A Study of Realistic 3D Sound Rendering using Sound Tracing for Mobile Virtual Reality

Eunjae Kim*, Juwon Yun*, Woonam Chung*, Youngsik Kim†,
Cheong Ghil Kim††, and Woo-Chan Park*
*Sejong University, Korea
†Korea Polytechnic University, Korea
††Namseoul University, Korea

***Abstract*— During the past decade, there has been a rapid development of audio and video technologies in many areas. Furthermore, the commencement of commercialization of 5G service ultimately comes down to the inherent belief that smartphone VR is going to be the primary driver in the market. Here, 3D audio technology which is mainly used to make computer games more realistic will soon be applied to the next generation of mobile devices and expected to offer a more expansive experience than its visual counterpart. This paper introduces the 3D sound technologies covering algorithms and systems for modeling based on multichannel methods and physics-based geometric algorithms. Also, it introduces an audio rendering pipeline to a scene graph-based virtual reality system and a hardware architecture to model sound propagation.**

I. Introduction

Nowadays, the importance of VR is being revisited in the mobile industry. Especially the commencement of commercialization of 5G [1] service ultimately comes down to the inherent belief that smartphone VR is going to be the primary driver in the market. As for the 3D sound product, 360 Reality Audio [2] sound technology with object-based audio space technology was introduced. This technology adds distance, angle, and location information to sound elements such as voices, choruses, and instruments, and scatters them appropriately in the space around the user. Fig. 1 shows the concept of spatial sound of 360 Reality Audio [2], in which spatial sound create multiple virtual speakers not like the conventional stereo in which the speakers are fixed in number and position.

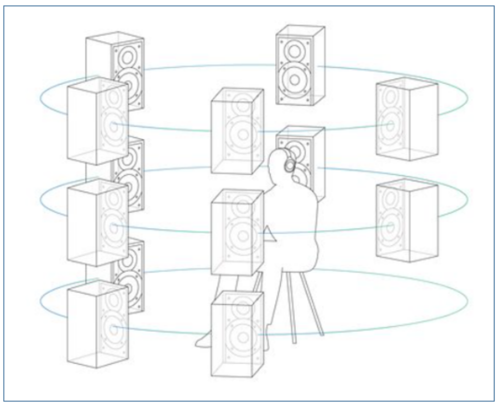

Fig. 1 The concept of spatial sound technique of 360 Reality Audio

VR technology, which moves reality and reproduces it in virtual space, is expected to remove the time and space constraints in video and game contents and industrial simulators. VR content represents a 360º full-angle view in real time. Therefore, sound must also be expressed in omnidirectional immersive sound technology.

Even though the recent interest in the virtual reality field is rapidly increasing, most VR technologies are focusing on the visual element. However, in order to support the realistic virtual reality environment, it is essential to reproduce the auditory space in addition to the visual space. Techniques used to reproduce auditory spatiality include multichannel audio systems, head related transfer functions, and sound rendering based on three-dimensional geometric models [3, 4].

Multichannel audio used for 3D sound reproduction has physical constraints such as space and cost for dedicated speaker systems. In addition, the HRTF used in the head mount display (HMD) does not consider the physical environment of the sound because it does not consider the environment and objects.

The sound rendering technology based on the 3D geometry model simulates by reflecting the position of the listener, the number and position of sound sources, and the surrounding objects and materials in the virtual space. This reproduces the physical properties of sound, such as reflection, transmission, diffraction, and absorption, providing a sense of auditory spatiality to the user. However, physics-based sound simulation of surrounding objects and materials in real time is expensive to calculate and consumes high power [5]. In mobile environments with limited resources, or in standalone HMDs, it becomes difficult to provide realistic sound immersion to users when it becomes difficult to process sound rendering in real time. Therefore, it is essential to reduce power consumption and computational cost in order to provide auditory space on various platforms.

Fig. 2 shows an overview of sound rendering which is simulating acoustic reality in a virtual world such as the goal of graphics is simulating visual reality in virtual world. As you can see, it consists of 3 major components and there are lot of things that need to be covered [6, 7].

