Research on the Application of Sound in Virtual Reality

Jiwen Tang*

Shanghai Conservatory of Music

*Corresponding author. Email: tangjiwen1999@foxmail.com

Abstract. Immersive sound is an essential element of any Virtual Reality (VR) experience. In recent years, increasing popularity of VR devices and numerous advancements in hardware and software technologies, the development of immersive interactive sound has become more potent. This article explores VR through the sonic perception, presents a review of the Method of combining and applying sound and VR, with the goal of providing more references and suggestions for the sound of VR. This includes the creative and technical work of music and spatial sound in practical projects in recent years. Some examples of musical VR are provided in this article to compare their characteristics. A review of the state of the art of sound simulations, focusing on the various factors that ensure a complete and immersive audio experience, while trying to strike an appropriate balance between realism and accuracy. This article finally discusses the future and challenges of sound in VR.

Keywords: Virtual reality, Sound, Music, Spatial sound.

1. Introduction

The offline retail market of virtual reality (VR) devices will continue to be affected by COVID-19 in 2022. However, with affordable VR devices gradually occupying the market, the number of users of VR devices is still growing increasingly. It is expected that the global VR user base will increase to 216 million by 2025. At the same time, because a large number of offline social production and activities were forced to stop during the pandemic, VR technologies have the opportunity to go deep into various industries and play their role., researchers are also constantly exploring the new potential of virtual reality, alter the way of observing, hearing, touching etc.

In a VR environment, sound is of great benefit to enhance users' immersion and presence (see Figure 1) [1]. Combined with stronger visual stimuli in VR, sound can play a greater role in enhancing the VR experience [2]. Sound can direct users' attention and behaviour in VR and elicit psychological anxiety or tension. Also, because hearing is based on Omnidirectional perception, users can perceive the changes of the surrounding environment through sound without turning our heads. Appropriate sound design and arrangement based on binaural signals can build sound externalization [3], make the perceived agreement with the user's expectation towards a corresponding real acoustic event, in order to create a virtual environment considered to be real. Therefore, the dynamic experience produced by the combination of multi-sensory interaction and immersive sound from VR makes the previous sound design and concept based on static experience less applicable.

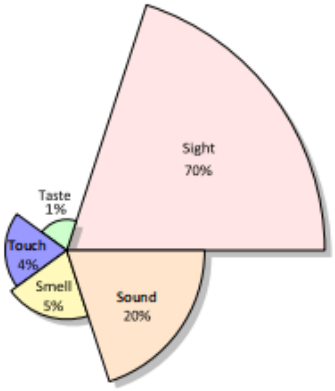


Figure 1 Comparison of the weighting importance of the five human senses.

Although sound has always been valued by researchers in VR, there is a general consensus that sound is an undeveloped sensor system in VR. By searching the relevant literature of virtual reality in recent years, this review summarizes how sound is designed and used in various forms of VR environment, how it is combined with VR to structure an interactive system, explore the latest audio technologies and how these technologies can further improve the interactive sonic experience of users.

2. Music In Virtual Reality

Considering non-visual sensory modalities in VR systems, as well as using sound and music as primary drivers of rendered realities, current concepts of VR systems and experiences need be extended. Thus, the next two aspects of musical VR are developed abstractly from audio-first and sensory modalities [4]. Finally, the applications of music education in VR are studied.

2.1. Audio-first

As part of sonic interactive systems, music can play a variety of roles, including narrative and non-narrative uses of music, recorded audio playback, triggered interactive events, and dynamic spatial audio that changes in response to head or location tracking.

As opposed to most scenarios where music plays a relatively passive or supporting role, in a VR experience that puts audio first, musical interactions can drive the user's continued immersion in the VR environment or trigger a series of events. Among the genre's possible sonic interactive applications are game-based musical scenarios, computer-generated musical engines, immersive performance environments, and virtual musical instruments.

When developing a composition, the composer is constrained by a number of internal and external factors. By utilizing virtual reality, the composer will be able to expand their music beyond these physical constraints based on the affordability and accessibility of the software. For example, Damian Mills, Franziska Schroeder, and others have developed a VRMI (virtual reality musical instrument) that helps people with disabilities play and make music [5]. The application provides a musician who lives with quadriplegia cerebral palsy the ability to play accessibly through visual and tactile interaction with the VRMI.

A project focusing on DMI (Digital musical instrument) explores the non-physicality of virtual environments and their impact on interactions with DMI [6]. Through the control of gravity in the virtual environment, users can get a different experience from using the real piano when interacting with the keyboard (see Figure 2). This project makes use of only one feature of the virtual environment and does not pay attention to immersion or presence, even though the user can participate in the interactive experience through hand tracking. Another virtual project produced a fully immersive DMI [7]. By creating galaxies and planets, the user can customize the sounds and compose their music.

