Application of Virtual Reality Technology in Computer-Aided Design of Electronic Music

Minmin Ge[1], Yan Tian[2] and Yifan Ge[3]

[1]College of preschool education, Jiangsu Second Normal University, Nanjing, Jiangsu 210000, China, geminmin61@163.com
[2]College of preschool education, Jiangsu Second Normal University, Nanjing, Jiangsu 210000, China, tianyan33@163.com
[3]School of Mechanical Electronic & Information Engineering, China University of Mining and Technology, Beijing 100083, China, sqt1900405071@student.cumtb.edu.cn

Corresponding author: Minmin Ge, geminmin61@163.com

Abstract. In order to make the computer better use in the auxiliary creation of electronic music, this paper proposes a music teaching system based on virtual reality technology. The system is based on audio processing and analysis, and brings a unique sense of immersion, interactivity and conception through virtual reality technology. 3DMax is used to create a virtual electronic music creation environment model, which has the function of a digital audio workstation, and calculates a reasonable minimum rendering complexity according to the network endurance, ensuring that the parallel system is fully utilized to improve rendering efficiency within the allowable range of network carrying capacity. At the same time, a learning mechanism is also adopted, and the use of spectrogram as a quality evaluation mechanism can effectively ensure the direction of auxiliary creation and provide reliable prior knowledge for the production of music. This system provides a new model for the computer-aided creation of electronic music, promotes the efficient creation of electronic music, and provides a faster creation space for music lovers. It is in line with the current computer composition system's development direction towards a comprehensive system. The electronic music business provides great convenience and realizes the rapid development of computer-aided creation.

Keywords: Virtual reality technology; Computer composition; Electronic music; Computer aided creation
DOI: https://doi.org/ https://doi.org/10.14733/cadaps.2022.S5.32-42

1 INTRODUCTION

With the development of Internet and information technology, Internet-assisted teaching has been applied in many disciplines. The basic purpose of music education is to cultivate aesthetic of students' interest and ability, including the training of basic knowledge and skills such as singing, performance, rhythm, appreciation and creation. It determines the particularity of music teaching methods, which are expressed as the fusion of sensibility and rationality. Therefore, the preparation for teaching, the process of teaching and the interaction with teaching can all be achieved through the use of Internet technology. Music teaching system based on virtual reality technology virtualizes the resources of teachers and teaching environment so as to make the use of resources more convenient and optimize the utilization rate of resources. Create the environment for music learning through the unique immersion, interactivity and imagination of virtual reality technology. Current VR technology has been able to create a realistic three-dimensional virtual space in which users can observe things in time and without restriction as if they were there. Here, the virtual speakers can be placed anywhere to present dynamic sound effects. Virtual surround technology can make use of the hardware of two-channel stereo to generate a simulated stereo sound field by processing the sound field signal by a circuit without adding channels and audio material to it. VR Virtual Sound can use its unique panoramic sound field simulation algorithm to accurately restore the ambient sound and virtual sound sources in three-dimensional space, while supporting the sound field reconstruction with the listener's head rotating and moving. Taken together, the VR system based on two-channel virtual surround technology enables listeners to obtain a better listening experience. The fundamental reason is that some acoustic effects are used to produce sound images from different directions, which are no longer limited to the horizontal direction.

This paper analyses the reliable application of virtual reality technology in computer-aided creation of electronic music. By using virtual surround sound technology in the creation of electronic music, you can use "three-dimensional sound image" to produce different electronic music effects than before, and even trigger the "three-dimensional sound image" rebuilding continuously by the author's displacement in the virtual world. This will lead to new breakthroughs in real-time interactive electronic music works, and even help to combine dance with electronic music creation to produce real-time interactive electronic music works.

2 RELATED STUDIES

Electronic music production, also known as MIDI production, is a music production product that can be completed on the computer with the rise of computers in the 1990s.In the most basic equipment we need, computers are essential, sound cards and synthesizers are also necessary. Making 3D animations is the same professional software that requires MAYA and Softimage XSI. MIDI also requires such highly professional software applications. Generally speaking, this kind of software which can make music under PC system, as well as some music making software which can be operated and used under MAC system, are all needed by us. Only after these basic hardware and software are assembled, can they maximize their ability, show their style and start to compile wonderful music works.

