

Research Article

Research on Action Recognition Method of Dance Video Image Based on Human-Computer Interaction

FenTian Peng¹ and Hongkai Zhang² 

¹Shangqiu Normal University College of Music, Hhnan 476000, China

²Physical Education Department of Guilin Teachers College, Guilin 541199, China

Correspondence should be addressed to Hongkai Zhang; zhg@mail.glnu.edu.cn

Received 27 July 2021; Revised 30 August 2021; Accepted 15 October 2021; Published 10 November 2021

Academic Editor: Bai Yuan Ding

Copyright © 2021 FenTian Peng and Hongkai Zhang. 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.

Human-computer interaction technology simplifies the complicated procedures, which aims at solving the problems of inadequate description and low recognition rate of dance action, studying the action recognition method of dance video image based on human-computer interaction. This method constructs the recognition process based on human-computer interaction technology, constructs the human skeleton model according to the spatial position of skeleton, motion characteristics of skeleton, and change angles of skeleton, describes the dance posture features by generating skeleton node graph, and extracts the key frames of dance video image by using the clustering algorithm to recognize the dance action. The experimental results show that the recognition rate of this method under different entropy values is not less than 88%. Under the test conditions of complex, dark, bright, and multiuser interference, this method can make the model to describe the dance posture accurately. Furthermore, the average recognition rates are 93.43%, 91.27%, 97.15%, and 89.99%, respectively. It is suitable for action recognition of most dance video images.

1. Introduction

Action recognition is one of the focus in current research studies. Many scholars have optimized and innovated the action recognition technology according to the characteristics of action changes and human body structure [1, 2]. Among them, the documents [3] optimized the action recognition method by using the improved deep convolution neural network and built a new recognition network by combining the Google Net network model with the idea of batch normalization transformation. The 4 documents [4] introduced the MEMS sensor network to collect the acceleration and angular velocity of gymnastics and recognized gymnastic movements by the classification model based on standard deviation, mean square error, and other classification feature of parameter setting, which has a high recognition rate.

However, when the background of the video image is complex, the light is weak, or there is multiuser interference,

the recognition rate of the previous research methods will be greatly reduced. In the broadest sense, human-computer interaction includes the interaction between human and machine, human and computer, and human and robot. In the past few years, human-computer interaction technology has been widely used in various industries and has good development in education, medical treatment, manufacturing, and other aspects, which can improve the display effect of action recognition effect. So, it is necessary to develop a new action recognition method of dance video image based on human-computer interaction technology.

2. Action Recognition Method of Dance Video Image Based on Human-Computer Interaction

2.1. Construction of Human Skeleton Model Based on Human-Computer Interaction. Our human skeleton consists of more than 200 skeleton nodes, and each node has its own degree of

freedom. In order to recognize dance action, the human skeleton model has a certain complexity. This study adds Kinect device to the action recognition method and replaces all skeleton information with joint points and then generates the joint point image which is consistent with the contour of the human body by human-computer interaction technology, which all aims to feedback the correlation between skeletons and joint nodes.

Generated by Kinect device, the nodes of hands, wrists, feet, and ankles are closer. In order to simplify the calculation, the nodes such as wrist and ankle are removed to optimize the human skeleton model. In view of the spatial position of skeletons, this paper studies the changes of skeleton position caused by dance movements at a certain time [5, 6]. Taking the two actions of raising hands and kicking as examples, the lower body is still when raising hands, while the upper body is relatively still when kicking. Based on this, this study regards the neck as a central node, takes the neck node as the center, constructs the spatial position structure of the model, establishes the space coordinate system by taking other skeleton nodes as the center, and then uses the coordinates of other points to do subtraction. It can get the following results:

$$H_{ab} = \{r_i - r_j\}, \quad i, j = 1, 2, \dots, N, i \neq j\}. \quad (1)$$

In the formula, H_{ab} represents an eigenvector whose coordinate is (a, b) ; r_i and r_j represent the horizontal coordinate of two adjacent skeleton nodes i and j on the same frame, respectively; and N represents the total number of nodes. Because there are differences in dancers' figure, the position from head to spine center is different. In order to minimize the influence, this design method calculates the absolute length from head to spinal center of identified dancer according to the calculation result of formula (1) and reduces the effects of figure difference. The formula is as follows:

$$H_{ab}' = \frac{H_{ab}}{L}. \quad (2)$$

In the formula, a represents the actual eigenvector under adaptive change and L represents the absolute length from head to spinal center. It is known that dancing is a dynamic movement process, so the constructed model needs to have the ability of dynamic change, which can generate matching recognition results according to the change of each frame in the video sequence [7–10]. Therefore, this method can calculate every skeleton node by subtraction based on the principle of difference operation between frames.

$$H_{bb}' = \{r_i^\alpha - r_j^\beta\}, \quad i, j = 1, 2, \dots, N\}. \quad (3)$$

