Interactive Sound Propagation in Virtual Environments using Geometric Acoustics and Numerical Wave Simulation with Signed Distance Functions

Srinivas Kaza (kaza@mit.edu)
Lukas Murmann (lmurmann) Fredo Durand (fredo@mit.edu)

***Abstract*—Geometrical Acoustics (GA) is a popular family of methods for simulating acoustics in virtual environments. Unfortunately, most GA techniques fail to take into account the effects of diffraction. As a result, these simulations tend to suddenly muffle occluded sources or receivers as they enter shadowed regions. There are several approaches to approximating the effects of diffraction within a geometrical acoustics framework, such as beam tracing or edge subdivision, but these techniques either create too many sampling rays or scale poorly with the surface area of the diffracted region. By representing the scene as a signed distance function, we propose a technique to model sound propagation involving a wave simulation near scene geometry in conjunction with sphere-tracing technique for other regions. Both simulation methods, the wave simulation and the sphere-tracing, run concurrently with a fixed timestep. We define a novel two-way coupling between the wave simulation and the geometrical acoustics, which extracts plane waves the periphery of the wave simulation and inserts spherical waves from incoming rays. Our system incorporates a form of spatial decomposition, in the form of a near-object envelope, as well as frequency decomposition, in the form of frequency binning, to improve its performance.**

I. Introduction

Engineering and artistic projects often require an efficient and realistic model of sound propagation. Simulating the acoustical properties of virtual environments can help architects and acoustic designers quickly evaluate the sound quality of spaces that they create. Game developers and 3D artists may also be interested in a more realistic sound rendering model to help improve immersion. In either case, the simulation must be able to run interactively.

This project details the design and implementation of such a system. In particular, our approach accurately simulates diffraction against arbitrary geometry, and handles materials with arbitrary acoustical properties. The former property is particularly important, because most real-time acoustical simulations currently do not handle “shadowed” regions behind occluding geometry.

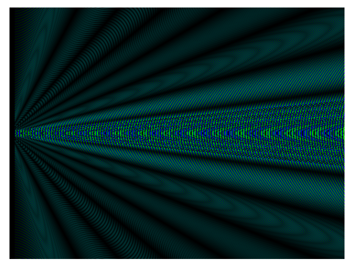

Fig. 1. Rendering diffraction in the single slit experiment.

Our approach uses a combination of geometrical acoustics and numerical wave simulation techniques to accurately model the means by which sound travels through a space. We believe that the strengths and weaknesses of both of these techniques complement each other well, and we introduce an implicit modeling scheme, the signed distance function (SDF), to determine which approach to use. The proposed method not only handles arbitrary geometry, but also *dynamic* geometry. Because this technique uses signed distance functions (SDFs) to perform most of our operations, changes to the scene geometry become visibile to the simulation after the SDF is regenerated.

II. Related Work

Most computational room acoustics techniques can be split into three groups – image source methods, ray-based methods, and wave simulations [1]. This paper focuses on the latter two of these three categories of acoustical rendering techniques. Ray-based techniques, within the family of geometrical acoustics (GA), have become popular recently as some graphical rendering techniques, such as bidirectional path tracing [2], have been adapted to model sound propagation. These techniques only model the energy histogram of the scene rather than directly recording the impulse response; as a result, phase information is not captured. Additionally, GA models usually do not attempt to capture the effects of diffraction. While simulations that use the Uniform Theory of Diffraction (UTD) or the Biot-Tolstoy-Medwin (BTM) diffraction model are capable of reproducing some of these effects within a GA framework [3], they are generally not robust to arbitrary scene geometry or scale poorly with the number of edges to be diffracted.

Numerical simulations which model the wave equation can also be used for sound propagation. Boundary element methods, including finite element methods (FEM) and finite difference time domain (FDTD) methods, are often used towards this end [1]. Like the ray-based GA techniques, these simulations work well for capturing both early and late reverberations. Unlike the previously mentioned geometric acoustics techniques, boundary element methods are usually not intended for interactive use. FDTD methods have the desirable property of operating in the time-domain, but typically do not run in real-time. In addition to these issues, determining directional information from an auralized wave simulation can be difficult. On the other hand, boundary element methods can model the wave-like properties of sound, such as wave phase and diffraction.

