SONEL MAPPING: A STOCHASTIC ACOUSTICAL MODELING SYSTEM

B. Kapralos[1,3], M. Jenkin[1,3] and E. Milios[2,3]

[1] Computer Science and Engineering, York University, Toronto, Ontario, Canada. M3J 1P3
[2] Faculty of Computer Science, Dalhousie University, Halifax, Nova Scotia, Canada. B3H 1W5
[3] Centre for Vision Research, York University, Toronto, Ontario, Canada. M3J 1P3

{billk, jenkin}@cs.yorku.ca eem@cs.dal.ca

ABSTRACT

Modeling the acoustics of an environment is a complex and challenging task. Here we describe the *sonel mapping* approach to acoustical rendering. Sonel mapping is a Monte-Carlo-based approach to modeling diffuse, specular, absorption and diffraction effects in an efficient manner. The approach models many of the subtle interaction effects required for realistic acoustical modeling, and the approach is computationally efficient allowing it to be used to acoustically model interactive virtual environments.

1. INTRODUCTION

Accurate acoustical modeling of even small, simple environments is a complex, computationally expensive and time consuming task for all but the simplest environments. As a result, accurately modeling of sound as it propagates through the environment is extremely difficult and beyond current analytical and computational reach except for certain simple scenarios. This is similar to the rendering problem in computer graphics, where fully modeling the interaction of light and the environment is complex and computationally expensive. Advances in computer graphics have resulted in sophisticated and efficient approximations to the task of modeling illumination effects. Can similar approaches be used for acoustical modeling? Although there are several key differences between light and sound waves, there are also several similarities. By accounting for the differences between the propagation of sound and light as well as the differences in how propagating sound waves interact when they encounter objects in the environment, an acoustical modeling method inspired by photon mapping [1] and the Huygens-Fresnel principle [2], has been developed. The method is termed *sonel mapping* and its goal is to model the propagation of sound within an environment, taking into consideration both specular and diffuse reflections, refraction, absorption and diffraction in an efficient manner.

The reminder of this paper is organized as follows. In Section 2 a brief overview of photon mapping is provided. The framework of the sonel mapping method is described in detail in Section 3. Experimental results are presented in Section 4 while a summary and future research directions are provided in Section 5.

2. PHOTON MAPPING

Photon mapping is a two-pass, Monte-Carlo particle-based, global illumination method developed by Jensen in 1995 [3]. Photon mapping is preferred over finite element techniques such as radiosity [4] for a variety of reasons (see [3]). Most importantly, the approach is independent of the scene geometry thereby allowing for the illumination of arbitrarily complex scenes to be computed. In addition, photon mapping relies on stochastic techniques such as Monte-Carlo integration methods and therefore, the solution can be made more accurate by increasing the number of samples at various points of the computation.

In the first pass of photon mapping, photons (the basic quantity of light) are emitted from each light source and traced through the scene until they interact with a surface. When photons encounter a diffuse surface, they are stored in a structure called a *photon map*. In the second stage, the scene is rendered using the information provided by the previously collected photon map to provide a quick estimate of the diffuse reflected illumination. Distribution ray tracing is employed to model specular and caustic effects [1].

3. THE SONEL MAPPING METHOD

Sonel mapping is the application of the photon mapping technique to modeling the acoustics of an environment and in particular, to the task of estimating the room impulse response. Sonel mapping uses the same basic approach as photon mapping but applies the technique to the task of acoustical modeling, taking into account the physical attributes of sound propagation, addressing the possible interactions when a propagating sound encounters a surface/object or obstruction in its path (e.g., specular or diffuse reflection, diffraction or absorption). Following the same strategy as used in photon mapping, rather than modeling the exact *mechanical wave* phenomena

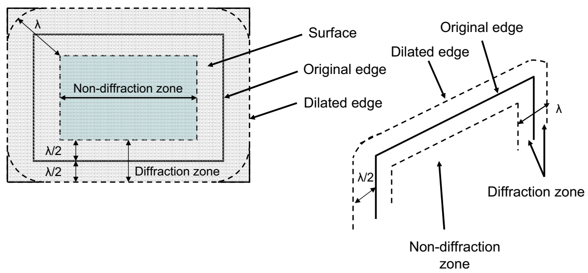

Fig. 1. Defining the diffraction and non-diffraction zones.

of sound propagation (e.g., particles in the medium as 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 sonel 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 traveled 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 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 will be reflected either specularly or diffusely or absorbed by the surface, the decision being made using a Russian Roulette strategy. 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.

In the second stage (the acoustic rendering stage), the room impulse response is estimated through the use of the previously constructed sonel map coupled with distribution ray tracing. A number of "visibility" rays are traced from the receiver into the scene where they may interact with any surfaces/objects they may encounter. 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. When a ray intersects a diffuse surface at point p, 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 impulse response. Specular reflections are handled using the same approach as in the sonel tracing stage whereby ideal specular reflections are assumed (e.g., angle of reflection is equal to the angle of incidence with respect to the surface normal). When a 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 impulse response. Diffraction effects that occur when an visibility ray encounters and edge are handled using a modified version of the Huygens-Fresnel principle. 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.

