Reflection properties influencing the precedence effect

Florian Wendt, Robert Höldrich
University of Music and Performing Arts Graz, Austria
Institute of Electronic Music and Acoustics, Email: {wendt, hoeldrich}@iem.at

Introduction

The precedence effect refers to a perceptual phenomenon whereby the direct sound dominates the perception. This allows us to localize sound sources in challenging reverberant environments in which the direct sound is followed by multiple reflections. Measurements on the precedence effect have traditionally been made by presenting two similar sound instances, one representing the direct sound and the other simulating a specular reflection on a large, smooth, and rigid wall, cf. [1, 2].

Very often the surface of a wall is not completely smooth but contains regular or irregular coffers, bumps or other projections. Depending on their sizes compared to the wavelength, the reflected sound field can be scattered. In this case we speak of a diffusely reflecting wall yielding a spatial and temporal widening of the reflection. Only little is known on the influence of diffuse reflections on the precedence effect and to the authors' knowledge, there is only one study considering temporal diffusion [3].

We present a model to simulate a diffuse reflection with the scattered energy assumed to be distributed according to Lambert's cosine law. Subsequently we sketch a listening experiment conducted over headphones studying the influence of reflection properties on the echo suppression, a measure of the precedence effect. After this, we discuss the results, which are modeled in the last section.

Simulating a Diffuse Reflection

In psychoacoustics studies on the precedence effect the lagging reflection is usually an exact copy of the leading direct sound. Such a *specular reflection* occurs when a plane wave is reflected on an infinitely large, smooth, and rigid wall to the specular direction with $(\theta_S, \phi_S) = (\theta_R, \phi_R)$, cf. Figure 1. The impulse response of such a reflection is independent of the constellation consisting of source S, receiver R, and wall A and can be described by a Dirac $\delta(t)$ with the delay t calculated from the distance between the corresponding image source and the receiver. A specular reflection represents one of two "extreme" conditions which have been identified for rigid surfaces [4]. The other extreme condition is a *diffuse reflection* which occurs when the reflected energy is scattered.

Let S be a point source with sound power Q, cf. Figure 1. We can determine the sound power $\mathrm{d}P_A$ reaching the element $\mathrm{d}A$ of a perfectly rigid wall by the angle of incidence θ_S and distance r_S:

$$\mathrm{d}P_A = \frac{Q\cos\theta_S}{4\pi\, r_S^2}\mathrm{d}A. \tag{1}$$

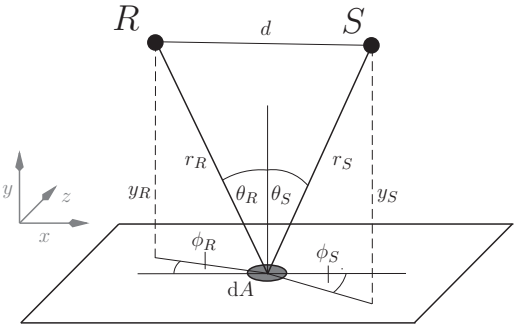

Figure 1: Schematic geometry of a reflection on the xz-plane.

For diffuse reflections the scattered sound power from the element $\mathrm{d}A$ is usually assumed to be proportional to the cosine of the angle of reflection $\cos\theta_R$ [4]. Thus, the intensity $\mathrm{d}I_A$ reaching the receiver R reflected by the wall element $\mathrm{d}A$ is defined by

$$\mathrm{d}I_A = \frac{Q\cos\theta_S\cos\theta_R}{4\pi\, r_S^2\, r_R^2}\mathrm{d}A. \tag{2}$$

Distances r_S and r_R are depending on corresponding angles θ_S and θ_R. This results in a spatial spread of diffuse intensities around the direction of a specular reflection (θ^*, ϕ^*). Neglecting all phase relations and the interference effects caused by them, the overall diffuse intensity I reaching the receiver R is obtained by the integration over the wall A yielding an elliptical integral, which we solve numerically. Figure 2 shows reflected intensities $\mathrm{d}I_A$ reaching the receiver R of a rigid wall in the xz-plane.

The temporal spread of the intensity at the receiver R is shown in Figure 3 as the energy-normalized intensity I as function of time for different constellations. The major difference of our envelopes compared to [3], in which diffuse refection responses are approximated the probability density function of a gamma distribution, is the instantaneous onset.

Experimental Setup and Conditions

Several perceptual phenomena relate to the precedence effect. The present study investigates the *echo suppression* for specular and diffuse reflections. Tested suppression thresholds were the *echo threshold*, defined as the minimum level of the lagging sound at which it was possible to detect a second auditory event, and *masked threshold*, defined as the minimum level at which it was possible to detect that a lagging sound was present at all.