Within the field of computer graphics, signed distance functions have been used to implicitly represent scene geometry. A signed distance function represents the signed distance from a position in the scene to the nearest surface. A raymarching technique known as *sphere tracing* is used to determine intersections between rays and geometry reprsented by the SDFs. Hierarchical representations of SDFs have also been developed [4] to reduce the memory consumption and computational overhead involved in traversing a volumetric data structure. Our approach uses SDF functions to determine the near-object wave "envelope", as well as to raymarch sound rays for the geometrical acoustics.

The notion of evaluating the wave simulation in the near-object envelope and performing geometric acoustics elsewhere, a *spatial decomposition*, has been proposed before [5]. Alternatively some techniques perform a *frequency decomposition* of the scene, in which low frequencies and high frequencies are processed separately (e.g the high frequencies via geometric acoustics and the low-frequencies via a numerical simulation) [6]. Some modern techniques alternatively use a combination of *frequency decomposition* and *spatial decomposition*. Approaches that perform one-way and two-way couplings at the ray-wave barrier have also been demonstrated. However, none of these techniques effectively handle dynamic geometry. The approach referenced by [5] involves an expensive precomputation phase.

III. Methods

Both boundary element methods and ray tracing approaches have advantages and disadvantages. Most geometric acoustics techniques do not model the wave properties of sound; as a result, they cannot track the phase of a sound wave. In lieu of these deficiencies, ray tracing techniques can be better directionally auralized than numerical wave simulations.

We propose a hybrid approach which performs a numerical simulation of the wave equation in tandem with a geometrical acoustics technique. The numerical wave simulation is performed in the vicinity of potentially-occluding geometry, which we refer to as the "near-object envelope", whereas the geometrical acoustics is performed elsewhere. Given that we use signed distance functions as a scene representation, we can describe the envelope's region as follows:

$$f(\vec{x}) < \frac{\lambda}{k}$$

where f is the signed distance function, λ is the wavelength, and k is the discretization coefficient for the numerical simulation. The wave equation is simulated with a finite difference time domain method using Yee's method. It runs in lockstep with the geometrical acoustics technique, iterating with the same timestep.

The regions of the simulation outside of the envelope are handled by a geometrical acoustics method. The scene's signed distance function is used to raymarch using the sphere tracing method, where the exact step length can be determined by

$$s(\vec{x}) = \min(f(\vec{x}),\ c * dt),$$

where s is the step length, x is the location of the ray, c is the speed of sound, f is the signed distance function, and dt is the timestep.

The wave simulation employs a perfectly matched layer (PML) to compensate for the energy dissipation at the wave-ray barrier. The contents of the PML region are ignored for the purpose of the wave-ray interface.

We also present a novel two-way coupling between waves and rays at the surface of the near-object envelope. When a ray intersects a grid cell within the envelope, a spherical wave is generated on the boundary of the wave simulation, adjusted for the pressure, phase,

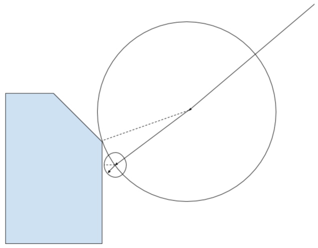

Fig. 2. Visualization of iterative sphere tracing – the signed distance function provides bounds on the length of the raymarch.

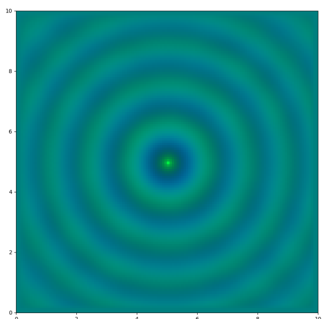

Fig. 3. Our FDTD wave simulation, modeling an exponential sine sweep signal.

and wavelength of the ray. Every timestep, the boundary sound flux is approximated by creating outgoing rays on the boundary of the simulation. Similarly, the phase of the wave on the boundary cells is used to determine the phase, pressure, and wavelength of the ray.