In the formula, α and β represent the frame sequence of two continuous changing actions, respectively. The generation of dance action depends on the skeleton angle, so the model also needs to have the basic characteristics of angle changes of skeleton. If the azimuth angle and elevation angle of the skeleton are set as θ and γ , it also can obtain the skeleton recognition angle of the model through the following equations:

$$\begin{cases} \theta = \arctan\left(\frac{r_i^z - r_j^z}{r_i^x - r_j^x}\right), \\ \gamma = \arctan\left(\frac{r_i^y - r_j^y}{\sqrt{(r_i^x - r_j^x)^2 + (r_i^z - r_j^z)^2}}\right). \end{cases} \quad (4)$$

Substituting the calculation results of formulae (2)–(4) into the human skeleton model, the recognition of human skeleton motion features in the dance video images under the human-computer interaction technology [11, 12] can be realized.

The model minimizes the influence of dancers' body difference through formula (2) and obtains more accurate action recognition results. With the support of formula (3), the model has the ability of dynamic change and generates matching recognition results according to the change of each frame of video sequence. The model has the basic characteristics of bone angle change under formula (4), so it can recognize the characteristics of human bone movement in dance video images based on human-computer interaction technology [11, 12].

2.2. Description of Dance Posture Characteristics. Dance movement can be seen as a combination of multiple postures on the timeline, and its complexity and duration determine the length of the posture sequence. The skeleton feature recognition nodes of different dance postures can be obtained based on the above model. Under the control of human-computer interaction technology, the human skeleton model generates the posture according to the distance characteristics between the specific joint point and the central joint, among which 10 nodes were the most stable feature joint points.

According to the characteristic joint node $g_i^t = (x_i^t, y_i^t, z_i^t)$ and the "SpineBase" node $g_0^t = (x_0^t, y_0^t, z_0^t)$, it can calculate the relative distance between them and describe the dance posture at t time according to the result, in which $i = 1, 2, \dots, 10$. The formula is as follows:

$$\left| \overrightarrow{G_0^t G_i^t} \right| = \sqrt{(x_i^t - x_0^t)^2 + (y_i^t - y_0^t)^2 + (z_i^t - z_0^t)^2}. \quad (5)$$

In the formula, G_i^t represents the spatial coordinates of the i -th characteristic joint point at the time t and G_0^t represents the initial node coordinate. In order to eliminate the movement differences caused by height, body shape, and other aspects, this model deals with the skeleton distance characteristics of different people by the equal proportion method and distinguishes the directional characteristics of dance movements of left and right, up and down, and so on [13, 14]. It is known that the ratio of length of joint skeleton to height is constant. If the distance feature is multiplied by the height scale factor λ , then $\lambda = 1/\Delta y$ can be obtained, in which Δy represents the absolute value of the y coordinate difference between "Head" joint and "FootLeft" joint [15–18]. Assuming that frame number of dance video image

is M , then the distance characteristic matrix of $n * M$ dimension is as follows:

$$f = \begin{bmatrix} G_1 \\ G_1 \\ \vdots \\ G_t \\ \vdots \\ G_M \end{bmatrix} = \lambda \begin{bmatrix} \left| \overrightarrow{G_0^1 G_1^1} \right| & \left| \overrightarrow{G_0^1 G_2^1} \right| & \cdots & \left| \overrightarrow{G_0^1 G_n^1} \right| \\ \left| \overrightarrow{G_0^2 G_1^2} \right| & \left| \overrightarrow{G_0^2 G_2^2} \right| & \cdots & \left| \overrightarrow{G_0^2 G_n^2} \right| \\ \cdots & \cdots & \cdots & \cdots \\ \left| \overrightarrow{G_0^t G_1^t} \right| & \left| \overrightarrow{G_0^t G_2^t} \right| & \cdots & \left| \overrightarrow{G_0^t G_n^t} \right| \\ \cdots & \cdots & \cdots & \cdots \\ \left| \overrightarrow{G_0^M G_1^M} \right| & \left| \overrightarrow{G_0^M G_2^M} \right| & \cdots & \left| \overrightarrow{G_0^M G_n^M} \right| \end{bmatrix}. \quad (6)$$

According to the above process, this model realizes the recognition of dance video movements and describes the characteristics of dance posture. The description of recognition node of dancers' skeleton feature on the timeline is shown in Figure 1.

According to Figure 1, it can predict human activity intention and performance with high accuracy by determining the posture and motion of the specific dancer's skeleton feature recognition node on a given timeline. Although different people have different sizes and shapes of skeleton, their motion performance is similar when they perform similar movements. Therefore, for a given movement, the speed and acceleration of joint points are more stable than the position information of dancers' skeleton feature recognition node.

