19th INTERNATIONAL CONGRESS ON ACOUSTICS
MADRID, 2-7 SEPTEMBER 2007

Acoustical Modeling with Sonel Mapping

PACS: 43.55.Ka

Kapralos, Bill[1]; Jenkin, Michael[2,3]; Milios, Evangelos[4]
[1]Faculty of Business and Information Technology, University of Ontario Institute of Technology
  2000 Simcoe St. North, Oshawa, Ontario, Canada, L1H 7K4; bill.kapralos@uoit.ca
[2]Department of Computer Science and Engineering, [3]Centre for Vision Research, York
  University, 4700 Keele St. West, Toronto, Ontario, Canada, M3J 1P3; jenkin@cse.yorku.ca
[4]Faculty of Computer Science, Dalhousie University, 6050 University Ave. Halifax, Nova Scotia,
  Canada, B3H 1W5; eem@cs.dal.ca

ABSTRACT
Acoustical modeling of even small, simple environments is a complex, computationally expensive, and time consuming task for all but the simplest environments. Given the similarities that exist between the fields of computer graphics (image synthesis) and acoustical modeling, this paper describes the application of suitably modified computer graphics and optics-based modeling methods and techniques to accurately model environmental acoustics. By accounting for the differences between the propagation of sound and light as well as differences in how propagating sound waves interact when they encounter objects in the environment, a sound synthesis method termed sonel mapping, is developed. Sonel mapping is a Monte-Carlo-based technique that models many of the complex effects that propagating acoustical signals encounter in the environment including diffuse and specular reflections, and diffraction effects. This modeling is performed in an efficient manner in contrast to available deterministic techniques. Results of various simulations demonstrate that the method conforms to theoretical results.

INTRODUCTION
Despite the benefits of spatial sound cues, they are often overlooked by the majority of immersive virtual environments where historically, emphasis has been placed on the visual senses instead [5]. Furthermore, when present, the spatial sound cues that are present do not necessarily reflect natural cues. Many systems that do convey sound information do so poorly, typically assuming that all interactions between a sound wave and objects/surfaces in the environment are specular reflections, despite that in our natural settings, acoustical reflections may be diffuse and there may also be diffractive and refracted components to the sounds we hear as well. Failure to accurately model all these phenomena leads to a decrease in the spatialization capabilities of the system, ultimately leading to a decrease in performance and a decrease in presence or immersion [15].

This work develops a probabilistic-based acoustical modeling method termed *sonel mapping*. Sonel mapping is inspired by the computer graphics (realistic image synthesis) method *photon mapping* [8] and is intended to estimate the time and frequency dependent *echogram* (the temporal acoustical energy distribution) of a particular environment.

BACKGROUND
There are two major computational acoustical modeling approaches: (i) wave-*based modelling*, and (ii) *geometric modelling*.

Wave-Based Room Impulse Response Modeling
The objective of wave-based methods is to solve the *wave equation* to recreate a particular soundfield. An analytical solution to the wave equation is rarely feasible [14] hence, wave-based methods use numerical approximations such as finite element methods, boundary element methods, and finite difference time domain methods instead [14]. Numerical approximations sub-divide the boundaries of a room into smaller elements. By assuming the pressure at each

of these elements is a linear combination of a finite number of basis functions, the boundary integral form of the wave equation can be solved [16]. The numerical approximations associated with wave-based methods are computationally prohibitive, rendering them impractical except for the simplest static environments. Aside from basic or trivial applications, such advanced techniques are currently beyond our computational ability for interactive virtual environment applications.

Ray-Based (Geometric) Room Impulse Response Modeling
Comparable to the field of geometrical optics, many acoustical modeling approaches adapt the hypothesis of geometrical acoustics that sound and rays behave the same way. The acoustics of an environment are modeled by tracing (following) these rays as they propagate through the environment while accounting for any interactions between the rays and any objects/surfaces they may encounter. Mathematical models are used to account for sound source emission patterns, atmospheric scattering, the medium's absorption of sound ray energy as a function of humidity, temperature, frequency, distance, and the interactions with any surfaces/objects the rays may encounter. At the receiver, the room impulse response is obtained by constructing an *echogram*, describing the distribution of incident sound energy (rays) over time. The echogram can be converted to an equivalent room impulse response function through a post-processing operation [13]. Geometric acoustic-based methods include *image sources* [1] *ray tracing* [12], *beam tracing* [6], *phonon tracing* [2], and *sonel mapping* [10].

