Exploration of Dance Choreography Strategies Based on Big Data Information Analysis Technology

Qin Wang
Lianyungang Normal College,Lian yungang ,Jiangsu china, 222006
Email: wwx13321@163.com

Abstract:

Intelligent choreography is the inevitable direction of the intelligent development of dance art. At present, there is the problem of insufficient accuracy of choreography feature extraction. This paper proposes an automatic music choreography algorithm combined with big data. Based on a large number of existing music and dance data, this algorithm uses machine learning algorithm to train the model, which can automatically and intelligently generate desired dance movements in combination with screening conditions. Moreover, combined with the idea of data mining, this paper puts forward a data generation strategy, which randomly adds irregular disturbances at the time series level and spatial level to the professional dance sequences in the data set, synthesizes non-professional dances, and constructs a training set that meets the requirements in scale and choreography complexity. In addition, this paper designs a two-stage framework: dance-music alignment stage. Through the analysis of the results, it can be seen that the choreography method combined with feature extraction proposed in this paper can effectively identify various dance features. In particular, it can learn from a large number of dance features, effectively improve the quality of dance choreography, and play a certain role in promoting the further development of subsequent dance choreography.

Keywords: choreography; big data; optimization; strategy

1 INTRODUCTION

In recent years, with the rapid development of new technologies such as artificial intelligence and virtual reality, virtual character dance has gradually emerged, which not only presents eye-catching stage effects, but also has been successfully applied to robot dance, auxiliary dance teaching, game character movement generation and other fields. Traditional choreography is time-consuming and laborious. It not only requires professional dancers to repeatedly and accurately dance to establish movement capture data, but also requires professional animators to correct the data, which seriously hinders the further development of virtual character dance. The development of neural networks makes it possible to train models based on massive data, but the scarcity of pairwise data of dance and music makes the research of intelligent choreography limited [1]. The diversity of dances requires a variety of dances and rich combinations of movements. Even for experienced movement capture companies, it is quite a difficult task to establish a large number of data sets that meet the requirements. Nowadays, digital platforms have become an increasingly important part of people's lives. Through the intelligent choreography and display system to realize the storage, classification, retrieval, storage and utilization of multimedia data such as dance can not only significantly promote the research and development of intelligent choreography, but also enrich people's lives. At present, most of the dance systems are dance-assisted teaching, and the management of multimedia digital assets is mostly used in museums and art institutions. However, the asset management and display system based on intelligent dance choreography has not yet been developed. Based on the design idea of "high cohesion and low coupling", the hierarchical architecture of the system simplifies the responsibilities of each layer, which can realize many quality attributes such as system reusability, portability, ease of maintenance, ease of understanding and replaceability [2].

This paper proposes an automatic music choreography algorithm combined with big data. Based on a large number of existing music and dance data, this algorithm uses machine learning algorithm to train the model, which can automatically and intelligently generate desired dance movements in combination with screening conditions. Moreover, combined with the idea of data mining, this paper puts forward a data generation strategy, which randomly adds irregular disturbances at the time series level and spatial level to the professional dance sequences in the data set, synthesizes non-professional dances, and constructs a training set that meets the requirements in scale and choreography complexity.

2 RELATED WORKS

(1) Automatic choreography of dance

The beauty and expressiveness of the dance depend largely on the choreography of the dance. Choreographing a dance often requires careful design and thought from a professional choreographer. From listening to music to choreographing a beautiful dance, it usually takes a lot of energy and the choreography process takes a long time. Applying computers to choreography can help dancers choreograph more quickly and efficiently. The first thing to consider in choreography is the matching of dance and music.

Reference [3] proposed an autoregressive codec network that can generate matched choreographic sequences for a given audio input. They extracted the joint coordinates of the dancers from the video and the music spectrum from the audio. The system can generate dances that match the music, and can express the music according to a piece of music that has never been heard before, making the music of choreography more extensive. Reference [4] proposed a new dance automatic generation framework, which consists of three parts, namely, music feature encoder, pose generator and music type classifier. The music feature encoder consists of a one-dimensional CNN, which converts the input music (cha-cha, rumba, tango, waltz) into a set of audio feature vectors, and then determines the music type by considering the entire audio feature vectors. The pose generator uses an autoregressive model to generate the next pose frame with the current output pose as the input of the next node. The framework can generate a single-person dance video dancing to a given music beat, and can generate appropriate dances according to the categorized types. Furthermore, the dance pattern learned from the training data may exhibit that a movement sequence different from that of the training data may be generated according to the ability of never hearing the input music dance. To achieve reasonable choreography, the most important thing is to deal with the matching problem of dance and music [5]. Reference [6] used music and movement information to extract key gestures, and used audio beats to extract key gestures from videos instead of movement information from videos. Then, it extracted features from RGB images of bones and Kimect, and used three classifiers: GMM (Gaussian Mixed Model), SVM (Support Vector Machine) and CNN to solve the pose recognition problem. Finally, it designed an Adavus-oriented recognizer using HMM.