2.3. Image Action Recognition by Extracting Key Frame.

The sample video sequence of dance movements is established based on the description results of dance posture features, which is represented by $\{p^{(1)}, p^{(2)}, \dots, p^{(M)}\}$. And there is $p^{(i)} \in W^{(M)}$. Here, $p^{(i)}$ represents the 3D coordinate position vector of the joint point in the i th frame of the sequence; $W^{(M)}$ represents the set of vectors; and $\{c^{(1)}, c^{(2)}, \dots, c^{(k)}\}$ represents the extracted sample centers whose number is k . k cluster centroids are selected randomly and represented by $\{o_1, o_2, \dots, o_K\}$, respectively,

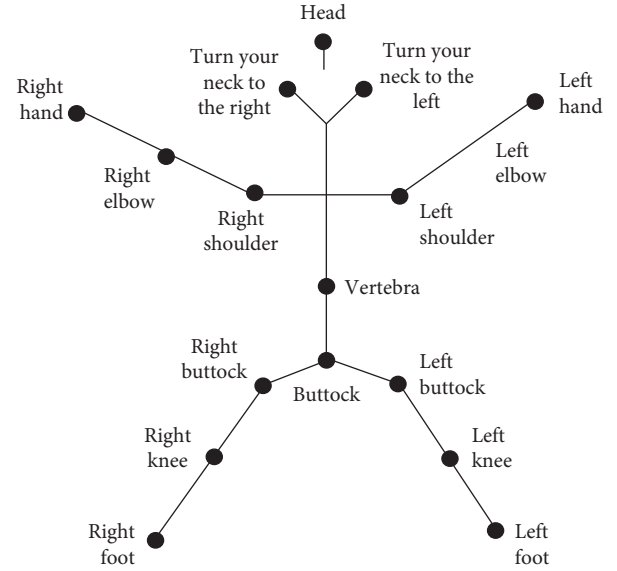


FIGURE 1: Identify nodes of dancer skeleton feature.

$\{o_1, o_2, \dots, o_K\} \in W^{(M)}$ [19, 20]. It can calculate that the minimum distance between the sample and k cluster centroid in the random sample $p^{(i)}$, and the formula is as follows:

$$D = \arg \min \sum_{i=1}^M \sum_{j=1}^K \|p^{(i)} - o_j\|^2. \quad (7)$$

In the formula, D represents the minimum distance between the sample and the centroid. It can classify the samples into a class set that should belong to a certain feature according to the above calculation results, and then the centroid is recalculated for each class.

$$o_j = \frac{\sum_{i=1}^M l_{ij} p^{(i)}}{\sum_{i=1}^M l_{ij}}. \quad (8)$$

In the formula, l_{ij} represents the feedback constant when the vector $p^{(i)}$ is classified; $l_{ij} = 1$ exists when the class is j ; otherwise, $l_{ij} = 0$ [21, 22]. k sample centers are returned according to the calculation results, and the above steps are repeated until the convergence is completed.

The frame clustering is completed according to the above process, and the Euclidean distance is calculated between the sample and the m th frame. The formula is as follows:

$$I = \left[\text{dis}[(x_{i1}, y_{i2}, z_{i3}), (x_{m1}, y_{m2}, z_{m3})], \dots, \text{dis}[(x_{ij}, y_{ij}, z_{ij}), (x_{mj}, y_{mj}, z_{mj})] \right]. \quad (9)$$

After calculating the Euclidean distance for all the frames of the dance video image, the minimum value of $\text{dis}[(x_{ij}, y_{ij}, z_{ij}), (x_{mj}, y_{mj}, z_{mj})]$ is recorded as 1, otherwise as 0. The extraction of key frames is completed so far. According

to the deep fusion feature of the result, namely, the fusion of the location feature and the angle feature of the joint points, the movement recognition of the dance video image is realized based on the human-computer interaction [23–25].

3. Case Testing

3.1. Setup Human-Computer Interaction Test Environment.

The design of the interactive interface can be realized through Windows Presentation Foundation. In order to ensure the stability of the test, this example selects a four-wheel robot for man-machine interaction, and the upper limb of the robot can simulate human arm movement. This example sets 10 kinds of dance movements according to the human-computer interaction technology. The action picture is shown in Figure 2.

The robot can recognize actions and make corresponding postures according to the ten groups of dance action images in Figure 2. The recognition device of the robot is shown in Figure 3.

The robot can recognize the movements of these dance video images and make corresponding gestures. The recognition device is shown in Figure 3.

In Figure 3, the speed control is a very important link in the robot driving process. The four-wheel self identification robot can adjust driving force according to the position signal collected by the recognition device. If the driving force is greater than the resistance, the robot will continue to accelerate at a certain acceleration. And it is easy to rush out as turn the corner when the speed is too high. In the same way, if the driving force is less than the resistance, the robot will continue to slow down. The robot tested in this example must keep the corresponding speed according to the actual situation, which is to ensure the unique variable in the process of human-computer interaction and simulation.

