Research Article

Application of Multilevel Local Feature Coding in Music Genre Recognition

Yangxin Xiao,1 Qiang Zhang,2 Meng Wu,3 and Dong Kailing4

1Music School, Jiangxi Normal University, Nanchang 330027, Jiangxi, China
2School of Music and Dance, China-Asean Art College, Chengdu University, Chengdu 610000, Sichuan, China
3Jiangxi Police Institution, Nanchang 330100, Jiangxi, China
4Chengdu Polytechnic College, ChengDu 610000, China

Correspondence should be addressed to Qiang Zhang; zhangqiang01@cdu.edu.cn

Received 20 January 2022; Revised 28 February 2022; Accepted 3 March 2022; Published 22 March 2022

Academic Editor: Naeem Jan

Copyright © 2022 Yangxin Xiao et al. 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.

When the current method is used to recognize music genre style, the extracted features are not fused, which leads to poor recognition effectiveness. Therefore, the application research based on multilevel local feature coding in music genre recognition is proposed. Features of music are extracted from timbre, rhythm, and pitch, and the extracted features are fused based on D-S evidence theory. The fused music features are input into the improved deep learning network, and the storage system structure is determined from the advantages of cloud storage availability, manageability, and expansibility. It is divided into four modules: storage layer, management layer, structure layer, and access layer. The model of music genre style recognition is constructed to realize the application research based on multilevel local feature coding in music genre recognition. The experimental results show that the recognition accuracy of the proposed method is always at a high level, and the mean square error positively correlated with the number of beats. After segmentation, the waveform is denser, which has a good application effect.

1. Introduction

The music genre is one of the most mentioned music labels. With the increase of Internet music library capacity, retrieving music by genre has become the mainstream method of music information retrieval [1], and it is also an important basis for a music service platform to recommend music to users. Automatic and accurate music genre recognition can effectively reduce labor costs. Commonly used music genre recognition models generally include two stages: training and testing [2]. In the training stage, firstly, the mathematical model is established to describe the digital characteristics of music genres with discrimination [3, 4]; then the digital features of music files are extracted by preemphasis, Mel filter, and cepstrum lifting; Finally, the classifier is trained based on the digital characteristics and distribution characteristics of different schools. In the testing stage, the same way as the training stage is used for digital feature extraction. The classifier obtained in the training stage is used to calculate the extracted digital features and evaluate the genre. Multilevel local feature coding is of great significance for music genre recognition.

In recent years, scholars try to use different methods to improve the application performance of music genre recognition, and constructing music genre recognition methods has gradually become a research hotspot. Jakubec and Chmulik [5] proposed automatic music type recognition for in-car infotainment. Automatic music type recognition is a basic tool for music retrieval, recommendation, and personalization in the intelligent infotainment system and music streaming media services. Such systems may be particularly helpful for in-car audio because the driver’s interaction with such infotainment systems may be the main object of his/her distraction. In order to better classify genres, two important tasks need to be considered, namely, classifier and audio feature extraction. In this system, timbre texture and pitch content features are used for genre classification. Wood texture includes Mel frequency cepstrum...
coefficient and other spectral features. For pitch content, features extracted from Chroma are selected. The purpose of this work is to explore the possibility of classifying music types from audio signals and to create an automatic music type recognition system in a MATLAB programming environment. A music type recognition system is developed on a gtzan data set, including ten different music types, such as rock, pop, classical, and so on. Ng et al. [6] proposed multilevel local feature coding fusion for music type recognition. Music-type recognition plays a basic role in music indexing and retrieval. Different from images, music genres are composed of immediate features, which are highly diversified and have different levels of abstraction. However, most representation learning methods for Mgr focus on global features and make decisions from the same level of features. In order to make up for these defects, the convolutional neural network is combined with netvlad and self-attention to capture different levels of local information and learn their long-term correlation. The Meta classifier is used to learn from the high-level features aggregated in different local feature coding networks, so as to carry out the final Mgr. Experiments indicate that this strategy outperforms other advanced models in terms of accuracy. Although the above methods have made some progress, they are not ideal in the field of music genre recognition. Therefore, application research based on multilevel local feature coding in music genre recognition is proposed, which can identify the local delicate music genre features and frequency rhythm of the music.

