Research Article

Research on Multiperson Motion Capture System Combining Target Positioning and Inertial Attitude Sensing Technology

Yifei Wang [1] **and Yongsheng Wang** [2]

¹*Faculty of Humanities, Shaanxi University of Technology, Shaanxi, China*
²*School of Art and Design, Lanzhou Jiaotong University, Gansu, China*

Correspondence should be addressed to Yongsheng Wang; wys_lz@126.com

The purpose of this study is to solve the problems of multiple targets, poor accuracy, and inability to obtain displacement information in motion capture. Based on fusion target positioning and inertial attitude sensing technology, Unity3D is employed to create 3D scenes and 3D human body models to read real-time raw data from inertial sensors. Furthermore, a gesture fusion algorithm is used to process the raw data in real time to generate a quaternion, and a human motion capture system is designed based on inertial sensors for the complete movement information recording of the capture target. Results demonstrate that the developed system can accurately capture multiple moving targets and provide a higher recognition rate, reaching 75%~100%. The maximum error of the system adopting the fusion target positioning algorithm is 10 cm, a reduction of 71.24% compared with that not using the fusion algorithm. The movements of different body parts are analyzed through example data. The recognition efficiency of "wave," "crossover," "pick things up," "walk," and "squat down" is as high as 100%. Hence, the proposed multiperson motion capture system that combines target positioning and inertial attitude sensing technology can provide better performance. The results are of great significance to promote the development of industries such as animation, medical care, games, and sports training.

1. Introduction

Human motion capture is widely utilized in film and television production, somatosensory games, sports training, medical rehabilitation, and human behavior analysis. This technology plays a vital role in promoting the development of relevant industries [1]. In the film and television production industry, the shaping of many characters uses human motion capture technology. Characters in movies are 3D virtual actors designed using computers. The 3D virtual animation actors are driven by capturing the movements and facial expressions of real actors, thereby making the actions of the virtual animator the same as the real actors' actions [2]. In somatosensory games, electronic games can be operated by swinging the handle if inertial sensors are added so that people can obtain visual stimulation and feel the happiness brought by electronic games [3]. In sports training, the real-time

monitoring and calculation of the athletes' limb movement process can help control the exercise amount of each athlete, achieving efficient and safe training [4]. In the medical rehabilitation field, human motion capture can provide accurate exercise programs by recording human body motion information for a long time, thereby assisting rehabilitation training for patients [5]. Therefore, human motion capture technology is a comprehensive interdisciplinary technology developed in recent years, involving computer graphics, ergonomics, communication technology, and other disciplines. It is currently a research hotspot in the field of human-computer interaction. However, the current human motion capture system faces problems such as more targets, poor accuracy, and inability to obtain displacement information, which severely restricts the development of related industries. Therefore, research on human motion capture technology plays a very important role in promoting the development of the industry.

Motion capture technology can accurately and in real time restore the human body's motion state in real life via a virtual 3D computer model [6]. Fusion methods are often utilized in motion capture. In China, research on the complexity of the fusion target positioning technology has started late. Also, the information processing level is low, which cannot reach the standard as a typical information fusion system. Valuable information fusion methods, such as compensation learning fusion method, active perception method, Mahle's random fusion method, the expert system method, physical model method, and parameter template method, are adopted [7]. The multiperson motion capture system often employs inertial sensor-based motion capture equipment, which can overcome the problem of insufficient light and shade in optical motion capture [8]. However, this device system faces a series of problems. The motion capture device based on inertial sensors needs to bind the inertial sensor nodes to the human joints to return the measurement information of the human joints during the movement in real time [9]. Besides, Xia et al. proposed that the inertial sensor's gyroscope could measure the angular velocity of the human body's joints relative to the human body coordinate system. The human body's current attitude could be obtained when the initial state of the human body was known. However, the gyroscope sensor chip had zero bias error, and long-term integration would cause the error to increase [10]. Liu et al. pointed out that as the time of the motion capture device increased, there would be zero offsets and drift. The system could not simultaneously capture multiperson motions [11].

Research significance: motion capture simplifies the human body into a multibody system, generally an articulated multirigid body system; input data are obtained by means of measurement or mode; the movement of the human body is simulated by describing the movement of the multibody system as the output result. The limitations of traditional motion capture (high price and restricted space) make the market urgently need affordable and portable motion capture equipment to be applied to various industries, such as improving the quality of physical education by capturing the sports characteristics of athletes and helping patients with rehabilitation by capturing normal human postures. Therefore, based on the existing motion capture methods, how to use a small number of sensors to accurately and stably capture motion has become one of the difficulties in this field.

Therefore, to solve the above problems, a multiperson motion capture system is designed and implemented using fusion target positioning and inertial attitude sensing based on analyzing relevant literature. Then, human behavior recognition is achieved on this basis. The innovative points are as follows: (i) the data of human inertial sensors are combined to generate quaternions, and the 3D human model is driven to precisely capture body movements. (ii) A multiperson motion capture solution combining target positioning and inertial attitude sensing is proposed. The fusion target positioning technology and inertial sensor positioning technology are combined to capture human spatial displacement information. (iii) Different equipment

and parameters are employed for continuous optimization. Finally, a multiperson motion behavior capture system with a better actual operation and higher accuracy can be obtained.

2. Research Method

2.1. Multiperson Sensor Capture System. The human motion capture system based on inertial sensors consists of hardware and software. The hardware acquires human motion information and transmits data, while the software realizes the functions of 3D animation display and human behavior recognition. The system includes the inertial sensor, the 3D animation display, the human body displacement positioning technology, and human behavior recognition. The inertial sensor captures human body movement information in situ. The inertial sensor measures nine data pieces at the same time, including a 3D magnetic field meter, a 3D gyroscope, and a 3D acceleration. These nine data pieces need to be fused first, and then the human body posture can be calculated. Three-dimensional animation display is to present the captured motion animation in real time in the form of 3D virtual characters on the computer. The human body displacement positioning technology introduces positioning technology to measure the displacement of the human body in space. It also improves the positioning accuracy as much as possible to obtain human displacement information during the capture process. According to the human behavior data captured by the inertial sensor human motion capture system, the human behavior recognition recognizes human movements through data processing, feature extraction, and machine learning. The specific structure is shown in Figure 1.

