Audio Engineering Society

Convention Paper

Presented at the 118th Convention
2005 May 28–31 Barcelona, Spain

Acoustical Modeling Using a Russian Roulette Strategy*

Bill Kapralos[1,3], Michael R. Jenkin[1,3], Evangelos Milios[2,3]

[1] *Dept. of Computer Science and Engineering, York University. 4700 Keele St. Toronto, Canada. M3J 1P3*

[2] *Faculty of Computer Science, Dalhousie University. 6050 University Avenue, Halifax, Canada. B3H 1W5*

[3] *Centre for Vision Research, York University. 4700 Keele St. Toronto, Canada. M3J 1P3*

Correspondence should be addressed to Bill Kapralos (`billk@cs.yorku.ca`)

ABSTRACT

One of the problems with geometric (ray) based acoustical modeling approaches is handling the potentially large number of interactions between a propagating sound ray and objects/surfaces it may encounter. A sound ray incident on a surface may be absorbed, reflected both specularly and diffusely, be refracted and diffracted. Typical solutions to modeling such effects include emitting several "new" rays at each interaction point. Such solutions are computationally expensive for all but very simple environments. Rather than using such deterministic strategies and following these generated rays until they leave the environment or become sufficiently reduced in power that they no longer contribute to the acoustical landscape, probabilistic techniques such as the Russian Roulette strategy can be applied instead. Russian Roulette ensures the path length of each acoustic ray is kept at a manageable size yet allows for paths of arbitrary size to be explored. Here we describe the application of a Russian Roulette approach to acoustic modeling. Experimental results are presented that demonstrate the ability of Russian Roulette to provide a computationally reasonable solution to room acoustical modeling.

1. INTRODUCTION

Many available acoustic modeling systems are based on geometric acoustics and therefore assume that sound behaves as rays. Such systems employ suitably modified light (visual) based modeling techniques to model the acoustics of an environment.

*The financial support of NSERC, NCE IRIS, CRESTech and a York University President's Dissertation Scholarship to B. Kapralos is gratefully acknowledged. The authors also acknowledge the feedback from Konstantinos Derpanis and Erich Leung.

Several of the more popular ray-based methods include *image sources* [1], *ray tracing* [9] and *beam tracing* [5]. One of the problems associated with such 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 be reflected both specularly and diffusely, be refracted and diffracted. Solutions to modeling such effects include the generation and emission of several "new" rays at each interaction point. These approaches can lead to exponential running times and the algorithm can quickly become computationally intractable for all but simple environments.

Another problem associated with such approaches is the determination of how to terminate a sound ray. One of the simplest solutions involves keeping track of the number of times a ray has been reflected (the "reflection count") and terminating the ray once its reflection count has exceeded some predefined threshold value. Another approach is to terminate the ray based on its energy content. This second approach is more representative of the real world whereby the termination of a sound is determined by the amount of energy it has lost and not the number of times it has been reflected [15]. One measure of energy attenuation is the energy discontinuity percentage (EDP). The EDP represents the percentage of the original ray energy that must be lost before the ray is terminated [4]. Regardless of whether a reflection count or an EDP criterion is used to terminate an acoustic ray, assuming specular and diffuse reflections only, upon encountering a surface three types of interaction may occur. A portion of the ray's energy may be absorbed by the surface, a portion reflected specularly and a portion reflected diffusely according to the following constraint [3]:

