INNER ROOM EXTENSION OF A GENERAL MODEL FOR SPATIAL PROCESSING OF SOUNDS

Shahrokh Yadegari
Center for Research in Computing and the Arts
California Institute of Telecommunications and Information Technology
Department of Theatre and Dance, University of California, San Diego

ABSTRACT

F. Richard Moore proposed a general model for spatial processing of sounds in 1982. [5] This model separates space in two nested areas. The outer room is an imaginary acoustic space within which the inner room (or the real performance space) is located. The inner room is denoted by the location of the speakers which simulate the sound heard in the inner room as if the speakers were "openings" connecting the inner and the outer room. The spatial impression is produced by diffusing simulated direct sound rays, early echos, and global reverbration of the sound sources as heard at each speaker location. The model does not allow sound to travel through the walls of the inner room, and thus, when sound sources are near the walls of the inner room, or travel through these walls, unexpected results may be heard in opposing speakers. Furthermore, the simulation of sound sources inside the inner room are not as convincing as simulation of sound sources outside the inner room. This paper discusses an inner room extension of the general model. This extension defines the inner room as multiple nested imaginary rooms, and it provides an improved ray intersection algorithm. It also slightly alters the algorithm by which the delay time and attenuation factors of the direct and reflected rays are calculated. This extension ameliorates a number of undesirable effects and, according to our subjective tests, provides a more convincing spatialization impression when the sound sources are inside the inner room.

1. INTRODUCTION

This paper describes an extension to F. Richard Moore's "general model for spatial processing of sound". Moore's general model simulates the most perceptibly recognized effects of room acoustic to produce the desired spatial impressions and it draws on the works of Gardner [2], Blauert [1], and Stevens [8], on psychophysics of spatial perception, and those of Schroeder [7] and Moorer [6], on simulation and the use of artificial reverberation.

The general model does not simulate spatial impressions for a specific listener position; however, it produces spatial impressions for a concert setting where audiences are located in various locations of the performance space. The algorithm separates the space in two nested rooms. The outer room is an imaginary acoustic space, inside which the inner room is located. The speakers are located on or near the perimeter of the inner room, and are considered to be "openings" connecting the inner and the outer room . No sound is to be propagated through the walls of the inner room. The sound of the imaginary outer room is heard through the sound propagated through the "openings" at the location of the speakers by which the simulated sound of the outer room is diffused inside the inner room. Thus Moore's model achieves spatial impressions which are minimally dependent on the location of the audience in the performance space.

The strength of this model is to realistically localize sound sources in the area inside of the outer room and outside of the inner room. If we were to follow the same algorithm for calculation and diffusion of direct rays and reflected rays for sources inside the room, as we did for sources outside the room, the simulated sound rays no longer would mimic a physically realistic scheme; meaning that when a source is inside the inner room, the general model cannot be applied to it. This is due to the fact that in a realistic situation, even if the speakers were openings to the outer room, the sound is no longer heard through these "openings." This situation causes the produced spatial effects simulated for sound sources inside the inner room not be as effective as those simulated for sources outside the inner room. Furthermore, since no sound travels through the walls of inner room, any time a sound source comes close to an inner wall or travels through it, undesired effects could result in the simulated sounds rays for speakers located opposite of this inner wall.

This paper discusses an extension of the general model which improves the impression of simulated spatial effects for sound sources located inside the inner room, and provides an algorithm to produce smooth curves for turning speakers on and off as the sound sources pass through the walls of the inner performance room. This extension makes two modifications to the original Moore's algorithm and provides an improved ray intersection algorithm. The modifications are as follows: 1) the inner room is defined as multiple nested imaginary rooms with imaginary speakers on their perimeter; when a source is inside the outmost inner room (the primary inner room), the largest imaginary room is chosen so that the source is outside of that room, and delay and attenuation factors are adjusted to diffuse the sound as if the sound was being

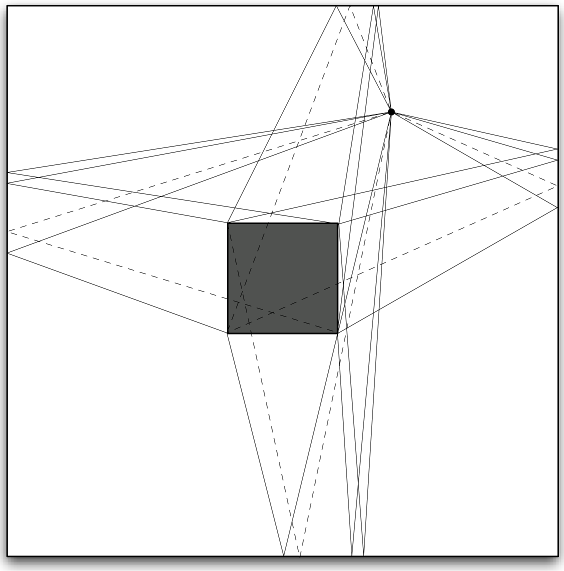

Figure 1. First order reflections off the four walls to 4 speakers located at the corners of the inner room. The paths indicated with dotted lines intersect with walls of the inner room.

