WILEY | Hindawi

Research Article

A New Kinect-Based Posture Recognition Method in Physical Sports Training Based on Urban Data

Dianchen He[1,2] **and Li Li** [1]

[1]*College of Sports Science, Woosuk University, Jeonju, Republic of Korea*
[2]*College of Sports Science, Shenyang Normal University, Shenyang, China*

Correspondence should be addressed to Li Li; 48241194@qq.com

Physical data is an important aspect of urban data, which provides a guarantee for the healthy development of smart cities. Students' physical health evaluation is an important part of school physical education, and postural recognition plays a significant role in physical sports. Traditional posture recognition methods are with low accuracy and high error rate due to the influence of environmental factors. Therefore, we propose a new Kinect-based posture recognition method in a physical sports training system based on urban data. First, Kinect is used to obtain the spatial coordinates of human body joints. Then, the angle is calculated by the two-point method and the body posture library is defined. Finally, angle matching with posture library is used to analyze posture recognition. We adopt this method to automatically test the effect of physical sports training, and it can be applied to the pull-up of students' sports. The position of the crossbar is determined according to the depth sensor information, and the position of the mandible is determined by using bone tracking. The bending degree of the arm is determined through the three key joints of the arm. The distance from the jaw to the bar and the length of the arm are used to score and count the movements. Meanwhile, the user can adjust his position by playing back the action video and scoring, so as to achieve a better training effect.

1. Introduction

Urban big data is a massive amount of dynamic and static data generated from the subjects and objects including various urban facilities, organizations, and individuals, which have been collected and collated by city governments, public institutions, enterprises, and individuals using a new generation information technologies. Big data can be shared, integrated, analyzed, and mined to give people a deeper understanding of the status of urban operations and help them make more informed decisions on urban administration with a more scientific approach, thereby, optimizing the allocation of urban resources, reducing the operating costs of the urban system, and promoting the safe, efficient, green, harmonious, and intelligent development of the cities as a whole.

Nowadays, the sports hardware facilities are also perfect with the development of the smart cities [1–3]. The quality of people's life has been guaranteed and improved. The human body has a rich and variety of movements [4]. In many applications, a more comprehensive analysis for human movement is needed, such as behavior monitoring, movement analysis, and medical rehabilitation. If the human body can be identified and tracked in real time, then the posture of the human body can be identified accurately, this process can make it more convenient to observe and learn human behavior [5]. Therefore, it is necessary to find a good way to recognize human posture.

In recent years, human motion recognition based on Kinect has shown great significance in the field of medicine, and many institutions are carrying out relevant researches. Enea et al. [6] installed a Kinect on a walker to extract information about the legs for medical gait analysis. Chen et al. [7] used depth information extracted from Kinect to detect the body's joints and used a random forest classifier to classify depth image pixels into multiple parts of the body. Thang

et al. [8] adopted human anatomy marks and the human body skeleton model to obtain a depth map and estimate the posture of the human body. The geodesic distance was used to measure the distance between the different parts of the body. Xu et al. [9] used the Kinect sensor to obtain human body images and identify 3D human body posture. Yang et al. [10] utilized a Kinect device to capture the scene and estimate the body's limb posture. Moreover, some other researchers had done the related works to improve the posture recognition [11–13]. These methods only take physical characteristics into consideration; it ignores the global features, which shows a poor effect on physical sports training. Although deep learning methods have been applied to many fields, it is not mature in physical training and rehabilitation. In this paper, we focus on the study of improving the Kinect technology for posture recognition.

Many medical experts have brought Kinect to medical rehabilitation because it is cheap and useful such as using Kinect for rehabilitation treatment. The basic idea is to use depth information and skeleton tracking technology to track the body limb and determine the position of the body. At the end, it can identify the movement of the body. In reference [14], Kinect rehabilitation training could effectively enhance the quality of rehabilitation. It not only assisted patients to recover motor function but also improved their psychological quality and reduced their negative emotions. Wang et al. [15] designed a rehabilitation system for the human shoulder that required the patient to touch his or her hand to a set point. However, it was impossible to measure the specific location of the joint in real time. Rehabilitation training usually does not require patients to carry out rapid and large movements, but requires high tracking accuracy of Kinect's human skeleton. If accurate tracking of the human hand and leg skeleton can be observed, then, the movements of patients can be identified more accurately, so as to achieve better recovery effect.

Our main contributions are as follows:

(a) We propose a new Kinect-based posture recognition method in a physical sports training system based on urban data

(b) First, Kinect is used to obtain the spatial coordinates of the human body joints

(c) Then, the angle is calculated by the two-point method and the body posture library is defined

(d) Finally, angle matching with posture library is used to analyze posture recognition

(e) We adopt this method to automatically test the effect of physical sports training and it can be applied to the pull-up of students' sports. This method can measure the angle between the skeleton in real time, improve the accuracy of posture matching, and can accurately identify the human posture. The new algorithm is simple with high efficiency

The remainder of the paper is organized as follows. In Section 2, we give an outline of the Kinect imaging principle. Section 3 describes the posture recognition method in detail.