The system hardware includes the inertial sensor node module, the information receiving module, and the positioning module. The inertial sensor node collects human body information; the information receiving module realizes the communication between the wireless inertial sensor node and the computer; the positioning module acquires human body displacement information. The inertial sensor node module is bound with human joints to collect human joint motion information in real time. The inertial sensor node module includes the nine-axis inertial sensor, Zigbee transmission module, and microprocessor and power circuit. The nine-axis inertial sensor module includes a three-axis magnetometer, a three-axis gyroscope, and a three-axis accelerometer. It collects the three-axis acceleration, three-axis magnetic force, and angular velocity generated by the human limbs in real time. The sensor nodes are wirelessly transmitted sending the collected data to the receiving module through the Zigbee transmission module. When the inertial sensor is bound to the human joints, elastic bands with Velcro are used. The elastic bands can be easily bound to the human joints. The Velcro is glued to the back of the inertial sensor nodes. During use, the elastic band is tied to the human body joint, and the inertial sensor node with Velcro is directly attached to the strap. The information receiving module realizes the communication between the wireless inertial sensor node and the computer. The
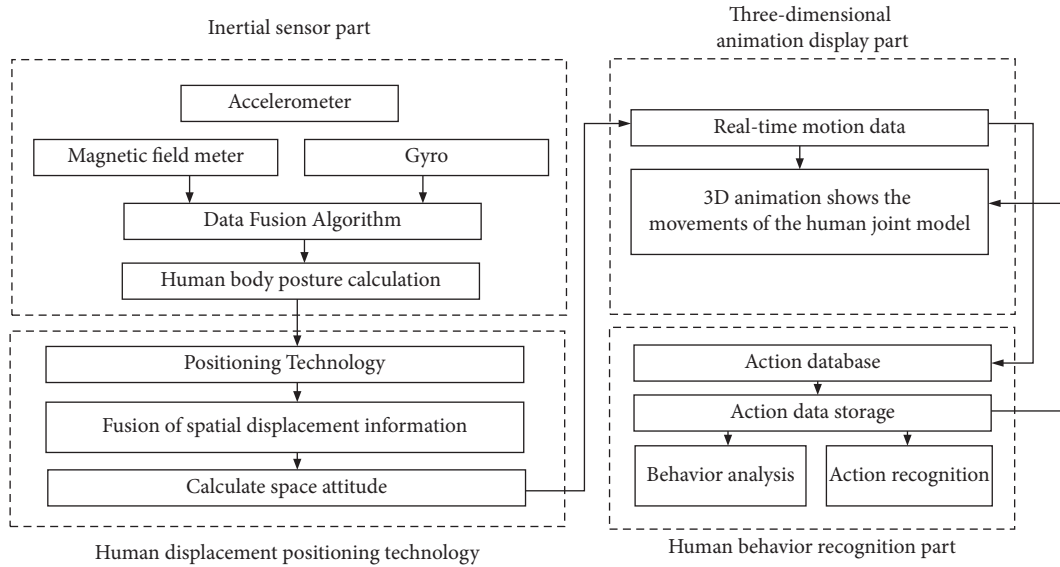
Inertial sensor part

Three-dimensional animation display part

Accelerometer

Magnetic field meter | Gyro

Data Fusion Algorithm

Human body posture calculation

Real-time motion data

3D animation shows the movements of the human joint model

Positioning Technology

Fusion of spatial displacement information

Calculate space attitude

Action database

Action data storage

Behavior analysis | Action recognition

Human displacement positioning technology

Human behavior recognition part

FIGURE 1: Overall design of the multiperson motion capture system.

information receiving module is connected to the computer through the USB interface. The information sent by the inertial sensor node is transmitted through the Zigbee wireless network, the information receiving module receives the node information and sends the received information to the computer through the serial port, and the data are processed in the computer.

The system software includes 3D scene creation, movement data storage, and human behavior recognition. Three-dimensional scenes are used for the reproduction of human motions; motion data storage stores captured human behavior data as files; human behavior recognition uses human motion information collected by inertial sensors and uses pattern recognition to identify human behaviors.

The process of reading data from the inertial sensor is displayed in Figure 2. After the information collection device is connected to the computer through the USB interface, the corresponding serial port is opened, and the baud rate is set to 115200. A thread for reading the inertial sensors and a thread for reading attitude data is created to solve the problem that the refresh rate of reading inertial sensor data does not match that of the human model animation.

The module of the inertial sensor node is bound to human joints for a real-time collection of human joint motion information. It includes human inertial sensors, Zigbee transmission, microprocessor modules, and power circuit modules. The principle of inertial sensor nodes is presented in Figure 3. The human body inertial sensor module includes a three-axis magnetometer, a three-axis gyroscope, and a three-axis accelerometer, which are used to collect the three-axis acceleration, three-axis magnetic force, and angular velocity generated by the movement of the

human body in real time. The sensor node is wirelessly transmitted through the Zigbee transmission module and sends the collected data to the receiving module.

2.2. Inertial Attitude Sensing. The attitude is calculated by magnetometer and acceleration. Despite the poor dynamic response, it will not produce a cumulative error. Therefore, the gyroscope, magnetometer, and accelerometer have complementary characteristics in the frequency domain. Hence, using complementary filters to fuse these three types of data in the inertial sensor can improve the measurement accuracy of the inertial sensor and the dynamic response performance of the system. The calculation steps are illustrated in Figure 4. First, the quaternion is calculated according to the initial state of the inertial sensor when fusing the inertial sensor attitude data. The gravity vector and magnetic field lines are then deducted to obtain and normalize the accelerometer and magnetometer data. After the matrices are multiplied, they are summed, and the proportional-integral controller is used to adjust the data. Finally, the quaternion is updated.