(2) Unity of rhythm and movement

The beat is a distinct feature of music. The correlation between dance movements and music beats is the basic condition for the performance of music in dance [7]. There should be a corresponding relationship between dance movements and music beats, so that dance and music are more closely linked. If the dance is out of sync with the beat, it will make the dance look weird. Moreover, depending on the strength of the music pitch, the strength of the dance movements will also vary. At present, researchers have done the following research on the correspondence between music beats and dance movements.

Reference [8] proposed a model that maps the beat structure to dance movements using Gaussian processes, and used a large number of dance movements on the Internet for training. However, this approach only focuses on the relationship of music beats to dance movements, and ignores the influence of other characteristics of music. Reference [9] believed that BGM (background music) can enhance the emotional impact of computer animation. By changing the relationship between background music and animation, the synchronization of background music and animation is realized. Then, feature points are extracted from MIDI files and movement data and matched, and the dynamic synchronization of music and movement is realized. The study also introduces how to generate a large number of new background music by traversing the music diagram, so that the generated music is more in line with the dance movement than the original music. Reference [10] added an acoustic feature extraction module to the LSTM model. The feature learning component is designed by using autoencoder. The LSTM-Autoencoder model can extract the mapping relationship between sound and movement features, and the output results are not only consistent with music, but also have rich expressiveness and good continuity. At the same time, they constructed a dataset of music and dance choreographed by four dances (waltz, tango, cha-cha and rumba). Reference [11] proposed a weakly supervised deep recursive method for real-time basic dance generation based on audio power spectrum input. The model uses convolutional layer and LSTM to process the audio input, and then the LSTM layer is used to predict the dance sequence. However, this approach lacks supervision, and this end-to-end coding method employs an adaptive coding configuration, which reduces the accumulation of erroneous feedback in large dance sequences. Reference [12] solved the problem of dance timing through two stages, and a movement detector was developed in the decomposition stage to extract movement beats from dance sequences. Then, it normalized each dance sequence into a series of dance units in time by using the previously extracted dance movement beats. Reference [13] proposed a musical movement GAN (Generative adversarial network). When a given input music piece runs, it first extracts the musical style and beat information, and then produces a dance sequence according to the musical style.

(3) Dance generation algorithm

In recent years, based on the research of "artificial neural network" and "deep learning" algorithms, the underlying logic of artificial intelligence algorithms has been continuously improved and deepened, and artificial intelligence can learn independently and derive similar results without human intervention or minimal intervention, thus being used for choreography [14]. Artificial intelligence dance generation technology is usually based on deep learning to design algorithm models. After training with massive data, it simulates the progressive learning mode of human brain neurons layer by layer to learn more advanced features in the input data, thus improving the accuracy of recognition and derivation, and enabling artificial intelligence to learn choreography after listening to music or watching videos [15]. For example, the generative adversarial network GGAN is a generative model based on deep learning, which generates new data by learning two neural networks: a generator and a discriminator. In dance generation, the GAN can generate new dance moves through the generator, and then use the discriminator to evaluate whether the generated dance moves conform to the characteristics of the original dance data. Through continuous iterative training, the generator can generate dance moves closer to the original dance data [16]. However, due to the strong artistic and personalized characteristics of dance, there is no universal dance data representation method at present. Therefore, different dance generation algorithms often require different data representation methods and training methods [17].

3 NON-PROFESSIONAL DANCE DATA GENERATION

Combined with the idea of data mining, this paper proposes a data generation strategy, which randomly adds irregular disturbances at the time series level and space level to the professional dance sequences in AIST + + data set, synthesizes non-professional dances, and constructs a training set that meets the requirements in terms of scale and choreography complexity.

3.1 Dance dataset

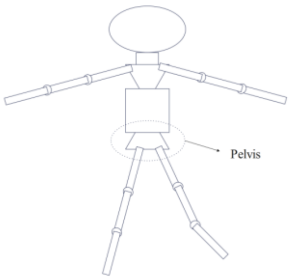

Figure 1 Hierarchical structure of human skeleton at joint points

The industry generally uses the following combinations to describe a human skeleton. First, a stationary skeleton as shown in Figure 1 is defined. The pose of the skeleton is T-shaped, and each child node has a displacement vector relative to its parent node, which is denoted as the initial displacement $O \in R^{J\times 3}$. Secondly, for the root node, its position on the X, Y and Z axes in three-dimensional space is represented by a vector, which is denoted as $Trans \in R^{T\times 3}$. Finally, for the rest of the nodes, the three-dimensional Euler rotation values on the X, Y and Z axes are expressed by vectors, which are denoted as $Rot \in R^{(J-1)\times 3}$. The three-dimensional Euler rotation Rot is obtained by equivalent transformation to obtain the three-dimensional rotation matrix RotMat[18].

