

The input signal is computed in the network in a forward direction: the very front of the network is the input, where each input sample corresponds to a definite known ideal output, while at the output at the very end of the network, the error information is formed between the predicted value and the ideal value, while the gradient information of the error signal can be passed backward from back to front according to the chain rule. At the end of a round of iterations, the new prediction results in an error value whose gradient information is fed back to the layers of the network

through the reverse conduction law, and the parameter values of each neuron will be corrected based on the error gradient according to the established update strategy. This cycle is repeated until the network reaches the accuracy index.

Since the professional basic action division in the tennis field has less ambiguity arising, it can be presumed that the similarity between feature vectors of similar actions is high and the clustering effect of action features is more obvious, i.e., the linear differentiability of the dataset is high. Definitely, the training data sample size is infinitely large, the model algorithm tends to be perfect, and we can obtain models with small bias and variance; however, in real engineering problems, this ideal situation often does not exist. Learning algorithms with low bias values tends to be more "flexible," responding to the higher complexity of the model and thus being able to fit the data very accurately. However, overly flexible learning algorithms will fit different training sets in completely different ways, resulting in higher variance values as well. This phenomenon is often also referred to as overfitting: that is, models that use too many parameters can bring the loss function values down to very low during training but instead have a higher error rate when predicting new samples.

Figure 3 shows the comparison diagram of action recognition process under traditional machine learning algorithm and deep learning algorithm, from which it can be found that compared with the traditional machine learning algorithm which requires a lot of manual feature extraction work, the deep learning algorithm often takes the original data as input directly, extracts the abstract features of the data layer by layer through the hierarchical structure of the network, and finally realizes the mapping to the target output. From the input of raw data to the acquisition of the task target, deep learning automatically completes the integrated work of feature representation, feature selection, and model learning.

The first step in a sports analysis study is to break down the underlying movements for the specific sport in the context of the project. This part often requires a combination of expertise in the field of sport. The most famous application of this aspect is the Laban dance score, which laid the foundation of human kinetics and was one of the first cases of using computer notation to record human movement and analyze it logically. The greater the continuity of movement and the greater the degree of freedom of the limbs, the more difficult it is to disassemble. Ideally, with an infinitely large sample size of training data and a near-perfect modeling algorithm, we could obtain a model with very small bias and variance, but in real engineering problems, this ideal situation does not often exist. Learning algorithms with low bias values tend to be more "flexible," responding to the complexity of the model and thus being able to fit the data very accurately. The vast majority of sports in the matter are far less difficult to disassemble than dance, so there is a well-established system of disassembling basic movements in the field of their teaching long ago. Under the premise of focusing only on the geometric nature of the movement, the human body can be reduced to a skeleton model, while completely ignoring muscle movement, trunk movements can mostly be described more accurately with a combined rigid body model, and only movements that are suitable for rigid-body modeling expression and more concerned with the movement process are suitable for the inertial analysis scheme. Under the rigid body kinematic model, inertial data is the most natural and suitable data for quantitative analysis of human movement form.

3.3. Experimental Verification and Conclusion. The application of human action data collected by inertial sensors to action recognition, whether online or offline, is a pattern recognition process; we can summarize the overall process specification as follows: first for the modeling of the motion background, to complete the basic action classification system, followed by the design of the acquisition and tagging scheme, in addition to recording the inertial data of each action sample, which must also record the matching action tags, in addition to using inertial motion capture devices to capture human body information, it is necessary to ensure that the devices have a certain accuracy and sampling rate to reflect the real action situation as realistically as possible. The specific capture device is called an inertial measurement unit, which captures the linear acceleration signal of the movement through an accelerometer, the rotation rate of the movement through a gyroscope, and in some cases a magnetometer for heading reference. A typical configuration has a single-axis accelerometer, gyroscope, and magnetometer on each of the three airframe axes (pitch, roll, and yaw). The three-axis IMU allows for the complete recording of point motion information at fixed parts of the body. In this way, the inertial sensor converts the rich and complex motion information into a finite-dimensional digital signal. Figure 4 illustrates the inertial data for two types of action examples in the tennis action dataset collected in this experiment, and observation of the above figure reveals that very little information can be obtained from the action curves. From a cognitive point of view, there is no intuitive connection between the curves and the specific "forehand lunge" and "forehand serve high" movements, although the raw signals collected by the inertial sensors are a faithful record of the real movements, which are complete and comprehensive enough. Some studies in motion modeling have shown that motion reduction can be achieved with inertial data. But the sensor data does not directly reflect the properties of the tennis action. A clear correspondence between it and the actual motion cannot be easily established at the human cognitive level; in other words, the correspondence between the raw data and the actual problem is difficult to understand, especially for algorithmic models that are less intelligent than humans.

