RAY ACOUSTICS USING COMPUTER GRAPHICS TECHNOLOGY

Niklas Röber, Ulrich Kaminski and Maic Masuch

Games Research Group
Department of Simulation and Graphics,
Otto-von-Guericke-University Magdeburg, Germany

niklas@isg.cs.uni-magdeburg.de

ABSTRACT

The modeling of room acoustics and simulation of sound wave propagation remain a difficult and computationally expensive task. Two main techniques have evolved, with one focusing on a real physical - wave-oriented - sound propagation, while the other approximates sound waves as rays using raytracing techniques. Due to many advances in computer science, and especially computer graphics over the last decade, interactive 3D sound simulations for complex and dynamic environments are within reach.

In this paper we analyze sound propagation in terms of acoustic energy and explore the possibilities to map these concepts to radiometry and graphics rendering equations. Although we concentrate on ray-based techniques, we also partially consider wave-based sound propagation effects. The implemented system exploits modern graphics hardware and rendering techniques and is able to efficiently simulate 3D room acoustics, as well as to measure simplified personal HRTFs through acoustic raytracing.

1. INTRODUCTION

Physically correct sound simulations of larger and more complex environments remain a difficult, if not impossible task. This is mainly due to the extensive nature of sound wave propagation, along its complex interaction with scene objects. Unlike light, the audible spectrum covers a large area of frequency bands (octaves), and is additionally, due to a slow propagation, highly time-dependent. Although, this introduces several complications, it also allows, in certain situations, to discard some of the wave phenomena, especially for the higher frequency bands. As a result, two main approaches have evolved for the simulation of sound wave propagation: The wave-based and the ray-oriented techniques, with the first one concentrating on the lower and the last one on the middle and higher frequency ranges. Here Section 2 has a closer look on both techniques and compares them in terms of efficiency and applicability. Although several improvements have been reported for both techniques, sound simulations are in general performed offline and are valid only for certain frequency ranges. Due to advances in computational power, as well as in computer graphics and acoustics, interactive and dynamic ray-based sound simulations are feasible also for complex and more difficult scenes.

Accelerated and driven by computer games and the demand for an even higher visual realism, computer graphics hardware has evolved tremendously over the last decade and nowadays outperforms the CPU in terms of computational capacity by several magnitudes. As of the easy availability of this processing power, graphics hardware has been exploited in a number of non-graphics calculations, such as solving differential equations, as well as for simulations and numerical analyses [1]. The GPU is, in general, very well suited for the computation of parallel problems and was also more recently employed as DSP for sound signal processing [2, 3]. In the area of sound simulations, the GPU was used to solve basic geometric room acoustics [4], as well as wave-based sound propagation using waveguide meshes [5]. Besides some physical differences, the propagation of sound and light share several similarities that make existing graphics rendering techniques exploitable to accommodate an acoustic energy propagation model.

The goal of this work is to build a foundation for ray-based sound simulations using an acoustic energy propagation model, and furthermore, to demonstrate its applicability and efficiency using modern graphics hardware and rendering techniques. We derive the acoustic rendering equations from global illumination models and radiometry used in computer graphics [6], and extend the existing model by time- and frequency dependencies. This paradigm is later employed in a GPU-based implementation to perform realtime sound simulations using ray-based techniques for the applications of room acoustics and personalized HRTF simulations. The audible spectrum is divided into 10 frequency bands, which are interpreted individually with respect to their wavelength and energy. The local energy contribution of each surface patch is evaluated separately per frequency band using functions of reflection, transmission/refraction, absorption and diffraction. Finally, the acoustic energy at the observers position is accumulated and filtered regarding direction and distance using HRTFs. The system allows us to simulate realtime interactive and dynamic environments with varying acoustic materials, but also to approximate individual HRTFs through ray-acoustic simulations.

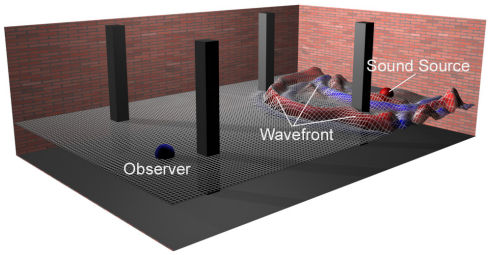
The paper is organized as follows: After this introduction, we review in Section 2 the existing approaches for sound simulations and compare their advantages and drawbacks. Section 3 follows up on the ray-based approach and develops a model for the propagation of acoustic energy in enclosures. This model studies the flow of acoustic energy from sound sources, its local interaction with objects and materials, as well as the measurement using a scene mounted listener. The following Section 4 maps the here developed concepts onto graphics primitives and rendering techniques, and discusses its implementation using modern programmable graphics hardware. Section 5 presents and discusses results using examples from room acoustic simulations and personalized HRTF measurements. The closing Section 6 summarizes the work and discusses several ideas for future improvements.

2. ACOUSTIC SIMULATION TECHNIQUES

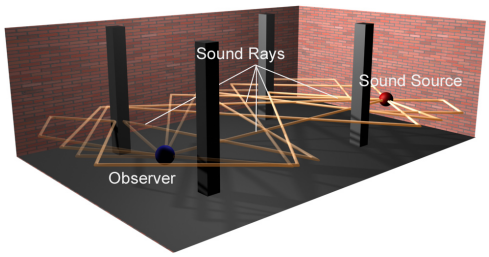
Auralization is defined as the simulation and reproduction of the acoustic properties describing a virtual scene, which has applications in many areas, including architectural design, sound and mu-

sic production and even audio-based computer games [7]. An accurate and efficient simulation is thereby still a difficult and computationally extensive task.