Various VR music creation platforms have basically matured in technology, but each function is not comprehensive enough to fully meet the professional standards of the electronic music creation industry. Ideally, a virtual music studio should be an all-round virtual studio that integrates creation, production, recording and performance, and can be operated by multiple people and online remotely. In such a music studio, the creator can select various virtual instruments and their timbre, enter notes or electronic sound samples in virtual tracks and sequencers, or record the performance of virtual instruments for creation. The basic principles of electronic music creation will not change much here. On this basis, different ways of playing different instruments require further code details processing by the developer, which is obviously

feasible. Streeter et al. [1] explored two different design environments, one using OpenGL and the other using Smalltalk-80 to implement squeak. Based on the principle of symmetry, Shi and Wang [2] made the computer-aided design system take into account the musical style based on the neural network algorithm. In early July 2017, the VR Music App Block RockingBeats landed on the HTC Vive platform. Block Rocking Beats, adapted to Oculus Rift and HTC Vive headset monitors, is actually a VR game but has music authoring capabilities. In this app, users can use drums, guitars and other instruments to create music, as well as post-production such as mixing, but its most special feature is to support multiple people to create at the same time, and to allow these people to communicate effectively and in real time, to share their own ideas in electronic music creation. In addition, the app supports users to build tracks in a visual audio timeline [3]. In addition to the creation mode, it can also be used for simple user interaction, that is, the "music anchor" mode, and its beautiful light design is also beautiful. Schumacher and Wanderley [4] proposed a solution to integrate gesture signals into the environment openmusic as music materials. Previously, they also proposed a model in which spatial sound rendering is regarded as a subset of sound synthesis, and spatial parameters are regarded as abstract music materials within the framework of global composition [5]. Leonard et al. [6] used genesis-rt as the virtual reality platform. A similar app is The Music Room, released by Chroma Coda Studio, which focuses on developing the artist's ability to use gestures in virtual spaces to perform instruments using gesture tracking technology. On the basis of a large number of virtual instruments, it also has a complete Bitwig studio and the ability of a digital audio workstation (DAW) for users to create professional music. Nowadays, users can use virtual reality content through Unity, such as 3D video games, architectural visualization, real-time 3D animation, etc. As one of the classic software developed by virtual reality system, has powerful function components [7]. It can quickly add the basic physical properties of things in the scene, and provide the corresponding function component interface for users. Users can create functions for specific objects with common languages such as JavaScript, C# according to their own needs. With the Unity engine, you can quickly create three-dimensional scenes and drive virtual models with animation [8, 9]. This brings great convenience to the development of three-dimensional virtual reality applications and is the mainstream engine of virtual scene development.

3 COMPUTER AIDED DESIGN OF ELECTRONIC MUSIC WITH VIRTUAL REALITY TECHNOLOGY

3.1 Virtual Reality Rendering System

Virtual reality technology is a technology full of sense of science and technology. It uses multimedia computer simulation technology to form a special environment, So that users can experience richer experiences than the real world. Although the traditional CAD system can produce complex, accurate and complete three-dimensional modeling [10], it is not developed for conceptual design and lacks the support of design methodology, so it does not reflect the creative process of conceptual design. At the same time, the traditional CAD system does not allow fast input and reproduction of incomplete conceptual modeling due to many constraints, resulting in it being basically a conceptual drawing tool after the basic design plan is finalized, rather than an auxiliary design tool.

A typical CAD software program consists of three basic parts: input part, application part and output part. For a virtual reality software system, although the complexity of the three components increases, the components remain basically the same. The client server architecture is used in the CAD system (Figure 1). A CAD system's PC is used to process data input, while the geometry modeling and output tasks run on the SGI workstation. With the development of virtual reality technology, the artistic works produced by virtual reality become more and more complex. The traditional single server rendering method for virtual reality works of art can no longer fulfill the complex rendering requirements well, so the rendering work has been improved by building a computer clustered network rendering method. In a network environment, due to the limited

rendering capabilities of CPU and GPU on network nodes, in order to improve rendering efficiency, a parallel cluster network rendering system can be used to connect with network resources. Cluster network rendering system is mainly composed of rendering server, file distribution server, central storage system, rendering management software Enfuzion, etc.