propagated from imaginary speakers on the perimeter of that imaginary room 2) in contrast to the original model in which delay and attenuation factors were calculated in reference to speaker locations, in this extension delay and attenuation factors are calculated in reference to the center of the inner room.

2. MOORE'S SPATIAL ALGORITHM

Moore's model is defined in two spatial dimensions; however, the concept could easily be applied to a three dimensional space as well. Based on the location of the source and geometry of inner and outer rooms, simple ray-tracing algorithms are used to calculate the direct and reflected rays to the speaker locations. Direct paths are simply straight lines to the speaker locations. Figure 1 shows the paths for the first order reflections of a source from each surface to each speaker location (here it is assumed that the speakers are located at corners of the inner room.) Other than continuous control over the location of the source, three other parameters are defined to characterize the diffusion pattern of the sound source. Thus, the radiation vector (RV) is defined as follows:

$$RV = (x, y, \theta, amp, back), \tag{1}$$

where x and y denote the location of the source with (0,0) being at the center of the inner room, θ is the source radiation direction, amp is the amplitude of the vector, and $back$ is the relative radiation factor in the opposite direction of θ ($0 \leq back \leq 1$). $Back$ and θ are used to denote the supercardiod shape for radiation pattern of the sound source. Setting $back$ to zero denotes a strongly directional source and setting $back$ to one denotes an omnidirectional source. The following equation is used to calculate the amplitude scale factor for a simulated sound ray:

$$r(\phi) = \left[1 + \frac{(back - 1)\,|\theta - \phi|}{\pi}\right]^2 \tag{2}$$

where $r(\phi)$ is the scale factor and ϕ is the direction of the ray being simulated. Subsequently, the final attenuation factor for each simulated sound ray is calculated based on the following equations:

$$\alpha_i = \rho_i K_i B_i D_i \tag{3}$$

$$D_i = \frac{1}{d_i^{\gamma}} \tag{4}$$

where α is the total attenuation factor, ρ is the amplitude scalar determined based on the radiation pattern of the sound source and the angle by which the sound ray leaves the source (see eqn 2), K is the "cut factor" (zero if a sound ray "cut"s through a wall of inner room, and one otherwise), B acounts for absorption at reflection points, D is the attenuation factor due to the length of the path calculated based on d, the distance that the ray has to travel, and γ denotes the power law governing the relation between subjective loudness and distance. Moore points out that "these attenuation factors would in reality be frequency-dependent, but the vividness of the spatialization effect is sometimes reduced if they are implemented as such." [3]

The delay values for each simulated sound rays is calculated by the relation

$$\tau_i = \frac{R \times d_i}{c} \tag{5}$$

where τ is the delay value, R is the sampling rate in Hz, d_i is the distance between the source and a speaker, and c is the speed of sound.

Moore made a partial, though fairly complete, practical and useful, implementation of the general model in the "space unit generator" of `cmusic`. [4] This implementation used a fixed 50 millisecond fade time for turning sound rays on and off based on the result of the "cut" factor of each ray and the inner walls. As long as the source is located in a considerable distance from the inner wall, this scheme could be practical; however, if the source comes close to a wall of the inner room or if it passes through that wall, turning on and turning off speakers with a fixed 50 millisecond delay in opposite speakers to the wall would be perceived as noticeable distraction. In the next section we shall discuss the extension offered to this model for the inner room which not only prevents the undesired effects explained above, but also provides a better spatial impression when the source is located inside the inner room.

3. NEW EXTENSIONS

The extension offered in this paper for Moore's general model are in three categories: 1) an improved ray intersection algorithm, 2) definition of nested imaginary inner rooms, and 3) slightly altered delay time and attenuation factor calculations.