$$\alpha + \delta(1-\alpha) + (1-\delta)(1-\alpha) = 1 \qquad (1)$$

where α is the incident surface absorption coefficient indicating the fraction of sound energy absorbed by the surface, δ is the incident surface diffuse reflection coefficient indicating the fraction of sound energy reflected diffusely and $(1-\delta)(1-\alpha)$ represents the amount of energy reflected specularly. Hence, at each point of incidence (provided $0 < \alpha < 1$ and $0 < \delta < 1$), two new rays are created; one that will be reflected specularly and the other diffusely. As a result, after M interactions, a total of 2^M rays may be generated. Such an approach is clearly impractical for real-time purposes except perhaps in some very simple environments. Rather than using such deterministic approaches to determine the type of interaction between an acoustic ray and an incident surface, probabilistic approaches, such as a Russian Roulette strategy [6], can be used instead. A Russian Roulette strategy ensures that the path length of each acoustic ray is maintained at a manageable size yet due to its probabilistic nature, allows for paths of an arbitrary size to be explored. Here we describe the application of a Russian Roulette based solution to the sonel mapping acoustic modeling algorithm [8]. Experimental results are presented to illustrate the ability of a Russian Roulette strategy to provide an accurate, yet computationally reasonable solution to room acoustical modeling.

This paper is organized as follows: Section 2 provides background information. In particular, an introduction to the sonel mapping acoustical modeling method and the Russian Roulette approach is presented. The use of the Russian Roulette approach in the framework of the sonel mapping acoustical modeling method is given in Section 3. Experimental results are presented in Section 4 while a summary and future research directions are provided in Section 5.

2. BACKGROUND

2.1. Sonel Mapping

Sonel mapping [8] is an application of the photon mapping image synthesis method [7] to auralization. Sonel mapping is a two-pass "particle-based", acoustical modeling method whose 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, allowing it to be used to model the acoustics of interactive virtual environments. Work with respect to this goal is ongoing and although diffraction and refraction have yet to be fully implemented, the current system supports specular and diffuse reflections in any combination.

In the first pass (the *sonel tracing* stage), sound elements known as *sonels* are emitted from each sound source and are traced through the scene until they interact with a surface. The 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 (x,y,z coordinates), incident direction and energy in addition to information specific to sound and sound propagation, including: distance traveled and frequency. The distribution of sound frequency in a given sound source is approximated by considering the center frequency of a fixed number of frequency bands. Each sonel holds the energy contained in each frequency band and each frequency band (center frequency) is considered separately at each stage of the simulation. When a sonel encounters a diffuse surface at some point x, it is stored in a structure called a *sonel map* while a "new" sonel is generated and reflected diffusely by choosing a random direction over the hemisphere centered about point x. Upon encountering a specular surface, the sonel is reflected specularly where the angle of reflection is equal to the angle of incidence (specularly reflected sonels are not stored).

In the second stage (the *rendering* stage), the room impulse response is estimated through the use of the previously constructed sonel map coupled with acoustic distribution ray tracing. The impulse response is estimated by emitting acoustic rays from each receiver and tracing them through the scene, recording their interaction with any objects/surfaces. When the ray intersects a diffuse surface at point x, the sonel is terminated and the sonel map is used to provide an estimate of the sound energy leaving point x 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 approach described for specular reflections in stage one. However, in contrast to stage one, when (if) a sound ray encounters a sound source, its energy is scaled to account for attenuation by the medium and added to the accumulating impulse response.

2.2. The Russian Roulette Approach

Russian Roulette is a Monte Carlo (stochastic) approach initially introduced to the field of particle physics simulation [6] to terminate random paths whose contributions were estimated to be small. Arvo and Kirk [2] introduced Russian Roulette to the field of computer graphics by incorporating it into their stochastic ray tracing method as a means of terminating recursive rays. Russian Roulette ensures that the path length (reflection count) is kept at a manageable size. Yet, due to its probabilistic nature, allows for paths of an arbitrary size to be explored. With respect to image synthesis, this allows for the generation of an unbiased image. Mathematically, given an integral $I = \int f(x)dx$, an estimator I_m for the integral I and an *acceptance probability* P, with Russian Roulette, a uniformly distributed random number $\xi \,\epsilon\, [0 \ldots 1]$ is introduced to determine whether the estimator I_m is to be evaluated or not. The Russian Roulette estimate I_r can then be made as follows [7]:

$$I_r = \begin{cases} I_m/P & \text{if } \xi < P \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

The resulting Russian Roulette estimate I_r is unbiased [7]. In other words, its expected value equals the expected value of the integral I, hence, the accuracy of the Russian Roulette estimate can be increased (and thus the variance reduced) by increasing the number of samples. Despite the increase in variance in comparison to the original estimator I_m, assuming that the probability P can be computed quicker than the estimate provided by I_m, the Russian Roulette estimate I_r can be computed faster than I_m [7].

