Design and Application of Yoga Intelligent Teaching Platform Based on Internet of Things

Dong Fu and Jian Wang

1Department of Physical Education, Shangluo University, Shangluo, Shanxi 726000, China
2Shangluo International Medical Center Hospital, Shangluo, Shanxi 726000, China

Correspondence should be addressed to Dong Fu; 226009@slxy.edu.cn

Received 4 July 2022; Revised 11 August 2022; Accepted 16 August 2022; Published 5 September 2022

Copyright © 2022 Dong Fu and Jian Wang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the rapid development of economy, science, and technology, people’s pursuit of quality of life continues to improve, leisure sports have gradually become a trend, and yoga, as a traditional and fashionable way of fitness, is favored by more and more people. However, the traditional training class yoga teaching time is fixed, which will have many inconveniences for professionals. However, learning according to the teaching video cannot guarantee the accuracy of the movement and may not achieve the training effect. Therefore, in this situation, intelligent e-yoga teaching is of great significance to improve the training level. This system is based on Internet of things technology, selects Kinect as the sensing carrier of human motion information, and designs an electronic yoga teaching system. The system integrates the functions of motion information acquisition and motion evaluation. First, standard yoga movements are collected through Kinect as a comparison template for yoga training. Second, the system collects the trainers’ action data, uses Hausdorff distance algorithm to evaluate the similarity of action flow, and identifies the action name based on the threshold. Third, through the motion evaluation algorithm based on joint point angle measurement, it points out the joint points whose actions are not in place. Based on the completion time of the action, the action speed is evaluated. Finally, the system outputs the evaluation results in the form of text and voice through text conversion technology. The system can carry out efficient posture recognition and can achieve the purpose of evaluating the training quality and giving guiding suggestions. It can meet the basic training needs of users and has great application value.

1. Introduction

At present, the world is in a period of rapid economic development. When people enjoy the convenience brought by social development, they are also under increasing pressure. In the highly competitive labor market, people put a lot of time and energy into their work. The high-intensity work makes many professionals have no time to exercise, which also leads to the continuous decline of their physical quality. Professionals are in a subhealth state, with a surge of people suffering from headache, insomnia, endocrine disorders, depression, and other diseases. At the same time, the harm caused by subhealth problems cannot be ignored. As we all know, physical exercise can effectively relieve and release pressure and has a positive role in promoting physical and mental health. Among many fitness methods, yoga, as a fashionable exercise method that is both healthy and healthy, is gradually favored by many professionals and school students. Yoga is a combination of body posture and psychological awareness to achieve the purpose of self-cultivation. Unlike physical exercise, yoga is mainly used to strengthen the body. The investigation of yoga practitioners for many years shows that yoga exercise can make the mental state of practitioners become peaceful, alleviate the anxiety caused by work competition, and reduce the frustration in life. It is of great significance to adjust the psychological state and physiological balance. At the same time, yoga has the power of curing diseases; long-term practice of yoga can prevent chronic diseases such as hypertension, varicose veins, and arteriosclerosis and can enhance human
immunity. The medical profession has confirmed that yoga can effectively regulate the endocrine system and nervous system of the human body, thus helping to promote the physical and mental health of the human body.

In contemporary society, people tend to participate in yoga training institutions as the main way of learning. Although this is an effective way, the learning environment and learning progress are generally controlled by teachers, and the time is fixed, so it brings a lot of inconvenience [1]. At the same time, there are few professional formal yoga education institutions. Many so-called professional yoga coaches can obtain the qualification certificate after short-term training in some training institutions. It is difficult to guarantee the professionalism of these yoga coaches. Many people also choose to download teaching videos from the resource-rich network to learn by themselves. Although this approach is convenient and easy to operate, it cannot guarantee the accuracy of action and achieve the expected effect without guidance. Second, improper yoga practice will cause great damage to people’s body. According to the statistics of sports injury cases in the hospital, the proportion of lumbar disc protrusion and spinal joint dislocation caused by improper exercise has increased significantly, and in these cases, the injury caused by yoga practice accounts for a large proportion. Therefore, blind practice without professional guidance will not only fail to achieve the effect of fitness but will bring damage and disease to the body in all aspects.