The determination of the threshold parameters is at the heart of the interception algorithm. Due to the diversity of action modes covered in the collection, the energy base of each action varies, and even the magnitude difference between different performers under the same class of actions is huge, so it is not practical to use a constant value as the threshold to complete the interception of all actions. The equipment mounting solution of fixing the measurement
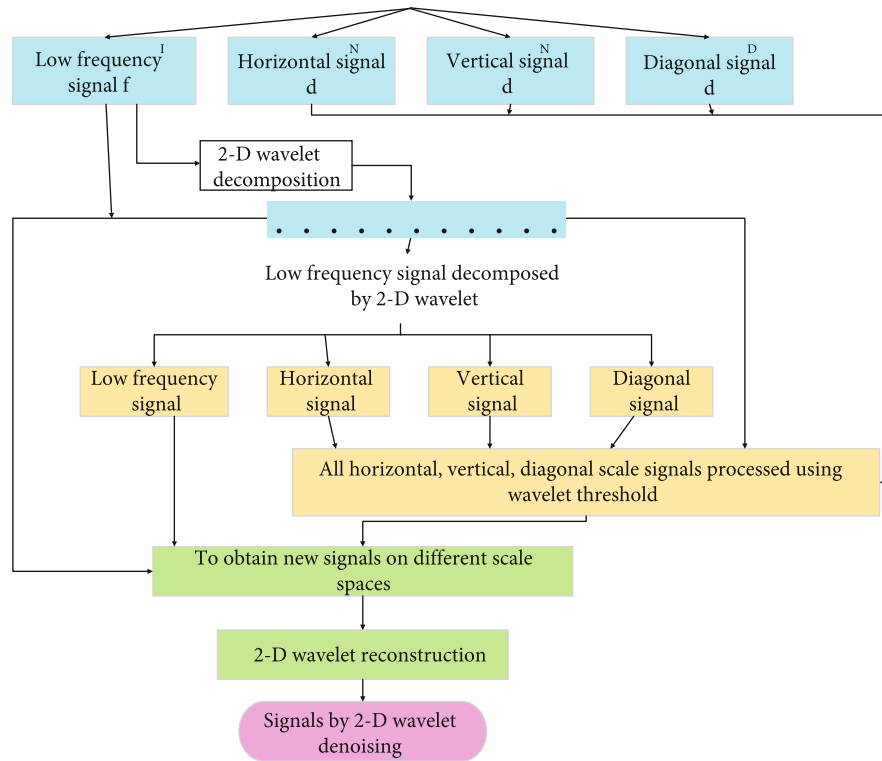
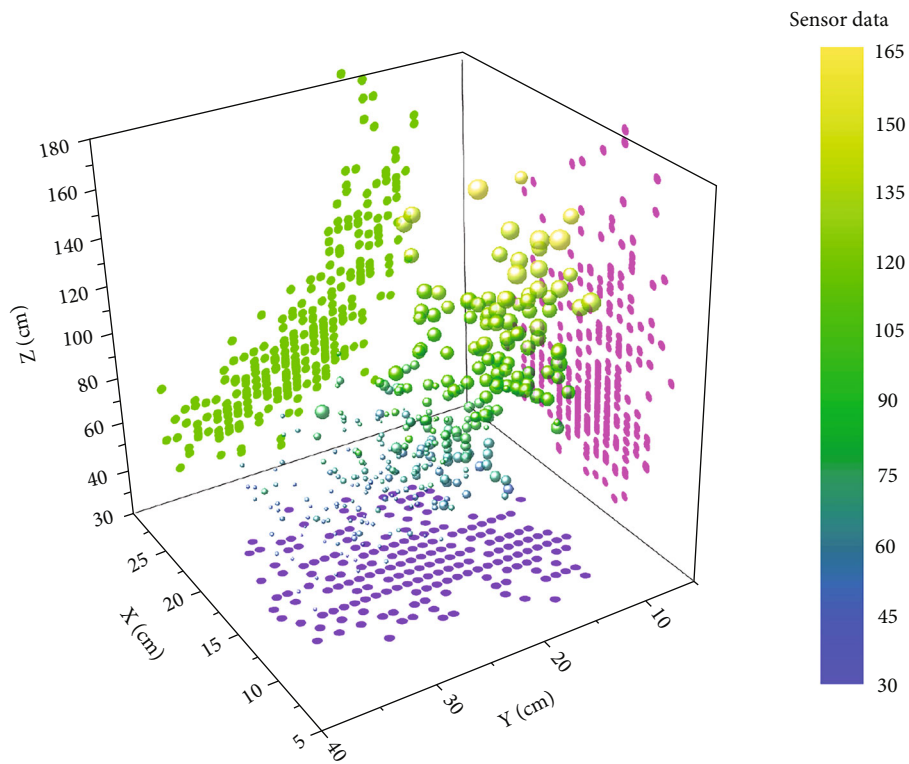FIGURE 3: Comparison of action recognition process under traditional supervised learning and deep learning.