The most often employed approaches are waveguide meshes and raytracing techniques, see also Figure 1. Figure 1(a) displays here a visualization of the waveguide technique, a more physically correct wave-based sound propagation model, based on time-domain finite difference meshes. The acoustic energy, eg. pressure, is distributed along sampling points using difference equations. Figure 1(b) shows a visualization of the ray-based approach that approximates sound waves through particles and acoustic energy, and were raytracing techniques are used to determine the virtual soundfield. As both techniques have their own advantages and limitations, the wave-oriented techniques are usually employed for the lower frequency end, while the ray-based techniques are used for the middle and higher frequency parts.



(a) Wave-based Approach.



(b) Ray-based Approach.

Figure 1: Acoustic Simulation Techniques.

2.1. Wave-based Acoustics

Wave-based room acoustics is concerned with the numerically evaluation of the wave equation in order to simulate sound wave propagation. Often employed techniques are finite element methods (FEM) and 3D waveguide meshes (time-domain difference models) [8, 9]. The 1-dimensional waveguide technique is a numerical solution to the wave equation and was first applied to simulate string-based musical instruments [10]. The digital waveguide mesh is an extension of the 1D technique and constructed by bilinear delay lines that are arranged in a mesh-like structure [8]. Higher dimensions are built by scattering junctions that are connected to the delay lines and act as spatial and temporal sampling points. The equations that govern the rectilinear waveguide mesh are based on difference equations derived from the Helmholtz equation by discretizing time and space [11]. Depending on the mesh's resolution and the internodal sampling distance, the simulations can be rather expensive. Due to advances in computing power, realtime wave-based room acoustics is feasible for smaller meshes.

Although, the simulations using waveguide meshes are very accurate, there are some drawbacks as well. The two major problems are a direction dependent dispersion error, and a finite mesh resolution to model a more complex boundary behavior [8]. Several approaches have been discussed to overcome these limitations and include higher tesselated meshes, different mesh topologies and frequency warping techniques [12, 13]. Additionally, the sampling frequency of the rooms impulse response needs to be oversampled, with previous research showing that a typical waveguide mesh gives a valid bandwidth only as far as $f_{update}/4$ [8]. Therefore, this technique is only practical to the very lower frequency end. However, recent research has shown that waveguide meshes can easily and efficiently be implemented using graphics hardware. Combined with a new sampling lattice, the performance increase was measured by a factor of 25, and even more for finer mesh resolutions [5].

2.2. Geometric Acoustics

Geometric acoustics is based on optical fundamentals and light propagation and approximates sound waves through particles moving along directional rays [14, 15]. These rays are traced through a virtual scene, starting at the sound source and towards a listeners position, at which the accumulated energy is later evaluated. As sound waves are now simplified as rays, wave phenomena and differences in wavelength are usually discarded and ignored. This method is therefore only applicable to frequencies whose wavelength are much shorter than the dimensions of the enclosure, or any object within, refer also to [16, 17].

Several articles have been published over the last years, which discuss the realtime possibilities of ray-acoustic sound simulations [18, 19]. The majority of implementations, however, employs raytracing only to determine specular reflections using ray/beam-tracing approaches and uses conventional 3D sound APIs for spatialization and sound rendering [14, 15, 4]. As raytracing is a long known area of research in computer graphics, several improvements and advancements to the original approach have been proposed, and were partially applied to ray-acoustics as well. Savioja et.al. have designed the DIVA auralization system based on a ray-acoustics approach, to examine modeling techniques for virtual acoustics, as well as for physically-based auralizations [20, 21].

Some of the more recent geometric acoustic implementations already utilize computer graphics hardware to increase the simulations efficiency. Jedrzejewski uses the GPU for simple 2D geometric room acoustics using rays and specular reflections [4], while Kapralos and Deines employ a particle-based system to adopt the phonon mapping technique towards a *phonon tracing* approach [22, 23, 24]. Although, this technique allows an accurate modeling of acoustic materials and sound propagation, it only permits static and non-changing environments. Interesting, from the perspective of a complete GPU-based sound simulation and rendering approach, is also the work by Gallo and Whalen [3, 2], who employ the GPU as DSP for sound signal filtering and synthesis.

3. ACOUSTIC ENERGY PROPAGATION

Sound is the propagation of mechanical energy in the form of pressure variations and can be described by attributes such as frequency, wavelength, speed of propagation etc. Light on the other hand is an electromagnetic radiation, which is described by similar, however, largely different quantities. The propagation of light

f_j	f_{range_j} (Hz)	f_{center_j} (Hz)	λ_{center_j} (m)
f_0	22 – 44	31.5	10.88
f_1	44 – 88	63	5.44
f_2	88 – 177	125	2.74
f_3	177 – 354	250	1.37
f_4	354 – 707	500	0.68
f_5	707 – 1,414	1,000	0.343
f_6	1,414 – 2,828	2,000	0.172
f_7	2,828 – 5,657	4,000	0.086
f_8	5,657 – 11,314	8,000	0.043
f_9	11,314 – 22,627	16,000	0.021

Table 1: Frequency Bands f_j.

energy and its interaction with objects can be measured and described by using techniques of radiometry, from which global illumination models used in computer graphics are derived [6]. The concepts of radiometry, along its properties and equations, can be mapped to the propagation of acoustic energy as well. This assumes that the propagation of sound waves can be simplified to a ray-based approach by largely neglecting characteristics such as wavelength, diffraction and interference. For middle- and higher frequencies, and depending on the rooms and enclosed objects size, this assumption is true to a certain degree. Especially at the lower frequency end wave-based effects become such prominent that they prevail. Therefore, the here discussed model also addresses these issues and incorporates the wavelength to approximate diffraction and interference effects. The following sections discuss the theories behind, and extend the concepts of radiometry towards a ray/energy-based acoustic propagation model suitable for sound wave simulations.