Despite being simple to implement, ray-based methods typically assume that all interactions between a sound ray (wave) and objects/surfaces in the environment are specular thus, ignore diffraction effects. As a result, these methods are valid only for higher frequency sounds where reflections are primarily specular and sound can, therefore, be considered along only straight-ray paths [4] (see [3,4,16] for examples of acoustical diffraction modeling methods). Another problem associated with ray-based approaches is handling the large number of potential interactions between a propagating sound ray and any objects/surfaces it may encounter. A sound ray incident on a surface may simultaneously experience specular and diffuse reflection, be refracted, and be diffracted. Typical solutions to modeling such effects include the generation and emission of several "new" rays at each interaction point. Such approaches can lead to exponential running times making them computationally intractable except for the most basic environments and only for very short time periods.

As an alternative to common deterministic approaches to estimate the type of interaction between an acoustical ray and an incident surface, probabilistic approaches such as a Russian roulette strategy may be used instead. Russian roulette ensures that the path length of each acoustical ray is maintained at a manageable size, yet due to its probabilistic nature, arbitrary size paths may be explored. Sonel mapping employs a Russian roulette solution in order to determine the type of interaction between a "sound ray" (known as a *sonel*) and a surface to determine when the sonel is terminated [11].

THE SONEL MAPPING METHOD
Following the same strategy as used in photon mapping, rather than modeling the exact *mechanical wave* phenomena of sound propagation (e.g., particles in the medium they move about in their equilibrium position), the process is approximated by emitting one or more sound elements (*sonels*) from each sound source. These sonels are traced through the scene until they encounter the surface of an object. Each sonel can be viewed as a packet of information propagating from the sound source to the receiver, carrying the relevant information required to simulate the mechanical wave propagation. The information carried by each sonel includes the information used by photons in the photon mapping approach: position, incoming incidence direction (at the point of intersection between the sonel and the surface), and energy in addition to information specific to sound and sound propagation, including: distance travelled and frequency.

Like photon mapping, sonel mapping is a two-pass Monte-Carlo particle-based technique. In the first pass (the *sonel tracing* stage), sonels are emitted from each sound source and traced through the scene until they interact with a surface. The distribution of sound frequency in a given source is approximated by considering the center frequency of a fixed number of frequency bands (channels). Each sonel represents the energy contained in one frequency

2

band (center frequency). For the purpose of handling the modeling of acoustical diffraction, as shown in Figure 1, each original surface is *dilated* in a frequency dependent manner by an amount equal to $\lambda/2$ (where, λ is the wavelength). The dilated surface is divided into two zones: i) the *diffraction zone,* and ii) the *non-diffraction zone*. The region on the dilated surface within a distance of $\lambda/2$ of the original (non-dilated) surface edge comprises the diffraction zone and the remainder of the surface comprises the non-diffraction zone (see Figure 1). The type of interaction experienced by the sonel will depend on which zone the sonel is incident upon. A sonel incident within the non-diffraction zone will be reflected either specularly or diffusely or absorbed by the surface, the decision being made using a Russian roulette strategy. When the reflection is specular, ideal specular reflection is assumed whereby, with respect to the surface normal vector, the angle of reflectance is equal to the angle of incidence. When the reflection is diffuse, the sonel is stored in the sonel map and a new sonel will be created and reflected diffusely from the interaction (intersection) point by choosing a random direction over the hemisphere centered about point p. When a sonel is incident within the diffraction zone, the sonel will be reflected in a random direction over the hemisphere centered about the diffraction point. Diffusely reflected sonels are stored in the sonel map.
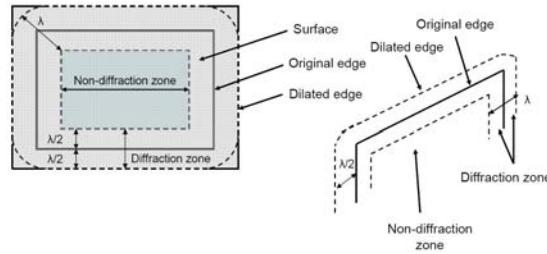

Figure 1. Diffraction and non-diffraction zones defined.

