Study on Application of Audio Visualization in New Media Art

Yingfang Zhang, Yi Pan, Junren Zhou
Sino-Korean School of Multimedia Design, Shanghai University of Engineering Science, Shanghai 201620, China
zyfang0502@163.com

Abstract. Audio Visualization is an objective interpretation and judgment for music representation, and an approach of representation used to understand, analyse and compare representation capacity and internal structure of music. As a way of representation for new media art, audio visualization is widely used and frequently appreciated, from large light show, stage performance to music video and so on. With Cyber Space as an example, this paper has explained how audio visualization is planned, designed and produced in new media art products. In addition to creativity idea origin, product naming, style definition, music selection and editing, scenography and interactive design, software and hardware resources used in product creation are also briefed.

1. Introduction
As new media technology is developed and multimedia is popularized, audio visualization has gradually entered into people's life. Its vivid and dynamic representation has exceeded traditional media for information communications where pure texts are mainly used. Cross-media contents and products represented by audio visualization have exploded in terms of quantity, enabling people to gradually get adapted to getting information through audio and video, and also transforming modern society to a post-text society. Audio visualization plays an irreplaceable role in current society.

During 1980s, birth of MIDI technology caused the realization of standardized communication among different digital instruments. With advancement of digital audio processing and audio post-editing technologies, creators can accurately and efficiently obtain required sound. Digital image technology and multimedia technology have allowed smart combination of images and sounds to form new media arts either focusing on image or sound. Emergence of sensing technology has facilitated and diversified controls over different devices and increasing computer performance have expanded creation room. Supported by these increasingly developed technologies, and with continuously updated ideas from creators, forms of stage performances are more and more diverse as creation room is expanded. [1]

New media, as a carrier, exists in various kinds of arts. It is an emerging discipline of art based on "optics" and "electronic media". The audio visualization mentioned herein refers to a form of representation where audio and video acting as carriers are integrated to digital media art.

2. Overview of Audio Visualization
Visualization refers to any technique or method used to create graphics, images, or animations to communicate information. Historically, it includes cave paintings, Egyptian hieroglyphs, etc. In
computer technology, visualization is a means of process presentation and is a tool for human-computer interaction, such as visual programming, visual modeling and so on.

Audio Visualization is an objective interpretation and judgment for music representation, and an approach of representation used to understand, analyze and compare representation capacity and internal structure of music. As an emerging visualized technology, music visualization has exerted an important influence on currently popular media players where feature extraction, image processing and virtual reality are widely focused in multimedia studies.

After features of music, such as wave form, frequency, pitch, loudness, rhythm, speed and tone are extracted, music visualization technology has them projected to respective visual effects. These visual effects are expressed in multiple ways, such as smoke, water wave and flame, even these complicated and new ones that are generated on the basis of computer graphic transformation, like changeful virtual roles (movements of small animals or people) and their virtual scenes, and abstract elements including musical note, loudness and rhythm, which are more vivid and touching. The visualized effect in Windows Media Player is a typical example for music visualization which also provides the initial solution to music visualization. The birth of music visualization makes seeing music possible. [2]

It is a new research direction in the fields of music, digital audio, graphics, image processing, virtual reality, etc., and has broad application prospects in entertainment, education, art and business.

2.1. Principles and Representation Forms of Audio Visualization
When feeling various sounds, people will associate the listening experience with other experience obtained through another non-auditory organ unconsciously and this phenomenon is called synaesthesia. Generally, musical notes at high voice are often linked to bright visual effect and active or happy emotions while musical notes at low voice, to dim visual effect, grieved or sad emotions. Slow rhythm easily enables people to think of open space or peaceful emotion while people are inclined to think of narrow space or agitating emotion when hearing urgent rhythm.

