

# An FPGA-based Sound Field Renderer for High-Precision Sound Field Auralization

Yiyu Tan<sup>†</sup>, Xin Lu<sup>†</sup>, Guanghui Liu<sup>\*</sup>, Peng Chen<sup>§</sup>, Truong Thao Nguyen<sup>§</sup>, Yusuke Tanimura<sup>§</sup>

<sup>†</sup> Department of Systems Innovation Engineering, Iwate University, Japan

<sup>\*</sup> F.Widjaja Foundation Inflammatory Bowel and Immunobiology Research Institute, Cedars-Sinai Medical Center, USA

<sup>§</sup> National Institute of Advanced Industrial Science and Technology, Japan

{tanyiyu, luxin}@iwate-u.ac.jp, guanghui.liu@cshs.org, {chin.hou, nguyen.truong, yusuke.tanimura}@aist.go.jp

## 1. INTRODUCTION

Realistic sound field auralization can significantly improve the sense of presence and immersion, situational awareness of users, and augment the visual sense in human-computer interaction realistic communication, virtual and augmented realities. In sound field auralization, highly accurate room impulse response is critical to achieve precise auralization results. To date, although many methods have been proposed to simulate room impulse response, geometrical methods are widely applied in real-time sound field auralization because of their low computational load at price of accuracy. In contrast, wave-based methods provide high accuracy, but require much higher computational capability since spatial grids are oversampled to suppress dispersion errors. Generally, the required computing capability is increased as the fourth power of the sampling frequency and proportionally with the volume of sound spaces. In this work, an FPGA-based sound field renderer was investigated for high-precision real-time sound field auralization, in which the wave-based rendering algorithm was applied to obtain room impulse response accurately, and dedicated hardware was developed to speed up computation.

## 2. SYSTEM DESIGN

As shown in Fig. 1, a sound field auralization system consists of the sound field renderer, input and output interfaces. The sound field renderer computes the room impulse response according to the features of a virtual scene like dimensions and boundary conditions. The room impulse response is convoluted with the input sound signals and the results are performed binaural auralization by convoluting with the head related transformations functions. The auralization results are finally output to drive speaker system. In this procedure, the sound field renderer is the affects the auralization performance significantly because it provides accurate room impulse response for further convolution.

**Sound field rendering algorithm.** The hardware-oriented FDTD algorithm [1-3] shown in Equation (1) is applied to update sound pressure of each node.

$$P^n(i, j, k) = D1 \times [P^{n-1}(i-1, j, k) + P^{n-1}(i+1, j, k) + P^{n-1}(i, j-1, k) + P^{n-1}(i, j+1, k) + P^{n-1}(i, j, k-1) + P^{n-1}(i, j, k+1) + 2P^{n-1}(i, j, k)] - D2 \times P^{n-2}(i, j, k) \quad (1)$$

where  $P$  is the sound pressure of a node  $(i, j, k)$ ,  $D1$  and  $D2$  are parameters chosen according to the position of a node. Equation (1) shows that sound pressures of a node and its six neighbors are required to compute the sound pressure of a node.

**System design and implementation.** Sound field rendering with FDTD methods is memory-bound, and the proposed system was designed using OpenCL. The spatial blocking and temporal blocking were applied to alleviate the required external memory bandwidth and reuse data, respectively. The system consisted of Data input module, computation engine, and Data output module.

The computation engine contained 16 processing elements (PEs). Each PE computed sound pressures of nodes in a spatial block at a time-step, and 16 PEs were cascaded to compute all sound pressures in the same spatial block at continuous 16 time-steps.

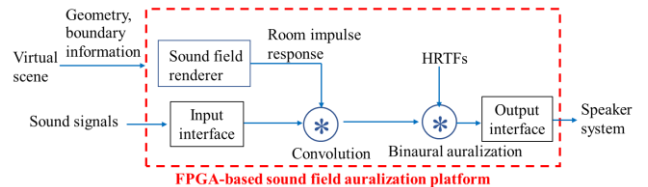


Fig. 1 Sound field auralization system

## 3. PERFORMANCE EVALUATION

Table 1 presents the rendering time per time step taken by the proposed FPGA-based sound field renderer implemented using the FPGA card DE10-Pro and software simulations performed on a desktop machine with 512 GB DRAMs and an Intel Xeon Gold 6212U 24-core processor running at 2.4 GHz. The sound space was a three-dimensional shoebox with dimension being 16m×8m×8m. The incidence was an impulse and the number of computed time steps was 32. The reference C++ codes in software simulations were compiled using the GNU compiler (version: 4.8.5) with the option `-O3` and `-fopenmp` to use all 24 processor cores. As shown in Table 1, the proposed FPGA-based sound field renderer outperforms the software simulations by 11 times even though it runs at about 350 MHz.

Table 1. Rendering Time Per Time Step (s)

FPGA	Software simulation
0.0486	0.5363

## ACKNOWLEDGMENTS

Thanks for Intel's donation of the FPGA card DE10-Pro and the EDA tools through University Program. This work was supported by JSPS KAKENHI Grant Number JP19K12092 and JP22K12123.

## REFERENCES

- [1] Y. Tan, T. Imamura, and M. Kondo. FPGA-based acceleration of FDTD sound field rendering. *Journal of the Audio Engineering Society*, 69[7/8] (2021), 542-556.
- [2] T. Yiyu, Y. Inoguchi, M. Otani, et al. A real-time sound field rendering processor. *Applied Science*, 8[35] (2018).
- [3] Y. Y. Tan, Y. Inoguchi, Y. Sato, et al (2014). A real-time sound rendering system based on the finite-difference time-domain algorithm. *Japanese Journal of Applied Physics*, 53 (2014), 07KC14.