FIGURE 4: Raw inertial sensor data plot.

device to the sports equipment, while minimizing the obstruction to the collector's movement, can also lead to the resulting tennis action dataset not being sensitive enough

to the distinction between grip styles, and based on this situation, mounting the motion acquisition equipment set on different sides of the racket is a viable solution. So, a
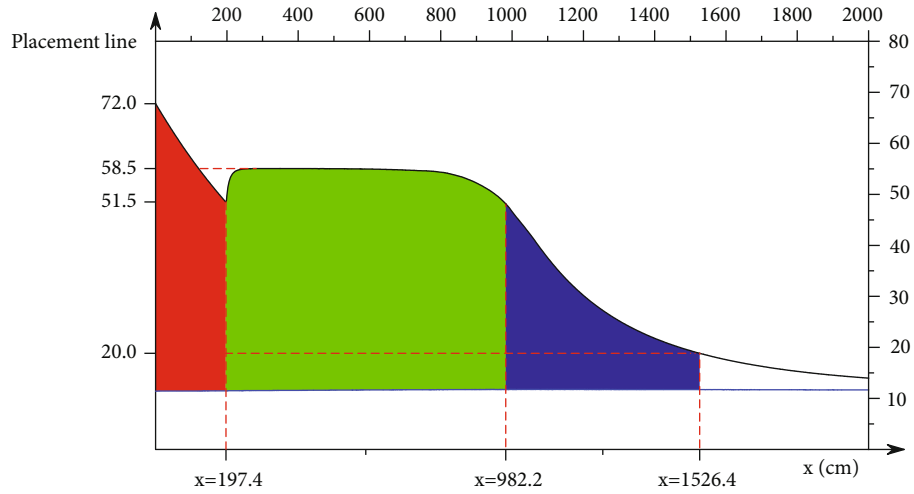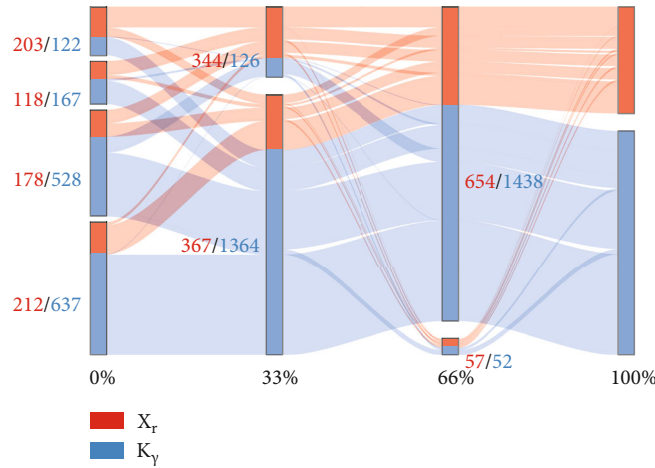
FIGURE 5: Gyroscope data energy curve.

