Research Article

Research on the Application of Hybrid Density Network Combined with Gaussian Model in Computer Music Choreography

Li-Juan Wang (ID)

Academy of Music, Jilin Normal University, Siping, Jilin 136000, China

Correspondence should be addressed to Li-Juan Wang; 13134463205@163.com

Dance, as a unique form of expression, is usually accompanied by music and presented to the audience visually, improving people's cultural and spiritual lives while also strengthening their creative energy. And dance choreography is usually created by a few skilled choreographers, either individually or together, with a high level of expertise and complexity. With the introduction of motion capture technology and artificial intelligence, computers can now do autonomous choreography based on music, and science and technology are changing the way artists produce art today. Computer music choreography must solve two fundamental issues: how to create realistic and creative dance moves without relying on motion capture and manual creation and how to improve music and dance synchronization utilizing appropriate music and movement elements and matching algorithms. This article employs a hybrid density network to generate dances that fit the target music in three steps, action generation, action screening, and feature matching, to address the aforementioned two concerns.

1. Introduction

Color role synthesis techniques are used extensively in the creation and postretouching of virtual characters in computer games, advertisements, and film productions in recent years. Virtual character action drawing techniques are a crucial part of computer graphics because they allow virtual characters to mimic the actions of real humans. They are an extremely active research area. The invention and widespread usage of motion capture technology has secured the authenticity and safety of the action, but it is still merely a duplicate of the data. In games, animation, virtual reality, and a variety of other applications, people have a strong desire to engage with virtual characters who seem like humans. What kinds of creative activities, such as dancing, await the dance movements of virtual characters, particularly user-made dance animation, in which the animator must manually alter the position and rotation of each bone in the model in a critical position frame? Completing this job not only takes time, but it also necessitates a high level of talent on the part of the animator, restricting the scope of virtual character dance animation [1]. As a result, a successful dance synthesis algorithm technique can be applied to a variety of sectors, such as music-assisted dance instruction, video game character movement generation, human behavior research, and virtual reality [2]. Based on the foregoing, this paper proposes a novel solution as follows: an automatic music choreography algorithm that uses dance data, a deep learning algorithm to train the training model, and a combination of filtering conditions to generate the training model automatically and intelligently to meet the expected dance movements and arrange the dance based on the matching of music and action clips. The algorithm can be modified to include unique and original dance motions, effectively replacing the traditional choreography algorithm and offering real value [3]. The algorithm is

primarily designed for 3D computer animation characters and video game roles. Animation synthesis, virtual reality, dance instruction, and other fields of interest are also important.

2. Related Concept Elaboration

2.1. Research Background. In recent years, China's economic development has accelerated, and residents' disposable income has increased year after year, propelling the rapid growth of tourism in tourist destinations. However, problems such as irregular tourism industry operation and low tourist satisfaction in tourist destinations remain fundamentally unsolved. So the direction and focus of future tourism research will be on how to establish a good tourism environment, maintain sustainable tourism development, deliver positive economic advantages, improve tourist travel experiences, and enrich the cultural life of tourist destination residents [4].

2.2. Expressions. Dance, as a unique form of expression, is typically accompanied by music and presented to the audience in a visual manner, enriching people's cultural and spiritual life while also encouraging their creative urge. Professional choreographers work alone or in teams to develop dance choreography that is both professional and complex. With the introduction of motion capture technology and artificial intelligence, computers can now do autonomous choreography based on music, and science and technology are changing the way artists produce art today. First is how to generate realistic dance movements without using motion capture; second is how to use appropriate music and movement features and then match algorithms to improve the synchronization of music and dance [5]. In three processes, motion creation, motion screening, and feature matching, a hybrid density network is used to generate dances that fit the target song. The dances generated in this research have improved in terms of movement coherence and realism when compared to previous investigations. Subjective user ratings suggest that the choreographic outcomes in this study match the music to a higher extent, based on the hybrid density network action generation algorithm [6]. This paper covers the steps for developing the music action dataset, the action classification premise, and the feature representation of the training data. The action generation model used in this paper's model structure, as well as parameter selections in the model training and prediction technique, is then discussed [7]. The parameter management approach and the coherence-based action filtering algorithm are described in detail when using the model for action generation, and tests are built to evaluate the algorithm and provide the experimental findings. Musical and movement components are included into the choreography. A multilevel music and movement feature matching method is developed to choreograph the generated dance motions based on the qualities of the target music. The whole music feature extraction system, including the BPM and matching algorithms, is first shown. The rhythm and intensity feature extraction technique, data segmentation and feature matching algorithm, and movement connection algorithm are then described as part of the dance synthesis method [8]. The appropriate experiments are then designed to verify the choreography results, and the experiment assessment criteria are introduced, as well as the experimental outcomes.

2.3. Research Model. Instead of relying on the user's manual production and motion capture data, the motion generation problem must be solved in order to design an effective computer choreography algorithm that ensures that the choreography is realistic and novel enough. This chapter implements the motion generating technique based on a hybrid density network [9]. To begin, we generate the music-action dataset, categorize the data, and use features to characterize the training data. This chapter builds the action generation model from the sequence generation model outlined in Chapter 2, finishes the model's training and action generation, and performs parameter control during the action creation process [10].This chapter chooses generated movement sequences based on coherence and gives a candidate movement library for the next choreography in order to ensure the quality of the generated movements and make them appropriate for further choreography.