The performance and robustness are evaluated in Section 4. The conclusion is drawn in Section 5.

2. Kinect Imaging Principle

Kinect emits a near-infrared light to get a depth map. Kinect actively tracks large objects regardless of the amount of light. The depth image of Kinect is captured by an infrared projector and camera. The projection and reception are overlapped. There are similar processes for transmitting, capturing, and computing visual representations.

Structural light is with specific patterns such as points, lines, and surfaces. The principle of depth image acquisition based on structured light is to project the structured light into the scene and the image sensor to obtain the pattern corresponding to the structured light [16]. Because the structure light will be changed according to the shape of the object and the depth information of each point in the image can be calculated by using the triangle principle and the obtained pattern.

The Microsoft Kinect depth concept uses the light coding technology, which is different from the traditional two-dimensional pattern projection method of structured light. Kinect's light-coding infrared transmitter emits three-dimensional depth coding. Laser speckle is the source of light coding and the result of diffuse reflection of the laser. That is to form random spots, and the spots are not the same anywhere in the space. So every time the light source is labeled and all random spots are saved. When an object is placed in this area, if a random speckle of the object's appearance is obtained, the object's position can be found. Thus, it can obtain a depth image of the scene.

3. Coordinate Transformation

Camera space refers to the 3D spatial coordinates used by the Kinect. The origin coordinate ($x = 0$, $y = 0$, and $z = 0$) is in the center of the Kinect's infrared camera. The x-axis is to the left of the Kinect irradiation direction. The y-axis is the upward direction along the Kinect irradiation direction. The z-axis is along the Kinect irradiation direction.

The coordinate system of the depth image is the origin of an infrared camera. The positive x-axis and y-axis directions are horizontal to the right and vertical downward, respectively. The z-axis is the spatial coordinate system of the camera axis direction and meets the right-hand spiral criterion, and its type is *DepthImagePoint* (x, y, and z). The point type of the bone tracking coordinate system is *SkeletonPoint (x, y,* and z). Where (x, y, and z) is the spatial coordinate system. Only the x and y values are used in the 3D coordinates of the bone joint points. The z value is related to the distance from the object to the Kinect. If the z value is smaller and the distance is closer, then, the bone image is larger. Because Kinect does not use the same camera to collect depth images and color images, because the corresponding coordinate systems are different, so they need to be transformed into coordinate systems. The KinectSensor the SDK provides *MapDepthToSkeletonPoint*, *MapSkeleton-PointToDepth*, *MapDepthToColorImagePoint*, and *MapSkele-tonPointToColor* to transform coordinate system.
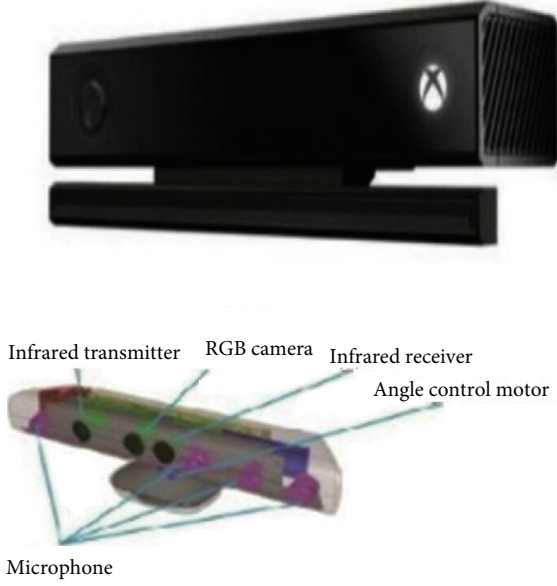
Infrared transmitter RGB camera Infrared receiver

Angle control motor

Microphone

FIGURE 1: Kinect structure.

4. Proposed Posture Recognition Method

The Kinect camera integrates devices such as infrared transmitter, RGB camera, and infrared receiver, as shown in Figure 1. The best working range is 1.2-3.5 m, the horizontal angle of the RGB camera is 57°, the vertical angle is 43°, and the shooting frequency is 30 Hz, which can ensure high accuracy in a fast scanning moment.

The human posture recognition algorithm is mainly composed of skeleton acquisition, angle measurement, angle matching, and posture recognition. The flow of the algorithm is shown in Figure 2. Firstly, the human skeleton is obtained and the spatial coordinates of the skeleton joints are calculated. Then, it calculates the distance between the joints and the angle between the joints. Finally, the calculated angle is matched with the angle template in the posture library to evaluate the posture recognition.

Kinect can provide the three-dimensional coordinates with 20 bone joints of the human body. Figure 3 is the skeleton diagram of the human body.