FIGURE 6: Schematic diagram of the variation pattern of quantile values.

threshold determination scheme with adaptive capability needs to be proposed. First, observe the gyroscope data energy profile for a sample action as shown in Figure 5. A series of quantile arrays are calculated for the energy sequence (before smoothing), and the quantile lines at different percentiles are plotted, from which it can be found that the distribution of energy values for a segment of the action is mainly concentrated in the smooth segment of the action, the reason being that the action signal in this segment is mainly caused by random body jitter of the wearer and the degree of fluctuation of the data points is high.

The experimental hypothesis for the variation pattern of the quantile values is that the quantile values increase dramatically at the beginning of the action segment. To confirm this hypothesis, the variation curves of the quantile values and their difference curves were plotted for uniform increases of X_r values from 1% to 100% as shown in Figure 6. By the experiment, it can be observed that there is a steep increase in the quantile values near 50%. Substituting the quantile values at this point into the inertial data plot to do the verification basically matches with the starting and ending thresholds of the data, and the pattern is verified on the data of other kinds of actions. The partic-

ular quantile point obtained throughout the hypothetical process experiment was then referred to as the maximized group clustering quantile value, in the sense that it maximizes the concentration effect of the low-amplitude motion segment and continues to increase the quantile increasing the spacing between quantile values significantly. The mathematical determination method of maximizing the cluster clustering quantile value not only requires first plotting the quantile value change curve but also requires the minimization of the squared difference as the objective function for the line fitting and taking its inflection point, and such a calculation process is undoubtedly very complicated in practical application.

The number of anchor nodes is the number of attributes in the localization decision. The anchor node ratio is the ratio of the number of anchor nodes to the total number of nodes in the localization area. Adjusting the anchor node ratio will affect the localization performance. Increasing the ratio of anchor nodes will increase the deployment cost. Therefore, the performance of the proposed algorithm with different anchor node ratios is simulated to find the best value that meets the localization accuracy requirements and saves energy. In performing the simulation, three neural
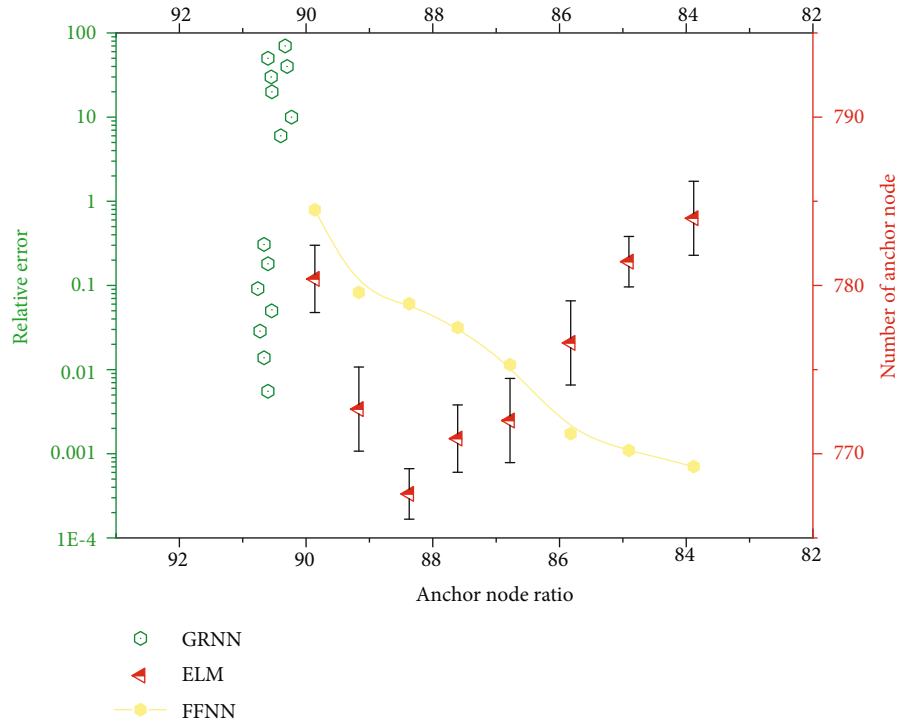
FIGURE 7: RLE of different algorithms with different anchor node ratios.

