A filtering model for efficient rendering of the spatial image of an occluded virtual sound source

William L. Martens, Jens Herder, and Yoshiki Shiba
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

Rendering realistic spatial sound imagery for complex virtual environments must take into account the effects of obstructions such as reflectors and occluders. It is relatively well understood how to calculate the acoustical consequence that would be observed at a given observation point when an acoustically opaque object occludes a sound source, but the interference patterns generated by occluders of various geometries and orientations relative to the virtual source and receiver are computationally intense if accurate results are required. In many applications, however, it is sufficient to create a spatial image that is recognizable by the human listener as the sound of an occluded source. In the interest of improving audio rendering efficiency, a simplified filtering model was developed and its audio output submitted to perceptual evaluation. Two perceptually salient components of occluder acoustics were identified that could be directly related to the geometry and orientation of a simple occluder. Actual occluder impulse responses measured in an anechoic chamber resembled the responses of a model incorporating only a variable-duration delay line and a first-order (one-pole, one-zero) filter. The filtering model has the attractive feature that either the reflecting or the occluding effects of an obstruction can be simulated through changes in continuously variable parameters, requiring no change in the filter structure.

Keywords: audio rendering, occluder, first-order reflection, human perception

INTRODUCTION

What does acoustical occlusion sound like? When an object is interposed between a source and a receiver, and that object is not acoustically transparent in the frequency range of interest, it typically reduces the sound pressure level (SPL) that can be measured at the receiver position. It also changes the spectral energy distribution, and usually spreads the energy over time as well. But when are the changes to the sound source audibly characteristic of the occlusion? In a related study by the authors [1], this general question was reduced to the specific experimental question, “When can a human listener confidently detect the presence of an occluder in a blind listening test?” Because the SPL at the source was varied from trial to trial, the SPL at the receiver position was an unreliable cue for the detection of occlusion. Under these circumstances, occlusion effects were somewhat difficult to identify. One goal of the research reported in this paper was to develop a filtering model that could effectively synthesize signals that would sound more clearly as if they resulted from occlusion. Another goal of this research was to develop efficient means for rendering the spatial image of an occluded virtual sound source in order to minimize computational cost while maximizing effect audibility.

In contrast to the goal of reducing rendering cost by selectively eliminating reflectors and occluders from the rendering equation [1], the motivation for this study was to simplify the rendering algorithm, and thereby allow more sonic effects to be rendered. An alternative approach to rendering the complex spectro-temporal phenomena associated with occlusion is to reduce the effects to simple changes in sound source gain. This approach was taken by Takala & Hahn [2], who set gain via a form factor proportional to the amount of occlusion. What is the rationale for such gross simplification? In many multimedia applications, as well as in conventional audio production for sound effects, the purpose of audio rendering is to illustrate an event rather than to attempt an acoustically accurate simulation [2]. Given that occlusion effects can be difficult to identify in blind listening tests [1], there seems to be some justification for such extreme rendering simplifications. On the other hand, there are changes in tone color of the direct sound that are characteristic of some occlusion effects. In contrast to highly detailed physical solutions, however, such as those based upon Fresnel zones [3], a solution was desired that was driven rather by perceptually-defined specifications. As a more detailed alternative to the simplification that reduces occlusion effects to changes in gain, this paper presents a filtering model of low computational cost that can produce distinctive auditory spatial images associated with identifiable occlusion effects.

The research project reported here consisted of three distinct phases. First, many acoustical measurements of an actual occluder were made in an anechoic chamber in order to characterize the stimulus. Second, a digital filtering model was designed to capture the variation in occluder impulse responses that was observed as the occluder size, position, and orientation were varied. Finally, a set of experimental stimuli were prepared using the completed filtering model, and listeners were asked to listen to all pairwise comparisons of the stimuli and to rate each pair on perceptual dissimilarity. The spatial imagery that results from the occlusion simulation model described herein was submitted to perceptual evaluation in order to document the variation in human response associated with variation in the parameters of the model.

ACOUSTICAL MEASUREMENTS OF OCCLUSION

The acoustical responses of obstructions of various sizes were measured at various positions and orientations relative to the location of a loudspeaker and a microphone. Fig. 1 shows the changes in the response of a small, rectangular board of 40 by 90 cm as its orientation varied in relation to a sound source and sink (receiver) that were each located 100 cm from the center of the board.

The curves in the left panel of Fig. 1 plot the normalized energy over time for the response at nine occluder angles. The 90° angle provided the maximum obstruction to sound in this set of measurements, while the 10° angle presented the smallest silhouette along the line of sight between source and sink. The vertical dashed line marks the

