3D Audio Down-Mixing System for Immersive Realistic Virtual Reality

Dukki Hong, Seiyoung Lee, Woo-Chan Park*
Department of Computer Engineering
Sejong University, Seoul, Korea
{dkhong, sylee} @rayman.sejong.ac.kr,
pwchan@sejong.ac.kr

***Abstract* –For realistic immersive VR environment, it is necessary to have high-quality reproduction of virtual space with visual as well as acoustic spaciousness. This paper proposes the architecture of a 3D audio down-mixing system based on sound-tracing. The proposed system can reproduce high-quality acoustic spaciousness by simulating the physical effects of sound between many sound sources and the listener in the virtual 3D space. The experimental results show that future implementation with ASIC could achieve real-time performance with dozens of sound sources.**

Keywords - 3D Audio, audio mixing system, sound rendering, ray tracing, FPGA

I. Introduction

To achieve a realistic immersive VR environment, it is necessary to have high-quality reproduction of virtual space with visual as well as acoustic spaciousness. The multi-channel audio system or 3D sound technology, which uses HRTF, is one method to reproduce this acoustic spaciousness [1, 2], but there are problems: adequate space is required for a dedicated installation; a low-quality sound could be reproduced.

Most of the 3D sound technologies based on HRTF have the limits to reproduce realistic sound because they do not simulate the physical effects of sound between many sound sources and the listener in the virtual 3D space. To overcome the limits, the leading world companies and universities are announcing 3D sound technology based on geometric method or numeric method [3-6]. 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.

In this paper, we propose a 3D audio mixing system based on sound-tracing. The proposed system can reproduce high-quality acoustic spaciousness by simulating the physical effects of sound between many sound sources and the listener in the virtual 3D space. With a description of a 3D scene, it is possible for users to listen to various reproduced sounds from even the same sources. The reason is that a VR environment and an object's material properties could be taken into account when simulating the physical effects of sound.

II. THE PROPOSED ARCHITECTURE

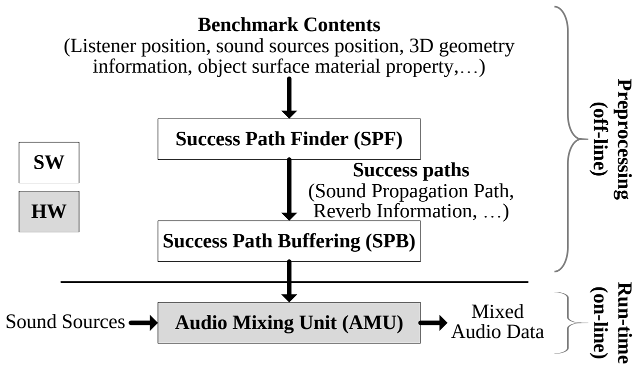

Figure 1. Overall architecture of proposed audio mixing system

Figure 1 shows the overall architecture of the proposed audio down-mixing system. This architecture is composed of a success path finder (SPF), success path buffering (SPB), and audio mixing unit (AMU). The SPF finds the path that the sound takes successfully to reach the listener. The SPB is a buffer to store sound paths found and information related. The AMU generates mixed audio data using stored path/reverb information. The SPF and SPB are processed on off-line mode by software; whereas, the AMU is processed on-the-fly by hardware.

Figure 2 shows the architecture of the AMU hardware. The AMU is composed of a crossover unit, path rendering unit (PRU), reverb rendering unit (RRU), and mixer. The crossover unit divides loaded PCM data from memory into

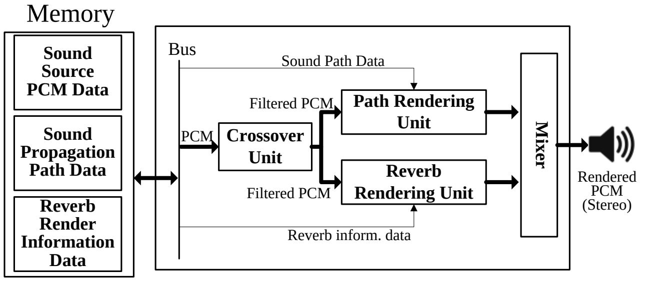

Figure 2. The proposed AMU H/W architecture

each frequency band range, and transfers the data to the PRU and RRU respectively. With this transferred data and success sound path, the PRU generates sound data that are applied with the physical properties of the sound. The RRU generates the reverb sound data using transferred data and reverb information. The sound data, which are output to speakers or headphones, are mixed in the mixer with the output data of the PRU and RRU.