The significance of this research is to apply Kinect to yoga teaching in combination with Internet of things technology. As a new computer software development carrier, Kinect not only has a high-resolution camera but also can collect user action information in real time and support voice input. This system uses this device to capture the action data flow of students and professionals. After matching and comparison, the system can give the matching degree of user actions and standard actions in real time and remind users to make targeted improvements. This computer-aided instruction system can make yoga lovers who have a strong demand for self-study adjustment in time according to the system prompts to ensure the accuracy of their movements. At the same time, this teaching platform is cheap and universally applicable, which solves the problems of time and space in yoga learning.

This study uses the mainstream Internet of things technology to design a yoga teaching system based on Kinect, which mainly includes the following research contents: building an electronic yoga teaching system based on Kinect. The system can automatically capture the yoga movements of trainers and can use Hausdorff algorithm to give the similarity of movements by comparing them with the standard movements in the template database. Motion evaluation is based on joint point angle. By measuring the angle between skeletons in real time and comparing it with the angle between skeletons of template action, we can more intuitively see the unqualified joint points.

2. Related Work

2.1. Human-Computer Interaction. Human-computer interaction (HCI) is a science used to study the interaction between computer systems and users. It has a close relationship with artificial intelligence, virtual reality, ergonomics, and other technical fields [2–7]. The design of human-computer interaction system is required to be simple and easy to operate with intuitive visualization function. Nowadays, human-computer interaction technology has been one of the key research contents in the computer field, and it is also the focus of competition in the field of science and technology.

As an important part of the computer field, human-computer interaction technology has developed from the initial human to adapt to the computer to the current computer to adapt to people. This process has gone through more than half a century and has also reaped great progress and achievements. Natural, efficient, and ubiquitous human-computer interaction technology is the intersection of research directions in many fields, such as intelligent science and technology, computer, mathematics, psychology, and even physiology [8–13]. It is one of the focuses of computer research at present and in the future, and it is also the development trend of information field in the future.

The human-computer interaction mainly goes through the following five stages: (1) Manual Operation Stage. The characteristic of the embryonic period of human-computer interaction is that the computer designers debug and use very clumsy large-scale computers through manual manipulation and binary code. (2) Language Command Interaction Phase. At this stage, programmers began to use interactive command language to manipulate the operation of the computer. There are many commands to remember at this stage. At that time, compared with the manual operation stage, the operation of the debugging program has been greatly simplified. (3) Graphical User Interface (GUI) Phase. The main features of GUI are desktop metaphor, windows, icons, menus, pointing devices, and other technologies such as direct manipulation and “WYSIWYG.” At this stage, the interface design is very simple, so even users who do not know much about the computer can operate skillfully according to the graphical instructions. At this stage, information technology has made unprecedented development. (4) Network User Interface Phase. This stage is represented by web browsers based on Hypertext Markup Language (HTML) and Hypertext Transfer Protocol (HTTP). At this stage, emerging technologies such as search engine, animation, and chat tools began to emerge. (5) Multichannel Intelligent Human-Computer Interaction Stage. The main achievements of this stage are the personification of computer system and the miniaturization of computer, for example, virtual reality technology, smart phones, and laptops. Multichannel interaction allows imprecise input, such as using human language, gestures, and actions to interact with the computer, which breaks away from the shackles of traditional human-computer interaction and makes human-computer interaction efficient and natural. The multichannel interaction technology has
developed rapidly in recent years. It not only meets the "people-centered" interaction criteria but also promotes the development of the information industry. The channels in multichannel interaction mainly include user action, expression intention, and perceptual information feedback, such as language, expression, human posture, smell, touch, and taste. At present, this technology has been successfully used in speech recognition, gesture recognition, pen interaction, digital ink, and so on.

With the development of human-computer interaction, the mode of human-computer interaction is constantly updated from the initial keyboard and mouse to touch technology, the emerging somatosensory devices, and virtual reality technology in recent years. The mode of human-computer interaction is more and more close to people's ideas. The development trend of human-computer interaction is as follows: (1) The Operation is More Free. It gets rid of the shackles of contact equipment, and the input and output methods are gradually dominated by voice, camera, and projection contactless input equipment, and human-computer interaction can be realized in any scene, which is no longer limited to indoor. (2) More Intelligent Interaction. The traditional human-computer interaction system requires users to transfer interaction information through contact sensors. Now, human-computer interaction can be realized through gesture recognition, speech recognition, and other technologies, realizing the naturalness of human-computer interaction. (3) The Operation is More Humanized. The traditional human-computer interaction technology requires people to adapt to the computer, while the current human-computer interaction technology is based on people's ideas, making full use of people's actions, language, and vision, and combining the characteristics of the computer to achieve a natural, efficient, and multichannel human-computer interaction.