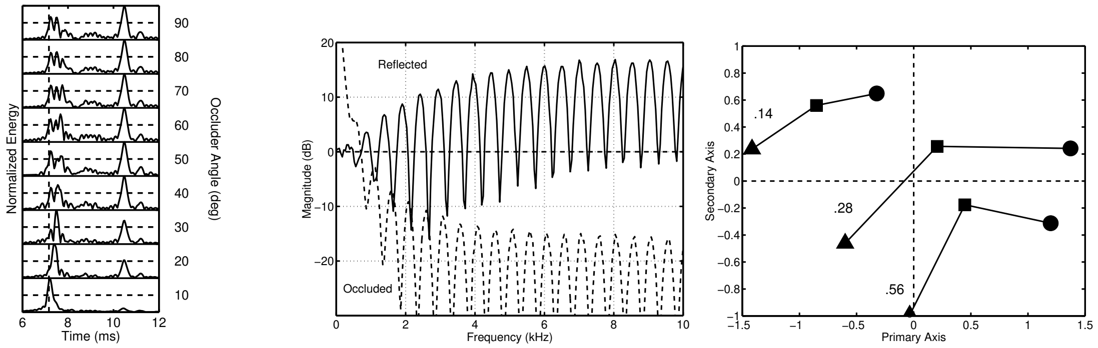

Figure 1: Time-domain responses at nine occluder angles (left panel); model magnitude response for reflected and occluded sound (center panel); MDS results for nine model outputs (right panel).

time at which the normalized energy peaks for an unobstructed sound arriving at the microphone position via the direct path from the loudspeaker. Note that the amplitude of the initial peak observed at around 7 msec decreases with occluder angle while that of the secondary peak observed at around 10.5 msec increases with occluder angle (with increasing obstruction). The absolute peak energy is also declining with elevation, though that detail is obscured in the figure by normalizing energy for each measurement according to the maximum of each curve.

The magnitude response over frequency for the nine occluder angles (not shown here) showed substantial attenuation in a broad band centered on 5 kHz, but showed virtually no attenuation at around 8 kHz. At the 10° angle, on the other hand, very little attenuation was observed at frequencies below 8 kHz. This small, rectangular occluder almost always shows a boost in gain at presents a very small silhouette along the line of sight between microphone and speaker.

FILTER DESIGN FOR OCCLUSION AND REFLECTION SIMULATION

The filtering model was designed to capture the sonic effects of a modeled obstruction, whether the sound source be reflected or occluded by that obstruction. Because obstructions located near the direct-sound path are always switching between these two states, a comprehensive solution was required. The filter operates by tapping the audio input buffer containing the sound-source audio samples either earlier or later than the tap for the direct-sound signal. If the obstruction is oriented as a reflector, then the obstruction signal is delayed relative to the direct sound, and receives a high-pass emphasis. If, however, the obstruction is oriented as an occluder, then the obstruction signal will lead the direct sound, and is given a low-frequency emphasis.

Characteristic magnitude responses that result when obstruction signals are combined with direct-sound signals are shown in the center panel of Fig. 1. For the relatively small obstruction simulated here, the upper, solid curve shows that the high-frequency portion of the direct sound is reflected at higher gain. In contrast, the lower, dashed curve shows that the low-frequency portion of the occluded direct sound is amplified in the model, somewhat exaggerating the increase in gain observed in measured responses. In both cases, the depth of the comb pattern in the magnitude response increases with increasing frequency. Attenuation of DC energy for early reflection simulation is common [4], but a single filter structure that can smoothly transform reflection effects into occlusion effects is novel and advantageous. processing

PERCEPTUAL EVALUATION OF SIMULATED OCCLUSION

In order to evaluate the spatial imagery associated with the simulated occlusion effects available using the filtering model, an exploratory listening experiment was executed. Nine "obstructed" versions of a single sound source were prepared using the filter at three delays and three low-pass settings. Listeners rated the perceptual dissimilarity of all pairwise comparisons of these sound stimuli on a 5-point scale. The right panel of Fig. 1 shows the results of multidimensional scaling (MDS) analysis of the dissimilarity data. Symbol shape codes obstruction signal delay (triangle: .5 msec; square: 1 msec; circle: 2 msec). The line segments connect symbols associated with sounds processed using identical zero frequency (expressed numerically in the plot as proportions of the Nyquist frequency). Stimuli in the lower right corner of the plot sounded darker, had more audible "combing" quality, and could be characterized as sounding "strongly occluded." Stimuli in the upper left corner of plot sounded brighter and had less "combing." These sounds did not give a strong impression of occlusion.

CONCLUSION

The filtering model described here allows for smooth transition from occluder to reflector state, easing processing control by handling the two states between which obstructions are always switching. Perceptual evaluation confirmed that the model enables the creation of a continuous range of obstruction effects that also produces satisfying auditory spatial imagery.

REFERENCES

[1] W. L. Martens and J. Herder, "Perceptual criteria for eliminating reflectors and occluders from the rendering of environmental sound," 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.

[2] T. Takala and J. Hahn, "Sound rendering," *Computer Graphics*, vol. 26, pp. 211–220, July 1992.

[3] N. Tsingos and J.-D. Gascuel, "Fast rendering of sound occlusion and diffration effects for virtual acoustic environments," *Proceedings of the Audio Engineering Society 104th Convention*, May 1998. Preprint No. 4699 (P4-7).

[4] G. S. Kendall, W. L. Martens, D. J. Freed, M. D. Ludwig, and R. W. Karstens, "Image model reverberation from recirculating delays," *Proceedings of the Audio Engineering Society 81st Convention*, 1986. Preprint No. 2408 (C-17).