5. Calculating the Distance between the Joints

In the above analysis, 20 key points of the human body have been obtained. Next, the distance between the two key joints is computed. Firstly, the scene depth information obtained by Kinect is used to calculate the actual distance between the person and the camera. In reference [17], the obtained depth value was used to calculate the actual distance d from the target to the Kinect sensor, i.e.,

$$d = K \tan (H d_{\mathrm{raw}} + L) - O, \quad (1)$$

where d_{raw} is the depth value. $H = 3.4 \times 10^{-4}$ rad, $K = 12.35$ cm, $L = 1.18$ rad, and $O = 3.7$ cm. The transformation formula

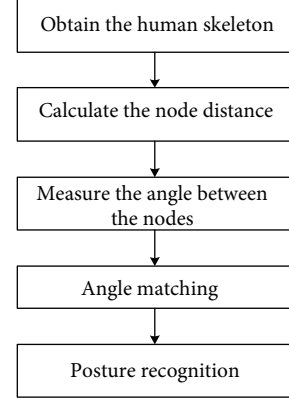FIGURE 2: Flowchart of the proposed algorithm for posture recognition.

from pixel coordinate $(x_{\mathrm{image}}, y_{\mathrm{image}}, \mathrm{and} z_{\mathrm{image}})$ of depth image to actual coordinate $(x_{\mathrm{real}}, y_{\mathrm{real}}, \mathrm{and} z_{\mathrm{real}})$ is

$$\begin{cases} x_{\mathrm{real}} = \left(x_{\mathrm{image}} - \dfrac{w}{2}\right)\left(z_{\mathrm{real}} + D'\right) F \cdot \dfrac{w}{h}, \\ y_{\mathrm{real}} = \left(y_{\mathrm{image}} - \dfrac{h}{2}\right)\left(z_{\mathrm{real}} + D'\right) F, \\ z_{\mathrm{real}} = d, \end{cases} \quad (2)$$

where $D' = -10$, and $F = 0.0021$ according to the abundant experiments. The resolution of Kinect is $w \times h = 640 \times 480$. $X(x_1, x_2, x_3)$ and $Y(y_1, y_2, y_3)$ are two points in the spatial coordinate system. Combining equations (1) and (2), the actual coordinate of the joint can be obtained. Then, we use the Euclidean distance to get the distance between the joints.

$$D(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2} \quad (3)$$

6. Calculating the Angle

The three-point method (three joints) is mainly used to solve the angles between human body connection points. The coordinates of the actual position of the key nodes calculated by the formula (2) are used to calculate the distance of the three key nodes related to the human body as shown in Figure 4. By using the cosine law (equation (5)), the angles between the connection points are calculated. The main disadvantage of this method in the recognition of human posture is that the instability of the closed joint has a great impact on the angle measurement during the measurement process resulting in inaccurate posture recognition. Figure 5(a) shows the angle measurement effect of the three-point method.

$$\begin{cases} a = D(B, C), \\ b = D(A, C), \\ c = D(A, B), \end{cases} \quad (4)$$
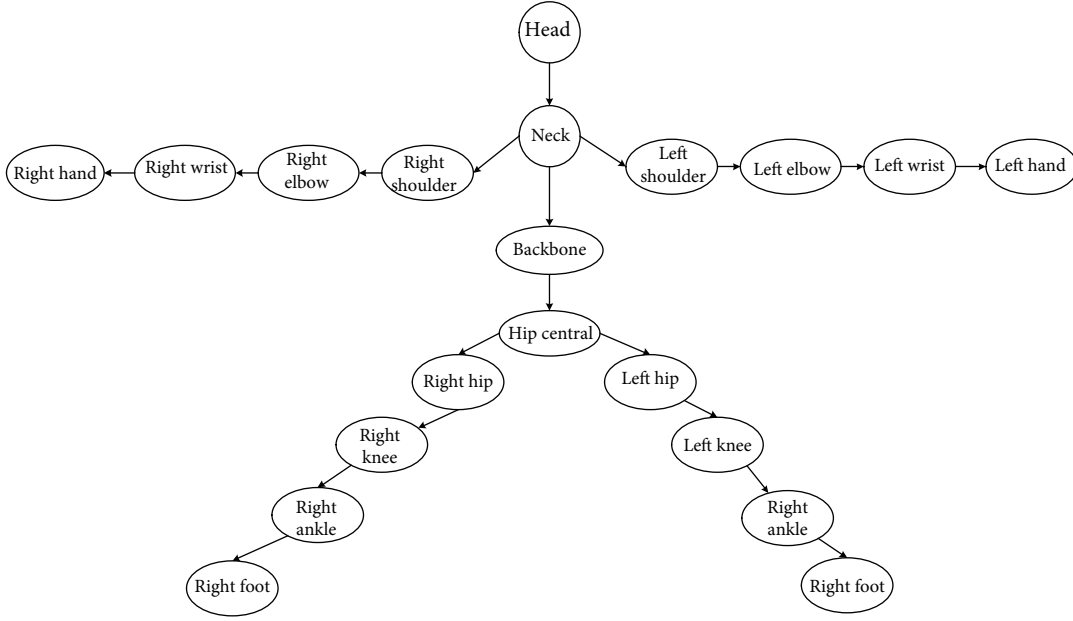
FIGURE 3: Joints of the human skeleton.

$$\theta = \cos^{-1} \frac{a^2 + c^2 - b^2}{2ac}. \tag{5}$$

7. Posture Definition

Equation (6) is used to define the angle condition of the joint.

$$P_A = \{P_1, P_2, \theta, \tau\}. \tag{6}$$

So it is centered on P_1. The angle between the joint P_2 and the x-axis is θ. τ is the set angle threshold value. The definition of more postures only needs to determine the angle relationship between the joints, and different thresholds can be set to meet different precision requirements. Set $\theta_i (i = 1, 2, 3, 4)$ as the angle of the joint. $\theta_1 = $ (left shoulder, left elbow), $\theta_2 = $ (left elbow, left wrist), $\theta_3 = $ (right shoulder, right elbow), and $\theta_4 = $ (right elbow, right wrist). τ is the threshold value. Then, the posture definition should satisfy the angle condition as follows: $\Delta = \{\theta_1, \theta_2, \theta_3, \theta_4, \tau\}$.