At present, with the advent of Microsoft Kinect, somatosensory interaction technology has become one of the most popular research directions in the field of human-computer interaction. Kinect gets rid of the traditional handheld remote control. It can achieve the purpose of human-computer interaction by capturing the user's actions, facial expressions, and voice sequences in real time. The emergence of Kinect has played an important role in promoting the development of contactless human-computer interaction. The research on its application in the field of human-computer interaction is also one of the key research topics in the future. To sum up, the development direction of human-computer interaction technology is to focus on natural and contactless interaction and realize real-time and natural interaction through user's actions, expressions, languages, etc.

2.2. Research Status of Posture Recognition. Human posture recognition is accomplished by many technical fields such as machine learning, sensor technology, artificial intelligence, and computer vision. According to the different ways of obtaining human posture information, the human posture recognition system is divided into wearable sensor-based and vision-based recognition systems.

Pose recognition system based on wearable sensors generally refers to embedding microsensors in some carry-on items or close-fitting clothes. This type of sensor can collect human motion data in real time, and the system can analyze and process the motion data in real time. The data can also be sent to the host through wireless network transmission, and the host can set a specific algorithm to analyze the motion data and identify the human posture. There are two kinds of embedded sensors: three-axis gyroscope and acceleration sensor. The gyroscope can detect the angular motion state. This kind of sensor mainly relies on the precession and axial orientation of gyroscope to measure the pose angle formed during human motion. Tapia and Haskell [14] proposed a method for real-time automatic recognition of human posture using wireless acceleration sensors and wireless heart rate monitors. The implementation process of this method is to fix the wireless acceleration sensor at the joints of limbs and hips respectively and then obtain the standard deviation, FFT kurtosis, and other features as processing signals. The collected actions are divided into three groups by dynamic Bayesian classification method, and the recognition rate of the three groups is 94.6%. Zhang et al. [15] used wearable acceleration sensors to capture human motion data and carried out fall detection based on support vector machine algorithm. After testing, the success rate of fall detection reached 96.7%. Bourke and Lyons [16] proposed a fall detection algorithm based on the threshold value of the dual-axis gyroscope sensor. The gyroscope sensor is installed on the experimenter's body to measure the angular velocity information during the action. After many times of experimental data statistics, the threshold value for distinguishing falls from other daily activities is obtained. The success rate of judging falls based on the threshold value is as high as 100%. The working principle of wearable sensor is to capture the data of human motion with the help of sensors placed at the joints of human body and process the data through analysis. The change rules of human body data in different postures are summarized, and then a specific algorithm is used to achieve the purpose of pose recognition. Although this kind of posture recognition system has high data analysis and recognition ability, this kind of contact sensor can only be placed on the body to collect human motion parameters, which brings great inconvenience to people. Compared with the contact pose recognition system based on wearable sensors, it is obvious that the noncontact recognition system has more universal applicability. This kind of recognition system is mainly based on computer vision technology to obtain human motion information from video sequences for pose recognition. Computer vision technology is actually the simulation of human visual function. With the rapid development of this technology, vision-based pose recognition will become the mainstream direction. The researchers use the pose recognition method of hidden Markov model. First, the cascaded hidden Markov model is used to model the human motion. Its special feature is that the expectation maximization algorithm is used to ensure the reliability and accuracy of the
model. At the same time, the Rao blackwelled particle filter approximate reasoning algorithm with high computational efficiency is also used to analyze human motion, which can effectively recognize human posture. By using the method of inductive reasoning, Granum and Moeslund [17] summarized human motion recognition into four processes: system initialization, bone tracking, pose estimation, and pose recognition. Moselund et al. summarized the research progress of recognition technology based on computer vision on the four processes of pose recognition.

As the most representative product in the field of computer vision, Kinect has made outstanding contributions to the field of computer vision. In particular, Kinect’s function of acquiring 3D depth images of human body in real time makes it a popular direction for researchers to analyze human posture based on depth information. At the same time, the advent of Kinect has also promoted the development of human posture recognition technology. Its accuracy and real-time performance are beyond the traditional recognition methods. At present, researchers have begun to apply Kinect to the research of gesture and human posture recognition.

