SIN-CHN SCIENTIFIC PRESS PTE.LTD.

Molecular & Cellular Biomechanics 2024, 21(4), 747.
https://doi.org/10.62617/mcb747

Article

Sports training posture recognition method based on Kinect body sensor and internet of things technology

Jie Rong[1], Huijuan Dong[2], Lei Wang[3], Hongxia Sun[4,*]

[1] Physical Education Department, Xinxiang Institute of Engineering, Xinxiang 453000, China
[2] Wushu Department, Hebei Sport University, Shijiazhuang 050021, China
[3] Department of physical education, Tangshan Normal University, Tangshan 063000, China
[4] Department of Physical Education and Research, Beijing City University, Beijing 100083, China
*** Corresponding author:** Hongxia Sun, shc_123@126.com

CITATION

Rong J, Dong H, Wang L, Sun H. Sports training posture recognition method based on Kinect body sensor and internet of things technology. Molecular & Cellular Biomechanics. 2024; 21(4): 747. https://doi.org/10.62617/mcb747

ARTICLE INFO

Received: 7 November 2024
Accepted: 21 November 2024
Available online: 9 December 2024

Abstract: Most of the routine sports training posture recognition uses the principle of image processing method, which has strong limitations, and there is a problem of data loss in the recognition process. The recognition error is large, which reduces the accuracy of the recognition results. Based on this, a new method of sports training posture recognition is proposed by introducing Kinect body sensor and Internet of Things technology. First, the mathematical description method is used to model the human body in three dimensions to represent the continuous posture changes of the trainer. Secondly, Kinect somatosensory sensor and the Internet of Things technology collect the sports training action information of trainers, track and capture the movement of limbs and the whole body, and extract the sports training behavior characteristics of the recognized people. On this basis, the recognition algorithm of sports training posture is designed to achieve the goal of sports training posture recognition. The experimental results show that after the application of the new method, the recognition error of sports training posture is small and the recognition accuracy is high.

Keywords: Kinect body sensor; internet of things; sports; training; posture

1. Introduction

Under the background of the reform and development of the current education system, influenced by the objectives of the curriculum reform and the teaching plan in many aspects, sports training has gradually become the key teaching content to improve the quality of sports teaching in various colleges and universities, and the quality level of sports teaching in most colleges and universities has been significantly improved [1]. Physical training can improve the physical quality of students in an all-round way, which is conducive to the growth and development of students. In physical education teaching and training, it covers a wide range of contents and sports training postures, such as running, basketball, football, shot put, high jump, swimming, javelin, etc. [2]. No matter what kind of sports training posture, there are corresponding normative actions. Some students do not grasp the sports training posture properly in the learning process, and there are problems of non-standard movement posture [3]. Once the movement and posture of sports training are not standardized, on the one hand, it will affect the sports performance of students, and in serious cases it may delay graduation; On the other hand, it will reduce the quality of sports training and cannot achieve the expected training effect [4]. In order to solve the problem of non-standard sports training posture, we should identify students' sports training posture in the training process, obtain the existing problems in time, and guide and correct them to achieve the best training effect [5].

At this stage, most of the commonly used sports training pose recognition methods use image processing principles to obtain two-dimensional images of human posture through ordinary cameras, import two-dimensional information into the three-dimensional information model, and obtain the recognition results [6]. However, this recognition method has a serious problem of data loss in the process of two-dimensional information to three-dimensional information reconstruction, which affects the accuracy of training pose recognition [7]. Kinect somatosensory sensor is increasingly mature under the current development and research. Its application in sports training posture recognition can improve the low accuracy of traditional recognition methods [8]. Kinect somatosensory sensor is a third-generation human-computer interaction device based on somatosensory technology and a new generation of somatosensory interaction device integrating many advanced visual technologies. It changes the traditional interaction mode and realizes the interaction experience different from the previous wearable sensors [9]. The hardware R&D of Kinect somatosensor integrates a number of advanced technologies, including sound, electricity, light and mechanics. There are nearly hundreds of main parts and components required for sensor operation. If it is divided into micro units, the number can reach thousands. It is an electronic product with extremely complex process [10]. The infrared camera of Kinect body sensor is CMOS structure, which can receive reflection spot images in real time, identify targets, collect depth information, and realize the synchronization of target image color data stream, depth image stream and audio stream [11]. Based on this, on the basis of traditional sports training pose recognition methods, this paper introduces Kinect body sensor, combines it with the Internet of Things technology, and proposes a high-precision sports training pose recognition method.