T- type (Posture starting): $\Delta T = (180, 180, 0, 0, 15)$;

Hands up: $\Delta D = (180, 90, 0, 90, 15)$;
Put down hands: $\Delta_0 = (270, 270, 270, 270, 15)$;
Raise left hand, flat right hand: $\Delta_{L_0} = (180, 90, 0, 0, 15)$;
Raise left hand, put down right hand: $\Delta_{L_1} = (180, 90, 270, 270, 15)$;
Raise right hand, flat left hand: $\Delta_{R_0} = (180, 180, 0, 90, 15)$;
Raise right hand, put down left hand: $\Delta_{R_1} = (270, 270, 0, 90, 15)$;

8. Human Body Posture Matching

In this paper, the threshold range of angle is set up when building a posture library. First, it traverses all the angles,
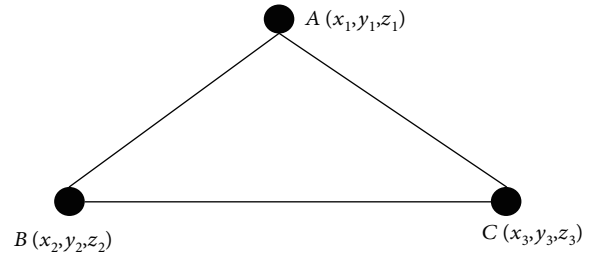

FIGURE 4: The calculation of joints angle.

then, it determines if the four angles are within the specified threshold. If YES, then the posture matching is successful. That is, all angles satisfy equation (7). If one of the angles is not satisfied, the matching is not successful and the match is resumed.

$$\max_{i \in N} |\theta_i - a_i| < T, \tag{7}$$

where θ_i is the measured angle, a_i is the set expected angle, and T is the threshold value.

9. Experiments and Analysis

In this section, the algorithm tests seven actions in physical education class from Shenyang Normal University. It can be seen from Table 1 that when the threshold value is set to 15°. The T-type, the raised hands are recognized with 100%. The recognition rate of putting down hands is recognized with 96%.

Participants make corresponding actions according to the prompts. It starts with a T-type with multiple stages. There is one action in link 1 and two consecutive actions in link 2. The rest can be done in the same manner. Only by

Angle1: 144.37621°

Angle2: 105.42531°

Angle3: 239.94973°

Angle4: 127.15227°

Angle1 :171.13171°

Angle2: 92.77521°

Angle3: 0.779486°

Angle4: 85.91438°

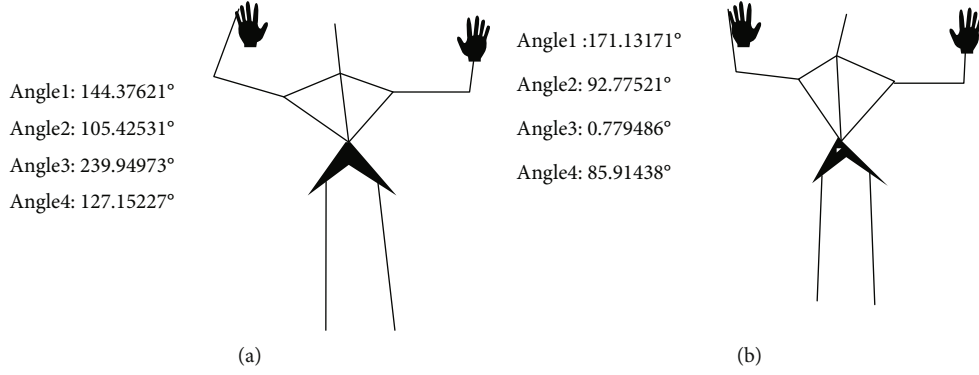(a)   (b)

Figure 5: Results of angle measurement. (a) Three-point method. (b) Two-point method.

Table 1: Statistical analysis of posture recognition experiments.

Action	Number of people	Test times/per capita	Correct recognition number	Recognition rate (%)
T-type	10	10	100	100
Raise hands	10	10	100	100
Put down hands	10	10	96	96
Raise left hand (flat right hand)	10	10	100	100
Raise left hand (put down right hand)	10	10	100	100
Raise right hand (flat left hand)	10	10	100	100
Raise right hand (put down left hand)	10	10	100	100

completing the low link, we can enter the high link. If this link is completed, and the next link does not meet the requirements, it will start from the first link again.