3. RUSSIAN ROULETTE AND SONEL MAPPING

As with image synthesis, and photon mapping in particular, a Russian Roulette strategy can be used in several stages of the acoustical modeling process. In the sonel mapping method, a Russian Roulette strategy is used at each sonel/surface interaction point in both the sonel tracing and the rendering stages, to decide whether the incident sonel is to be reflected or absorbed by the surface and if reflected, the type of reflection. These decisions are collectively decided based on the value of a uniformly distributed random number $\xi \,\epsilon\, [0 \ldots 1]$ as follows:

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

where, s_a and δ_a are the specular and diffuse surface coefficients respectively, computed by averaging the specular and diffuse surface coefficients of each frequency bandwidth as shown below:

$$s_a = \frac{\sum_0^{N_B-1} s_i}{N_B}, \qquad \delta_a = \frac{\sum_0^{N_B-1} \delta_i}{N_B} \tag{3}$$

where s_i and δ_i are the specular and diffuse surface coefficients respectively for the center frequency of band i and N_B is the total number of bands considered. In the sonel tracing stage, in the event of a diffuse reflection, (e.g., $\xi \,\epsilon\, [0 \ldots \delta_a]$) the sonel will be stored in the sonel map and a new sonel will be created and reflected diffusely from the interaction (intersection) point. In contrast to the reflection count or energy termination criterion, the energy of the sonel will not be attenuated to account for surface absorption. Absorption with a Russian Roulette strategy is handled when the sonel is absorbed at the surface (e.g., $\xi \,\epsilon\, ((\delta_a + s_a) \ldots 1]$). When the reflection is specular, (e.g., $\xi \,\epsilon\, (\delta_a \ldots (\delta_a + s_a)]$), a new sonel is created and reflected specularly where the angle of reflection is equal to the angle of incidence. If the sonel is absorbed, tracing of the incident sonel is terminated. A similar approach is taken in the rendering stage except that in the event of a diffuse reflection, as described in Section 2.1, the sonel map is used to provide an estimate (using density estimation techniques).

3.1. Justification for the Use of Russian Roulette

Given the slow propagation speed of sound in air, time is an important component in any acoustical modeling system [13]. Since the probability of tracing arbitrarily long paths decreases as the number of sonel/surface interactions increases, the probability that a sonel is not terminated also decreases with time. An inaccurate representation of the estimated impulse response (echogram) will result since the latter portion of the estimated impulse response will contain very few samples leading to lower estimated reverberation times than predicted by Sabine's formula (see [15]). That being said, the use of Russian Roulette in acoustical modeling must be evaluated in the context of other possible approaches.

Limitations of a Reflection Count Termination Criterion: A termination criterion based on a reflection count has its limitations as well. In particular, the reflection count must be set to a large value to ensure paths of arbitrary length are traced. Setting the reflection count to a very large value is clearly impractical due to memory and computation speed limitations (e.g., an increase in the reflection count leads to an increase in the memory and computation time requirements). As a result, when used as a termination criterion, the reflection count must be kept at a manageable level and therefore there will be paths that are not traced at all, also leading to a non-linear decay of sound pressure level over time and therefore shorter than predicted reverberation times. With a Russian Roulette strategy, although the probability of tracing a particular path of length n decreases as the number of times the sonel is reflected increases, paths of arbitrarily long lengths can nevertheless be traced given the probabilistic nature of the algorithm.

