Real-Time Acoustic Impulse Response Generation Using Static Ray Tracing for Anomaly Detection

Oleksandr Terletskyy[1], Valeriy Trushevskyy[1]

[1] *Ivan Franko National University of Lviv, Universytetska 1, 79000 Lviv, Ukraine*

Abstract

This paper presents a real-time method for simulating sound propagation using ray tracing techniques, with applications in cybersecurity for detecting and analyzing acoustic threats. By modeling acoustic wave behavior in virtual environments, this approach enables accurate simulation of reflections, diffractions, and reverberations, which can be leveraged for tasks such as audio-based anomaly detection and spatial audio analysis in secure facilities. The integration of impulse response generation with real-time convolution processing allows for precise acoustic profiling, dynamically responding to changes in the environment to enhance situational awareness. This method provides a foundation for advanced acoustic monitoring systems, supporting applications ranging from secure communications to threat detection. Results demonstrate the method's ability to simulate complex acoustic interactions with high fidelity and responsiveness, making it a valuable tool for real-time audio surveillance and cybersecurity solutions.

Keywords

Sound propagation, ray tracing, convolution, impulse response, real-time audio, spatial audio

1. Introduction

Realistic sound propagation is essential for immersive environments in virtual reality, gaming, architecture, and cybersecurity. Acoustic behavior depends on reflections, diffractions, and reverberations, which traditional methods cannot adapt to in real time. This study presents a real-time simulation method that uses ray tracing to model sound paths and FFT based convolution to generate impulse responses [1]. The approach supports cybersecurity applications such as acoustic anomaly detection and source localization by accurately reproducing sound behavior in complex spaces. To ensure stability, the system employs static rays, reducing response fluctuations and audio artifacts. A running index mechanism updates impulse responses continuously, maintaining smooth playback as environmental data changes. This method provides a stable and adaptive solution for high-quality sound simulation and real-time acoustic analysis in both digital and security contexts

2. Static Ray Tracing Approach

Accurate impulse response generation is essential for simulating sound propagation and identifying acoustic patterns in areas such as cybersecurity, where it supports surveillance, anomaly detection, and secure communication [2]. The proposed method models sound using ray tracing, where sound rays travel through the environment, reflect from surfaces, and return to the listener. Each ray carries information on distance, direction, and energy decay, forming a time-dependent impulse response (IR) that represents the environment's acoustic signature. To maintain stability, static rays are used—each frame reuses the same set of directions. This avoids random fluctuations in the impulse response and ensures consistent data for real-time analysis. A running index updates the impulse response incrementally, dividing it into small segments for efficient processing. This

[1]Corresponding authors

oleksandr.terletskyi@lnu.edu.ua (O. Terletskyy); valeriy.trushevsky@lnu.edu.ua (V. Trushevskyy)

0000-0000-0000-0000 (O. Terletskyy); 0000-0001-8672-7152 (V. Trushevskyy)

structure supports long impulse responses and enables smooth adaptation to environmental changes with minimal computation.

2.1. Impulse Response

The impulse response $h(t)$ describes how an environment reacts to a short sound, defined by the convolution of the input signal $x(t)$ with $h(t)$ [3]:

$$y(t) = x(t) * h(t) = \int_{-\infty}^{\infty} x(\tau)h(t-\tau)d\tau, \tag{1}$$

where $y(t)$ is the output signal, $x(t)$is the input signal, $h(t)$ is the system's impulse response, and denotes the convolution operator, or in discrete form:

$$y[n] = \sum_{k=0}^{N-1} x[k] \cdot h[n-k], \tag{2}$$

where y(t) (or y[n]) is the output signal, x(t) (or x[k]) is the input signal, h(t) (or h[n]) is the impulse response, and * denotes convolution. Each ray contributes to h(t) based on its path length and energy loss, and summing all contributions forms an accurate model of the space's acoustic behavior.

2.2. Static Ray Tracing

Ray tracing simulates sound by following rays as they reflect and attenuate [1]. By introducing static rays with fixed directions across frames, the impulse response can be kept stable without noise artifacts. This precision is vital in acoustic monitoring and intrusion detection, where slight deviations can indicate security threats. Each ray originates from the sound source, reflects according to the surface normal, and loses energy over distance. Rays reaching the listener contribute to the impulse response with recorded travel time and intensity. The resulting data captures the environment's spatial characteristics for real-time anomaly detection.