In this paper, the BPMs of the corresponding music for each type of dance in the dataset are extracted, and the mean and the most values are shown in Table 1. From Table 1, the average BPM of street dance is 150.83; the average BPM of folk dance is 120.62. The maximum values of BPM for street dance and folk dance were 176.17 and 142.39, respectively, and the minimum values were 100.64 and 86.45, respectively. The skewness of BPM for both street dance and folk dance is less than 0, indicating that the distribution of these two indices is left-skewed, and the kurtosis of both is less than 3. The concomitant probability value of JB-statistic for all variables is 0.0000, which indicates that the original hypothesis of "the series obeys normal distribution" is rejected at 1% significance level for all variables of JB-test; therefore, the serial variables of BPM for street dance and folk dance do not obey normal distribution. It can be found that for both dance styles, the mean value of BPM for street dance is larger, and the mean value for folk dance is the smallest. Looking at the dance movements of each type, it can be found that the overall speed of street dance movements is faster than that of folk dance. In other words, the faster dance should have a larger mean value of BPM, which is consistent with the BPM of different dance speeds and with the intuitive audiovisual perception. However, the BPM range is not concentrated, and the feature value alone is not sufficient to describe the overall characteristics of the music.

2.4. Objectives and Algorithm Performance. The goals of this study are to develop a neural network-based, music-driven, and computer-automated choreography that improves the novelty and coherence of the generated dance movements, improves the harmony between the generated movements and the music, and allows users to control the generated dance results according to their preferences. The study focuses on datasets of architectural music movement.

TABLE 1: Descriptive statistical test of BPM sequences of street dance and folk dance.

Variables	Mean value	Maximum value	Minimum value	Standard deviation	Skewness	JB statistic	Probability
Street dance	150.83	176.17	100.64	0.2656	-0.6143	583.1392	0.0000
Folk dance	120.62	142.39	86.45	0.2171	-0.3720	266.8481	0.0000

Although there are some publicly available motion datasets, the majority of them are sports data, such as running and playing ball. Dance movement data that is accompanied by music is rare. Action and music data are particularly essential types of training in deep learning-based studies on the association between music and dance movements [11]. As a result, the VMD action file and WAV music file retrieved from the network and its supporting files with built-in music-dance movement dataset of 192 segments, 1057344 frames, and approximately 587 frames are used in this research. A new strategy is proposed in the generation process of actions, parameter control algorithm, and its application of consistency-based action filtering algorithm, with the goal of improving the authenticity and consistency of created actions and ensuring the consistency of action quality. The iliac bone position assesses the coherence of the action based on the rate of velocity change of each joint in neighboring frames during action production, and the average output of a Gaussian model utilizing a hybrid density network is employed as the skeleton during action filtering [12]. When compared to other control methods, the averaging method produces more realistic movements. The generated raw action data was substantially more consistent than the on-screen action data. Suggestions include integrating overall feature matching with local feature matching to create a multilevel music and activity feature matching algorithm [13], as shown in Table 1 and Figure 1.

This paper argues that the more notes a piece of music contains in an average bar, namely, the more frequent note changes, the richer the musicality of the music, and the more varied the corresponding dance movements should be. Based on this idea, this paper calculates the average duration of notes (frames) of the whole piece. The total number of frames of music signal is denoted as the total number of changing notes [14]. The specific steps of the algorithm are as follows: $X(n)$ is obtained by CQT transformation of music signal $X(r \bigcirc)$, which represents the frequency amplitude of music signal in each semitone of each frame. The frequency amplitude of the KTH semitone of music signal in f frame is represented. The stability test of each asset price is carried out by ADF unit root test.

When the target music is entered, the overall characteristics are extracted first to determine the dance style and dance speed most likely corresponding to the target music, such as BPM and the comparison table of the average duration of changing notes and the values listed in Table 2, and then, the corresponding movement generation model is selected to generate movements [15].

This paper believes that the overall speed of music should be positively correlated with the speed of movement, and the speed of movement can be measured by the speed of obvious local body parts (arms, legs, etc.) in addition to the overall movement of the whole body [16].

In addition, as can be seen from the above, the shorter the average duration of changing notes is, it means that the notes are changing, as shown in Table 2 and Figure 2.

3. Empirical Analysis

There are two approaches to music segmentation: one is that music structure is composed of several repeated patterns, and music can be segmented by extracting repeated patterns; another approach is to take a fixed rhythmic length of the segment as a musical segment [16]. For the first idea, the pattern of repetition may be represented by the repetition of the same passage by different instruments, so the method of structural analysis should depend on the order of the notes and should not be influenced by the timbre of the instrument or the voice. But this idea is difficult in practice [17].

This is because the timbre of each musical instrument has a basic feature, which is always composed of the keynote and its overtone (the frequency is an integer multiple of the fundamental frequency) [18]. When different musical instruments play the same note, although it is basically the same, the energy distribution of overtone is different, so it is difficult to directly extract the accurate repetition pattern in the frequency domain. Second, not all musical melodies have strict repeating patterns, and even if they do, they may be far away from each other [19]. Therefore, the length difference of music fragments after segmentation may be great, which is not conducive to the matching of subsequent action segments. For the second idea, it is actually more in line with the reality. For example, in dance teaching, music and movement are often segmented with several eight-beat rhythm lengths. The length of music fragments obtained by this segmentation method is more uniform, which is convenient for matching and choreography of subsequent music and movements. According to this idea, when music is segmented, the metronomic period is firstly extracted, the metronomic length of music Tmaz is estimated, and music is segmented according to several lengths [20].

On the other hand, it is generally believed that for the music with faster rhythm and speed, the corresponding choreography movement changes more quickly, and the duration of music sections is also shorter in the process of choreography and dance teaching. We already know the speed marker of BPM songs; BPM is inversely proportional to the length of BPM songs [21], as shown in Table 3 and Figure 3.