2. Design of recognition method for sports training posture

2.1. 3D modeling of human body

The mechanical characteristics and laws of human movement reveal the mechanical mechanisms of bones, muscles, joints, and their interactions. It is pointed out that human movement is generated by the transmission of muscle force through bones and joints, and is influenced by gravity, inertia, and external resistance. In the recognition process, priority is given to joints and angles such as the knee and hip joints (crucial for gait stability and efficiency), as well as the shoulder and elbow joints (crucial for fine motor performance), because they play a critical role in understanding and predicting human motion patterns. By applying these biomechanical principles to 3D modeling and recognition, human motion can be more accurately simulated, providing scientific basis for fields such as motion analysis, rehabilitation training, and human-computer interaction.

In the sports training posture recognition method based on Kinect body sensor and Internet of Things technology designed in this paper, firstly, according to the characteristics of human motion posture, mathematical description method is used to carry out three-dimensional modeling of human body, and the problem of human motion posture description is transformed into the problem of human static posture description. Through the three-dimensional model of the human body, the continuous posture sequence of the trainer in a period of time is characterized, and then the corresponding posture of the human body at a specific time point in the process of sports training is obtained [12].

Determine the joint points of the human body and connect them to represent the static posture of the trainer [13]. Based on the camera perspective, obtain the three-dimensional coordinates of the trainer in the perspective, establish a three-dimensional spatial coordinate system, and then establish a three-dimensional human model [14]. The three-dimensional model of human body based on joint points designed in this paper is shown in **Figure 1**.

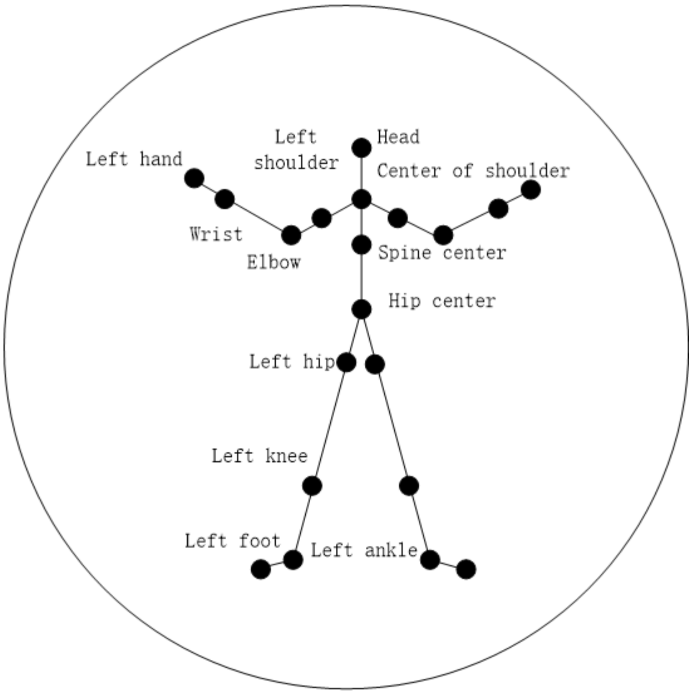

Figure 1. 3D model of human body based on joint points.

As shown in **Figure 1**, in the 3D model of human body designed in this paper, there is symmetry between the left joint point and the right joint point. Through this model, the characteristics of the human body movements and postures of the trainers are described. Special attention should be paid to that the characteristics of human movement posture should include complete information of all movement training actions. For a certain training action, its corresponding action characteristics should be distinctive, robust and cumulative [15]. Among them, the robustness of posture features in human movement training means that when the trainer is in different perspectives of the camera, the camera features are not vulnerable to interference, and the extracted posture features can accurately describe the movement training action [16].

On the basis of directly obtaining the three-dimensional coordinates of 20 joints, the model was simplified according to the topological structure of the human body and the characteristics of different limb flexibility [17]. The simplified model does not include the bones of the legs. The main reason is that in the motion training posture recognition method based on Kinect sensor and Internet of Things technology, posture recognition is carried out in the context of human-computer interaction, and the motion training action is a control command [18] with the action of the hand as the core. Secondly, in different movements, the human shoulder is relatively still, with the middle of the shoulder as the origin, and the shoulder is in the horizontal axis of the horizontal axis and the vertical axis of the three-dimensional coordinate system [19]. The three-dimensional coordinate system is combined with the three-dimensional model of the human body to represent and reflect the dynamic changes of human motion posture during the process of sports training [20].

2.2. Collect sports training action information based on Kinect body sensor and internet of things technology

After completing the three-dimensional modeling of the human body, next, use Kinect body sensor and Internet of Things technology to collect the sports training action information of the trainers, track and capture the movements of the limbs and the whole body, quantify the obtained sports training action indicators, and obtain the human sports training posture. First, design the operating parameters of Kinect body sensor, as shown in **Table 1**.

Table 1. Kinect sensor operating parameter settings.