The sensor network can obtain the motion data of different feature points of the human body. The human body is abstracted and divided. Every movement of a limb can be regarded as a limb movement relative to the joint of the parent node. Since the hip joint has a small motion range, it is selected as the root node of the entire joint tree. The model of the human body's principal joints is presented in Figure 5. The corresponding sensors are placed at each joint to form a sensor network. The central controller of human body attitude data is controlled by different control modules. It can
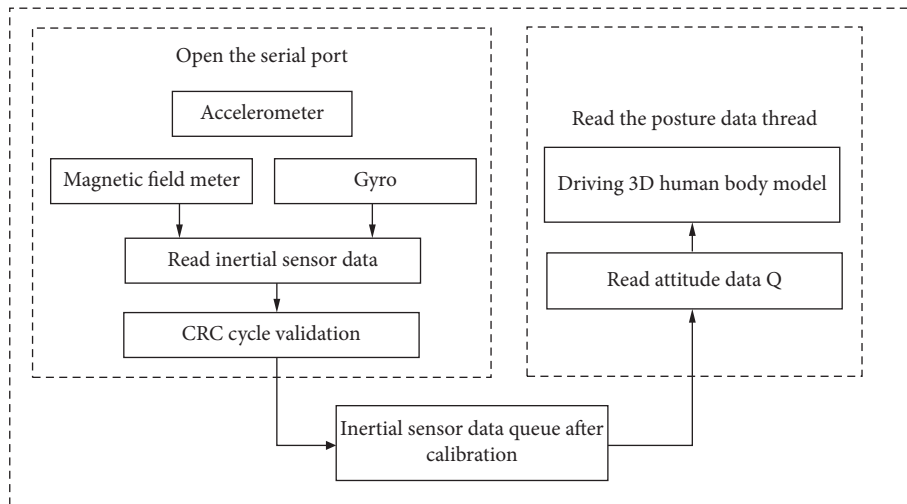
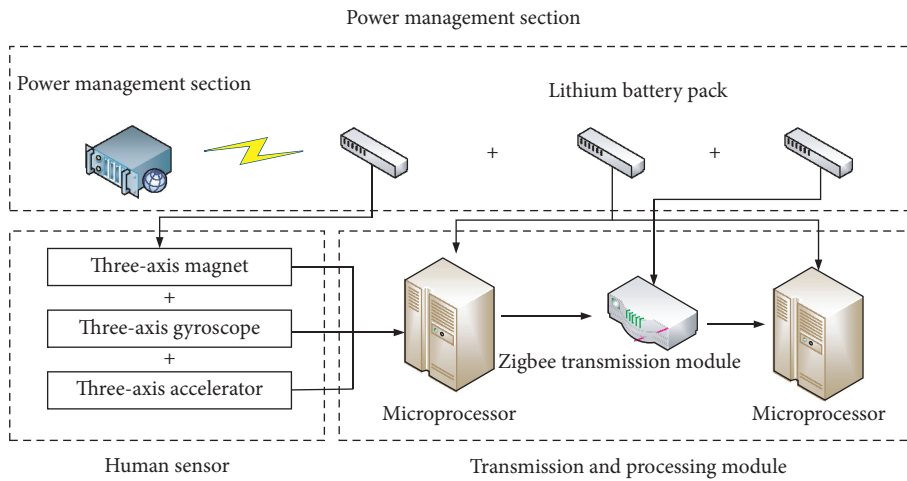FIGURE 2: The flowchart of inertial sensor data reading.
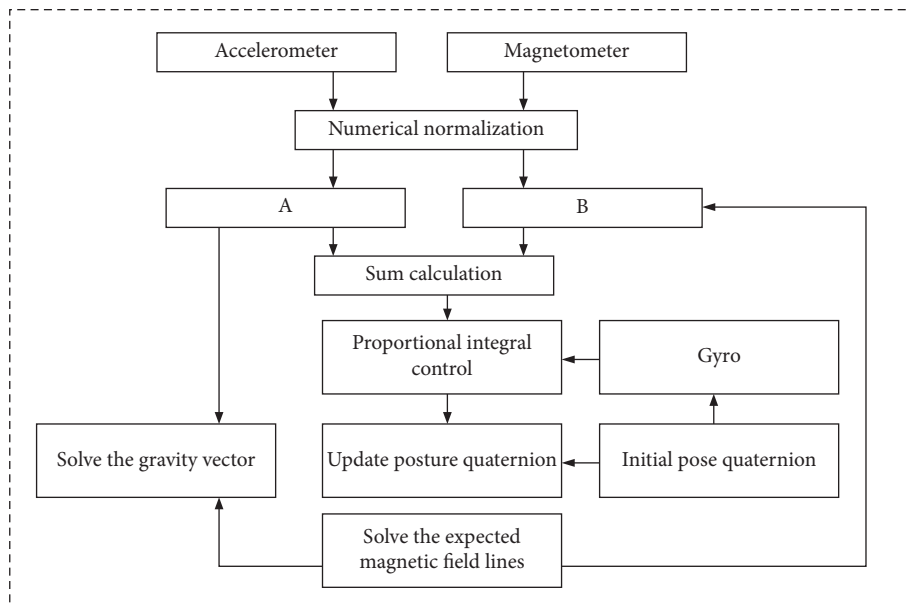

FIGURE 3: Inertial sensor nodes.

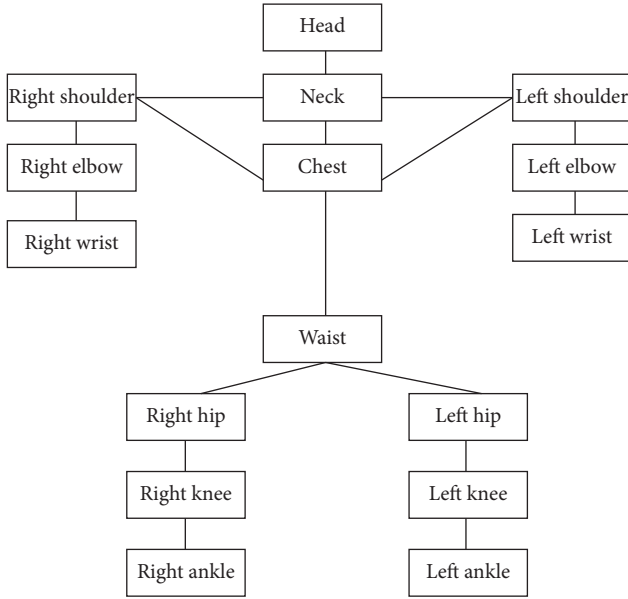FIGURE 4: The flowchart of the attitude fusion algorithm.

Figure 5: Human joint structure.

control different sensor nodes, obtain the data of the nodes, and integrate the wireless transmission module through programming to transmit the data wirelessly.

2.3. Human Posture Positioning. Because inertial sensor motion capture systems do not have an external reference point, they cannot directly obtain position information. Therefore, most inertial motion capture systems only estimate the posture of the human body in situ and cannot obtain accurate position information of the human body in global coordinates. Target positioning technology obtains the 3D position information of the captured target and records the complete movement information of the captured target through the data fusion algorithm. To solve the problem that inertial sensors cannot capture the displacement of the human body, researchers have proposed a variety of solutions, such as a Kalman filter-based inertial navigation system and radio frequency identification system tightly coupled method for indoor pedestrian positioning and navigation. A motion capture system that integrates inertial sensors, GPS, pressure sensors, cameras, and theodolites is employed to capture the 3D dynamics and kinematics information of the human body during alpine skiing competitions. The fusion target positioning technology and inertial sensor positioning technology are combined because the two sensors have good complementarity. In terms of indoor positioning, the fusion target positioning technology has higher positioning accuracy than other positioning methods. However, the fusion target positioning technology is easily affected by the nonline-of-sight effect during the positioning process. The nonline-of-sight effect is that two devices are blocked during the communication process, and the nonline-of-sight effect affects the positioning accuracy. The inertial sensors can make up for the shortcomings of ultra-wideband positioning

technology, thereby improving the positioning accuracy in the human motion capture system.