3.1. Matching Analysis. The outputs of each synthesized dance style are pleasing to users, demonstrating that the music choreography algorithm described in this research is effective. Both types of dancing have mean values of
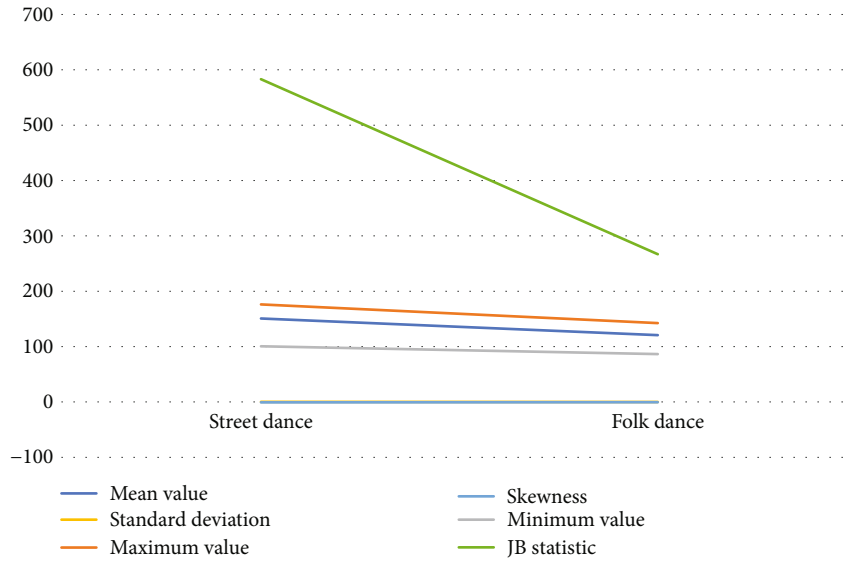
Figure 1: Descriptive statistical test of BPM sequences of street dance and folk dance.

Table 2: BPM and variation note sequential ADF unit root tests.

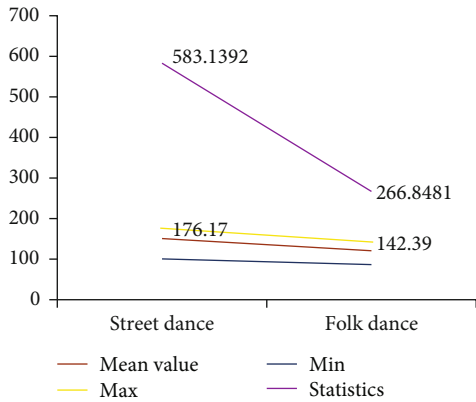| Variable	No intercept and no trend		With intercept		With intercept and trend	
	T-statistic	P values	T-statistic	P values	T-statistic	P values
	2.037026	0.9905	1.558638	0.5038	1.186690	0.9122
	0.311912	0.7758	1.796122	0.3829	1.619738	0.7854
	57.51557	0.0001	57.57892	0.0001	57.59223	0.0000
	75.32716	0.0001	75.32366	0.0001	75.32531	0.0001

Figure 2: BPM and variation note sequential ADF unit root tests.

coherence, authenticity, and degree of matching with music that are higher than the score, showing that consumers are satisfied with the synthesized outcomes. The three indexes of street dance have the greatest ratings, as seen in the observation graph. The rhythm of street dance style movements is evident, and the amplitude of the movements is larger, whereas the amplitude of folk dance movements is often tiny, and it might be difficult for users to distinguish between a little movement and jitter data, affecting perception [22]. The poor synthesis findings of otaku dance compared to street dancing are due to the greater diversity of otaku dance actions and lower concentration of actions, which makes it difficult to train and learn the action generation model in the action dataset developed in Chapter 3 of this research. Only one of the 42 users who took part in the scoring misjudged the styles of street dance and otaku dance, and the rest of the users' assessments were correct, as indicated by the optimal lag order test and cointegration test following the follow-up visit.

From the test results of maximum lag order 4, it can be seen that LR, FPE, AIC, SC, and HQ show that the optimal lag order is 3. According to the majority principle, 3 is chosen as the optimal lag order, and the VAR(3) model is established [23], as shown in Tables 4 and 5 and Figure 4.

As can be seen from Table 4, the original hypothesis of "no cointegration vector" cannot be rejected at the 10% level for either the trace statistic or the maximum eigenroot statistic, so the London Brent crude oil futures and gold future price series are not cointegrated [24].

From the JJ cointegration test, it can be seen that there is no cointegration relationship, so the VEC model cannot be established, so the VAR model should be established after smoothing the variables differentially. Since the difference variables dlgf and dllco are both smoothed, a VAR model can be established. The model is $dlgf(t) = 0.000178 - 0.113479dlgf(t-1) - 0.091914dlgf(t-2) - 0.004051dllco(t-1) + 0.005440dllco(t-2) + \varepsilon 1t$.

TABLE 3: Comparison of choreographic changes of different dance styles.

Dance style	Music and dance matching	Coherence	Authenticity	Accuracy of style judgment
Street dance	4.39	4.18	4.15	97%
Folk dance	4.29	4.06	4.05	96%

4. User Study

A multilevel choreography algorithm based on music and movement characteristics is proposed in this chapter, with the goal of improving the ensemble of music and choreographic motions. Music overall feature extraction and matching, local rhythm and intensity feature extraction and matching, and intermediate frame interpolation processes make up the algorithm. Based on the visual effect of the synthesized dance, the choreographic outcomes are visually assessed in order to analyze the algorithm's effectiveness and the effect of the synthesized dance. Experiment 6 is utilized to test if the hierarchical music and action feature matching method is effective; the dance synthesis impact is studied to see if the music and action feature matching algorithm based on rhythm and intensity features is effective [25].

