SOURCE DIRECTIVITY AND SPATIAL AUDIO FOR INTERACTIVE WAVE-BASED SOUND PROPAGATION

Ravish Mehra   Lakulish Antani   Dinesh Manocha

UNC Chapel Hill

ABSTRACT

This paper presents an approach to model time-varying source directivity and HRTF-based spatial audio for wave-based sound propagation at interactive rates. The source directivity is expressed as a linear combination of elementary spherical harmonic sources. The propagated sound field due to each spherical harmonic source is precomputed and stored in an offline step. At runtime, the time-varying source directivity is decomposed into spherical harmonic coefficients. These coefficients are combined with precomputed spherical harmonic sound fields to generate propagated sound field at the listener position corresponding to the directional source. In order to compute spatial audio for a moving and rotating listener, an efficient plane-wave decomposition approach based on the derivatives of the sound field is presented. The source directivity and spatial audio approach have been integrated with the Half-Life 2 game engine and the Oculus Rift head-mounted display to enable realistic acoustic effects for virtual environments and games.

Figure 1: Directional source and spatial audio system integrated with Oculus Rift HMD and Half-Life 2 game engine.

1. INTRODUCTION

Sound is ubiquitous in the physical world and forms the basic medium of human communication (language) and of human artistic expression (music). It is a pressure wave produced by vibration of a surface and transmitted through a medium. Sound propagation describes the behavior of sound waves as they are emitted by the source, interact with the environment, and reach the listener. Mathematically, the process of sound propagation can be expressed as a second-order partial differential equation called the *acoustic wave equation* in time-domain and the *Helmholtz equation* in frequency-domain. Over the years, sound propagation has emerged as a powerful tool to enhance the realism of virtual environments and user's sense of presence and immersion in these environments. It augments the visual sense of the user increasing his/her situational awareness and localization abilities [1]. Studies in audio-visual cross-modal perception have shown that high quality sound rendering can increase the quality perception of visual rendering [2]. Sound propagation has also been used for noise modeling and acoustic prediction in the field of noise control, room acoustics, and urban planning.

Existing sound propagation techniques can be broadly classified into geometric and wave-based techniques. Geometric techniques assume rectilinear propagation (ray-like behavior) of sound waves, trace rays or beams from the source, and accumulate con-

tributions at the listener position. This rectilinear propagation assumption is valid at high frequencies of sound and therefore it is hard to model wave-effects such as diffraction, interference, that happen at low frequencies, using geometric techniques. Wave-based techniques numerically solve the acoustic wave equation and can accurately perform sound propagation at all frequencies. However, these techniques are computationally expensive and have high runtime memory requirements. The state-of-the-art in wave-based sound propagation is mostly limited to offline techniques. These include standard numerical solvers such as finite difference time domain (FDTD), boundary element method (BEM), and finite element method (FEM). Due to the high computational and memory requirements, these techniques are limited to small acoustic spaces. Recently, there has been some work in developing interactive wave-based techniques for free-space sound radiation [3], first order scattering from surfaces [4], and sound propagation in small-to-medium sized scenes [5, 6]. However, wave-based sound propagation for interactive applications still faces many challenges:

- **Directional sources**
  Most sound sources we encounter in the real world, for e.g., human voice, speakers, and musical instruments, are directional sources. Current interactive, wave-based techniques can only model source directivity during the preprocessing stage. As a result, the directivity gets baked into the final solution, and it is not possible to modify the directivity pattern at runtime, for e.g. for time-varying directivity such as a rotating siren or a person covering his/her mouth.
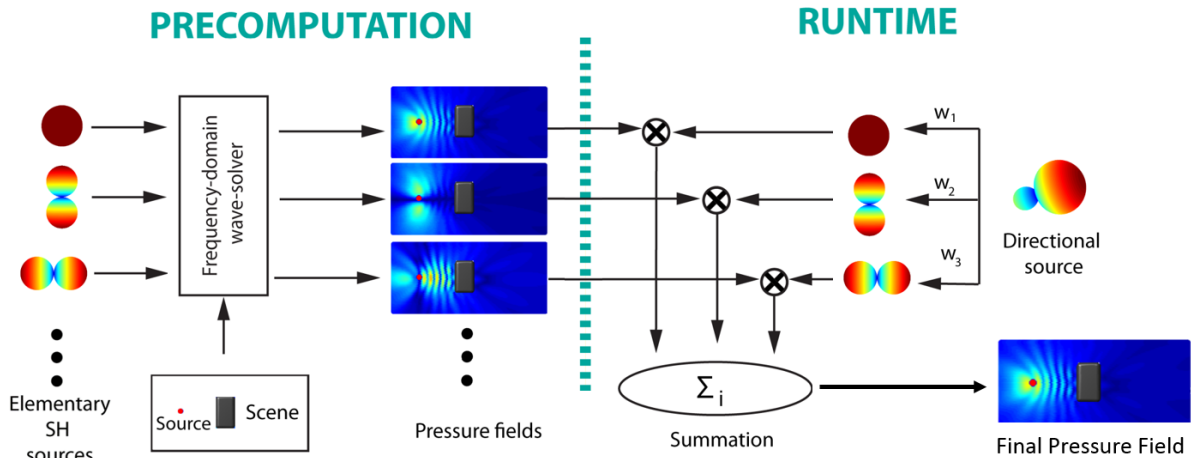
Figure 2: Overview of the source directivity approach.

- **Spatial audio**

  Accurate computation of spatial audio corresponding to the listener's motion and head rotation is a challenging task. Prior wave-based techniques for computing spatial audio are computationally expensive and applicable only for offline applications. Interactive techniques resort to coarse approximations based on simplified models which are not accurate for localization and externalization, both of which are necessary for immersion in virtual environments [7].