The fusion target positioning technology fuses and merges multiple methods, including the fusion of sensors, image recognition algorithms, and feature algorithms, to accurately position target actions and behaviors [12]. The integral method is adopted to calculate the position $X_{\text{SIM},n}$ of the human body at time step n in the x-dimension. The equation is as follows:

$$
X_{\text{SIM},n} = \begin{bmatrix} x_{\text{SIM},n} \\ y_{\text{SIM},n} \\ z_{\text{SIM},n} \end{bmatrix} = X_{P,n} + \frac{\delta_{xP,n}}{\delta_{xP,n} + \delta_{\text{IMU},n}} \left(X_{\text{IMU},n} - X_{P,n} \right).
\tag{1}
$$

In (1), $x_{\text{SIM},n}$, $y_{\text{SIM},n}$, and $z_{\text{SIM},n}$ refer to the 3D coordinates of the human body position at time step n, $X_{P,n}$ corresponds to the human body position coordinates at time step n in the x-dimension of the fusion target positioning, $\delta_{xP,n}$ represents the human body position coordinates of the inertial sensor in the x-dimension at time step n, $X_{\text{IMU},n}$ stands for the standard deviation of fusion target positioning in the x-axis to provide position estimation, and $\delta_{\text{IMU},n}$ denotes the standard deviation of the inertial sensor. Since it is difficult to predict the standard deviation of the inertial sensor, the three dimensions are all set to the same value. The following process is utilized to deal with the influence of the integration error. At time step n, if the currently estimated ultra-wideband (UWB) positioning standard deviation is less than the inertial sensor standard deviation of $n-1$, it is represented by

$$
\delta_{k,P,n} < \delta_{\text{IMU},n-1} \forall k \in \{x, y, z\}.
\tag{2}
$$

In (2), $\delta_{k,P,n}$ represents the position deviation of fusion target positioning at time step n, and $\delta_{\text{IMU},n-1}$ denotes the standard deviation of the inertial sensor at time step $n-1$. Then the standard deviation of the inertial sensor will be reset to

$$
\delta_{\text{IMU}} = 0.05.
\tag{3}
$$

Otherwise, the standard deviation of the inertial sensor is set as

$$
\delta_{\text{IMU},n} = \frac{0.05}{\sec^2 \cdot t_{sr}^2}.
\tag{4}
$$

In (4), t_{sr}^2 refers to the sampling frequency of the inertial sensor. Therefore, according to the error standard deviation of fusion target positioning, the infinite increase in the error standard deviation of the inertial sensor is limited, and the integral error is reset under the visual distance condition. Under the condition of nonvisual distance, if the error caused by nonvisual distance is less than the integral error, the integral error is reset. The flowchart of the fusion algorithm is shown in Figure 6. The workflow is as follows.

Here, Unity3D is adopted as a development tool to create a 3D scene and a 3D human body model. Besides, the 3D
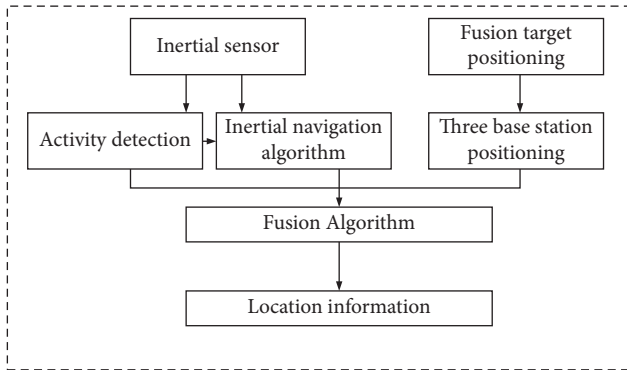
FIGURE 6: The flowchart of fusion target positioning algorithm.

human body model is employed to display the captured human movements in real time. Unity3D supports most mainstream 3D animation technologies and has a visual design environment, a scene editor that is easy to learn, and a convenient design process. Unity3D can well support 3D model files, saving the time of creating 3D scenes. The raw human motion data captured by the inertial sensor in real time is processed to generate an attitude quaternion. The motion parameters input to the virtual 3D human body model are quaternion data, and the quaternion data are converted into angle rotation parameter information in the bone pipeline.

Human body displacement positioning technology: there is no external reference point during human motion capture, and the spatial displacement information cannot be obtained because the inertial sensor is a self-contained system. Therefore, positioning technology is required to obtain the displacement information during human motion capture. Wireless positioning technology includes wide-area positioning technology and short-distance positioning technology [13]. Wide-area positioning is a large-scale technology, and the commonly used approaches include satellite positioning and mobile positioning. The short-range positioning technology has many kinds, including Wireless Local Area Network (WLAN), Radio Frequency Identification (RFID), Global Positioning System (GPS), UWB, Bluetooth, and ultrasonic. Table 1 summarizes the accuracy and scale of various positioning methods.

Ultra-wideband positioning technology is very different from traditional communication technology. Traditional communication technology uses carrier waves, while ultra-wideband positioning technology uses nanosecond-level narrow pulses with GHz-level bandwidth. Ultra-wideband positioning technology has been successfully applied to many fields, such as valuables storage positioning, mine personnel positioning, and parking lot positioning. Compared with other positioning technologies, ultra-wideband positioning technology has many advantages, such as strong penetration, low power consumption, resistance to multipath effects, and high positioning accuracy. As shown in Figure 7, the ultra-wideband positioning technology is used to obtain human body displacement information in the process of human body motion capture. To effectively improve the detection range and accuracy, a variety of

positioning combinations are prepared in this project. The system of this project will first call different positioning modules according to the corresponding distance so that different calling distances can be positioned in real time.

2.4. Behavior Recognition and Display. The 3D animation display includes three parts: 3D scene creation, action data storage, and recognition. The 3D scenes are used for the reproduction of human actions. Human action recognition uses human action information collected by inertial sensors. By limiting the range of motion of each joint in the human joint model, the captured human motion can be more in line with normal human motion. When capturing human motion, the inertial sensor node is bound to the designated joint of the body first. It is determined according to which joint movements need to be collected. Increasing the number of inertial sensor nodes can improve the accuracy of motion capture and make the captured human movements more accurate. According to the principle that each joint of the human body exists independently and interacts, the motion properties of the human body are restricted by the joints, and the motion of the human body is composed of several key joints. Therefore, it is unnecessary to consider the part of the joint that has less impact during human movement. The main joint model of the human body adopts a tree-like hierarchical structure, and the entire joint model of the human body is called a joint tree. The entire joint tree consists of a root node and multiple leaf nodes to form a parent-child relationship. In the process of human movement, every movement of a limb can be regarded as a limb movement relative to the joint of the parent node. Because the hip joint has a relatively small range of motion, it is selected as the root node of the entire joint tree.