Visual effects in qualitative studies can verify the algorithm's performance, but this indicator alone does not provide a complete quantitative evaluation of the trial outcomes. In the field of computer-assisted music choreography, it is difficult to scientifically and quantitatively analyze the choreographic effect, and there is no uniform objective and quantitative evaluation index, so subjective evaluation criteria are frequently used to examine the experimental results [26]. In this paper, 35 graduate students were invited to conduct a user experience study using a manual user rating method. In order to ensure that the users have certain aesthetic cognition of music and dance and sufficient musical sense and to ensure the reliability of the rating, a user ability test experiment was designed during the questionnaire survey. In the user ability test experiment, two musical dances from the training dataset were presented to the participating testers: one with music that matched the dance and the other with the same music that matched the other unmatched dance. The participants were asked to select the mismatched dance segments, and only those who selected the correct ones had enough musical sense to perceive the rhythm of the music and the movements and to judge the degree of matching the music and the movements. A novel framework for automatic music choreography is provided in this research in order to generate dance motions that are both novel and coherent and match the target music, as well as to ensure that the choreography system has appropriate robustness and generalization capacity. The framework is divided into four sections: model training and movement creation, choreography and synthesis, and dance visualization utilizing 3D character animation are the most critical steps, with model training and movement generation and choreography based on music and movement features being the most important [27], as shown in Figure 5.

In this study, we present a parameter management approach and a coherence-based motion filtering strategy to increase the authenticity and coherence of the generated motions. According to the experimental results, the mean value technique boosts the realism of the manufactured actions, and the coherence of the filtered action data is considerably improved when compared to the generated original action data. With the purpose of strengthening music and action unity and coherence, this work proposes a multilevel music and action feature matching method that combines overall feature matching with local feature matching. To match dancing movements, overall music features are used first, followed by rhythm and intensity characteristics to match local music-movement fragment features. When a control based on overall music characteristics is added to the final synthesis result, the speed and other qualities of each movement fragment are more consistent, and the entire choreography is more aesthetic [28]. In this paper, we look at the complete process of computer music choreography and propose a computer music choreography framework that provides a fresh solution to the problem. The framework includes a movement dataset construction module and a model training module.

The framework is made up of four modules: movement dataset building, model training and movement generation, dance choreography and synthesis, and 3D character animation visualization, all of which ensure the authenticity, uniqueness, and compatibility with the music of the dance fragments. Compared with the traditional algorithm, the synthesized choreographies are more novel and diverse, and the choreography system in this paper has stronger stability and generalization ability than the music-movement mapping model obtained by machine learning algorithm only. In addition, the user's requirements can be reflected in the choreography results, which makes the proposed choreography system more practical.

Although the algorithm proposed in this paper can obtain better computer music choreography effect and ensure the novelty and consistency of the choreographed movements as well as the conformity with the target music, the research on the direction of feature extraction, feature matching, and dance evaluation of music and movements is not sufficient. The shortcomings of this paper and the directions of subsequent work are mainly as follows: the existing research has less research on action high-level features, and the screening algorithm based on action features proposed in this paper mainly uses action bottom-level features. In future research, we can try to analyze the high-level features of the action; when measuring the match between local music and action, only the rhythm and intensity features of both are considered. In future studies, we can
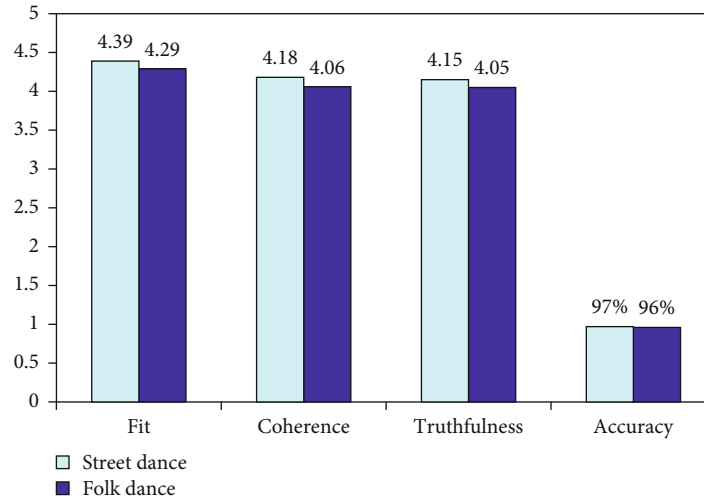
FIGURE 3: Comparison of choreographic changes of different dance styles.

TABLE 4: Optimal lag order test.

Lag	LogL	LR	FPE	AIC	SC	HQ
0	2496.306	NA	0.001263	-0.998521	-0.995912	-0.997607
1	34872.66	64713.82	2.97e-09	-13.95783	-13.95000	-13.95509
2	34909.9	74.41330	2.93e-09	-13.97114	-13.95809	-13.96657
3	34931.41	42.96391*	2.91e-09*	-13.97815*	-13.95989*	-13.97175*
4	34934.62	6.403435	2.91e-09	-13.97783	-13.95435	-13.96960

TABLE 5: JJ cointegration test.

Original hypothesis	Trace statistic (probability)	Maximum characteristic root statistic (probability)
No covariance vector	8.424707	5.976167
	(0.4212)	(0.6163)
There is at most one cointegrating vector	2.448539	2.448539
	(0.1176)	(0.1176)

try to include other more abstract features common to music and movement, such as emotion and style; there is not enough research on methods to evaluate the effect of music choreography. In the existing studies, the effect of dance synthesis is often measured by visual effects, such as subjective ratings by professionals on snapshots of dance movements and synthesized dance videos. A common objective quantitative index for the assessment of the effect of music choreography has not yet been proposed, and further research is needed.