Source directivity and spatial audio have a significant impact on the sound generated by the auditory displays and 3D audio systems. Source directivity can change the acoustics of the environment [8]. This is noticeable in everyday life: positioning of musical instruments in an orchestra [9], good sound places (sweet spots) in front of the television in the living room. Spatial audio based on listener directivity also plays a key role in enhancing user's immersion in the auditory environment by providing binaural cues that correspond to the direction of the incoming sound [7]. Therefore, it is important to incorporate realistic sound effects produced by both source directivity and spatial audio into auditory displays and other 3D audio systems. These would help improve the sense of presence of users in the environment by providing them with improved spatialization, localization, and externalization cues, thereby enriching their experience.

In this work, we develop a set of efficient techniques for handing directional sources and computing spatial audio for wave-based sound propagation methods. The source directivity can be analytical, data-driven (i.e. based on directivity measurements of real world sources), rotating or time-varying function. Figure 2 gives an overview of our source directivity approach. Given a scene and a source position, a set of pressure (sound) fields due to elementary spherical harmonic (SH) sources is precomputed using a frequency-domain, wave-based sound propagation technique. Next, these pressure fields are encoded into basis functions (e.g. multipoles or equivalent sources) and stored for runtime use. Given the time-varying source directivity at runtime, a SH decomposition of the directivity is performed to compute the corresponding SH coefficients. The final pressure field at the listener position is computed as a summation of the pressure fields due to SH

sources weighted by the appropriate SH coefficients.

Our spatial audio framework supports listener's motion and head rotation, and allows the use of personalized HRTF. In Figure 3, we give a broad overview of our spatial audio approach. In order to compute spatial audio for wave-based techniques, an efficient plane-wave decomposition technique is proposed based on the derivatives of the pressure field. Acoustic responses for both the ears are computed at runtime by using this plane-wave decomposition and the HRTF-based listener directivity. These binaural acoustic responses are convolved with the (dry) audio to compute the spatial audio at the listener position. Our approach enables accurate wave-based sound propagation for directional sources and listeners in interactive applications.

2. PRIOR WORK

We discuss the related work in the field of sound propagation, directional sources, and spatial audio.

2.1. Acoustic wave equation

Sound propagation in the frequency-domain can be expressed as a boundary value problem using the Helmholtz equation:

$$\nabla^2 p + \frac{\omega^2}{c^2} p = 0 \ \ \text{in } \Omega, \tag{1}$$

where $p(\mathbf{x}, \omega)$ is the complex-valued pressure field at position $\mathbf{x}$ with angular frequency $\omega = 2\pi\nu$, c is the speed of sound, Ω is the propagation domain, and ∇^2 is the Laplacian operator. The behavior of p at infinity must be specified using the *Sommerfeld radiation condition* [10]. In order to complete the problem specification, either a Dirichlet, Neumann or mixed boundary condition is specified at the boundary of the domain.

2.2. Wave-based Techniques

Wave-based propagation techniques can be classified into *frequency-domain* and *time-domain* techniques. The most common among frequency-domain techniques include the finite element method (FEM) [11, 12] and the boundary element method
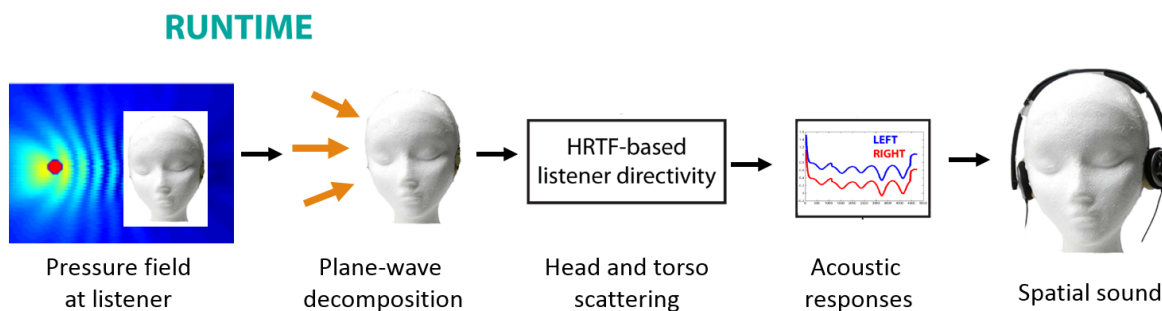
RUNTIME

Figure 3: Overview of our spatial audio approach.

(BEM) [13]. BEM expresses the global acoustic field as the sum of elementary radiating fields from monopole and dipole sources placed on a uniform, sub-wavelength sampling of the surface. Traditional BEM scales as the square of the surface area but recent research on the fast multipole method for BEM (FMM-BEM) [14] has improved the complexity to linear in surface area by creating a hierarchical clustering of BEM monopoles and dipoles, and approximating their interactions compactly using high-order multipole Green's functions. Offline FMM-BEM solutions are infeasible for interactive applications because of the large number of monopole and dipole sources in the final solution that need to be stored and summed at runtime.

In time-domain, the most popular technique is the finite difference time domain (FDTD) method [15], which requires a spatial discretization of the entire volume of the 3D scene. The computational and memory costs of these methods scale linearly with the volume of the scene. Faster methods like the pseudospectral time domain (PSTD) [16] and the adaptive rectangular decomposition (ARD) [17] have been proposed and achieve good accuracy with a much coarser discretization.