For the purposes of computing outgoing ray properties, we want to use the pressure and particle velocity state from the (non-PML) edges of the grid. However, Yee's method operates on a grid that is spatially and temporally staggered. Thus, to compute the presure at time t, we take an average of the neighbors of the cell at times $t - \frac{1}{2}$ and $t + \frac{1}{2}$, and then advance the ray by $t = \frac{1}{2}$ to compensate for the lost time. We also continuously compute a STFT on the barrier cells, to approximate the frequency of the outgoing rays.

We use an exponential sine-sweep to determine the response of the system. The volume of the envelope depends on the wavelength, so several frequency bins are chosen in advance. The volume of the envelope, as well as the volume of every grid cell used in the numerical integration, is determined by the average frequency in the bin. Hence, the first sweep may encompass wavelengths from 20 meters to 5 meters, the second encompasses wavelengths 5 meters to 50 cm, etc.

IV. Emergent Properties

This model can describe both early and late reverberations by taking indirect acoustical contributions into effect. Because we perform steps of forwards and backwards path tracing, we consider this a bidirectional path tracing technique. While there is prior work [1] on bidirectional paht tracing for the geometrical acoustics, our approach is somewhat simpler because the wave simulation regions can be treated as an already-illuminated surface.

The ingress model described earlier relies on the Huygens-Fresnel principle. Every point on a wavefront (i.e in this case, a plane wave) is a source of spherical wavelets. This property helps hide the aliasing from incoming rays onto the surface of the wave simulation region.

The neighborhood around every sound source, regardless of whether or not there is any occluding geometry nearby, is within the region of the wave simulation. Thus, we have no issues with aliasing the noise associated with forwards path tracing in these regions. The outgoing rays produced by the geometrical acoustics model then collide with wave simulation regions, entering those numerical simulations. This step works much like the "wave-shooting" pass in bidirectional path tracing. As a result, we can not only capture the effects of early and late reverberations, but also more complex interactions like acoustical caustics.

At the same time, we continuously compute direct contributions from the periphery of all of the wave simulations to the listener. This is the equivalent of the "gathering" pass in bidirectional path tracing. Using geometrical acoustics for auralization is much simpler (and, as mentioned in the Methods section, more directionally accurate) than directly on the numerical wave simulation.

An interesting side-effect of this technique is that this simulation is not limited to performing a fixed number of ray bounces, as most path tracers are. Regions which are dominated by indirect acoustical interactions may experience many more ray bounces than other regions.

V. Proposed Evaluation Methodology

While formal tests of this system were never completely conducted (see the Results for some discussion), we did propose an evaluation technique for a handful of simple cases. We found several issues with the testing methodology used in other acoustical simulation papers. To begin with, some papers directly compared the impulse response of a system using a ground-truth simulation to the impulse response generated by their simulation. While this technique might seem like an simple approach to comparing different acoustical simulations, small discrepancies in the impulse response are perceptually significant, and not necessarily in an obvious manner. Other evaluation strategies that compared the output of auralization could be easily affected by the chosen auralization technique.

The proposed evaluation technique involves computing a fine-grain acoustical FDTD simulation over the entire simulation space and comparing pressure measurements across the space to the corresponding pressure measurements within the wave-region of our hybrid technique. In particular, we would create a space with a variety of obstacles with different acoustical properties. The "ground truth" would be a wave simulation of the entire region, with a small spatial and temporal timestep. These results would be compared to pressure measurements taken within the hybrid solution's wave region.

If the wave simulation is faulty, then the reuslts of this experiment may not be as meaningful. However, this property need not be a cause for concern if the aim of evaluating the system is to determine whether or not the hybrid wave-geometrical acoustics system represents the same thing as a wave simulation of the entire space.

VI. Results and Discussion

Once implemented, the above model was not an accurate representation of sound representation. For the simple tests mentioned in the previous section, the egress rays did not look correct. However, we believe that with a handful of changes, this model could work.

The main issue with this model is that the particle velocity of the finite-difference time-domain simulation is not the best approximation of the direction of the wavefront. A wavefront is a the movement of planes of equivalent phase, an emergent property of the simulation pressure and particle velocity. Thus, merely approximating a plane wave via the instantaneous particle velocity at the periphery of the wave simulation would only be correct in the sense that it would be *aligned* with the direction of wave propagation. In other words, the particle velocities are collimated, but will periodically reverse.