The goal of combining feature matching is to reduce the influence of the target song's overall qualities on choreography while also improving the quality, consistency, and harmony of music and movement. First, the target music's overall note density and beat per minute (BPM) of note density are extracted using a constant Q-transform for initial matching with features like movement speed, and then, local music and movement fragment data are matched using rhythm and intensity. The multilevel feature matching algorithm is used to construct more thematic dances by synthesizing dance movement fragment sequences with more uniform features such as pace. A new solution is proposed by combining the hybrid density network-based motion creation algorithm with the multilevel music motion feature matching algorithm user control. A computer music choreography framework is designed to present a fresh notion for solving the challenge of computer music choreography. The framework, which includes a movement dataset construction module, a model training and movement generation module, a dance choreography and synthesis module, and a 3D character animation visualization module, can ensure the authenticity, novelty, and musical harmony of dance clips at the same time. Only action mapping is based on the projection model, which offers great stability and generalization ability when compared to music produced by deep neural networks. In addition, in the choreography
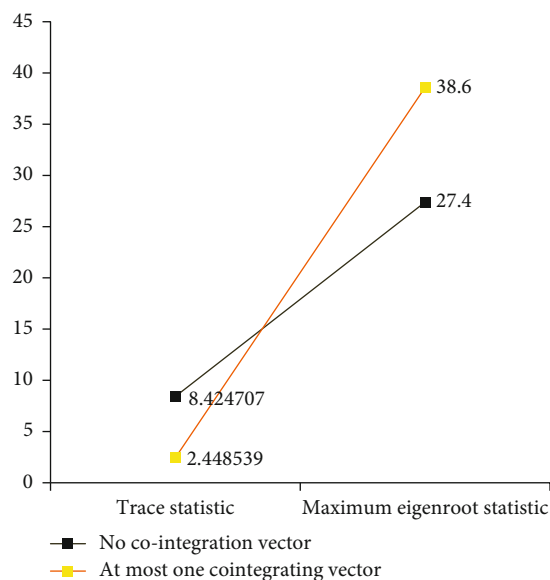
FIGURE 4: JJ cointegration test.

module, the user can control the movement speed and dance space of the local skeleton so that it can still dance with automatic computer choreography; in order to control the choreography results according to preference, this framework is more useful.

In order to improve the authenticity and diversity of computer-automated music choreography, this chapter proposes a motion generation algorithm based on a hybrid secret method degree network. First, it introduces that the construction of the music motion dataset is done by manual labeling based on the classification, and the motion data is preprocessed to obtain the feature representation of the model training data. Then, the action building is completed, and the generative model is trained to compare the effect of action generation under different training time. Then, three parameter control methods are introduced for the actions in the generation process, and the generated original actions are filtered based on the action consistency to ensure the consistency and authenticity of the validation data. The algorithm can generate sufficiently realistic and diversified dance moves, according to the findings of the experiments. The average method produces the most stable movements; as training time goes on, the human skeletal structure of the generated movements becomes more realistic, and the relative relationship between joints becomes more stable; the consistency-based movement filtering algorithm also achieves the expected effect, and the movement generation algorithm introduced in this chapter produces the best results for subsequent music and movement feasibilities, as shown in Figure 6.

The system can be divided into three phases: model training, motion generation, and music arrangement. The system uses prebuilt datasets to train the motion generation models, obtains motion generation models of different styles and speeds, and stores the models with optimal results. Before music choreography, the system uses various action

generation models to generate a certain number of action clips and builds various candidate action databases through coherence filtering. Users can choreograph music directly on this basis or generate their own candidate movements and then choreograph music. It is a flowchart for the use of model training, a flowchart for the use of action generation and screening, and a flowchart for music choreography based on a multilevel music and action feature matching algorithm. In addition, the system introduces user control in the steps of the overall music feature matching algorithm, provides a user control interface, and gives two optional methods: one is to use the system default parameters for action matching, and the other is for the user to set the local bone velocity threshold and spatial threshold of the action segment and then perform action matching. In other words, the system can adjust the matching results of music and movements according to the user's requirements for dance characteristics, thus adjusting the choreography results for targeted control according to the user's preferences. If the user wants the final choreography to have distinct movement characteristics of local body parts (arms, legs, etc.), he can set a higher bone speed threshold and order to match movement segments with bone speed greater than the threshold, as shown in Figure 7.

The feature extraction and matching algorithm of the music and action of the previous festival has obtained multiple action segments with the matching result of the target music, and the connectivity constraints are satisfied between adjacent segments. Based on this, this section will analyze the adjacent power responses to produce transition connection segments, solve the action mutation problem, and splice the action segments into a series of complete actions to complete the final dance arrangement. The action mutation discussed in this section pertains to the mutation generated by the action segment pony and its nearby actions; the segment is known as the action at the connection, and the fixed distance will cause incoherence of the immediately connected actions, impacting the dance's visual appearance. The intermediate frame interpolation algorithm is utilized in this section to interpolate K frames and action segments at the conclusion of action segment m based on the interpolation weights and to perform interpolation between the intermediate actions of n to generate the final interpolated action. This algorithm's interpolation action allows for a natural link of two action segments. Simultaneously, some of the properties of the prior action segment that ended the action are kept. The interpolated motions maintain movement continuity, but there may be unrealistic movements, such as foot sliding, and the K value should not be too high to prevent the interpolated movements from continuing too long and damaging the appearance.

