Perceptual criteria for eliminating reflectors and occluders from the rendering of environmental sound

William L. Martens and Jens Herder
Spatial Media Group
University of Aizu
Fukushima-ken 965-8085; Japan
e-mail: {wlm,herder}@u-aizu.ac.jp
www: http://www.u-aizu.ac.jp/~{wlm,herder}

Summary

Given limited computational resources available for rendering spatial sound imagery, effective means are needed for guiding selection of the components of the rendering that provide the most audible differences in the results. Rather than beginning with an analytic approach that attempts to predict audible differences on the basis of objective parameters, this project began with subjective tests of how audibly different occluded and reflected sound is when compared to unobstructed sound. Single-channel recordings of 90 short speech sounds were made in the presence and absence of these two types of obstructions under otherwise anechoic conditions. The angle of those obstructions was varied over a 90° range, and listeners were asked to rate their confidence that the recording of each of these 90 stimuli included an obstruction. In another experiment, listeners produced confidence ratings for the presence of an occluder when the stimulus presentation included simulated reverberation. Because the sound source varied from trial to trial in these experiments, confidence ratings were scattered, but clearly modulated by obstruction parameters. Although obstruction identification was already difficult under anechoic listening conditions, embedding the test signal in simulated reverberation did not reduce identification performance. Thus, at least for obstruction effects that operate within the first few milliseconds of direct sound arrival time, the obtained confidence ratings provide a basis for an evaluation function that predicts which reflectors and occluders might be most important for rendering.

Keywords: audio rendering, sound spatialization resource management, level of detail, occluder, first-order reflection, human perception

INTRODUCTION

The rendering of sound is computationally intense, especially if the goal is synthesis of realistic spatial auditory imagery for a complex environment. As computational resources are always limited, there is a need for sound spatialization resource management [1]. For some realtime auditory display applications, dynamic allocation of computational resources for synthesizing indirect sound can be quite critical. Besides the indirect sound associated with the walls of a model enclosure, spatial audio rendering for complex environments must include components for objects located within the enclosure, such as reflectors and occluders. In contrast to enclosure-related reflections that arrive at the listening position with longer delays, the contribution of these sonic obstructions to auditory spatial image formation is predominantly focussed on characteristics of the source, such as tone coloration [2].

One principle of resource management is to focus attention on the most perceptually salient components of the rendering equation, such as those that affect the tone color of direct sound. The complementary goal is to eliminate those components with consequences that are largely inaudible, such as the the impact of low-amplitude signals reflected by small obstructions at non-specular angles. The goal of the research reported here is to establish means towards perceptually-informed resource management for rendering the sonic effects of obstructions that operate within the first few milliseconds of the direct sound arrival time. Besides objects that occlude the direct sound (via interposition between sound source and sink, or receiver), objects that are near the direct sound path between source and sink produce reflections that arrive with these short delays. Both occlusion effects and these short-delay reflection effects share the characteristic that their angular incidence nearly matches that of the direct sound. Early reflections associated with low-order reflections from the walls of an enclosure arrive at the listening position with longer delays and at incidence angles not matching those of the direct sound. Their audibility is not based on changes in the tone color of the direct sound as much as it is upon changes in spatial impression, though similar issues of resource management apply here as well [3].

Two questions were addressed in the experiments reported here. The first question concerns the audibility of obstruction effects when correlated parameters such as loudness are unreliable cues to the presence or absence of an obstruction: Are obstruction effects *per se* identifiable in blind listening tests? The second question concerns the potential masking effects that subsequently arriving indirect sound might have on the audibility of obstruction effects: Are obstruction effects less identifiable when the presentation includes simulated reverberation?

METHOD

For the first two experiments, a set of 360 stimuli were constructed for use in four separate listening sessions. Each listening session comprised a set of 90 anechoically-recorded speech samples (Japanese numbers ranging from 1 to 90, spoken by one of five talkers). Within each session, the angle of the obstruction (i.e., reflector or occluder) was varied over a 90° range. The presence or absence of these obstructions was randomly varied from trial to trial, there being a 50% probability that a given trial would contain only the unobstructed direct sound. The stimuli were presented in an anechoic chamber over a small loudspeaker located directly ahead of the listener at a distance of 230 centimeters. In the third experiment, two laterally positioned loudspeakers were used to add simulated reverberation to the stimulus presentation. In all experiments, listeners were asked to rate their confidence regarding the presence or absence of an obstruction using a scale from 1 to 5.

RESULTS AND DISCUSSION

The left panel of Fig. 1 shows confidence ratings made by four listeners in the presence (circular symbols) and absence (diamond symbols) of a reflector, plotted as a function of reflector angle. A third-order polynomial was fit to the average ratings for each of these conditions, and the distance between these two curves gives an indication of how easily the sound of the reflector could be identified. Note that the highest ratings were not observed for specular reflection, but rather at an incidence angle of around 40°.