Iain Findlay-Walsh uses the theory of VR to explain the virtual space and sense of presence felt when listening to pop music [8]. Even though, as noted, the main driver of audio-first VR is sound interaction, the analog narrative world of digital pop music offers a range of possible spaces as well as narrative perspectives for the listener to consider. Sound and users can interact in more internal ways than simply triggering events. But based on the multi-sensory nature of VR, presence and interaction of virtual reality can also be approached from diverse perspectives.



Figure 2 Artificial physics can be simulated using the virtual keyboard. By controlling key collisions and orientation, creating diverse sounds.

2.2. Sensory Modalities

A fundamental characteristic of VR systems and experiences is the simultaneous presentation of stimuli via multi-sensory. Additionally, different sensorial inputs, each representing a different aspect of the non-real environment or experience itself, provide a strong sense of immersion and presence in VR, as well as providing artists and developers with a variety of abilities to create memorable and captivating user experiences. Sensory modalities are extremely significant in collaborative music performances that involve virtual instruments. These modalities make it possible to define a participant's contribution and, consequently, their degree of engagement in the performance [9].

Alena Miskinis and Xiangxu Lin explore a new application of stress therapy that reduces the effects of stress by combining VR technology and music therapy practices [10]. And in addition to the two sensory interactions from music and video, the application offers further interaction options through touching. An additional application is Hyperreal Instruments, in which users are able to experience tactile stimulation by combining a digital representation of the instrument with a physical replica of the instrument made from the same prototype. A sensory fusion across visual, acoustic, and tactile modalities embeds the virtual experience in the performer's physical world while providing a sense of presence and immersion for the performer.

VR offers new opportunities for the transfer of creative ability to users. In the context of music interactive design, participatory systems are not novel; however, VR allows the designers to assign the users specific abilities that can be flexibly customized and applied in ways that transcend the realm of physical reality. Through integrating senses into how the user explores and engages with the virtual environment, stimulation of sound to users extends to the whole body, establishing a deeper relationship between users and the musical VR system.

2.3. Education

Education as a fundamental domain, combined with VR has also attracted many researchers to explore it [11]. Most VR musical education software or platforms are designed to popularize music knowledge and guide beginners. The content covers different areas from basic theory knowledge, instrumental music, vocal music, etc. Some of the recent VR musical education applications will be discussed in the following three aspects.

2.3.1. Immersion

Immersion means that learners keep immersive in the virtual music environment created by virtual reality technology, feeling that they are part of this environment, surrounded by the virtual music world, and more focused during the experience. Wei Chen proposes an intelligent music teaching system design [12] that provides a stage for singers and dancers in a virtual environment that is not constrained by physical factors. The authors build the platform by constructing four systems including model embedding, face tracking, gesture tracking, and camera tracking. Through these systems, the teacher can observe in real-time the student's attention to music in the virtual environment, the state

of continuity of music learning, etc. Immersion can be used as a measure of learning efficiency, which facilitates better teacher supervision.

2.3.2. Imaginative

Imaginative means that the use of virtual reality technology can combine learners' rational and perceptual cognition of music and stimulate learners' musical imagination. For example, in the process of teaching foreign music appreciation or minority music appreciation, learners can be brought into a specific virtual music environment and experience the differences between the music through the period characteristics of the foreign music or minority music scenes to be appreciated. Sati Doganyigita and Omer Faruk Islim investigated the use of virtual reality in vocal training [13]. They suggested that when a musician is playing, he or she needs to first imagine the pitch and the possible changes in sound, and then control his or her breathing based on this imagery. They recreated the abstract pitch changes in a VR environment with a roller coaster ride of ascent and descent.

2.3.3. Interactivity

Interactivity refers to the need for users to interact naturally with the virtual music environment using devices such as virtual joysticks instead of traditional tools such as mice and keyboards while receiving feedback from the music environment. research by Sebastian Rutkowski and Mateusz Adamczyk et al has demonstrated that training people in VR environments driven by several interactions can improve musicians' hand-eye coordination responses and reaction times [14]. Ko Okasaki and Mitsunori Makino have developed a VR-based self-training portable system for piano in which learners play a virtual keyboard by manipulating virtual fingers [15]. With the current lack of haptic support in the field of VR instrumental music teaching, precise haptic feedback between the fingers and the instrument could greatly enhance interactivity and make virtual teaching more effective.

Overall, VR musical education is growing rapidly, and many schools are already experimenting with this new educational model. The combination of VR and music education has been studied in different educational directions of VR environments, which have a focus on immersion, imagination, and interaction, but no mature VR music education platform has emerged due to technical limitations and the high price of VR equipment. Even so, researchers still have a great interest in VR musical education and continue to explore the future of VR musical education.