$$Rot_{t,j} = \begin{bmatrix} \theta_z & \theta_y & \theta_x \end{bmatrix} \tag{1}$$

$$RotMat_{t,j} = \begin{bmatrix} cos\theta_z & -sin\theta_z & 0 \\ sin\theta_z & cos\theta_z & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} cos\theta_y & 0 & -sin\theta_y \\ 0 & 1 & 0 \\ sin\theta_y & 0 & cos\theta_y \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & cos\theta_x & -sin\theta_x \\ 0 & sin\theta_x & cos\theta_x \end{bmatrix} \tag{2}$$

After the three-dimensional spatial positions of all nodes at all time points are calculated, $P \in R^{T \times J \times 3}$ (T is the number of frames and $J = 21$ is the number of joint points) is obtained, which represents the three-dimensional spatial position of an entire movement sequence under the control of the combination mode.

$$P_{t,j} = \begin{cases} RotMat_{t,j} \cdot O_j + P_{t,j_{parent}} & ifnotroot \\ Trans_t & else \end{cases} \tag{3}$$

This paper proposes a non-professional dance data generation strategy based on key frames. Figure 2 briefly describes the implementation steps of this data generation strategy. When the data generation strategy is implemented, the following three conditions are met. (a) The content of choreography remains unchanged during the data generation process. (b) Timing perturbation factor: it is used to adjust the timing alignment between the movement beat and the music beat. (C) Spatial disturbance factor: It is only used to adjust the completion of the movement.

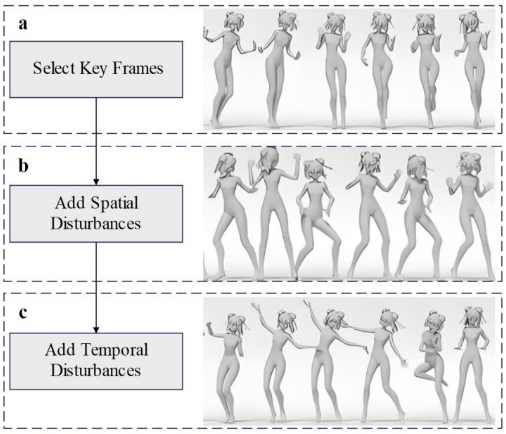

Figure 2 Data generation

3.2 Recognition of dance movement features

When learning to dance, amateur dancers are usually more likely to notice the conspicuous changes in movements (such as pauses and circles), which usually occur with the beat of the movement. Based on this observation, this paper defines the movement beat as the frame sequence number with a large speed change to assist data generation.

In order to obtain the movement beat, the speed $V \in R^{T \times J \times 3}$ of the movement sequence (namely, the first partial derivative of the displacement with respect to time) is first calculated according to the following formula[19]:

$$V = \frac{\partial p}{\partial t} \tag{4}$$

After the velocity at each time is obtained, the angle between the velocity direction at the current time and the previous time is calculated at each time point according to the following formula, and the minimum value (and the maximum value of the direction change) among all joint points is taken as the velocity change at the current time. After all time points are calculated, the change $\hat{V} \in R^T$ of velocity is obtained.

$$\hat{V}_t = \min_j \sum_{axis} \frac{v_{t,j,axis}}{|v_{t,j,axis}|} \cdot \frac{v_{t-1,j,axis}}{|v_{t-1,j,axis}|} axis \in \{x,y,z\} \tag{5}$$

After obtaining the movement beats of the dance sequence, the selection of key frames still needs to go through the following specific steps. First, the time interval t_{pad} (t_{pad} is 3 seconds in the specific implementation) is used to sample evenly over the entire movement sequence, and the frame numbers obtained by sampling are used as the initial key frame numbers. Then, with the initial key frame numbers as the origin, the movement beats closest to them are searched as candidate key frames. Finally, the candidate key frames with a time interval of less than 1 second are discarded to determine the final key frame sequence $K \in R^{N}$. The last step is to ensure the authenticity of the adjusted movement.

Therefore, this strategy defines a spatial factor $S' \in R^{N \times J}$ for the key frame sequence K above, and randomly generates the value of the spatial factor using an approximate inverse Gaussian distribution (formula 6) to control the degree of spatial perturbation at all joint points]:[20]

$$S_n^{'} = tanh\left(S_n^{'} * d\right) * \alpha + \beta \tag{6}$$

Among them, $S_n^{'} : U\left(-3,3\right)$ is the random value of the nth key frame that satisfies the uniform distribution, α and β are parameters used to control the shape of the inverse Gaussian distribution (in the specific implementation, α is 1.1 and β is 1.3, which can ensure that the randomly generated value is between 0.2 and 0.7 and conforms to the actual situation). $d \in \{-1,1\}$ is a randomly generated binary parameter used to control whether to reduce or expand the movement amplitude, $d = 1$ means expansion, and vice versa.

Then, using linear interpolation, $S_n^{'}$ will propagate to the whole spatial factor sequence $S \in R^{T \times J}$ $\left(T > N\right)$, so as to control the spatial disturbance of all frames.

