IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 11 August 2023, accepted 4 September 2023, date of publication 6 September 2023, date of current version 26 September 2023.

*Digital Object Identifier 10.1109/ACCESS.2023.3312560*

RESEARCH ARTICLE

# Efficient Propagation and Remapping of Sound Through a Geometric Approach in Virtual Environments and Terrains

**JONG-HYUN KIM[1], JUNG LEE[2], AND SUN-JEONG KIM[3]**

[1]Department of Design Technology, College of Software and Convergence, Inha University, Michuhol-gu, Incheon 22212, South Korea
[2]Department of Computer Engineering, Hanbat National University, Yuseong-gu, Daejeon 34158, South Korea
[3]Department of Convergence Software, Hallym University, Chuncheon 24252, South Korea

Corresponding author: Sun-Jeong Kim (sunkim@hallym.ac.kr)

This work was supported in part by the National Research Foundation of Korea (NRF) grant funded by the Korea Government (MSIT) under Grant RS-2023-00254695 (Contribution Rate: 30%), in part by the Research and Development Program for Forest Science Technology under Project 2021390A00-2323-0105 (Contribution Rate: 30%) provided by Korea Forest Service (Korea Forestry Promotion Institute), and in part by Basic Science Research Program through the NRF funded by the Ministry of Education under Grant 2022R1F1A1063180 (Contribution Rate: 40%).

**ABSTRACT** In this paper, we propose an efficient sound propagation and remapping technique based on geometry to enhance immersive sound effects in virtual environments, and suggest a method to interactively represent sound by considering the altitude and slope of the terrain. The proposed technique can represent real-time patterns of sound such as waves and flows of sound, as well as diffraction that occur in a physical environment. Our approach involves identifying the position of obstacles from the sound source and calculating the new position of sound that has been refracted or diffracted due to the obstacles. Using a ray-tracing technique, we determine whether there is a collision between obstacles and sound, and calculate the reflected and refracted vectors due to the collision to predict the magnitude of sound that agents beyond the obstacles would hear. In this process, the number of reflected and refracted rays affects the magnitude of sound, allowing us to model sound attenuation based on the distance to the obstacle and the material of the obstacle. In this process, we control the magnitude of sound based on the altitude difference from the sound source and the gradient difference depending on the shape of the surrounding terrain. As a result, we efficiently and quickly represented diffraction patterns expressed in a physically-based approach, and demonstrated that the magnitude of sound diffuses naturally according to the deformed diffraction pattern as the obstacles move and rotate. Furthermore, we realistically represented the sound heard by the audience in terrains with altitude differences, and demonstrated that the magnitude of sound varies depending on the terrain.

**INDEX TERMS** Sound diffraction, sound flow, sound wave, synthesizing sound, virtual environment.

## I. INTRODUCTION

Recently, physically-based approaches that represent virtual environments like reality have been considered important, and there have been attempts to improve user immersion based on various sensory inputs [1], [2], [3], [7]. Among those attempts, the most prominent sense is vision. Similarly, various technologies are being developed to maximize the visual effects in games and VR (Virtual Reality) contents. Recently, there have been studies proposed to enhance the immersion of users, regardless of whether they use HMD (Head-mounted display) or not. But most of them have mainly focused on visual elements. Just focusing on visual elements in virtual environments has limitations in realistically representing sensations [4].

When visual cues are blocked or objects within the visual field cannot be perceived, hearing becomes the next most important sense. When a random sound occurs in darkness, people generally react to it and turn their head in the direction of the sound. In addition, even if the line of sight is obstructed by obstacles, sound can be transmitted to the audience behind the obstacles through diffraction effects [5], [6]. However, most VR contents do not take into account these auditory features, and especially do not reflect changes in

The associate editor coordinating the review of this manuscript and approving it for publication was Xiaogang Jin.

sound according to interaction with objects and the physical characteristics of sound that occur in reality. If the sound is consistently blocked, even with excellent visual elements, the user's immersion may be reduced and they may feel cognitive dissonance due to the unnaturalness of the auditory experience. In addition to visual elements, if auditory elements are well represented, immersion can be enhanced. Furthermore, by utilizing various auditory cues, new types of auditory-focused contents can be created for users with disabilities, providing them with additional information and enhancing their overall immersive experience. In addition, if auditory elements are well represented, it is possible to enhance immersion in VR environments by incorporating more complex interactions based on the patterns of sound diffraction and scattering across terrains.

Recently, there have been continuous efforts to apply physics-based sound synthesis techniques to animation or virtual reality [8]. Generally, audio used in VR is based on a stereo environment, so it cannot match the immersive sound quality of a multichannel audio system in terms of representing sound directionality or frequency expression. To alleviate this problem and provide realistic audio sound, research using binaural rendering techniques has also been proposed [9]. However, the linear convolution involved in binaural rendering techniques is computationally intensive, making it difficult to execute in real-time. Recently, Kim implemented binaural rendering in the frequency domain, performing separate binaural rendering in both low and high frequency ranges [10]. In most stereo headphone-based VR services that utilize this approach, a constant sound is delivered regardless of changes in the user's position (azimuth, distance, etc.). As a result, these characteristics create another mismatch between sound and scene, leading to a decrease in immersion. To address this, SCC (Sound Scene Control) technique for binaural sound has been proposed to apply to changes in the user's azimuth, or a technique to control sound level for changes in the user's distance [11]. The aforementioned methods have mostly focused on stereo sound and have not represented the refraction or diffraction of sound or controlled the sound intensity through them. Ultimately, the inadequate representation of sound results in a decrease in immersion in virtual environments, and in most cases, techniques aimed at improving it require complex data structures and resources, making it difficult to use them in interactive environments.

### A. PROBLEM STATEMENT

In this section, we will review the representative approaches for generating sound in virtual environments and discuss the differentiation from existing techniques in terms of constructing effective immersive sound. Furthermore, we will discuss the problem that this paper aims to address.

#### 1) WAVE-BASED TECHNIQUES

These techniques allow for accurate sound propagation as they are based on the fundamental physics of sound waves. There is a study that has improved computational efficiency by pre-calculating the propagation pattern of sound waves [12]. There are also approaches to efficiently controlling sound propagation: There is a study that has calculated sound propagation using wave-based methods in small and medium-sized scenes or scenes where the sound source changes dynamically [13], a study that directly controls the propagation of sound in large-scale scenes [14], a study that represents sound propagation when the sound source changes dynamically or is added [15]. These approaches have the advantage of accurately representing various types of frequencies (reflection, diffraction, interference), but they require a very long computation time for high frequencies, making them unsuitable for real-time application in virtual environments.

#### 2) GEOMETRIC TECHNIQUES

These techniques represent sound propagation by solving sound waves in a way similar to ray-tracing. These techniques operate stably even at high frequencies and propagate sound using existing rendering technologies such as ray-tracing [16], [17], [18], image sources [19], [20], beam-tracing [21], [22], and photon-tracing [23], [24]. In addition, diffraction of sound is approximated by combining UTD (Uniform Theory of Diffraction) and ray-tracing. Furthermore, these approaches have been expanded and developed using techniques such as Frustum-tracing [25], [26], [27], Acoustic radiance transfer [28], [29], Conservative frusta tracing [30], Diffuse reflections/higher order diffraction [31], [32], Diffraction kernel and mobile devices [33], [34]. This approach accurately represents the propagation of high-frequency sound waves, but as the reflected energy diffuses, the propagation of sound waves is approximated in low frequencies, resulting in accumulated errors. Compared to the wave-based approach, this approach requires less preprocessing time, but the computation cost increases rapidly as the number of triangles and reflection depth increases.

#### 3) HYBRID TECHNIQUES

This approach attempts to utilize the advantages of both wave-based techniques (efficient for low frequencies but computationally expensive for high frequencies) and geometric-based techniques (inaccurate for low frequencies but accurate for high frequencies), and is relatively accurate and efficient for all frequencies [35], [36], [37], [38], [39]. This approach allows for mostly accurate calculations of sound reflection, diffraction, and interference, and the preprocessing time is not lengthy.

#### 4) TIME COMPLEXITY OF RAYTRACING

To approximate sound diffraction and reflection based on ray tracing, at least three types of rays (primary ray, diffraction ray, reflection ray, etc.) must be generated in large numbers [40]. To realistically convey sound diffraction to the audience behind obstacles, the calculation process of

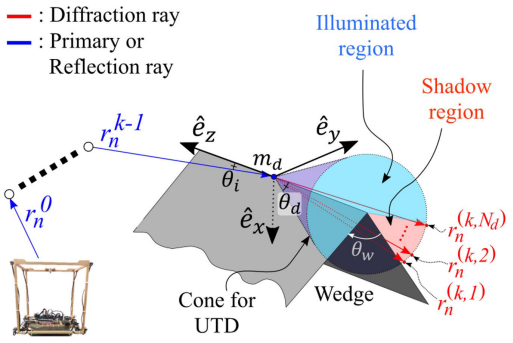



**FIGURE 1.** Representation of diffraction and reflection by tracing multiple rays [40].

repetitive reflection and refraction that cause changes in the sound must be added. This process has a significant impact on simulating sound diffraction, and the number of iterations can greatly increase depending on the complexity of the scene (see Figure 1).

In this paper, we propose a method of efficiently generating results that closely approximate the actual sound diffraction effect that occurs due to reflections and refractions in the surrounding environment, by calculating the diffraction positions of sound approximated through rays. This method can be efficiently utilized in real-time virtual environments without preprocessing because it simplifies the complex ray-tracing process. This method can be efficiently used in real-time virtual environments without the need for global illumination or preprocessing because it simplifies the complex ray-tracing process.