In the second stage (the *acoustic rendering* stage), the echogram is estimated through the use of the previously constructed sonel map coupled with distribution ray tracing. A number of *acoustical visibility rays* are traced from the receiver into the scene where they may interact with any surfaces/objects they may encounter. A sonel incident within the non-diffraction zone will be reflected either specularly or diffusely or absorbed by the surface, the decision once again, being made using a Russian roulette strategy. When the interaction at point p is a diffuse reflection, tracing of the ray terminates and the sonel map is used to provide an estimate of the acoustic energy leaving point p and arriving at the receiver using a *density estimation* algorithm. The energy is scaled to account for attenuation by the medium and added to the accumulating echogram. Specular reflections are handled using the same approach as in the sonel tracing stage whereby ideal specular reflections are assumed. When an acoustical visibility ray encounters a sound source, the fraction of energy leaving the sound source and arriving at the receiver is determined, scaled to account for attenuation by the medium and the added to the accumulating echogram. Diffraction effects that occur when an acoustical visibility ray encounters an edge are handled using a modified version of the Huygens-Fresnel principle [7]. Provided the sound source remains static, the information contained in the sonel map does not need to be updated and therefore, to account for the changing soundfield arriving at the receiver, only the acoustical rendering stage needs to be re-computed.. The direct sound reaching the receiver is determined by sending *shadow rays* towards the sound source in order to test for possible occlusion with any objects.

Edge Diffraction
Given a sound source (S) and receiver (R) in free space (e.g., no obstacles between them), having originated at S at time t = 0 with an amplitude E_0, at time t' the wave will have propagated a distance ρ. This expanding wavefront is divided into a number of ring-like regions, collectively known as *Fresnel zones* [7]. The boundary of the ith Fresnel zone corresponds to the intersection of the wavefront with a sphere of radius $r_0 + i \times \lambda/2$ centered at the receiver where, r_0 is equal to the distance between the receiver and the expanding wavefront after it has traversed a distance of ρ from the sound source, and λ is the wavelength of the sonel. In other words, the distance from the receiver to each adjacent zone differs by half a wavelength ($\lambda/2$). The total energy E_t reaching the receiver can be determined by summing the energy reaching

the receiver from each zone. This is approximately equal to one half of the contribution of the first zone E_1 (e.g., $E_t \approx |E_1|/2$) [7].

Essentially, given a sound source, receiver and edge, the energy reaching the receiver is determined by considering the energy arriving at the receiver from the first Fresnel zone as in the unoccluded scenario described above. To account for diffraction effects, a *visibility factor* is introduced. The visibility factor represents the fraction of the first zone visible from the receiver and is denoted by v_1. In essence, positions on the first zone are uniformly sampled and ray casting is used to determine the fraction of the zone visible to the receiver. The total visibility of the zone is equal to the fraction of sampled positions where a clear path between the sampled position and the receiver exists (n_{vis}), versus the total number of positions sampled (N_{vis}), given mathematically as $v_1 = n_{vis}/N_{vis}$. Greater details regarding the modeling of acoustical diffraction effects with sonel mapping are available in [9].

SIMULATIONS: COMPARISONS TO PHYSICAL ACOUSTICAL PROPERTIES
In this section, the results of a series of simulations are presented in the form of color-filled contour plots. This is accomplished by illustrating the sound energy propagation as simulated for a particular environment (room) for various sound source, receiver, and occluder (edge) configurations and providing comparisons to theoretical results. The simulations were conducted in a simulated enclosure (room) of 10.00m × 8.00m × 10.00m (Figure 2(a)). The frequency of the sound source was 250Hz (λ = 1.37m), and sound source level was 90dB. The sound source was positioned at location (0.69m, 4.00m, 4.83m) and remained stationary throughout all scenarios considered while the position of the receiver was varied across the x-z plane (e.g., y-axis remained constant at y = 4.00m) in equal increments equal to λ/2 or 0.685m (the x coordinate ranged from 1.37m to 8.93m while the z coordinate ranged from 0.35m to 8.97m). A flat surface (occluding plane) 3.50m × 5.00m was positioned such that it formed a plane along the "y-z" axis (e.g., constant x). The coordinates of the vertices comprising the edge were (3.45m, 0.00m, 3.45m), (3.45m, 5.00m, 3.45m), (3.45m, 0.00m, 6.45m), and (3.45m, 0.00m, 6.45m). In all scenarios considered, unless stated otherwise, the number of sonels emitted in both the sonel tracing and acoustical rendering stages was 30,000.