This strategy defines a standard standing movement u, and linearly interpolates each movement frame with it (Equation 8) to obtain the perturbed non-professional dance sequence.

It should be noted that the dance sequences in this paper are all described by local directional expression, and the standard standing posture is no exception. The local direction expression is defined as the direction of the displacement vector of the current node from its parent node. The specific calculation method is to use the position of the current joint point, subtract the position of its parent node, and organize it into a unit vector, as shown in the following formula 7:

$$p_{t,j} = \frac{p_{t,j} - p_{t,j_{parent}}}{\left|p_{t,j} - p_{t,j_{parent}}\right|} \tag{7}$$

$$p'_{t,j} = \left|p_{t,j}\right| \cdot \left(\frac{p_{t,j}}{\left|p_{t,j}\right|} \cdot S_{t,j} + u_j \cdot \left(1 - S_{t,j}\right)\right) \tag{8}$$

For ease of understanding, the above standard standing posture can be simplified into three parts: limbs, spine, and connection points. For the joint points on the limbs, $u_j = \left(0,0,-1\right)$ means that their direction is vertical to the ground downward. For the joint point on the spine, $u_j = \left(0,0,1\right)$ means that its direction is vertical to the ground upward. For the joint points (shoulders and crotch) located at the connection points, they are not affected by the standard standing posture and remain in their original state.

Time disturbance: This data generation strategy will adjust the temporal relationship between professional dance sequences and their accompaniment music to obtain non-professional movement sequences with weak rhythm. Therefore, this strategy defines a time factor $Q' = R^{N}$ for the key frame sequence K above, and randomly generates a value through the inverse Gaussian distribution (Equation 9) to control the time disturbance range of N key frames.

$$Q_n^{'} = tanh\left(q_n^{'}\right) * \alpha + \beta \tag{9}$$

The process of timing perturbation is as follows. First, for each key frame n, its offset frame $Q_n^{'}$ is moved to the new time node $T \in R^N$. When $Q_n^{'}$ is negative, it means forward offset, and when $Q_n^{'}$ is positive, it means backward offset. Secondly, between every two adjacent key frames, linear interpolation is used to calculate the frame sequence number after the timing offset corresponding to each moment, and the frame sequence number sequence $Q \in R^T$ is obtained. Finally, since the sequence of key frames may change after the timing offset, the sequence Q of key frames should be checked to ensure its monotonicity and avoid the inversion of the adjusted movement.

The adjusted movement sequence $P' \in R^{T \times J \times 3}$ (using the local direction expression described above) is calculated from formula 10:

$$p_t^{'} = \left|Q_t - \lfloor Q_t \rfloor\right| \cdot p_{\lceil Q_t \rceil}^{'} + \left|Q_t - \lceil Q_t \rceil\right| \cdot p_{\lfloor Q_t \rfloor}^{'} \tag{10}$$

3.3 Dance sequence professionalism enhancement network

A two-stage framework is designed. The first is the dance-music alignment stage: by adjusting the density of the movements in a local range, the beat is aligned with the music beat. The second is the dance professionalism enhancement stage: for the dance movements that have completed the time sequence alignment, with the support of the decoder and loss function, the movement completion is adjusted, and the choreography is ensured not to change. Figure 3 shows the two-stage structure of the framework.

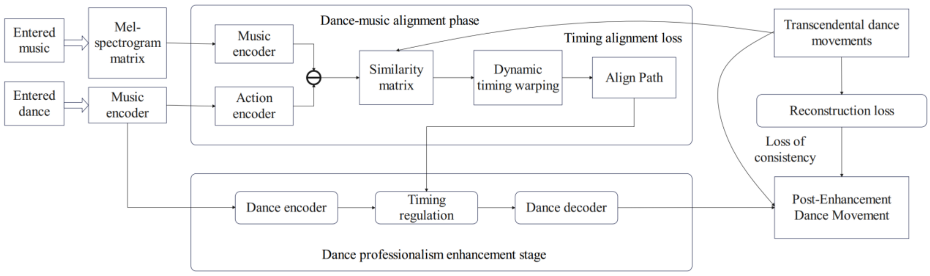

Figure 3 Two-stage dance professionalism enhancement network

The upper blue box is the dance-music alignment stage, and the lower blue box is the dance professionalism enhancement stage.

For the input T-frame music signal, this paper uses librosa audio analysis library to calculate the Mel scale sound spectrum matrix $G \in R^{T \times B}$, where T is the number of music frames and B is the number of sound spectrum channels. For the input T-frame dance movement, this paper first calculates the relevant joint positions, and then calculates the velocity and acceleration of each joint point in each frame in the X, Y, and Z directions, and there is $K \in R^{T \times C}$. Because the velocity and acceleration each have 3 dimensions, $C = J \times (3+3)$.