In this section, we test student sports pull-ups using our method. Kinect's core technology is bone tracking, which allows the device to better capture human motion and extract deep information. Microsoft Kinect adopts the depth measurement technology of Light Coding. To obtain the spatial position of the key joints, the human body in the depth image needs to be segmented through the machine learning method. Finally, the depth image is transformed into the bone image [18–20].

The conversion process from depth image to bone image requires three steps: body recognition, body part recognition, and joint recognition. The Kinect SDK can track the 3D coordinates of the 25 bone points for 30 frames in real time. Kinect 2.0 can identify the location of six people at the same time and give complete skeletal information of two people at the same time. With each joint of the three states: TRACKED, NOT TRACKED, and INFERRED can get a complete skeleton of the human body connected by 25 nodes.

The scoring module firstly extracts the user's motion information by using the Kinect module and then extracts the angle features and position information of the data [21, 22]. Since there is no time limit during the pull-up test, the score of each movement is calculated by presetting the scoring criteria and rules to judge whether the user's arm is straight and the position relationship between the jaw and the bar. There are two indicators to evaluate the pull-up, namely, the distance between the lower jaw and the bar when the body is at the highest point and the bending angle of the arm at the end of the movement.

By calculating the distance difference between the height (hi_i) of the mandible and the height of the actual crossbar (Δhi_i) when the human body is at the highest point, we can calculate the score $Ascore_i$, namely,

$$Ascore_i = \Delta hi_i \times \delta_i, \tag{8}$$

where δ_i is the variable used to determine the threshold of scoring within each distance interval. It is similar to calculate the bending angle of the arm and the distance between the lower jaw and the bar. The score $Bscore_i$ is set by measuring the angle difference $\Delta Angle_i$ between the bending angle of the i-th elbow joint and the angle at full extension, and the angle of the left and right elbow is averaged to set the score $Bscore_i$. The relation between the two angles can be expressed by the function $f(x)$, namely,

$$Bscore_i = f(\Delta Angle_i). \tag{9}$$

Since in the actual test, it is impossible to ask everyone to fully extend their arms, so we let the function $f(x)$ be a segmented function to set a certain threshold range. In this system
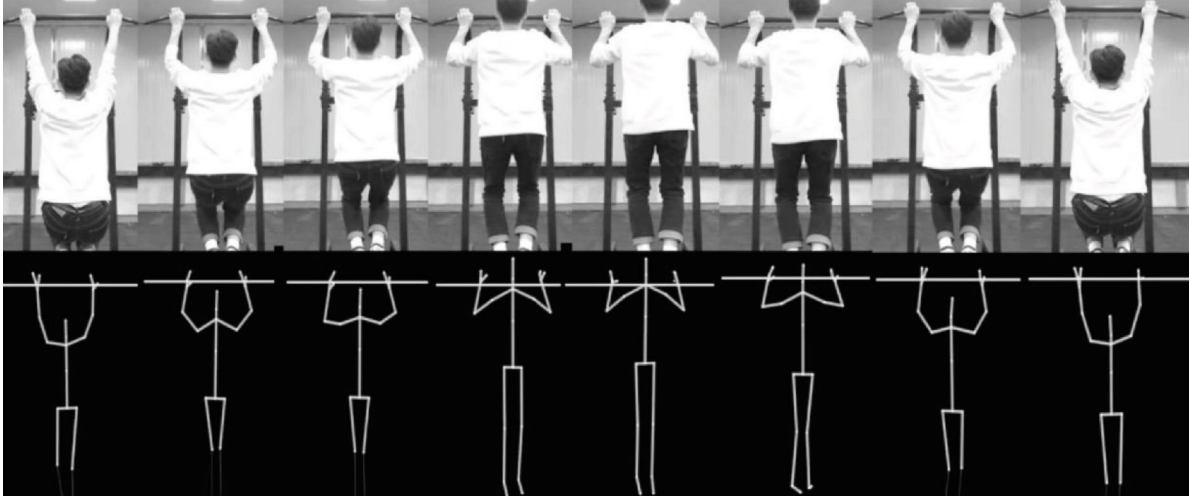
FIGURE 6: Lower crossing the jaw with arms straight.

$$f(x) = \begin{cases} 10, x \in [0, 20), \\ 5, x \in [20, 40), \\ 0, x \in [40, 90). \end{cases} \tag{10}$$

According to the national physical test standards, it is necessary to take the distance between the lower jaw and the crossbar and the straightening degree of the arm into comprehensive consideration to score the pull-up. Therefore, the weighted value of both should be taken to obtain the final score, i.e.,

$$\text{Score}_i = \alpha A\text{score}_i + \beta B\text{score}_i, \tag{11}$$

where α and β are the weight coefficients of $A\text{score}_i$ and B score$_i$, respectively. The weights of the two indexes can be changed by selecting different coefficients. In this system, $\alpha = \beta = 1$.

This project counts and scores according to the position relationship between the user's jaw and the crossbar and the bending angle of the arm. When the lower jaw crossing the bar is detected and the arm bending angle is within a certain threshold range then a count is made and the corresponding score is given. Other cases are not counted, and the corresponding score will be given. The following observations provide the counting and scoring results in several situations (The full score is 10).