Interactive wave-based techniques In recent years, we have seen increasing interest in developing interactive wave-based techniques for sound propagation in indoor and outdoor spaces. Sound radiation from a single vibrating object in free space can be efficiently modeled using precomputed acoustic transfer functions [3]. Tsingos et al. [4] solve the boundary integral formulation of the Helmholtz equation subject to the Kirchhoff approximation in real-time. Raghuvanshi et al. [6] rely on a volumetric sampling of acoustic responses on a spatial grid and perceptual encoding based on the acoustic properties of indoor spaces. Recent work has shown that FDTD simulations can run in real-time on the GPU for small spaces that span a few meters across [5].

2.3. Geometric Techniques

The image source method [18] is the most commonly used geometric technique for modeling specular reflections, and there has been much research on improving its performance [19]. Other techniques based on ray tracing [20] or radiosity [21] have been developed for modeling diffuse reflections, but these energy-based formulations may not model phase accurately. Techniques based on acoustic radiance transfer [22] can model arbitrary surface interactions for wide-band signals, but cannot accurately model wave phenomena such as diffraction. The two main approaches for modeling diffraction in a geometric acoustics framework are the uniform theory of diffraction (UTD) [23] and the Biot-Tolstoy-

Medwin (BTM) formulation [24]. UTD is an approximate formulation, while the BTM is an offline technique that yields accurate results with a significant performance cost.

2.4. Directional sources

Real-world sound sources have a characteristic directivity that varies with frequency (see Figure 4). Meyer et al. [9] measure the directivity of brass, woodwind and string instruments in an anechoic chamber. This data contains magnitude information only, measured in terms of relative sound pressure levels as a function of directions in different octave bands. Otondo and Rindel [25] measured the tone-specific directivity of a trumpet, French horn, and clarinet. Recently, directivities of male and female singing voices have also been measured [26]. Source directivity has a significant effect on the propagation of sound in an environment [8]. Vigeant's PhD thesis [8] is an excellent reference for more details on the effect of source directivity on room acoustics. Interactive geometric acoustic (GA) techniques can incorporate high-frequency source directivities at runtime [27]. These methods essentially involve enumerating sound propagation paths from the source to the listener. The directions of rays emitted from the source (received at the listener) can be used to apply attenuation to the corresponding propagation paths based on the source directivity. However, the geometric techniques are not accurate in the low frequency range (e.g. 20-1000Hz), as sound waves can diffract (bend) around obstacles and undergo interference and other wave effects. Interactive wave-based propagation techniques [6, 28] can handle elementary directional sources such as monopoles, dipoles, quadrupoles, and their linear combinations. Other techniques have been proposed to incorporate measured directivities in wave-based techniques [29, 30]. However, these techniques are unable to handle sources with time-varying directivity at runtime.

2.5. Spatial Audio

The human auditory system obtains significant directional cues from the subtle differences in sound received by each ear, caused by the scattering of sound around the head [1, 7]. These effects are represented using the *head-related transfer function* (HRTF). HRTF measurements are performed on individuals in controlled environments and the recorded data is available online [31, 32, 33]. Interactive GA techniques can incorporate high-frequency HRTF-based listener directivity at runtime. However, handling listener directivity effects at low frequency arising due to the wave nature

of sound (e.g. diffraction, interference) remains a significant challenge. Integrating HRTFs into wave-based techniques involves computation of propagation directions using the *plane wave decomposition*. Prior plane-wave decomposition techniques either use spherical convolution [34, 35] or solve a linear system [36], and are computationally expensive. Interactive wave-based techniques resort to simpler but approximate listener directivity models based on a spherical head and a cardioid function [6].

3. DIRECTIONAL SOURCE

In this section, we give an overview of our approach for handling directional sources in a frequency-domain, wave-based sound propagation technique. We start by presenting a source formulation that incorporates the radial and directional radiation characteristics of sound sources. Next, we discuss a far-field representation of this source formulation that can be used to efficiently handle rotating or time-varying source directivities. Based on the linearity of the Helmholtz equation, we present a novel approach that incorporates our source representation into a general frequency-domain, wave-based propagation technique. All the variables used in this paper, except the SH basis functions, positions, and speed of sound, are frequency dependent. For the sake of brevity these dependencies are not mentioned explicitly.

3.1. Source Formulation

The radiation pattern of a directional source can be expressed using the one-point multipole expansion [37] as:

$$s(\mathbf{x}, \mathbf{y}) = \sum_{l=0}^{L-1} \sum_{m=-l}^{l} b_{lm} \ h_l^2(2\pi\nu r/c) \ Y_{lm}(\theta, \phi), \quad (2)$$

where $s(\mathbf{x}, \mathbf{y})$ is the (sound) pressure field at point $\mathbf{x}$ due to the directional source centered at point $\mathbf{y}$, $h_l^2(.)$ are the spherical Hankel functions of the second kind, $Y_{lm}(.)$ are the complex-valued spherical harmonic (SH) basis functions [38], (r, θ, ϕ) is the vector $(\mathbf{x} - \mathbf{y})$ expressed in spherical coordinates, b_{lm} are the weights of the expansion, L is the order of the expansion, c is the speed of sound in the medium ($343m/s$ for air at standard temperature and pressure), and ν is the frequency. This source formulation is valid in both near and far-fields[1].

3.2. Source Representation

Our choice of the source representation for directional sources is motivated by the measured directivity data that is currently available for real world sound sources. Most measurements have been collected by placing sources in an anechoic chamber and recording their directivity by rotating microphones every few degrees at a fixed distance from the source. Typically, these measurements are carried out at a distance of a few meters, which corresponds to the far-field for the frequencies emitted by these sources[2]. Keeping this in mind, we chose a source representation that corresponds to the far-field radiation pattern of a directional source. Under

---

[1]Far-field refers to the region of space where the distance d of any point in that region to the source is greater than the wavelength λ of the sound emitted by the source. The complementary region is the near-field [39, p. 165].