2.2.1. Gesture Recognition. Kinect is used to obtain the depth image information of the hand, and the depth vector segmentation technology is used to detect the fingertip and palm center. The sign language recognition of deaf mutes based on Kinect was proposed [18]. The human hand movement information was obtained through Kinect and analyzed and compared with the movement data in the standard library, and the recognition results were displayed in the evaluation interface. In the research of hand gesture recognition based on depth image information, the region of hand gesture is extracted from the depth image by using the difference of gray value, and the classification decision tree is established by combining the contour characteristics to realize hand gesture recognition.

2.2.2. Gesture Recognition. A vision-based tracking system [19], which obtains the patient’s lower limb movement information by installing Kinect on the wheelchair, analyzes the patient’s condition by using the obtained movement data and customizes the training plan for the patient. Shotton et al. obtained the depth image information of the human body through Kinect [20, 21], classified various parts of the human body using the random forest algorithm, and then calculated various joint points of the human body. Schwarz et al. [22] used Kinect to obtain the depth data of human motion for pose estimation and used geodesic distance and depth data of optical flow to track each joint point of human body. In recent years, Kinect has also been widely used in the field of medical rehabilitation to assist patients with limb disorders in rehabilitation training. Da Gama et al. [23] developed a natural interaction system based on Kinect for sports rehabilitation training, developed a scoring mechanism to measure patient performance, set different training parameters for different patients, and prevented patients from secondary injury during rehabilitation training. Bonnechère [24] developed different game systems based on Kinect to assist patients in rehabilitation training and improve their enthusiasm for training. Experiments show that the Kinect-based rehabilitation training system can not only reduce patients’ negative emotions during training but also exercise patients’ psychological quality.

At present, there is relatively little research on Kinect in yoga teaching. Only reference [25] evaluated the accuracy of dance movements based on Kinect’s bone gap tracking technology. From the above analysis, it can be seen that Kinect can be used not only for local research such as human part recognition and gesture recognition but also for overall research such as human posture recognition. In terms of research methods, different recognition tasks are based on different data types. Gesture recognition mainly relies on depth image data to segment human hand parts from the depth image obtained by Kinect and recognize gesture names. Gesture recognition is more inclined to use the human bone node data provided by Kinect for comparison and analysis with template data.

3. Action Flow Evaluation Algorithm

In order to analyze and evaluate the action of trainers when the system is offline, the system needs to be able to record the coordinate information of bone nodes captured by Kinect in real time. Kinect, a somatosensory device, can capture up to 30 frames of data per second. Assuming that the time of an action is 3 seconds, the system will record up to 90 frames of data, and each frame is a collection of three-dimensional coordinate values of 25 joint points. Taking the 90-frame 3D coordinate point set thus formed as a set of sample data, after matching it with the actions in the standard library, the distance between the two sets of point sets can be calculated. The size of this distance is used as the evaluation standard of training. In this study, the Hausdorff distance measurement algorithm is used to evaluate the matching between the sample data point set and the template data point set.

3.1. Data Preprocessing. The yoga teaching system designed in this study mainly realizes the function of collecting and evaluating the exercise information of trainers. After experimental tests, Kinect will capture some data information with little correlation, namely, interference points, at the beginning of each motion capture, which will reduce the accuracy of the evaluation results. Therefore, it is necessary to regularize the training data before evaluation and intercept the middle data segment of the action data stream as a new action point set to participate in the final matching process. In addition, after the motion capture starts, the captured data points are continuous.

The motion comparison of this system is based on the three-dimensional coordinate information of 25 joint points of the human skeleton model obtained in the process of motion capture. However, differences in body shape or distance from Kinect among different practitioners will lead to great differences in bone data information obtained each time. Therefore, the data obtained in the process of motion
capture cannot be directly used for motion matching, and the obtained bone data model must be normalized first.

3.2. Hausdorff Distance Algorithm. Hausdorff distance is a similarity measure of two given data point sets. This algorithm does not need to establish point-to-point correspondence. It only needs to calculate the maximum distance between two point sets, which can effectively deal with the matching problem between point sets with more feature points. The obtained bone data model must be normalized first. Capture cannot be directly used for motion matching, and the algorithm does not need to establish point-to-point correspondence. It only needs to calculate the maximum distance between two point sets, which can effectively deal with the correspondence. It only needs to calculate the maximum distance between two point sets, which can effectively deal with the matching problem between point sets with more feature points.