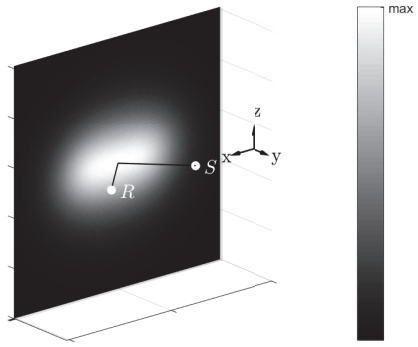

Figure 2: Diffusely reflected intensities $\mathrm{d}I_A$ coded in gray-scale of a rigid wall approached by Lambert's cosine law. The sound propagation path of a specular reflection is indicated as solid line.

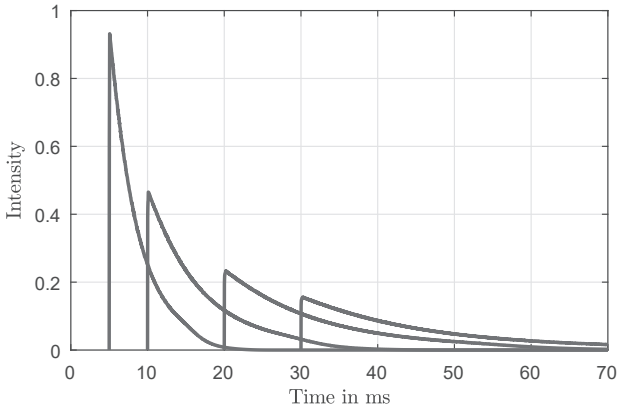

Figure 3: Energy-normalized intensities I arriving at the receiver R as function of time of all simulated conditions c.

Table 1: Constellation of simulated conditions derived from the specular delay T^* and the specular direction $(\theta^*, \phi^*) = (45°, 0°)$.

cond.	T^*	d	y_0	$A_x \times A_z$
c_5	5 ms	4.1 m	2.0 m	10 m × 8.5 m
c_{10}	10 ms	8.3 m	4.1 m	20 m × 17 m
c_{20}	20 ms	16.5 m	8.3 m	40 m × 34 m
c_{30}	30 ms	24.8 m	16.5 m	60 m × 51 m

The echo threshold was measured by a method of adjustment in which listeners were instructed to reduce the level of the lagging reflection down to the point where it was as faint as possible while still remaining audible as a second auditory event. The masked threshold was measured by a two-alternative forced choice paradigm in a 3-down 1-up adaptive procedure [5]. Listeners were instructed to decide which of the two consecutively presented stimuli included a reflection. The two stimuli could be repeated at will and feedback was given after each response. Loudness cues were removed by roving the level of both stimuli and tonal coloration is expected to be the most relevant cue.

The stimuli signals consisted of a sequence of a 50-ms-long white noise burst (instant on- and offset) followed by a 200-ms-long pause. For the masked threshold experiments, the sequence was repeated four times for each stimulus. For the echo threshold an ongoing sequence was presented until the listener logged in his or her answer, with the level faded in by 2 seconds to build up the precedence effect at the beginning of each trial. The listening experiments were conducted over headphones. Binaural signals were generated by convolving direct sound and reflection with head related transfer functions (HRTFs).

In contrast to the specular reflection, the spatial and temporal distribution of a diffuse reflection response depends on the setup of the constellation, i.e. source-to-receiver distance and respective distances to the reflecting wall. The simulated setups in the experiment consisted of a reflecting wall at the left of source and receiver, which results in delays T^* of corresponding specular reflections. Conditions c were created by choosing the distance between source and receiver d and their distance to the wall y_0 in a way, that the incoming direction of specular sound is $(\theta^*, \phi^*) = (45°, 0°)$, cf. Figure 2.

Obviously, Lambert's cosine law requires an infinite large wall. For the numerical solution, the size of the wall $A = A_x \times A_z$ was chosen in a way that approximately 95% of the overall energy is taken into account by the simulation of an uniformly sampled wall at 8500 elements $\mathrm{d}A$ with the specular direction at its center. The intensities $\mathrm{d}I_A$ were encoded into Ambisonics (order $N = 17$) and decoded using a max-$\boldsymbol{r}_\mathrm{E}$ weighting [6] resulting in 632 envelope signals describing the diffuse reflection. Obtained envelope signals were used as a model to produce a reflection that emulated the characteristics of the diffusive wall, but with a white spectrum. White Gaussian noise was multiplied with each of the 632 envelopes to generate reflection waveforms resulting in a colored reflection. An iterative whitening procedure similar to the Hilbert transform approach described in [7] was then applied to avoid coloration in the listening experiment. Table 1 lists tested conditions with specular delays T^* derived from the respective constellation. The energy-normalized intensities I arriving at the receiver R of all conditions is shown in Figure 3.