3.1. Emission and Radiation

In order to study and describe the propagation of sound waves using raytracing techniques, an adequate propagation model that incorporates time- and frequency dependencies needs to be defined. This can be realized in analogy to the physics of light transportation and global illumination models [6], which now have to be extended and adopted towards acoustic properties and an acoustic energy propagation [25].

Whereas the wavelength of the visible spectrum ranges only between $380\ nm$ to $780\ nm$, the wavelength in acoustics spreads from $17\ mm$ at $20\ kHz$ up to $17\ m$ at a frequency of $20\ Hz$. The frequencies in the audible spectrum are classified and described by frequency bands (octaves) according to human psychoacoustics. In the following sections f_j describes a certain frequency band, with j being the index number and $j+1$ the next higher octave. Table 1 provides an overview of the different frequency bands, along their index number, frequency range f_{range_j}, center frequency f_{center_j} and center wavelength λ_{center_j}. The audible spectrum $A_{spectrum}$ is therefore defined as the sum of these 10 frequency bands:

$$A_{spectrum} = A_s = \sum_{j=0}^{9} f_j. \tag{1}$$

Similar to light, acoustic energy can be described as the amount of pressure variations per unit volume and time, or more accurately, by the changes in velocity of air particles contained in a volume element per unit time. The quantity for describing and measuring acoustic energy is radiant power Φ, or flux, and measured in $Watt$ or $Joule/sec$ [6]. The intensity is thereby described as the amount of acoustic energy flowing from/to/through a surface element per unit time:

$$I(t) = \frac{d\Phi}{dA}dt. \tag{2}$$

The transfer of acoustic energy using a participating media (air) is characterized by the energy transport theory. The energy density in the medium of propagation is hereby the sum of the kinetic and potential energy per unit volume dV and time $E(t) = E_{kin}(t) + E_{pot}(t)$ [25]. The kinetic energy density is defined as the pressure of a sound wave as:

$$E_{kin}(t) = \frac{1}{2}\frac{Mv^2}{V_0}dt = \frac{1}{2}\rho_0 v^2 dt, \tag{3}$$

with v being the average velocity of air particles, ρ_0 the average media density and $\frac{M}{V_0}$ its mass per unit volume V_0. The potential energy density can be derived from the gas law as:

$$E_{pot}(t) = \frac{\int pdp}{c^2\rho_0}dt = \frac{1}{2}\frac{p^2}{c^2\rho_0}dt, \tag{4}$$

with p as the pressure of the sound wave and c as the speed of sound in this medium, and therefore defines the total amount of acoustic energy density [25] as:

$$E(t) = E_{kin}(t) + E_{pot}(t) = \frac{1}{2}(\rho_0 v^2 + \frac{p^2}{c^2\rho_0})dt. \tag{5}$$

Equation 5 is valid at any position and time within the virtual auditory environment and serves as basis to describe an acoustic energy propagation model. In order to quantitatively measure flux per unit projected surface area and per unit angle, radiance is introduced with:

$$L(x,\Theta) = \frac{d^2\Phi}{d\omega dA cos\theta}, \tag{6}$$

which varies with position x and the ray's direction Θ. By incorporating the wavelength λ_j of the frequency bands used (ref. Table 1), Equation 6 is redefined to:

$$L(x,\Theta,f_j) = \int_{A_s} L(x,\Theta,f_j)d\lambda. \tag{7}$$

The acoustic energy interacting with a surface element can be further differentiated in incident E_i (incoming) and exitant E_e (outgoing) energy, and is also measured in $Watt/m^2$:

$$E_i = \frac{d\Phi}{dA}, E_e = kE_i. \tag{8}$$

The scalar k is herby defined over $[0,1]$ and describes the reflectivity of the surface with $E_{surface} = E_i - E_e$ and is affected by the surface material definition. Using a lossless participating media, the exitant radiance at one point $L(x_1 \rightarrow \Theta)$ is exactly the same as the incident radiance at another point receiving this amount of energy $L(x_2 \leftarrow \Theta)$ [6]. Using a density function and volume elements, $p(x)dV$ defines the physical number of *sound particles* carrying an *acoustic energy quant*. If moved in time dt across a differential surface area dA, and by using the direction ω and speed of propagation c; $N = p(x,\omega,f_j)cdtdAcos\theta d\omega d\lambda$

describes the number of particles flowing through this surface element. The radiance per unit volume is accordingly redefined to:

$$L(x,\Theta,f_j) = \int_{A_s}\int p(x,\omega,f_j)h\frac{c}{\lambda_j}d\lambda. \quad (9)$$

An energy/sound source emits acoustic energy that is propagated through and by the participating media. The energy radiates through an emittance pattern, which can be homogenous in any direction, eg. spherically, or direction dependent, such as a cone. As with light, also acoustic energy attenuates with distance using the familiar inverse square law. Furthermore, atmospheric absorption occurs, at which certain frequencies are absorbed by the propagating media. However, this factor is very small and can safely be ignored for smaller enclosures, but becomes more prominent with increasing distances.

An observer, or listener, can be placed anywhere within the scene to *record* the acoustic energy present at this location. The listener does not interfere or participate in the energy propagation, but, if required, such as for binaural listening, an additional geometry can be placed nearby to simulate head-shadowing effects. The incoming rays are then weighted and filtered using HRTFs regarding the ray's direction and delay.

3.2. Local acoustic Energy Exchange

The most interesting part in a ray-based acoustic simulation is the interaction and exchange of acoustic energy with objects and surface elements. Depending on the objects size and the acoustic material parameters specified, some of the incoming energy might get absorbed, reflected, refracted or transmitted, with the total amount of energy according to Equation 8 being constant.