Open Environment
In this scenario, the occluder was absent and the absorption coefficient (α) of each of the six surfaces comprising the room was assigned a value of one (e.g., α = 1). Since there was no occluder present, there was a direct path between the sound source and each of the receiver positions thus only direct sound could reach the receiver. The resulting contour plot is shown in Figure 2(b) where, receiver level (dB) is given as a function of position across the plane of constant y. Receiver level decreases with increasing distance from the sound source. In Figure 2(c), the energy (W/m^2) across a line of positions along the x-axis of constant z (z = 4.80m) is shown (dashed blue line) illustrating the exponential decrease in energy with increasing sound source distance (energy as opposed to level is shown to illustrate the exponential decrease in energy with increasing sound source distance). Included in Figure 2(c) is the plot of the actual (theoretical) results of the attenuation of sound energy due to absorption by the air (solid red line). The difference between the results of the simulated and actual measurements are not significantly different from each other (T value 0.28 P = 0.78 Degrees of Freedom 22).


| (a) Room set-up. | (b) Contour plot. | (b) Comparison to theoretical results. |
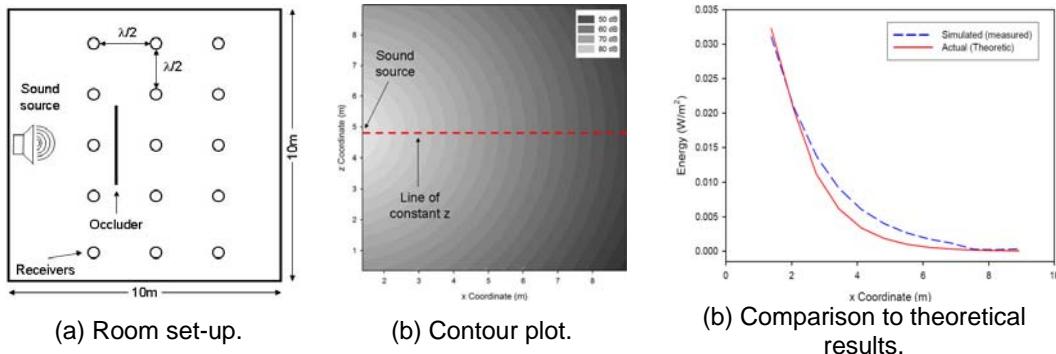
Figure 2. Results of the open environment simulation.

4

Energy Propagation in the Presence of an Occluder without Diffraction
This simulation is similar to the simulation described above except for the presence of the occluder. As with the six surfaces comprising the room, the absorption coefficient of the surfaces of the occluder (both sides) was assigned a value of one (e.g., $\alpha = 1$). The resulting plot of receiver level (dB) as a function of position across the plane of constant y is shown in Figure 3(a).
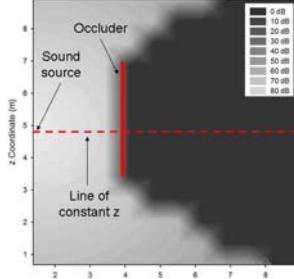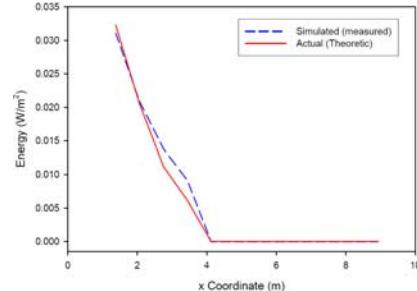
| (a) Contour plot.	(b) Comparison to theoretical results.