Existing approaches are difficult to use in real-time applications due to their high computational complexity, and they are limited in their applicability to certain scenes because they do not handle sound reflection, refraction, diffraction effects in complex terrains, or a variety of height-level sound sources. The technical contributions of this paper to overcome these limitations are as follows:

- Efficient calculation of diffraction patterns represented by obstacles using refracted rays
- Controlling the intensity of sound based on sound reflection and diffraction
- Conveying sound style to various sound sources
- Controlling the intensity of sound based on elevation differences

## II. RELATED WORK

AR (Augmented Reality) users are almost unrestricted in their movements such as walking, as they can see the real world as it is. However, as some of the information in the real environment is replaced by virtual information, users may feel some awkwardness in perceiving and making decisions about certain objects [11]. For users experiencing VR, since the real environment is not visible, to avoid collisions with actual obstacles, a bounding box is displayed in the actual movable range or the position of actual obstacles identified using an externally mounted camera is overlaid in the virtual environment [41]. However, some obstacles are perceived not only through visual cues but also through auditory cues, so there may be problems with obstacle recognition if the sound or auditory system is incomplete: changes in sound due to reflection and diffraction in invisible spaces.

Considering how users perceive and are affected by sounds in virtual environments is crucial to increasing their immersion in the virtual world. Sound not only conveys a sense of spatial depth to the user, but also enhances the perception of the entire space by allowing for more immersive interaction with objects within the virtual environment. In addition, playing synchronized sound effects according to the changes in the virtual environment can reduce symptoms such as VR motion sickness, compared to when the sound and the visuals are not synchronized.

Many researchers have studied the impact of auditory stimuli on VR experiences. Fujioka et al. conducted research on the integration of multimodal sensory information in material perception [42]. In their study, they conducted a visual and auditory experiment that combined the visual appearance of six materials with the impact sound of eight materials, demonstrating that visual and auditory stimuli are closely related to the perception of objects. Kern and Ellermeier conducted an experiment on how the presence of synchronized footstep sounds with the user's movements affects the user's VR experience, and showed that footstep sounds enhance immersion and realism [43]. According to these studies, in virtual environments where visual information dominates spatial perception, users can use auditory information as an auxiliary tool to more accurately perceive the virtual environment, and when visual information is ambiguous, auditory information becomes a dominant factor in this perceptual process.

Because sound has time-related properties, the mere existence of sound itself can provide important information that something is happening. The representative example of a system that uses sound to perceive the environment and avoid obstacles is the cane training experience for the visually impaired. Lahav and Mioduser presented a method for visually impaired individuals to acquire spatial mapping and orientation in a virtual environment [44]. Inman et al. performed spatial awareness training in VR for people with visual impairments, and showed that sound can help detect and navigate around obstacles [45]. In addition, it has been shown through experiments that listening to footstep sounds can help identify various terrain materials and predict movement speed. Siu et al. proposed a sound and vibration-based exploration system to enable visually impaired individuals to navigate complex virtual environments [46].

Vorlaender and Vorlaender proposed a method for calculating and simulating the propagation paths of sound, taking into account reflection, diffraction, and attenuation due to terrain and obstacles [48]. This study applied acoustic rendering techniques based on propagation paths to reproduce realistic sound environments. One of the advantages of this algorithm

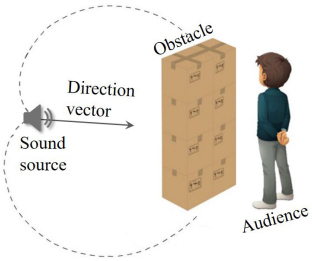



**FIGURE 2.** Scene configuration (dotted line: propagation area of sound).

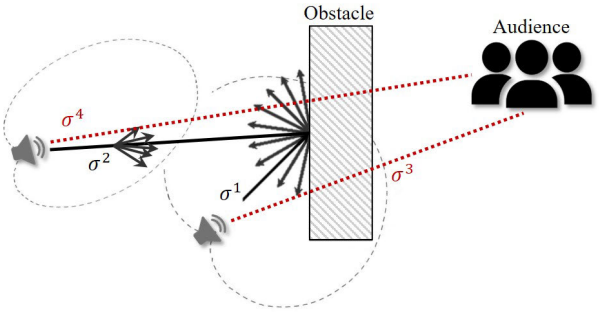



**FIGURE 3.** Rays generated in a hemispherical integration area, taking into account the relationship between the sound source, obstacles, and users.

is its ability to accurately model sound propagation in outdoor environments. Through this process, it is possible to simulate the propagation paths and characteristics of sound, as well as predict real-world outdoor sound environments in advance. This can be effectively applied in various fields such as urban planning, environmental noise assessment, and acoustical design, among others. However, one of the drawbacks of this method is its high computational cost. Outdoor environments consist of complex terrains and various obstacles, requiring extensive calculations to model sound propagation. As a result, it may be challenging to employ this method in interactive environments, necessitating the utilization of optimized algorithms and hardware to address this issue. Liu and Liu proposed an outdoor sound propagation technique based on Adaptive Finite-Difference Time Domain-Parabolic Equation (FDTD-PE) [49]. This algorithm possesses both accuracy and computational efficiency as its advantages. The adaptive FDTD handles modeling in the high-frequency range, while the PE is responsible for modeling in the low-frequency range. This allows for accurate calculation of sound propagation across various frequency ranges. This approach requires a high level of technical expertise due to the complex mathematical models and computational methods involved in adapting FDTD and PE. Furthermore, the accuracy of the modeling results can vary significantly depending on the input data. In this method, five materials that make up the virtual environment (tile, concrete, metal, wood, carpet) were selected, and the system was designed so that the user could receive sound and vibration by tapping a controller in the shape of a cane.

## III. PROPOSED FRAMEWORK

### A. REFLECTION AND REFRACTION OF SOUND DUE TO OBSTACLES

In this paper, the following assumptions are made to calculate sound reflection and refraction caused by obstacles: 1) The ground is flat, 2) Sound has anisotropy, 3) The amplitude of sound is constant. Based on this assumption, an obstacle and the user (listener) are arranged as shown in the Figure 2).

The purpose of designing such experimental setups in this paper is as follows: The reason for the first and third assumptions is to reduce computational complexity. In other words, assuming a constant pitch and a flat terrain allows us to efficiently control the sound volume by considering only the distance from the sound source and the sound's reflection/refraction/diffraction. The calculated sound map allows us to efficiently control sound in complex terrains and extend the solver to accommodate audio sources with varying pitches. The second assumption is made to efficiently compute sound propagation. Physically interpreted sound is represented as an isotropic energy potential field, and to calculate its directional properties, it requires numerous ray-tracing and complex computational processes. In this paper, to efficiently reduce the computational complexity of this process, sound propagation is assumed to have directional properties similar to energy. Indeed, in real-life scenarios, when measuring the sound level emitted from a mobile device, it is observed that the sound level differs between the front and back sides of the speaker, even when the positions are the same. Therefore, the assumption of isotropic sound in this paper is not a physically inaccurate approach. In other words, the mentioned assumptions aim to reduce computational complexity and consider the characteristics of terrain and various elevations of sound in solver extension, making it suitable for real-time applications such as games. However, when sound is measured and analyzed based on physical principles, our method may appear somewhat inaccurate.

To determine the presence of obstacles, we first shoot rays from the sound source towards a hemispherical area centered on the direction of the user, as shown in the Figure 2. In this paper, up to 20 rays are generated through random sampling with the distance from the sound source to obstacles and the user as weights (see Figure 3). In the figure, $\sigma$ represents the distance from the sound source to each entity, and we use this value to control the hemisphere-shaped ray generation area, which integrates energy. $\sigma^1$ and $\sigma^2$ represent the distances between the sound source and the obstacles. $\sigma^1$, which is closer, has a higher chance of being blocked, so the integration area for it is set as the entire hemisphere, while the integration area for $\sigma^2$, which is relatively further away, is set smaller. When the distance is far, the probability of being blocked is relatively low, so it is possible to calculate sufficient sound energy even without integrating over an entire hemisphere. If there is no obstacle, the integration

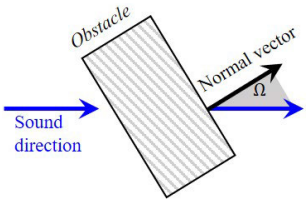
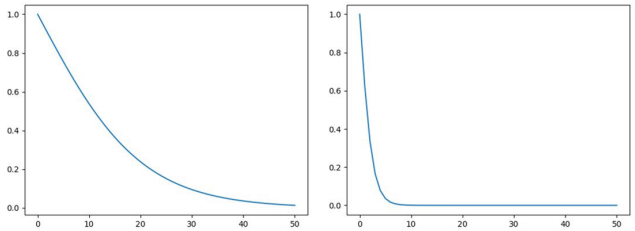



**FIGURE 4.** Calculating range(Ω) of refracted vector.



(a)$\alpha$ : 0.1 (b)$\alpha$ : 0.8

**FIGURE 5.** Sound intensity as $\alpha$ value changes.

area is controlled using $\sigma^3$ and $\sigma^4$, which are the distances between the sound source and the user.

Adjusting the integration area based on ray-tracing is equivalent to adjusting the propagation area of sound. In the case of $\sigma^2$, rays are generated in a relatively narrow area, so the propagation area is set to be small. On the other hand, $\sigma^1$ is the opposite, where the propagation area is set to be larger. We used this method to efficiently handle ray-tracing calculations.

Just like light, sound also undergoes reflection and refraction when it collides with objects. Therefore, we implemented the ray generated from the sound source to refract based on the collision point with obstacles. To represent this feature, the refracted direction after the collision with an obstacle follows Snell's law, and this type of ray represents the refraction of sound energy. Finally, the changes in sound propagation due to obstacles are taken into account to approximate the diffracted sound from refracted rays. The approximated diffracted sound based on the refracted sound from the original sound is used to generate the sound heard by the audience located behind the obstacles. In this paper, to more accurately consider the relationship between the slope of obstacles and the direction of sound, we generated the refracted sound vector between the normal vector of the obstacle surface and the direction vector of the sound (see Figure 4). Refraction vectors resulting from collisions with obstacles are transformed by refraction within the Ω region, creating another rays.