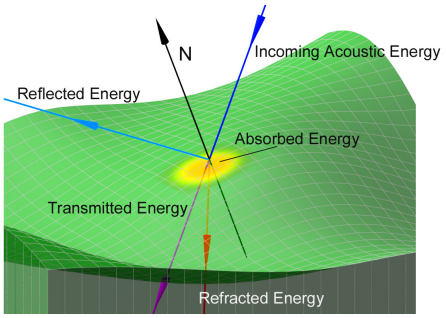


Figure 2: Local acoustic Energy Exchange.

Figure 2 shows a schematic of the local acoustic energy exchange. The four effects of absorption, reflection, refraction and transmission are described in more detail in the remainder of this section. Every ray that is cast into the scene contains, depending on the sound source emittance of course, the energy of all frequency bands. The energy contribution of each ray is evaluated at the point of intersection with the surface patch using the ray's length, as well as the surface material properties defined.

Some of the incident acoustic energy is thereby usually absorbed, converted into heat and dissipated back into the system. The absorption is frequency dependent and characterized by a frequency band coefficient α_{f_j}:

$$L_{e_{absorbed}}(x \leftarrow \Theta) = \sum_{j=0}^{9} E_{i_j}\alpha_{f_j}. \quad (10)$$

Transmission is defined as the energy that passes through an object. We redefine this term to describe the frequency-weighted amount of energy that passes through an object *unaltered* and without refraction. In acoustics, objects smaller than the wavelength of an incoming sound wave do not interfere, instead the wave simply diffracts around the object and continues unchanged. An according frequency dependent modeling of energy transmission can be realized using an objects bounding box or sphere that simply transmits all acoustic energy whose wavelength is equal or above the objects size:

$$L_{e_{transmitted}}(x \rightarrow (\pi+\Theta)) = \sum_{j=0}^{9} E_{i_j}\tau_{f_j}. \quad (11)$$

Here $L_{e_{transmitted}}(x \rightarrow (\pi+\Theta))$ describes the amount of exitant energy per ray for all bands, which simply pass along the direction opposite to the incoming ray, i.e. the ray's original direction. The term τ_{f_j} is used for a finer modeling and a frequency-weighting of the transmission effects.

Reflection and diffuse scattering are probably the two most important qualities in acoustic raytracing and can be very well described using bidirectional reflection distribution functions (BRDF) [6]. A BRDF is defined for a point x as the ratio of the differential radiance reflected in an exitant direction Θ_e and the differential irradiance incident through an incoming angle Θ_i:

$$brdf_{reflected}(x,\Theta_i \rightarrow \Theta_e) = \frac{dL(x \rightarrow \Theta_e)}{dE(x \leftarrow \Theta_i)}. \quad (12)$$

The BRDF is frequency dependent, but direction independent, eg. $f_r(x,\Theta_i \rightarrow \Theta_e) = f_r(x,\Theta_e \rightarrow \Theta_i)$ [6, 26]. Diffuse scattering uniformly reflects the incoming acoustic energy in all directions. In acoustics, this behavior is largely influenced by the surface roughness, which can be used to determine a specular reflection coefficient that describes the ratio between specular and diffuse reflections. Using a complete diffuse scattering, the radiance is independent from the angle of exitance and the BRDF defined as:

$$brdf_{reflected}(x,\Theta_i \leftrightarrow \Theta_e) = \frac{\rho_{diffuse}}{\pi}, \quad (13)$$

in which the reflectance $\rho_{diffuse}$ represents the fraction of incident energy reflected at the surface. Pure specular reflection on the other hand diverts all incident energy in only one direction R, which can be simply computed using the law of reflection and the surface normal N: $2(N(\pi+\Theta_e))N-(\pi+\Theta_e)$. A frequency dependent BRDF for acoustic raytracing can be modeled through:

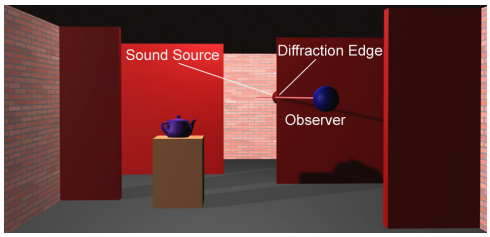
$$L_{e_{reflected}}(x \leftarrow \Theta_i) = \sum_{j=0}^{9} E_{i_j}\upsilon_{f_j}, \quad (14)$$

in which υ_{f_j} is a weighting factor per frequency band f_j. The majority of materials, however, exhibit a sort of *glossy* surface, a combination of specular reflection and diffuse scattering.

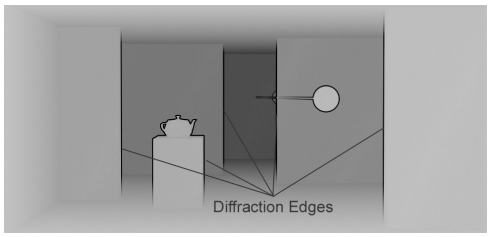
Refraction occurs at the crossing of two different isotropic media and can be computed similar to the reflection term in Equation 12, except that the outgoing angle Φ of the refracted ray is determined using Snell's Law: $sin\Phi = \frac{\eta_2}{\eta_2}$. Here η_1 and η_2 are the refraction indices of their respective media. A frequency band weighted refraction can be defined similar to Equation 14 by using ν_{f_j} as weighting coefficient per frequency band.