The common ways of storing motion data include motion file storage, video file storage, and database storage. Motion file storage generally stores the motion information saved by animation editor software. The database storage can store the captured motion information in the database system; the commonly used database systems include MySQL and Oracle. Table 2 presents a comparison of various motion data storage methods. According to the needs of this project, the storage of human motion capture data requires a small space occupation and a simple storage structure. Besides, the third-party database software is not adopted to make other software more convenient to call the stored motion information. Bounding Volume Hierarchy (BVH) files occupy a small space and are easy to read and store. The BVH file uses Euler angles to describe the rotation data of the bone and limb joints, and the human body model is displaced in 3D space by changing the parameters of the hip node.

Human behavior recognition module: the multiperson motion capture based on fusion target positioning and inertial attitude sensing technology is a new research field, which overcomes many shortcomings and limitations of traditional motion recognition based on video sequences and has higher operability and practicality. The human motion information collected by inertial sensors uses BVH

TABLE 1: Accuracy and positioning scale of various positioning methods.

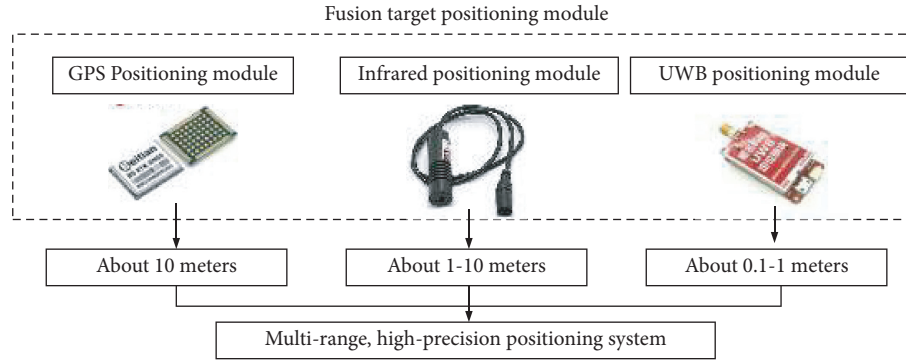| Positioning methods	Positioning accuracy (m)	Positioning scale
GPS	About 10	Outdoor wide-area positioning
Mobile positioning	About 10	Outdoor wide-area positioning
RFID, infrared, ultrasonic	1∼10	Outdoor local positioning
WLAN, Bluetooth	1∼10	Indoor positioning
UWB	0.1∼1	Indoor positioning

FIGURE 7: Human displacement positioning technology.

TABLE 2: Comparison of various motion data storage methods.

Storage methods	Storage types	Space occupation	Storage structure	Whether special software is required
BVH format	Motion file	Small	Simple	No
FBX format	Motion file	Small	Simple	Yes
Maya format	Motion file	Small	Simple	Yes
Database system	Database	Large	Complicated	Yes
Video format	Video	Large	Complicated	Yes

files to store human behaviors. Human behavior recognition preprocesses data and extracts features based on the data stored in BVH files. Moreover, it uses algorithms to recognize human behaviors. The process of human behavior recognition is to extract and select those feature vectors that have a large contribution to classification and recognition from BVH motion data. However, too many features will also cause redundancy between features and reduce the recognition accuracy. Hence, the principal component analysis is required. Finally, the appropriate motion classification algorithm is selected for behavior recognition.

2.5. Data Performance Comparison.

(1) Data source: data are read from the inertial sensor. After the information collection device is connected to the computer through the USB interface, the corresponding serial port is opened, and the baud rate is set to 115200. To solve the problem that the refresh rate of reading inertial sensor data does not match the refresh rate of the human body model animation, a thread for reading inertial sensor and a thread for reading posture data are created. The read inertial sensor thread realizes that the original data read from the inertial sensor consists of three parts,

namely, a three-axis accelerometer, a three-axis magnetometer, and a three-axis gyroscope, and then uses CRC16 to check the data. After verification, the data fusion algorithm is used to fuse the data into posture data and then put it into the Queue; the read posture data thread reads the posture data from the Queue and uses the posture data to drive the 3D human body model.

(2) Experimental environment: the behavioral information of eight healthy school students is collected during the experiment. The participants in the experiment include four females and four males, with ages of 22–27 years, height of 158–180 cm, and weight of 45–80 kg. Each participant in the experiment performs 16 actions, which were required by the experimenter. The types of actions are divided into three parts: upper limbs, lower limbs, and trunk according to the human joints. Each action is composed of these three parts. The main types of movements can be determined according to different movements. For example, the movement of waving a hand is mainly the movement of the upper limbs, and the movements of the trunk and lower limbs are relatively simple; thus, this type of movement belongs to the type of upper limbs. Before the behavior

collection, the subject of the purpose of the collection action is informed, and the behavior is demonstrated to the subject. After the subjects put on the inertial sensors, the behavior collection begins. Each behavior is repeated four times, and each behavior is required to be completed within 5 to 15 seconds.

(3) Performance comparison: the method of [14] is used to conduct a comparative analysis to further verify the effectiveness of the proposed model. A motion capture system based on multiple cameras is developed in the literature. To improve the capture accuracy, the multicamera layout method is used as a motion capture system for various animals with significant differences in appearance characteristics and motion behavior. Motion capture determines the demand boundary based on the object's appearance type (shape and space volume) and behavioral characteristics, thereby deriving a typical matching principle. A multicamera system locked in a 3D force measurement platform is used, its semicircular layout model shows that the error of the system is only below 3.8%, and the capture deviation rates are 3.43% and 1.74%, respectively.

3. Results and Discussion

In the result section, the appropriate model kinematics method is determined first, and then the method is used to conduct experiments under different conditions to determine the performance of the model. The effectiveness of the method is further determined by analyzing the human body motion trajectory. To prove that the proposed method has better performance than other models, it is compared with the method in [14], and finally, a human body model for motion capture is given.