No	Name	Parameter
1	Visual angle	Vertical direction 43°, horizontal direction 57°
2	Vertical movable angle	± 28°
3	Image frame	30FPS
4	Depth Image Format	QVAG
5	Color Image Format	VGA
6	Audio Format	16 kHz,16 bit
7	Infrared camera	CMOS construction; Emit infrared light beams to form light reflection.
8	Color camera	VNA3842956CMOS; Color image with a resolution of 1280 × 960, 12 frames per second, 640 resolution × 480, 30 frames per second.
9	Control motor	Adjust the Kinect viewing angle range.
10	Array Microphone	Accepts a 16 Hz audio signal, PCM mono.

As shown in **Table 1**, it is the operating parameters of the Kinect body sensor designed in this paper. On this basis, the body sensor is used to collect the movement information of sports training. At the key part of the trainer, fixed sensors are used to receive the corresponding sports training information [21]. At the same time, based on the magnetic detector fixed on the trainer, the magnetic field information around the sports training site is captured in real time, and the position of the training object [22] is determined by using the coupling relationship between the electromagnetic signals. The sensors on the trainer can follow the dynamic changes of movement and posture, and the magnetic field can penetrate the obstacles in front of the trainer without blocking, with a high degree of freedom. The marked points that can reflect light are coated and attached to the key points of the human body to be identified, and the light reflected by the marked points is captured by the camera with photosensitive effect [23]. Then the collected data is sent to the computer through the transmission line for corresponding pre-processing, so as to obtain human movement training action information.

2.3. Extraction of sports training behavior characteristics

After the collection of sports training action information based on Kinect somatosensory sensor and Internet of Things technology, the next step is to extract the sports training behavior characteristics of the identified person, providing a reference for subsequent sports training posture recognition.

Generally, when recognizing objects, it is necessary to distinguish the external contour of objects first. Generally, it is to recognize video image sequences. Image processing technology can easily obtain the contour edges of objects [24]. This method is simple and does not need to model the human body structure. Shape information can include contour, aspect ratio, area perimeter, curvature and projection histogram. The feature extraction model of sports training behavior designed in this paper is shown in **Figure 2**.

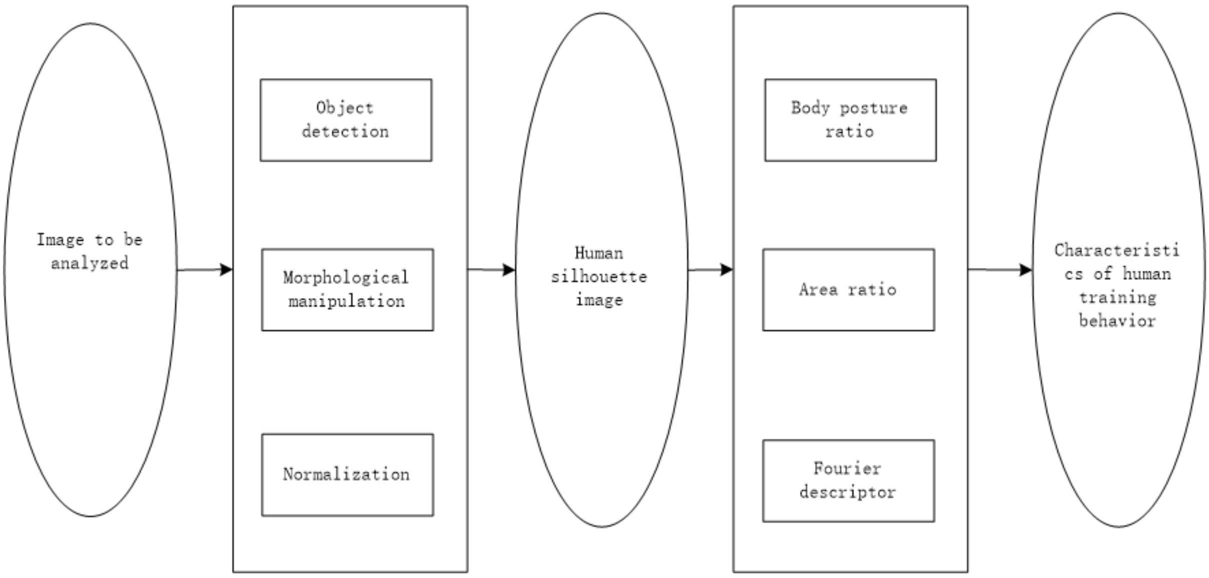

Figure 2. Extraction model of sports training behavior characteristics.

As shown in **Figure 2**, first use the Gaussian modeling method to extract the human body contour, then extract the edge feature data from the normalized image, establish the human action behavior feature vector through Fourier transform, and finally use the hierarchical recognition method and classifier to analyze and extract the human sports training behavior characteristics. The feature extraction model only needs to extract the edge contour features to build a model describing the human posture, which greatly reduces the computational complexity and storage. In the process of feature extraction of sports training behavior, it is necessary to simulate the multidimensional features of pixels, monitor the maximum value of spatial data, compare each sports training behavior sampling point and all its adjacent points, and observe the size of adjacent pixels [25]. The vector angle of human joints is selected as one of the main criteria for recognizing human motion. By fitting polynomials, locating joint positions and calculating their scales, the edge noise of joint positions is removed, and the vector direction distribution characteristics of key points are used to determine their respective orientations, so that the motion training behavior feature points have rotation invariance. Finally, the parameters of the operator are obtained through pixel matching to describe the local motion training behavior information in the recognition image.