network-based localization algorithms, GRNN, FFNN, and ELM, are selected as the comparison algorithms. As can be seen in Figure 7, the RLE of all four localization algorithms in different localization areas decreases as the proportion of anchor nodes increases. With the transition from underground parking lots and indoor office areas to relatively less crowded areas such as campus roads and open activity areas, the reduction in crowd density reduces electromagnetic interference and small-scale fading in the surrounding environment, leading to a decrease in the RLE of all algorithms. From the figure, it can be seen that the relative localization error of the algorithm in this paper is the smallest, which is better than the three comparison algorithms and shows a stable decreasing trend in different regions. It indicates that the algorithm in this paper has the best localization performance; GRNN is slightly inferior, while ELM and FFNN have the worst performance.

In a multiarea localization scenario, the population density and geographic location of buildings affect the ambient noise level in the localization space, while the localization performance of the algorithm varies with the noise standard deviation. A larger noise standard deviation indicates a more disturbing environment and a harsher wireless environment in which it is located. To verify the adaptability of the localization algorithm to different regions in the localization space and the robustness to environmental interference, the variation of RLE with noise standard deviation in different regions is simulated. The simulation results are shown in Figure 8. In the four localization regions, the RLE of all four algorithms increases more significantly with the increase of the noise standard deviation. From the figure, it can be seen that the relative error of the FFNN algorithm fluctuates the
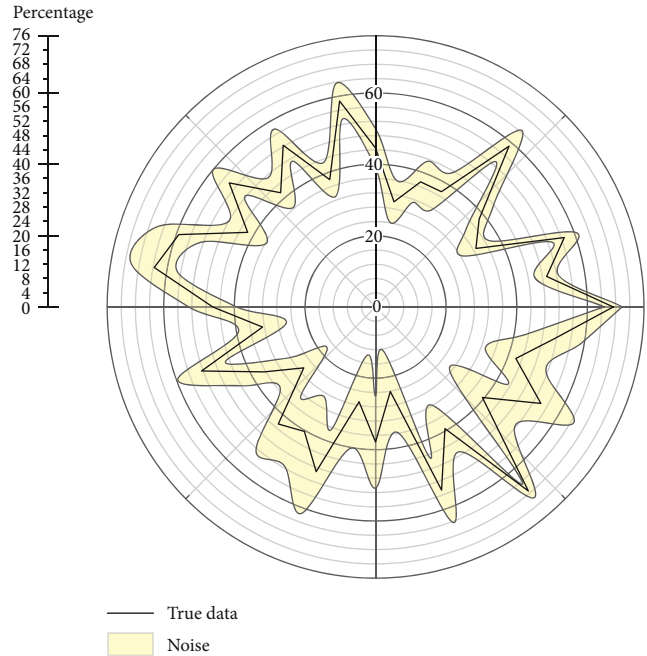

FIGURE 8: The effect of noise on the algorithm RLE.

most in the four regions, and the rising trend is more obvious. The localization error of the ELM algorithm also increases rapidly with the increase of noise standard deviations, especially in the underground sports field and the open region where the stability is poor. In contrast, the relative localization error (RLE) of the GRNN algorithm and the algorithm in this paper grows steadily. The RLE of the
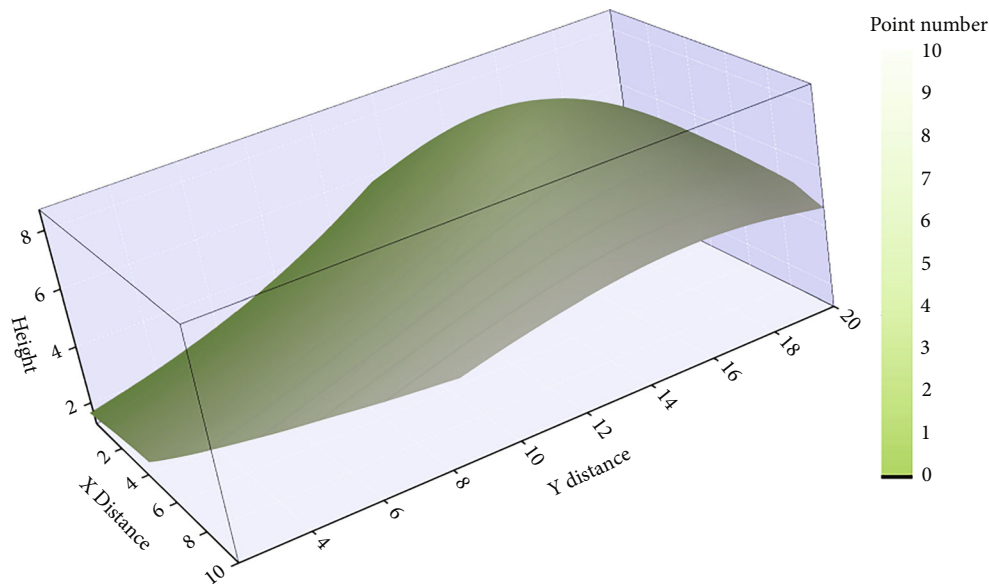
FIGURE 9: Accelerometer output graph.