The human-computer interaction interface is shown in Figure 4.

In the human-computer interaction display interface, the first step is to record the dance video. During this process, the recorder should stand within the effective recognition range of the Kinect sensor. At this time, the skeleton image and dance video image can be shown in the interface. And the name of the recorded action in ComboBox control is selected. The second step is to click "Motion capture," and the human-computer interface display comes into 3S countdown. The recorder can start to show the dance movements after timing. The interface uploads the action feature data to the buffer in real time and starts to enter the action recognition program, while recorder can check action recognition effect. If the dance movements are not standard or the collected data are not comprehensive, it needs to record again. The last step is to record 10 kinds of dance movements one by one according to the above process, then click "Save to Template.txt" to store the data, establish the action template library, update the system interface, open "Load Action Template," and finally enter the action recognition test stage after generating all the data.

In order to setup human-computer interaction test environment, it is necessary to turn on the switch of the robot, connect the Bluetooth of the robot's main controller with the Bluetooth adapter of the computer, and select the communication serial port and baud rate. Then, it is also important to click the "Open/Close COM" button to establish effective communication between the robot and the computer. The recorder can control the robot through its

own actions. According to the recognition results of the video image of the recorder's dance movements, this test evaluates the action recognition method based on human-computer interaction and analyzes the overall performance and robustness of the method.

3.2. Overall Performance Evaluation. It is known that different joint points have different abilities to describe actions. So, its contribution can be described by information entropy, among which the joint points are lower than the threshold value that needs to be eliminated. This study determines the threshold value according to the recognition rate and gets information entropy threshold. The formula is as follows:

$$s = \frac{n_1}{n_1 + n_2}. \quad (10)$$

In the formula, s represents the average recognition rate of 10 dance movements and n_1 and n_2 represent the times of recognition successes and recognition failures, respectively. Assuming that the threshold is K , the interval is set to 0.05 according to the equal interval value method when $K \in [0.05, 0.95]$. The recognition rate under different entropy values is shown in Figure 5.

According to the test results in Figure 5, the recognition rate is the highest when $K = 0.25$ and $s = 0.9762$. When the value of K continues to increase, the value of s decreases. Because the entropy value of some joint points with large contribution will be less than the threshold value when the value of K is too large, it always filters out these joint points. So, it will influence the action display effect of the robot in the process of action recognition. According to the test results shown in Figure 5, with the increase in K value, the overall result is above 0.88 although the value of average recognition rate s decreased somewhat, which indicates that this method is practical and can be applied to the action test.

3.3. Robustness Testing. In order to achieve a good test effect of action recognition in different dance video images, the test verifies whether the recognition method can successfully recognize the dance movement under the condition of different background environment, different ambient light, and the presence of multiuser interference.

Figure 5 shows the action recognition effect based on the above test conditions.

According to Figure 6, it can be seen that under different background environments, the recognition methods based on human-computer interaction can use the constructed human skeleton model to describe the characteristics of dance posture accurately.

This test sets 10 groups of actions and calculates the average recognition effect of 10 groups of actions in four test environments. Taking the example of lifting the left arm, swinging the right hand backward, and lifting the left leg, the simulation diagram is shown in Figure 7.

For the simulation recognition results obtained under the human-computer interaction technology and the quantization processing, it is processed quantitatively. The quantization results are shown in Table 1.



FIGURE 2: 10 groups of dance actions in the experiment.

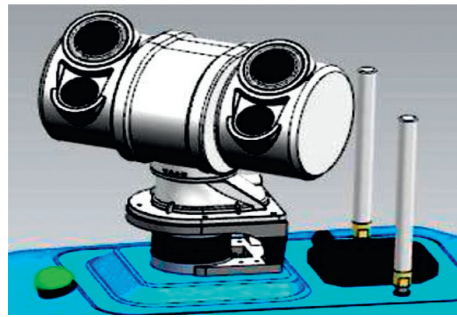


FIGURE 3: Recognition device of four-wheel robot.