2.4. Design of recognition algorithm for sports training posture

After the above sports training behavior feature extraction is completed, on this basis, a sports training posture recognition algorithm is designed to carry out the next step of posture recognition. This paper adopts the principle of joint point angle measurement method. First, set the coordinates of two points in space as $X(x_1, x_2, x_3)$, $Y(y_1, y_2, y_3)$, calculate the European distance between two points, the Equation is:

$$D(X,Y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2} \tag{1}$$

Through calculation, the Euclidean distance of two points in space is obtained. On this basis, in order to obtain the included angle of any joint in the process of sports training of the identified person, this paper obtains the three point coordinates of the joint. The schematic diagram of joint angle calculation is shown in **Figure 3**.

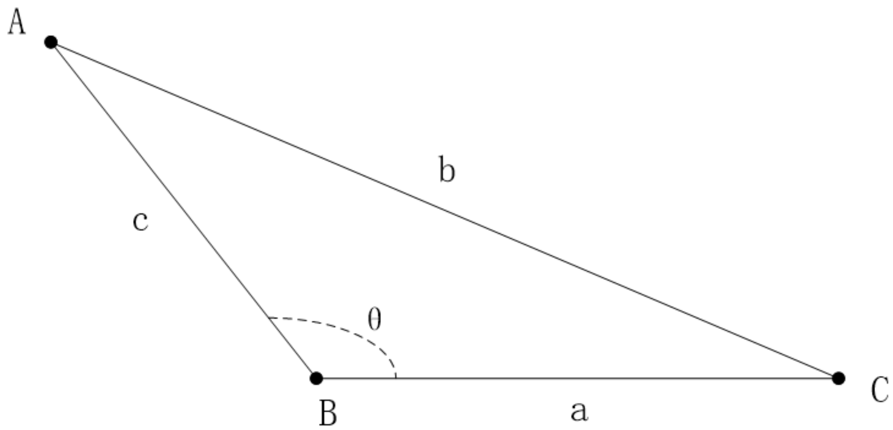

Figure 3. Schematic diagram of joint angle calculation.

According to the schematic diagram of joint point angle calculation in **Figure 3**, combined with the Euclidean distance obtained by Equation (1), the distance between joint points is calculated by using the cosine theorem. The Equation is:

$$\begin{cases} a = D(B,C) \\ b = D(A,C) \\ c = D(A,B) \end{cases} \tag{2}$$

$$\theta = cos^{-1}\frac{(a^2 - b^2 + c^2)}{2ac} \tag{3}$$

On this basis, in order to reduce the calculation error of the included angle of the joint points, the positive direction of the X axis is taken as the reference line, and the line between the two joint points is taken as the line to be measured. The line to be measured takes the outward direction of the central axis of the human body as the positive direction, and the outward direction of the transverse axis of the shoulder as the positive direction. The angle between the line to be measured and the reference line is calculated in a counterclockwise order. Define the angle as the angle of the two joint points. Ensure that the line to be measured and the reference line are relatively stable, thus improving the accuracy of the included angle calculation. The recognition expression of the joint point angle of the recognized person's sports training posture is:

$$P = \{P_1, P_2, \theta, \tau\} \tag{4}$$

Among them, P_1, P_2 they respectively represent two joint points of the identified person's sports training posture; θ indicates the included angle between the line connecting two joint points of the identified person and the reference line; τ indicates the angle adjustment threshold of the recognized joint points. The threshold value is adjusted by the angle between the joints to control the requirements of different recognition accuracy for motion training posture. Set the threshold range of the corresponding pose angle, read all the angle values first, and judge whether the angle value of the joint is within the preset threshold range. If all angles are within the threshold range, pose recognition is considered to be successful. If any angle is not within the threshold range, pose recognition is considered to be failed and requires re recognition. In sports training, the recognition of motion state is slightly different from the recognition of static state. The motion state presents a higher dynamic characteristic and requires a higher attitude angle. Through the observation of the user's posture and the division of the gait cycle, the angle obtained by the leg sensor changes the most. The sensors on the left and right thighs and calves are used to collect data, and the number of steps, step frequency and step speed in the process of sports training are counted to achieve the goal of sports state recognition in sports training. Finally, the recognition error is obtained by fusing the results of training pose recognition, and an error correction scheme is designed to improve the accuracy of training pose recognition results.