The research is organized as follows: The multilevel local music feature extraction and feature fusion are presented in Section 2. Section 3 analyzes the music genre cloud storage and style recognition model. The experiment and its results are discussed in Section 4 of the study. Finally, the research job is completed in Section 5.

2. Multilevel Local Music Feature Extraction and Feature Fusion

2.1. Music Feature Extraction. The music genre recognition method based on multilevel local feature coding extracts music features from three aspects: timbre, rhythm, and pitch.

2.1.1. Music Timbre Feature Extraction. Frequency usually affects the probability of musical timbre. The short-time Fourier transform method is used to calculate the characteristics of timbre. Let $f_m$ represent the time domain corresponding to the original audio signal, Fourier transform $f_m$ to obtain the timbre feature sequence $G_g$ and the extracted timbre features are as follows:

(1) Spectral Centroid. Assuming that $D_u$ represents the centroid of the music spectrum [7], it mainly describes the centroid of the spectrum, which can be calculated by the following formula:

$$D_u = f_m \times G_g \times B_r.$$  

In formula (1), $B_r$ represents the center frequency corresponding to the $t$ frequency band [8, 9]. Frequency is usually a random variable, and the normalized amplitude is the probability density of frequency:

$$M_d = \frac{R_e}{E_s} \times D_u.$$  

In formula (2), $R_e$ represents the predicted probability item value, and $E_s$ represents the actual probability item value.

(2) Spectral Diffusion. Spectral diffusion describes the diffusion degree of spectral centroid in the spectrum, which can be calculated by the following formula:

$$P_{KS} = M_d \times V_{th}.$$  

In formula (3), $V_{th}$ represents the diffusion velocity of the spectral centroid in the spectrum.

(3) Spectral Skewness. Spectral skewness refers to the degree of data asymmetry, which belongs to a statistical concept. Spectral skewness describes the degree and direction of spectral amplitude skew, which can be calculated by the following formula:

$$F_X = P_{KS} \times m_3.$$  

In formula (4), $m_3$ represents the third-order center distance corresponding to the frequency.

(4) Spectral Kurtosis. Spectral kurtosis represents the flatness of frequency distribution near the center, and the spectral kurtosis is calculated by the following formula:

$$m_4 = \sum (F_X - D_u) \times P_{KS}.$$  

In formula (5), $m_4$ represents the fourth-order center distance corresponding to the frequency.

(5) Spectral Flux. The change of frequency amplitude can be reflected by spectral flux:

$$F_{lux} = Y_t(n) - W_{io} \times m_4.$$  

In formula (6), $Y_t(n)$ represents the $t$ frame sequence after Fourier transforms, and $W_{io}$ represents the spectral flux eigenvalue. Music timbre feature extraction consists of five characteristic factors: spectral moment center, spectral diffusion, spectral skewness, spectral kurtosis, and spectral flux. The average amplitude difference function dynamic threshold strategy is used to segment the audio, the base audio rate is extracted through the autocorrelation function, and the audio segments are matched according to the pitch base frequency. After the weighted synthesis of the Euclidean distance algorithm and DTW (dynamic time warping) distance similarity matching algorithm [10, 11], the music feature recognition results can be obtained; the recognition results of the operation are shown in Figure 1.

As can be seen from Figure 1, the components are mainly input module, preprocessing module, feature extraction
module, similarity matching module, and result output module. The relevant calculation steps in the system are described as follows:

1. The algorithm of the preprocessing module includes noise reduction [12], weighting, and framing windowing to process the input audio signal;
2. The feature extraction module calculates the fundamental frequency of each processed frame signal, divides the audio, and obtains the pitch information and fundamental frequency period of each note after calculation;
3. In the similarity matching module, the music that best matches the pitch contour of the note of the audio segment is searched in the music database to obtain the music with the highest similarity.

2.1.2. Music Rhythm Feature Extraction. Music genre style can be reflected by the strength and speed of rhythm, find the most significant periodicity in the music score signal and realize the extraction of music rhythm characteristics. Based on multilevel local feature coding, in music genre style recognition, combined with the low-frequency characteristics of beat, the beat of the music is obtained. The specific process is as follows:

Wavelet coefficients are obtained by malat algorithm [13], and there are sampling processes and low-pass and high-pass filtering in each layer:

\[
\begin{align*}
x_G(n) &= \sum_{k=0}^{n} (2n - k)g(k), \\
x_D(n) &= \sum_{k=0}^{n} (2n - k)d(k).
\end{align*}
\]  

In formula (7), \(g(k)\) represents a high pass filter and \(d(k)\) represents low-pass filter. High-frequency wavelet coefficients and low-frequency wavelet coefficients can be obtained by filtering. The filtered signal is decomposed by a wavelet. The signal decomposition process is shown in Figure 2.