The influence of diffuse and specular reflection on the echo threshold is investigated with all conditions c listed in Table 1 and two repetitions yielding 16 adjustment tasks (8 diffuse + 8 specular) to be evaluated by each listener in random order. The masked threshold was studied with conditions $c_{(5,20)}$ and one repetition yielding 4 randomly conducted adaptive tasks. All stimuli were generated in advance by convolving the excitation signal with respective impulse responses. After this, the signals were normalized to their RMS value for level equalization and binaural stimuli were created by convolving them with corresponding HRTFs of the Neumann KU100 dummy head [8].

The listening experiment employed Matlab and PureData running on a personal computer with a M-Audio MobilePre sound card and

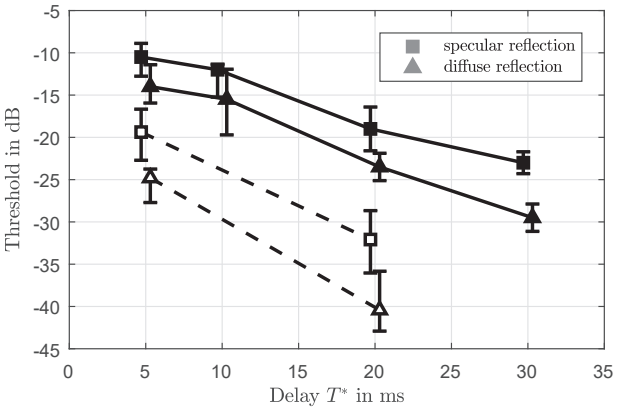

Figure 4: Means and 95% confidence intervals of echo threshold (filled symbols) and masked threshold (open symbols) for both reflection types.

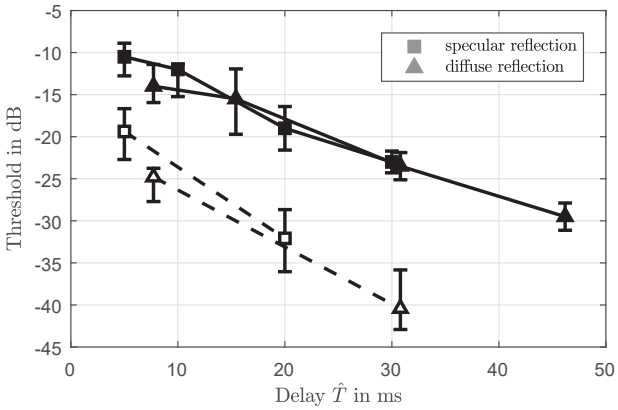

Figure 5: Means and 95% confidence intervals of echo threshold (filled symbols) and masked threshold (open symbols) for both reflection types. The data is temporally aligned to the centroid of energy of the corresponding impulse response.

Beyerdynamics DT770 headphones. The playback level was fixed at 65dB(A). Twelve listeners (all male; age 26-55 years) participated in the experiment. All of them are experienced listeners and reported normal hearing acuity.

Experimental Results

The results are given in Figure 4 as mean values by filled symbols for the echo threshold and open symbols for the masked threshold with corresponding 95% confidence intervals. Thresholds are plotted separately for the specular reflection (squares) and the diffuse reflection (triangles) as a function of the delay.
A statistical analysis of variance (ANOVA) of the data reveals both thresholds, the condition, and the reflection type to be highly significant parameters ($p \leq 0.02$). It can be seen that the suppression decreases progressively with increasing delay. Masked thresholds are lower then echo thresholds and their progression resembles those obtained by similar studies on the echo suppression [9, 10].
Interestingly, the diffuse reflection weakens the echo suppression and in each condition the diffuse threshold was lower than the corresponding specular threshold. A pairwise ANOVA of both reflections types reveals conditions $c_{(20,30)}$ of the echo threshold and both conditions of the masked threshold to be significantly different ($p \ll 0.01$). This is in contrast to findings obtained by [3], for which the reflection type was not found to be significant for most conditions. A reason therefore could be the temporal alignment of diffuse reflection responses to energy centroid in their experiments.

The difference between corresponding suppression thresholds in our experiment is not constant along condition. Linear regression lines reveal different slopes for the echo threshold (specular: $-0.53\,\mathrm{dB/ms}$, diffuse: $-0.65\,\mathrm{dB/ms}$) as well as for the masked threshold (specular: $-0.85\,\mathrm{dB/ms}$, diffuse: $-1.04\,\mathrm{dB/ms}$).
The size of 95% confidence intervals provides evidence that listeners performed similarly for both reflection types in each test and no significance is obtained. A reason for the larger confidence intervals of the masked threshold compared to the echo threshold is due the fact that conditions of the latter were tested twice.