3.1. Model Kinematics Method Determination. The forward kinematics of the human body is the posture of the human body displayed when the rotation angle of each joint of the human body is known. The position of any joint of the human body can be calculated by the connection of the human body joint, the joint length, and the rotation angle. There is only one general forward kinematics solution. Human body inverse kinematics gives the parameters of each joint based on the known posture. Therefore, it is difficult to solve the problem of inverse kinematics, and the solution result is not unique. In the process of human motion capture, the legs have fewer joints, simple movements, and relatively small calculations. Therefore, it is more suitable. The inverse kinematics of the legs calculates all the joint variables corresponding to the limbs based on the position and posture of the limbs. In this movement, the foot is the end, and the trigonometric function can be used to calculate the movement data of the thigh and calf when the foot is moving. Because the motion range of the human foot joints is limited, the foot and calf are treated as a whole in the process of using inverse kinematics motion capture. The four-position movement is involved in the process of body rotation so that quaternary data are used for

calculation. Figure 8 shows the analysis result of leg kinematics information data, and the left leg is stepped forward. Figure 8(a) shows the calculation result of quaternary data using inverse kinematics, and Figure 8(b) shows the fourth data of forwarding kinematics. In the meta-calculation results, as can be seen from the figure, the motion capture accuracy of inverse kinematics is higher than that of forwarding kinematics. In the same data image with different spatial coordinates, the highest model accuracy of inverse kinematics is 0.9178, and the highest model accuracy of inverse kinematics is 0.5562%. The two are quite different so that the inverse kinematics method is finally selected for image analysis.

3.2. Model Performance under Different Conditions. Figures 9(a)–9(f) display the results of the system capture distance under different line-of-sight algorithms. The positioning data of the X-axis and Y-axis fluctuate under the line-of-sight conditions; overall, however, they tend to be stable. Under the nonline-of-sight conditions, the model data without the fusion positioning algorithm fluctuates greatly. The fluctuation ranges of X and Y are [100–150] and [110–150], respectively. After the fusion algorithm is adopted, the fluctuation trend of positioning data is significantly reduced. The fluctuation ranges of X and Y are [125, 135] and [125, 140], respectively. Apparently, the maximum error of using fusion target positioning alone is 35 cm, while the error of fusion target positioning and the inertial sensor is 10 cm, a reduction of 71.24%. Moreover, the accuracy of model positioning is significantly improved.

3.3. Human Body Motion Trajectory Analysis. Figure 10 shows the results of the model positioning motion trajectories under different positioning systems.

Figure 10(a) presents the fusion target positioning system, and Figure 10(b) shows the method of combining fusion target positioning with inertial sensors. In the entire process, both methods can provide a clear positioning of the target. However, for the model without inertial sensors, the motion trajectory is narrower, showing a large gap with the actual moving process. In contrast, the fusion model can provide a higher positioning accuracy, and all the moving processes within the range can be shown.

3.4. Results of Human Gesture Recognition. Figures 11(a) and 11(b) display the comparison result of the proposed method and the classification algorithm of [14] on human behavior recognition. The behavior with the worst recognition effect is "draw X" and "draw circle," and the recognition result is 62.5%; the behavior with the best recognition effect is "wave," "throw," "crossover," "pick up," "walk," and "squat down," and the recognition result reaches 100%. In terms of the method in [14], the behavior with the worst recognition result is "draw X," "draw circle," and "forward," and the recognition result is 75%; the behavior with the best recognition result is "wave," "crossover," "triangle," "pick up," "run," "walk," and "squat down," and the recognition result reaches 100%. Both classification algorithms have the best recognition results for
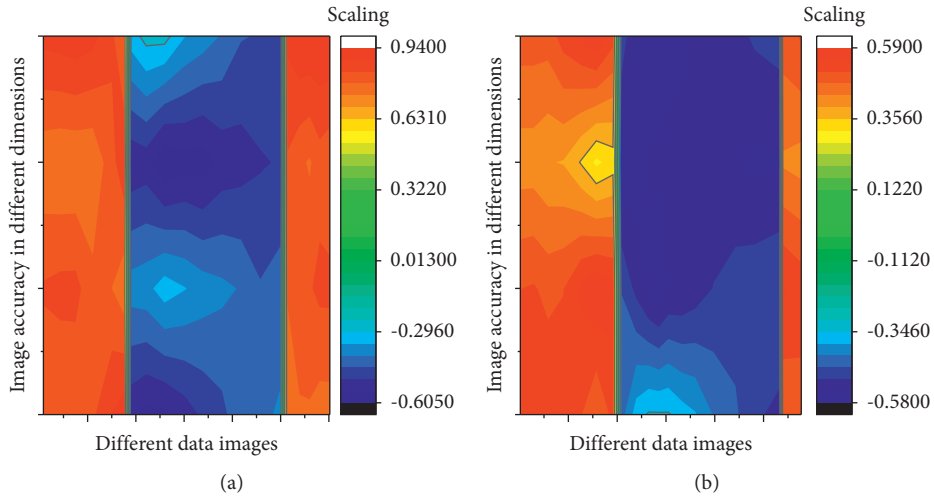
(a)

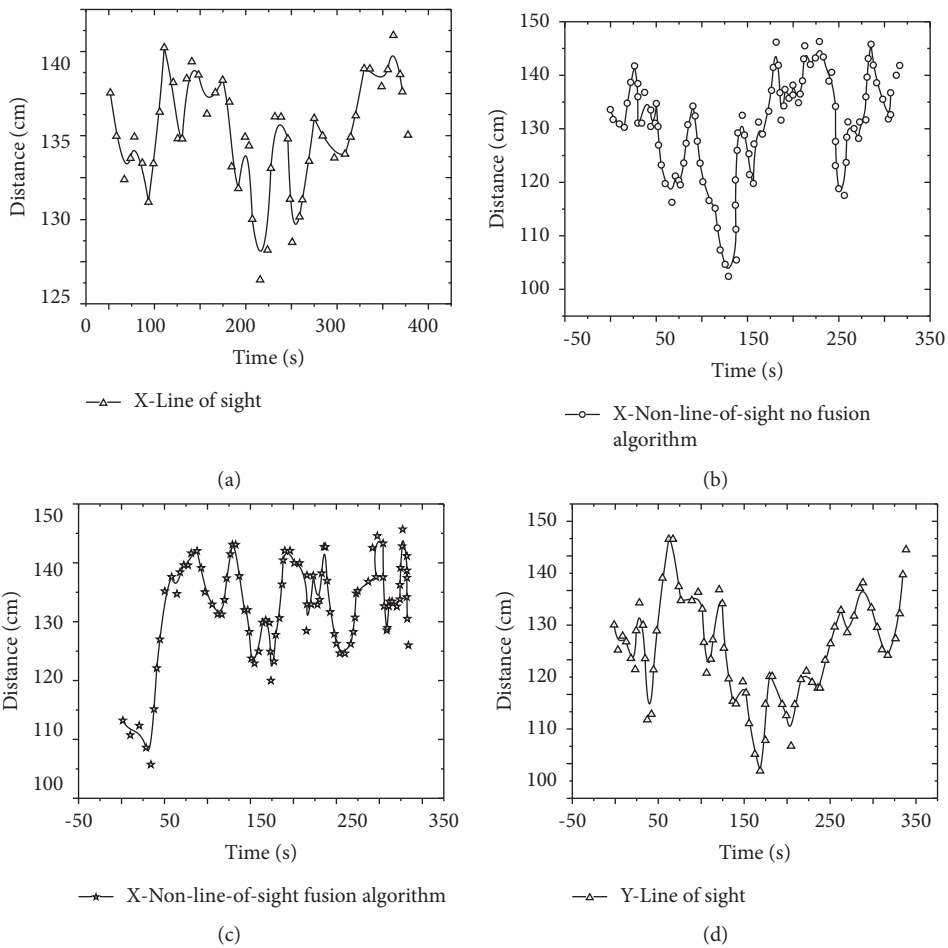(b)

Figure 8: Different kinematic leg data information.