[2]For a distance $d > 3.43$m, corresponds to far-field for all the frequencies $\nu > 100$ Hz (for all wavelengths $\lambda = c/\nu < 3.43$m).

far-field approximation, $h_l^2(z) \approx i^{l+1}e^{-iz}/z$ where $i = \sqrt{-1}$ resulting in the following source representation [40]:

$$s(\mathbf{x}, \mathbf{y}) = \frac{e^{-i2\pi\nu r/c}}{r} \sum_{l=0}^{L-1} \sum_{m=-l}^{l} a_{lm} Y_{lm}(\theta, \phi), \quad (3)$$

$$= \frac{e^{-i2\pi\nu r/c}}{r} D(\theta, \phi), \quad (4)$$

where $a_{lm} = b_{lm} i^{l+1} c/(2\pi\nu)$ are the SH coefficients and $D(\theta, \phi) = \sum \sum a_{lm} Y_{lm}(\theta, \phi)$ is the directivity function at frequency ν. This directivity function can be either specified at each frequency analytically or measured/simulated at discrete sample directions. Depending on the data, the directivity function can be complex-valued (both magnitude and phase) or real-valued (magnitude only). Typically, the measured data is magnitude only and available as directivities averaged over octave-wide frequency bands [41].

We compute the coefficients a_{lm} of the source representation from the directivity function $D(\theta, \phi)$ as follows.

Analytical Given an analytical expression for the directivity function $D(\theta, \phi)$, we can compute the coefficients a_{lm} by SH projection:

$$a_{lm} = \int_0^{2\pi} \int_0^{\pi} D(\theta, \phi) \ Y_{lm}(\theta, \phi) \ \sin\theta \ d\theta \ d\phi. \quad (5)$$

This expression can be evaluated either symbolically or numerically, depending on $D(\theta, \phi)$ [42].

Data-driven Given the directivity function $D(\theta, \phi)$ at sampled locations $\{(\theta_1, \phi_1), (\theta_2, \phi_2), ..., (\theta_n, \phi_n)\}$, we can fit the SH expansion to this function in the least-square sense, by solving an over-determined linear system ($n > L^2$) to compute the coefficients a_{lm}:

$$\sum_l \sum_m Y_{lm}(\theta_k, \phi_k) \ a_{lm} = D(\theta_k, \phi_k) \quad \text{for } k = 1, ..., n. \quad (6)$$

Our source formulation and the corresponding SH representation can handle both complex-valued and real-valued directivities. If the directivity function is real-valued as in the case of widely available measurement data, real-valued SH basis functions [42] are used in the aforementioned expressions.

3.3. Frequency-Domain Sound Propagation

In this section, we describe our approach for incorporating the directional source representation in a general frequency domain, wave-based sound propagation technique. The steps outlined below are repeated for frequency samples in the range $[0, \nu_{\max}]$, where $\nu_{\max}$ is the maximum frequency simulated.

Directional sources The linearity of the Helmholtz equation implies that the pressure field of a linear combination of sources is a linear combination of their respective pressure fields [10]. Our source representation for directional sources is a linear combination of elementary spherical harmonic sources $s_{lm}(\mathbf{x}, \mathbf{y}) = \frac{e^{-i2\pi\nu r/c}}{r} Y_{lm}$ with different weights a_{lm} (equation 3). Therefore, for a given scene, if we compute the pressure field $p_{lm}(\mathbf{x})$ corresponding to each of the elementary sources $s_{lm}(\mathbf{x}, \mathbf{y})$, then the pressure field due to any arbitrary directional source $s(\mathbf{x}, \mathbf{y})$ can be expressed as the linear combination of the precomputed
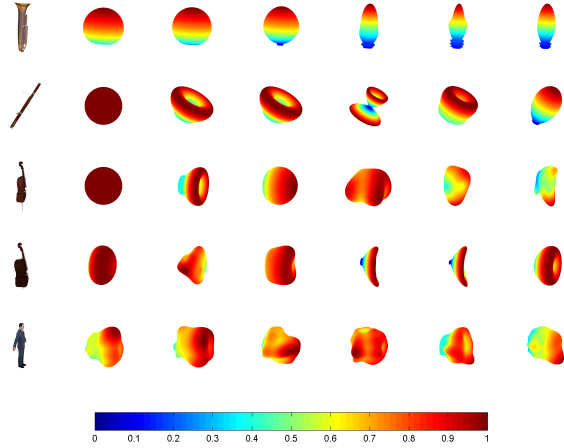
Figure 4: Source directivity of musical instruments and a human singer corresponding to different frequency octaves: $[0 - 176]$ Hz, $[177 - 354]$ Hz, $[355 - 710]$ Hz, $[710 - 1420]$ Hz, $[1420 - 2840]$ Hz and $[2840 - 5680]$ Hz. These directivity patterns are visualized by fitting our spherical harmonic-based source representation (error threshold 5%) to the real world measurement data provided by Meyers et al.[41].

pressure fields of the elementary sources with the same weights a_{lm}:

$$\underbrace{\sum_l \sum_m a_{lm} s_{lm}(\mathbf{x}, \mathbf{y})}_{s(\mathbf{x},\mathbf{y})} \longrightarrow \underbrace{\sum_l \sum_m a_{lm} p_{lm}(\mathbf{x})}_{p(\mathbf{x})}.$$

The pressure fields for elementary sources can be computed using any underlying numerical solver such as BEM or ESM [13, 28]. In the case of interactive applications, this computation is performed during the preprocessing stage, and the resulting pressure field data is efficiently encoded and stored as multipole weights (for BEM) or equivalent source weights (for ESM). This pressure field data completely defines the acoustic response to any directional source at the given position (up to the SH approximation order). At runtime, the specified source directivity $D(\theta, \phi)$ is decomposed into a spherical harmonic-based representation, using Equations (5) or (6), and the resulting weights a_{lm} are used to compute the final acoustic response $p(\mathbf{x_0})$ at listener position $\mathbf{x_0}$, as described above.