The traditional mathematical description of Hausdorff distance is as follows: two point sets \( a \) and \( B \) are given. Then, the Hausdorff distance between \( a \) and \( B \) is calculated as follows:

\[
H(A, B) = \max \{ h(A, B), h(B, A) \},
\]

\[
h(A, B) = \max_{a \in A} \left( \min_{b \in B} \| a - b \| \right),
\]

\[
h(B, A) = \max_{b \in B} \left( \min_{a \in A} \| b - a \| \right).
\]

In the formula \( \| \cdot \| \) represents the distance norm. The calculation process of \( H(a, b) \) is to calculate the minimum distance from each point \( a \) in point set \( a \) to point set \( B \) and sort all the minimum distances obtained. The maximum value in the sequence represents the one-way Hausdorff distance from point set \( a \) to \( B \). Similarly, the one-way Hausdorff distance \( h(B, a) \) from point set \( B \) to point set \( a \) can be obtained. The bidirectional Hausdorff distance \( h(a, b) \) is the larger of \( H(a, b) \) and \( H(b, a) \); that is, it represents the distance between two sets. According to the size of \( H(a, b) \), the matching degree of two-point sets can be determined. The smaller the Hausdorff distance of two-point sets, the higher the similarity between the two sets of point sets. Therefore, this algorithm can be used to measure the similarity between the template action and the trainer action.

Although the traditional Hausdorff distance algorithm simplifies the process of calculating the matching degree between point sets, the motion matching problem involved in this system is based on the three-dimensional coordinate points captured by Kinect. It is inevitable that frame skipping occurs in the process of motion data acquisition by body sensors, and the collected point sets may be mixed with noise points. If the traditional Hausdorff distance algorithm is used, it may cause large errors, so the improved Hausdorff distance algorithm is explored and verified in this design.

Mean Hausdorff distance (MHD) can further improve the matching accuracy. Similar to the traditional Hausdorff distance and partial Hausdorff distance calculation methods, the mean Hausdorff distance will still calculate the larger one-way Hausdorff distance between two point sets as the final Hausdorff distance, but the calculation form of one-way Hausdorff distance will be different:

\[
H_{MHD}(A, B) = \max(h_{MHD}(A, B), h_{MHD}(B, A)).
\]

\[
\begin{align*}
h_{MHD}(A, B) &= -\frac{1}{N_A} \sum_{a \in A} -\min_{b \in B} \|a - b\|, \\
h_{MHD}(B, A) &= -\frac{1}{N_B} \sum_{b \in B} -\min_{a \in A} \|b - a\|.
\end{align*}
\]

To sum up, the one-way Hausdorff distance between two point sets is calculated by the average distance from all points in one point set to another point set. This calculation form fully considers the negative impact of interference points, makes all feature points in the set participate in the calculation of the final result, has a certain average effect, and ensures the anti-interference performance of the system. The mean Hausdorff distance algorithm solves the problem of mixing noise points in the process of action acquisition to some extent. Therefore, this algorithm is selected as one of the algorithms to evaluate the matching degree between the template point set and the training point set in this design.

4. Intelligent Yoga Teaching System Based on Internet of Things

The designed yoga teaching system aims to provide a simple, easy to operate, and reliable family teaching service without time and space constraints for the majority of yoga lovers. The system can make an objective evaluation of the training effect, give users more intuitive guidance in the form of voice, and make it more suitable for the training process in the natural environment.

4.1. System Design Principles

4.1.1. Reliability. As the system is a real-time human-computer interaction system, it is the primary condition to maintain stable and reliable operation. Reliability here mainly includes two meanings: first, the system can identify the user’s posture as accurately as possible and give a timely and correct response. Because when different users complete the same action, there will be differences in duration and action amplitude. Moreover, even if the same user performs the same action in different time periods, the action data flow is different. Therefore, a good recognition system can recognize the correct name of the action when the difference between the duration and the action amplitude is small. Second, the system can correctly understand the user’s intention as much as possible and avoid identifying some unconscious operations of the user.

4.1.2. Ease of Use. The teaching system designed in this study is for yoga practitioners in various fields and ages. Therefore, the teaching platform should have a simple and easy-to-use interface and conform to the operating habits of general users, so that users can adapt independently. At the same time, each action in the template library should be well distinguished from other actions to reduce the false
recognition rate, and the platform has an action evaluation module so that users can timely understand their own training.