III. Experiment Results

Table 1 shows the experimental results of the proposed system. The proposed system was implemented on a Xilinx FPGA Artix-7 XC7A200T. The benchmark 5.1 channel sound sources for performance estimation were extracted from the Lord of Rings: The Fellowship of the Ring movie on DVD. The sources of sample rate, format, and length are 48 KHz, signed PCM 16-bit, and 10 seconds, respectively. Using the SPF with extracted sound sources, 10 success sound paths were extracted from each sound source.

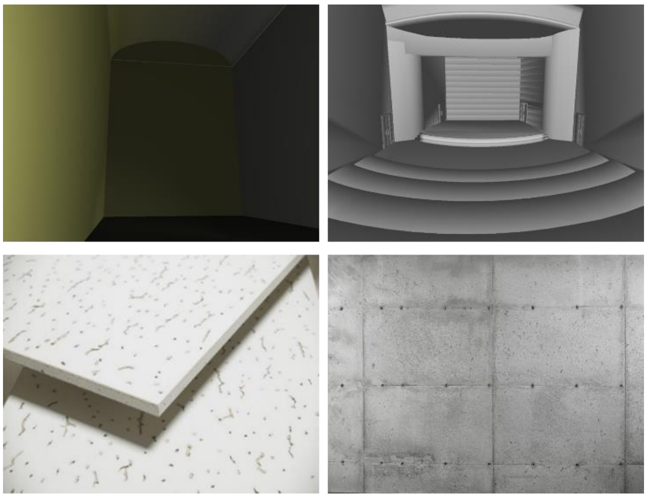

Figure 3. The benchmark scenes and materials: Dome, Concert Hall, Gypsum Board, and Concrete

Table 1 Experimental result of the proposed architecture

Item	Description	
FPGA Chip	Xilinx Artix-7 XC7A200T	
FPGA Clock Speed	50 MHz	
FPGA Utilization	Logic Utilization	34 %
	BRAM Utilization	53 %
Processing Time (5 sound sources)	114.29 ms (4.19M sample per sec, 59.59 fps)	
Sound Effects	Diffraction, Reflection, Transmission, Absorption, Reverberation	
Supported Materials	Metal, Wood, Concrete, etc.	

As the experimental results show, the proposed FPGA-based system can support up to five sound sources in real-time performance. Based on current implementation, we expect that the ASIC of the proposed system is able to process dozens of sound sources in real-time.

IV. Conclusion

We proposed the audio down-mixing system based on sound-tracing. The proposed system can reproduce high-quality acoustic spaciousness by simulating the physical effects of sound between many sound sources and the listener in the virtual 3D space. The current version of AMU H/W has relatively high BRAM utilization that we plan to reduce BRAM utilization in the future.

Acknowledgement

This research was supported by the MSIP(Ministry of Science, ICT and Future Planning), Korea, under the ITRC(Information Technology Research Center) support program (IITP-2016-R2718-16-0017) supervised by the IITP(Institute for Information & communications Technology Promotion). The CAD tools were supported by IDEC.

References

[1] M. Vorlander, Auralization: Fundamentals of Acoustics, Modelling, Simulation, Algorithms and Acoustic Virtual Reality, Springer, 2008.

[2] A. Febretti et al., "Cave2: a hybrid reality environment for immersive simulation and information analysis," in *Proceedings SPIE Electronic Imaging 2013*, Burlingame, CA , 2013, 864903-864903-12.

[3] L. Antani et al., "Next-Gen Sound Rendering with OpenCL," in *ACM SIGGRAPH 2011 BOF OpenCL*, Vancouver, BC, 2011.

[4] B. Hook and T. Smurdon, "An Intro to Virtual Reality Audio," in *GDC 2015*, San Francisco, CA, 2015.

[5] N. Ward-Foxton, "Environmental Audio and Processing for VR," in *GDC 2015*, San Francisco, CA, 2015.

[6] NVIDIA. (2016, May). *NVIDIA VRWorksAudio* [Online]. Available: https://developer.nvidia.com/vrworks