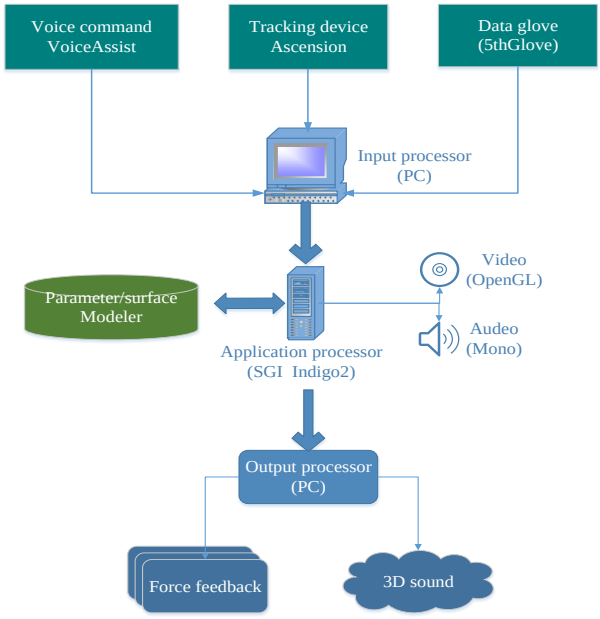

Figure 1: The structure of the virtual reality technology rendering system.

Figure 2 shows the rendering management software Enfuzion task submission process. First, task is submitted from user to the control computer on the client. The control computer distributes the task to different work nodes according to the load balance. At the same time, the control computer also stores the rendering results generated by the work nodes. Finally, the rendering results are returned to the user.

Markov chain has been widely used in the field of algorithmic music creation. Computers can intelligently create music according to the possibility of the next note provided by the conversion table. We can build conversion tables to create different music styles. Through the collection and statistics of specific style music works, such as composers or time periods, the corresponding conversion table can be constructed. The conversion table is actually the quantification of melody style. Markov chains are mainly used to produce a certain style of melody. This method can simulate the composer's thinking when creating electronic music and control the computer to create electronic music. We can progress by using existing melodies as input parameters or by customizing the possibility of each note appearing to produce a new style of music as output. For example, as shown in table 1, it is a Markov chain, that is, the style of music melody is represented by statistical methods. It can query a note by converting the table to any of the following charts of the probability that a note will appear. Not only can we define notes as the state of a Markov chain, but also rhythm, commonly used song techniques, and the structure of a song can be defined as a Markov chain, and then preserve its style templates, resulting in a unique style. This way the electronic music melodies are created with richer lines.

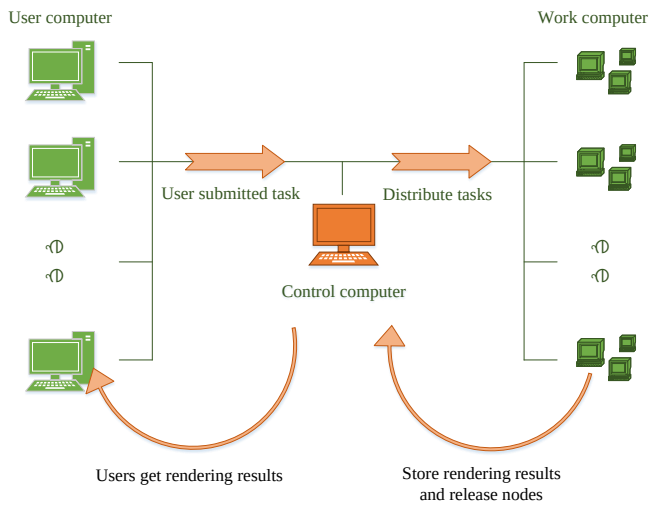

Figure 2: The structure of the virtual reality technology rendering system.

	whole note	minims	crotchets	Quavers	semiquaver
whole note	0	0	0	0.16	0
minims	0.57	0	0.47	0.19	0
crotchets	0.33	0.74	0	0	0.19
Quavers	0.1	0.26	0.53	0.37	0.38
semiquaver	0	0	0	0.28	0.43

Table 1: Markov chain with probability of note occurrence.

3.2 Characteristic Analysis of Electronic Music Signal

The most natural and direct method for signal analysis is to use time as an independent variable. The typical time domain characteristics of music signals use the most short-term energy and short-term average zero-crossing rate. Short-time energy is one of the most commonly used time-domain characteristics in music signal analysis. It is relatively simple to calculate. Many methods of note detection use short-time energy characteristics. Figure 3 gives a short-term energy map of a piece of music, from which you can see that the short-term energy can well reflect the change of amplitude.