3.3. Diffraction and Interference

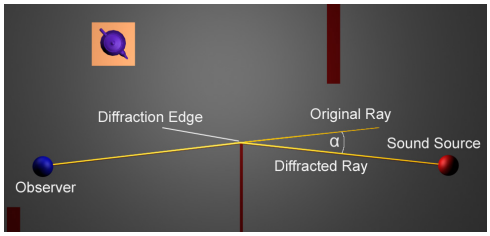
Edge diffraction and interference are acoustic phenomena that can be modeled accurately using wave-based techniques, but do not fit well into the concept of ray-acoustic simulations. However, both are very important and prevail especially in the lower frequency ranges. Therefore, and in order to obtain a more *realistic* simulation, these effects have to be included, or at least approximated. Generally, this is done by combining wave-based and ray-based approaches and by choosing a certain threshold as boundary frequency. But to a certain degree, these effects can also be approximated within ray-acoustics.



(a) Scene Rendering as seen from Listener's Position.



(b) Combined Depth/Edge Map.



(c) Top View with original and diffracted Ray.

Figure 3: Ray Acoustic Diffraction Simulation.

Sound waves with larger wavelength simply bend around edges, such as if an additional sound source was placed at the diffraction edge. Diffraction effects are in ray/energy acoustics simply modeled through ray-bending, according to the ray's length and its associated frequency band f_j. As diffraction is dependent on the objects size and the ray's wavelength, the amount of energy that is diffracted is determined individually per frequency band f_j. The maximum possible diffraction angle was herby determined experimentally using a wave-based sound propagation system [5]. Figure 3 visualizes the concept of the implemented diffraction system. It shows a virtual scene from the listener's perspective (Figure 3(a)), the constructed edge map (Figure 3(b)) and the by angle α diffracted ray from a listener to a sound source (Figure 3(c)). For each edge in Figure 3(b), additional rays are cast into the scene for diffraction simulation.

Interference describes the superposition of two or more sound waves and the resulting changes in amplitude. Using a ray-acoustic sound simulation, interference effects can only be approximated roughly using the ray's length and the center wavelength λ_{center_j} of the current frequency band f_j. By using an additional scalar associated with each ray, also the modeling of phase-preserving and phase-reversing reflections are possible. The next section focusses after these theoretical discussions on the implementation of the here described acoustic energy propagation model using efficient computer graphics hardware.

4. GRAPHICS-BASED RAY ACOUSTIC SIMULATIONS

While the last section discussed the propagation of acoustic energy and its interaction with objects and materials, this section maps the there developed concepts onto computer graphics primitives and rendering equations. The presented framework implements a ray-based acoustic simulation system that exploits modern computer graphics hardware. The system is designed along current GPU-based raytracing systems [27, 28, 29], which were extended towards the acoustic energy propagation model as discussed in the last section. The advantages and applicabilities of such an implementation can be summarized as:

- Efficient ray-based acoustic simulation system that incorporates wave phenomena,
- Realtime implementation that exploits graphics hardware,
- Built-in visualization of sound wave propagation,
- An eclectic modeling and design of acoustic materials, with
- Applications for impulse response measurements and general room acoustics, as well as to
- Approximate individualized HRIRs.

The system takes any 3D polygonal mesh as input, which is pre-processed into a more efficient accessible structure. It allows an interactive sound simulation for meshes of up to 15,000 polygons. Using a short pulse as sound signal, room impulse response (RIR), as well as head-related impulse response (HRIR) measurements are possible. Alternatively, a monaural sound file can be used as input signal, resulting in a spatialized binaural representation with the virtual rooms imprint. The sound source/listener positions, as well as the acoustic material definitions can be changed and adjusted interactively. All sound signal processing, including HRTF convolution and delay filtering, is realized using fragment shaders onboard the graphics hardware.

4.1. Auralization Pipeline

The auralization pipeline employed in our system stretches over the CPU and GPU systems, but the majority of computations is carried out in graphics hardware. Figure 4 shows an overview of the pipeline, along its partition in CPU and GPU related tasks. As initialization, 3D scene data, as well as sounds and frequency band decomposed HRTFs are loaded into texture memory. The sound data is also decomposed into 10 bands and assigned a position and emittance pattern within the virtual room. Rays are now cast into the scene starting at the listeners position, and the per frequency band received acoustic energy is accumulated and stored within so called cubemaps. This cubemap is later evaluated and the sound data is filtered and delayed using HRTFs according to their position and the ray's length. The binaural mixdown is performed

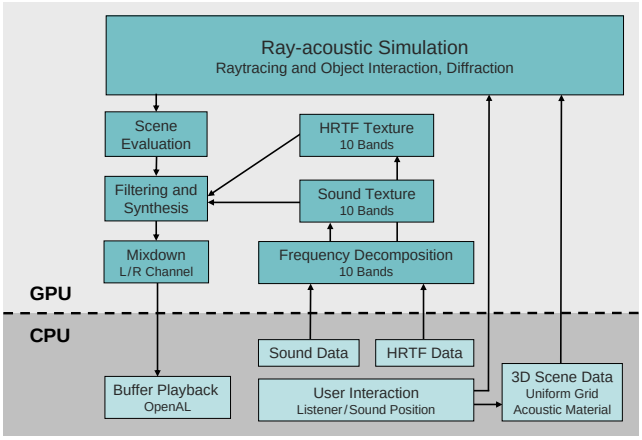


Figure 4: Auralization Pipeline.

using a two-channel floating point texture, which is streamed back to the CPU and fills a native OpenAL stereo buffer for sound playback.

4.1.1. *Uniform Grid Structure*

In a pre-processing step, the 3D scene is converted into a uniform grid structure that subdivides 3D space and groups neighboring triangles together in a voxel-based topology. These voxels are of uniform size and axis aligned. This space subdivision is necessary in order to efficiently determine ray/object intersections, as now only the triangles grouped in one voxel element have to be tested [27, 28, 29]. Care has to be taken in defining the voxel's size, as with very detailed objects the number of polygons can easily exceed the number of possible shader instructions.