The rest of the paper is organized as follows. In Section 2, we review the basic concept of 3D sound and system. Section

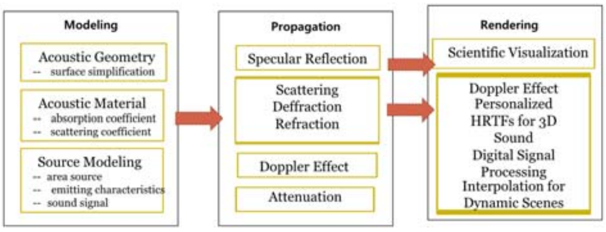

Fig. 2 Three sound rendering components

3 covers computational modeling approach based on geometrical simulations method with a new hardware architecture for sound propagations unit. Section 4 concludes this paper.

II. General Instructions

A. Multi-channel Audio System

In general, the multi-channel audio system or 3D sound technology with HRTF (head-related transfer function), is one method to reproduce this acoustic spaciousness [3, 8]. That is, HRFT can describe process that an ear receives a sound from a point where sounds are generated in space. However, it is known that HRTF based 3D sound technologies have some limitations to reproduce realistic sound properly. This is because they do not simulate the physical effects of sound between many sound sources and the listener in the virtual 3D space.

B. Sound Tracing

To overcome the limits of multi-channel audio system, 3D sound technology based on geometric method has been introduced [3, 7]. In the geometry methods, the convergence method combining ray-tracing in 3D graphics and sound-processing is called sound-tracing. Sound-tracing is a method to create the realistic sound by tracing propagation paths between listener and sound source. The basic sound tracing pipeline consists of two stages of acoustic modeling and auditory display shown in Fig. 3.

The input to an auralization system is a description of a virtual environment, an audio source location, an audio receiver location, and an input audio signal. The auralization system computes a model for the propagation of sound waves through the environment and constructs digital filter(s) that encode the delays and attenuations of sound traveling along different propagation paths. Convolution of the input audio signal with the filter(s) yields a spatialized sound signal for output with an auditory display device [3, 9].

C. Sound Propagation

In general, modeling sound propagation is a process of finding a solution to an integral equation expressing the wave field of points in space in terms of surrounding surfaces. There are difficulties with modeling sound scattering waves in a 3D environment. Sound waves traveling from a source to a destination travel along a variety of propagation paths with different sequences of reflections, diffractions, and refractions

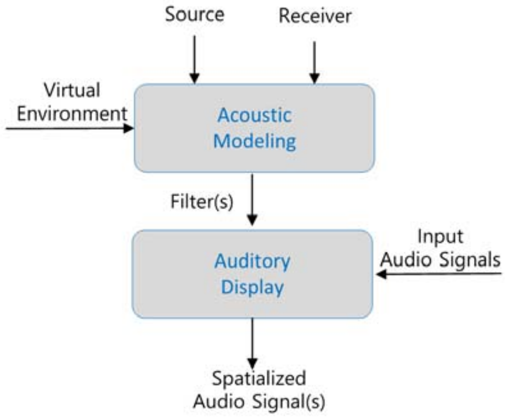

Fig. 3 Sound tracing pipeline

at surfaces of the environment. And we can notice that sound and light are both wave phenomena, modeling sound propagation is similar to global illumination.

The effect of these propagation paths includes reverberation to the original source signal as it reaches the receiver. So, auralizing a sound for a particular source, receiver, and environment can be achieved by applying filter(s) to the source signal that model the acoustical effects of sound propagation and scattering in the environment [10] shown in Fig. 4.

There are two ways to calculate sound propagation, the numerical method and the geometric method [9]. The numerical method is to directly solve wave equations [11, 12] in the form of quadratic partial differential equations. This is the most accurate method, but the cost of solving the equation is not suitable for interactive applications such as games. In spite of these difficulties, methods for solving sound propagation have been proposed to accelerate numerical methods, such as the finite-difference time-domain (FDTD) method [12] and the frequency-domain [13] method. However, these methods are limited to static scenes.

Finally, the sound generation step regenerates the input sound based on the configuration of the speaker to the listener,

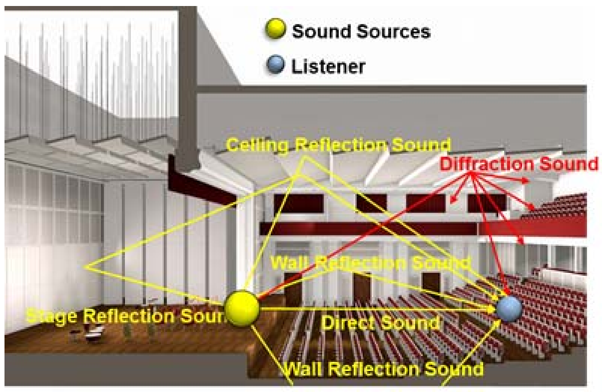

Fig. 4 Ray tracing method

using the physical characteristics of the sound such as the effective path and reflection, transmission, absorption coefficients, distance attenuation, and time obtained in the previous step. It's a step. Of the three phases in the sound tracing pipeline, the sound propagation phase is expensive and there is a need to accelerate the sound propagation phase.