Short-time average zero-crossing rate is the simplest feature in time-domain analysis of music signals. Its rate can reflect the spectral properties to some extent, and a rough estimation of the spectral characteristics can be obtained by short-time average zero-crossing rate. The formula for short-term average zero-crossing rate is:

$$Z_n = \sum_{i=n-N+1}^{n} \left|\mathrm{sgn}[x(i)] - \mathrm{sgn}[x(i-1)]\right| 2N^{-1} w^2(n-i) \tag{1}$$

Short-time average zero-crossing rate can simply distinguish the style of music. Similar styles of music do not have a large difference in short-time average zero-crossing rate, while different styles of music should have a large difference in short-time average zero-crossing rate.

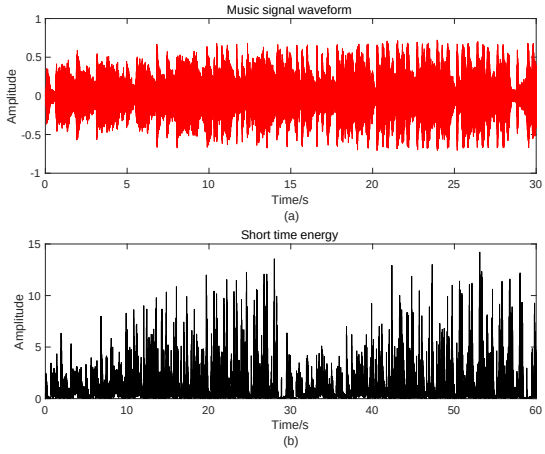

Figure 3: The structure of the virtual reality technology rendering system.

Short-term average zero-crossing rate can distinguish mute segments of music and can be considered mute stages of music when the value is less than a certain threshold. Figure 4 show an average short-time zero-crossing rate graph for a piece of music, where the ordinate coordinates represent the number of times a signal passes through zero points within each sampling interval.

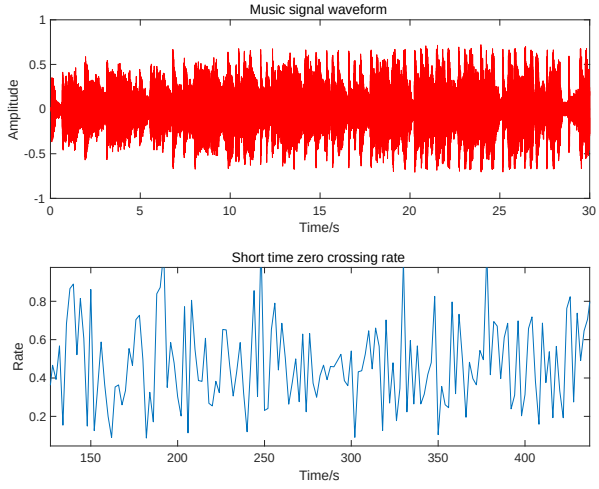

Figure 4: The structure of the virtual reality technology rendering system.

3.3 Frequency Domain Analysis and Audio Processing

The process of human perception of music signals is closely related to the spectrum analysis function of human auditory system. Therefore, spectrum analysis of music signals is an important method to understand and process music signals. Fourier transform is an important algorithm in the field of digital signal processing. Fourier transform can convert time-domain signal into frequency-domain signal, and then inverse transform can convert the processed frequency-domain signal back to time-domain signal. The time representation of a signal has the highest time domain resolution. It can distinguish any two near impulse signals, but it has no capability of frequency

analysis at all. Continuous Fourier Transform (CFT) of a signal has the highest frequency resolution. It can distinguish two complex sinusoidal signals with similar frequencies, but no time domain resolution at all. Short-time Fourier Transform (STFT), as a time-frequency analysis tool, has a certain time and frequency resolution. It is made up of Fourier transforms for each hour. Specifically, the signal is associated with time, so the transformation of signal spectrum can be tracked in time.