X-Line of sight

(a)

X-Non-line-of-sight no fusion algorithm

(b)

X-Non-line-of-sight fusion algorithm

(c)

Y-Line of sight
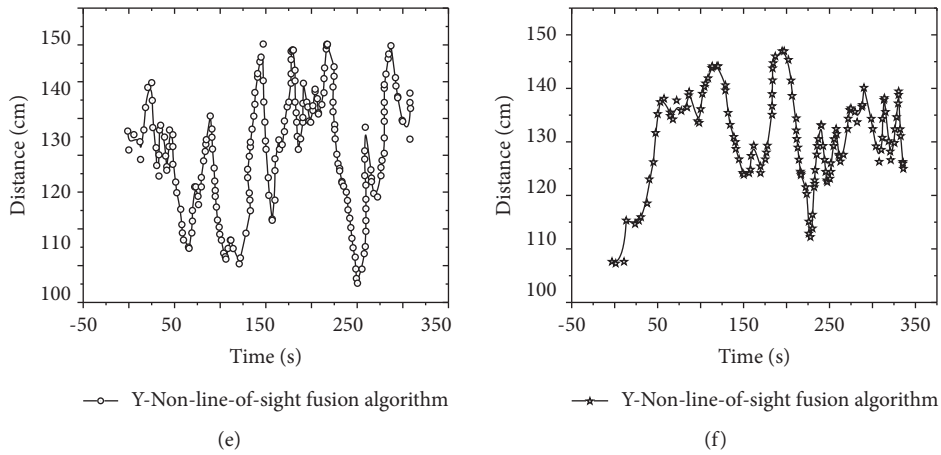
(d)

Figure 9: Continued.

(e)

(f)

FIGURE 9: System capture distance results under different line-of-sight algorithms.
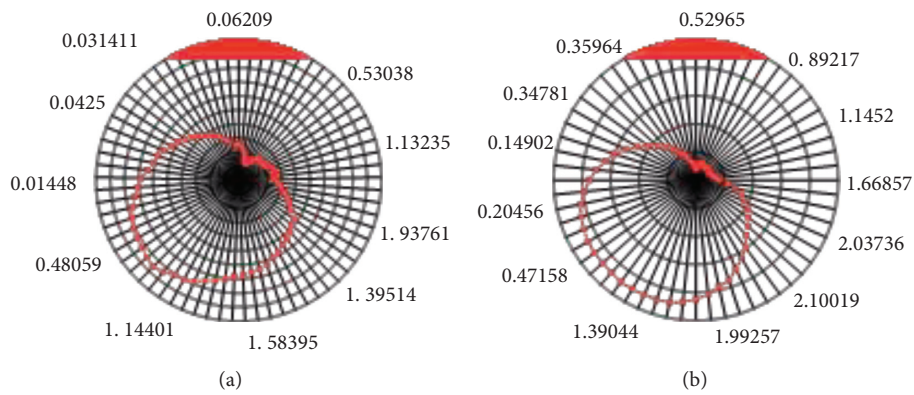

(a)

(b)

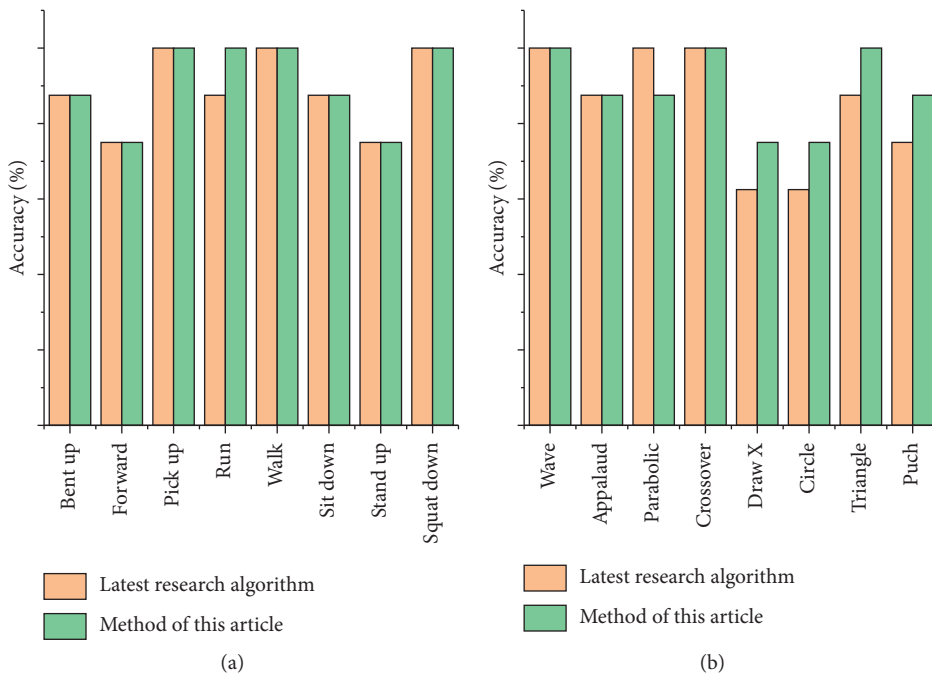FIGURE 10: Model positioning motion trajectory results under different positioning systems.

(a)

(b)

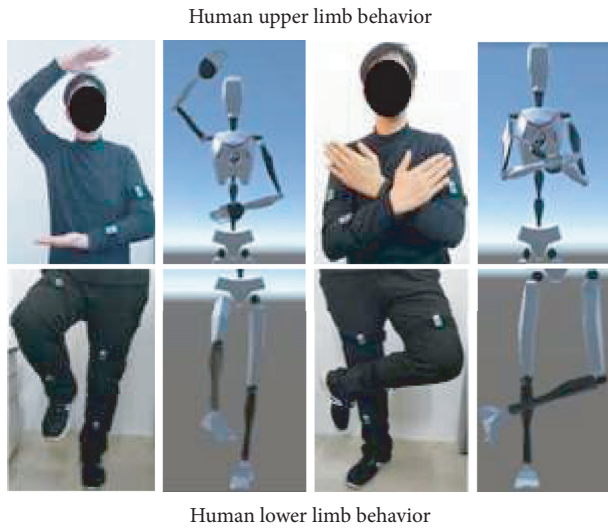FIGURE 11: Different behavior recognition results.