3. Experimental test

3.1. Test preparation

The above content is the whole design process of the sports training posture recognition method proposed in this paper using Kinect body sensor and Internet of Things technology. Before the proposed method is put into actual sports training, its feasibility and recognition effect are tested, and the following experimental tests are conducted. First, set up the environment required for this experiment test, and select a Microsoft Kinect for Windows 2.0 and a computer equipped with Windows 8 system. Establish a connection between Kinect and the computer, and ensure the reliability of the connection through USB. In the experimental test, Kinect somatosensory sensor is used to collect two-dimensional and depth images of human sports training posture in real time, and computer is used to continuously analyze and process the collected data.

Set the Kinect sensor parameters as shown in **Table 1**, which will not be elaborated here. At the same time, it is stipulated that the bone tracking accuracy of the Kinect sensor is 25 joint points with an error of less than 5 mm. The proposed recognition algorithm parameters are: (1) Feature extraction threshold: a joint angle change of more than 15 degrees is considered a valid action; (2) Time window size: 0.5 s (used for smoothing and reducing noise); (3) Pose matching similarity threshold: 85% (above this value, it is recognized as a correct pose). The above parameters were determined through experimental optimization and expert evaluation. The selection of resolution and frame rate for depth sensors is to obtain sufficiently clear image data while ensuring real-time performance. The accuracy of bone tracking is set based on the official technical specifications of the Kinect sensor and error testing in practical applications to ensure accurate capture of subtle movements of trainees. The feature extraction threshold and time window size are designed to filter out unnecessary noise while preserving key motion features. The similarity threshold for pose matching is obtained based on a large amount of test data to balance recognition accuracy and false alarm rate. The size and diversity of the training dataset ensure that the algorithm can generalize to different individuals and poses, improving the accuracy and robustness of recognition.

In order to correct the physical training posture of the trainers, a standard physical training posture must be used as a template for comparison in this experiment. By obtaining the information of professional sports coaches' sports training posture, occlusion points can be recovered to form standard posture data. In this paper, professional sports coaches were invited to randomly demonstrate several groups of standard sports training postures, and Kinect sensors were used to collect the data of sports training postures. Based on the occlusal point repair algorithm, the occlusal joints were visualized. Mark the data of each group of sports posture, match the data with the names of sports training sports posture one by one, save them as sports information in a unified way, and store them in the database. The database contains 20 groups of sports training action postures, and each training action postures also includes 5 detailed decomposition actions, that is, 100 action postures, which can fully meet the needs of sports training. This experiment takes running, high jump, long jump, basketball, football, volleyball sports training action posture as an example, uses the methods proposed in this paper to identify sports training posture, and test the recognition results.

The sports training posture image dataset used in this experiment and its introduction are as follows:

(1) Human pose estimation dataset: In this dataset, this article mainly utilizes annotated images containing exercise poses such as high jump, sit ups, push ups, and walking. These images are finely annotated with the positions of human body keypoints, such as keypoints. During the processing, the image is first preprocessed through size adjustment, normalization, and other methods to ensure the consistency of the input data. Then, the training set is used for model learning, the validation set is used to adjust the hyperparameters of the model, and the test set is used to evaluate the final performance of the model. By using three data augmentation techniques of rotation, scaling, and translation, the diversity of data can be further increased and the generalization ability of the model can be improved.

(2) MPII Human Pose Dataset: This dataset provides over 25,000 images containing various human poses, with each image detailing the joint positions of the

human body. When processing these data, this article first filters out images related to motion posture, and then preprocesses these images, including denoising, grayscale or color preservation. Next, generate a heat map of the joint points using the annotated information as input for model training. By implementing precise data segmentation and cross validation strategies, the model ensures accurate recognition of various human postures.

(3) UCF Sports Action Dataset: In this dataset, this article focuses on video clips of sports such as basketball, tennis, swimming, etc. These videos are first segmented into multiple action segments, each corresponding to a specific motion action. Then, preprocess each segment, including frame rate adjustment, size normalization, etc. In order to extract motion features from videos, this article uses Kinect motion sensors and IoT technology methods. Finally, utilizing these features for action recognition and behavior modeling, evaluate the performance of the model in different difficulty levels and scenarios.

(4) J-HMDB: This dataset contains videos of various human activities, many of which showcase demonstrations of various movement postures. When processing this data, this article first annotates each video to determine the keyframes and corresponding action categories. Then, human detection and pose estimation are performed on each keyframe to extract joint position and contour information of the human body. These pieces of information are used as inputs for model training to learn feature representations for different motion postures. By combining temporal contextual information, the accuracy of the model in recognizing human behavior can be further improved.

(5) YT-BB: This dataset is sourced from video clips on YouTube and contains rich sports content. When processing these data, this article first filters out video clips related to motion and annotates and classifies them. Then, each segment is preprocessed, including debounce, color correction, etc., to improve video quality. Next, deep learning models are used to extract spatiotemporal features from videos, which are then utilized for tasks such as behavior recognition and pose estimation. Through large-scale data training and model optimization, ensure that the model can accurately recognize various complex motion postures and behavioral patterns.