Time-varying and rotating directivity For a source with time-varying directivity, the spherical harmonic decomposition (6) of the source directivity function $D(\theta, \phi)$ is computed at runtime at interactive rates using fast linear solvers. For the special case of a rotating directional source, the new SH coefficients after rotation can be computed by applying the SH rotation matrix to the original SH coefficients [42].

3.4. Moving sources

Our system handles moving sources by precomputing the SH pressure fields at regularly sampled grid of 3D positions. The choice of sampling resolution is guided by runtime memory considerations. In practice, a sampling of 1-2 m is shown to generate good results

for auralization [43]. At runtime, as the source moves, interpolation is performed to compute the pressure field at the source position. The memory requirement of this approach scales with the size of the scene. However, perceptual compression techniques similar to Raghuvanshi et al. [43], can be used to compress the acoustics responses further.

4. SPATIAL AUDIO

We now describe an efficient method for computing spatial audio for a frequency-domain, wave-based sound propagation technique. First, we compute the plane-wave decomposition of the pressure field by expressing it into the SH basis and the use the pressure field derivatives at the listener position to compute the SH coefficients. Next, the head-related transfer function (HRTF) is also expressed in the SH basis. At runtime, spatial audio can be computed as a dot product between the SH coefficients of the plane wave decomposition and the HRTF.

4.1. Plane wave decomposition using derivatives

In the frequency domain, the global sound field can be expressed as a superposition of pressure due to plane waves [44]. This basis is also known as the *Herglotz wave basis*, with the basis functions centered at listener position $\mathbf{x}_0$:

$$\psi_{\mathbf{s}}(\mathbf{x}) = e^{ik\mathbf{s}\cdot(\mathbf{x}-\mathbf{x}_0)}, \tag{7}$$

where $\mathbf{s} = (s_x, s_y, s_z)$ is the unit vector in the direction of plane wave propagation, $k = 2\pi\nu/c$ is the wave number, ν is the frequency, and c is the speed of sound. In terms of these basis functions, the total pressure at $\mathbf{x}$ can be determined by integrating over all directions:

$$p(\mathbf{x}) = \frac{1}{4\pi} \int_S \psi_{\mathbf{s}}(\mathbf{x})\mu(\mathbf{s})d\mathbf{s}, \tag{8}$$

where $\mu(\mathbf{s})$, also known as the *signature function* [36], specifies the complex-valued amplitude of the plane wave traveling along direction $\mathbf{s}$, for a given frequency ν. The signature function can be further decomposed using the complex-valued SH basis functions, yielding:

$$p(\mathbf{x}) = \frac{1}{4\pi} \sum_l \sum_m \alpha_{lm} \int_S \psi_{\mathbf{s}}(\mathbf{x})Y^*_{lm}(\mathbf{s})d\mathbf{s}, \tag{9}$$

where $Y^*_{lm}(\mathbf{s})$ is the complex conjugate of the SH basis functions $Y_{lm}(\mathbf{s})$ and α_{lm} is the corresponding SH coefficient. The plane wave basis functions, shown in equation (7), have an important property: all of the basis functions have zero phase at the listener position $\mathbf{x}_0$. In other words, $\psi_{\mathbf{s}}(\mathbf{x}_0) = 1, \forall \mathbf{s}$. Based on this property and the orthonormality of the SH basis, it can be proven that the plane-wave decomposition of the pressure field can be computed up to any order by computing the derivatives of the pressure field up to the same order. The detailed derivation can be found in Mehra et al. [45].

Theorem : *Given the polynomial expression of the nth order and q^{th} degree SH*

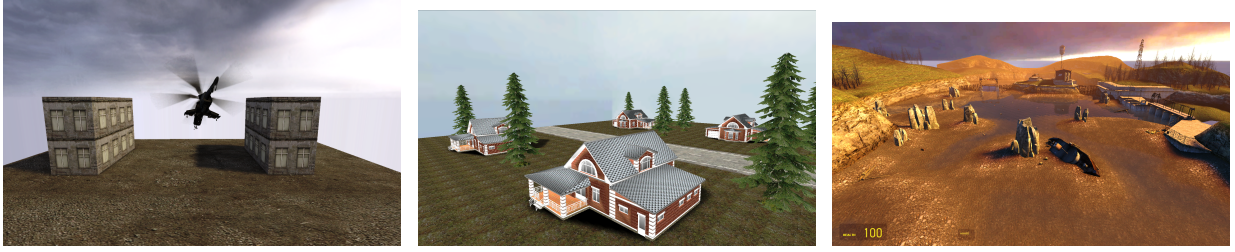$$Y_{nq}(s) = A \sum_{(a,b,c)} \Gamma_{a,b,c}\ s_x^a\, s_y^b\, s_z^c, \tag{10}$$

Figure 5: We demonstrate realistic acoustic effects produced by our technique for directional sources and listeners in the following scenes: (left) parallel walls, (center) urban street, and (right) reservoir.

where $a \geq 0, b \geq 0, c \geq 0$ and $a + b + c = n$, the corresponding SH coefficient in the plane-wave decomposition of the pressure field is

$$\alpha_{nq} = A \sum_{(a,b,c)} \Gamma_{a,b,c} \quad \frac{4\pi}{(ik)^{a+b+c}} \quad p^{(a,b,c)}(\mathbf{x}_0), \qquad (11)$$

where $p^{(a,b,c)} = \frac{\partial^{a+b+c} p}{\partial x^a \partial y^b \partial z^c}$. The above expression and equation 9 gives the plane-wave decomposition of sound field in terms of the pressure and its high-order derivatives at point $\mathbf{x}_0$.