4.1.2. *Frequency Decomposition and Synthesis*

A frequency-based acoustic raytracing has many advantages, as now some of the wave-based propagation effects can be approximated, as well as it allows a more realistic frequency-dependent definition of acoustic materials. Currently we employ 10 frequency bands, grouped into octaves as known from psychoacoustics, see Table 1. For the frequency decomposition of sound data and HRTFs, we employ a time-based convolution using windowed sinc filters, with their cutoff frequencies specified as the bands respective border frequencies. These 10 bands are loaded as floating point textures into graphics hardware. To remain data precision, we currently employ 3 RGBA textures to hold the sound data, although, using data compression, two should be sufficient for 16 bit sound data. Ray/object interactions are evaluated per frequency band and the contributions from each ray are accumulated and also stored individually. The final auralization is a binaural sound signal that is generated by filtering the original sound texture using HRTFs according to the simulations result.

4.2. Acoustic Raytracing and Diffraction Simulation

The authoring of 3D scenes can be conveniently performed using 3D Studio MAX, where a custom-built plugin is used to assign acoustic material definitions to each object. This acoustic material defines the wavelength specific energy exchanges for each surface patch. Although, all materials are assigned per vertex, no interpolation of neighboring material attributes is performed yet.

The raycasting and acoustic energy accumulation is carried out using so called cubemaps. One cubemap is hereby centered around the observers position and a ray is cast into the scene per cubemap texel. Figure 5 shows a visualization of this cubemap raycasting approach. Each ray cast is traced through the virtual scene and its acoustic energy accumulated and stored per frequency band. At points of ray/object intersection, the local surface acoustic energy exchange is evaluated according to Section 3.2. Newly generated rays from refraction, transmission and/or reflection are further traced, until their possible energy contribution falls below a certain threshold ϵ. The cubemap not only stores all incoming acoustic energy per frequency band, but also the ray's direction and length. This information is later used for the final binaural sound signal synthesis.

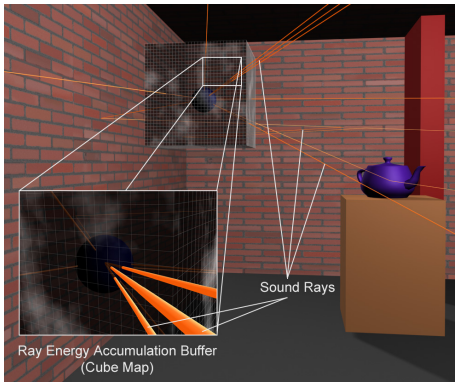


Figure 5: Ray Energy Accumulation Buffer.

4.2.1. *Diffraction Simulation*

The ray acoustic simulation also incorporate diffraction effects on edges and object borders. To find possible diffraction locations, a depth/edge map is employed, which highlights these edges, see also Figure 3(b). These maps are created by using the scenes depth buffer and an image-based edge detection algorithm. If a ray is cast close to a diffraction edge, the ray is bend according to the diffraction of sound waves [16, 17]. Here the ray's energy is attenuated, depending on the angle and the ray's wavelength. Another wave-based phenomena is interference, which can be roughly approximated by using the ray's length and the frequency bands center wavelength λ_{center_j}, see Table 1. Although, this is a very simple approximation, it would also allow the modeling of phase-reversing and -preserving boundary reflections, as well as to use this information for interference effects of the same and/or different frequency bands.

4.3. Implementation

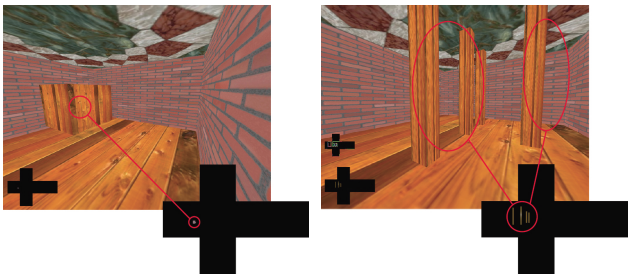
Today's graphics hardware, and especially the new generation with its unified shader architecture, can be seen as powerful parallel processing machines, which can very efficiently execute small programs - so called shaders - in parallel. Shaders are freely programmable using high level shading languages such as GLSL and Cg. As graphics applications typically require the processing of huge amounts of data, graphics hardware has been optimized to support

this with a highly parallel design. Combined with a fast and optimized memory architecture for accessing and storing the data, this makes this hardware very interesting for any computationally intensive and parallelizable task.

All convolutions and sound synthesis are carried out using fragment shaders on graphics hardware, with a single shader for each task. The data, eg. sounds, geometry and material definitions are stored within textures and accessed during the rendering task from within the shaders. The results of the simulation are again stored as textures, from which they are read back to the CPU for sound playback.

5. RESULTS AND DISCUSSION

This section discusses some results of the ray-acoustics simulation system. The implementation is based on *n*vidia type graphics hardware and uses Cg as shading language. The current experiments were performed with three different graphics hardware generations, showing that only the newest one (GeForce8800GTX) was also able to additionally perform a realtime auralization of the results besides the sound simulation. Frame rates of up to 25 fps could be achieved using a detailed model of a living room (1,500 polygons) including a binaural auralization of the scene, ref. Figure 7(b).



(a) Transmission and Refraction. (b) Diffraction around Pillars.