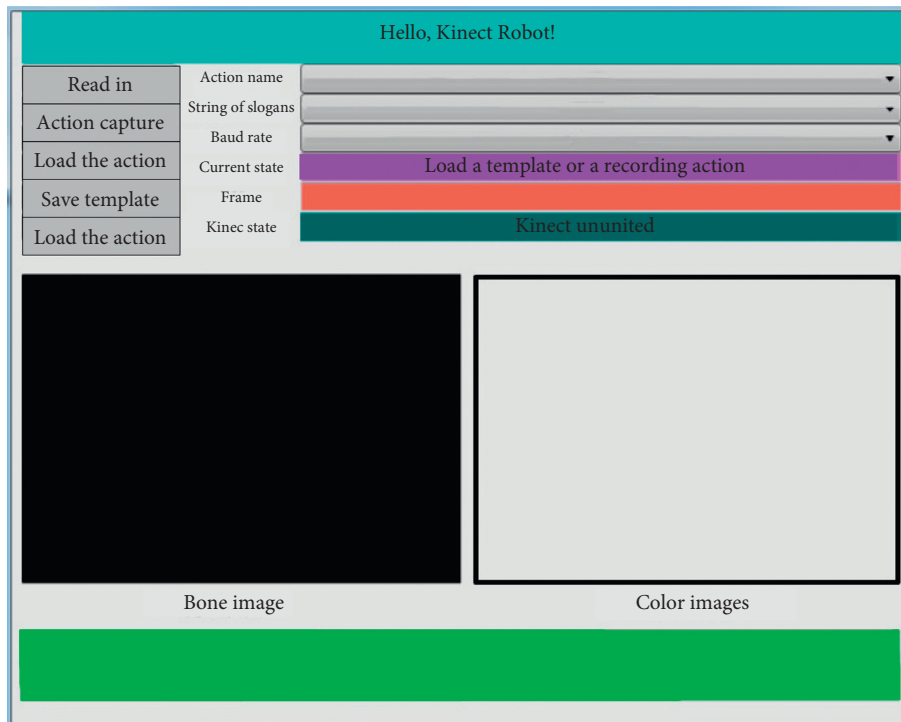


FIGURE 4: Display interface of human-computer interaction.

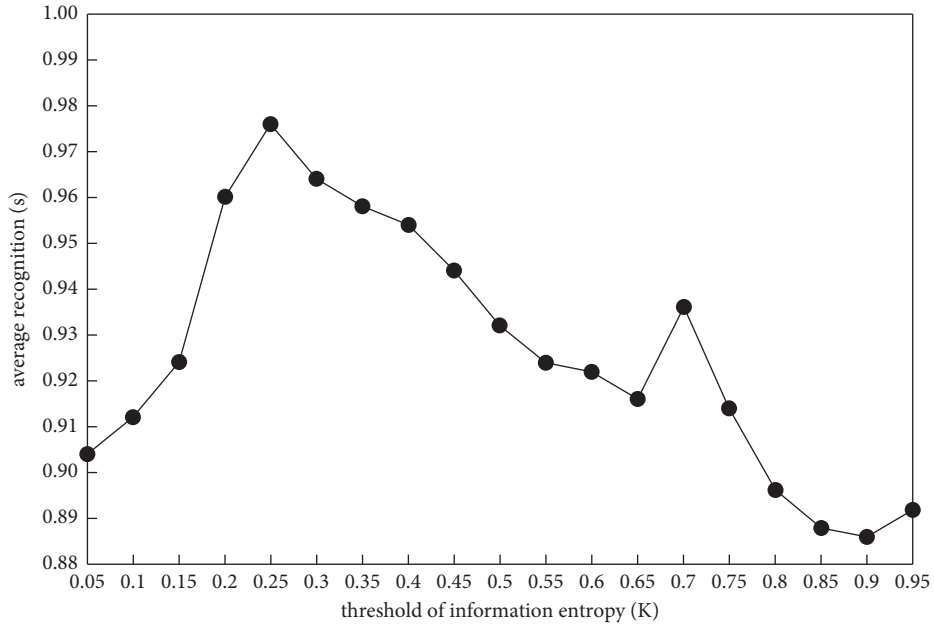
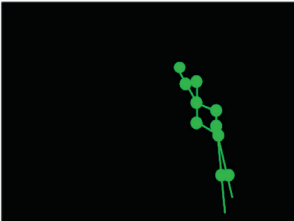



FIGURE 5: Recognition rate under different entropy values.

Hello, Kinect Robot!

Read in	Action name	1
Action capture	String of slogans	COM1
Load the action	Baud rate	9600
Save template	Current state	Gesture # right arm up added, recognizing
Load the action	Frame	
	Kinect state	Kinect uninitiated



Bone image



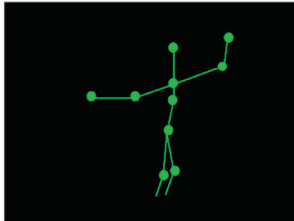
Color images

Identification result: # DTW distance of right arm lifting: 0.110976

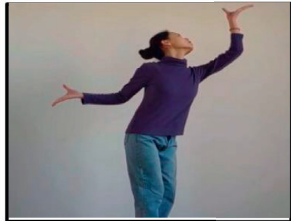
(a)

Hello, Kinect Robot!

Read in	Action name	2
Action capture	String of slogans	COM1
Load the action	Baud rate	9600
Save template	Current state	Gesture # left arm raised added, recognizing
Load the action	Frame	
	Kinect state	Kinect uninitiated



Bone image



Color images

Recognition results: # left arm lifting DTW distance: 0.473928

(b)

FIGURE 6: Continued.