With physical wave forms of music files, music control systems classify a piece of music into several periods and automatically identify its basic emotional features such as shocking, missing, expressing deep feeling, joyful, inflaming, sorrowful, delightful or ardent, for further deciphering and encoding. Following this, the music is transformed into control signals which are outputted to specific control units after synchronous processing. For example, when the music is represented as a fountain, control units will manipulate various actuators like submersible pumps, electromagnetic valves, underwater colored lights and variable frequency drivers according to water shape rules to keep fountain shape synchronous with changing lights and music, thereby delivering a perfect combination of fountain shape, colors and musical emotions. This allows the generation of changeful waterscapes where video and audio are fused to better reflect theme and internal elements delivered by the music. [3]

The audio visualization, a representation form that uses computer programming to present electronic music through geometric elements, e.g., points, lines and planes, has aroused broad attention from new media art and digital art. In these art products dominated by electronic music, audio visualization is often used as the core for performance. In these products, visual effects are changed with music rhythm, which can be immediately generated or pre-set in control systems. If images are immediately generated by computer calculation, it is required to consider overall product style when programming and then design specific contents. In the case of pre-set images, the process similar to making a film, including dividing shots and making scripts becomes necessary.

The American electronic music producer Karma Fields's musical work The Hex (Trailer2) is dominated by music. The performer plays in the hexahedron made by the screen. The imagery on the curtain changes along with the melody of the music, as shown in Figure 1. [4] New media art works such as this combined with installation art, digital art, and performing arts are emerging, and are also the forms of audio visualization in the art field.
2.2. Applications of Audio Visualization

Audio visualization can be applied in education, communications, medical care and even military areas, also with pretty wide applications in stage performance, such as pop music concerts, awarding ceremonies, TV programs, music festivals, and some traditional musical performances like operas and symphony concerts. [5] As a part of scenography in stage performance, audio visualization mostly serves to coordinate with style and theme, combined settings, other design elements such as scenes and lights, and actors/actress.

Many products have pre-set audio which are used to produce visual effects and images. For these products, it is necessary to carefully analyse each part of music, then portray detailed scripts and shots, and eventually produce video using post-processing software AfterEffect and Promiere. Finally, audio and video are combined to form a completed product. [6] When producing visual effects for these products, it is important to have visual effects close to themes, helping audiences better understand music. For instance, visual effects for traditional rock music should focus on dark colors with high contrast and purity, which are easily linked to ardent and strong elements like flames and steels. However, when it comes to folk music, gray and light colors presenting story-like experience should be mainly used in visual effects.

In recent years, as electronic music has been playing an increasingly significant role in mainstream music, many stage performance styles based on electronic music are generated. Unlike pop music, modern electronic music is mostly pure music without singers, which has changed the performer part that most attracts people's attention in original musical stage performance. As a result, scenography becomes more important.

3. Audio Visualization in Cyber Space

This product is inspired by the Cyber concept in computer culture. Normcore and electronic style are used as the main keys, which are consistent with its name of "Cyber Space". In terms of software and hardware selection, as a product of new media art, Cyber Space uses software and hardware, such as processing, osc and kinect, to present stage performance with audio visualization as the carrier.

3.1. Plan of the Project

Different themes match with different main styles and productions, which naturally requires different musical styles. When you want to present a large scene with great momentum, it is mostly to select classic music or symphony. Tranquil and gentle scenes often require a piece of solo or soft music. If you would like to create exciting and passionate atmosphere, rock is a good choice. Electronic music
generally goes with popular products with distinct rhythms. The first step for planning Cyber Space is extracting key words. Following brainstormed thinking and collection, Digital, Image, Scenes, interaction, design, geometry, Minimalist, Generate, Electronic, Technology, Electronic and Performing are finalized as the key words.

Regarding concept determination, Cyber Space theme is explained as follows: Cyber Space is originated from cybernetics and spaces, two words initially created by a Canadian sci-fiction author William Gibson in his short fiction Burning Chrome and promoted and popularized in later fictions. "Cyber" in modern society, is relating to characteristic of the culture of computers and artificial intelligence. "Cyber Space" is an abstract concept in areas of philosophy and computer, meaning the virtual reality in computers and computer networks. Naming this work as "Cyber Space" notably demonstrates that this product is closely related to computer science and digital art, therefore electronic music which is also closely linked to development of computer and digital art is selected as the background.

3.2. Design of Audio Visualization
After planning, visual effects can be designed. Images expressed in combined forms of points, lines and planes and real-time music sampling through application of processing work together to create real-time images for product performance. The design idea of using points, lines and planes is mainly originated from overall music style and product's theme "Cyber Space".