Figure 3. Results of the energy propagation in the presence of an occluder without diffraction.

As Figure 3(a) illustrates, no sound energy could reach any receiver whose position was such that the occluder obstructed the direct path to the sound source. As a result, the sound level at such receiver positions was zero or in other words, these receivers were within the "shadow region" (regions of black on the plot). In Figure 3(b), the energy (W/m^2) across a line of positions along the x-axis of constant z (z = 4.80m) is shown (dashed blue line) illustrating the decrease of energy in the shadow region due to the presence of the occluder and hence the occlusion of the direct sound. Included in Figure 3(b) is the plot of the actual (theoretical) results of the attenuation of sound energy due to absorption by the air (solid red line). In this scenario, given the presence of the (ideal) occluder, any measurements whose x-coordinate was greater than 3.45m (e.g., greater than the x coordinate of the occluder) resulted in an energy measurement of zero. The difference between the results of the simulated and actual measurements are not significantly different from each other (T value 0.90 P = 0.93 Degrees of Freedom 22).

Energy Propagation in the Presence of an Occluder with Diffraction
In contrast to the previous simulation, here diffraction was permitted off of the occluder. The absorption coefficient of each of the six surfaces comprising the room in addition to the surfaces comprising the occluder was assigned a value of one (e.g., $\alpha = 1$). The resulting plot (receiver level (dB) as a function of position across the plane of constant y) is shown in Figure 4(a).
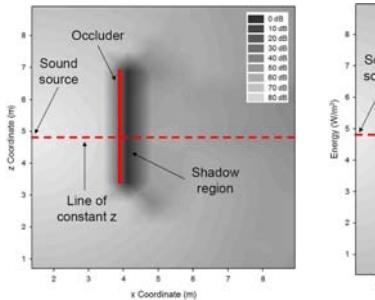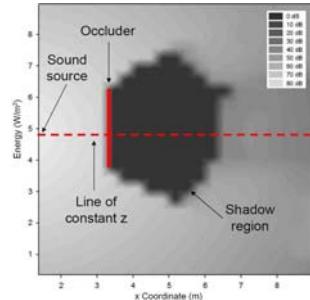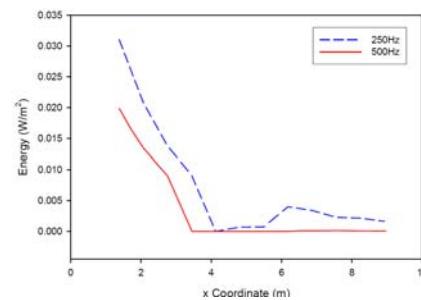
| (a) 250Hz.	(b) 500Hz.	(c) Comparison to theoretical results.

Figure 4. Results of the energy propagation in the presence of an occluder with diffraction simulation.

To illustrate the inverse relationship between diffraction and sound frequency, this demonstration was repeated for a sound source frequency of 500Hz. The theoretical diffraction model dictates an inverse relationship between frequency and diffraction. Given a greater sound source frequency, diffraction effects should be smaller or in other words, the shadow

zone should greater.   In addition to increasing the frequency of the sound source, the dimensions of the occluder (across the z-axis) were also decreased from 3.5m to 2.5m to further illustrate the inverse relationship between frequency and diffraction. The resulting plot (receiver level (dB) as a function of position across the plane of constant y) is shown in Figure 4(b). The "shadow region" (250Hz) found in Figure 4(a) although still present, is much smaller that the "shadow region" (500Hz) in Figure 4(b). The energy of a line of receiver positions across the x-axis of constant z for the energy propagation in the presence of an edge and in the presence of diffraction simulation for both the 250Hz and 500Hz frequencies are illustrated in Figure 4(c).
.