In this paper, an attention masking matrix (formula 11) is added to the Transformer coding layer to mask irrelevant values in the information processed by the module[21].

$$B_a(i,j) = \begin{cases} 0 & |i-j| < \delta \\ -\infty & otherwise \end{cases} \tag{11}$$

Among them, δ is a parameter used to control the size of the neighborhood (in the specific implementation, $\delta = 50$).

After obtaining the implicit feature sequences f_G and f_K of music and dance respectively, this paper can calculate the similarity of the two implicit feature sequences. For example, in the specific implementation, this paper uses the Euclidean distance (formula 12) to calculate the distance between the implicit feature sequence frames and obtain a similarity matrix of size $T \times T$:

$$F(i,j) = \left| f_G(i) - f_K(j) \right|_2^2 \tag{12}$$

Among them, i is the sequence number of the music frame and j is the sequence number of the movement frame.

In this paper, the dynamic programming idea is used to find the optimal alignment path. The shortest distance D from the starting point (upper left) of each point on the similarity matrix is calculated by formula 13, and the forward point P_a of the shortest path from the starting point to each point is obtained by formula 14. After all points in the matrix are calculated, the shortest path distance from the starting point to the end point is obtained. Then, the optimal alignment path $W \in R^{T \times T}$ is converted into the optimal alignment matrix using formula 15.

$$D_{i,j} = min\left(D_{i-1,j-1}, D_{i,j-1}, D_{i-1,j}\right) + F_{i,j} \tag{13}$$

$$P_{a_{i,j}} = mIndexin\left(D_{i-1,j-1}, D_{i,j-1}, D_{i-1,j}\right) \tag{14}$$

$$W_{i,j} = \begin{cases} 1 & P_{a_{i,j}} \neq 0 \\ 0 & otherwise \end{cases} \tag{15}$$

The optimal alignment matrix is a timing synchronization scheme calculated for the input music and dance.

3.4 Training process and loss function

In this paper, three loss functions are designed for the success of the training process, namely, the synchronization loss function in the dance-music alignment stage, the reconstruction loss function in the dance professionalism enhancement stage and the consistency loss function.

Synchronization loss function: It is known that the professional dance sequence and its corresponding music sequence are synchronized in time series. This paper defines a synchronization loss function on the similarity matrix. The triplet loss of the music feature frame, positive sample and negative sample is calculated as the synchronization loss (Formula 6)[22]:

$$L_{triplet} = \sum_{t}^{T}\left[\left|f_G(t) - f_K\left(\hat{\Phi}(t)\right)\right|_2^2 - \left\|f_G(t) - f_K(r)\right|_2^2 + a\right] \tag{16}$$

Among them, f_G and f_K are music features and dance features, t is the music frame number, r is the randomly selected frame number, and $\hat{\Phi}(t)$ is the dance frame number corresponding to the music frame t.

A reconstruction loss function (formula 17) is defined herein:

$$L_{recon} = \sum_{t}^{T}\sum_{j}^{J}\left|p_{t,j} - \dot{p}_{t,j}\right| \tag{17}$$

Among them, $p_{t,j}$ is the j-th joint point of the t-th frame of the output sequence, and $\dot{p}_{t,j}$ is the j-th joint point of the t-th frame of the transcendental professional dance movement.

Consistency loss function: In order to ensure fluency of the enhanced dance sequence, this paper defines a consistency loss function, which is used to minimize the difference between the output and the prior professional dance in higher dimensions and ensure the fluency of movements. The consistency loss function can be described as (formula 18):

$$L_{cons} = \sum_{t}^{T}\sum_{j}^{J}\left|v_{t,j} - \hat{v}_{t,j}\right| \tag{18}$$

Among them, v and $\hat{v}$ are the velocities of the output sequence and the prior sequence, respectively.

In order to enable the network framework of this paper to deal with sequences with different lengths, this paper firstly uses the method of supplementing 0 after the sequence to unify the length of the input sequence. Then, a key value masking vector M_{kp} is used to mask the useless padding values in the input sequence:

$$M_{kp_i} = \begin{cases} 0 & iforiginsequence \\ 1 & otherwise \end{cases} \tag{19}$$

Two-stage training mode (Figure 4): In order to ensure the training effect, the two-stage frameworks in this chapter are trained separately. In order to enable the dance-music alignment stage to find the optimal time-series alignment matrix, this paper uses a priori alignment matrix $^{1}W_{GT}$ and alignment loss function during training[23].

Hyperparameter setting: After continuous tuning, this paper finally determines that the number of samples in each batch (batch size) is 64, the learning rate is 0.0001, and Adam Optimizer 2 is used for backward propagation.