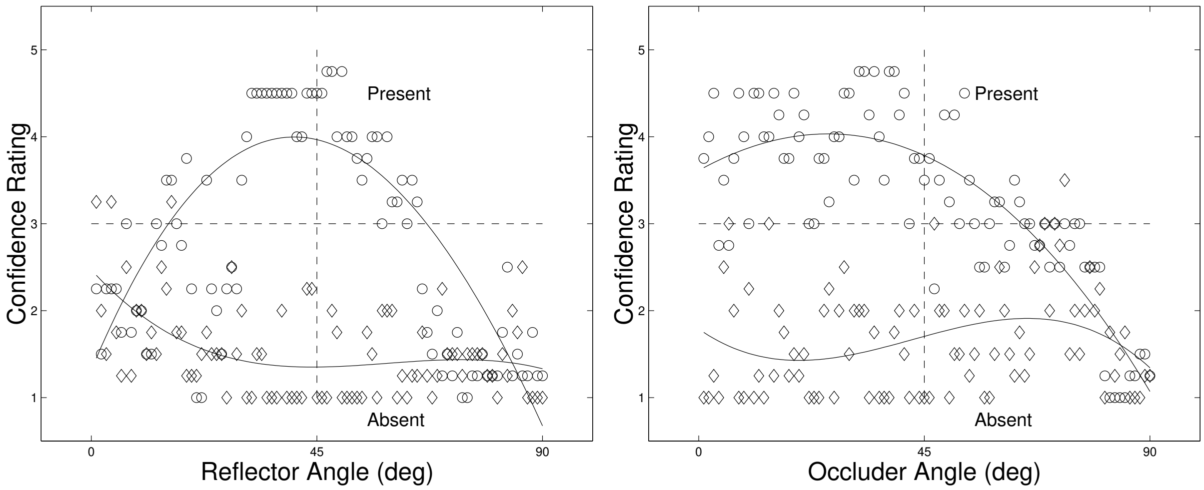

Figure 1: Identification of Reflector (left) and Occluder (right) Presence

The right panel of Fig. 1 shows confidence ratings made in the presence and absence of an occluder. However, in this case, the 0° angle corresponds to the spatial configuration for which the maximum width of the occluder is interposed between source and sink. The smooth curves fit to these data begin to converge for angles above 45°, where blocking by the occluder becomes negligible.

Though listeners learned to identify the occluder, the task required more concentrated effort that did the reflection identification task. Because the level and spectral distribution were changed for each trial, listeners could not base identification on subtle changes to a familiar sound source. For the third experiment, a unique sound source was again employed on each trial; however, only four obstruction configurations were included in the stimulus set. Here again, listeners rated their confidence that an occluder had been present or absent, but they heard the sound stimulus in three reverberant contexts: under anechoic conditions, with the addition of reverberation simulating a large room, and with reverberation simulating a small room. Fig. 2 shows the relative frequency of each response for four listeners. With no evidence of a ceiling or floor effect, it appears that identification performance is not degraded by the presence of simulated reverberation.

How can these data be used to determine which reflectors and occluders are most important to render? As stated in the introduction, the goal is to render only those that combine with the direct sound to provide an identifiably different spatial image in comparison to the image of the direct sound alone. An arbitrary threshold confidence rating of "3" (marked by the horizontal dotted lines in Fig. 1) can be set to select which obstructions should be rendered, but this approach does not consider the number of rendering elements. An alternative is to choose the obstructions for rendering from a set containing both reflectors and occluders, using the standardized confidence rating scale as a common means for setting their priority.

CONCLUSION

Auditory displays should be able to avoid using computationally-expensive extended rendering techniques when the audible difference associated with such extensions are negligible.

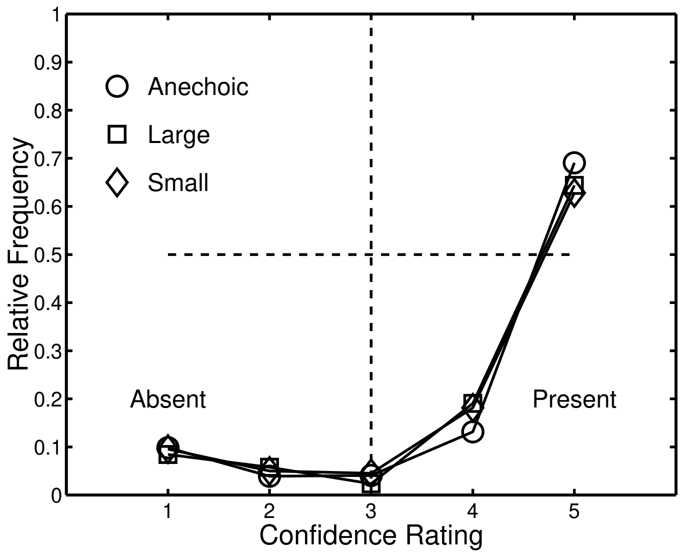

Figure 2: Identification of Occluder Presence in Reverberation

Studies such as these are useful in guiding the development of auditory display applications and in refining auditory display technology. The results of this study guided the implementation of efficient rendering [4] of obstructions for the Sound Spatialization Framework [5] developed at the University of Aizu.

REFERENCES

[1] J. Herder and M. Cohen, "Sound Spatialization Resource Management in Virtual Reality Environments," in *ASVA'97 — Int. Symp. on Simulation, Visualization and Auralization for Acoustic Research and Education*, (Tokyo, Japan), pp. 407–414, The Acoustical Society of Japan (ASJ), April 1997.

[2] M. F. E. Barron, "The subjective effects of first reflections in concert halls - the need for lateral reflections," *Journal of Sound and Vibration*, vol. 15, pp. 475–94, 1971.

[3] D. R. Begault, "Audible and inaudible early reflections: thresholds for auralization system design," in *Audio Engineering Society 100th Convention*, (Copenhagen), 1996. Preprint 4244.

[4] W. L. Martens, J. Herder, and Y. Shiba, "A filtering model for efficient rendering of the spatial image of an occluded virtual sound source," in *137th Meeting of the Acoustical Society of America and the 2nd Convention of the European Acoustics Association*, (Berlin), Acoustical Society of America (ASA), and European Acoustics Association (EAA), March 1999.

[5] J. Herder, "Sound spatialization framework: An audio toolkit for virtual environments," *Journal of the 3D-Forum Society, Japan*, vol. 12, pp. 17–22, September 1998.