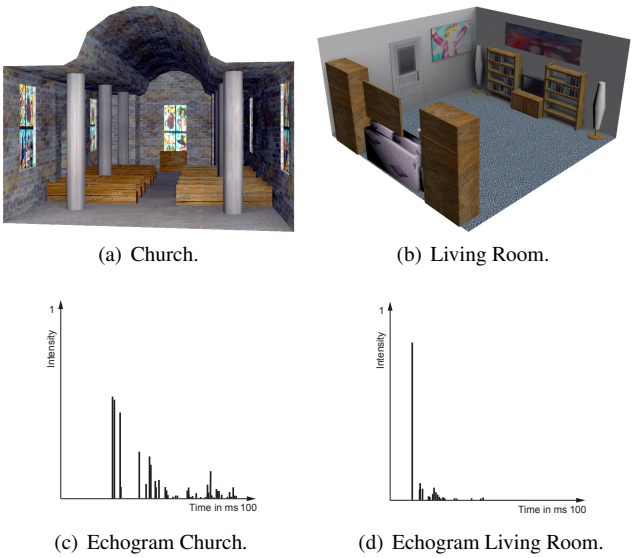
Figure 6: Evaluation of Sound Propagation Effects.

Figure 6 shows two visualizations of sound propagation effects. Here Figure 6(a) displays the transmission and refraction parts of the simulation, whereas Figure 6(b) shows diffraction effects of several pillars. In both cases the sound source is hidden and the simulation results are visible in the unfolded cubemaps below. Both cubemaps show a red/braun shifting of the color, denoting a stronger transmission/diffraction in the lower frequencies.

5.1. Example 1: Room Acoustics

The first example shows two different rooms along their echograms. Figure 7(a) displays thereby a small church, while Figure 7(b) shows an average living room. The echogram of the church, ref. Figure 7(c) shows strong and late echoes, while the echogram in Figure 7(d) shows that nearly all acoustic energy, except the direct line, was absorbed by walls and furniture. Both echograms clearly visualize the rooms acoustic properties. Each room has been modeled using 3D Studio MAX, in which for each surface a different materials has been specified. The properties for the acoustic material definitions were taken from the CARA database[1]. The

[1] `http://www.cara.de`



(a) Church. (b) Living Room.

(c) Echogram Church. (d) Echogram Living Room.

Figure 7: Room Acoustic Example Scenarios.

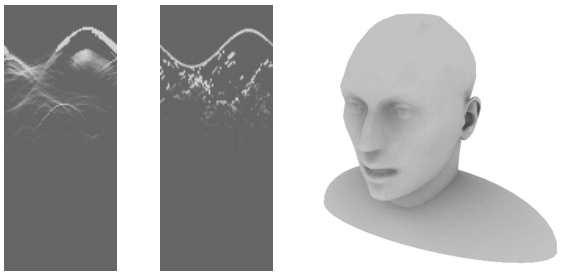
echograms show the early reflections, as well as late reverberation and diffraction effects.

5.2. Example 2: Personalized HRTF

The second example shows an HRIR simulation of the horizontal plane using our ray-acoustics approach. The simulation was performed using 72 sound sources, each 1.2 m apart from the head at a 5 degree interval. Although, the simulation does not exhibit all effects of a regular measured HRIR, it shows the most prominent features. The simulation was performed using a 3D model of the KEMAR mannequin. Figure 8 shows two different simulation results, along the original 3D model used. Here Figure 8(a) displays an HRIR simulation of the system from [30], while Figure 8(b) show the results of the here presented ray-acoustics system. Thereby roughly 18 million rays were traced per sound source, resulting in a simulation time of 22 seconds per sound source. Although the most important features are clearly present, several effects are still missing. This is partially due to the fact that we only consider one diffraction per ray. Also, a more detailed fine tuning of parameters along the material definitions for the head, torso and ear will yield better results. The goal is to combine an individualized HRTF simulation with room acoustics, to yield a realtime personalized binaural room simulation. Better results with geometric models have been achieved by [31], but, however, also with a much longer simulation time.

6. CONCLUSIONS AND FUTURE WORK

We have presented a realtime graphics-based implementation of a ray acoustic simulation system that is based on an acoustic energy propagation model. This underlying model is founded on sound propagation, as well as global illumination models, and the ray/energy approach used therefore valid and its implementation using graphics hardware and techniques viable. The current results clearly show the possibilities of this system and motivate a further research in this area.



(a) HRIR 1. (b) HRIR 2. (c) KEMAR Model.

Figure 8: HRIR Simulation and 3D Model.

The current implementation already exhibits good and very promising results, yet some ideas are left for future improvements. One extension would be an enhanced space partitioning structure, such as kD-Trees that allow a non-uniform subdivision of 3D space. Future work also includes more and finer partitioned frequency bands for a more accurate studying and modeling of wave-based propagation effects. Another beneficial extension would be a higher incorporation of radiosity techniques, although one has to be careful to not impede here with realtime simulations and dynamic environments. Additionally, ambisonics and their implementation using spherical harmonics in realtime computer graphics might be an interesting path to explore.

7. ACKNOWLEDGEMENTS

The authors would like to thank Yuvi Kahana, Brian FG Katz, as well as Richard O. Duda for providing a polygonal mesh of the KEMAR head for the HRIR simulations.

8. REFERENCES

[1] John D. Owens, David Luebke, Naga Govindaraju, Mark Harris, Jens Krüger, Aaron E. Lefohn, and Timothy J. Purcell, "A survey of general-purpose computation on graphics hardware," in *Eurographics 2005, State of the Art Reports*, Aug. 2005, pp. 21–51.

[2] Sean Whalen, "Audio and the Graphics Processing Unit," Tech. Rep., 2005, `http://www.node99.org/projects/gpuaudio/`.

[3] Emmanuel Gallo and Nicolas Tsingos, "Efficient 3D Audio Processing with the GPU," in *GP2, ACM Workshop on General Purpose Computing on Graphics Processors*, 2004.

[4] Marcin Jedrzejewski, "Computation of Room Acoustics Using Programmable Video Hardware," in *International Conference on Computer Vision and Graphics*, 2004.