algorithm proposed in this paper is significantly better than the comparison algorithm in four regions, and the difference value between regions is the smallest. The fluctuation range of RLE of the algorithm in this paper is smaller in the indoor sports area, underground sports field, campus sports field, and open area. It indicates that the algorithm in this paper has better robustness in different regions, can adapt to the changes of environmental noise, and has better stability of positioning accuracy.

The sequence length of tennis action data is unified to 128 samples by the resampling algorithm, i.e., each segment of action data is saved in the form of a 12×128 matrix, and the data matrix is expanded and spliced into a one-dimensional vector to be fed into the network learning. In addition, due to the translation property of sliding window segmentation, a segment of tennis action may be segmented into multiple data windows. A common data form is the basis for sharing research results. This paper gives a common inertial device standard, motion recording scheme, and data storage form and establishes a simple error calibration scheme for MEMS devices in motion capture application scenarios and a data cleaning method for the low automation of the data acquisition process. Observing the grayscale plot of the confusion matrix, it can be found that both recognition schemes are relatively easy to cause misjudgment for two types of tennis actions: forehand lunge and backhand lunge. From a practical perspective, this is because the two types of tennis actions are relatively close to each other, the trajectory of the racket is a lunge action, and the difference only lies in whether the player's grip is forehand or backhand. Figure 9 shows an example of the accelerometer output curve for the two motions.

Such experimental results reflect the fact that the device mounting solution of fixing the measurement device to the sports apparatus, while minimizing the hindrance to the collector's movement, can also lead to the resulting tennis action dataset being less sensitive to the differentiation of grip patterns, and based on this situation, mounting the motion acquisition device set on different lateralities of the racket is a viable solution.

4. Conclusion

This paper focuses on the study of motion recognition algorithms based on inertial motion capture schemes through wireless sensors. Since most of the current motion analysis is a shallow use of general algorithmic models, often not combined with expertise in the field of inertial guidance to target the characteristics of inertial data, and the research results are limited to small-scale motion datasets, this paper establishes a standardized motion recognition research process that best fits inertial motion capture schemes. This paper establishes a standardized motion recognition research process that best fits the inertial motion capture scheme, including a summary of data processing experience and a generalization of ideas for decomposing the emotion recognition task.

The main research work is divided into the following parts: (1) Acquisition and preprocessing of inertial datasets: a common data form is the basis for sharing research results; this paper gives a common inertial device standard, motion recording scheme, and data storage form and establishes a simple error calibration scheme for MEMS devices in motion capture application scenarios and a data cleaning method for the low automation of data acquisition process. (2) Motion interception algorithm research: from the offline recognition and online recognition of two research modes, focus on the needs of this paper and the implementation of the motion interception algorithm under the event window and motion window, respectively. To accurately detect the starting and ending points of motion, a stable motion amplitude indicator function is established using the Teager

operator combined with Gaussian smoothing filtering, a parametric modeling method for motion thresholds is derived, and an adaptive threshold determination scheme based on energy peaks is determined, which can accurately intercept the effective signal segments of various motions. Based on the professional research foundation in the field of statistics and signal processing, we designed a set of feature calculation schemes that can cover the motion characteristics to the maximum extent, including a total of 19 types of features under statistical features, signal time-frequency features, and system modeling features, and proposed a set of scientific feature contribution evaluation indexes based on the principle of information gain, and optimized and adjusted the applied feature combination scheme by combining the tennis action dataset, and the feature dimension was reduced by 20.78% under the streamlined combination. The feature dimensionality was reduced by 20.78%, while the classification accuracy only decreased from 97.99% to 97.60%.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] M. A. K. Quaid and A. Jalal, "Wearable sensors based human behavioral pattern recognition using statistical features and reweighted genetic algorithm," *Multimedia Tools and Applications*, vol. 79, no. 9-10, pp. 6061–6083, 2020.