4.1.3. Scalability. When designing an interactive system, it is necessary to consider the subsequent scalability of the system, so as to expand the template library and provide more complex functions to cope with changes in requirements.

4.2. Overall System Architecture. The hardware of the system is composed of a Kinect and a Windows8.1 computer. Kinect connects to the computer through USB3.0. Kinect is responsible for collecting human motion data stream, and then the computer is responsible for data analysis and processing. The software of this system is developed based on MFC architecture under the integrated development environment of Visual Studio 2013. The system uses kinect2.0 sensor to monitor the surrounding environment and converts the collected depth image into bone data. After the PC control end starts to work, it receives the bone data transmitted by kinect2.0 sensor through USB3.0 by means of “pulling” and draws it to the picture control of the dialog interface based on MFC architecture through the two-dimensional graphic API-direct2d for display. The system first normalizes the collected s-dimensional bone data of coaches and students’ actions and then compares and analyzes them using the Hausdorff algorithm to obtain the similarity of the two groups of action sequences. At the same time, the system will calculate the angle of some joint points in real time. According to the similarity and the angle difference of joint points, students can further improve their actions. The overall scheme design of the system is shown in Figure 1.

4.3. System Function. The designed system contains several functions, as listed in Table 1.

4.3.1. Software Interface Design. This article uses SQL server 2013 database engine to store the login information of management trainers. Trainers need to create new personal information when using the system for the first time. During registration, the user name and password entered by the user are written into the created database table through ADO function. When the user logs in, the sales interest entered by the user is compared with the information in the database table. If there is no error, the user logs in successfully and enters the module selection interface. The module selection interface of the system includes the basic classification of yoga postures, as well as the operations of entering the system, exiting, and logging in. Because there are many yoga postures, yoga can be further divided into standing posture, sitting posture, kneeling posture, lying posture, and hand support according to different postures. When users train, they only need to select the category according to the training posture, enter the system, and only compare it with the posture library. After classification, they can improve the overall recognition rate and work efficiency of the T system and reduce the pressure of system action analysis. Users need to click to enter the system, and the main interface window will pop up. The main interface is realized by MFC nonmodal dialog box, which controls the startup of the system and realizes the function of human-computer interaction. In the test interface, the largest control displays the human bone image detected by Kinect.

4.3.2. Data Acquisition Module. The system obtains bone data, color data, and depth data through the relevant APIs in the Kinect SDK. Their acquisition is encapsulated into classes for the convenience of calling. All data acquisition is completed by the same Kinect. During encapsulation, the Kinect sensor class will provide some interfaces to manage the acquisition of three types of Kinect data and the switch of the FT body, because if a class object releases resources when the dump is over, the use of other classes will be invalid.

4.3.3. Module Library Creation Module. In order to evaluate the standard degree of user actions, the system must
establish a library containing standard actions as an evaluation template. The design standard library is established by inviting professional yoga instructors to demonstrate yoga movements, using Kinect to collect yoga movement data, saving the data in txt files, and naming them with the corresponding action names, so as to facilitate comparison and call.

4.3.4. Direct2d Drawing Module. The system draws the joint point information collected by Kinect into the picture control through the two-dimensional graphics API-direct2d. Because Kinect collects h-dimensional coordinate information, it needs to use a coordinate mapper to convert 3D coordinates into 2D coordinates. Direct2d enriches the visual effect of Windows applications and enables 2D graphics to be presented with high quality and performance.

When using direct2d to draw bone data into the dialog box, the first task is to create an id2dlfactory interface object. The interface is the root object of direct2d and the root of all other resources. After the interface is successfully created, direct2d resources can be created. Basically, all D2D resources are created by the id2dlfactory interface. Second, the system needs a render site rendertarget. Here, id2dlhwrendertarget is created and then used to create renderer targets and perform rendering operations.

4.3.5. Action Recognition Module. Motion recognition first extracts the bone data stream of the trainer’s action through Kinect and then calculates the matching distance between the trainer’s action sequence and the action sequence in the standard library. The distance value is saved in the array until the matching distance with all actions in the standard library is calculated, and then all matching distances are compared with the preset threshold. Therefore, in the process of matching two groups of action sequences, there must be a strict threshold as the evaluation standard. For the setting of this threshold, we collect a series of algorithms are used to comprehensively evaluate the standard degree of their actions.