### B. CALCULATING DIFFRACTION OF SOUND CAUSED BY OBSTACLE

This paper utilizes rays, as explained earlier, to calculate sound diffraction. The approximated magnitude of the diffracted sound perceived by the listener was attenuated based on the number of collisions with obstacles relative to the original sound intensity. For example, depending on how much the obstacle blocks the sound source, the degree of sound attenuation should be different for cases where the obstacle is partially obstructing, completely obstructing, or not obstructing the sound source at all. In this paper, to reflect such phenomena, the approximated intensity of the diffracted sound, $\Gamma_{cp}$, is calculated as follows:

$$\Gamma_{cp} = 2\frac{1}{1+e^{\alpha h}} \tag{1}$$

where $h$ is the number of rays that collided with the obstacle, and $\alpha$ represents the degree of sound attenuation according to the material of the obstacle, with a value between 0 and 1. This value controls the degree of sound attenuation, and ultimately affects the sound intensity (see Figure 5).

The approximated position of the diffracted sound, Gamma, is calculated as follows (see Equation 2).

$$\Gamma^p_{cp} = \frac{\sum_{k=0}^{\|V_h\|} V_h^k + \left(\sum_{k=0}^{\|V_{nh}\|} V_{nh}^k w\right)}{n} \tag{2}$$

where $n$ represents the number of refracted vectors, $V_h$ and $V_{nh}$ represent the vectors that are refracted by colliding with the obstacle and those that are not, respectively. In addition, $w$ represents the degree of refraction according to the material of the obstacle. Through this, we can predict the refraction flow of sound that is generated beyond obstacles by calculating the diffracted sound, attenuation, and the refractive position according to the direction of the sound. After calculating $\Gamma_{cp}$ and $\Gamma^p_{cp}$, the user's position relative to the obstacle is determined with respect to the obstacle. If the user is located beyond the obstacle, only the diffracted sound is considered for calculating the sound level. If the user is not beyond the obstacle, only the original sound is also considered for the sound level calculation. While our method is not a fully physics-based approach, its computational process is highly efficient, allowing for real-time generation of results that are similar to actual sound diffraction patterns.

### C. PRACTICAL SOUND REMAPPING

Generally, style transfer technique refers to the application of the visual style appearing in one image to another image in the fields of computer vision, image processing, and the like. In this paper, we adopt the idea of style transfer to explain a method for easily remapping the style of one audio source to another audio source. The previously described sound intensity map was calculated solely through ray-tracing, and such a map shows a distribution of sound energy obtained through the rays. As mentioned earlier, we modified the ray-tracing technique commonly used in the rendering field (see Equation 3). In this paper, we controlled the integration area in a hemisphere shape to calculate sound refraction and diffraction (see $H^2$ in Equation 3). The sound energy varies depending on Ω value, which represents the

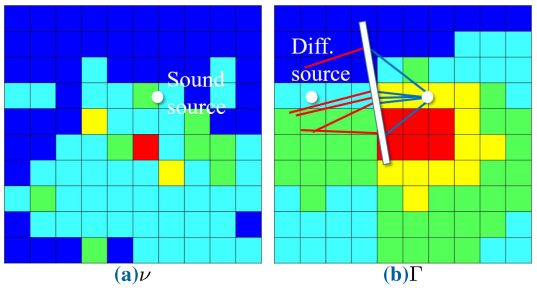



**FIGURE 6.** Sound maps with different characteristics (source file of $\nu$: Variations On The Canon).

material of the obstacle.

$$L(x,\omega) = L_e(x,\omega) + \int_{H^2} f_r\left(x,\omega' \rightarrow \omega\right) \times L\left(x^*\left(x,\omega'\right), -\omega'\right) cos\theta' d\omega' \tag{3}$$

In this paper, to efficiently represent the style of an audio source from the computed sound intensity map, we remap its style while preserving the sound energy using the following equation (see Equation 4).

$$\nu^t = \sum_i \nu_{0,i}, \tag{4a}$$

$$\Gamma^*_{i,cp} = \frac{\Gamma_{i,cp}}{max\left(\Gamma_{i,cp},\ldots,\Gamma_{n-1,cp}\right)}, \tag{4b}$$

$$\Gamma^\dagger = \sum_i \Gamma^*_{i,cp}, \tag{4c}$$

$$\Upsilon = \frac{\Gamma^*_{cp}}{\Gamma^\dagger}\nu^t, \tag{4d}$$

where $\nu_0$ represents the intensity level of the input sound, and it is used to maintain the total energy of the given input audio. This process is similar to the way mass and momentum conservation are handled in physics-based simulations. Therefore, even if the sound diffraction pattern changes depending on or obstacles or the distance from the sound source, the intensity of the sound is preserved during the simulation through $\nu^t$. Equation 4 adjusts the degree of diffraction based on the total energy of the input sound, and it is readjusted based on the distance between the sound source and the user or the refraction and diffraction caused by or obstacles. Therefore, the total energy of the input sound does not change, so the energy is conserved. This adjustment method can be applied in real-time applications as it is rescaled through the diffraction map $\Gamma$ calculated in the preprocessing stage.

As can be seen in the two maps of Figure 6, they represent sounds with different characteristics. In this paper, Equation 4 is used to control sound in order to efficiently represent these characteristics. The $\nu$ map is the result of calculating the sound intensity based on the Euclidean distance from the sampling position to the origin of $\nu$, without taking into account any obstacles.

**TABLE 1.** An example where the magnitude of the input sound is preserved.

| *Symbol* | $i \rightarrow 0$ | ... | ... | $i \rightarrow 3$ | *Sum* |
|---|---|---|---|---|---|
| $\nu_{0,i}$ | 4 | 2 | 1 | 3 | |
| $\nu^t$ | 10 | | | | 10 |
| $\Gamma_{i,cp}$ | 0.1 | 0.3 | 0.4 | 0.2 | |
| $\Gamma^*_{i,cp}$ | 0.25 | 0.75 | 1 | 0.5 | |
| $\Gamma_\dagger$ | 2.5 | | | | |
| $\Upsilon$ | 1 | 3 | 4 | 2 | 10 |

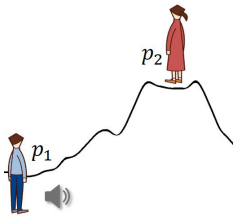



**FIGURE 7.** Elevation effect of terrain.

Table 1 is an example to provide a more intuitive understanding of Equation 4. The total sum of $\Upsilon$ confirms the calculated diffraction pattern of the sound while maintaining the total sum of $\nu^t$, which is the total intensity of the input sound. We can say that the overall shape of the sound has been transformed while the total energy is conserved, as the total sum of the energy, which is represented as the size of $\nu^t$, is maintained at 10.

## IV. SOLVER EXTENSIONS FOR CONSIDERING SHAPE AND ELEVATION OF TERRAINS

In this section, we will describe some extension cases of the sound propagation algorithm discussed so far. The method described in the previous section assumed a flat surface. In this section, we propose a synthesizing technique to efficiently extend our method to complex environments such as terrain. First, we approximate the elevation and slope of the terrain as 2D maps, and then we use these maps along with $\Gamma$ to composite sound data, thereby calculating the intensity of sound that takes into account the geometric features of the terrain.

### A. CALCULATING ELEVATION MAP

In this paper, sound is controlled by taking terrain into account using two weighted characteristics that affect sound propagation. The first one is the elevation of the terrain, and the second one is the slope of the terrain. The reason for considering the elevation of the terrain is that the sound heard changes depending on the difference in height. For example, comparing a listener $p_1$ at a height similar to the generated sound with a listener $p_2$ at a relatively higher position, the sound heard at $p_2$ is attenuated more due to the difference in elevation compared to $p_1$ (see Figure 7). To efficiently calculate this phenomenon, this paper proposes a method

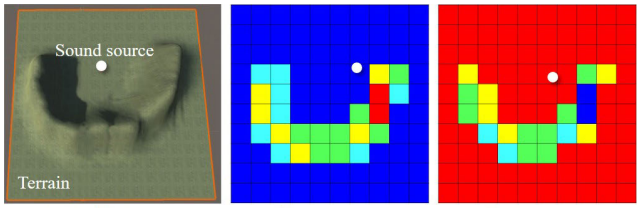



**FIGURE 8.** An elevation map that takes into account the location of the sound source.

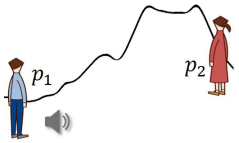



**FIGURE 9.** Slope effect of terrain.

of approximating 3D geometry through 2D maps, rather than directly manipulating the computationally intensive 3D geometry.

To implement an algorithm that takes into account the elevation of the terrain, it is necessary to determine the shape of the terrain and the location of the sound source. The elevation map is calculated based on the difference between the elevation value of the terrain, $e^{ter}_{i,j}$, and the height of the sound source, $s^p_y$: $\left\| e^{ter}_{i,j} - s^p_y \right\|$. As shown in the Figure 8, we can observe that the elevation map varies depending on the position of the sound. Figure 8b shows the case where the sound is located in a high area of the terrain, and thus, the higher regions are calculated with larger values while the lower regions far from the sound location are calculated with lower values. If the sound is moved to a lower location in the same scene, the opposite pattern can be observed (see Figure 8c). This experiment demonstrates the change of the elevation map according to the height in an intuitive manner.

### B. CALCULATING SLOPE MAP

In flat areas, sound propagates uniformly in all directions, but in terrains with irregularities, the terrain shape acts as an obstacle causing sound to diffract and sometimes slow down in propagation speed (see Figure 9). So we consider not only the elevation but also the slope of the terrain as a factor that affects the measurement of sound, and uses the gradient to apply it (see Equation 5).

$$\nabla e^{ter}_{i,j} = \frac{e^{ter}_{i+\frac{1}{2}h,j} - e^{ter}_{i-\frac{1}{2}h,j}}{h}, \frac{e^{ter}_{i,j+\frac{1}{2}h} - e^{ter}_{i,j-\frac{1}{2}h}}{h} \tag{5}$$

Figure 10 shows the result of the calculation of weights based on slope. In the case of a flat or peak terrain where there is almost no gradient, values close to 0 were obtained (see blue nodes in Figure 10b). Comparing region (A) and (B), we can see that even for sloped areas, the result can vary depending on the degree of the slope.