3. Spatial Sound In Virtual Reality

Imitating the real is required in a virtual environment. For playback in virtual environments, considerable repositories of audio are often constructed to record real-world sounds. Generally, when we hear these sounds played back in virtual environments, we may recognize them and consider them to be "realistic". Sound characteristics are only static during recording, and cannot change with the dynamic virtual environment. Many of the captured sounds do not correspond to what the listener might expect to hear. The physical characteristics of these sounds have not been restored.

The audio techniques that are employed for interactive multimodal systems differ significantly from offline techniques, which aim solely at precision and authenticity. Finding techniques that strike the right balance between them remains a challenge. This section provides an overview of the state of the art of sound technologies in virtual environments including sound localization and sound propagation.

3.1. Sound Localization

In many cases, headphones may prove sufficient as a playback medium for the HMD acoustic environment, however, loudspeaker playback may be preferred to preserve the user's head-related transfer function or to allow the user to utilize hearing aid devices or other ear-mounted devices. HMDs have been shown to alter acoustic signals in current research. As sound travels into the ear, the shape of the ear, head, and body affects the sound quality. This results in differences between the

two ears' level, interaural time, and spectral characteristics, which are used to identify the sound source [16].

3.1.1. Binaural Rendering

The ability to process sounds arriving at both ears into complex auditory scenes is called binaural hearing (see Figure 3). There are three types of headphone simulations, according to Shinn-Cunningham and Shilling, that are diotic displays, dichotic displays, and spatialized audio. [17].

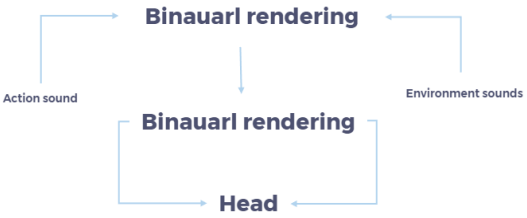


Figure 3 A schematic representation of the way of interpretation of sound.

3.1.2 Head-Related Transfer Function (HRTF)

HRTF is an acoustic function that reflects the structure of the head and is often used with Ambisonics to build high-performance binaural renderers, but its performance is not very satisfactory for generic HRTF, so there is an increasing amount of research based on individual HRTF. Mengfan Zhang and Tianyi Guan presented a modeling individual's HRTF [18]. They obtained anthropometric parameters by taking some photos of individuals and used deep neural networks to model the morphology of individuals.

3.2. Sound Propagation

It is the simulation of sound's movement from the source to the ears of the receivers that is referred to as sound propagation. As part of the acoustic modeling process of sound propagation, the major difficulty is computing reflection paths from a sound source to a receiver. Due to the fact that there are many paths through which sound may travel from source to receiver, accurate modeling is a very computationally expensive process. Room impulse response (RIR) convolved with the dry signal is a simple but limited approach to simulating the space. There have been several geometric, wave-based, and hybrid approaches to modeling sound propagation. In order to use wave-based techniques in real-time simulations, more efficient solutions must be found that can reduce computational and mathematical costs.

Carl Schissler, Gregor Mückl, and Paul Calamia presented a fast diffraction pathfinding method for sound propagation based on geometric acoustic simulations. The method uses a unique mesh preprocessing pipeline and probabilistically explores possible paths, except that the method only calculates direct diffraction between source and receiver without considering reflections and diffusions, and in addition has some inherent drawbacks of the UTD diffraction model. Rui Ding and Shiguang Liu presented a simulation method for sound propagation based on underwater scenes in a virtual environment [19]. Mixing the normal mode method and ray-tracing method. Dividing the scenes in the preprocessing stage, different methods are used for calculation in different regions. However, this method lacks support for scenes with irregular boundaries.

4. Conclusion

Our environment is constantly filled with sounds, from street noise to music playing on the radio, but these are often neglected in interactive environments. Interactive immersive environments have the ability to lead to enhanced degrees of engagement and focus, as sound links multiple senses in an integrated manner. Many forms of entertainment rely on graphics for engaging and communicating with their audiences. However, the subtle experience of sound in an interactive environment is something we cannot ignore.

Firstly, this paper summarizes the characteristics and applications of musical VR. For music, virtual reality has the greatest potential when it facilitates experiences that cannot be replicated in the real world. Furthermore, novel musical experiences can be created through the use of VR technologies beyond conventional musical interfaces, such as analog or digital musical instruments. Although there have been many outstanding musical VR applications. Musical VR as a field is still in its infancy. It is important to consider how different stakeholders to research interaction with virtual elements and spaces in musical contexts to design future musical VR systems. Of particular importance is the perception of sound within multimodal VR environments, how the interaction between objects and the virtual bodies affects our sense of self and sense of control over musical events in virtual environments and how we communicate subtle cues and affectations in those environments with each other.