We present a multilevel choreography algorithm based on music and action features in this research, with the goal of enhancing the harmony of efficient actions in music and choreography. The method includes steps for music and movement feature extraction and matching, local rhythm and intensity feature extraction, matching, and intermediate frame interpolation in order to assess the algorithm's effectiveness and the effect of the synthesized dance.
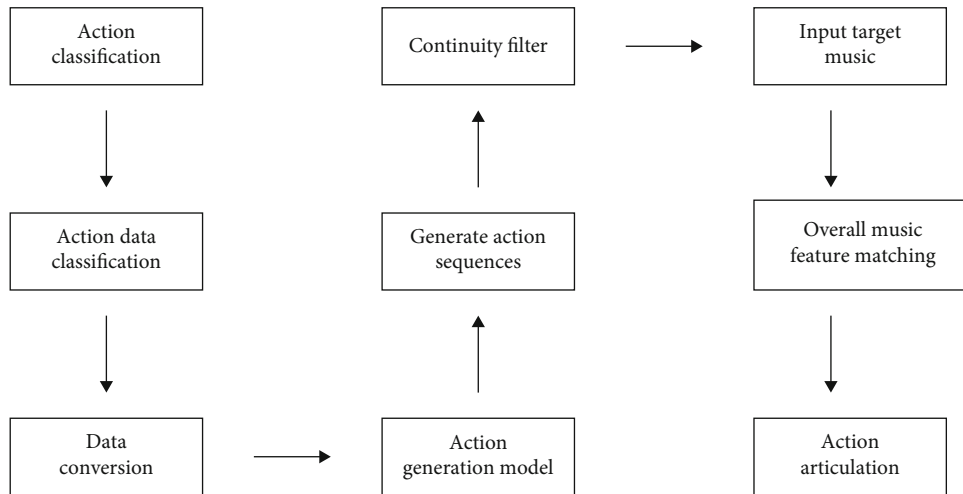
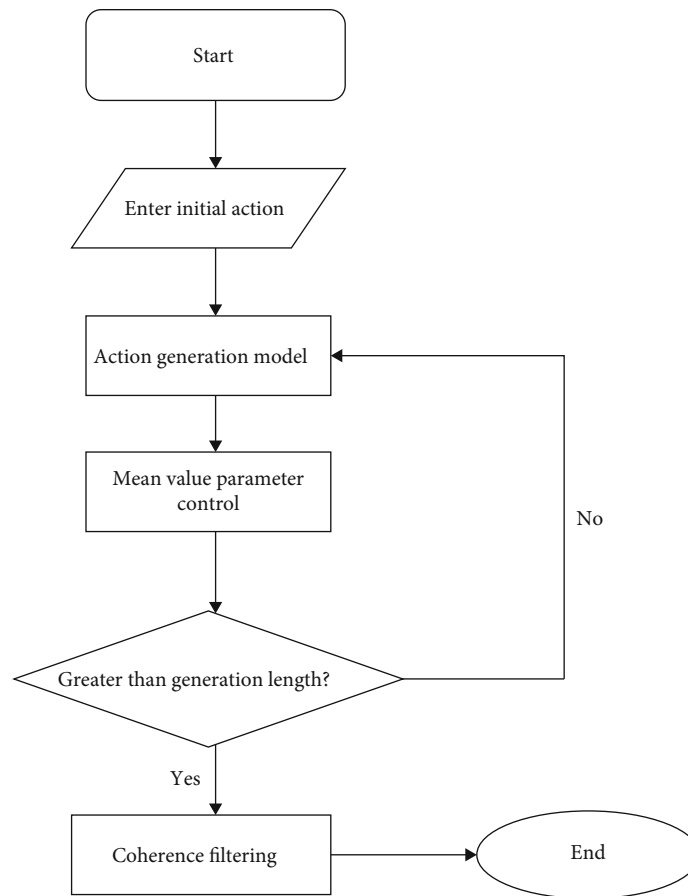Figure 5: Framework of automatic music choreography system.

Figure 6: Motion generation algorithm of method degree network based on mixed secret.

The problem of computerized music choreography has been studied by previous authors and the corresponding computerized automatic sound system; music choreography system has been proposed. The systems can be broadly divided into two categories: one is based on Shirati (2006) [1] represented by traditional people designed by engineers of music and movement features and feature matching algo-

rithms to select the target music from the constructed movement database, and the movement database usually consists of motion capture data; the other is based on Alemi (2017) [2] based on machine learning algorithms, directly constructing music dance mapping models, and generally, the mapping relationship between music and movement features is obtained through model training, so that the dance
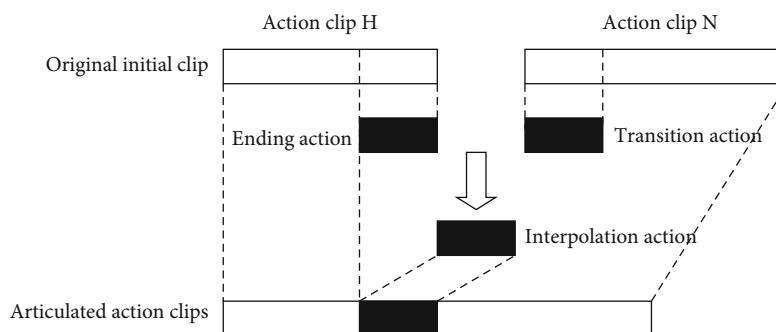
Figure 7: The system adjusts music and actions according to the user's requirements for dance characteristics.
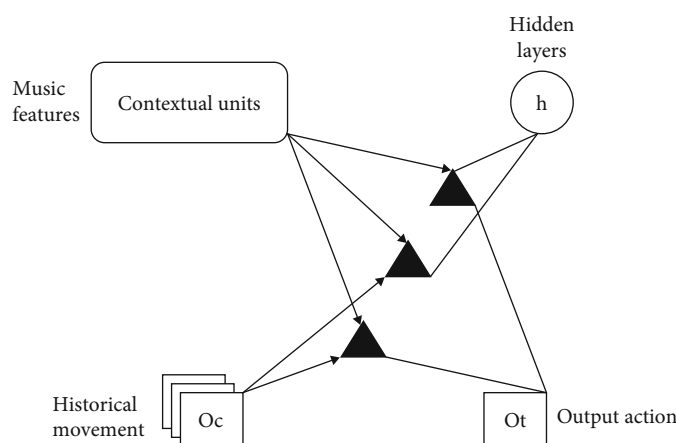

Figure 8: Music choreography system framework.

movements corresponding to the target music can be calculated, and the following figure shows the framework of the music choreography system, as shown in Figure 8.