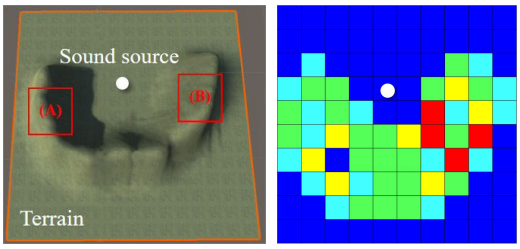



**FIGURE 10.** Visualization of slope map.

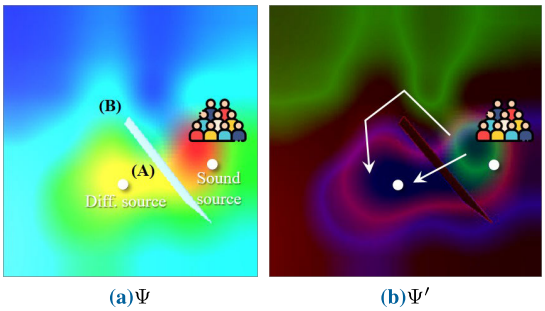



**FIGURE 11.** Sound intensity map generated by our method (people icon: audience).

### C. COMPOSITING MAPS

To summarize, in this study, the following three feature maps were used to represent the sound intensity considering the terrain: 1) $\Gamma^E$ is a feature map that calculates sound intensity on a flat plane, 2) $\Gamma^e$ is an elevation map, which represents the difference between the elevation of the terrain and the height of the sound source, 3) $\Gamma^s$ is a slope map, which represents the gradient of the terrain and its effect on sound propagation. The final sound intensity map, $\Psi$, which takes all of these factors into account, is calculated as follows (see Equation 6).

$$\Psi(x) = k_c \Gamma^E(x) \Gamma^e(x)(1 - \Gamma^s(x)) \tag{6}$$

where $x$ represents the sampling position, $\Gamma^{E,e,s}$ represent the sound intensities obtained from the three previously calculated maps, and these values are all normalized. To reflect the actual phenomenon that the sound attenuation increases in areas with steep slopes, the equation was set to be inversely proportional to $\Gamma^s$ as the slope becomes steeper. In this paper, to prevent the phenomenon of $\Psi$ becoming too small or too large, we added $k_c$ to control the sound intensity, and after several experiments, this value was set to 0.8.

## V. RESULTS

We conducted experiments on a computer equipped with an Intel i7-7700k 4.20GHz CPU, 32GB RAM, and an NVIDIA GeForce GTX 1080 Ti graphics card. We used Unity3D as our IDE.

To evaluate the effectiveness of our method, we need to verify whether the generated diffracted sound produces patterns similar to the sound flow beyond the obstacle. We assumed that there was an audience at each node of the

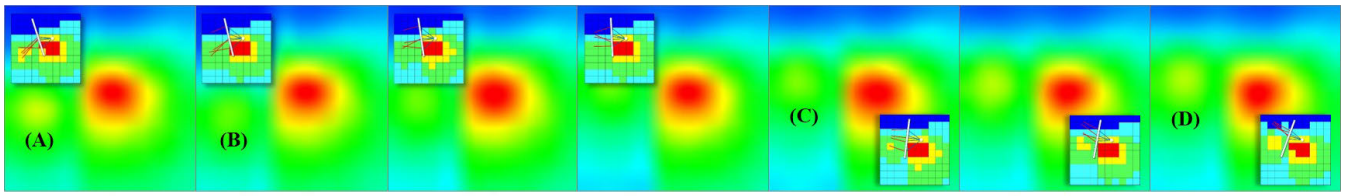



**FIGURE 12.** Sound intensity map that varies with the rotation of the obstacle (inset image: simulation scene).

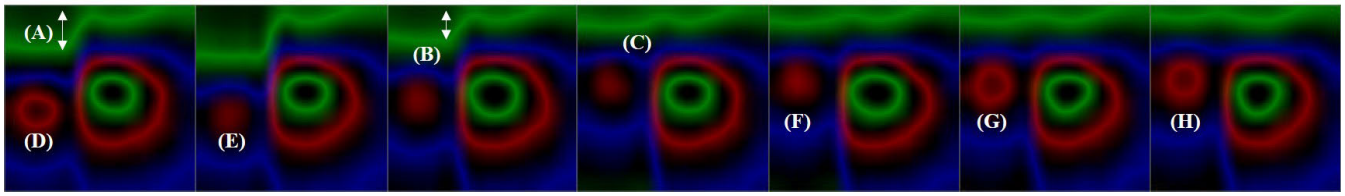
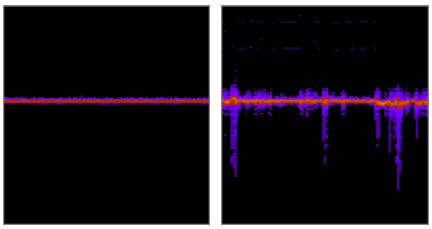



**FIGURE 13.** Filtered results of Figure 12.

grid, and calculated and compared the sound levels at all nodes. During this process, we compared the sound intensity calculated at each node taking into account how the sound changes depending on the distance to obstacles and how much the sound is blocked by obstacles. In simple terms, we confirmed through experiments that even when measured at the same location, the sound can be perceived differently depending on how much the sound source is obstructed by obstacles, with a softer sound when heavily obstructed and a louder sound when slightly obstructed. We also confirmed whether the sound was perceived in a large stereo format from a specific direction as the position of the diffracted sound changed depending on the slope of the obstacle. To intuitively visualize these analysis results, we used a map in the form of a rainbow color table.

Figure 11 shows the intensity map of the sound generated through the method proposed in this paper. The white wall representing the obstacle separates the sound source location and the approximated diffraction sound in the intensity map generated by our proposed method. As explained earlier, it can be seen that the sound intensity is also represented in the parts hidden by obstacles due to the refraction of rays generated from the source location (see (A) in Figure 11a). When constructing the scene, we discretized the grid space into low resolution, and when sampling the sound, we used the smoothed map to utilize the transformed sound intensity. Looking at the area near (B) in Figure 11a, the diffraction effect is more clearly represented. In this area, the convex protrusion that matches the shape of the obstacle is represented in the potential field in the part where it is blocked by the obstacle, which means that the sound is also transmitted to the audience beyond the obstacle. Figure 11b is the result of filtering (a) to obtain a clearer contour field of sound intensity.

Despite using a computationally reduced ray-tracing method to calculate sound refraction and diffraction, our



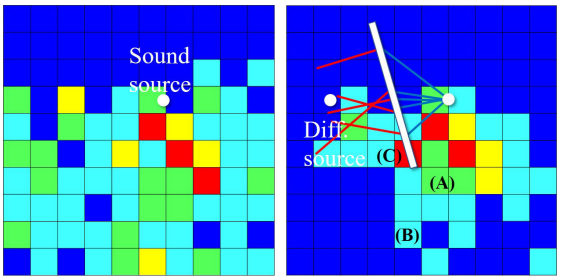
(a)Measuring pos. : (B) in Figure 11 (b)Measuring pos. : (A) in Figure 11

**FIGURE 14.** The sound spectrum visualized by our method for $\Psi$ (spectrum range: 5.5KHz∼70Hz).

method successfully approximates diffraction patterns that are physically-based. In the traditional method, reflecting and refracting rays are used to simulate diffraction, but this requires a complex recursive tracing process using many rays. However, in this paper, we efficiently calculated the sound intensity beyond obstacles by approximating the diffraction location based on refracted rays using $\Gamma^{p}_{cp}$, thus excluding the need for a complex recursive process. From the user's perspective, obtaining controlled sound intensity is made easier by sampling pre-calculated $\Psi$ based on the current position. Our method has a low computational cost in the pre-processing stage and can produce all results in real-time.

Figure 14 shows the analysis of the sound spectrum heard in Figure 11-(A,B). As shown in Figure 11a, sound was transmitted through refraction in the (A) area, while in the (B) area represented by diffraction, relatively smaller sound was clearly visible through the analyzed spectrum (The slight noise included in the results is caused by the sound conversion process by the CODEC).

Figure 12 shows how the shape of diffraction changes and how the flow of sound is altered as obstacles rotate. Sound diffraction varies depending on the shape of obstacles and the direction of sound propagation, and our proposed




(a)$\nu$ : original sound (b)$\Upsilon$ (our method)

**FIGURE 15.** The result of applying our method to 'Canon Variations'.

method effectively represents these characteristics. Since the refracted rays are randomly sampled in the area $\Omega$ between the sound direction and the normal vector of the obstacle, the sound is relatively loud in the area across the obstacle in Figure 12a-(A), and the sound decreases as the obstacle rotates (see (B) in Figure 12b). The sound does not necessarily decrease as the obstacle rotates; in the case of a revolving structure like a revolving door, the sound gradually increases again (see from Figure 12e-(C) to Figure 12g-(D)). In reality, when rotating a mobile device, we can experience a difference in the loudness of sound heard at the same location, depending on the direction of the device's speaker. The results of this paper effectively capture these characteristics.

Figure 13 is a contour field representation of the sound propagation pattern in the scene depicted in Figure 12, which was created for a detailed analysis. The pattern of how the angle affects the sound level, which was observed in Figure 12-(A∼D), can be more clearly confirmed through the color-coded contour field (see from (D) to (H) in Figure 13). Furthermore, the rotation of obstacles affects the surrounding sound diffraction pattern, and the difference can be clearly seen (see from (A) to (C) in Figure 13). In the existing geometric approach, similar patterns required many rays and numerous recursive processes to be created. In contrast, the proposed method is more efficient as it directly calculates the sound intensity at the approximated diffraction position, $\Gamma^p_{cp}$.