4.2. Sound Rendering Using HRTFs

The HRTF relates the pressure received at each ear of the listener to the pressure at a point positioned at the center of the listener's head, and accounts for scattering from the head. We perform a SH decomposition of the left and right ear HRTFs, $H_L(\mathbf{s}) = \sum_{l'} \sum_{m'} \beta^L_{l'm'} Y_{l'm'}(\mathbf{s})$ and $H_R(\mathbf{s}) = \sum_{l'} \sum_{m'} \beta^R_{l'm'} Y_{l'm'}(\mathbf{s})$. Similar to Rafaely et al. [35], the pressure received at left and right ear (p_L and p_R) can be expressed as a dot product of the SH coefficients of the sound field and the HRTFs.

$$p_L = \frac{1}{4\pi} \sum_l \sum_m \beta^L_{lm} \alpha_{lm}, \quad p_R = \frac{1}{4\pi} \sum_l \sum_m \beta^R_{lm} \alpha_{lm}. \quad (12)$$

4.3. Accuracy

The derivatives of pressure field are computed by differentiating the pressure basis functions analytically rather than using a finite difference stencil. Therefore, higher-order derivatives do not suffer from numerical instabilities, allowing the SH coefficients of the sound field to be computed to any desired order. Rafaely et al [35] conducted a study on the effect of SH order on spatial perception. The results indicate that a SH order of 1-2 is sufficient for spatial perception of frequencies up to 1 kHz; a SH order of 3 suffices up to 2 kHz; and a SH order of 3-6 suffices up to 8 kHz. Higher SH order results in better spatial resolution, but computing the higher-order derivatives also increases the computational cost. Therefore, the SH order may be determined based on a performance-accuracy trade-off.

4.4. Interactive Applications

Our method allows spatial audio to be computed efficiently using two dot products (left and right ear) of complex-valued vectors. This enables the use of HRTF-based listener directivity for interactive wave-based sound propagation techniques as compared to

previous approaches which are offline. Our approach also provides the flexibility of using individualized (user-specific) HRTFs without recomputing the sound simulation and the plane wave decomposition. Only the SH coefficients of the individualized HRTFs need to be updated; the SH coefficients of the sound field remain the same. Additionally, our method enables head-tracking for a moving and rotating listener at interactive rates by simplifying the head rotation computation using SH rotation matrices.

5. IMPLEMENTATION AND RESULTS

Our source directivity and spatial audio framework has been integrated with the Valve Source Game Engine. We use the Oculus Rift head mounted display (HMD) DK-1 as the display device. It has a resolution of 640x800 per eye and 110 degree diagonal field of view. The Oculus HMD tracks the head orientation information which is used in the 3D game rendering as well as for generating spatial audio using our technique. It uses a combination of 3-axis gyros, accelerometers, and magnetometers to calculate the head orientation of the user. The position of the user in the game is controlled using the Xbox 360 wireless game controller. 3D spatial audio is delivered to the user using the Bayerdynamic DT 990 headphones. Figure 1 shows our system. The entire system is setup on a desktop machine with Intel Xeon E5-2687W CPU (3.1 GHz), 64 GB RAM, and NVIDIA Geforce GTX Titan GPU.

Figure 5 shows the scenes we used to demonstrate acoustics effects produced by our technique. Please refer to the supplementary video to listen to the auralization results on these scenes. In the *parallel walls* scene, we have a helicopter (directional source) flying between and around two parallel buildings. As the helicopter goes behind one of the buildings and out of sight of the listener, the sound gets low-passed due to diffraction. The second scene *urban street* consists of a radio playing behind one of the houses. The dominant acoustic effects in the scene are diffraction, reflection, and scattering of sound waves. In the *reservoir* scene, we have two directional source sources: helicopter and a rotating cello. In all the scenes, the spatial audio is played to the listener using the generic KEMAR HRTF-based listener directivity.

6. CONCLUSION AND FUTURE WORK

In this paper, we have proposed a novel approach for handling time-varying source directivity and HRTF-based listener directivity in frequency-domain, wave-based sound propagation technique. In comparison to geometric techniques, our approach can model wave-effects from low frequency directional sources and handle sound diffraction around the human head. In contrast to

prior interactive wave-based techniques [6, 28], our approach can handle rotating or time-varying source directivity at runtime. Also, we have proposed an efficient plane-wave decomposition approach to perform spatial audio rendering at interactive rates as compared to prior computationally expensive, offline approaches.

Our approach has few limitations. It would be expensive to handle directional sources with sharp directivity patterns using our current approach. This would require a very high-order SH expansion. In future, we would like to explore other basis functions (for e.g. wavelets) for handling sharp directivities. In the current implementation, we have used magnitude-only directivity data. But our approach can easily support complex data (magnitude and phase) as well, which we plan to test in the future. Our source formulation is general enough to handle both near-field and far-field sound radiation by directional sources. However, in order to model near-field directivity, a set of dense measurements of complex frequency responses very close to the source at twice the Nyquist rate is required. We hope that such a dataset becomes available in the future. We would also like to perform quantitative and qualitative comparison of our technique with other source directivity and spatial audio approaches used for wave-based techniques. Lastly, we would like to perform hybridization of the wave-based techniques with the geometric approaches to handle source directivity that spans the complete audible frequency range.