This paper conducted tests from the above datasets, and the specific test results are as follows.

3.2. Test results

Use the auxiliary training system to collect the actual coordinates of each joint point of the trainer's sports training posture, as shown in **Table 2**.

Table 2. Actual coordinates of joint points of sports training movement posture.

No	Position	Abscissa	Ordinate
1	Head	−81.2	215.4
2	Neck	−24.6	30.5
3	Left shoulder	−149.8	−51.4
4	Left elbow	−382.4	−52.6
5	Left hand	−614.7	−61.8

Table 2. (*Continued*).

No	Position	Abscissa	Ordinate
6	Right hand	94.3	−8.7
7	Right elbow	81.5	224.6
8	Right shoulder	44.7	512.6
9	Right knee	−106.5	−849.7
10	Right foot	−67.3	−1125.3
11	Left knee	−234.9	−962.4
12	Left foot	−315.2	−1142.8

As shown in **Table 2**, it is the actual coordinates of joint points of sports training posture. Select the manual node coordinate recognition error as the evaluation index of this experiment. Under the same sports training action posture, use the body sensor to measure and identify the hand node coordinates of the trainer, label them as 01–10, compare the recognition results of the hand node coordinates of the trainer with the actual coordinates of the hand node of the sports training action posture, and obtain the recognition error. Compare the manual node coordinate recognition error between the recognition method proposed by the Institute of Statistics and the single Kinect motion sensor method, and draw a comparison chart of the recognition error, as shown in **Figure 4**.

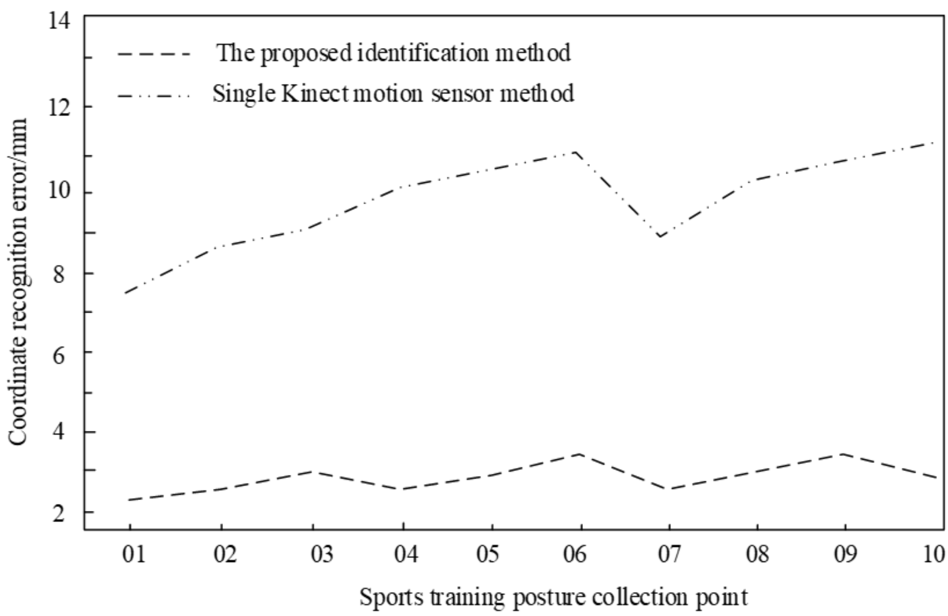

Figure 4. Comparison of posture recognition errors in two methods of sports training.

It can be seen from the comparison results in **Figure 4** that there is more or less difference between the trainers' sports training posture and the standard training posture. After using the sports training posture recognition method based on Kinect somatosensory sensor and Internet of Things technology proposed in this paper, the recognition error of the coordinates of hand nodes corresponding to 10 groups of sports training posture acquisition points is less than the recognition error of traditional methods, which can analyze and identify the trainers' sports training posture according to the trainers' joint point coordinates and the angle of each joint. The recognition result is of high precision, and targeted training suggestions are given to help trainers adjust sports training actions reasonably and achieve the goal of assisting sports training.

In order to evaluate the impact of simplified 3D human models (excluding leg bones) on the overall recognition accuracy of full body movements in motion training posture recognition methods based on Kinect sensors and IoT technology, two models were defined for experimentation. Among them, the complete model is constructed using the three-dimensional coordinates of all 20 joints to create a complete three-dimensional model of the human body. Simplify the model by constructing a simplified human 3D model using only the 3D coordinates of the upper body joints (excluding leg bones). Using accuracy (the proportion of correctly recognized actions to the total number of actions) as an evaluation metric, the accuracy of two models in recognizing whole-body movements was compared, and the results are shown in **Table 3**.

Table 3. Action recognition accuracy before and after model simplification.