Figure 15 shows the result of controlling the sound of 'Canon Variations' using the sound remapping technique proposed in this paper. Figure 15a shows the result of controlling the intensity of sound based on the distance between each node and the sound source, without considering the surrounding environment. The reason why the sound propagated to each node did not have a constant intensity is that we used sound sources with varying heights. This style of audio is difficult to control because the distance from the sound source and the pitch of the sound itself can affect the result. However, our method can efficiently control such sounds because it utilizes the precomputed $\Gamma$ according to refraction and diffraction. Figure 15b visualizes the final controlled sound intensity, and unlike Figure 15a, it shows the refraction and diffraction patterns of the sound. This method may not have the same level of accuracy as image-based style transfer techniques used in artificial intelligence or computer vision, but it allows for easy real-time control of sound for arbitrary audio sources.

Figure 19 shows the spectrum analysis of the sound heard at positions A, B, and C in Figure 15. In contrast to Figure 14, this result used actual audio sources with varying pitch, resulting in a sound intensity with diverse levels. The sound at position (A) is perceived to be almost the same intensity as the original sound, while at positions (B) and (C), the intensity of the sound is relatively lower or higher due to refraction and diffraction. This pattern matched the results shown in the intensity map of the sound, $\Upsilon$.

Figure 16 shows the result of sound synthesis based on the calculation of $\Gamma^e$ and $\Gamma^s$ of the terrain using a diffusion map of constant sound intensity, $\Gamma^E$. The sound source is the ground surface, so in $\Gamma^e$, higher weight values were calculated closer to the ground surface. The final intensity map of the sound, Psi, takes into account not only the distance to the sound source but also the terrain slope, allowing us to observe the effect of terrain shape on the sound propagation (see Figure 16e).

In Figure 17, the sound is generated in the flat area in the middle of the terrain. Therefore, the sound propagated evenly at a moderate height, and this characteristic was clearly distinguished from Figure 16 (see $\Gamma^e$ in Figures 16 and 17).

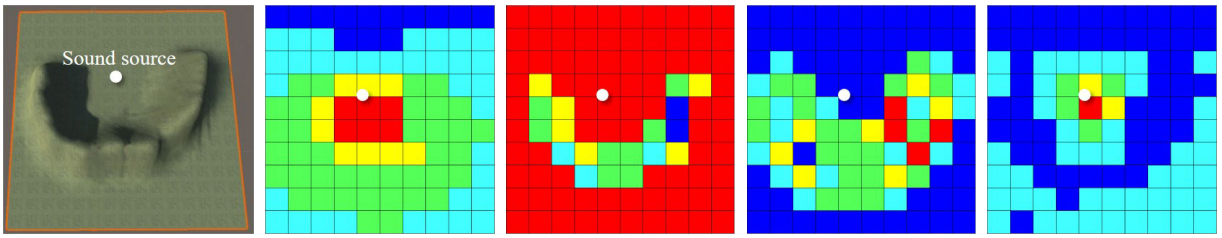
The experiment conducted in Figure 18 is different from the previous ones in that it was conducted on a terrain with relatively gentle undulations overall, and the sound source was created at an altitude somewhere between the ground and the peak. In $\Gamma^e$, which only considers elevation, a relatively large weight value was measured near the peak, but in $\Gamma^s$, which takes slope into account, a large weight value was measured in gentle areas, and a small weight value was observed near the peak with obstacles. In $\Psi$, which is composed of these feature maps, the sound was measured to be louder in the gentle terrain near the sound source, and relatively smaller sounds were measured in the vicinity of obstacles such as peaks (see Figure 18e). The approach proposed in this paper is not a perfect physically-based method, but it can quickly generate sound patterns that are similar to real phenomena, making it widely applicable in real-time or interactive applications. A list of symbols used in this paper can be found in Table 2

## VI. DISCUSSION

In this section, we provide a detailed explanation of the algorithm proposed in this paper, the impact of the computations on the resulting feature maps, and the experimental results analyzing the effects of sound source position on the outcome. In addition, we analyze the results of the interaction between sound and NPC using our method.

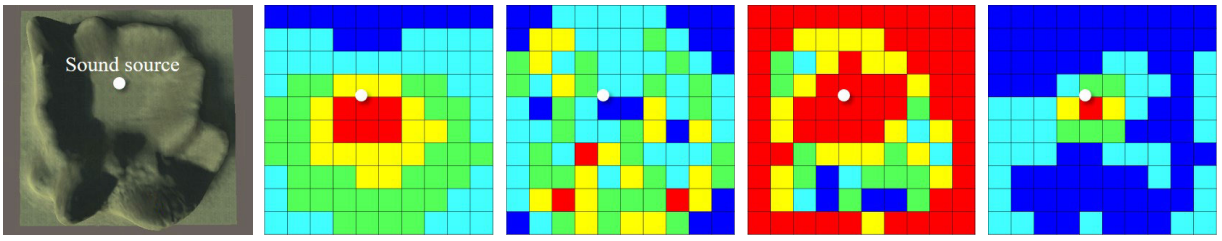
### A. DETAILS FOR COMPOSITING MAPS

We used an Equation 6 to composite three features. This equation is structured to scale the map of sound propagation calculated on a flat plane, which is the $\Gamma^E$, to match the



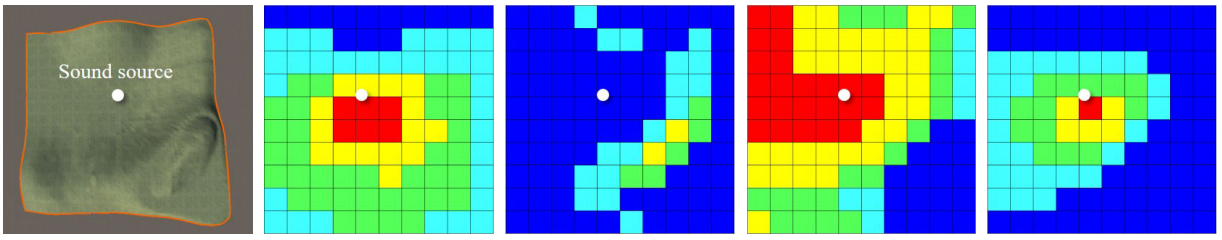

(a)Initial scene (b)$\Gamma^E$ (c)$\Gamma^e$ (d)$\Gamma^s$ (e)$\Psi$

**FIGURE 16.** Sound synthesis result 1 considering Terrain.



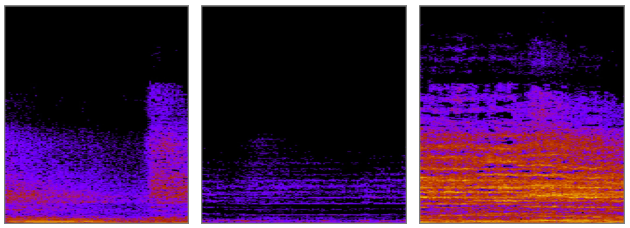

(a)Initial scene (b)$\Gamma^E$ (c)$\Gamma^e$ (d)$\Gamma^s$ (e)$\Psi$

**FIGURE 17.** Sound synthesis result 2 considering Terrain.




(a)Initial scene (b)$\Gamma^E$ (c)$\Gamma^e$ (d)$\Gamma^s$ (e)$\Psi$

**FIGURE 18.** Sound synthesis result 3 considering Terrain.



(a)Measured at (A) in Figure 15 (b)Measured at (B) in Figure 15 (c)Measured at (C) in Figure 15

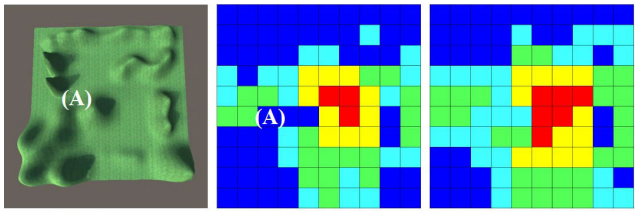
**FIGURE 19.** Visualization of sound spectrum on $\Upsilon$ with our method (*X*-axis: time, *Y*-axis: frequency, frequency range: 14.5KHz∼440Hz). The magnitude of the frequency is represented with color spectrum, from low (purple) to high (yellow).

**TABLE 2.** List of symbols used in the paper.

| Name | Description |
|---|---|
| $\Gamma_{cp}$ | Approximated diffraction sound |
| $h$ | Number of rays that collided with obstacles |
| $\alpha$ | Attenuation factor of sound based on obstacle material |
| $\Gamma^p_{cp}$ | Location of the approximated diffracted sound |
| $n$ | Number of refraction vectors |
| $V_h$ | Refraction vectors colliding with obstacles |
| $V_{nh}$ | Refraction vectors that do not collide with obstacles |
| $\Omega$ | Integrate area for calculating refraction and diffraction |
| $\nabla e^{ter}_{i,j}$ | Elevation's gradient |
| $\Gamma_E$ | Map of sound loudness calculated in flat-plane |
| $\Gamma_e$ | A map that takes into account the elevation of the terrain |
| $\Gamma_s$ | A slope map that takes into account the gradient of the terrain |
| $k_c$ | Coefficients that control the loudness of the sound |

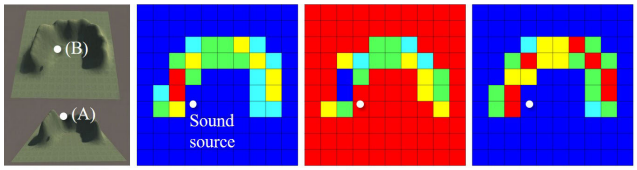
features of the terrain. If we denote the three feature maps of the terrain as $\Gamma^{tp}$, then the Equation can be expressed as follows: $k_c\Gamma^E\Gamma^{tp}$. In this process, we can simply scale the two maps using the subtraction operator, but this can lead to a failure to consider the terrain in the measured sound or the appearance of discontinuous parts (see Figure 20).