Limitations of an Energy Termination Criterion: A termination criterion based on a minimum energy content such as the EDP shares the same limitations of the reflection count termination criterion (when assuming diffuse reflections only, an EDP value can be directly converted to a corresponding reflection count (see [4])). The EDP must be set to a large enough value to ensure all paths of arbitrary length are traced. However, once again, setting the EDP to a very large value is clearly impractical due to memory and computation speed limitations. Hence, as with the use of a reflection count termination criterion, there will be paths that are not explored.

Increasing Accuracy by Increasing the Number of Sonels: Consider an environment where the surface absorption coefficient of each surface is α. Using a Russian Roulette strategy, the probability of an incident sonel being reflected ρ_{ref} (diffusely or specularly) is $\rho_{ref} = 1 - \alpha$. Hence, at each surface/sonel interaction, the incident sonel will either be reflected (either specularly or diffusely) or absorbed. Which of these two interactions will occur can be described by a *Bernoulli trial* [12] and therefore, the probability that a particular sonel will be consecutively reflected n times (P_{ref_n}) can be described mathematically as:

$$P_{ref_n} = \prod_{i=1}^{n} \rho_{ref} \quad (4)$$

This corresponds to generating a sequence of n consecutive random numbers ξ_i, with each $\xi_i \leq \rho_{ref}$. Clearly, assuming $\rho_{ref} < 1$, the probability of generating this sequence decreases as n increases. With Russian Roulette, the probability of tracing a path of length n therefore decreases as n increases. However, this can be "counter-balanced" by increasing the number of sonels emitted by a sound source. Replacing each sonel originally emitted at the sound source with M sonels, the probability of generating a path of length n becomes:

$$P_{ref_n} = M \times \prod_{i=1}^{n} \rho_{ref} \quad (5)$$

For a given probability level P_{ref_n}, it is possible to choose a value M such that the probability of generating a path of length n reaches P_{ref_n}. As previously described, as n is increased, the probability of tracing a path of length n decreases, becoming zero as n approaches infinity. With respect to any practical application, the path length n will be finite. In other words, a sonel (sound) will not propagate indefinitely but will eventually lose all its energy after a portion of it is absorbed at each reflection point and by the medium. Hence, an appropriate M can always be found to ensure the probability of tracing a sonel until its energy is negligible, exceeds a pre-defined threshold value.

Reduced Computation Time: The use of a Russian Roulette approach can lead to a reduction in the required computation time while still allowing arbitrarily long paths to be traced. An exponential increase in the number of sonels to be traced (assuming two new sonels are reflected at each interaction between a sonel and a surface) is clearly impractical for any real-time applications except perhaps in certain simple, trivial environments.

4. RESULTS

In this section, the applicability of a Russian Roulette strategy to acoustical modeling applications and in particular, the sonel mapping method, is demonstrated. Initially, using an EDP termination criterion, the reverberation time (the time required for the total energy emitted by a sound source to drop to one millionth (or 60dB) of its initial value [10]) is estimated for a simple rectangular enclosure (room) and sound source/receiver configuration, for various EDP settings as shown in Table 1 (typical EDP values range from 90 to 99% [4]). All reflections were assumed diffuse (e.g., "perfectly diffuse field" [3]) hence, at each sonel/surface interaction point, a single sonel was reflected instead of multiple sonels that would occur, for example, if both specular and diffuse reflections were permitted. The number of sonels initially emitted from the sound source during the sonel tracing stage (stage one) was constant $(10,000)$. Similarly, the number of acoustic rays emitted during the rendering stage (stage two) was also constant $(10,000/10 = 1000)$.

For each of the previously computed EDP-based reverberation times, the simulation was repeated using a Russian Roulette termination criterion also assuming diffuse reflections only. The number of sonels initially emitted from the sound source was adjusted such that the resulting Russian Roulette based reverberation time estimate was equal to the corresponding EDP based reverberation time estimate. The number of receiver rays emitted in the rendering stage was $1/10$ the number of sonels emitted in the sonel tracing stage. The difference between the time taken to compute the EDP based reverberation time estimate and the corresponding Russian Roulette based reverberation time estimate is taken as the measure of performance.