The system detects that when the user's body is at the highest point, the lower jaw crosses the crossbar, the bending angle of the left arm is 172°, and the bending angle of the right arm is 163°. It counts once and determines the scoring interval according to the bending angle of the arm. We synthesize the $A\text{score}_i$ and $B\text{score}_i$, it obtains ten points as shown in Figure 6.

The system detects that when the user's body is at the highest point, the lower jaw crosses the crossbar, the bending angle of the left arm is 151°, and the bending angle of the right arm is 148°. It does not count and determines the scor-

ing interval according to the bending angle of the arm. We synthesize the $A\text{score}_i$ and $B\text{score}_i$, it obtains five points as shown in Figure 7.

The system detects that when the user's body is at the highest point, the lower jaw does not cross the crossbar, the bending angle of the left arm is 171°, and the bending angle of the right arm is 160°. It does not count and determines the scoring interval according to the bending angle of the arm. We synthesize the $A\text{score}_i$ and $B\text{score}_i$, it obtains six points as shown in Figure 8.

The system detects that when the user's body is at the highest point, the lower jaw does not cross the crossbar, the bending angle of the left arm is 158°, and the bending angle of the right arm is 161°. It does not count and determines the scoring interval according to the bending angle of the arm. We synthesize the $A\text{score}_i$ and $B\text{score}_i$, it obtains two points as shown in Figure 9.

Under natural conditions in the laboratory, 100 groups of experimental data are selected for the experiment. In this paper, the accuracy and real time of body recognition are tested, respectively. The experimental results are shown in Tables 2 and 3. It can be seen from the table that the recognition accuracy of the proposed method in this paper is over 88%.

This proposed method is compared with the other three body recognition methods containing DTW [23], IKS [24], and ConvNets [25]. Indicators include accuracy and time. Accuracy refers to the proportion of the correct sample in the total test samples.

10. Conclusions

This algorithm is developed by combining Microsoft Visual Studio 2010 with Microsoft Kinect SDK 1.7. The experiment shows that this method can measure the angle between the skeleton in real time and identify the posture of the human body accurately. The algorithm is simple and accurate. Moreover, different angle ranges can be set according to the requirements of different postures, so the reusability is
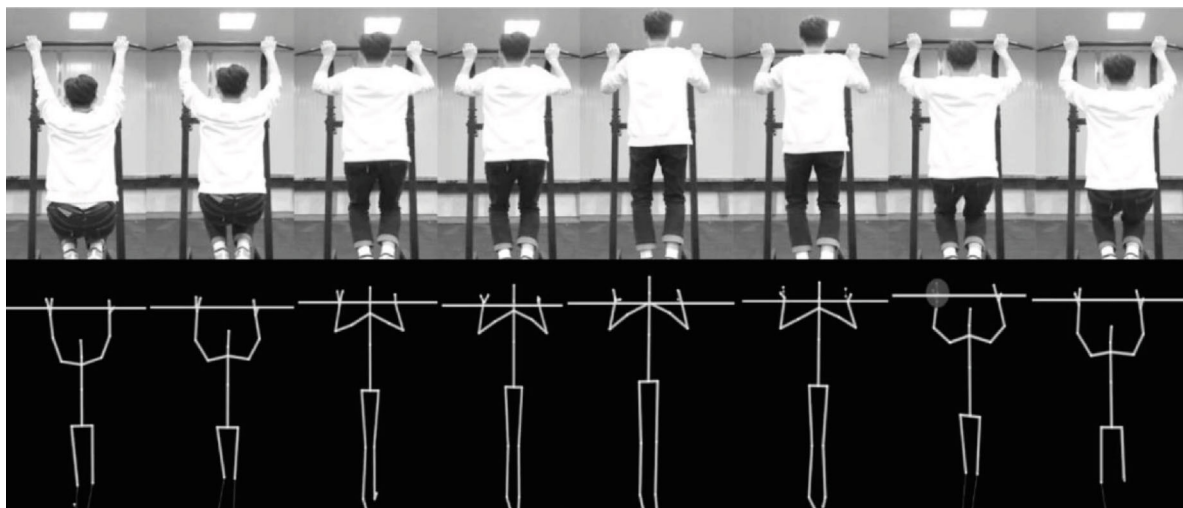
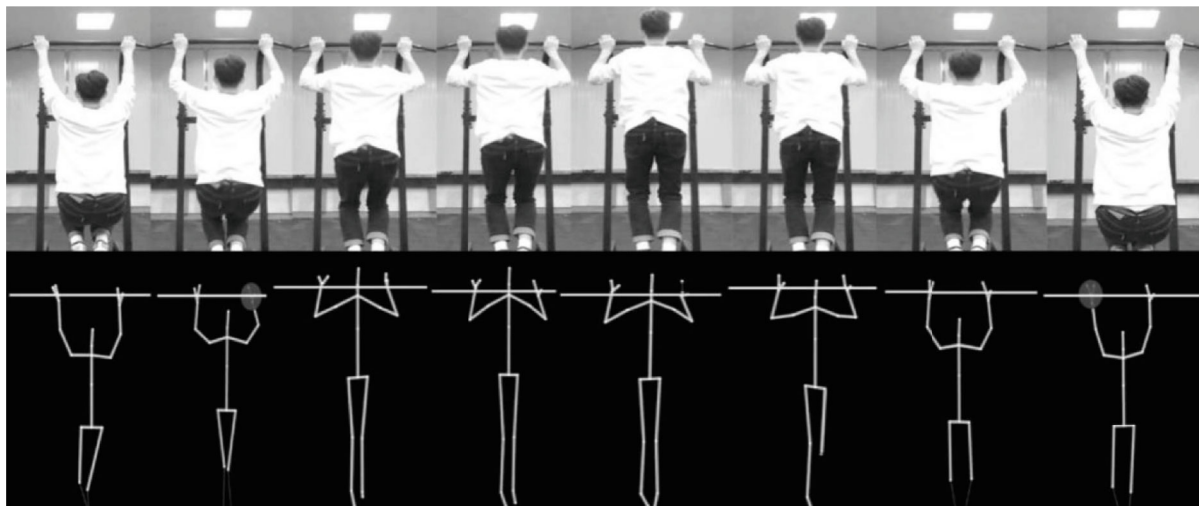Figure 7: Lower crossing the jaw with arms not straight.