III. PROPOSED ARCHITECTURE

A. *Overall System Architecture*

This section introduces the overall hardware architecture of sound tracing shown in Fig. 5. It consists of Sound Propagation Unit (SPU) and Success Path Buffer (SPB). SPU is a unit that performs GA based sound propagation. SPB stores sound propagation paths found by SPU and writes to external memory.

B. *Sound Propagation Unit*

Fig. 6 shows the overall architecture of the proposed Sound Propagation Unit (SPU). The SPU datapath consists of a unit that performs a general ray tracing algorithm and a unit that performs sound propagation processing. The ray tracing part includes Ray Setup Processor, Ray Generator, Traversal & Intersection Test Unit (T & I), and Hit Point Calculator. The sound processing part includes Propagation Path Validator, Reverb Geometry Collector, Path Cache Validator, Path Impulse Response (IR) Calculator, and Reverb IR Calculator. The memory part consists of T&I cache memory (Node/TriVertex cache, List cache), Path Cache, Sort Working Memory, Success Path Buffer (SPB), and Internal Register File. The external memory is configured with acoustic geometry data, acceleration structure data, sorted LR triangle data, path IR data, and reverb IR data.

IV. SIMULATION RESULTS

The proposed Sound Tracing Unit was implemented into D

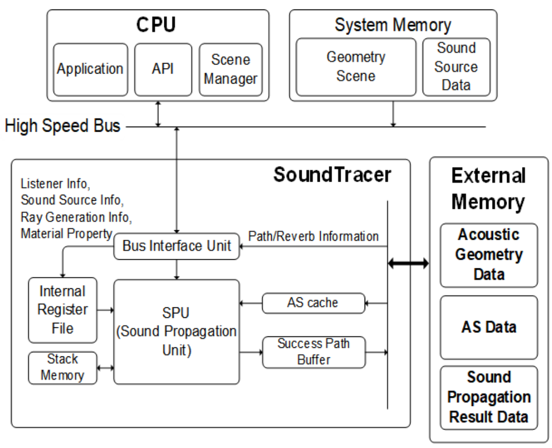

Fig. 5 Sound tracing hardware architecture

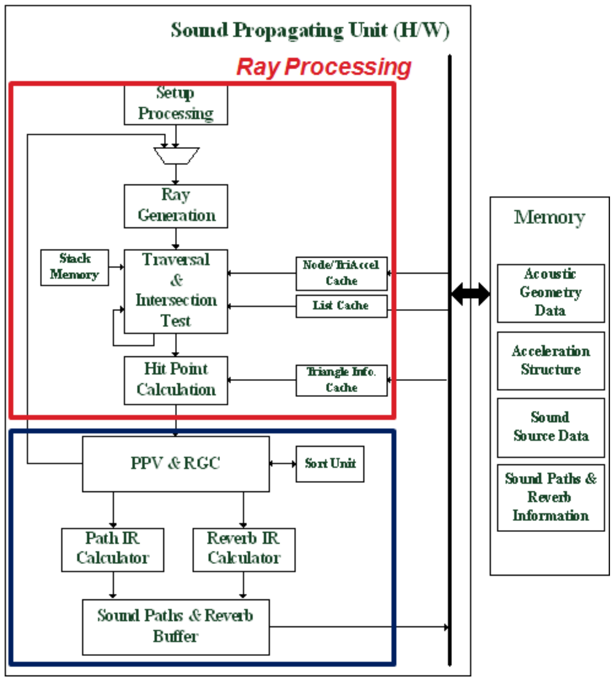

Fig. 6 Sound Propagation Unit

ynalith Systems iMPRESS board containing Xilinx Artix7-200T FPGA chip shown in Fig. 7 and performed an emulation test by combining the proposed SPU to the commercial game engine. It can operate with low power and high performance. As a result of performance experiments, our sound rendering system achieved performance that can be processed in real time when implemented in ASIC.

