Influences of the Floor Reflection on Auditory Distance Perception

Mike D. Ebelt, Johannes M. Arend, Christoph Pörschmann
TH Köln – University of Applied Sciences, Institute of Communications Engineering, 50679 Köln, Germany
Email: mike.ampeso@gmail.com, johannes.arend@th-koeln.de, christoph.poerschmann@th-koeln.de

Introduction

Human auditory distance perception relies on several cues, such as intensity, the direct-to-reverberant ratio, spectral, and binaural cues. Moreover, in enclosed spaces, distance perception is highly affected by early reflections and reverberation. Different studies in the field of lateral and vertical reflections indicated a particular influence of the floor reflection on sound source localization and distance perception. Reflection pattern has influence on timbre of a sound signal. Especially the interferences of strong single reflections cause comb filter effects which change the perception significantly [1]. Guski [2] for example, suggested that the floor reflection improves localization accuracy. Furthermore Bech [3, 4] has shown that the natural signal level of some early reflections, including the floor reflection, exceed the human threshold of detection (*TD*). In a more recent study, Gourévitch and Brette [5] analyzed the influence of floor reflections on binaural cues by means of numerical models. They came to the similar hypothesis that the floor reflection may potentially provide additional cues for distance and elevation estimation. However, investigations on the influence of a single floor reflection on distance perception are very rare. Therefore, we decided to examine this issue by means of a listening experiment in a headphone-based virtual acoustic environment (*VAE*). Subjects compared a reflection-free sound signal (reference signal) to a superimposed one comprising a direct sound with a single floor reflection (comparison signal). The subjects' task was to rate relative perceived distances between both stimuli. Our results reveal that the floor reflection has a significant influence on distance perception.

Materials

We generated binaural room impulse responses (*BRIRs*), which are basically the superimposition of direct sound and one floor reflection. Thus, perceived acoustic information resulted from sound incidence of the direct signal and the first-order floor reflection exclusively. We summed up two single head related impulse responses (*HRIRs*), according to the angle of the incident direct signal and reflection. Indicated in black, Figure 1 displays the binaural room transfer function (*BRTF*) with Δt = 0.25 ms, δ_R = –30°. The first dip is clearly visible at 2 kHz and the first peak at about 4 kHz. In the range < 1 kHz a general level increase appears. The interference pattern, or comb filtering, influences timbre of a perceived sound signal. With increasing time difference, not only the number of peaks and dips increases, they also become more extensive towards lower frequency ranges, as shown in Figure 2. For comparison, the graphic display also shows the head related transfer function (*HRTF*) of the reference signal drawn in grey. We used HRIRs based on a full-spherical HRIR dataset by artificial head measurements [6]. We applied a first-order high shelf filter (fc = 1 kHz, G = –6 dB) to the reflection, according to slight absorption properties. The signal processing caused a slight fluctuating direct signal above 1 kHz (± 0.5 dB). To avoid distance perception based on level differences between both stimuli, due to energy increases summing up two coherent sound signals, we matched the energies of the two incident signals. The generated BRIRs were based on non-individual HRTFs of a Neumann KU100 dummy head. The source signal was white noise with a duration of 750 ms (10 ms linear on- and offset ramp).

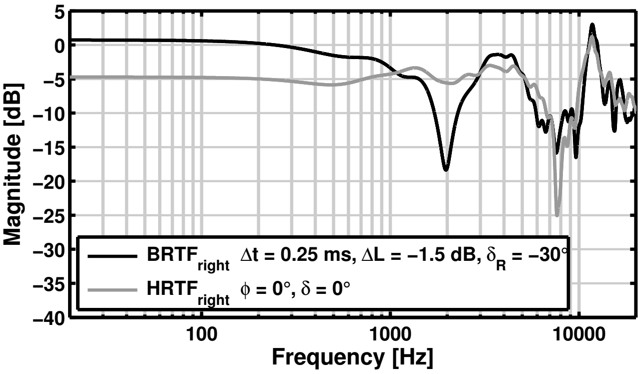

Figure 1: BRTF – comparison signal (Δt = 0.25 ms, δ_R = −30°, black). HRTF – reference signal (φ = 0°, δ_R = 0°, grey).

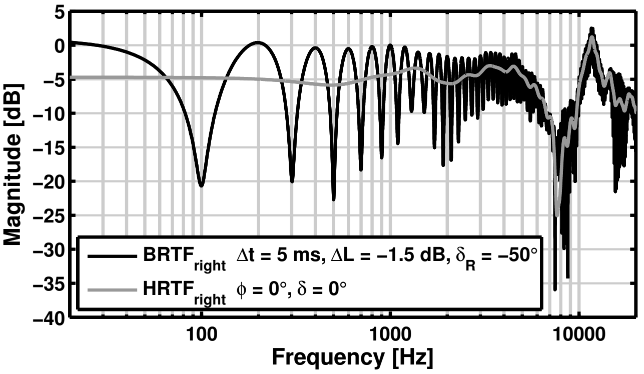

Figure 2: BRTF – comparison signal (Δt = 5 ms, δ_R = –50°, black). HRTF – reference signal (φ = 0°, δ_R = 0°, grey).