Action type	Complete model accuracy/%	Simplified model accuracy/%	Difference/%
Arm waving	95	94	−1
Boxing	92	91	−1
Jump	88	86	−2
Jogging	85	80	−5
Balance exercise	90	89	−1

In **Table 3**, the difference is the percentage decrease in accuracy of the simplified model relative to the complete model. According to **Table 3**, in upper body dominant movements such as arm swinging, boxing, and balance exercises, the difference in accuracy between the simplified model and the complete model is small (both within 1%), indicating that leg bones have a relatively small impact on overall recognition accuracy in these movements. In movements involving full body movement such as jumping and jogging, the accuracy of the simplified model has decreased compared to the complete model (2% and 5%, respectively), but still remains at an acceptable level. This may be because leg movements play a certain role in these movements, but the simplified model can still recognize the overall movement trend well by retaining key information of the upper body. In summary, simplifying the model by removing leg bones reduces the complexity of data processing and computation, and improves the real-time performance and efficiency of the system. In motion training posture recognition methods based on Kinect sensors and IoT technology, hand centered control commands are the main interactive method. Therefore, simplifying the model by focusing on upper body movements is more in line with the needs of human-computer interaction. Although the accuracy of the simplified model has decreased in some whole-body movements, considering the needs of resource optimization and human-computer interaction, this trade-off of accuracy is reasonable. Especially in practical applications, adding Kinect sensors can further improve recognition accuracy.

Using a motion training posture recognition method based on Kinect sensors and IoT technology (as proposed in this paper), a vision based method, and a wearable sensor based method, data were captured from 10 volunteers performing specific motion training actions such as squats, push ups, jumps, etc. Analyze the comprehensive performance of three methods based on computational complexity, time delay, processing time, and system requirements. The results are shown in **Table 4**.

Table 4. Comprehensive performance of three methods.

Method	**Computational complexity**	**Time delay/ms**	**Single action processing time/s**	**System requirements**
This paper's method	Medium	50	0.2	Kinect sensor, medium configuration server
Visual based approach	High	150	1.0	High speed camera, high-performance computer
Method based on wearable sensors	Low	30	0.1	Multiple inertial sensors, low-power processor

According to **Table 4**, the motion training posture recognition method based on Kinect sensor and IoT technology exhibits significant advantages in computational complexity, real-time performance, processing time, and system requirements. This method utilizes Kinect sensors to obtain concise and clear three-dimensional coordinate data of human joints, with moderate computational complexity, and IoT technology makes data processing and transmission more efficient. In terms of real-time performance, this method performs excellently with a time delay of only 50ms, meeting the needs of most real-time applications, such as real-time monitoring and feedback of exercise training postures. In terms of processing time, the average processing time for a single action is 0.2 s, which is fast. The system requirements are moderate, requiring only Kinect sensors and moderately configured servers, reducing hardware and maintenance costs while improving system flexibility and scalability. Compared with vision based methods and wearable sensor based methods, the method proposed in this paper has advantages in computational complexity, real-time performance, processing time, and system requirements, and is suitable for most real-time application scenarios.

4. Conclusion

Based on the above experiments, it can be concluded that the sports training posture recognition method proposed in this paper, which is based on Kinect motion sensor and IoT technology, is feasible. In real-world sports training environments, the system is implemented through the deployment of Kinect body sensors, which are responsible for collecting exercise data from trainees. The hardware requirements include at least one Kinect sensor and a data processing unit that supports IoT technology (such as a medium-sized server or embedded system). This solution has good scalability and can easily add more sensors or data processing nodes to adapt to larger scale sports training scenarios. For different users or environments, the system needs to adjust the position and angle of the sensors and perform necessary software parameter settings for preliminary calibration to ensure that the Kinect sensor can accurately capture and recognize the user's motion training posture. Through the research in this paper, the recognition error is reduced, and the trainers' sports training posture can be analyzed and recognized according to the coordinates of the trainers' joint points and the angles of each joint, and then the results of sports training posture recognition are compared with the standard sports training posture to find out the corresponding problems, guide and correct the training posture pertinently, and comprehensively improve the quality of sports training.

Author contributions: Conceptualization, JR and HD; methodology, HS; formal analysis, JR; investigation, JR; data curation, LW; writing—original draft preparation, JR; writing—review and editing, JR. All authors have read and agreed to the published version of the manuscript.

Ethical approval: Not applicable.

Conflict of interest: The authors declare no conflict of interest.