3.1. Improved ray Intersection algorithm

As part of a real-time implementation of Moore's general model, the author offered a simple frequency independent ray intersection algorithm for fading in/out sound rays in speakers smoothly as a sound source moves in the space. [9] Instead of producing binary "cut" factors, in this algorithm fractional "cut" factors are calculated based on a distance between the edge of an inner wall and the intersection point, a diffraction threshold, and a crossfade factor. If a ray intersects with multiple walls, the final "cut" factor is calculated as the product of the "cut" factors with each wall, according to the following relations:

$$k_{i,s} = \begin{cases} 0 & \text{when } \delta_{i,s} > \mathit{TH}, \\ \left(\dfrac{\mathit{TH} - \delta_{i,s}}{\mathit{TH}}\right)^{\mathit{CF}} & \text{when } TH > \delta_{i,s} > 0. \end{cases} \tag{6}$$

$$K_i = \prod_{s=1}^{S} k_{i,s} \tag{7}$$

where $k_{i,s}$ is the diffraction attenuation factor for ray i intersecting with surface s, $\delta_{i,s}$ is the distance between intersection point and the corner of the wall, *TH* is the diffraction threshold variable (TH could be defined as a constant or as a fraction of the size of the wall), *CF* is the crossfade exponential factor, S is the number of surfaces of the inner room, and K_i is the final "cut" factor for ray i.

3.2. Imaginary inner rooms

Moore's algorithm simulates spatial impressions based on the model that the sounds of an outer room are heard inside an inner room. Thus, the model's results are more convincing when the source is outside the inner room. In most applications it is not desired for the inner room to have a physical presence; meaning that when a sound source passes through a wall, it is usually not meant to be heard as such. Furthermore, when the source is inside the room the model can no longer be applied realistically, hence the undesirable effect of speakers in opposite side turning on and off abruptly when a sound source passes through an inner wall.

The extension offered in this paper suggests that the inner room be configured as multiple nested imaginary rooms. The specific imaginary inner room chosen for calculations is to be the largest inner room based on the location of the sound source, so that a source is always outside of an inner room, as shown in figure 2. The innermost imaginary inner room is a point at the center of the room and has dimensions of zero. Imaginary speakers are defined to be at the intersections of the lines drawn from the center of the room to each real speaker location, and the walls of the imaginary room. Attenuation factors and delay times for each ray will be calculated in relation to the chosen imaginary room. Thus, the sound diffused by each

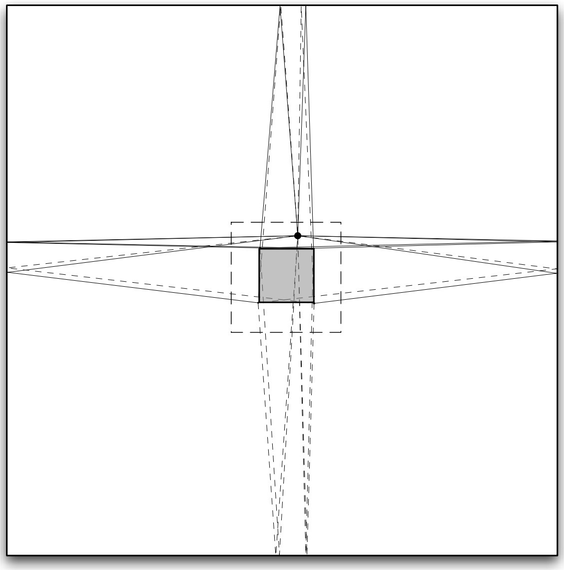

Figure 2. First order reflections off the four walls to 4 the corners of the imaginary inner room. The paths indicated with dotted lines intersect with walls of the imaginary inner room.

speaker located at the perimeter of the real inner room is attenuated and delayed in proportion to the distance between the center of the room and the location of the imaginary speaker that the real speaker is shadowing.

The diffraction threshold factor TH and the crossfade factor CF, explained in the previous section, will also be set for each imaginary room. Various implementations can offer linear or exponential scaling of TH and CF for imaginary inner rooms (e.g., the closer we get to the center of the room, the smaller the TH factor could get). Keeping TH constant for all the imaginary rooms will cause the walls of smaller imaginary rooms to be more translucent. Thus, when a source travels from outside of the main inner room towards the center of the room, the speakers located on the opposite wall gradually become louder as the source approaches the center of the room. In such a scenario, when the source gets closer the center of the inner room, the closer the "cut" factors get to one.