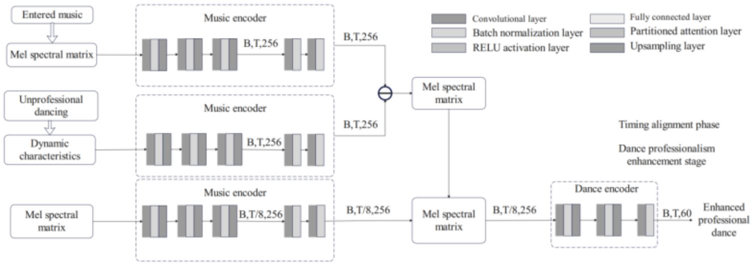

Figure 4 Network details of the two-stage framework

4 SYSTEM DESIGN AND TESTING

4.1 System design

The application system of 4-layer structure constitutes the intelligent choreography and display system of dance in this paper, as shown in Figure 5. The access layer is a user interaction platform based on Unity. The second layer is an interface layer based on Nginx, uWSGI and Django, which mainly handles the interaction requests of the access layer. The logic layer mainly distributes programmed functional modules. The last layer is the data layer based on MySQL. The user directly faces Unity in the access layer. Unity sends the user request to Django, Django processes the business logic and sends the generated result to uWSGI, and uWSGI processes the request sent by the browser and returns the response result. Nginx not only acts as a relay server, but also as a reverse proxy server to accept connection requests from clients with different network addresses and forward them to the server, and return the results processed by the server to the client.

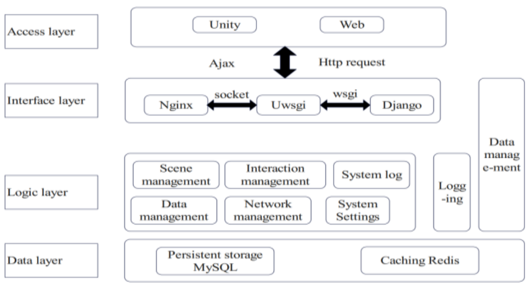

Figure 5 System hierarchical structure design

According to the principle of low coupling in software design, the functional modules are designed as scene management module, network management module, interaction management module and data management module, as shown in Figure 6.

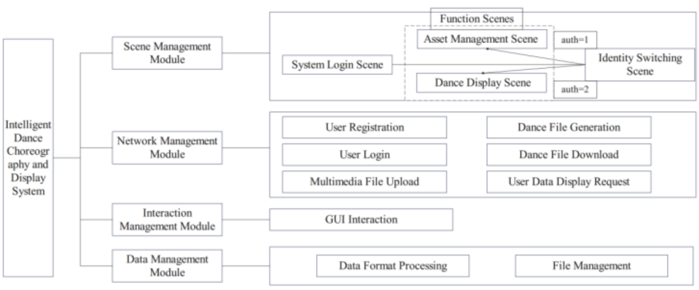

Figure 6 Functional modules of intelligent dance choreography and display system

The intelligent choreography method of this system is based on the directed graph neural network model, including three steps. First, music and movement features are extracted, and a 1D-CNN-based music classifier is trained to determine the input music type. Then, the music type is fused into the music features and encoded as a latent vector. Finally, the latent vector and movement features are input into the improved time-series directed graph neural network and the spatial directed graph neural network as control signals, and the movement prediction module outputs the predicted movement sequence. The overall framework is shown in Figure 7.

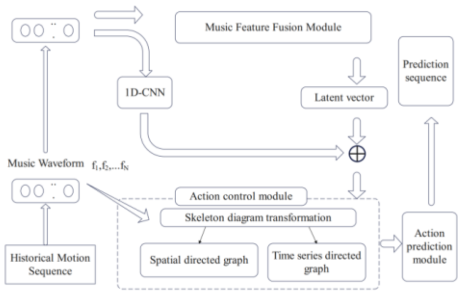

Figure 7 Network structure of intelligent choreography method

4.2 Results

The purpose of the experiment is to extract the speed features of the arm skeleton movements and evaluate the effect of the feature extraction algorithm through the visual effects of the movement clips. The movement screening is performed based on the average arm speed of each movement clip, as shown in Figure 8.

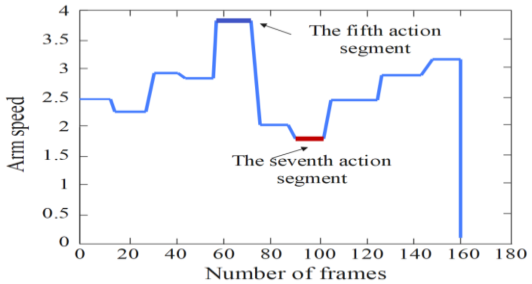

Figure 8 Average arm velocity of choreographed movement segments

The spatial measurement of each movement segment in this experiment is shown in Figure 9. In order to verify whether the judgment is accurate, the movement paths projected by the root nodes of the two movement segments on the ground are drawn respectively, as shown in Figure 10, where the blue path corresponds to the first movement segment and the red path corresponds to the first movement segment.