CONCLUSIONS

This paper presented the sonel mapping algorithm. Sonel mapping overcomes many of the fundamental problems associated with deterministic approaches (e.g., exponential running times) by employing Monte-Carlo methods. Instead of relying on a deterministic approach, sonel mapping employs a Russian roulette approach to determine which type of interaction does occur at each sonel/surface interaction point. Using a Russian roulette approach, a single interaction occurs at each sonel/surface interaction point as opposed to multiple interactions inherent with many deterministic approaches. Sonel mapping enjoys various advantages over deterministic based techniques, including i) the ability to handle arbitrary geometry, ii) low memory consumption, iii) the ability to handle procedural geometry, iv) the ability to handle any type of reflection model, and v) it does not require a pre-computation of the representation for the solution. Furthermore, the use of a Russian roulette approach allows for the possibility of exploring arbitrarily long paths that may not necessarily be explored using deterministic approaches while eliminating the potential exponential running times inherent in many deterministic approaches.   Moreover, with Russian roulette, the accuracy of the simulation can be improved by increasing the number of samples initially emitted from the sound source. Although this leads to an increase in computation time, an efficiency vs. accuracy trade-off can nevertheless be made.   Diffraction effects are approximated in a very simple and efficient manner allowing computation at interactive rates using a modified version of the Huygens-Fresnel principle. Future work includes experiments with human subjects.

REFERENCES:

[1] J. B. Allen and D. A. Berkley. Image method for efficiently simulating small-room acoustics. J Acoust Soc Am, 1979 **65(4)**, 943-950.

[2] M. Bertram, E. Deines, J. Mohring, J. Jegorovs and H. Hagen. Phonon tracing for auralization and visualization of sound. *IEEE Visualization 2005*, Minneapolis, MN USA.

[3] P. T. Calamia, U. P. Svensson and T. A. Funkhouser. Integration of edge-diffraction calculations and geometrical-acoustics modeling. *Forum Acusticum 2005*, Budapest, Hungary, August 29-September 2 2005.

[4] P. T. Calamia and U. P. Svensson. Fast time-domain edge-diffraction calculations for interactive acoustic simulations. EURASIP Journal on Applied Signal Processing, Special Issue on Spatial Sound and Virtual Acoustics, 2007. Article ID **63560**, 10 pages.

[5] S. Carlile. *Virtual Auditory Space: Generation and Application*. R. G. Landes Company,1996, TX USA.

[6] T. Funkhouser, N. Tsingos, I. Carlbom, G. Elko, G. Pingali, M. Sondhi and J. E. West. A beam tracing approach to acoustic modeling for interactive virtual environments. *SIGGRAPH 1998*, July 19-24 1998 FL USA.

[7] E. Hecht. *Optics*. Pearson Education Inc. 4th edition, 2002. San Francisco, CA. USA.

[8] H. W. Jensen. *Realistic Image Synthesis Using Photon Mapping*. A. K. Peters, 2001, Natick, MA USA.

[9] B. Kapralos, M. Jenkin and E. Milios. Diffraction modeling for interactive virtual acoustical environment. *GRAPP 2007*, Barcelona, Spain, March 8-11 2007.

[10] B. Kapralos, M. Jenkin and E. Milios. Sonel mapping: A stochastic acoustical modeling system. *IEEE ICASSP 2006*, Toulouse, France, May14-19 2007.

[11] B. Kapralos, M. Jenkin, and E. Milios. Acoustical modeling using a Russian roulette strategy *118th AES Convention*, 2005, Barcelona, Spain, May 28-31 2005.

[12] A. Krokstad and S. Strom and S. Sorsdal. Calculating the acoustical room response by the use of a ray tracing technique. J. Sound Vib, 1968, **8(1)**, 118-125.

[13] K. H. Kuttruff. Auralization of impulse responses modeled on the basis of ray-tracing results. J Audio Eng Soc, 1993, **41(11)**, 876-880.

[14] L. Savioja. Modeling Techniques for Virtual Acoustics. PhD Dissertation, Helsinki University of Technology, Telecommunications Software and Multimedia Laboratory, Helsinki, Finland, 1999,

[15] U. P. Svensson and U. R. Kristiansen. Computational modeling and simulation of acoustic spaces". 22nd *Int. Conf. on Virtual, Synthetic and Entertainment Audio*, June 15-17 2002, Espoo, Finland, 11-30.

[16] N. Tsingos, T. Funkhouser, A. Ngan and I. Carlbom". Modeling acoustics in virtual environments using the uniform theory of diffraction. *SIGGRAPH 2001*, August 12-17 2001, Los Angeles, CA USA