7. REFERENCES

[1] J. Blauert, *Spatial hearing: the psychophysics of human sound localization*. Mit Press, 1983. [Online]. Available: http://books.google.com/books?id=rrC5QgAACAAJ

[2] R. L. Storms and M. Usa, "Auditory-visual cross-modal perception," in *In ICAD*, 2000.

[3] D. L. James, J. Barbič, and D. K. Pai, "Precomputed acoustic transfer: output-sensitive, accurate sound generation for geometrically complex vibration sources," in *ACM SIGGRAPH 2006 Papers*, ser. SIGGRAPH '06. New York, NY, USA: ACM, 2006, pp. 987–995. [Online]. Available: http://doi.acm.org/10.1145/1179352.1141983

[4] N. Tsingos, C. Dachsbacher, S. Lefebvre, and M. Dellepiane, "Instant Sound Scattering," in *Rendering Techniques (Proceedings of the Eurographics Symposium on Rendering)*, 2007. [Online]. Available: http://www-sop.inria.fr/reves/Basilic/2007/TDLD07

[5] L. Savioja, "Real-Time 3D Finite-Difference Time-Domain Simulation of Mid-Frequency Room Acoustics," in *13th International Conference on Digital Audio Effects (DAFx-10)*, Sept. 2010.

[6] N. Raghuvanshi, J. Snyder, R. Mehra, M. C. Lin, and N. K. Govindaraju, "Precomputed Wave Simulation for Real-Time Sound Propagation of Dynamic Sources in Complex Scenes," *SIGGRAPH 2010*, vol. 29, no. 3, July 2010.

[7] D. R. Begault, *3D Sound for Virtual Reality and Multimedia*. Academic Press, 1994.

[8] M. C. Vigeant, "Investigations of incorporating source directivity into room acoustics computer models to improve auralizations," *ETD collection for University of Nebraska - Lincoln*, no. AAI3315882, 2008. [Online]. Available: http://digitalcommons.unl.edu/dissertations/AAI3315882

[9] J. Meyer and U. Hansen, *Acoustics and the Performance of Music (Fifth edition)*, ser. Lecture Notes in Mathematics. Springer, 2009. [Online]. Available: http://books.google.com/books?id=ruA3AQAAIAAJ

[10] A. D. Pierce, *Acoustics: An Introduction to Its Physical Principles and Applications*. Acoustical Society of America, 1989. [Online]. Available: http://www.worldcat.org/isbn/0883186128

[11] O. C. Zienkiewicz, R. L. Taylor, and P. Nithiarasu, *The finite element method for fluid dynamics*, 6th ed. Butterworth-Heinemann, Jan. 2006. [Online]. Available: http://www.worldcat.org/isbn/0750664312

[12] L. L. Thompson, "A review of finite-element methods for time-harmonic acoustics," *The Journal of the Acoustical Society of America*, vol. 119, no. 3, pp. 1315–1330, 2006. [Online]. Available: http://dx.doi.org/10.1121/1.2164987

[13] A. Cheng and D. Cheng, "Heritage and early history of the boundary element method," *Engineering Analysis with Boundary Elements*, vol. 29, no. 3, pp. 268–302, Mar. 2005. [Online]. Available: http://dx.doi.org/10.1016/j.enganabound.2004.12.001

[14] N. A. Gumerov and R. Duraiswami, "A broadband fast multipole accelerated boundary element method for the three dimensional Helmholtz equation," *The Journal of the Acoustical Society of America*, vol. 125, no. 1, pp. 191–205, 2009.

[15] S. Sakamoto, A. Ushiyama, and H. Nagatomo, "Numerical analysis of sound propagation in rooms using finite difference time domain method," *Journal of the Acoustical Society of America*, vol. 120, no. 5, p. 3008, 2006.

[16] Q. H. Liu, "The PSTD algorithm: A time-domain method combining the pseudospectral technique and perfectly matched layers," *The Journal of the Acoustical Society of America*, vol. 101, no. 5, p. 3182, 1997. [Online]. Available: http://dx.doi.org/10.1121/1.419176

[17] N. Raghuvanshi, R. Narain, and M. C. Lin, "Efficient and Accurate Sound Propagation Using Adaptive Rectangular Decomposition," *IEEE Transactions on Visualization and Computer Graphics*, vol. 15, no. 5, pp. 789–801, 2009. [Online]. Available: http://dx.doi.org/10.1109/TVCG.2009.28

[18] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," *The Journal of the Acoustical Society of America*, vol. 65, no. 4, pp. 943–950, 1979.

[19] 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 *ACM SIGGRAPH*, 1998, pp. 21–32.

[20] T. Lentz, D. Schröder, M. Vorländer, and I. Assenmacher, "Virtual reality system with integrated sound field simulation and reproduction," *EURASIP J. Appl. Signal Process.*, vol. 2007, no. 1, p. 187, 2007. [Online]. Available: http://dx.doi.org/10.1155/2007/70540

[21] N. Tsingos and J. D. Gascuel, "A general model for the simulation of room acoustics based on hierachical radiosity," in *ACM SIGGRAPH 97*, ser. SIGGRAPH '97. New York, NY, USA: ACM, 1997. [Online]. Available: http://dx.doi.org/10.1145/259081.259236

[22] S. Siltanen, T. Lokki, S. Kiminki, and L. Savioja, "The room acoustic rendering equation." *The Journal of the Acoustical Society of America*, vol. 122, no. 3, pp. 1624–1635, Sept. 2007. [Online]. Available: http://dx.doi.org/10.1121/1.2766781

[23] N. Tsingos, T. Funkhouser, A. Ngan, , and I. Carlbom, "Modeling acoustics in virtual environments using the uniform theory of diffraction," in *Computer Graphics (SIGGRAPH 2001)*, August 2001.