3.1. Determining the type of interaction

When a sonel encounters (is incident upon) a surface, its perpendicular distance from each of the dilated edges (see Figure 1) on the surface is calculated. If this distance is greater than λ for each edge then the sonel is within the *non-diffraction zone* and will be either completely absorbed, reflected specularly or reflected diffusely. Which of these three interactions actually occurs is determined stochastically using a Russian Roulette strategy [5]. These decisions are collectively decided based on the value of a uniformly distributed random number $\xi \,\epsilon\, [0 \ldots 1]$ as follows:

$$\begin{aligned} \xi \,\epsilon\, [0 \ldots \alpha] &\quad\rightarrow\quad \text{absorption} \\ \xi \,\epsilon\, (\alpha \ldots (\alpha + \delta)] &\quad\rightarrow\quad \text{diffuse reflection} \\ \xi \,\epsilon\, ((\alpha + \delta) \ldots 1] &\quad\rightarrow\quad \text{specular reflection} \end{aligned}$$

where, α and δ are the absorption and diffuse surface coefficients respectively. In both the sonel tracing and acoustical rendering stages, in the event of absorption, (e.g., $\xi \,\epsilon\, [0 \ldots \alpha]$) the sonel will be absorbed and tracing of the sonel is terminated. Russian Roulette provides a comparable solution yet at a fraction of the computation time and sonel cost when considering the early portion of the impulse response [5]. Russian Roulette also allows for the possibility of exploring arbitrarily long paths that may not necessarily be explored when using other deterministic termination criteria. In addition, with

Russian Roulette, the accuracy of the solution can be improved by increasing the number of sonels 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.

When the distance between the incident sonel and one of the dilated surface edges is less than λ, then the sonel is within the *diffraction zone* and will therefore be diffracted off the corresponding edge (if the distance between the incident sonel and more than one edge is less than λ, the sonel will be diffracted off of the edge to which it is closest).

3.1.1. Specular reflection

When the reflection is specular, (e.g., $\xi \epsilon ((\alpha + \delta) \ldots 1])$, ideal specular reflection is assumed whereby, with respect to the surface normal vector, the angle of reflectance is equal to the angle of incidence.

3.1.2. Diffuse reflection

When the reflection is diffuse, (e.g., $\xi \epsilon (\alpha \ldots (\alpha + \delta_a)])$, in stage one, 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. A similar approach is taken in the rendering stage except that in the event of a diffuse reflection, the sonel map is used to provide an estimate (using density estimation techniques).

3.1.3. Edge diffraction

Diffraction is modeled using a modified version of the Huygens-Fresnel principle [2]. The Huygens-Fresnel principle, states that every point on the primary wavefront can be thought of as a continuous, direction dependent emitter of secondary wavelets (sources) and these secondary wavelets combine to produce a new wavefront in the direction of propagation [2]. Given a sound source and receiver 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* [2]. The boundary of the i^{th} Fresnel zone corresponds to the intersection of the wavefront with a sphere of radius $r_0 + i\lambda/2$ 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. 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 form each zone. This is approximately equal to one half of the contribution of the first zone E_1 (e.g., $E_t \approx |E_1|/2$) [2].

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 $v_1 = n_{vis}/N_{vis}$

In order to determine the energy arriving at the receiver from the i^{th} Fresnel zone (including the first zone), the position of one of the secondary sources within the i^{th} zone is required in order to calculate the remaining sampled positions within the zone. As previously described, a sonel will be diffracted if it falls within the diffraction zone and therefore, may not necessarily be incident on the edge itself. However, whether or not the sonel is actually incident on the edge itself or close to the edge, if the sonel is to be diffracted, its position will be projected on to the edge (p_{edge}). Since the position of both the sound source and p_{edge} are known, the distance between them r_{se} can be determined. The radius of the primary wavefront is then set to this distance (e.g., $\rho = r_{se}$). Being on the edge itself, p_{edge} will be located on the surface of the sphere representing the initial wavefront and is assumed to be the position of one of the secondary sources in this particular Fresnel zone (Z_{init}). Although only the first zone is of interest, given the position of a secondary source in any other zone, simple geometry allows the position of a secondary source in the first zone to be easily determined [6]. Once the position of a secondary source in the first Fresnel zone Z_1 has been determined, the total energy E_t reaching the receiver from the first zone can be determined by scaling the uuobstructed energy with the visibility factor $E_t = v_1 \times |E_1|/2$ [6].

4. EXPERIMENTAL RESULTS

To test the ability of the sonel mapping algorithm to model the acoustics of an environment, the quasi-cubic enclosure illustrated in Figure 2 was simulated for various frequency values (125Hz, 250Hz, 500Hz, 1000Hz, 2000Hz and 4000Hz) and for various number of sonels emitted from the sound source ($10,000$, $100,000$, $500,000$, $1,000,000$ and $2,000,000$) while measuring the total sound level arriving at the receiver over a brief time interval and recording the time taken to compute the simulation. The dimensions of the room were 70m $\times$ 15m $\times$ 70m while the position (x, y, z coordinates, in meters) of the single omni-directional sound source and single receiver were $(15, 10, 55)$ and $(60, 9, 60)$ respectively. Sound source level was 90dB and its energy was divided equally amongst all sonels. The surfaces of the enclosure (four walls,

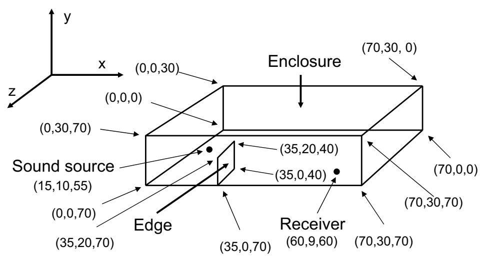

Fig. 2. Experimental set-up of room enclosure.