The short-time Fourier transform of the signal x n(e) is defined as:

$$X_n(e^{jw}) = \sum_{-\infty}^{+\infty} x(i)w(n-i)\exp(-ijw) \tag{2}$$

In the above formula, $w(n)$ is a window function. Spectrogram, the square of STFT amplitude. The spectrogram is actually the energy density spectral function $P_n(w)$. Define as follows:

$$x_n^2(e^{jw}) \tag{3}$$

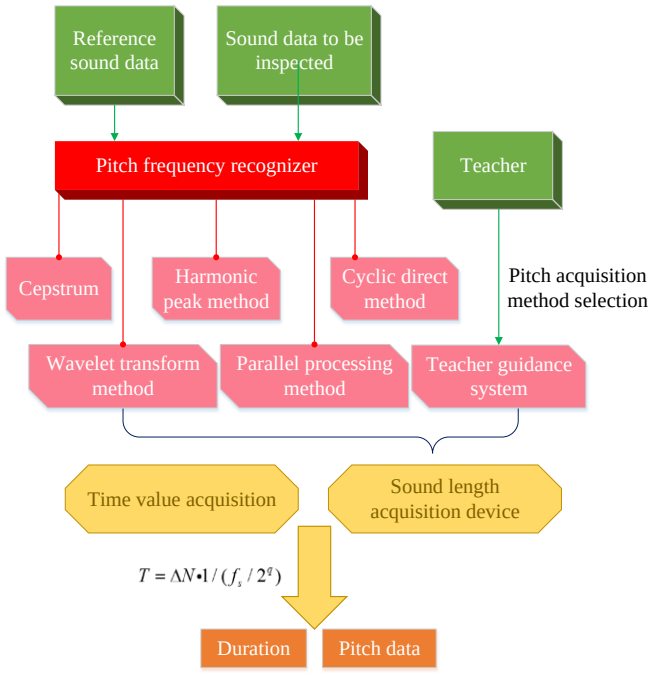

Figure 5: The structure of the virtual reality technology rendering system.

Audio processing system includes feature extraction module and feature processing module. The feature extraction module includes a base audio frequency recognizer and a length acquisition, as shown in Figure 5. The built-in recognition methods of baseline audio frequency are cepstrum, harmonic peak, cyclic straight, wavelet transform and parallel processing. Based on the method selected by the teachers in the teacher's guidance system to get the pitch, the characteristics of the pitch frequency data are extracted. The length acquirer uses an algorithm to extract the duration value, where Delta N is the number of samples between two note endpoints, q is the scale of wavelet decomposition, and FS is the frequency at which the signal is initially sampled.

4 EXPERIMENT AND ANALYSIS

4.1 Virtual Reality Scene Architecture

Create a model of virtual learning environment using 3DMax and import the model into Unity 3D for rendering. In order to simulate the behavior changes of mobile objects after collision in real life, trigger collision detection is combined with traditional collision detection, and the corresponding behavior change components are designed and developed. When a dynamic object enters the specified trigger area, OnTriggerEnter (Collidercollider) function is called to change the value of the state parameters and activate the behavior components. Make appropriate behavioral changes; When a dynamic object leaves the trigger area, the OnTriggerExit (Collidercollider) function is called to change the value of the state parameter and make corresponding behavioral changes.

Camera is the key node of the human-computer interaction module of the whole virtual reality panoramic display platform, which is equivalent to the user's eyes in the virtual world. In Unity3D, the configuration of the camera needs to be combined with many modules. The main code for the specific perspective transformation component is as follows:

```
//Get Current Location
x=Input.GetAxis("Horizontal") * Time.deltaTime * speed; //Move left and right
z=Input.GetAxis("Vertical") * Time.deltaTime * speed; //Move forward and backward
transform.Rotate(new Vector3(x,y,0)); //Rotation angle (increase)
//Camera position
Rotation=Quaternion.Slerp(transform.rotation, targetRotation, Time.deltaTime * 3)
```

The result of audio processing is displayed dynamically in the virtual scene, which increases the interactivity of learning in the virtual scene, and the result prompt information of audio processing is displayed in the virtual scene through Unity3D development.

```
void Awake()
{//Initialize Alhpa to make fonts and colors
Color alpha = mText.color;
alpha.a = DefAlpha;
mText.color = alpha;  }
void Start()
{
mText.text = Texts[CurrentTextIndex]; //Assign a value to mText
Fade(true);
}
```

Figure 6 lists three commonly used film and television digital media image sizes. The larger the image, the larger the amount of data transmitted in real time and the greater the network speed required. However, for a specific network, its data transmission capacity is certain. In Table 1, it is assumed that the stable and continuous effective data transmission capacity of the network is 200 Mbps. This assumption is reasonable and is proposed after fully considering redundancy on the actual cluster network platform. It calculates a reasonable rendering complexity threshold (lowest value) according to the network bearing capacity, which can ensure that the parallel system is fully utilized within the allowable range of the network bearing capacity to effectively improve the rendering efficiency. Assume that there are 5 nodes participating in the rendering operation in the cluster network.