The extraction of music rhythm features is divided into four steps: full-wave rectification, low-pass filtering, down sampling, and deaveraging. The corresponding expression is as follows:

\[
\begin{align*}
y(n) &= |x(\alpha n)|, \\
y(n) &= (1 - \alpha)x(n) + \alpha y(n - 1), \\
y(n) &= x(\alpha n), \\
y(n) &= x(n) - E[x(n)].
\end{align*}
\]  

In formula (8), the function of parameters is to remove high-frequency noise and smooth the waveform of music rhythm characteristics, which is generally 0.99. When the peak passes, it will appear in the state with the same period and delay. The peak of the waveform after autocorrelation processing is monitored to obtain the beat histogram.

2.1.3. Extraction of Music Pitch Features. Autocorrelation operation is performed on the basis of signal periodic features to detect the peak value of autocorrelation output and extract the pitch features of music, similar to the process of extracting rhythm features. The extraction process of pitch features is shown in Figure 3.

2.2. Fusion of Music Features. Music feature fusion based on multi-level local feature coding uses D-S evidence theory to fuse the above extracted features [14]. The sample space
\( \{ a_1, a_2, \cdots, a_n \} \) is used to describe the music features extracted in the above process, and the sample space \( \{ a_1, a_2, \cdots, a_n \} \) is processed through \( Q(a) \) to obtain the probability distribution function \( M_1(a_1), M_2(a_2), \cdots, M_n(a_n) \), which describes the music feature fusion texture feature.

The target multivariate probability distribution function is fused by the following formula:

\[
\begin{align}
Q(a) &= \sum M(a_i), \\
Q(\bar{a}) &= \sum 1 - M(a_i).
\end{align}
\]  

(9)

In formula (9), \( Q(a) \) and \( Q(\bar{a}) \) represent the probability distribution of target fusion proposition.

Let \( P_f(A) \) represents the fusion likelihood function and \( B_f(A) \) represents the fusion trust function, and their expressions are as follows:

\[
\begin{align}
P_f(A) &= 1 - B_f(-A) \times 2^D, \\
B_f(A) &= \sum M(B).
\end{align}
\]  

(10)

In formula (10), \( 2^D \) indicates that it is composed of all subsets of \( D \).

The music features are judged through the maximum class probability function \( f(A) \) to realize the fusion of music features:

\[
S_f(A) = B_f(A) + \left| \frac{C}{O} \right| \left[ P_f(A) - B_f(A) \right].
\]  

(11)

In formula (11), \( |C| \) and \( |O| \) describe the number of elements in \( C \) and \( O \), respectively.

3. Music Genre Cloud Storage Characteristics and Style Recognition Model

3.1. Research on Cloud Storage Characteristics and System Structure of Music Genres. Music genre cloud storage is the product of the development of cloud computing, and cloud storage further promotes the development of cloud computing, and the two promote each other. Cloud storage integrates cluster applications and distributed file systems. It is a technology that uses the network to provide data storage and access. Its characteristics are shown in Table 1:

Music genre cloud storage should not only have the traditional storage function but also analyze the actual application needs of users. We should not only make reasonable planning for software and hardware resources but also respond to users’ requests for resource storage in real time. This makes the music genre cloud storage system not only have internal storage management performance but also include application service performance. The cloud storage structure model of the music genre is shown in Figure 4.

The functions of music genre cloud storage structure are divided into the following four levels:

3.1.1. Storage Layer. It is mainly responsible for providing storage data, which is composed of physical storage and device management system, and belongs to the lowest part of music genre cloud storage system. Storage devices are generally connected by massive remote storage systems using networks. The main work of the equipment management system is to manage the physical storage equipment, including state detection, redundancy management, and the like, to ensure the barrier-free operation of the underlying equipment.
### Table 1: Cloud storage characteristics of music genres.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Description of music genre characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data availability</td>
<td>Availability is judged according to the fault-free running time of the system. Music genre cloud storage must provide users with continuous data storage services, not only to ensure data security and avoid data damage but also to ensure network stability.</td>
</tr>
<tr>
<td>Manageability</td>
<td>This feature can be expressed using the effective management ability at the same resource cost. For storage hosts distributed in different areas, comprehensive management is carried out to avoid system failure.</td>
</tr>
<tr>
<td>Storage efficiency</td>
<td>This feature indicates how to efficiently use the original storage, which is an important feature of music genre cloud storage. Compress and format convert the data to reduce the data storage space. In addition, you can find and delete duplicate data to improve storage efficiency.</td>
</tr>
<tr>
<td>Access mode</td>
<td>That is, the protocol used by music genre cloud storage service, and the method used by users to access the stored content. At this stage, most methods are based on web services.</td>
</tr>
<tr>
<td>Expansibility</td>
<td>Expanding storage capacity can further meet the diversified needs of users and improve service quality.</td>
</tr>
</tbody>
</table>

![Figure 4: Structural model of music genre cloud storage system.](image)

3.1.2. Management. The main work of this level is to schedule and manage storage operations, which is the most critical link. In order to provide integrated services for storage systems distributed in different regions, the management must use cluster and distributed technology. In addition, the management also needs to have data backup and disaster recovery technology to ensure the security of traditional music storage and provide users with more secure storage services.

3.1.3. Interface Layer. It is an interface designed for applications and a bridge connecting the management layer and the access layer. Music genre cloud storage services can design the application interface in combination with the actual business types.

3.1.4. Access Layer. The access layer is the interface between users and storage services. Authorized users can log in to the system through the specified public interface and enjoy specific storage services. The access mode is generally browser or cloud desktop.

3.2. Construction of Music Genre Style Recognition Model. In order to improve the accuracy of music genre recognition and personalized recognition, a music genre style recognition model is constructed. A matrix is used to describe the user’s scoring data, and the matrix is used as the input of the music genre style recognition model. The selection relationship between users and resources can reflect the interesting trend among users, and the relationship between resource nodes and user nodes in the network is displayed by the dual-mode network in the system. For a specific user group, when users are associated with multiple network activities, multidimensional connections will be generated among users in the group, forming a multimode network of selection relationship between users and resources.

![Figure 4: Structural model of music genre cloud storage system.](image)

Based on the application of multilevel local feature coding in music genre recognition, the fused music features are input into the improved deep learning network, and the music genre style recognition model is constructed to realize music genre recognition. The specific steps are as follows.

Step 1: Input the fused music genre features into the network, initialize the weight matrix of the deep learning network, and initialize the offsets $a$ and $b$ of the hidden layer and the visible layer to 0.

Step 2: Assign a value to the visible layer neural unit, forward transmit the input music features, and obtain the corresponding activation probability of forward propagation in the deep learning network.

Step 3: The activation probability value $p(h|v)$ corresponding to the neuron in the hidden layer is usually a real number, and it is binarized.

Step 4: Back propagate the probability value corresponding to the hidden layer unit in the deep learning network to obtain the reconstruction value $v'$ of the neural unit in the visible layer. Let $p(v'|h)$ represent the activation probability of neural unit $v'$ in the visible layer, and its calculation formula is as follows:

$$p(v'|h) = \prod_{i=1}^{m} p(v_i|h),$$  \hspace{1cm} (12)

Step 5: In the improved deep learning network, the $v'$ reconstructed by forward propagation is obtained, and the activation probability and back propagation probability corresponding to the neural unit $h'$ in the hidden layer are obtained in the above way.

Step 6: Obtain the increment of the offset $b$ in the visible layer by reconstructing the activation probability of the visible layer $v'$ and the activation probability
corresponding to the original visible layer \( v \). Similarly, the increment corresponding to the offset \( a \) in the hidden layer \( h \) is obtained by the activation probability corresponding to the hidden layers \( h' \) and \( h \). The increment of weight matrix \( W \) can be obtained by calculating the back propagation probability and forward propagation probability \([15]\). During each iteration, bias and weights are updated with the learning rate \( \epsilon \):

\[
\begin{align*}
W_t &= W_{t-1} + \epsilon \left[ \nabla p(h|v) - \nabla p(h'|v') \right], \\
b_t &= b_{t-1} + \epsilon \left[ p(v|h) - p(v'|h') \right], \\
a_t &= a_{t-1} + \epsilon \left[ p(h|v) - p(h'|v') \right].
\end{align*}
\]