Figure 9 Spatial measurement of movement segments

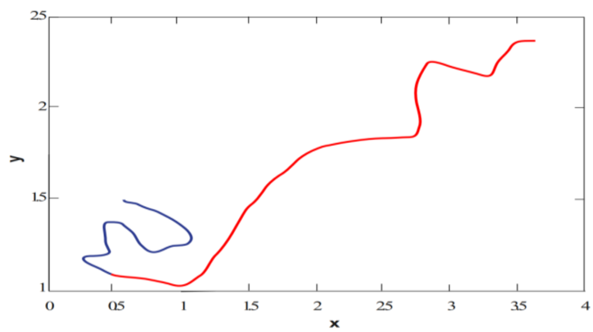

Figure 10 Movement path of root nodes

User ratings are used to analyze the dance synthesis effects of three styles. The style of the music is analyzed according to the overall characteristics of the target music (BPM and varying note duration), and the corresponding choreographic movements are generated. In this experiment, several target music pieces are analyzed, and three target music pieces suitable for generating hip-hop dance, house dance and modern dance are selected and choreographed. The test users judge the dance styles of the three segments respectively, and evaluate the matching degree between music and dance, the coherence and authenticity of dance movements. The results are shown in Table 1 and Figure 11.

Table 1 Manual scoring of three styles of dance

Dance style	Matching degree	Coherence	Authenticity	Style judgment calculation accuracy
Street dance	4.41±0.09	4.19±0.13	4.14±0.08	97.23%
House dance	4.37±0.15	4.08±0.14	4.05±0.11	97.15%
Modern dance	4.35±0.12	4.10±0.07	4.11±0.13	99.80%

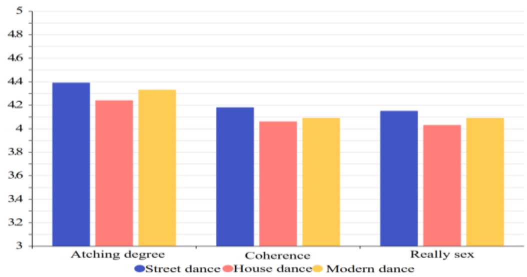

Figure 11 Comparison of scores of different types of dances

4.3 Analysis and discussion

In this paper, an intelligent choreography and display system is designed and implemented. Based on a new intelligent choreography method, this method uses improved time series directed graph and spatial directed graph to learn movement features and music features respectively, and intelligently choreographs three-dimensional dance sequences that are visually real and match the music beat by arbitrary input music. At the same time, the optical motion capture system is used to collect two kinds of music and dance movement capture data, and a high-quality music and dance data set is constructed. In terms of system design, this paper starts with the actual needs of asset management, analyzes and studies the needs of users and the functions of the system, and adopts an independent modular development model with a layered architecture to make the program designed in this paper easier to browse, correct, reuse and supplement. Finally, in order to meet the requirements of basic functions, the user interaction experience is optimized.

As shown in Figure 8, the fifth movement segment has the largest average arm moving speed, and the seventh movement segment has the smallest average arm moving speed.

As shown in Figure 9, according to the numerical judgment, the spatiality of the first movement segment should be weak, and the spatiality of the second movement segment should be stronger. By observing Figure 10, it can be found that the motion trajectory range of the second movement segment is indeed larger, indicating that the above judgment is accurate, and the movement segment can be spatially described by the spatial measurement of the movement segment, so the spatial feature extraction algorithm proposed in this paper is effective.

From Table 1 and Figure 11, it can be seen that the evaluation results of our model on the actual matching of different styles of dance and music are very good, which verifies the evaluation effect of our model and can play an important role in the analysis of dance fluency. Moreover, according to the user evaluation results, our model has received good feedback in the user group. And it can achieve reliable intelligent choreography effects in dance choreography, which has a certain promoting significance for the intelligent development of dance

From the above analysis, we can see that the dance choreography method combined with feature extraction proposed in this paper can effectively identify various dance features. In particular, it can learn from massive dance features, effectively improve the quality of dance choreography, and promote the further development of subsequent dance choreography.

5 CONCLUSION

In this paper, an intelligent choreography and display system is designed and implemented. Based on a new intelligent choreography method, this method uses improved time series directed graph and spatial directed graph to learn movement features and music features respectively, and proposes a key frame-based non-professional dance data generation strategy. Moreover, this paper designs a two-stage framework. The first stage is the dance-music alignment stage, which adjusts the density of the movements in a local range to align the beat with the music beat. The second stage is the dance professionalism enhancement stage, which adjusts the movement completion of the dance movements that have completed the time sequence alignment with the support of the decoder and loss function, and ensures that the choreography is not changed. It can learn from massive dance features, effectively improve the quality of dance choreography, and promote the further development of subsequent dance choreography.

The model in this paper applies big data analysis and learning ideas. Therefore, it is necessary to further promote the establishment of large-scale dance datasets, improve data expansion methods, obtain richer and more realistic non-professional dance datasets, or optimize the motion capture process technically to obtain more realistic dance data.