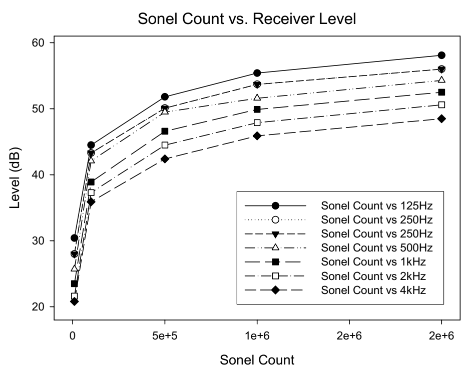

Fig. 3. Sonel count vs. receiver level.

ceiling and floor) were each assigned an absorption coefficient value of $\alpha = 0.15$. The diffuse and specular coefficients were each set to a value of $(1 - \alpha)/2$.

The simulations were performed using a Linux-based PC with a Pentium III 500MHz processor and 512Mb RAM. A summary of the experimental results are displayed in Figures 3 and 4. In Figure 3, a plot of sonel count vs. sound level at the receiver (computed over a 3s time interval) is shown for each for the frequencies considered. Generally, a decrease in level is observed as frequency is increased (and hence wavelength decreases). This is to be expected given the inverse relationship between wavelength and diffraction (in the sonel mapping method, as wavelength is decreased, the surface diffraction zone is decreased and therefore, the likelihood of diffraction also decreases). In addition, an increase in sound level is also observed with increasing sonel count (that does begin to level off). This too is also expected given that the likelihood of a sonel interacting with a receiver as the number of propagating sonels is increased. However, increasing the sonel count leads to a direct increase in the computation time. This is illustrated in Figure 4 where sonel count is plotted against simulation time (with error bars). Simulation time for each sonel count was averaged across each frequency channel. As shown, there is a linear direct relationship between sonel count and simulation time. With the computer used for this simulation, the average time to emit and trace one sonel is 0.13ms.

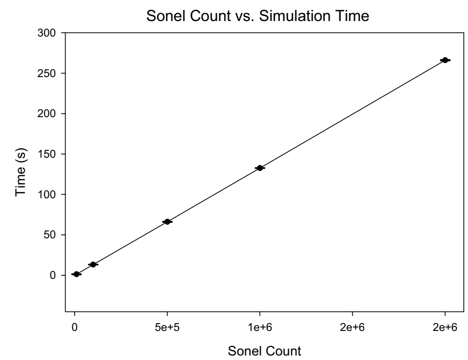

Fig. 4. Sonel count vs. average simulation time.

5. SUMMARY AND CONCLUSIONS

Sonel mapping is a two-pass stochastic, particle-based acoustical modeling method inspired by computer graphics and optics based techniques. Given its stochastic nature (and in particular Russian Roulette), it allows for the possibility of exploring arbitrarily long paths that may not necessarily be explored using other, deterministic approaches. In addition, 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 will lead to an increase in computation time, an efficiency vs. accuracy trade-off can nevertheless be made. Preliminary results based on the simulation of a quasi-cube enclosure with a diffracting edge for several frequencies and various sonel counts, confirm this. On average, it took 0.13ms to emit and trace a single sonel using a Pentium III, 500MHz computer. It is anticipated that performance will increase as machine performance is improved. Future work includes greater, more extensive experimental verification, including tests with human subjects whereby the efficiency vs. accuracy trade-off can be examined in order to determine just how many sonels are required to simulate a particular environment.

6. REFERENCES

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

[2] E. Hecht, *Optics*, Pearson Education Inc., San Francisco, CA. USA, fourth edition, 2002.

[3] H. W. Jensen and N. J. Christensen, “Photon maps in bidirectional monte carlo ray tracing of complex objects,” *Computers and Graphics*, vol. 19, pp. 215–224, 1995.

[4] M. F. Cohen and J. R. Wallace, *Radiosity and Realistic Image Synthesis*, Morgan Kaufmann Publishers, Inc., San Francisco, CA. USA, 1993.

[5] B. Kapralos, M. Jenkin, and E. Milios, “Acoustical modeling using a Russian Roulette strategy,” in *Proc. 118th Conv. of the AES*, Barcelona, Spain, May 28-31 2005.

[6] B. Kapralos, M. Jenkin, and E. Milios, “Acoustical diffraction modeling utilizing the Huygens-Fresnel principle,” in *Proc. IEEE HAVE 2005*, Ottawa, Ontario, Canada, 2005.