Figure 20 shows the experimental results of how $\Gamma^E$ and $\Gamma^{tp}$ affect the results. Figure 20b shows the results of applying the subtraction operator to the weight maps, and Figure 20c shows the results of applying the multiplication operator. When the subtraction operator is applied, as in Figure 20a-(A), even in areas where the slope is gentle, the sound is interrupted because the terrain features are not properly




(a) Initial scene (b) $\Gamma$ - $\Gamma^{tp}$ (c) $\Gamma\Gamma^{tp}$ (our method)

**FIGURE 20.** Effect of the relationship between $\Gamma^{E}$ and $\Gamma^{tp}$ on the results.




(a) Initial scene (b) $s_y^p$ : (A), highest peak (c) $s_y^p$ : (B), lowest peak (d) $s_y^p$ : The midpoint between (A) and (B)

**FIGURE 21.** Results of $\Gamma^{e}$ computed at various locations.

reflected. However, in this paper, this problem was easily resolved by using the multiplication operator, and as shown in Figure 20c, the terrain features are well represented.

In this paper, we employed parameter $w$ to control sound properties based on the material of the objects. However, we understand that the description of the specific characteristics being represented was not sufficiently clear. To address this, we have provided additional explanations to clarify the intended representation. Using sound-absorbing materials generally reduces sound reflection. In this paper, the parameter $w$ was used to represent the characteristics of solid materials, such as sound absorption. According to Snell's law, besides solid materials, there are environmental factors that affect the refraction of sound. In particular, at high temperatures, sound waves tend to refract downward, while at low temperatures, they refract upward. However, the parameter $w$ used in this paper does not capture these specific characteristics of sound refraction.

### B. EFFECT OF THE LOCATION OF THE SOUND ON THE RESULT

Of course, our method is highly dependent on the position of the sound. Especially when conducting experiments on terrains with significant differences in altitude, even if the (X, Z) coordinates are the same, the sound will propagate differently depending on the difference in altitude (Y coordinate). Figure 21 shows $\Gamma^{e}$ computed from various sound source locations. As mentioned earlier, the experiments were conducted by changing only the Y coordinate while keeping the same (X, Z) coordinates. (A) represents the highest location near the peak, while (B) represents the lowest location near the ground surface. As shown in the figure, the results vary significantly depending on the height difference of the sound.

Figure 21b visually demonstrates that sound is more prominent in areas close to the peak, while it is represented with a lower intensity in regions relatively far away or near the ground. Figure 21c shows the opposite pattern to Figure 21b. Figure shows the results of sound generation at a position roughly midway between (A) and (B). The figure effectively reflects the change in sound location, where the area that had a moderate sound intensity (yellow region) in Figure 21b is strongly expressed in red.

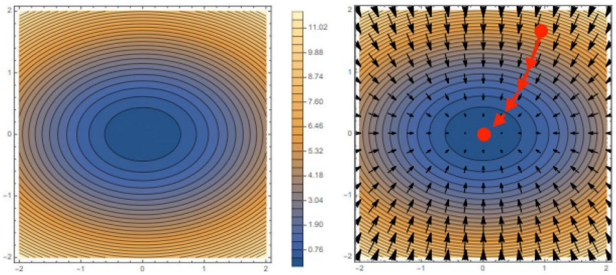
In geometric-based techniques for sound generation and propagation, it is common to construct the scene using a triangular mesh and estimate sound energy based on ray-tracing. The process of recursively computing the direction and number of rays for integrating sound energy is crucial, and it requires a significant amount of computational resources. Furthermore, the complexity of the terrain and the size of the sound source can also increase the computational requirements in this process. In contrast, the proposed method in this paper efficiently handles the aforementioned process through the following steps: 1) To reduce the computational requirements of the sound propagation calculation process, which increases with varying elevation and terrain complexity, the proposed method reduces the algorithm complexity by minimizing the dependencies on various characteristics when computing the sound map. In this process, sound propagation operates based on the elevation and slope of the terrain. The solver has been extended to efficiently remap the sound source to various elevations. This approach is more efficient in terms of computational requirements and memory usage compared to traditional techniques that handle sound propagation in 3D scenes. Furthermore, sound diffraction is typically represented by reflection and refraction, which require a large number of rays and recursive processes, resulting in increased computational requirements. However, our method minimizes the recursive process and reduces the computational workload by approximating the position for representing sound diffraction.

### C. CONTROLLING NPC MOVEMENT WITH SOUND

Based on the sound map shown earlier, we conducted an experiment to control the NPC to find the source of sound. The goal of this experiment was to confirm whether sound propagation was working properly by controlling the movement based solely on changes in sound intensity without any additional information. The sound field proposed in this paper has a similar form to the general potential field, and it can also be used for autonomous driving such as robot toys that navigate to the source of the sound.

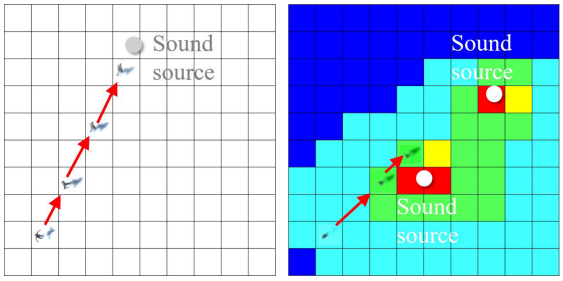
The movement of the NPC was implemented using the NavMesh feature provided by Unity3D, and the movement direction was determined based on the gradient of $\Psi$. Since $\nabla\Psi$ flows in the direction of the fastest increase at each point, in this paper, we controlled the movement of NPCs by using the direction of the fastest decrease, $-\nabla\Psi$ (see Figure 22).

Figure 23 shows the result of controlling the movement path of the NPC using $\Psi$ generated by the proposed method.



(a)Ψ (b)$-\nabla\Psi$

**FIGURE 22.** Gradient field of Ψ for controlling the NPC's path.




(a)One sound (b)Two sounds

**FIGURE 23.** The result of controlling the movement of an NPC using only the loudness information of the sound.

Figure 23a shows that the NPC is listening to and moving towards a single sound source. Since $-\nabla\Psi$ indicates the direction of the fastest decrease, the NPC moves closer to the shortest path. Figure 23b shows the result of an experiment with two sound sources, where the NPC was controlled to move towards the direction of the sound source with relatively high intensity from its current position. Since we are using a navigation mesh, even if there are obstacles, the NPC can move stably by specifying the path towards the sound source direction. Although the direction of $-\nabla\Psi$ is fixed, it can be easily integrated with navigation meshes that consider terrain and obstacles. Therefore, it can be easily applied to existing pathfinding approaches.

### D. REAL-WORLD DIFFERENCES AND LIMITATIONS

The proposed method aims to efficiently utilize sound effects in interactive applications based on virtual environments, such as games. Our method is important to note that directly comparing it to the representation of sound propagation in real physical spaces is challenging. The beam-tracing technique is one of the geometric-based ray-tracing methods [47], similar to our approach. Therefore, it is expected to represent a similar energy field to our method. However, unlike previous approaches that involve recursive tracing of numerous rays, our method efficiently represents diffraction by repositioning the sound using an approximated position. In conclusion, our method effectively demonstrates diffraction patterns that are generated based on the interaction between sound and obstacles. However, conducting a validation test that involves placing sound in real physical spaces, similar to what Thomas et al. did, may pose challenges [47]. Typically, conducting such experiments would require more than just a software-based comparison. It would involve additional devices such as sensors to be placed in the real physical space, which adds complexity to the experimental setup. However, we believe that it can be claimed that the ray-tracing approach for sound propagation is similar to the sound representation in real physical spaces.

## VII. COMPARISON AND USER EVALUATION

The differences between our method and previous approaches are as follows: 1) The calculation space for representing sound energy is 2D instead of 3D in our method. 2) Our method is based on a geometric approach known as ray-tracing, but instead of directly tracing numerous rays for reflection, refraction, and diffraction, we efficiently represent sound diffraction by calculating approximated positions. 3) Additionally, we have expanded the algorithm to enable the propagation of arbitrary sound by introducing a sound remapping method within the 2D sound map. 4) Lastly, we have proposed a terrain-based sound control technique based on the 2D sound map approach. In contrast to previous techniques that utilize actual 3D terrain models as input, our paper approximates the geometric features of the terrain (elevation and slope) into a 2D map format to control the volume of sound. As emphasized in 'Problem Statement', this study introduces a novel approach to depicting sound reflection/diffraction. It achieves a high level of computational efficiency compared to traditional 3D methods, while effectively representing sound propagation results in a manner akin to physics-based approaches. Furthermore, in this section, the proposed method's ability to depict natural sound propagation in virtual environments is evaluated through a survey analysis.

### A. USER STUDY DESIGN

Once a brief explanation of the gameplay mechanics is provided, participants proceed to conduct experiments on the following aspects: 1) How natural does the sound change in response to positional variations within the game/virtual environment? 2) How does the sound change based on the terrain variations? 3) Is the sound variation natural across different audio sources? Each gaming session consisted of experiments in three different scenes. The order of gameplay was randomized, and each game session was allotted a playtime of 5 minutes. For each game, we compared the sound effects embedded within the game engine (Unity3D) with the results of our method to investigate which one provides a more natural sound effect.

Participants were allowed to freely navigate through the experimental space where sound sources were placed on the terrain. In the first experiment, sounds were placed on a flat plane with no elevation changes, and participants observed changes in sound as NPCs moved freely. Before starting the game, participants were provided with explanations about the terrain structure and the form of sound sources.