3.3. Delay and attenuation factor calculations

In original Moore's model delay values and their corresponding attenuation factors are calculated based on the distance between the source and speaker location. This paper is proposing that delay values be calculated based on the distance between the source and a specific speaker (which could be an imaginary one) plus the distance between that speaker and the center of the room. Accordingly the attenuation factor D_i used in equation 3 would also be calculated based on the distance between the source and the center of the room, as follows:

$$\tau_i = \frac{R \times (d_i + \lambda_j)}{c} \quad (8)$$

$$D_i = \frac{1}{(d_i + \lambda_j + \Lambda)^{\gamma}} \quad (9)$$

where τ is the delay value, R is the sampling rate in Hz, d_i is the distance between the source and the speaker on the chosen inner room, j is the speaker number of chosen inner room, λ_j is the distance between that speaker and center of the room, c is the speed of sound, Λ is a constant distant factor added to further control the attenuation factors due to distance, and γ denotes the power law governing the relation between subjective loudness and distance.

In the original Moore's model, the sound diffused by a speaker would be louder if a source were located right on that speaker than the resulting diffused sound when a source is in the middle of the room. It is so, due to the fact that the delay and attenuation factors were calculated based on the distance between the source and the speaker. This matter further complicates the simulated spatial impression of a sound source inside of the inner room. The calculation of delay times and attenuation factors in relation to the center of the room not only solves the above problem but also seamlessly accounts for the delay time simulation of imaginary speakers at the perimeter of an imaginary room by the physical speakers located at the perimeter of the primary inner room. Due to the fact that in this extended model attenuation factors are calculated in relation to the distance between the source and the center of the room, when a source is at the center of the room, all speakers are engaged with the least amount of attenuation. Thus, the resulting impression could perceptually become much stronger compared to when the source is at a considerable distance from the center. Even though γ in equation 9 could be used to control the relation between attenuation of sounds due to distance, we found the addition of Λ as a constant to be useful to balance the sound levels between the simulated impressions when the source is in the center of the room or when it is further away.

4. CONSIDERATIONS

The spatial impression produced by this model for a source located inside of the inner room, is optimal for a listener located at the center of the room. It is our understanding that by using loudspeakers, it is impossible to create the same spatial impression of a source inside the inner room for all listening locations. The model suggested in this paper is meant for performative situations. The above shortcoming could be dealt with compositionally so that the different perceived spatial impressions would carry general meaningful musical connotations.

If the speakers are not located at an equal distance to the center of the room, calculation of the delays in relation to the center of the room will cause exaggeration of any inequalities in the distance of the speakers to the center of the room. While this effect could be used musically, if the simulated impressions are to follow the original model, to use this extension, speakers should be placed in equal distance to the center of the room.

5. SUMMARY

In this paper we have proposed an extension to F. Richard Moore's general model for spatialization of sound sources. This extension produces a better spatial impression when the source is inside the inner room and ameliorates a number of undesirable effects which are produced when a source passes through the walls of the inner room. The new extension defines the inner room as multiple nested imaginary rooms. To accommodate this extension, delay and attenuation calculations were slightly altered compare to the original model, and a new configureable ray intersection algorithm was offered for turning speakers on and off.

6. ACKNOWLEDGMENTS

I would like to thank F. Richard Moore and Miller Puckette for their suggestions and contributions to this work.

7. REFERENCES

[1] Jens Blauert. *Spatial hearing : the psychophysics of human sound localization*. MIT Press, Cambridge, Mass., rev. edition, 1997.

[2] M. B. Gardner. Binaural detection of single frequency signals in presence of noise. *Journal of Acoustical Society of America*, 34:1824–1830, 1962.

[3] F. R. Moore, M. V. Mathews, and J. R. Pierce. Spatialization of sounds over loudspeakers. In *Current directions in computer music research*, pages 89–103. MIT Press, Cambridge, MA, USA, 1989.

[4] F. Richard Moore. The computer audio research laboratory at UCSD. *Computer Music Journal*, 6(1):18–29, 1982.

[5] F. Richard Moore. A general model for spatial processing of sounds. *Computer Music Journal*, 7(6):6–15, 1983.

[6] James A Moorer. About this reverberation business. *Computer Music Journal*, 3(2):13–28, 1979.

[7] M. R. Schroeder. Binaural dissimilarity and optimum ceilings for concert halls: more lateral diffusion. *Journal of Acoustical Society of America*, 65:958, 1979.

[8] S. S. Stevens. Neural events and the psychophysical law. *Science*, 170:1043, 1970.

[9] Shahrokh Yadegari, F. Richard Moore, Harry Castle, Anthony Burr, and Ted Apel. Real-time implementation of a general model for spatial processing of sounds. In *Proceedings of International Computer Music Conference*, pages 244–247, San Francisco, 2002. ICMA.