Combining points, lines and planes in different forms to create letters, words, phrases and other shapes (as shown in Figure 2), generating different visual effects that enable audiences to intuitively feel theme and connotation of this product. Real-time sampled audio is projected to each visual effect whose shape is changed as musical rhythm changes. The letter S (as shown in Figure 2(c)) formed by curves is scaled up, scaled down and drawn again as musical rhythm changes, showing strong visual rhythm.

After overall design style is determined, processing programming-based visual effects are generated. Then, with extracted music features, changes of visual effects, including fonts, line thicknesses, line movements and colors are produced to harmonize music rhythm and melody. Concerning color selection, black, white and gray tones are mainly used to present the normcore idea.
In addition to processing programmed visual effects, visual effects from human images captured by kinect are also designed in this product. In this case, kinect captures movements from performers who are delivering performance behind the screen, and design visual effects of these captured images. The selected backgrounds are also dominated by points, lines and planes. With respect to specific visual design, many effects are designed and then tested, including line movements like falling rains, human 3D mirroring, geometric spreading movements, which can provide interactions in consistent with performers' movements.

In the stage design, the audience can see the kinect image projected on the screen, the silhouette behind the screen, and the visual effect as a background (see Figure 3).

### 3.3. Audio Visualization Technology

As described above, the visual effects of the background are generated using computer algorithms. Because the project itself requires a very strong flexibility, the work is mainly developed using the Processing programming language which is lightweight, highly flexible, and rapidly developed.

The audio is entered into the computer as data and then processed into an audio object via the open source audio processing library Minmin, which contains the sampled audio data. After the fast Fourier transform (FFT) algorithm is processed, a distribution array of high and low audio frequencies is obtained. Bass numbers are used as the basis for recognizing drums, thus enabling the recognition and tracking of beats. The project also uses a Geomaticative library that can process text and SVG graphics into coordinate arrays.
For Kinect driver and human body motion recognition processing, the project uses the OpenNI SDK for development and uses SimpleOpenNI as a middleware between Processing (Java) and OpenNI (C++).

The project will extract some of the key parameters affecting the screen, design an interactive interface that can be used to modify the visual effects and is in accordance with the habits of the theater staff, and use the iPad device as an external controller to communicate with the computer using the OpenSound Control (OSC) protocol to make the interactive interface is connected with other parts.

4. Conclusion
It is found during planning, designing and producing Cyber Space that audio visualization, as a part widely applied in stage performance and new media art, can be designed in a growing number of ways. The obtained audio visualization effects by dividing an audio into several segments, editing these segments, extract features and the produce images and videos against to-be-presented theme, are more enriched and can be presented with metaphoric symbols which are more easily understandable by and acceptable to audiences, and better combine video and music to help audiences to understand connotation of the product.

As a product in new media art, Cyber Space demonstrates electronic music in the way of audio visualization to allow audiences to feel delivered contents and emotions visually and audibly. Unlike the uniqueness and novelty from independent products in new media art, audio visualization in exhibition art and scenography are closer to popular aesthetics and focus more on representation effect. With emergence of more and more image and audio processing technologies, audio visualization will witness an extreme broad space for development and design in cultural communications and art areas. Therefore, after mastering application technologies, difficult point in this field has changed from technological transformation to creativity. An outstanding new media art product must be supported by sufficient cultural contents and represent emotions and ideas concealed in the product in proper ways. For this, each process matters so much.

References
[1] QU Tianxi, HUANG Dongjun, TONG Kana. A Survey of Audio Visualization[J]. Computer Science, 2007, 09:16-22.
[2] Zhang Yanpeng. Audio Visualization Design Method Based on Multi-Audio Features[D]. Harbin University of Science and Technology, 2013.
[3] Zhang Tian, Li Yu, Gao Chang, Qiu Rongfa, Li Haifeng. Analysis and Visualization of Digital Media Content Based on Audio[J]. Journal of Yanshan University, 2010, 02:100-105.
[4] Ravenkwok, Karma Fields - The Hex (Trailer 2) on Vimeo[EB/OL], vimeo.com/184869751
[5] Zhang Xiaofu, Zhang Zhaoxia, Liu Yanbiao. Three Important Categories in the Study of Chinese New Media Art [J]. Journal of Renmin University of China, 2013, 01:44-51.
[6] Chai Yuanyuan. Real-time Visualization of Digital Audio[J]. Electronic Technology and Software Engineering, 2013, 16: 105.