**TABLE 3.** Questionnaire (Q1 Q6, Q9, Q10: Satisfied(5), Average(3), Not Satisfied(0), Q7, Q8: Easy(5), Normal(3), Difficult(0)).

| No. | Question |
|---|---|
| $Q1$ | - Is the sound effects represented through our method more realistic than the sound effects provided by the game engine (Unity3D)? Please provide your answer for Scene 1, Scene 2, and Scene 3, respectively. |
| $Q2$ | - Does the sound represented through our method enhance the immersion of the content compared to the default sound effects provided by the game engine? Please provide your answer for Scene 1, Scene 2, and Scene 3, respectively. |
| $Q3$ | - Is the volume variation of the sound that increases or decreases as the NPC moves on the plane natural? |
| $Q4$ | - Is the volume variation of the sound that increases or decreases based on the terrain's elevation changes natural? |
| $Q5$ | - Is the volume variation of the sound that increases or decreases based on distance changes natural when using fast-paced music as the source? |
| $Q6$ | - Is the volume variation of the sound that increases or decreases based on distance changes natural when using slow-paced music as the source? |
| $Q7$ | - Could you easily locate the position of the sound source using only sound when nothing was visible? |
| $Q8$ | - Compared to the sound effects provided by the game engine, was it easier to locate the position of the sound source using our method when nothing was visible? |
| $Q9$ | - Was the hint information about the sound more accurate with our method compared to the sound provided by the game engine when nothing was visible? |
| $Q10$ | - Were you able to perceive changes in sound volume due to sound diffraction? |

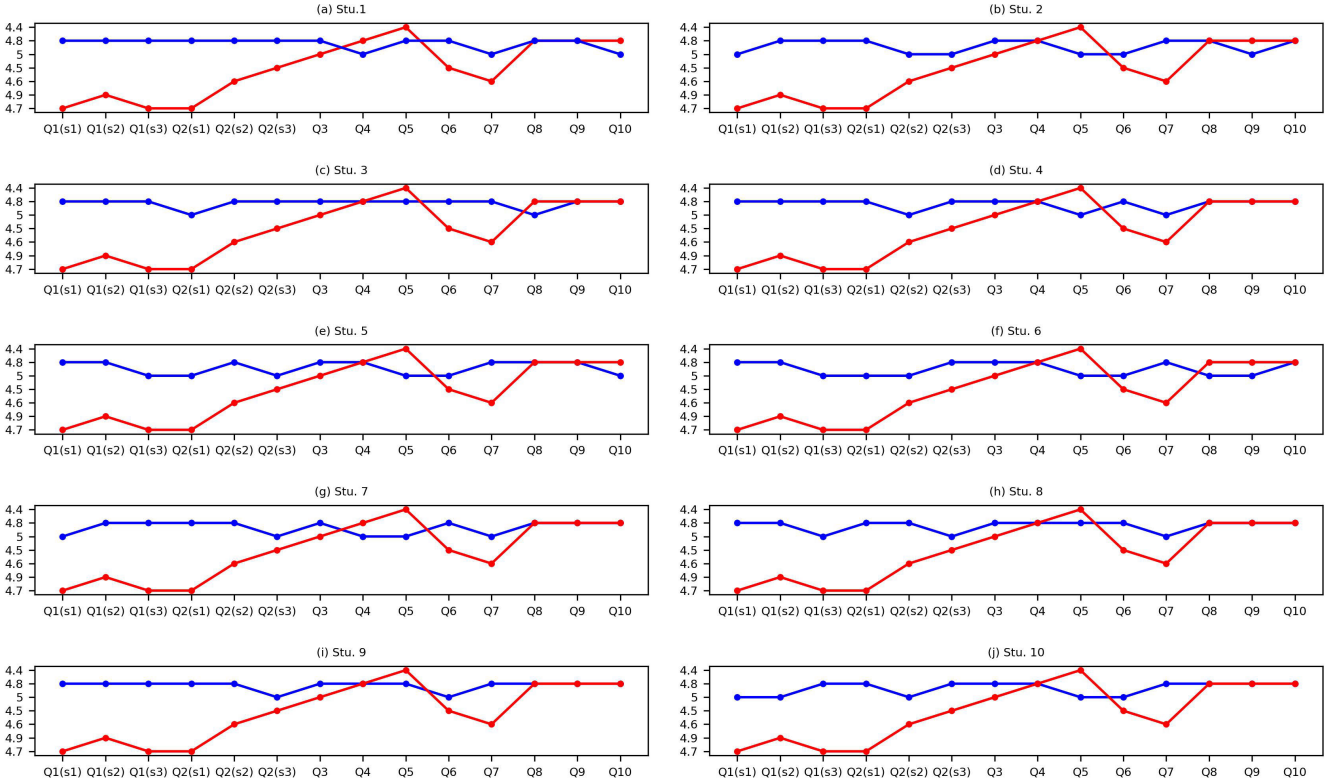



**FIGURE 24.** Survey participants (5 males (Stu. 1-5) and 5 females (Stu. 6-10)), (blue line: score for the question, red line: Avg. score for the question).

They were also given the freedom to choose NPC actions (attack, jump, run) as they wished. The game difficulty was not altered, and no specific time limits or game-ending events were implemented. In the second experiment, sounds were placed on a flat plane with elevation changes, and participants observed changes in sound as NPCs moved freely. The NPC was guided to move in accordance with the terrain's elevation, and its movement was constrained to allow participants to intuitively perceive changes in sound based on altitude variations. In the third experiment, we examined whether various sound volumes, along with obstacles and reflection/diffraction effects, were naturally represented. We examined whether the changes in sound volume were due to variations in the audio source itself, the effects of reflection/diffraction from obstacles, and whether these characteristics were naturally conveyed as the NPC moved.

### B. SUMMARIZATION

Participants were asked to play the game in each experimental scene both before (using sound effects embedded in Unity3D) and after applying our method. Subsequently, they were

requested to complete a questionnaire (see Table 3). Through this questionnaire, we aimed to investigate the impact of our method on realistic sound representation and the immersion of the content.

In addition to the naturalness and immersion of the sound, the survey in this paper also divided participants into male and female groups to gather opinions on the provided interface. Figure 24 presents the results of a survey conducted among the participants. The average score for all questions was 65.8 out of 70, indicating that most participants were satisfied with our interface. Examining the survey results, it is apparent that prior to the experiment, many participants had low expectations regarding the application of sound reflection/diffraction in the virtual environment, as they were unsure of how it would be implemented. Despite this, many participants expressed the opinion that the newly proposed sound effects in this study, rather than the sound effects embedded in the game engine, appeared to represent effects such as terrain elevation and diffraction caused by obstacles in a more natural and accurate manner. Furthermore, many participants also mentioned that our sound effects enhanced immersion and spatial perception, resulting in a more engaging experience. These positive evaluations are well reflected in Figure 24, and for Question 5, the average score was higher than individual scores, indicating a favorable trend. The reason for such results in Question 5 is likely due to the rapid variations in modulation of the fast-paced music used in that question. Participants may have found it challenging to distinguish whether these changes were a result of the proposed algorithm or not. Nevertheless, the study received favorable evaluations for most of the questions. The participants in the experiment were composed of 5 males (Stu. 1-5) and 5 females (Stu. 6-10). The survey results did not show significant differences or characteristics based on gender, and the common opinion among participants was that natural sound and immersion were generally improved.

## VIII. CONCLUSION AND FUTURE WORK

This paper proposes an efficient method for simulating sound diffraction and refraction based on a geometric approach, as well as a method for remapping sound to different forms. We also demonstrate the efficient control of sound using feature maps based on the height and slope of the terrain. In addition, to reduce the computational cost of the recursive process required in the ray-tracing process, we approximated the diffraction position, $\Gamma_{cp}^{p}$. As a result, we were able to generate diffraction and refraction patterns in sound propagation that are similar to those in the real world, and we obtained stable sound control results in various scenarios.

Due to the representation of the 3D terrain as a height map-based form, there can be instances where the same location has duplicate height values. Furthermore, in order to apply it to a typical triangular mesh, algorithmic improvements are necessary, and the current method has limitations in its applicability. This issue can be partially alleviated by employing a view-dependent approach, where the triangular mesh is projected onto the space for calculating the sound map. However, since it does not involve 3D geometry analysis, there may be a decrease in accuracy.

However, there are some limitations to our research: 1) The feature map is 2D, so it cannot perform accurate 3D geometric analysis. 2) Smoothing the entire area of the sound field can increase the computational cost. To address these issues in the future, we plan to consider the curvature and saliency of the terrain, and develop techniques to adaptively refine the smoothing area. Additionally, we plan to consider environmental factors that affect sound propagation, such as wind and temperature. In conclusion, this paper controlled the movement of the NPC in a 2D space, but additional methods will be necessary to control the NPC in a 3D space. To address this issue, we plan to design an efficient approach to remap the 2D feature map into a 3D space.

## REFERENCES


[1] Y. Wang, U. Jayaram, S. Jayaram, and K. Lyons, "Physically based modeling in virtual assembly," in *Proc. 21st Comput. Inf. Eng. Conf.*, vol. 1, Sep. 2001, pp. 295–305.

[2] V. Angelov, E. Petkov, G. Shipkovenski, and T. Kalushkov, "Modern virtual reality headsets," in *Proc. Int. Congr. Hum.-Comput. Interact., Optim. Robotic Appl. (HORA)*, Jun. 2020, pp. 1–5.

[3] J. Lin, X. Guo, J. Shao, C. Jiang, Y. Zhu, and S.-C. Zhu, "A virtual reality platform for dynamic human-scene interaction," in *Proc. SIGGRAPH ASIA Virtual Reality meets Phys. Reality, Modeling Simulating Virtual Humans Environ.*, Nov. 2016, pp. 1–4.

[4] J. Gugenheimer, E. Stemasov, J. Frommel, and E. Rukzio, "ShareVR: Enabling co-located experiences for virtual reality between HMD and non-HMD users," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, May 2017, pp. 4021–4033.

[5] K. van den Doel, P. G. Kry, and D. K. Pai, "FoleyAutomatic: Physically-based sound effects for interactive simulation and animation," in *Proc. 28th Annu. Conf. Comput. Graph. Interact. Techn.*, Aug. 2001, pp. 537–544.

[6] G. De Poli and D. Rocchesso, "Physically based sound modelling," *Organised Sound*, vol. 3, no. 1, pp. 61–76, Apr. 1998.

[7] Q. Zhang, C.-H. Kim, and H. W. Byun, "Multi-finger-based arbitrary region-of-interest selection in virtual reality," *Int. J. Hum.–Comput. Interact.*, pp. 1–15, Aug. 2022.

[8] S. Dirk, *Physically Based Real-Time Auralization of Interactive Virtual Environments*, vol. 11. Berlin, Germany: Logos Verlag Berlin GmbH, 2011.