In the case of a spherical wavefront, this ambiguity is not a problem. For the wavefront to remain a sphere, energy must be flowing outwards from the source in the center. On the other hand, even if the axis of propagation of a plane wave were determined, the *direction* of propagation would be unknown. Because we intend to model a two-way coupling at the periphery, we cannot easily determining whether the plane wave at the periphery is incoming or outgoing.

Given that our frequency decomposition constrains the frequency of the wave, we can use the angular frequency and the wavenumber from the FDTD simulation to determine the group velocity of the plane wave at the boundary. While this model can be used to generalize the egress plane waves into wave packets, we believe that a plane wave is a better representation of wave propagation in empty space. For the above model to work, the gradient of the dispersion relation must be uniform across the periphery of the simulation. If that is the case, then the wave packets will represent plane waves.

VII. Conclusion and Future Work

One of the major challenges mentioned in the literature of "hybrid" ray-wave approaches is bridging the gap between the model of the numerical wave simulation and the geometrical acoustics model. As mentioned before, there are a plethora of numerical simulation techniques that can describe complex near-object interactions. Most wave properties can be determined through the internal parameters of the numerical simulation. The geometrical acoustics side of the model can be as simple as a measure of incoming energy on the ray axis, with no measure of frequency or phase, or as complex as a wave packet. It is difficult to connect parameters in the numerical wave simulation to the geometrical acoustics and vice versa. In general, previous literature would precompute terms from the numerical simulation, which is unacceptable for our purposes. Our proposal was to recompute the entire system's impulse response as the environment changed or the user moved. While this operation would be expensive, agressive spatial and frequency decomposition could be used to reduce the burden.

In terms of optimization, both the signed distance function update and the wave simulation are best performed in parallel on the GPU. In particular, our CPU implementation described the forward differences in Yee's method as 2D convolutions which could be easily performed on a GPU. This property originates from the fact that most of the FDTD operations are either matrix additions, multiplications, or convolutions. Most of the operations in the numerical simulation would not only be cache friendly, but could also take advantage of shared memory. While the number of rays may not have strict bounds, the signed distance function update is effectively a large matrix multiply with a single non-linear operation.

Most new visual and acoustic rendering systems are developed to address a salient real-world quality that is not easily representable with a common model. We believe that wave properties, such as diffraction and caustics, are perceptually important. Including wave properties in an interactive simulation need not be limited to simple amendments to geometrical acoustics, such as the uniform theory of diffraction. We can benefit from a real-time numerical wave simulation by limiting the simulation region to a small envelope around geometry of interest.

REFERENCES

[1] S. Siltanen, T. Lokki, and L. Savioja, "Rays or waves? understanding the strengths and weaknesses of computational room acoustics modeling techniques."

[2] C. Cao, Z. Ren, C. Schissler, D. Manocha, and K. Zhou, "Interactive sound propagation with bidirectional path tracing," *ACM Transactions on Graphics (TOG)*, vol. 35, no. 6, p. 180, 2016.

[3] P. T. Calamia and U. P. Svensson, "Fast time-domain edge-diffraction calculations for interactive acoustic simulations," *EURASIP Journal on Applied Signal Processing*, vol. 2007, no. 1, pp. 186–186, 2007.

[4] S. F. Frisken, R. N. Perry, A. P. Rockwood, and T. R. Jones, "Adaptively sampled distance fields: A general representation of shape for computer graphics," in *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*. ACM Press/Addison-Wesley Publishing Co., 2000, pp. 249–254.

[5] P. T. Calamia and U. P. Svensson, "Fast time-domain edge-diffraction calculations for interactive acoustic simulations," *EURASIP Journal on Applied Signal Processing*, vol. 2007, no. 1, pp. 186–186, 2007.

[6] S. F. Frisken, R. N. Perry, A. P. Rockwood, and T. R. Jones, "Adaptively sampled distance fields: A general representation of shape for computer graphics," in *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*. ACM Press/Addison-Wesley Publishing Co., 2000, pp. 249–254.