ACKNOWLEDGE:

General Project of Philosophy and Social Sciences in Jiangsu Province's Universities《Research on the integration and construction of " three systems " of preschool education dance under the action of aesthetic education infiltration》（2024SJYB1383）

REFERENCES

[1] Jiang, H., & Yan, Y. (2024). Sensor based dance coherent action generation model using deep learning framework. Scalable Computing: Practice and Experience, 25(2), 1073-1090.

[2] Pang, Y., & Niu, Y. (2023). Dance video motion recognition based on computer vision and image processing. Applied Artificial Intelligence, 37(1), 2226962-2226970.

[3] Luo, W., & Ning, B. (2022). High-Dynamic Dance Motion Recognition Method Based on Video Visual Analysis. Scientific Programming, 2022(1), 6724892-6724899.

[4] Ma, M., Sun, S., & Gao, Y. (2022). Data-Driven Computer Choreography Based on Kinect and 3D Technology. Scientific Programming, 2022(1), 2352024-2352035.

[5] Sun, D., & Wang, G. (2024). Deep learning driven multi-scale spatiotemporal fusion dance spectrum generation network: A method based on human pose fusion. Alexandria Engineering Journal, 107(2), 634-642.

[6] Bhuyan, H., Killi, J., Dash, J. K., Das, P. P., & Paul, S. (2022). Motion recognition in Bharatanatyam dance. IEEE Access, 10(1), 67128-67139.

[7] Ji, Z., & Tian, Y. (2024). IoT based dance movement recognition model based on deep learning framework. Scalable Computing: Practice and Experience, 25(2), 1091-1106.

[8] Ding, S., Hou, X., Liu, Y., Zhu, W., Fang, D., Fan, Y., ... & Zhou, F. (2023). DanceTrend: An Integration Framework of Video-Based Body Action Recognition and Color Space Features for Dance Popularity Prediction. Electronics, 12(22), 4696-4704.

[9] Kritsis, K., Gkiokas, A., Pikrakis, A., & Katsouros, V. (2022). Danceconv: Dance motion generation with convolutional networks. IEEE Access, 10(2), 44982-45000.

[10] Ma, X., & Wang, K. (2022). Dance action generation model based on recurrent neural network. Mathematical Problems in Engineering, 2022(1), 8455961-8455970.

[11] Shao, H. (2022). Recognition and Prediction Method of Music and Dance Movements Based on G-ResNet SPP and Attention Mechanism. Mobile Information Systems, 2022(1), 1763262-1763272.

[12] Aristidou, A., Yiannakidis, A., Aberman, K., Cohen-Or, D., Shamir, A., & Chrysanthou, Y. (2022). Rhythm is a dancer: Music-driven motion synthesis with global structure. IEEE Transactions on Visualization and Computer Graphics, 29(8), 3519-3534.

[13] Geng, J. (2022). The collection and recognition method of music and dance movement based on intelligent sensor. Computational Intelligence and Neuroscience, 2022(1), 2654892-2654903.

[14] Hung, N. V., Loi, T. Q., Binh, N. H., Nga, N. T. T., Huong, T. T., & Luu, D. L. (2024). Building an online learning model through a dance recognition video based on deep learning. Информатика и автоматизация, 23(1), 101-128.

[15] Lan, J. (2022). Research on Automatic Dance Generation System Based on Deep Learning. Mathematical Problems in Engineering, 2022(1), 6366507-6366518.

[16] Zhang, A. (2023). Optimization Simulation of Match between Technical Actions and Music of National Dance Based on Deep Learning. Mobile Information Systems, 2023(1), 1784848-1784860.

[17] Wang, S., & Tong, S. (2022). Analysis of high-level dance movements under deep learning and internet of things. The Journal of Supercomputing, 78(12), 14294-14316.

[18] Odefunso, A. E., Bravo, E. G., & Chen, Y. V. (2022). Traditional African dances preservation using deep learning techniques. Proceedings of the ACM on Computer Graphics and Interactive Techniques, 5(4), 1-11.

[19] Yuan, X., & Pan, P. (2022). Research on the Evaluation Model of Dance Movement Recognition and Automatic Generation Based on Long Short-Term Memory. Mathematical Problems in Engineering, 2022(1), 6405903-6405912.

[20] Dong, L. (2022). Optimization Simulation of Dance Technical Movements and Music Matching Based on Multifeature Fusion. Computational Intelligence and Neuroscience, 2022(1), 8679748-8679759.

[21] Wang, J. (2024). Research on dance teaching methods based on emotion recognition and emotion analysis. Journal of Electrical Systems, 20(3), 1023-1036.

[22] Kim, T. (2024). Multi-Modal Beat Alignment Transformers for Dance Quality Assessment Framework. Journal of Multimedia Information System, 11(2), 149-156.

[23] Zhao, X. (2024). Emotion Analysis and Expression Algorithm of Dance Action Based on Machine Learning. Journal of Electrical Systems, 20(6s), 1468-1481.