[9] J. Breebaart, J. Herre, C. Jin, K. Kjörling, J. Koppens, J. Plogsties, and L. Villemoes, "Multi-channel goes mobile: MPEG Surround binaural rendering," in *Proc. Audio Eng. Soc. Conf., 29th Int. Conf., Audio Mobile Handheld Devices*, 2006, pp. 1–13.

[10] K. Kim, "A study on complexity reduction of binaural decoding in multi-channel audio coding for realistic audio service," *Contemp. Eng. Sci.*, vol. 9, no. 1, pp. 11–19, 2016.

[11] K. Kim, "Sound scene control of multi-channel audio signals for realistic audio service in wired/wireless network," *Int. J. Multimedia Ubiquitous Eng.*, vol. 9, no. 2, pp. 253–258, Feb. 2014.

[12] D. L. James, J. Barbic, and D. K. Pai, "Precomputed acoustic transfer: Output-sensitive, accurate sound generation for geometrically complex vibration sources," *ACM Trans. Graph.*, vol. 25, no. 3, pp. 987–995, 2006.

[13] N. Raghuvanshi, J. Snyder, R. Mehra, M. Lin, and N. Govindaraju, "Precomputed wave simulation for real-time sound propagation of dynamic sources in complex scenes," *ACM Trans. Graph.*, vol. 29, no. 4, pp. 1–11, Jul. 2010.

[14] R. Mehra, N. Raghuvanshi, L. Antani, A. Chandak, S. Curtis, and D. Manocha, "Wave-based sound propagation in large open scenes using an equivalent source formulation," *ACM Trans. Graph.*, vol. 32, no. 2, pp. 1–13, Apr. 2013.

[15] R. Mehra, A. Rungta, A. Golas, M. Lin, and D. Manocha, "WAVE: Interactive wave-based sound propagation for virtual environments," *IEEE Trans. Vis. Comput. Graphics*, vol. 21, no. 4, pp. 434–442, Apr. 2015.
[16] A. Krokstad, S. Strom, and S. Sørsdal, "Calculating the acoustical room response by the use of a ray tracing technique," *J. Sound Vibrat.*, vol. 8, no. 1, pp. 118–125, 1968.
[17] D. A. Pontarelli, R. B. Wise, and O. H. Olson, "Mark sensing in optical computations," *J. Sci. Arts*, vol. 48, no. 7, pp. 502–503, 1958.
[18] K. H. Kuttruff, "Auralization of impulse responses modeled on the basis of ray-tracing results," *J. Audio Eng. Soc.*, vol. 41, no. 11, pp. 876–880, 1993.
[19] J. Borish, "Extension of the image model to arbitrary polyhedra," *J. Acoust. Soc. Amer.*, vol. 75, no. 6, pp. 1827–1836, Jun. 1984.
[20] B.-I. Dalenbäck, P. Svensson, and M. Kleiner, "Room acoustic prediction and auralization based on an extended image source model," *J. Acoust. Soc. Amer.*, vol. 92, no. 4, p. 2346, Oct. 1992.
[21] T. Funkhouser, I. Carlbom, G. Elko, G. Pingali, M. Sondhi, and J. West, "A beam tracing approach to acoustic modeling for interactive virtual environments," in *Proc. 25th Annu. Conf. Comput. Graph. Interact. Techn. (SIGGRAPH)*, 1998, pp. 21–32.
[22] T. Funkhouser, N. Tsingos, I. Carlbom, G. Elko, M. Sondhi, J. E. West, G. Pingali, P. Min, and A. Ngan, "A beam tracing method for interactive architectural acoustics," *J. Acoust. Soc. Amer.*, vol. 115, no. 2, pp. 739–756, Feb. 2004.
[23] B. Kapralos, M. Jenkin, and E. Milios, "Sonel mapping: Acoustic modeling utilizing an acoustic version of photon mapping," in *Proc. 2nd Int. Conf. Creating, Connecting Collaborating through Comput.*, 2004, pp. 1–6.
[24] M. Bertram, E. Deines, J. Mohring, J. Jegorovs, and H. Hagen, "Phonon tracing for auralization and visualization of sound," in *Proc. VIS IEEE Visualizat.*, Oct. 2005, p. 20.
[25] C. Lauterbach, A. Chandak, and D. Manocha, "Interactive sound rendering in complex and dynamic scenes using frustum tracing," *IEEE Trans. Vis. Comput. Graphics*, vol. 13, no. 6, pp. 1672–1679, Nov. 2007.
[26] A. Chandak, C. Lauterbach, M. Taylor, Z. Ren, and D. Manocha, "AD-frustum: Adaptive frustum tracing for interactive sound propagation," *IEEE Trans. Vis. Comput. Graphics*, vol. 14, no. 6, pp. 1707–1722, Nov. 2008.
[27] M. Taylor, A. Chandak, Q. Mo, C. Lauterbach, C. Schissler, and D. Manocha, "ISound: Interactive GPU-based sound auralization in dynamic scenes," Tech. Rep. TR10-006, 2010.
[28] S. Siltanen, T. Lokki, and L. Savioja, "Acoustic radiance transfer method for room acoustic modeling," in *Proc. Int. Congr. Acoust.*, 2017.
[29] L. Antani, A. Chandak, M. Taylor, and D. Manocha, "Direct-to-indirect acoustic radiance transfer," *IEEE Trans. Vis. Comput. Graphics*, vol. 18, no. 2, pp. 261–269, Feb. 2012.
[30] A. Chandak, L. Antani, M. Taylor, and D. Manocha, "FastV: From-point visibility culling on complex models," *Comput. Graph. Forum*, vol. 28, no. 4, pp. 1237–1246, Jun. 2009.
[31] C. Schissler, R. Mehra, and D. Manocha, "High-order diffraction and diffuse reflections for interactive sound propagation in large environments," *ACM Trans. Graph.*, vol. 33, no. 4, pp. 1–12, Jul. 2014.
[32] L. Savioja and U. P. Svensson, "Overview of geometrical room acoustic modeling techniques," *J. Acoust. Soc. Amer.*, vol. 138, no. 2, pp. 708–730, Aug. 2015.
[33] C. Schissler, *Efficient Interactive Sound Propagation in Dynamic Environments*. Chapel Hill, NC, USA: Univ. North Carolina at Chapel Hill, 2017.
[34] A. Rungta, C. Schissler, N. Rewkowski, R. Mehra, and D. Manocha, "Diffraction kernels for interactive sound propagation in dynamic environments," *IEEE Trans. Vis. Comput. Graphics*, vol. 24, no. 4, pp. 1613–1622, Apr. 2018.
[35] R. Kirby, "Modeling sound propagation in acoustic waveguides using a hybrid numerical method," *J. Acoust. Soc. Amer.*, vol. 124, no. 4, pp. 1930–1940, Oct. 2008.
[36] S. M. Pasareanu, R. A. Burdisso, and M. C. Remillieux, "A numerical hybrid model for outdoor sound propagation in complex urban environments," *J. Acoust. Soc. Amer.*, vol. 143, no. 3, pp. 218–224, 2018.
[37] J. Nordström and J. Gong, "A stable and efficient hybrid method for aeroacoustic sound generation and propagation," *Comp. Rendus Mécanique*, vol. 333, no. 9, pp. 713–718, Sep. 2005.
[38] K. Hamiche, S. L. Bras, G. Gabard, and H. Bériot, "Hybrid numerical model for acoustic propagation through sheared flows," *J. Sound Vibrat.*, vol. 463, Dec. 2019, Art. no. 114951.
[39] H. Yeh, R. Mehra, Z. Ren, L. Antani, D. Manocha, and M. Lin, "Wave-ray coupling for interactive sound propagation in large complex scenes," *ACM Trans. Graph.*, vol. 32, no. 6, pp. 1–11, Nov. 2013.
[40] I. An, D. Lee, J.-W. Choi, D. Manocha, and S.-E. Yoon, "Diffraction-aware sound localization for a non-line-of-sight source," in *Proc. Int. Conf. Robot. Autom. (ICRA)*, May 2019, pp. 4061–4067.
[41] D. Połap, K. Kęsik, K. Książek, and M. Woźniak, "Obstacle detection as a safety alert in augmented reality models by the use of deep learning techniques," *Sensors*, vol. 17, no. 12, p. 2803, 2017.
[42] W. Fujisaki, N. Goda, I. Motoyoshi, H. Komatsu, and S. Y. Nishida, "Audiovisual integration in the human perception of materials," *J. Vis.*, vol. 14, no. 4, pp. 12–22, 2014.
[43] A. C. Kern and W. Ellermeier, "Audio in VR: Effects of a soundscape and movement-triggered step sounds on presence," *Frontiers Robot. AI*, vol. 7, p. 20, Feb. 2020.
[44] O. Lahav and D. Mioduser, "Multisensory virtual environment for supporting blind persons' acquisition of spatial cognitive mapping—A case study," in *Proc. 4th Int. Conf. Disability, Virtual Reality Assoc. Tech.*, 2001, pp. 1046–1051.
[45] D. P. Inman, K. Loge, and A. Cram, *Teaching Orientation and Mobility Skills to Blind Children Using Computer Generated 3D Sound Environments*. Atlanta, GA, USA: Georgia Institute of Technology, 2000.
[46] A. F. Siu, M. Sinclair, R. Kovacs, E. Ofek, C. Holz, and E. Cutrell, "Virtual reality without vision: A haptic and auditory white cane to navigate complex virtual worlds," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, Apr. 2020, pp. 1–13.
[47] T. Funkhouser, N. Tsingos, I. Carlbom, G. Elko, M. Sondhi, and J. West, "Modeling sound reflection and diffraction in architectural environments with beam tracing," in *Forum Acusticum*, 2002, p. 8.
[48] M. Vorländer and M. Vorländer, "Simulation and auralization of outdoor sound propagation," in *Proc. Auralization, Fundam. Acoust., Modeling, Simulation, Algorithms Acoustic Virtual Reality*, 2020, pp. 225–234.
[49] S. Liu and J. Liu, "Outdoor sound propagation based on adaptive FDTD-PE," in *Proc. IEEE Conf. Virtual Reality 3D User Interface (VR)*, Mar. 2020, pp. 859–867.