In this study, first, the principle of motion flow capture using Kinect is described. The somatosensory device has the advantages of noninvasive, not restricting human movement, and the capture accuracy can meet the basic requirements. In the analysis and evaluation of motion flow information, this study uses Hausdorff measure algorithm. The system also evaluates the movement with the help of the joint point angle. This method can more intuitively find the joint points where the movement is not in place. At the same time, this design realizes the voice prompt function of the system through the text-to-voice conversion technology. The biggest advantage of this system is that it will neither be affected by time and space, nor will it involve the privacy of users. It can work for a long time. In the future, the recognition effect of the algorithm can be further improved to further enhance the classification performance.

5. Conclusion

The traditional yoga learning can only be completed under the guidance of a fixed time, environment, and coach, which cannot meet the requirements of some professionals. Under this background, this study develops an electronic yoga teaching system based on Kinect, which can capture the movement information of trainers through Kinect. Then, a series of algorithms are used to comprehensively evaluate the trainers’ actions, so that the trainers can timely understand the standard degree of their actions.

In this study, first, the principle of motion flow capture using Kinect is described. The somatosensory device has the advantages of noninvasive, not restricting human movement, and the capture accuracy can meet the basic requirements. In the analysis and evaluation of motion flow information, this study uses Hausdorff measure algorithm. The system also evaluates the movement with the help of the joint point angle. This method can more intuitively find the joint points where the movement is not in place. At the same time, this design realizes the voice prompt function of the system through the text-to-voice conversion technology. The biggest advantage of this system is that it will neither be affected by time and space, nor will it involve the privacy of users. It can work for a long time. In the future, the recognition effect of the algorithm can be further improved to further enhance the classification performance.

Data Availability

The dataset can be obtained from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors thank the Path of College Sports Serving Community under the Trend of Life Study Shaanxi Sports Bureau Study Projects (no. 2021320), Research on the Construction and Governance Mechanism of Rural Sports Ecosystem under the Background of Rural Revitalization Strategy, Science and Technology Department of Shaanxi Province, Science and Technology Plan Project (no. 2022KRM140) in research, and Shaanxi Science and Technology Innovation Team Science and Technology Plan Project (no. 2022TD-56).
References

[1] A. Ross and S. Thomas, “The health benefits of yoga and exercise: a review of comparison studies,” Journal of Alternative & Complementary Medicine, vol. 16, no. 1, pp. 3–12, 2010.

[2] J. Long and J. Dowell, “Conceptions of the discipline of HCI: craft, applied science, and engineering,” HCI, vol. 89, no. 747, pp. 9–32, 1989.

[3] Z. Jia, Y. Lin, X. Cai, H. Chen, H. Gou, and J. Wang, “SStemotionnet: spatial-spectral-temporal based attention 3D dense network for eeg emotion recognition,” in Proceedings of the 28th ACM International Conference on Multimedia, pp. 2909–2917, Seattle, WA, USA, October 2020.

[4] J. M. Carroll, HCI Models, Theories, and Frameworks: Toward a Multidisciplinary science, Elsevier, Amsterdam, Netherlands, 2003.

[5] J. Zimmerman, J. Forlizzi, and S. Evenson, “Research through Design as a Method for Interaction Design Research in HCI,” in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 493–502, San Jose, California, USA, April 2007.

[6] Z. Jia, Y. Lin, J. Wang, Z. Feng, X. Xie, and C. Chen, “HetEmotionNet: two-stream heterogeneous graph recurrent neural network for multi-modal emotion recognition,” in Proceedings of the 29th ACM International Conference on Multimedia, pp. 1047–1056, China, October 2021.

[7] K. Kuutti and L. J. Bannon, “The Turn to Practice in HCI: Towards a Research agenda,” in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 3543–3552, Toronto, Ontario, Canada, April 2014.

[8] Z. Jia, Y. Lin, J. Wang, K. Yang, T. Liu, and X. Zhang, “MCCNN: A Multi-branch Multi-Scale Convolutional Neural Network for Motor Imagery classification,” in Proceedings of the Joint European Conference on Machine Learning and Knowledge Discovery in Databases, pp. 736–751, Springer, Cham, 2020.

[9] Z. Li, J. Wang, Z. Jia, and Y. Lin, “Learning Space-Time-Frequency Representation with Two-Stream Attention Based 3D Network for Motor Imagery classification,” in Proceedings of the 2020 IEEE International Conference on Data Mining (ICDM), pp. 1124–1129, IEEE, Sorrento, Italy, November 2020.