2.3. Listener-Directed Diffuse Rain

The Listener-Directed Diffuse Rain model improves diffusion efficiency by sending rays toward the listener instead of scattering them randomly [4]. The energy loss is determined by the alignment between the incoming ray and listener vector:

$$E_{\text{loss}} = max((d \cdot \frac{l-p}{\parallel l-p \parallel})^2, 0), \tag{3}$$

where **d** is the ray direction, **p** the hit point, and **l** the listener position. The adjusted energy and direction are:

$$\mathrm{E}_{\text{adjusted}} = \mathrm{E}_{\text{original}} \times E_{\text{loss}}, \tag{4}$$

The direction of the adjusted ray is then normalized towards the listener:

$$d_{\text{new}} = \frac{l-p}{\parallel l-p \parallel}, \tag{5}$$

where **y(t)** (or **y[n]**) is the output signal, **x(t)** (or **x[k]**) is the input signal, **h(t)** (or **h[n]**) is the impulse response, and * denotes convolution. Each ray contributes to **h(t)** based on its path length and energy loss, and summing all contributions forms an accurate model of the space's acoustic behavior.

2.4. Impulse Response Generation

Each ray adds energy to the impulse response based on travel distance and attenuation. Arrival time:

$$t = \frac{d}{c}, \tag{6}$$

where d is the total distance traveled by the ray and c is speed of sound. As sound travels, it loses energy.

$$E_d = \frac{1}{d^2 + \epsilon}, \tag{7}$$

where ϵ is a small constant to avoid division by zero. The amplitude contribution of a ray is:

$$A = E_r \times E_d \times s, \tag{8}$$

where Er is the ray's remaining energy and s a phase-shift factor. After summing all rays, the impulse response is normalized to fit a standard range. The resulting $h(t)$ encodes the environment's acoustic behavior and can be convolved with sound for realistic and responsive spatial audio.

3. Auralization

Auralization is the process of converting simulated acoustic data into audible sound, allowing listeners to perceive an environment's acoustic characteristics. It relies on the impulse response (IR), which describes how a space modifies sound through reflections and reverberation. The original signal $x(t)$ is convolved with the IR $h(t)$ to produce the auralized output $y(t)$:

$$y(t) = x(t) * h(t), \tag{9}$$

where * denotes convolution. This operation embeds the environment's time and frequency features into the sound, recreating realistic spatial effects

3.1. Running Index Method

In real-time simulation, impulse responses must adapt continuously to moving sources and changing environments. Recomputing them each frame through multiple FFTs is computationally expensive. The running index method reduces this cost by dividing the impulse response into smaller segments and updating only a few per frame. A running index cycles through these segments, gradually refreshing the IR while maintaining stable sound quality. This incremental update approach balances accuracy and performance, ensuring smooth, uninterrupted auralization in real-time applications

4. Results

The simulation was conducted in a 50 × 50 × 20 m room with a source at the center and a listener positioned 1 m to the right. The room was perfectly reverberant, and the impulse response was recorded over 3 seconds at 41 kHz (123,000 samples). This setup was used to evaluate how the number of rays and reflections affects reverberation accuracy and density uralization is the process of converting

4.1. Number Of Rays

Impulse responses were generated using 10, 100, and 1000 rays, each with 8 reflections. As shown in Figure 1, increasing the ray count improves reflection density and the smoothness of decay over time. With 10 rays, the response is sparse and dominated by early reflections. Using 100 rays adds mid and late reflections, producing a more natural reverberant tail. At 1000 rays, the response becomes dense and continuous, accurately modeling the diffuse reverberant field. A higher ray count therefore yields a more complete representation of acoustic energy decay, essential for realistic spatial audio and precise acoustic analysis.

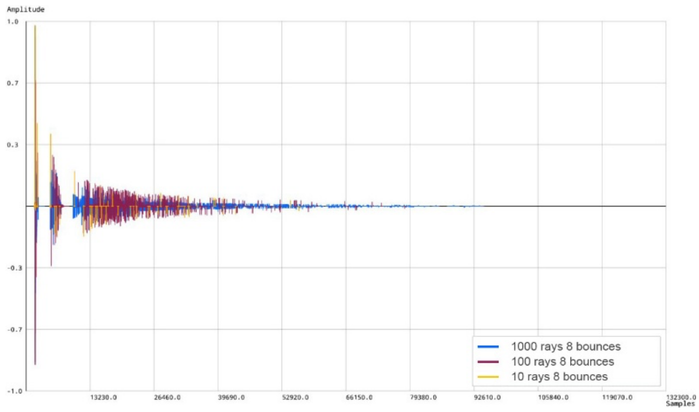

Figure 1: Impulse responses with 10, 100, and 1000 rays (8 reflections).