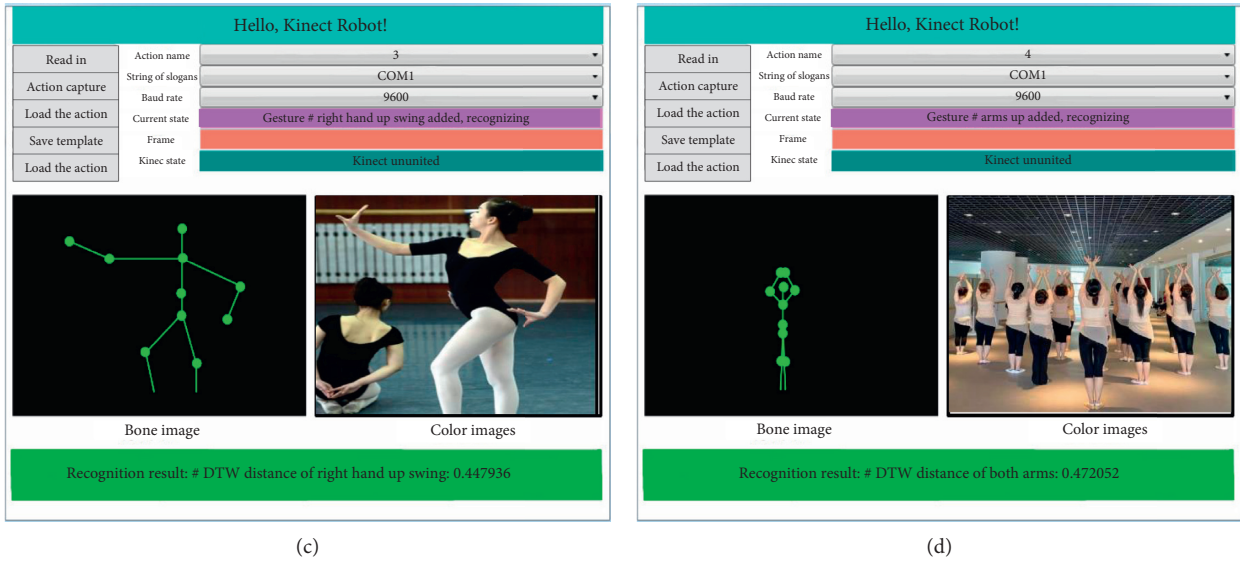


FIGURE 6: Recognition effect: (a) complex environment; (b) dark environment; (c) bright environment; (d) multiuser interference environment.