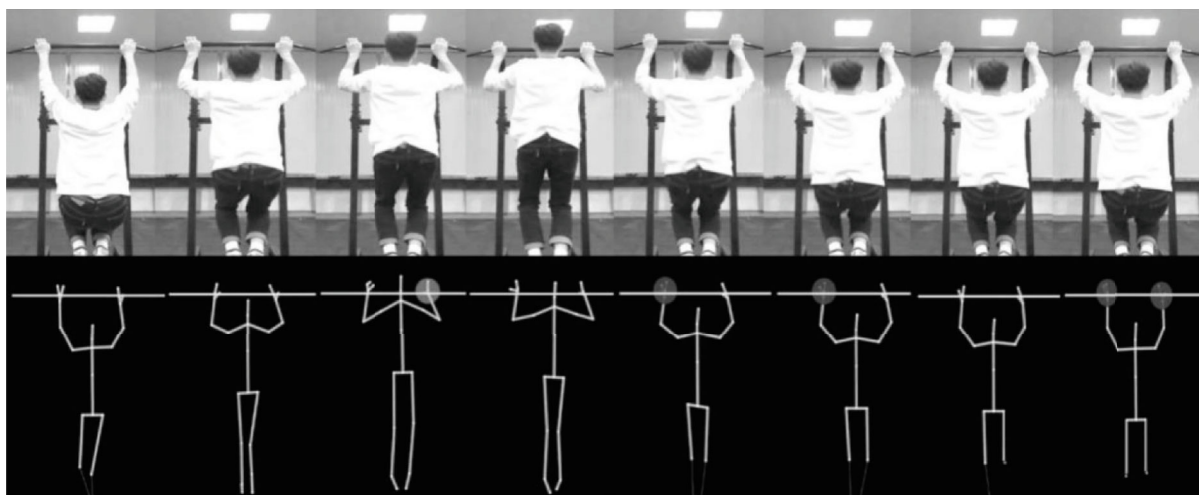Figure 8: Lower not crossing the jaw with arms straight.

Figure 9: Lower not crossing the jaw with arms not straight.

TABLE 2: Body recognition results.

Posture	Recognition number			
	Wave	Bow	Walk	Squat
Wave	98	92	96	98
Bow	—	90	—	—
Walk	—	—	89	—
Squat	—	—	—	97

TABLE 3: Results comparison with different methods.

Method	DTW	IKS	ConvNets	Proposed
Accuracy (%)	85.4	89.6	92.7	94.6
Time (s)	0.208	0.124	0.088	0.032

strong. Although the Kinect sensor can obtain the depth information of the human body and calculate the spatial position of the human body, it is not accurate enough to identify such problems as the coincidence of the joints. Therefore, while paying attention to the development of human behavior analysis, we should study problems such as skeleton correction to further improve the accuracy of skeleton. In the future, we will adopt deep learning and artificial intelligence method to perfect the quality of physical for all national persons.

Data Availability

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

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

References

[1] P. Li, Z. Chen, L. T. Yang, Q. Zhang, and M. J. Deen, "Deep convolutional computation model for feature learning on big data in internet of things," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 2, pp. 790–798, 2018.

[2] E. Gomede, F. Gaffo, G. Briganó, R. de Barros, and L. Mendes, "Application of computational intelligence to improve education in smart cities," *Sensors*, vol. 18, no. 1, p. 267, 2018.

[3] Q. Zhang, L. T. Yang, Z. Chen, and P. Li, "Incremental deep computation model for wireless big data feature learning," *IEEE Transactions on Big Data*, 2019.

[4] P. Li, Z. Chen, L. T. Yang, J. Gao, Q. Zhang, and M. J. Deen, "An incremental deep convolutional computation model for feature learning on industrial big data," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 3, pp. 1341–1349, 2019.