Secondly, this paper presented an overview of the state of the art of spatial sound for virtual environments. Many sound events have been synthesized from the point of view of the source model, and real-time simulations are available as well. The commercial software engines normally used by the VR community have not yet been incorporated into these sound simulations. In spite of the limitations of recordings, such as the fact that they do not show the same flexibility as simulations, they are still the first choice for many music designers.

References

[1] B. Kenwright, There's More to Sound Than Meets the Ear: Sound in Interactive Environments, in IEEE Computer Graphics and Applications, vol. 40, no. 4, 2020, pp. 62-70, 1. doi: 10.1109/ MCG. 2020.2996371.

[2] W Brinkman, A.R. Hoekstra, The Effect Of 3D Audio and Other Audio Techniques on Virtual Reality Experience. Studies in health technology and informatics, 2019, pp.44-48. https:// doi.org/10.1016/j.jbtep.2016.03.004

[3] V Best, R Baumgartner, M Lavandier. Sound Externalization: A Review of Recent Research. Trends in Hearing. January 2020. doi:10.1177/ 2331216520948390.

[4] L. Turchet, R. Hamilton and A. Çamci, Music in Extended Realities, in IEEE Access, vol. 9, 2021, pp. 15810-15832, doi: 10.1109/ACCESS.2021 1.3052 931.

[5] D. Mills, F. Schroeder, J. D'Arcy, GIVME: Guided Interactions in Virtual Musical Environments: In NIME 2021. https://doi.org/10.21428/92fbeb44. 5443652c

[6] A. Hofmann, Study for virtual keyboard instrument and hand tracking in a VR environment. In NIME 2021. pp. 261-267. https://doi.org/10.21428 /92fbeb44.2e31abfe.

[7] W. Costa et al., Songverse: A Music-Loop Authoring Tool Based on Virtual Reality, 2019 21st Symposium on Virtual and Augmented Reality, 2019, pp. 216-222, doi: 10.1109/SVR.2019.00046.

[8] I. Findlay-Walsh, Virtual auditory reality: Inhabiting digital pop music as simulated space. Sound Effects - An Interdisciplinary Journal of Sound and Sound Experience, 10(1), 2020, pp.71–90. https://doi.org/10.7146/se.v10i1.124199

[9] A. Çam, R. Hamilton, Audio-first VR: New perspectives on musical experiences in virtual environments, Journal of New Music Research, 2019 49, 1, pp.1-7, DOI: 10.1080/0929821 5.2019.1707234

[10] A. Miskinis, Virtual Harmony: Music Interaction with Virtual Reality to Reduce Stress. https://oaks.kent.edu/node/10374

[11] H Luo, G Li, Q Feng. Virtual reality in K-12 and higher education: A systematic review of the literature from 2000 to 2019. Computer Assist Learn. 2021, 37 pp.887–901. https: //doi.org/ 10.1111/jcal.12538

[12] W. Chen Research on the Design of Intelligent Music Teaching System Based on Virtual Reality Technology. Comput Intell Neurosci. 2022 Apr 7; 2022:7832306. doi: 10.1155/2022/7832306.

[13] S. Doganyigit, Islim O. F., Virtual reality in vocal training: a case study, Music Education Research, 2021, pp. 465-471. DOI: 10.1080/146138 08.2021.1879035

[14] S. Rutkowski; M. Adamczyk, A. Pastuła, E. Gos, Training Using a Commercial Immersive Virtual Reality System on Hand–Eye Coordination and Reaction Time in Young Musicians: A Pilot Study. Int. J. Environ. Res. Public Health 2021, 18, 1297. https:// doi.org/10.3390/ijerph18031297

[15] Ko Okasaki, Mitsunori Makino, A VR-based piano self-training portable system on standalone HMD," Proc. SPIE 11766, International Workshop on Advanced Imaging Technology 2021, pp.1176-1183; https://doi.org/10.1117/12.2591031

[16] S. Serafin Sonic Interactions in Virtual Environments. In Encyclopedia of Computer Graphics and Games. Springer, Cham. 2016, pp. 580-589. https://doi.org/10.1007/978-3-319-08234-9_243-1

[17] R. D. Shilling, B. Shinn-Cunningham, Virtual auditory displays. Handbook of virtual environment technology, 2002, pp. 65-92. DOI: 10.1109/HVET.1.2002.747064.

[18] C. Schissler, G. Mückl, P. Calamia. Fast diffraction pathfinding for dynamic sound propagation, ACM Transactions on Graphics, Article 138 2021, pp.13-26. https://doi. org/10.1145/3450626.3459751

[19] R. Ding, S. Liu, Underwater sound propagation for virtual environments. Visual Computer 37, 2021, pp. 2797–2807. https://doi.org/10.1007/s00371-021-02175-6