Step 7: Set the termination conditions and output the music genre style recognition results. The model has the following advantages:

3.2.1. Fast Recognition Speed. This feature is manifested in three aspects: fast data generation, large-scale expansion of massive data, and fast data disappearance. Digitization brings about the mobility of life. Therefore, it is necessary to strengthen the timeliness of information processing:

3.2.2. Recognition Immediacy. Compared with traditional protection methods, recognition models usually have a special relationship with users’ immediate actions. Intangible cultural heritage is dependent on the existence of “inheritors,” which is the same as that of big data;

3.2.3. Integration of Personalization and Socialization. The music genre style recognition model is the product of the combination of personalization and socialization. In essence, it is the extension and development of personalization, which also symbolizes new changes. Music genre style recognition model is closely related to human and social activities. The data can no longer be distributed according to the physical structure but needs to be reorganized in combination with the relationship between individual and society; thus, the application of multilevel local feature coding in music genre recognition is completed.

4. Experimental Analysis

Through the above research on the application of multilevel local feature coding in music genre recognition, in order to verify this method and its effect on music genre recognition, an experiment is designed to verify it. Set the experimental parameters as shown in Table 2:

<table>
<thead>
<tr>
<th>Project</th>
<th>Parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Processor</td>
<td>Inter core i3 M 350@2.27 GHz</td>
</tr>
<tr>
<td>Operating system</td>
<td>Windows10</td>
</tr>
<tr>
<td>Sound card</td>
<td>High fidelity audio sound card</td>
</tr>
<tr>
<td>Software tool</td>
<td>MATLAB 7.0 2018</td>
</tr>
<tr>
<td>Microphone</td>
<td>Dicsong stereo headphones</td>
</tr>
<tr>
<td>Hard disk</td>
<td>160 G</td>
</tr>
</tbody>
</table>

Based on the application of multilevel local feature coding in music genre recognition, four different styles of electronic music resources are identified, and the recognition speed and accuracy of each resource as shown in Figure 5 are obtained through recording and calculation. Automatic recognition performance of different electronic music signals is shown in Figure 5.

In Figure 5, the cylindrical recognition speed shows that the recognition method greatly speeds up the recognition speed of music signals, and the recognition speed decreases slightly with the increase of music duration. From the curve trend of mean square error value, it can be seen that the recognition accuracy of this method is always at a high level for different styles of electronic music signals, and the mean square error value positively correlated with the number of beats. The reason is that this method uses the Gaussian activation function with good analytical performance to replace the commonly used sigmoid activation function, so that the network complexity does not increase with the increase in input variables and has good smoothness and radial symmetry; Based on Windows platform, the automatic recognition process of C++ language architecture is adopted to realize the automatic recognition of target electronic music signals through preemphasis, windowing, and framing.

In order to verify the effectiveness of the application in music genre recognition constructed in this article, evaluation factors need to be constructed because evaluation factors can affect the general set of various factors. After investigation, the evaluation level indicators are as follows: very good, good, general, poor, and very poor. The corresponding level indicators and scoring data are shown in Table 3:

<table>
<thead>
<tr>
<th>Project</th>
<th>Parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Processor</td>
<td>Inter core i3 M 350@2.27 GHz</td>
</tr>
<tr>
<td>Operating system</td>
<td>Windows10</td>
</tr>
<tr>
<td>Sound card</td>
<td>High fidelity audio sound card</td>
</tr>
<tr>
<td>Software tool</td>
<td>MATLAB 7.0 2018</td>
</tr>
<tr>
<td>Microphone</td>
<td>Dicsong stereo headphones</td>
</tr>
<tr>
<td>Hard disk</td>
<td>160 G</td>
</tr>
</tbody>
</table>

Table 2: Experimental parameters.

In formula (14), the element of row \( i \) and column \( j \) in matrix \( J \) indicates that by observing the factor \( ui \) of an evaluated player, it is the fuzzy subset membership of \( vj \) level..
In the evaluation system of live music competition scoring, the primary indicators are singing skills and arrangement skills, and the secondary indicators are note sequence, pitch period, note segmentation, and pitch sequence. According to formula (14), the results are 0.34, 0.15, 0.21, and 0.30, respectively. Pitch period and pitch sequence describe music pitch; note sequence and note segmentation describe music melody.