Modeling the Echo Suppression

The question is whether the different results obtained for the specular and diffuse reflection can be explained by characteristic metrics of respective impulse responses.

A simple predictor considering only the temporal spread of the reflection response is the energy centroid. It is defined as the time $\hat{T}$ where the center of mass of the energy envelope is located. We obtain the energy centroid by

$$\hat{T} = k\,T^*, \tag{3}$$

with the constant $k = 1.54$ for diffuse reflections and $k = 1$ for specular reflections. Figure 5 shows the temporally aligned results based on the corresponding centroid of energy. It reveals that our model works qualitatively well, especially for the echo threshold at delays $T^* \geq 10\,\mathrm{ms}$. For the masked threshold, which is sampled with two conditions only, we do not see such a neat overlapping. A linear regression model reveals $R^2 = 0.98$ close to 1 for both echo and masked threshold and p-values close to 0 (echo: $p \ll 0.01$, masked: $p = 0.01$) and we conclude that a significant linear regression relationship exists between the model and our results. Optimal constants towards maximizing R^2 and minimizing the error variance are found at $k = 1.51$ for the echo threshold and $k = 1.65$ for the masked threshold.

Obviously, for specular reflections the constant $k = 1$ is independent of the constellation. For constellations with a fixed direction (θ^*, ϕ^*) like ours, the simulation of different diffuse responses reveals the diffuse constant k to be depending on the angle θ^*, e.g. $k(30°) = 1.38$, $k(15°) = 1.25$.

Conclusion

We presented a model for simulating diffuse reflections based on Lambert's cosine law. The modeled reflection yields a spatial and temporal smearing of the reflected sound field. The perception of diffuse responses were evaluated in a listening experiment on the echo suppression. In comparison to a specular reflection a diffuse reflection weakens the precedence effect: less level is necessary to hear an echo and less level is necessary to hear that a

reflection is present at all. Finally, we showed that the energy centroid of the energy envelope of diffuse responses is a suitable model for simulating their suppression. Obviously, this model considers only the temporal spread of a diffuse reflection indicating that the temporal characteristics of reflection responses play a major role in the precedence effect.

Acknowledgments

Our research was partly funded by the Austrian Science Fund (FWF) project nr. AR 328-G21, Orchestrating Space by Icosahedral Loudspeaker (OSIL). We thank all listeners taking part in the experiment.

References

[1] B. Rakerd and W. M. Hartmann, "Localization of sound in rooms, II: The effects of a single reflecting surface.," *The Journal of the Acoustical Society of America*, vol. 78, pp. 524–533, 1985.

[2] R. Litovsky and B. Shinn-Cunningham, "Investigation of the relationship among three common measures of precedence: Fusion, localization dominance, and discrimination suppression," *The Journal of the Acoustical Society of America*, vol. 109, no. 1, pp. 346–358, 2001.

[3] P. W. Robinson, A. Walther, C. Faller, and J. Braasch, "Echo thresholds for reflections from acoustically diffusive architectural surfaces," *The Journal of the Acoustical Society of America*, vol. 134, no. 4, p. 2755, 2013.

[4] H. Kuttruff, *Room Acoustics*. Taylor & Francis, 5 ed., 2009.

[5] H. Levitt, "Transformed Up-Down Methods in Psychoacoustics," *The Journal of the Acoustical Society of America*, vol. 49, no. 2B, pp. 467–477, 1971.

[6] J. Daniel, *Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia*. PhD thesis, Université Paris 6, 2001.

[7] A. Kohlrausch, R. Kortekaas, M. van der Heijden, S. van de Par, A. J. Oxenham, and D. Püschel, "Detection of Tones in Low-noise Noise: Further Evidence for the Role of Envelope Fluctuations," *Acta Acustica united with Acustica*, vol. 83, pp. 659–669, 1997.

[8] B. Bernschütz, "A Spherical Far Field HRIR/HRTF Compilation of the Neumann KU 100," *Fortschritte der Akustik – AIA-DAGA 2013*, pp. 592—-595, 2013.

[9] J. Blauert, *Spatial hearing - the psychophysics of human sound source localization*. The MIT Press, 1983.

[10] B. Rakerd, W. Hartmann, and J. Hsu, "Echo suppression in the horizontal and median saggittal planes," *Journal of the Acoustical Society of America*, vol. 107, no. 2, pp. 1061–1064, 2000.