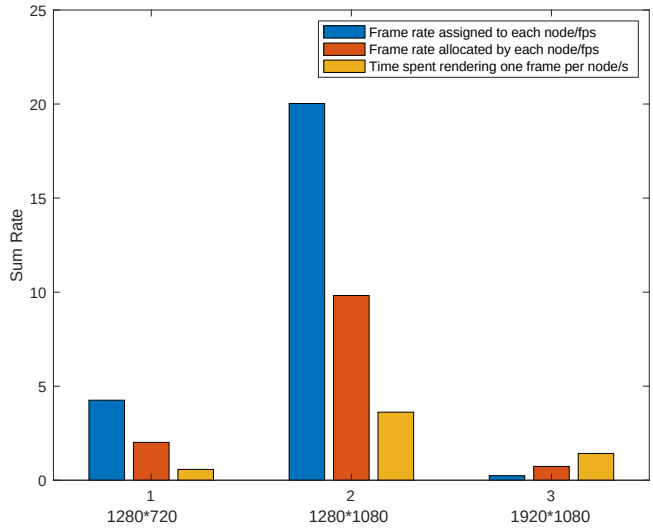

Figure 6: The relationship between network performance, virtual image size and rendering algorithm complexity.

4.2 Analysis of Experimental Results

There must be a quality evaluation mechanism to judge whether the music fragments produced by algorithmic composition comply with the traditional music theory, which is a very important part in any computer-aided design system. It can help detect whether the content of auxiliary creation forms a specified style under our control. We use the method of comparative spectrum to analyze the advantages and disadvantages of music style.

The spectrogram is a graph showing the changes of the music spectrum over time. The vertical axis is frequency and the horizontal axis is time. The strength of any given frequency component at a given moment is represented by the grayscale or hue of the corresponding point. The spectrogram shows a large amount of information related to the sentence characteristics of the music. It combines the characteristics of the spectrogram and the time-domain waveform, and clearly shows the change of the music spectrum over time, or a dynamic spectrum. In music signal processing, the spectrogram plays a very important role. Because there are many frequency components at each moment of music, the spectrogram can give the duration of each frequency component, and it can also intuitively see the notes in the music. Changes. Perform time-frequency analysis on a segment of electronic music created with computer aids, and get a spectrogram as shown in Figure 7.

It can be seen from Figure 7 that the size of a certain frequency remains basically unchanged for a period of time. This is because when the instrument is sounding, each note has a stable frequency spectrum and it will last for a period of time. Therefore, the energy peak in the frequency spectrum will remain at a stable frequency position for a period of time.

The application of virtual reality conforms to the direction of the current computer composing system towards a comprehensive system. Among them, due to the adoption of a learning mechanism, it can effectively ensure the direction of system evolution and provide reliable prior knowledge for the production of music; the use of selective interactive technology can not only ensure that the music has a certain musicality, it can also be appropriately reduce the amount of manual participation, will not affect the speed of system evolution, improve the efficiency of the system.

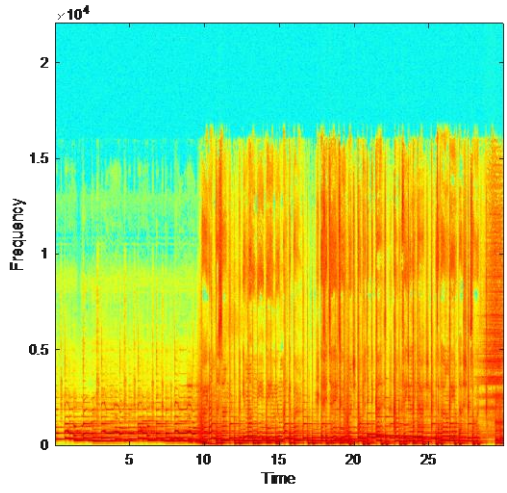

Figure 7: Spectrogram of computer-aided design electronic music.