V. CONCLUSIONS

his paper introduces the 3D sound technologies covering algorithms and systems for modeling based on multichannel methods and physics-based geometric algorithms. Also, it introduces an audio rendering pipeline to a scene graph-based virtual reality system and a hardware architecture to model sound propagation. The proposed system was implemented on

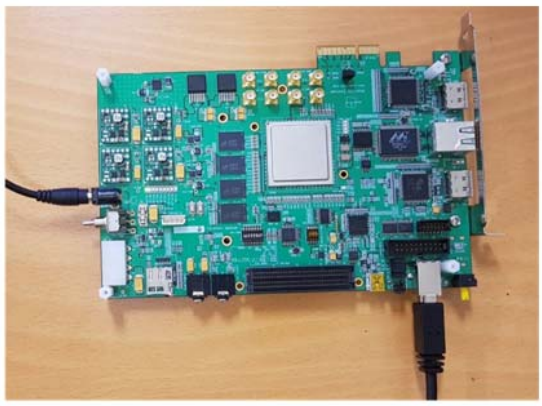

Fig. 7 Implementation with Dynalith Systems iMPRESS

an FPGA chip with a real-time rate and low power consumption. Conclusions

REFERENCES

[1] IEEE 802 5G/IMT-2020 Standing Committee [Online]. Available: http://www.ieee802.org/Stand_Com/5G/
[2] 360 Reality Audio [Online]. Available: https://www.sony.com/electronics/360-reality-audio
[3] M. Vorlander, *Auralization: Fundamentals of Acoustics, Modelling Simulation, Algorithms and Acoustic Virtual Reality*, Springer, 2008.
[4] A. Febrettia, A. Nishimotoa, T. Thigpena, J. Talandisa, L. Longa, J. Pirtlea, T. Peterkaa, A. Verloa, M. Browna, D. Plepysa, D. Sandina, L. Renambota, A. Johnsona, and J. Leigha, "Cave2: a hybrid reality environment for immersive simulation and information analysis," in Proceedings SPIE Electronic Imaging, vol. 8649, p. 864903, 2013.
[5] D. Hong., T. H. Joo, W.-C. Park. "Real-time sound propagation hardware accelerator for immersive virtual reality 3D audio," Proceedings of the 21st ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games, p. 20, 2017.
[6] Z. Y. Zhou, A. D. Cheok, Y. Qiu, and X. Yang, "The Role of 3-D Sound in Human Reaction and Performance in Augmented Reality Environments", IEEE Transaction on Systems, MAN, and Cybernetics—PART A: Systems and Humans, 37 (2) (2007) 262-272
[7] F. Keyrouz, K. Diepold, "An Enhanced Binaural 3D Sound Localization Algorithm", 2006 IEEE International Symposium on Signal Processing and Information Technology, (2006) 662-665
[8] J. Herre, J. Hilpert, A. Kuntz, J. Plogsties, MPEG-H Audio—The New Standard for Universal Spatial/3D Audio Coding, JAES Volume 62 Issue 12 pp. 821-830; December 2014
[9] T. Funkhouser, N. Tsingos, J. Jot, "Survey of Methods for Modeling Sound Propagation in Interactive Virtual Environment Systems, [Online]. Available: http://www-sop.inria.fr/reves/Basilic/2003/FTJ03
[10] C. G. Kim, A Study of 3D Sound Modeling based on Geometric Acoustics Techniques for Virtual Reality," *Journal of satellite, information and communications*, 11 (4) 2016, 102-106
[11] L. Savioja, "Real-time 3D finite-difference time-domain simulation of low-and mid-frequency room acoustics," 13th Int'l Conf. on Digital Audio Effects, vol. 1, p. 75, 2010.
[12] R. Mehra, N. Raghuvanshi, L. Antani, A. Chandak, S. Curtis, D. Manocha, "Wave-based sound propagation in large open scenes using an equivalent source formulation," *ACM Transactions on Graphics (TOG)*, vol. 32, no. 2, pp 19:1-19:13 2013.
[13] T. Funkhouser, I. Carlbom, G. Elko, G. Pingali, M. Sondhi, J. West, "A beam tracing approach to acoustic modeling for interactive virtual environments," In Proceedings of the 25th annual conference on Computer graphics and interactive techniques, pp. 21-32, 1998.