In this paper, a new framework for automatic music choreography systems is investigated to address the shortcomings of previous work in order to make the generated dance movements both novel and coherent and to analyze and match them with the target music while ensuring that the choreography system has sufficient generalization capability. The framework mainly includes early dataset construction, model training, and training. The framework is divided into four parts: movement generation, choreography, synthesis, and dance visualization using 3D character animation; the core steps are model training and movement generation, as well as dance arrangement based on music and movement features.

Downscaling technical actions, Gaussian processes, Hidden Markov Models, and other machine learning-based motion generating algorithms have all been used in dance studies. Identify potential connections between musical and dance movement characteristics. High-dimensional properties of motions can be mapped to a low-dimensional space that can be utilized to capture potential correlations behind joint rotations in motion capture data using dimensionality reduction techniques. The approach, however, necessitates preparation and processing procedures such as sequence alignment and set data length, and the timeliness of the

motion data cannot be immediately examined for modeling, limiting its use to real dance movement data. Gaussian latent variable model process postvariation model can effectively summarize the changes in human cloud power, but it is not suitable for real-time model generation because of the large computational and memory resources required to overcome the limitations of these two previous models have been mentioned, but their ability to capture data changes is limited.

Music-driven dance generation should not only consider cross-modal sequence-to-sequence mapping but also emphasize the complexity of music-to-dance mapping. The relationship between music and dance movements is arbitrary, influenced by the performer's style and expertise and personality traits. In addition, the mapping relationship between music and dance changes from short-term synchronization of posture and rhythm to long-term synchronization, and the formation of dance patterns presents a complex hierarchy of temporal structures. From this perspective, the neural network has better expressive power than the HMM model. The patterns were learned using LSTM-1, and dance sequences were generated. They used six hours of contemporary dance data captured by Microsoft Kinect to train the model. The trained network can extract dance style (the expression of the dancer's movements), syntax (the language of the piece or choreography), and semantics (the general theme of the dance piece). However, the

algorithm does not provide any control over the process and results generated how to do it, and the whole process does not have any musical accompaniment. The structure of the conditional Boltzmann machine constrained Boltzmann machine (FCRBM) is suitable for controlling the properties of the generated data and allows generating operations in real time. Alemi (2017) suggests that the GrooveNet does not depend on the classification or segmentation of the audio signal FCRBM model can learn continuous cross-modal mapping from audio information to action data in an unsupervised manner. With only 23 minutes of music, FCRBM is trained using a small dataset of dance movements, and the resulting model can be independent, generating basic dance movements based on audio, and you can also learn and generate movements based on training songs. However, this does not have obvious limitations in the poor performance of unheard music and is not very practical and should propose a music-oriented dance synthesis method based on the LSTM model and autoencoder model. The model takes sound features as input and outputs the final dance composition by extracting the mapping between sound and motion features.

5. Main Results of the Paper

(1) This work improves the realism and coherence of the generated motions by introducing a parameter control algorithm and a coherence-based movement filtering mechanism to the movement generating process. According to the results of the experiments, the mean value technique improves the realism of manufactured actions and improves the coherence of the filtered action data when compared to the generated original action data

(2) This work presents a multilevel music and action feature matching method that combines overall feature matching with local feature matching in order to increase music and action unity and coherence. After the dancing motions were matched based on overall music characteristics, the local music movement fragments were matched based on rhythm and intensity. According to the trial results, adding a control based on overall music characteristics makes the pace and other features of each movement fragment in the final synthesis result more consistent and the overall choreography more aesthetic

(3) This study examines the entire process of computer music choreography in detail and suggests a computer music choreography framework as a novel solution to the challenge. The framework is made up of four modules: movement dataset construction, model training and movement generation, dance choreography and synthesis, and 3D character animation visualization, all of which can ensure the authenticity, novelty, and musical compatibility of dance fragments at the same time. Compared with traditional algorithms, the synthesized choreography

is more novel and diverse, and the choreography system in this paper has stronger stability and generalization ability than the music-movement mapping model obtained by machine learning algorithms only. In addition, the user's requirements can be reflected in the choreography results, which makes the choreography system proposed in this paper more practical

6. Problems and Future Prospects

Although the algorithm proposed in this paper can obtain better computer music choreography effect and ensure the novelty and consistency of the choreographed movements and the compatibility with the target music, the research on the direction of feature extraction, feature matching, and dance evaluation of music and movements is not sufficient. The shortcomings of this paper and the subsequent future research directions are mainly as follows:

(1) There is less research on action high-level features in the existing studies, and the action feature-based screening algorithm proposed in this paper mainly utilizes action bottom-level features. In the future research, we can try to analyze the high-level features of the action

(2) In this paper, when considering the matching degree of local music and action, only the rhythm and intensity features of both are considered. In future studies, attempts can be made to include other more abstract common features of music and movement, such as emotion and style

(3) This paper does not adequately examine the methods for assessing the effects of music choreography. At present, when assessing the effect of dance synthesis in existing studies, it is often measured through intuitive visual effects; for example, when assessing the effect of dance synthesis in currently available studies, it is often measured through intuitive visual effects, if professionals are invited to subjectively set music to snapshots of dance movements, synthetic dance videos, etc. The industry has not yet proposed general objective quantitative indicators to evaluate the effect of music choreography, and future research is needed

Data Availability

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

Conflicts of Interest

The author declares no competing interests.