[10] Z. Jia, Y. Lin, J. Wang et al., “Multi-view spatial-temporal graph convolutional networks with domain generalization for sleep stage classification,” IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 29, pp. 1977–1986, 2021.

[11] W. Xu, J. Wang, Z. Jia, Z. Hong, Y. Li, and Y. Lin, “Multi-level spatial-temporal adaption network for motor imagery classification,” in Proceedings of the ICASSP 2022-2022 IEEE international conference on acoustics, speech and signal processing (ICASSP), pp. 1251–1255, IEEE, Singapore, May 2022.

[12] Z. Jia, J. Junyu, X. Zhou, and Y. Zhou, “Hybrid Spiking Neural Network for Sleep EEG Encoding,” Science China Information Sciences, vol. 65, no. 4, pp. 1–10, 2022.

[13] Z. Jia, X. Cai, and Z. Jiao, “Multi-modal physiological signals based squeeze-and-excitation network with domain adversarial learning for sleep staging,” IEEE Sensors Journal, vol. 22, no. 4, pp. 3464–3471, 2022.

[14] E. M. Tapia, S. S. Intille, W. Haskell et al., “Real-time Recognition of Physical Activities and Their Intensities Using Wireless Accelerometers and a Heart Rate monitor,” in Proceedings of the 2007 11th IEEE International Symposium on Wearable Computers, pp. 37–40, IEEE, Boston, MA, USA, October 2007.

[15] T. Zhang, J. Wang, L. Xu, and P. Liu, “Fall detection by wearable sensor and one-class SVM algorithm,” Intelligent Computing in Signal Processing and Pattern Recognition, pp. 858–863, Springer, Berlin, Heidelberg, 2006.

[16] A. K. Bourke and G. M. Lyons, “A threshold-based fall-detection algorithm using a bi-axial gyroscope sensor,” Medical Engineering & Physics, vol. 30, no. 1, pp. 84–90, 2008.

[17] T. B. Moeslund and E. Granum, “A survey of computer vision-based human motion capture,” Computer Vision and Image Understanding, vol. 81, no. 3, pp. 231–268, 2001.

[18] K. F. Li, K. Lotthrop, E. Gill, and S. Lau, “A Web-Based Sign Language Translator Using 3D Video processing,” in Proceedings of the 2011 14th International Conference on Network-Based Information Systems, pp. 356–361, IEEE, Tirana, Albania, September 2011.

[19] R. Z. L. Hu, A. Hartfiel, J. Tung, A. Fakih, J. Hoey, and P. Poupart, “3D Pose Tracking of walker Users’ Lower Limb with a Structured-Light Camera on a Moving platform,” in Proceedings of the CVPR 2011 WORKSHOPS, pp. 29–36, IEEE, Colorado Springs, CO, USA, June 2011.

[20] J. Shotton, R. Girshick, A. Fitzgibbon et al., “Efficient human pose estimation from single depth images,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 35, no. 12, pp. 2821–2840, 2013.

[21] J. Shotton, A. Fitzgibbon, M. Cook et al., “Real-time human pose recognition in parts from single depth images,” in Proceedings of the CVPR 2011, pp. 1297–1304, IEEE, Colorado Springs, CO, USA, June 2011.

[22] L. A. Schwarz, A. Mkhitaryan, D. Mateus, and N. Navab, “Human skeleton tracking from depth data using geodesic distances and optical flow,” Image and Vision Computing, vol. 30, no. 3, pp. 217–226, 2012.

[23] A. Da Gama, T. Chaves, L. Figueiredo, and V. Teichrieb, “Poster: Improving Motor Rehabilitation Process through a Natural Interaction Based System Using Kinect sensor,” in Proceedings of the 2012 IEEE Symposium on 3D User Interfaces (3DUI), pp. 145-146, IEEE, Costa Mesa, CA, USA, March 2012.

[24] B. Bonnechère, “Serious games in physical rehabilitation,” DOI, vol. 10, pp. 978–983, 2018.

[25] D. S. Alexiadis, P. Kelly, P. Daras, N. E. O’Connor, T. Boubakeur, and M. B. Moussa, “Evaluating a dancer’s performance using Kinect-based skeleton tracking,” in Proceedings of the 19th ACM International Conference on Multimedia, pp. 659–662, Scottsdale, Arizona, USA, November 2011.