[5] Niklas Röber, Martin Spindler, and Maic Masuch, "Waveguide-based Room Acoustics through Graphics Hardware," in *ICMC*, 2006.

[6] Philip Dutre, Philippe Bekaert, and Kavita Bala, *Advanced Global Illumination*, AK Peters Ltd., 2003.

[7] Niklas Röber and Maic Masuch, "Leaving the Screen: New Perspectives in Audio-only Gaming," in *Proceedings of 11th ICAD*, 2005, Limerick, Ireland.

[8] S. VanDuyne and J.O. Smith, "Physical Modelling of the 2-D digital Waveguide Mesh," in *Int. Computer Music Conference*, Tokyo, Japan, 1993, pp. 40–47.

[9] Y. Kahana, P.A. Nelson, M. Petyt, and S. Choi, "Numerical Modelling of the Transfer Functions of a Dummy-Head and of the external Ear," in *Audio Engineering Society, 16th International Conference*, Rovaneimi, 1999, pp. 330–345.

[10] Julius O. Smith, "Physical modelling using digital Waveguides," *Computer Music Journal*, vol. 16, no. 4, pp. 75–87, 1992.

[11] Lauri Savioja and Tapio Lokki, "Digital Waveguide Mesh for Room Acoustic Modelling," in *ACM SIGGRAPH Campfire: Acoustic Rendering for Virtual Environments*, Utah, USA, 2001.

[12] Mark J. Beeson and Damian T. Murphy, "Roomweaver: A digital Waveguide Mesh based Room Acoustics Research Tool," in *7th Int. Conference on Digital Audio Effects*, Italy, 2004, pp. 268–273.

[13] F. Fontana and D. Rocchesso, "Signal-Theoretic Characterization of Waveguide Mesh Geometries for Models of Two-Dimensional Wave Propagation in Elastic Media," *IEEE Transactions on Speech and Audio Processing*, vol. 9, no. 2, pp. 152–161, Februar 2001.

[14] Wolfgang Mueller and Frank Ullmann, "A Scalable System for 3D Audio Ray Tracing," in *Proceedings of the IEEE Int. Conference on Multimedia Computing and Systems*, 1999, pp. 819–823.

[15] Thomas Neumann, "Eine Geometrie-Engine zur Berechnung von 3D-Sound mit Raumakustik-Effekten in Echtzeit," M.S. thesis, Tchnische Universität Braunschweig, Institut für Compuergraphik, 2004.

[16] F. Alton Everest, *The Master Handbook of Acoustics*, vol. 4th edition, McGraw-Hill/TAB Electronics, 2001.

[17] Heinrich Kuttruff, *Room Acoustics*, Spon Press, London, 2000.

[18] Thomas Funkhouser, Nicolas Tsingos, Ingrid Carlbom, Gary Elko, Mohan Sondhi, and James West, "Modeling Sound Reflection and Diffraction in Architectural Environments with Beam Tracing," *Forum Acusticum*, 2002.

[19] Emmanuel Tsingos, Nicolas Gallo and George Drettakis, "Perceptual Audio Rendering of Complex Virtual Environments," in *Proceedings of ACM SIGGRAPH*, 2004.

[20] Lauri Savioja, Tapio Lokki, and Jyri Huopaniemi, "Auralization applying the Parametric Room Acoustic Modelling Technique - The DIVA Auralization System," in *Proceedings of ICAD*, Japan, 2002.

[21] Lauri Savioja, *Modelling Techniques for Virtual Acoustics*, Ph.D. thesis, Helsinki University of Technology, Finnland, 2000.

[22] B. Kapralos, M. Jenkin, and E. Milios, "Sonel Mapping: Acoustic Modeling utilizing an acoustic Version of Photon Mapping," in *IEEE International Workshop on Haptic Audio Visual Environments and their Applications*, 2004.

[23] Martin Bertram, Eduard Deines, Jan Mohring, Jevgenij Jegorovs, and Hans Hagen, "Phonon Tracing for Auralization and Visualization of Sound," in *IEEE Visualization*, Minneapolis, USA, 2005.

[24] Eduard Deines, Martin Bertram, Jan Mohring, Jevgenij Jegorovs, Frank Michel, Hans Hagen, and Gregory M. Nielson, "Comparative Visualization for Wave-based and Geometric Acoustics," *IEEE Transactions on Visualization and Computer Graphics*, vol. 12, no. 5, September/October 2006.

[25] Leo L. Beranek, *Acoustics*, Amer Inst of Physics, 1986.

[26] Zhiyun Li, Ramani Duraiswami, and Nail A. Gumerov, "Capture and Recreation of higher Order 3D Sound Fields via Reciprocity," in *ICAD*, 200f.

[27] Timothy J. Purcell, Ian Buck, William R. Mark, and Pat Hanrahan, "Ray Tracing on Programmable Graphics Hardware," in *Proceedings of ACM SIGGRAPH*, 2002, pp. 703–712.

[28] Timothy J. Purcell, *Ray Tracing on a Stream Processor*, Ph.D. thesis, Stanford University, March 2004.

[29] Gabriel Moreno-Fortuny, "Distributed Stream Processing Ray Tracer," Tech. Rep., University of Waterloo, `http://www.cgl.uwaterloo.ca/~gmoreno/streamray.html` 2004.

[30] Niklas Röber, Sven Andres, and Maic Masuch, "HRTF Simulations through acoustic Raytracing," Tech. Rep., Fakultät für Informatik, Otto-von-Guericke Universität Magdeburg, 2006.

[31] V.R. Algazi, R.O. Duda, R. Duraiswami, N.A. Gumerov, and Z. Tang, "Approximating the Head-related Transfer Function using simple Geometric Models of the Head and Torso," *Journal of the Acoustical Society of America*, , no. 5, pp. 2053–2064, November 2002.