Three contestants A, B, and C were selected as experimental personnel. Among them, A had received professional training in music and sang some music on the stage; B is an amateur, loves music, and will practice at rest; C has no interest in music and seldom sings. When singing the same song, the time is set to the time of the whole song. Then, select five professional judges to score the three experimental personnel. In order to ensure the effectiveness of the data, the on-site environment should remain unchanged from the beginning to the end and then compare the scoring data of the judges with the scoring data of the system to observe the effect of the constructed system. The specific results are shown in Tables 4–6:

By observing Tables 4–6, it can be seen that the average score of experimenter A is 85, and the singing level is excellent. The average score of experimenter B was 64, and his singing level was good. The average score of experimenter C is 49, indicating that the singing level is poor. In this article, the application method of multilevel local feature coding in music genre recognition is very close to the scoring data of the judges, which can meet the real scoring needs, meet people's subjective feelings, and have stronger objectivity.

The length of sound is the concrete embodiment of the duration of music, which can most mobilize the perception of the human body. In the recognition of music genres, the segmentation length is the first step of music note segmentation recognition, which is used to improve the segmentation accuracy of music notes. The original waveform of music is shown in Figure 6:

Rhythm is the relationship between all notes in the music field and the absolute height between each note. From this point of view, rhythm is constantly changed and improved with the development of society. The ultimate purpose of change is to make the connection between each note closer and present better music effect. At present, the rhythm is mainly divided into three categories: pure rhythm, five-degree phase law, and twelve average law. Each rhythm represents different forms of music singing, of which the twelve average law is the most commonly used. It mainly divides the notes into twelve semitones according to the frequency ratio, and each semitone represents the average law of a rhythm. The performance of notes is accurate to a certain extent, which lays a foundation for the segmentation and recognition of notes. The waveform after segmentation is shown in Figure 7:

It can be seen from Figure 7 that the original waveform is sparse and the waveform after segmentation is dense. We understand the physical characteristics of music notes. By extracting the characteristics of music notes, we lay the foundation for intelligent segmentation and recognition of notes. Based on the theory of pitch and sound level, the pitch
represents the frequency doubling degree of a note, and the sound level represents the treble level of a note. The pitch and scale of a note form a tone. The core of the feature extraction method is to analyze the music note structure and then further classify the note features according to the harmonic information represented by the notes to complete the refinement. The key of the feature extraction method is the sound-level mapping of the notes. This mapping process not only reduces the energy of the notes but also simplifies the note data frame to retain the key information of the notes, eliminate the redundant information, which will not change the training of notes.

In conclusion, the application of multilevel local feature coding in music genre recognition has good performance.

5. Conclusions and Prospects

5.1. Conclusion. For different styles of electronic music signals, the recognition accuracy of this method is always at a high level, and the mean square error is positively correlated with the number of beats.

In this article, the application method of multilevel local feature coding in music genre recognition is very close to the scoring data of the judges, which can meet the real scoring needs, meet people’s subjective feelings, and have stronger objectivity. The original waveform is sparse, and the segmented waveform is dense. The note data frame is simplified, the key information of the note is retained, and the redundant information is eliminated. This process will not change the note training.

5.2. Prospects. There are still some deficiencies with this article’s research material. In future research, I intend to improve the following three points:

1. The future research focus can be on learning the local features of different abstract levels and their long-term dependence. It shows excellent application prospects in the task of music genre recognition. We can try to provide better performance by optimizing the aggregation network of local features and global features at the same time and train the architecture in the way of multitask learning.

2. Learning local features at different levels of abstraction, but how these local features are. Future research work can study different filter visualization techniques to interpret filters and extend the proposed method to other tasks, such as audio event classification, emotion prediction, and music tagging.

3. In the future, we need to further explore the performance of the proposed method on larger data sets. In the future, we plan to add large data sets to the experiment of this article in order to obtain a more complete verification of the proposed music genre style recognition model.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This study was supported by The general project of Humanities and Social Sciences Research of the Ministry of Education, "Chinese music research from the perspective of western scholars" (19YJC76027), the general project of Humanities and social sciences research in Jiangxi universities, "Research on Chinese red music culture from the perspective of western scholars" (YS20216), and the Cultivation Fund of High-level Scientific Research Project of Humanities and Social Sciences of Chengdu University.

References