[2] S. S. Tabrizi, S. Pashazadeh, and V. Javani, "Comparative study of table tennis forehand strokes classification using deep learning and SVM," *IEEE Sensors Journal*, vol. 20, no. 22, pp. 13552–13561, 2020.

[3] E. E. Cust, A. J. Sweeting, K. Ball, and S. Robertson, "Machine and deep learning for sport-specific movement recognition: a systematic review of model development and performance," *Journal of Sports Sciences*, vol. 37, no. 5, pp. 568–600, 2019.

[4] M. Batool, A. Jalal, and K. Kim, "Telemonitoring of daily activity using accelerometer and gyroscope in smart home environments," *Journal of Electrical Engineering and Technology*, vol. 15, no. 6, pp. 2801–2809, 2020.

[5] A. Jalal, S. Kamal, and C. A. Azurdia-Meza, "Depth maps-based human segmentation and action recognition using full-body plus body color cues via recognizer engine," *Journal of Electrical Engineering and Technology*, vol. 14, no. 1, pp. 455–461, 2019.

[6] J. Sena, J. Barreto, C. Caetano, G. Cramer, and W. R. Schwartz, "Human activity recognition based on smartphone and wearable sensors using multiscale DCNN ensemble," *Neurocomputing*, vol. 444, pp. 226–243, 2021.

[7] M. Rana and V. Mittal, "Wearable sensors for real-time kinematics analysis in sports: a review," *IEEE Sensors Journal*, vol. 21, no. 2, pp. 1187–1207, 2021.

[8] G. Lan, W. Xu, D. Ma, S. Khalifa, M. Hassan, and W. Hu, "Entrans: leveraging kinetic energy harvesting signal for transportation mode detection," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 7, pp. 2816–2827, 2020.

[9] A. Switonski, H. Josinski, and K. Wojciechowski, "Dynamic time warping in classification and selection of motion capture data," *Multidimensional Systems and Signal Processing*, vol. 30, no. 3, pp. 1437–1468, 2019.

[10] M. R. Keyvanpour, S. Vahidian, and M. Ramezani, "HMR-vid: a comparative analytical survey on human motion recognition in video data," *Multimedia Tools and Applications*, vol. 79, no. 43-44, pp. 31819–31863, 2020.

[11] P. Wang, G. Li, J. Liu, Z. Hou, C. Meng, and S. Guo, "Flexible, freestanding, ultrasensitive, and iontronic tactile sensing textile," *ACS Applied Electronic Materials*, vol. 3, no. 5, pp. 2195–2202, 2021.

[12] X. Li, Y. J. Fan, H. Y. Li et al., "Ultracomfortable hierarchical nanonetwork for highly sensitive pressure sensor," *ACS Nano*, vol. 14, no. 8, pp. 9605–9612, 2020.

[13] X. Chen, L. Gong, L. Wei et al., "A wearable hand rehabilitation system with soft gloves," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 2, pp. 943–952, 2021.

[14] X. He, W. P. Tay, L. Huang, M. Sun, and Y. Gong, "Privacy-aware sensor network via multilayer nonlinear processing," *IEEE Internet of Things Journal*, vol. 6, no. 6, pp. 10834–10845, 2019.

[15] B. Ning and L. Na, "Deep spatial/temporal-level feature engineering for tennis-based action recognition," *Future Generation Computer Systems*, vol. 125, pp. 188–193, 2021.

[16] M. Oshita, T. Inao, S. Ineno, T. Mukai, and S. Kuriyama, "Development and evaluation of a self-training system for tennis shots with motion feature assessment and visualization," *The Visual Computer*, vol. 35, no. 11, pp. 1517–1529, 2019.