The dimensions of the room were 4m $\times$ 4m $\times$ 4m, the position (x, y, z coordinates, in meters) of the single omni-directional sound source was $(3.5, 3.5, 3.5)$ and the receiver was positioned at $(1.0, 1.0, 1.0)$. The absorption coefficient of each surface i (α_i) was set to 0.1 (a single frequency band was considered). The diffuse reflection coefficient (δ_i) of surface i was obtained as: $\delta_i = 1 - \alpha_i$. The reverberation time for this room as predicted by Sabine's formula, taking absorption by the medium into consideration is 1.03s. Reverberation times were estimated by computing a linear regression on the -5 to -35dB portion of the sound pressure decay curve [11]. The decay curve itself was obtained from the estimated impulse response using Schroeder's backwards integration method [14].

A summary of the results obtained using the EDP termination criterion is provided in Table 1. For each EDP setting, the corresponding reflection count ("Ref. Count") [4], time taken to compute the estimate ("Time") and the estimated reverberation time ("RT_{60}") are provided. The results obtained using a Russian Roulette termination criterion are summarized in Table 2 where, for each of the estimated reverberation times, the number of sonels required to compute it ("Num. Sonels"), the maximum reflection count ("Max. Ref. ") encountered by any of the emitted sonels, the time taken to compute the solution ("Time") and the percent difference ("$\%dif$") between the time taken to compute the corresponding reverberation time estimate with an EDP termination criterion t_{edp} and the time to compute the reverberation time estimate with a Russian Roulette criterion t_{rus} are listed (the 91.0 and 92.0 EDP values resulted in the same reverberation time of 0.19s). The percent difference is computed as:

$$\%dif = \frac{t_{edp} - t_{rus}}{t_{rus}} \times 100 \qquad (6)$$

where a positive difference indicates $t_{edp} > t_{rus}$ and a negative difference indicates $t_{edp} < t_{rus}$.

For all measurements, the percent difference is positive indicating the EDP based method takes more time to compute. The percent differences range from 509.79 to 3572.72, decreasing as reverberation time increases due to the greater number of samples required to compute a more accurate solution (e.g., a solution closer to the predicted reverberation time). In addition to being computationally more efficient, Russian Roulette does allow paths of greater length to be explored as opposed to an EDP based approach when considering typical EDP values. Referring to Tables 1 and 2, the maximum reflection count for each of the reverberation time estimates computed using Russian Roulette are larger than the corresponding EDP based measures.

EDP	**Ref. Count (n)**	**Time (s)**	**RT_{60}**
90.0	17	4.04	0.16
91.0	19	4.52	0.19
92.0	20	4.46	0.19
93.0	21	5.01	0.20
94.0	22	5.26	0.21
95.0	23	5.50	0.22
96.0	25	6.00	0.25
97.0	27	6.53	0.26
98.0	30	7.23	0.30
99.0	36	8.72	0.35

Table 1: Reverberation time estimates using an energy discontinuity percentage (EDP) termination criterion. The number of sonels initially emitted from the sound source during the sonel tracing stage (stage one) was constant $(10,000)$. The number of acoustic rays emitted during the rendering stage (stage two) was also constant $(10,000/10 = 1000)$.

5. SUMMARY

This paper presented incorporating the Russian Roulette strategy to the sonel mapping acoustical modeling system. For many acoustical modeling applications, and in particular ones that require real-time updating, it is the early portion of the impulse response that is of interest and actually computed while the latter portion is approximated due to practical considerations [15]. This results in a non-linear sound pressure decay curve leading to lower reverberation times than predicted by Sabine's formula [15]. Results based on a comparison between the estimated reverberation times computed using both an energy discontinuity percentage (EDP) and Russian Roulette termination criterion indicate Russian Roulette can provide a comparable solution yet at a fraction of the computation time and sonel cost when considering the early portion of the impulse response. Russian Roulette also allows for the possibility of exploring arbitrarily long paths that may not necessarily be explored with an EDP or reflection count termination approach. In addition, with Russian Roulette, the accuracy of the solution 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.