[24] U. P. Svensson, R. I. Fred, and J. Vanderkooy, "An analytic secondary source model of edge diffraction impulse responses ," *Acoustical Society of America Journal*, vol. 106, pp. 2331–2344, Nov. 1999.

[25] F. Otondo and J. H. Rindel, "The influence of the directivity of musical instruments in a room," *Acta Acustica*, vol. 90, no. 6, pp. 1178–1184, 2004. [Online]. Available: http://www.ingentaconnect.com/content/dav/aaua/2004/00000090/00000006/art00017

[26] H. Jers, "Directivity of singers," *The Journal of the Acoustical Society of America*, vol. 118, no. 3, pp. 2008–2008, 2005. [Online]. Available: http://link.aip.org/link/?JAS/118/2008/2

[27] T. Funkhouser, N. Tsingos, and J.-M. Jot, "Survey of methods for modeling sound propagation in interactive virtual environment systems," *Presence and Teleoperation*, 2003. [Online]. Available: http://www-sop.inria.fr/reves/Basilic/2003/FTJ03

[28] 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.*, Apr. 2013. [Online]. Available: http://doi.acm.org/http://dx.doi.org/10.1145/2451236.2451245

[29] H. Hacihabiboglu, B. Gunel, and A. Kondoz, "Time-domain simulation of directive sources in 3-d digital waveguide mesh-based acoustical models," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 16, no. 5, pp. 934–946, 2008.

[30] A. Southern and D. Murphy, "Low complexity directional sound sources for finite difference time domain room acoustic models," in *Audio Engineering Society Convention 126*, 5 2009. [Online]. Available: http://www.aes.org/e-lib/browse.cfm?elib=14960

[31] F. L. Wightman and D. J. Kistler, "Headphone simulation of free-field listening. i: Stimulus synthesis," *The Journal of the Acoustical Society of America*, vol. 85, p. 858, 1989.

[32] V. R. Algazi, R. O. Duda, D. M. Thompson, and C. Avendano, "The cipic hrtf database," in *Applications of Signal Processing to Audio and Acoustics, 2001 IEEE Workshop on the*. IEEE, 2001, pp. 99–102.

[33] B. Xie, X. Zhong, D. Rao, and Z. Liang, "Head-related transfer function database and its analyses," *Science in China Series G: Physics, Mechanics and Astronomy*, vol. 50, no. 3, pp. 267–280, 2007.

[34] M. Park and B. Rafaely, "Sound-field analysis by plane-wave decomposition using spherical microphone array," *The Journal of the Acoustical Society of America*, vol. 118, no. 5, pp. 3094–3103, 2005. [Online]. Available: http://link.aip.org/link/?JAS/118/3094/1

[35] B. Rafaely and A. Avni, "Interaural cross correlation in a sound field represented by spherical harmonics," *The Journal of the Acoustical Society of America*, vol. 127, no. 2, pp. 823–828, 2010. [Online]. Available: http://link.aip.org/link/?JAS/127/823/1

[36] D. N. Zotkin, R. Duraiswami, and N. A. Gumerov, "Plane-wave decomposition of acoustical scenes via spherical and cylindrical microphone arrays," *IEEE Trans. Audio, Speech and Language Processing*, vol. 18, 2010.

[37] M. Ochmann, "The source simulation technique for acoustic radiation problems," *Acustica*, vol. 81, pp. 512–527, 1995.

[38] E. W. Hobson, *The Theory of Spherical and Ellipsoidal Harmonics*. New York, NY, USA: Cambridge University Press, 1955.

[39] G. Kino, *Acoustic waves: devices, imaging, and analog signal processing*, ser. Prentice-Hall Signal Processing Series. Prentice Hall PTR, 1987. [Online]. Available: http://books.google.com/books?id=hcsYAQAAIAAJ

[40] M. Menzies, Dylan; Al-akaidi, "Ambisonic synthesis of complex sources," *J. Audio Eng. Soc*, vol. 55, no. 10, pp. 864–876, 2007. [Online]. Available: http://www.aes.org/e-lib/browse.cfm?elib=14175

[41] PTB, "http://www.ptb.de/cms/en/fachabteilungen/abt1/fb-16/ag-1630/room-acoustics/directivities.html (last viewed Oct. 23, 2012)," 1978. [Online]. Available: http://www.ptb.de/cms/en/fachabteilungen/abt1/fb-16/ag-1630/room-acoustics/directivities.html

[42] R. Green, "Spherical Harmonic Lighting: The Gritty Details," *Archives of the Game Developers Conference*, Mar. 2003. [Online]. Available: http://www.research.scea.com/gdc2003/spherical-harmonic-lighting.pdf

[43] N. Raghuvanshi, R. Mehra, D. Manocha, and M. C. Lin, "Adaptive rectangular decomposition: A spectral, domain-decomposition approach for fast wave solution on complex scenes," *The Journal of the Acoustical Society of America*, vol. 132, no. 3, pp. 1890–1890, 2012.

[44] R. Duraiswami, D. N. Zotkin, Z. Li, E. Grassi, N. A. Gumerov, and L. S. Davis, "High order spatial audio capture and binaural head-tracked playback over headphones with hrtf cues," in *119th AES Convention*, 2005.

[45] R. Mehra, L. Antani, S. Kim, and D. Manocha, "Source and listener directivity for interactive wave-based sound propagation," *IEEE Transactions on Visualization and Computer Graphics*, vol. 20, no. 4, pp. 495–503, Apr. 2014. [Online]. Available: http://dx.doi.org/10.1109/TVCG.2014.38