4.2. Number Of Rays

Figure 2 compares impulse responses for 2, 5, and 8 reflections using 1000 rays. With 2 bounces, only early reflections are present, and the decay is short. Increasing to 5 bounces extends the reverberant tail and improves density. At 8 bounces, late reflections fill the decay profile, approaching a continuous reverberant field. More bounces capture additional sound interactions, enriching realism and supporting applications that require accurate modeling of reflection persistence.

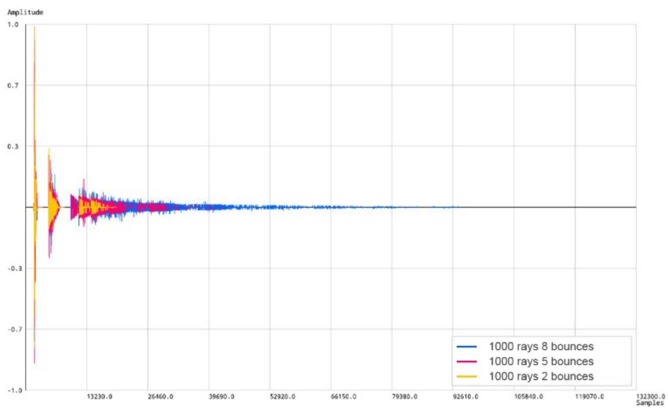

Figure 2: Impulse responses with 1000 rays and 2, 5, and 8 reflections.

4.3. Comparison with Sabine's Formula

The reverberation time $T60$ predicted by Sabine's formula [6]:

$$T_{60} = \frac{0.161 \times V}{A \times \alpha}, \tag{10}$$

where V is the room volume, A is the total surface area, and α is the average absorption coefficient. * denotes convolution was compared with simulated results for 2–10 bounces. For the given room (α = 0.1), the theoretical value is approximately 8.05 s. As shown in Figure 3, higher bounce counts produce decay times increasingly close to the theoretical $T60$. Around 8–10 bounces, the simulated results align closely with Sabine's prediction, confirming that additional reflections improve accuracy in modeling highly reverberant.

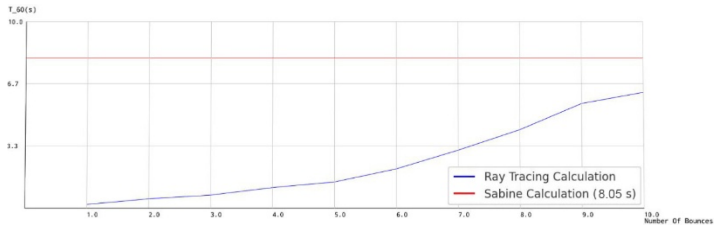

Figure 3: Simulated $T60$ versus Sabine's theoretical estimate

5. Conclusions

The results demonstrate that increasing both ray count and reflection depth enhances the precision and realism of impulse response simulations. Higher ray counts yield denser reverberation fields, while additional bounces extend the decay and capture complex reflections. These factors are critical not only for immersive spatial sound but also for acoustic analysis in cybersecurity, where detailed impulse responses can reveal subtle environmental changes or anomalies. The alignment with Sabine's theoretical reverberation time validates the simulation's accuracy. Such methods can improve acoustic monitoring, secure communication testing, and environmental profiling by providing reliable real-time detection of abnormal sound behavior. Overall, the static-ray approach combined with the running index method delivers stable, efficient, and physically consistent acoustic simulation suitable for both immersive and security-focused applications.

References

[1] A. Krokstad, S. Strom, and S. Sorsdal, “Calculating the acoustical room response by the use of a ray tracing technique,” J. Sound Vib., vol. 8, 1968, pp. 118–125.

[2] E. Ouellet-Delorme, H. Venkatesan, E. Durand, and N. Bouillot, “Live ray tracing and auralization of 3D audio scenes with vaRays,” in Proc. 18th Sound Music Comput. Conf., 2021, pp. 1–10.

[3] B. A. Farina, "Convolution of anechoic music with binaural impulse responses," Presented at the 15th Audio Engineering Society Conference on Audio in Large Spaces, Copenhagen, Denmark, 1993.

[4] S. Pelzer, D. Schröder, and M. Vorlaender, “The number of necessary rays in simulations based on geometrical acoustics using the diffuse rain technique,” in Proc. DAGA, 2011.

[5] Meta, “Acoustic ray tracing in Meta XR Audio SDK,” 2023. URL: https://developers.meta.com/horizon/documentation/unity/meta-xr-acoustic-ray-tracingunity-overview/

[6] L. Beranek, “Analysis of Sabine and Eyring equations and their application to concert hall audience and chair absorption,” J. Acoust. Soc. Am., vol. 120, 2006, pp. 1399–1410.