This work is ongoing and currently emphasis is

RT_{60}	Sonel Count	Max. Ref.	Time (s)	%dif
0.16	400	63	0.11	3572.72
0.19	500	81	0.14	3128.57
0.20	550	83	0.16	3031.25
0.21	600	66	0.17	2994.12
0.22	650	81	0.18	2955.56
0.25	850	73	0.26	2207.69
0.26	2000	73	0.57	1045.61
0.30	3000	85	0.85	750.59
0.35	5000	79	1.43	509.79

Table 2: Russian Roulette termination criterion to compute the corresponding reverberation times obtained using an EDP termination criterion.

on the modeling of diffraction effects through the implementation of a modified version of the Huygens-Fresnel Principle. Future work will consider the inclusion of this diffraction component to the sonel mapping Russian Roulette framework. Another approach being investigated is the combination of an EDP based approach and Russian Roulette. The EDP can be used to model the early portion of the impulse response while Russian Roulette can be used to model the latter portion thus incorporating the benefits of both approaches.

6. REFERENCES

[1] J. B. Allen and D. A. Berkley. Image method for efficiently simulating small-room acoustics. *Journal of the Acoustical Society of America*, 65(4):943–950, 1979.

[2] J. Arvo and B. Kirk. Particle transport and image synthesis. In *Proc. ACM SIGGRAPH 1990*, pages 63–66, Dallas, USA, Aug. 1990.

[3] B. Dalenbäck, M. Kleiner, and P. Svensson. A macroscopic view of diffuse reflection. *Journal of the Audio Engineering Society*, 42(10):793–806, 1994.

[4] S. M. Dance and B. M. Shield. Modelling of sound fields in enclosed spaces with absorbent room surfaces. part I: performance spaces. *Applied Acoustics*, 58:1–18, 1999.

[5] T. Funkhouser, N. Tsingos, I. Carlbom, G. Elko, M. Sondhi, J. E. West, G. Pingali, P. Min, and A. Ngan. A beam tracing method for interactive architectural acoustics. *Journal of the Acoustical Society of America*, 115(2):739–756, 2004.

[6] J. M. Hammersley and D. C. Handscomb. *Monte Carlo Methods*. Chapman and Hall, New York NY. USA, 1964.

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

[8] B. Kapralos, M. Jenkin, and E. Milios. Sonel mapping: Acoustic modeling utilizing an acoustic version of photon mapping. In *Proc. IEEE HAVE 2004*, Ottawa, Canada, Oct. 2-3 2004.

[9] A. Krokstad, S. Strom, and S. Sorsdal. Calculating the acoustical room response by the use of a ray tracing technique. *Journal of Sound and Vibration*, 8:118–125, 1968.

[10] H. Kuttruff. *Room Acoustics*. Spon Press, London, England, fourth edition, 2000.

[11] Y. W. Lam. A comparison of three different diffuse reflection modeling methods used in room acoustic computer models. *Journal of the Acoustical Society of America*, 100(4):2181–2192, 1996.

[12] A. Papoulis and U. Pillai. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill, New York, NY. USA, fourth edition, 2001.

[13] R. Rabenstein, O. Schips, and A. Stenger. Acoustic rendering of buildings. In *Proc. 5^{th} Int. Conf. on Building Simulation*, Prague, Czech Republic, Sept. 8-10 1997.

[14] M. R. Schroeder. New method for measuring reverberation time. *Journal of the Acoustical Society of America*, 37:409–412, 1965.

[15] L. N. Yang and B. M. Shield. Development of a ray tracing computer model for the prediction of the sound field in long enclosures. *Journal of Sound and Vibration*, 229(1):133–146, 2000.