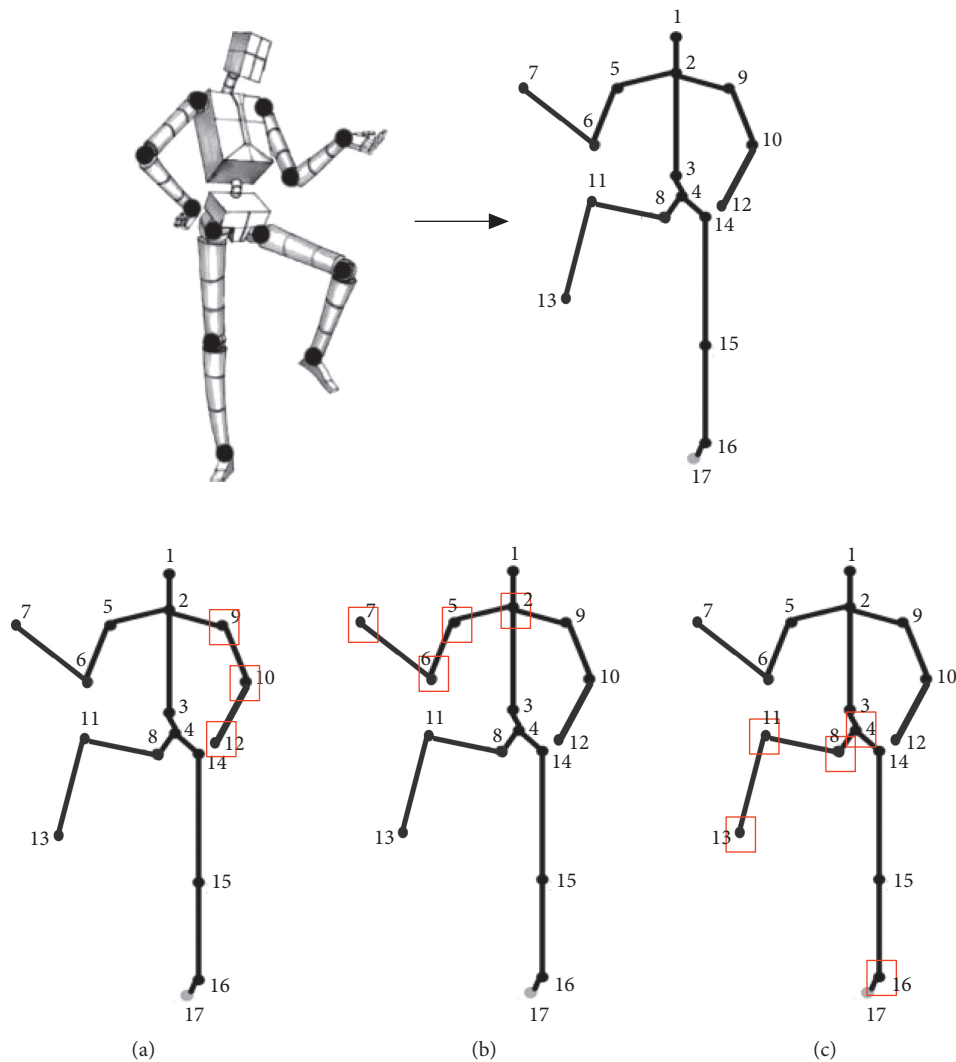


FIGURE 7: Simulation recognition results based on human-computer interaction technology: (a) the right hand swings backward; (b) the left arm up; (c) the left leg up.

TABLE 1: Test result of average recognition rate (%).

Dance movements	Complex environment	Dark environment	Bright environment	Multiuser interference environment
Left arm raised	95.2	90.3	97.5	90.4
Right arm raised	92.6	92.4	97.4	89.8
Arms raised	93.8	90.5	96.8	91.4
Arms flatted	94.1	90.3	97.2	90.5
Arms downward	92.5	90.4	97.6	89.2
Arms crossed	93.3	90.5	96.8	88.7
The left hand swings forward	93.0	92.7	96.9	89.5
The right hand swings backward	92.5	91.4	96.9	90.1
Clap	94.7	93.6	97.5	92.3
Arms akimbo	93.2	91.5	97.2	88.4
Arms on the shoulder	92.8	90.4	96.9	89.6

According to Table 1, it can be seen that the average values of four groups of test results are 93.43%, 91.27%, 97.15%, and 89.99%, respectively. Based on the above test results, it shows that the action recognition method of dance video image based on human-computer interaction can obtain more accurate recognition results under different test backgrounds.

4. Conclusion

This research takes human-computer interaction technology as the innovation point, studies a new action recognition method, and completes the recognition of specific dance video image action through human-computer interaction. The experimental results show that

- (1) The joint points of the human are mainly distributed in the limbs. The human skeleton model can be built combined with the joint points of the upper limbs, lower limbs, head, neck, and shoulders. Its principle is to accurately record the movement of each joint point and recognize the movements of the dancer in the process of various movements so as to output the correct human motion skeleton. With the human motion skeleton model, the recognition accuracy and efficiency of the computer vision system can be significantly improved. The human skeleton model based on human-computer interaction technology can accurately describe the characteristics of dance posture and make the recognition rate of the method in the range of 0.8865–0.9762 under different entropy values.
- (2) Using the clustering algorithm to extract key frames of dance video image can reduce the influence of interference environment on recognition effect. In the experimental results, the average recognition rates in complex environment, dark environment, bright environment, and multiuser interference environment are 93.43%, 91.27%, 97.15%, and 89.99%, respectively. The above data prove that the proposed method is robust and suitable for motion recognition of most dance video images.

This research takes human-computer interaction technology as the innovation point, studies a new action

recognition method, and recognizes the action in dance video image by human-computer interaction. However, as a result of the test time is insufficient and experimental test preparation is hasty, there are some certain limitations in terms of test location and setting different test environment. In the future research, we can set more experimental test conditions, but not limited to the different gender, body shape of the recording object, and dance actions of different difficulties, so as to provide more detailed and accurate research results for the development of new recognition technology.

Data Availability

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding this work.

References

- [1] K. Zhang, T. Liu, Z. Liu, Y. Zhuang, and Y. Chai, "Multimodal human-computer interaction technology for emotion regulation," *China Journal of Image and Graphics*, vol. 25, no. 11, pp. 2451–2464, 2020.
- [2] X. Fan, J. Fan, F. Tian, and G. Dai, "Human-computer interaction and artificial intelligence: from competition to integration," *Scientia Sinica Informationis*, vol. 49, no. 3, pp. 361–368, 2019.
- [3] S. Chen, W. Wei, B. He, S. Chen, and J. Liu, "Action recognition based on improved deep convolutional neural network," *Application Research of Computers*, vol. 36, no. 3, pp. 945–949+953, 2019.
- [4] J. Sun, X. Meng, H. Liang, H. Duan, and Z. Zhan, "Gymnastics action recognition based on MEMS sensor," *Journal of Electronic Measurement and Instrument*, vol. 34, no. 3, pp. 94–99, 2020.
- [5] D. Zhao, L. Yu, G. Liu, and Z. Dai, "Human-machine interaction intention recognition technology based on D-S evidence theory," *Modular machine tool and automatic processing technology*, no. 3, pp. 60–63, 2019.
- [6] L. Jing, L. Zhu, and X. Gou, "Kinematics analysis of lower limb rehabilitation exoskeleton mechanism based on human-