Distributed architecture can improve the cooperation between various parts of the system, and a new type of "symbolic" encoding is used, which breaks through the limitation of poor controllability of traditional binary encoding. Finally, through the experimental verification, a satisfactory result is obtained.

5 CONCLUSION

At present, the research and development of computer-aided creation electronic music system is almost blank in China, and research in this area abroad is very active. Based on virtual reality technology, this paper virtualizes music resources and creative environment resources by computer-aided creation, so that the use of resources is more convenient and the utilization of resources is more optimized. Unity3D engine is the basic development platform, combined with other modeling tools to build an efficient virtual reality electronic music creation system. Through this system, users can not only learn music in a virtual environment anytime and anywhere, but also can create an environment required for music learning through the unique immersion, interactivity, and conception of virtual reality technology. Develop an easy-to-master algorithmic composing software with complete intellectual property rights, noted musical notation mutual translation, artificial intelligence and expert system, and at the same time, develop a virtual reality system. On this basis, add spectrum analysis, slice synthesis, and digital signal processing. A complete set of application music intelligent and complete composition system, mixed drawing system, music production system. With the continuous maturity of virtual reality modeling technology, virtual surround sound technology, virtual reality hardware, etc., the prospects for virtual reality to be applied to electronic music creation are getting better and better. The future work of the pen is to complete the basic structure of the system, and other functions will be gradually improved.

Minmin Ge, https://orcid.org/0000-0002-6756-9758
Yan Tian, https://orcid.org/0000-0003-1396-8216
Yifan Ge, https://orcid.org/0000-0001-5372-7922

REFERENCES

[1] Streeter, E.; Davies, M.; Reiss, J.-D.; A Hunt, A.; Roberts, C.: Computer aided music therapy evaluation: Testing the Music Therapy Logbook prototype 1 system, Arts in Psychotherapy, 39(1), 2012, 1-10. https://doi.org/10.1016/j.aip.2011.11.004

[2] Shi, N.; Wang, Y.: Symmetry in computer-aided music composition system with social network analysis and artificial neural network methods, Journal of Ambient Intelligence and Humanized Computing, 11(8), 2020: 1-16. https://doi.org/10.1007/s12652-020-02436-7

[3] Ariza, C.: Two pioneering projects from the early history of computer-aided algorithmic composition, Computer Music Journal, 35(3), 2011, 40-56. https://doi.org/10.1162/COMJ_a_00068

[4] Schumacher, M.; Wanderley, M.-M.: Integrating gesture data in computer-aided composition: A framework for representation, processing and mapping, Journal of New Music Research, 46(1), 2017, 87-101. https://doi.org/10.1080/09298215.2016.1254662

[5] Schumacher, M.; Bresson, J.: Spatial sound synthesis in computer-aided composition, Organised Sound, 15(3), 2010, 271-289. https://doi.org/10.1017/S1355771810000300

[6] Leonard, J.; Cadoz, C.; Castagné, N., Florens, A.; Luciani, A.: A virtual reality platform for musical creation: GENESIS-RT, International Symposium on Computer Music Multidisciplinary Research. Springer, Cham, 2013, 346-371. https://doi.org/10.1007/978-3-319-12976-1_22

[7] Chang, Y.-S.; Chen, Y.-C.; Chuang, M.-J.; Chou, C.-H.: Improving creative self-efficacy and performance through computer-aided design application, Thinking Skills and Creativity, 31, 2019, 103-111. https://doi.org/10.1016/j.tsc.2018.11.007

[8] Ariza, C.: Two pioneering projects from the early history of computer-aided algorithmic composition, Computer Music Journal, 35(3), 2011, 40-56. https://doi.org/10.1162/COMJ_a_00068

[9] Agostini, A.; Ghisi, D.: Real-time computer-aided composition with bach, Contemporary Music Review, 32(1), 2013, 41-48. https://doi.org/10.1080/07494467.2013.774221

[10] Dahlstedt, P.; McBurney, P.: Musical agents: toward computer-aided music composition using autonomous software agents, Leonardo, 39(5), 2006, 469-470. https://doi.org/10.1162/leon.2006.39.5.469