[5] G. Jim, "Inside the race to hack the Kinect," *New Scientist*, vol. 208, no. 2789, pp. 22-23, 2010.

[6] E. Cippitelli, S. Gasparrini, S. Spinsante, and E. Gambi, "Kinect as a tool for gait analysis: validation of a real-time joint extraction algorithm working in side view," *Sensors*, vol. 15, no. 1, pp. 1417–1434, 2015.

[7] X. Chen, Z. Cao, Y. Xiao, and Z. Fang, "Hand pose estimation in depth image using CNN and random forest," in *Tenth International Symposium on Multispectral Image Processing and Pattern Recognition (MIPPR2017)*, Xiangyang, China, 2017.

[8] N. D. Thang, T. S. Kim, Y. K. Lee, and S. Lee, "Estimation of 3-D human body posture via co-registration of 3-D human model and sequential stereo information," *Applied Intelligence*, vol. 35, no. 2, pp. 163–177, 2011.

[9] D. Xu, X. Xiao, X. Wang, and J. Wang, "Human action recognition based on Kinect and PSO-SVM by representing 3D skeletons as points in lie group," in *2016 International Conference on Audio, Language and Image Processing (ICALIP)*, pp. 568–573, Shanghai, China, July 2016.

[10] Y. Yang, F. Pu, Y. Li, S. Li, Y. Fan, and D. Li, "Reliability and validity of Kinect RGB-D sensor for assessing standing balance," *IEEE Sensors Journal*, vol. 14, no. 5, pp. 1633–1638, 2014.

[11] W. Wang, J. Chen, J. Wang, J. Chen, and Z. Gong, "Geography-aware inductive matrix completion for personalized point of interest recommendation in smart cities," *IEEE Internet of Things Journal*, 2019.

[12] W. Wang, J. Chen, J. Wang, J. Chen, J. Liu, and Z. Gong, "Trust-enhanced collaborative filtering for personalized point of interests recommendation," *IEEE Transactions on Industrial Informatics*, p. 1, 2019.

[13] Q. Zhang, C. Bai, L. T. Yang, Z. Chen, P. Li, and H. Yu, "A unified smart Chinese medicine framework for healthcare and medical services," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2019.

[14] T. Hoang, H. Dang, and V. Nguyen, "Kinect-based virtual training system for rehabilitation," in *2017 International Conference on System Science and Engineering (ICSSE)*, pp. 53–56, Ho Chi Minh City, Vietnam, July 2017.

[15] Q. Wang, P. Markopoulos, B. Yu, W. Chen, and A. Timmermans, "Interactive wearable systems for upper body rehabilitation: a systematic review," *Journal of Neuroengineering and Rehabilitation*, vol. 14, no. 1, p. 20, 2017.

[16] Q. Zhang, C. Bai, Z. Chen et al., "Deep learning models for diagnosing spleen and stomach diseases in smart Chinese medicine with cloud computing," *Concurrency and Computation: Practice and Experience*, 2019.

[17] S. Samoil and S. N. Yanushkevich, "Depth assisted palm region extraction using the Kinect v2 sensor," in *2015 Sixth International Conference on Emerging Security Technologies (EST)*, pp. 74–79, Braunschweig, Germany, September 2015.

[18] L. Xie and H. J. Liao, "A posture recognition method based on skeletal node and geometric relation using Kinect," *Applied Mechanics and Materials*, vol. 543-547, pp. 2879–2883, 2014.

[19] B. Li, B. Bai, C. Han, H. Long, and L. Zhao, "Novel hybrid method for human posture recognition based on Kinect V2," in *Computer Vision. CCCV 2017. Communications in Computer and Information Science, vol 771*, Springer, Singapore, 2017.

[20] Z. Zhang, Y. Liu, A. Li, and M. Wang, "A novel method for user-defined human posture recognition using Kinect," in *2014 7th International Congress on Image and Signal Processing*, pp. 736–740, Dalian, China, October 2014.

[21] J. Gao, P. Li, Z. Chen, and J. Zhang, "A survey on deep learning for multimodal data fusion," *Neural Computation*, vol. 32, no. 5, pp. 829–864, 2020.

[22] J. Gao, P. Li, and Z. Chen, "A canonical polyadic deep convolutional computation model for big data feature learning in internet of things," *Future Generation Computer Systems*, vol. 99, article S0167739X19307393, pp. 508–516, 2019.

[23] N. Li, Y. Dai, R. Wang, and Y. Shao, "Study on action recognition based on kinect and its application in rehabilitation training," in *2015 IEEE Fifth International Conference on Big Data and Cloud Computing*, pp. 265–269, Dalian, China, August 2015.

[24] M. Eltoukhy, J. Oh, C. Kuenze, and J. Signorile, "Improved kinect-based spatiotemporal and kinematic treadmill gait assessment," *Gait & Posture*, vol. 51, pp. 77–83, 2017.

[25] S. Neili, S. Gazzah, M. A. El Yacoubi, and N. E. Ben Amara, "Human posture recognition approach based on ConvNets and SVM classifier," in *2017 International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*, pp. 1–6, Fez, Morocco, May 2017.