Acknowledgments

References

[1] A. Ns and B. Sf, "Data analysis of music preferences of web users based on social and demographic factors[J]," *Procedia Computer Science*, vol. 198, pp. 730–735, 2002.

[2] N. N. Siphocly, A. Salem, and E. El-Horabty, "Applications of Computational Intelligence in Computer Music Composition[J]," *Egypts Presidential Specialized Council for Education and Scientific Research*, vol. 1, 2021.

[3] X. Yuehua and L. Wei, "Computer analysis and automatic recognition technology of music emotion," *Mathematical Problems in Engineering*, vol. 2022, Article ID 3145785, 9 pages, 2022.

[4] Z. Xinhao, "The practice of string sound source in computer music production–take pop music production as an example," *Art and Performance Letters*, vol. 2, no. 4, 2021.

[5] S. Gang, "Research on architecture for long-tailed genre computer intelligent classification with music information retrieval and deep learning," *Journal of Physics: Conference Series*, vol. 2033, no. 1, article 012008, 2021.

[6] J. Xiangli, "Computer music query by humming considering subsequence matching algorithm," *Journal of Physics: Conference Series*, vol. 2037, no. 1, article 012028, 2021.

[7] L. Maxence, "An analysis of "nyx" (2017), a computer music work by Kerry Hagan," *Computer Music Journal*, vol. 44, no. 2-3, pp. 118–132, 2020.

[8] Y. Hong, "Research on the application of computer music software in piano rhythm teaching," *Journal of Physics: Conference Series*, vol. 1992, no. 3, article 032072, 2021.

[9] L. Mina, "Research on the application value of orff music education system in college music education based on computer new media technology," *Journal of Physics: Conference Series*, vol. 1992, no. 3, article 032032, 2021.

[10] X. Pan, "Research on the application of computer music software in college traditional music course," *Journal of Physics: Conference Series*, vol. 1992, no. 2, article 022178, 2021.

[11] C. Wei, "Research on school-enterprise cooperation training model of music performance professionals in colleges and universities based on repertoire creation and computer multimedia technology," *Journal of Physics: Conference Series*, vol. 1992, no. 2, article 022033, 2021.

[12] Y. Yunyu, "Research on the reform of vocal music teaching by computer multimedia technology," *Journal of Physics: Conference Series*, vol. 1992, no. 2, article 022089, 2021.

[13] H. Elisabeth, "Affective brain-computer music interfaces—drivers and implications," *Frontiers in Human Neuroscience*, vol. 15, 2021.

[14] W. Lingdan, "Application of computer music technology in college music teaching[J]," *Journal of Physics: Conference Series*, vol. 1915, no. 3, article 032073, 2021.

[15] C. Shi, "Research on computer generation technology of fractal music," *Journal of Physics: Conference Series*, vol. 1915, no. 4, article 042015, 2021.

[16] X. Shuai, "Research on the visualization of music stage performance based on the context of computer digital media," *Journal of Physics: Conference Series*, vol. 1915, no. 2, article 022027, 2021.

[17] X. Ning, "Design of online popular music teaching platform based on embedded computer network and virtual reality," *Journal of Ambient Intelligence and Humanized Computing*, 2021.

[18] L. Jie and L. Liang, "A study on the application of computer music in mass music tutoring," *Journal of Physics: Conference Series*, vol. 1865, no. 4, article 042051, 2021.

[19] F. Yuebin, "Application of computer technology in vocal music teaching," *Journal of Physics: Conference Series*, vol. 1881, no. 2, article 022050, 2021.

[20] L. Chuanli, W. Lizhi, and C. Libin, "Research on the application of computer technology in music creation," *Journal of Physics: Conference Series*, vol. 1883, no. 1, article 012031, 2021.

[21] Z. Meili and H. Kewen, "Research on the application of computer music making technology in new media environment," *Journal of Physics: Conference Series*, vol. 1871, no. 1, article 012142, 2021.

[22] R. Mihaela and C. Ciprian, "Discovering music cryptograms in the works of Viorel Munteanu through computer aided methods," *Artes. Journal of Musicology*, vol. 23, no. 1, pp. 165–173, 2021.

[23] K. C. Tseng, "Electrophysiological correlation underlying the effects of music preference on the prefrontal cortex using a brain–computer interface," *Sensors*, vol. 21, no. 6, p. 2161, 2021.

[24] W. Libo, "Research on the reform of music education mode in colleges and universities based on computer music technology," *Journal of Physics: Conference Series*, vol. 1744, no. 3, article 032149, 2021.

[25] S. Rui, "Research on the application of computer multimedia music system in college music teaching," *Journal of Physics: Conference Series*, vol. 1744, no. 3, article 032214, 2021.

[26] S. Yue, L. Yang, and S. Liu, "Application of computer technology in professional music teaching in colleges," in *Basic & Clinical Pharmacology & Toxicology*, vol. 128, WILEY, 111 RIVER ST, HOBOKEN 07030-5774, NJ USA, 2021.

[27] I. Fujishiro and A. Kobayashi, "[Invited paper] Ambient music co-player: generating affective video in response to impromptu music performance:special section on advanced imaging and computer graphics technology," *ITE Transactions on Media Technology and Applications*, vol. 9, no. 1, pp. 2–12, 2021.

[28] L. Ding and D. Lulu, "Application of computer music production technology and computer multimedia system in college sight-singing and ear training," *Journal of Physics: Conference Series*, vol. 1648, no. 2, article 022174, 2020.