We defined a virtual sound source according to its direct signal and one reflection on the basis of their time of incidence, elevation and intensity. Due to different sound paths shown in Figure 3 we shifted the HRIR from the floor direction in time. Therefore, we graduated the influencing factor *time difference* Δt between direct signal and floor reflection to simulate different sender-receiver distances. To add another factor we also graduated the *reflection angle* δ_R. Furthermore, we reduced the floor reflection's intensity to gain a fixed level difference $\Delta L = -1.5dB$ between direct signal and reflection. We also defined the angle of the incident direct signal φ to a fixed value of 0° according to frontal sound incidence. This procedure was carried out for every horizontal head orientation in graduations of 1° in order to gain a 360° BRIR dataset for each virtual sound source. These parameters served as basis to create appropriate conditions for an estimation method applying the listening experiment.

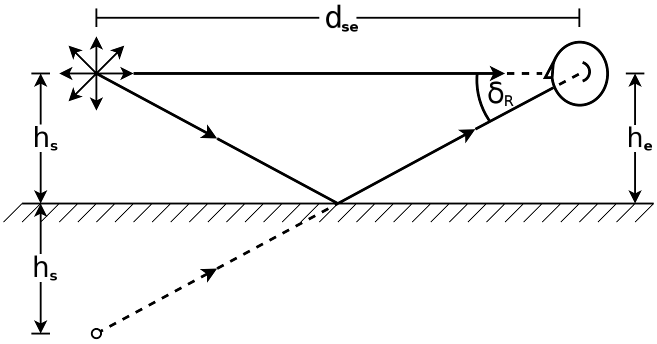

Figure 3: Geometric model of reflections. Spatial definition of an artificial BRIR. h$_s$ – height sender, h$_e$ – height receiver, d$_{se}$ – sender-receiver distance, δ_R – reflection angle.

Environment

Listeners performed the experiment in an anechoic chamber, which provides low background noise. For head-tracking, we used the *Polhemus FASTRAK* system. The *SoundScape Renderer* [7] was used for binaural rendering. The listening experiment was implemented and controlled with the MATLAB based software *Scale* [8]. Subjects entered their answers on a tablet PC (iPad).

Procedure

Within each trial, two different stimuli were presented successively. Received signals represented a direct signal free from reflections (HRIR) or rather one consisting of the same direct signal superimposed by a certain floor reflection (BRIR). The subjects' task was to specify whether the second auditory event was perceived as *closer* or *further* away than the first one. Presented at random order, the first stimulus was either an auralisation based on a HRIR (direct signal only) or a BRIR (same direct signal superimposed with floor reflection). The presented stimuli-pair was repeated twice before the subject was forced to make a choice.

Fifteen male and five female subjects aged from 19 to 34 with an average and median age of 25 participated in the listening experiment. The binaural system supported horizontal head rotations whereas vertical head movements were neglected. Subjects were encouraged to perform head rotations during the experiment. Within the experimental design, we defined the following conditions. We graduated the factor *time difference* in steps of Δt = 0.25 ms, 0.5 ms, 1 ms, 2 ms, 3 ms, 4 ms, 5 ms and 8 ms. For the *reflection angle* we set the parameters to δ_R = –5°, –30° and –50°. Time differences of 1 – 4 ms represent typical delays for floor reflections, as does an incident reflection angle of –30°. Greater distances between source and receiver were represented by 0.25 ms und 0.5 ms as well as the flat reflection angle of –5°. Providing observations about the behaviour of longer time differences we added 5 ms and 8 ms along with a typical –50° floor reflection angle for short sender-receiver distances. Table 1 shows a summary of all factors and their graduations. Resulting from the combination of all factors, subjects completed all 24 conditions in 16 trials (*within-subject design*). Three expert listeners appropriately adjusted the overall playback level to a comfortable audibility.

Table 1: Factors time difference Δt and reflection angle δ_R with their factor steps.

Factor	Factor steps	Count
Δt [ms]	0.25/0.5/1/2/3/4/5/8	8
δ_R [°]	–5 / –30 / –50	3

Results

The two answer categories *closer* and *further* were translated into a metrically scaled numerical value between –1 and +1. In case of an inverse presented succession within a stimuli-pair, the estimates were automatically inverted as well. In a first step we averaged the subjects' estimations over the repeated conditions. Secondly we averaged these subjects' mean values over the number of subjects. Thus we had a mean value per condition between –1 and +1.