Human upper limb behavior

Human lower limb behavior

Figure 12: Results of human motion capture.

the behaviors of "wave," "crossover," "pick up," "walk," and "squat down," and the recognition results for the behaviors of "draw X" and "draw circle" are the worst. On the whole, the classification result of [14] is better than the classification result of the method. This is because the collected behavior data samples are more balanced, and the classification needs to divide the behavior data into 16 groups, which is a multiclassification problem. Therefore, the classification result of the algorithm in [14] is better. According to the results of human behavior recognition, the behavior with a recognition rate of 100% has a larger motion range and is obviously different from other behaviors, such as "squat," "walk," and "pick up." The behaviors with low recognition rates are caused by other similar behaviors. For example, the behaviors of "draw circle" and "draw X" are very similar so that the recognition rate is low. Therefore, the design of the human body behavior recognition method can better recognize the behaviors with larger motion amplitude; besides, the motion is relatively single, and the recognition rate for behaviors with small changes in human joints is low.

3.5. Motion Capture Result Display. As shown in Figure 12, many experiments on the upper and lower limbs of the human body find that the motion attitude of the 3D human body model can capture the motion state of the upper and lower limbs of the human body in real time and accurately. Besides, the constructed model can capture the human body attitude in real time. In particular, complicated motions can also be effectively recognized. When the human body moves, muscles contract, and inertial sensors can also move joints through the jitter of nodes, proving the effectiveness of the proposed model.

4. Conclusion

The existing human motion capture systems based on inertial sensors and related research on human behavior recognition methods are studied. On this basis, a human motion capture system based on wireless inertial sensors is designed and implemented to recognize human behaviors. Moreover, a scheme that fuses inertial sensors and ultra-wideband positioning is proposed. The ultra-wideband positioning technology and inertial sensor positioning technology are combined together to capture human spatial displacement information. The captured human motion data are analyzed in depth based on motion capture. Experimental steps are designed, including human behavior data collection, human behavior data storage, dataset analysis, feature extraction, and pattern recognition. Based on the extracted feature parameters, different algorithms are used to recognize human behaviors. The algorithm has a recognition rate of 62.5%~100% for 16 human behaviors, which can better realize the recognition of human behaviors. Although a suitable multiplayer motion capture system has been constructed, there are still several shortcomings. First, in human motion recognition, the angular velocity is time-integrated to evaluate the angle of the joint. However, this is easily affected by the magnetic field and surrounding equipment. The algorithm of the fusion inertial sensor can be analyzed more profoundly to reduce unnecessary interference. Second, in the process of human behavior feature training, because there is no suitable dataset for training, the training data is less, and the accuracy of the model is not very high. In the following research, these two aspects will be analyzed and researched in depth to further improve the model of the multiperson motion capture system.

Data Availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] E. van der Kruk and M. M. Reijne, "Accuracy of human motion capture systems for sport applications; state-of-the-art review," *European Journal of Sport Science*, vol. 18, no. 6, pp. 806–819, 2018.

[2] T. Y. Mou, "Keyframe or motion capture? Reflections on education of character animation," *Eurasia Journal of Mathematics, Science and Technology Education*, vol. 14, no. 12, pp. 1649–1669, 2018.

[3] B. Yang, H. Dong, and A. E. Saddik, "Development of a self-calibrated motion capture system by nonlinear trilateration of multiple Kinects v2," *IEEE Sensors Journal*, vol. 17, no. 8, pp. 2481–2491, 2017.

 [4] J. A. Goldsmith, C. Trepeck, J. L. Halle et al., "Validity of the open barbell and tendo weightlifting analyzer systems versus the optotrak certus 3D motion-capture system for barbell velocity," *International Journal of Sports Physiology and Performance*, vol. 14, no. 4, pp. 540–543, 2019.

 [5] Z. Yang, M. H. Rafiei, A. Hall et al., "A novel methodology for extracting and evaluating therapeutic movements in game-based motion capture rehabilitation systems," *Journal of Medical Systems*, vol. 42, no. 12, pp. 255–263, 2018.

 [6] M. Yoshimura, M. Kitazawa, Y. Maeda, M. Mimura, K. Tsubota, and T. Kishimoto, "Smartphone viewing distance and sleep: an experimental study utilizing motion capture technology," *Nature and Science of Sleep*, vol. 9, pp. 59–65, 2017.

 [7] G. Liu, L. Shi, J. Xun, S. Chen, L. Zhao, and Y. Shi, "An orientation estimation algorithm based on multi-source information fusion," *Measurement Science and Technology*, vol. 29, no. 11, Article ID 115101, 2018.

 [8] A. Filippeschi, N. Schmitz, M. Miezal, G. Bleser, E. Ruffaldi, and D. Stricker, "Survey of motion tracking methods based on inertial sensors: a focus on upper limb human motion," *Sensors*, vol. 17, no. 6, pp. 1257–1265, 2017.

 [9] M. Rigoni, S. Gill, S. Babazadeh et al., "Assessment of shoulder range of motion using a wireless inertial motion capture device-A validation study," *Sensors*, vol. 19, no. 8, pp. 1781–1793, 2019.

[10] S. Xia, L. Gao, Y.-K. Lai, M. Z. Yuan, and J. Chai, "A survey on human performance capture and animation," *Journal of Computer Science and Technology*, vol. 32, no. 3, pp. 536–554, 2017.

[11] Y. Liu, Y. Zhang, and M. Zeng, "Sensor to segment calibration for magnetic and inertial sensor based motion capture systems," *Measurement*, vol. 142, pp. 1–9, 2019.

[12] W. Zhao, S. Han, R. Q. Hu, W. Meng, and Z. Jia, "Crowdsourcing and multisource fusion-based fingerprint sensing in smartphone localization," *IEEE Sensors Journal*, vol. 18, no. 8, pp. 3236–3247, 2018.

[13] R. Bharadwaj, S. Swaisaenyakorn, C. G. Parini, J. C. Batchelor, and A. Alomainy, "Impulse radio ultra-wideband communications for localization and tracking of human body and limbs movement for healthcare applications," *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 12, pp. 7298–7309, 2017.

[14] W. Zong, Z. Wang, Q. Xing et al., "The method of multi-camera layout in motion capture system for diverse small animals," *Applied Sciences*, vol. 8, no. 9, 2018.