- machine closed chain,” *Journal of Engineering Design*, vol. 26, no. 1, pp. 65–72+109, 2019.
- [7] G. Zuo, Z. Xu, J. Lu, and D. Gong, “A structure-optimized DDAG-SVM action recognition method for upper limb rehabilitation training,” *Acta Automatica Sinica*, vol. 46, no. 3, pp. 549–561, 2020.
- [8] B. Jia and P. Li, “Number gesture recognition method in human-computer interaction,” *Journal of Huaqiao university (Natural Science Edition)*, vol. 41, no. 2, pp. 260–267, 2020.
- [9] H. Zhang, D. Fu, and K. Zhou, “Video action recognition method based on timing enhancement,” *Pattern Recognition and Artificial Intelligence*, vol. 33, no. 10, pp. 951–958, 2020.
- [10] G. Liu, N. Wang, Y. Zhou, W. Wang, and Y. Yue, “Dumbbell action recognition based on improved ReliefF algorithm,” *Science technology and engineering*, vol. 19, no. 32, pp. 219–224, 2019.
- [11] J. Yang, M. Shi, F. Chao, and C. Zhou, “The dancing-robot for action imitation based on deep learning,” *Journal of Xiamen University*, vol. 58, no. 5, pp. 759–766, 2019.
- [12] L. Chen, “Research on recognition technology of human-body action in dance video image,” *Modern Electronic Technology*, vol. 40, no. 3, pp. 51–53, 2017.
- [13] H. Li, “Research on motion recognition method in dance video images,” *Video Engineering*, vol. 42, no. 7, pp. 34–37, 2018.
- [14] Y. Xu, “Research on dynamic recognition of gait silhouettes of dance movement based on AVI video,” *Modern Electronic Technology*, vol. 43, no. 16, pp. 119–121, 2020.
- [15] D. Li, “Design of retrieval system of specific action segments in high dynamic dance video,” *Modern Electronic Technology*, vol. 41, no. 5, pp. 97–101, 2018.
- [16] S. Lu, Q. Cao, and M. Sun, “Hardware design of mechatronic integration robot joint and its driving control system,” *Journal of Chongqing University of posts and Telecommunications (Natural science edition)*, vol. 47, no. 1, pp. 45–49, 2021.
- [17] N. Ma, X. Shi, D. Qin, and C. Liu, “A method of key frame extraction for music and dance video,” *Journal of System Simulation*, vol. 30, no. 7, pp. 2801–2807, 2018.
- [18] X. Chen, Y. Wang, Z. Wu, and J. Li, “Research and realization of a stage augmented reality system,” *Computer Engineering and Applications*, vol. 54, no. 3, pp. 11–17, 2018.
- [19] Q. Cai, E. Li, J. Jiang et al., “Study on the tea identification of near-infrared hyperspectral image combining spectra-spatial information,” *Spectroscopy and Spectral Analysis*, vol. 39, no. 8, pp. 2522–2527, 2019.
- [20] E. Yang, A. Zhang, R. Yang, and C. Wang, “Automatic recognition algorithm of three-dimensional image of surface cracks on ballastless track of high speed railway,” *Railway Transaction*, vol. 41, no. 11, pp. 95–99, 2019.
- [21] S. Dong, X. Sun, S. Xie, and M. Wang, “Automatic defect identification technology of digital image of pipeline weld,” *Natural Gas Industry*, vol. 39, no. 1, pp. 113–117, 2019.
- [22] W. Liu, H. Tian, J. Xie, E. Zhao, and J. Zhang, “Identification methods for forest pest areas of UVA aerial photography based on fully convolutional networks,” *Journal of Agricultural Machinery*, vol. 50, no. 3, pp. 179–185, 2019.
- [23] X. Zhou, X. He, and C. Zheng, “Recognition of radio signal based on image deep learning,” *Journal of Communications*, vol. 40, no. 7, pp. 114–125, 2019.
- [24] 魏.彤. Wei Tong and 周. Yin-he, “Blind sidewalk image location based on machine learning recognition and marked watershed segmentation,” *Optics and Precision Engineering*, vol. 27, no. 1, pp. 201–210, 2019.
- [25] Z. Zhong, “A method of image graying for roadway line identification,” *Journal of Tongji University*, vol. 47, no. 1, pp. 178–182, 2019.