References

1. Liu J, Tsujinaga S, Chai S, Sun H, Tateyama T, Iwamoto Y. Single Image Depth Map Estimation for Improving Posture Recognition. IEEE sensors journal 2021; 21(23):26997-27004.
2. Jiang H, Tsai S B. An Empirical Study on Sports Combination Training Action Recognition Based on SMO Algorithm Optimization Model and Artificial Intelligence. Mathematical Problems in Engineering 2021; 2021:1-11.
3. Li C, Cui J. Intelligent Sports Training System Based on Artificial Intelligence and Big Data. Mobile Information Systems 2021; 2021(1):1-11.
4. Wu M, Du D, Li Y, Bai W, Liu W. Multi-cascade Perceptual Human Posture Recognition Enhancement Network. IEEE Access 2021; 9: 64256-64266.
5. Xu Y. A Sports Training Video Classification Model Based on Deep Learning. Scientific Programming 2021; 2021(5):1-11.
6. Envelope M, Rlpdm B, Vfdlj B. Gesture recognition of wrist motion based on wearables sensors. Procedia Computer Science 2022; 210:181-188.
7. Li J, Ray S, Rajanna V, Hammond H. Evaluating the Performance of Machine Learning Algorithms in Gaze Gesture Recognition Systems. IEEE Access 2022; 10:1020-1035.
8. Xia Z, Xing J, Wang C, Li X. Gesture Recognition Algorithm of Human Motion Target Based on Deep Neural Network. Mobile Information Systems 2021; 2021(3):1-12.
9. Tu M. Gesture Detection and Recognition Based on Pyramid Frequency Feature Fusion Module and Multiscale Attention in Human-Computer Interaction. Mathematical Problems in Engineering 2021; 2021(7):1-10.
10. Fronteddu G, Porcu S, Floris A, Atzori L. A dynamic hand gesture recognition dataset for human-computer interfaces. Computer Networks 2022; 205:108781.
11. Yu Z, Zhao J, Wang Y, He L, Wang S. Surface EMG-Based Instantaneous Hand Gesture Recognition Using Convolutional Neural Network with the Transfer Learning Method. Sensors 2021; 21(7):2540.
12. Nallabolu P, Li Z, Hong H, Li C. Human Presence Sensing and Gesture Recognition for Smart Home Applications with Moving and Stationary Clutter Suppression Using a 60-GHz Digital Beamforming FMCW Radar. IEEE Access 2021; 9: 72857-72866.
13. Silva B, Melo D, Pombo N, Xu L. Internet of Things Assisted Monitoring based on a Ultrasound-based Gesture Recognition Contactless System. IEEE Access 2021; 9:90185-90194.
14. Peng Y, Wang J, Pang K, Liu L, Meng J, Li B. A Physiology-Based Flexible Strap Sensor for Gesture Recognition by Sensing Tendon Deformation. IEEE Sensors Journal 2021; 21(7):9449-9456.
15. Chmurski M, Zubert M, Bierzynski K, Santra A. Analysis of Edge-Optimized Deep Learning Classifiers for Radar-Based Gesture Recognition. IEEE Access 2021; 99:74406-74421.
16. Singh D K. 3D-CNN based Dynamic Gesture Recognition for Indian Sign Language Modeling. Procedia Computer Science 2021; 189:76-83.

17. Qiang B, Zhai Y, Zhou M, Yang X, Peng B, Wang Y. SqueezeNet and Fusion Network-Based Accurate Fast Fully Convolutional Network for Hand Detection and Gesture Recognition. IEEE Access 2021; 9:77661-77674.
18. Sun Y, Xiong H, Tan D K P, Han T X, Du R, Yang X. Moving Target Localization and Activity/Gesture Recognition for Indoor Radio Frequency Sensing Applications. IEEE sensors journal 2021; 21(21):24318-24326.
19. Ninos A, Hasch J, Alvarez M E P, Zwick T. Synthetic Radar Dataset Generator for Macro-Gesture Recognition. IEEE Access 2021; 9:76576-76584.
20. Zhang W, Wang J, Lan F. Dynamic Hand Gesture Recognition Based on Short-Term Sampling Neural Networks. IEEE/CAA Journal of Automatica Sinica 2021; 8(1):110-120.
21. Magrofuoco N, Roselli P, Vanderdonckt J. Two-dimensional Stroke Gesture Recognition: A Survey. ACM Computing Surveys 2021; 54(7):1-36.
22. Zeng X, Wu C, Ye W B. User-Definable Dynamic Hand Gesture Recognition Based on Doppler Radar and Few-Shot Learning. IEEE sensors journal 2021; 21(20):23224-23233.
23. Guo L, Lu Z, Yao L. Human-Machine Interaction Sensing Technology Based on Hand Gesture Recognition: A Review. IEEE transactions on human-machine systems 2021; 51(4):300-309.
24. Yu Z, Zhou B, Wan J, Wang P, Chen H, Liu X. Searching Multi-Rate and Multi-Modal Temporal Enhanced Networks for Gesture Recognition. IEEE Transactions on Image Processing 2021; 30:5626-5640.
25. Yang X, Zhu Y, Chen J. Intelligent Recognition of Multi Pose Face Images Based on Distance Constrained Optimization. Computer Simulation 2022;39(1):200-203.