The statistical analysis involved a two-way repeated measures ANOVA ($\alpha = 0.05$), which was GG-corrected for tests with more than one degree of freedom in the numerator (where GG is appropriate). The ANOVA showed a main effect for *time difference* ($F(7,133) = 72.41$, $p < .001$, $\epsilon = .32$, $\eta_p^2 = .80$) as well as a significant interaction effect among *time difference* and *reflection angle* ($F(14,266) = 3.42$, $p = .002$, $\epsilon = .50$, $\eta_p^2 = .15$). The main effect of *reflection angle* was not significant ($F(2,38) = 2.24$, $p = .14$). Further analysis revealed an ordinal interaction; thus, the main effect of *time difference* can be interpreted. Moreover, the main effect of *time difference* yielded a much higher effect size than the interaction effect. Following the variance analysis, we applied a one-sample t-test per condition to test whether the mean value was significantly different from zero. The alpha level was corrected according to Hochberg [9]. The evaluation shows 19 mean values, which are significantly different from zero for an alpha level of $\alpha = 0.05$ as shown in Figure 4. For perceived distances the plot shows three significantly *closer* estimations for the conditions with $\Delta t = 0.25$ ms. Interestingly, the curve rises between $\Delta t = 0.5$ ms and 2 ms showing two other conditions with significantly *further* estimates with $\Delta t = 1$ ms. In the range from $\Delta t = 2$ ms to 5 ms the 12 distances estimated as significantly *further* show an approximately straight line. The curve falls slightly for $\Delta t = 8$ ms with two other significantly *further* perceived distances.

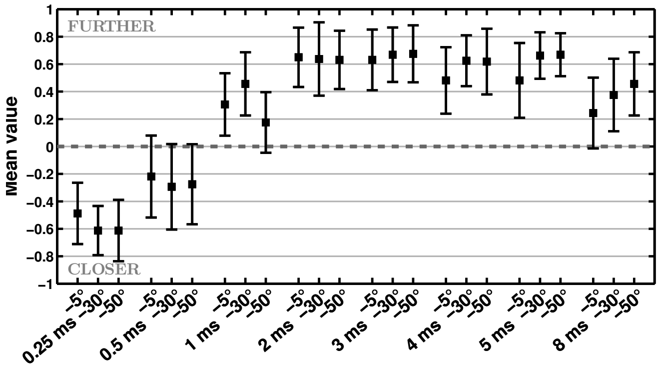

Figure 4: Mean values with their corresponding 95% confidence intervals.

Conclusion

We investigated auditory distance perception influenced by the floor reflection. By conducting a listening experiment subjects compared a reflection-free reference signal with a reverberant comparison signal. Results have shown that the distance perception is mainly influenced by the *time difference* between direct signal and the floor reflection. Because of sound level adaptation, distance perception was restricted to the auditory process based on spectral cues of the superimposed direct signal. The statistical analysis showed 16 conditions in which sound sources were frequently perceived as *further* away, whereas three conditions frequently caused a perception of *closer* sound sources. It is interesting to observe the mean values of evaluation increasing from *closer* perceived distances with low time differences to *further* perceived distances with larger time differences. Subjects' estimations were not significantly affected by the graduated factor *reflection angle*. Therefore, spectral changes relating to changes in the incident reflection angle did not significantly affect distance perception. Only the strong comb filter structure relating to the interfering floor reflection affected the perception significantly. With increasing time differences the wave movements, triggered by the reflection phase shift, clearly displace towards lower frequency domains. Different reflection delays affect timbre of the direct signal. Our results suggest that this leads to most perceptions of increased distance in the range of $\Delta t > 2$ ms. Frequency patterns for $\Delta t = 0.5$ ms and above 8 ms do not appear to be very useful for auditory processes that determine distance under preconditions presented in this study. Results are in agreement with Gourévitch und Brette's [5] hypothesis that monaural properties affect interaural time differences and interaural level differences, in particular through the floor reflection. These properties are used as binaural cues to determine distance. Our results are in agreement with Bech's [3, 4] findings that the natural signal level of four early reflections out of 17 ($\Delta t < 22$ ms), including the floor reflection ($\Delta t = 1.64$ ms; $\Delta L = -1.36$ dB; $\delta_R = -28°$), exceed the human TD unlikely, assuming the as *further* perceived distances indicate a gain of spatial impression. He postulated that the perceived spaciousness is affected in particular by the floor reflection.

These scientific findings might thus be a useful contribution in future virtual room design or rather in generating BRIRs. Additionally, they give an impetus for other research questions relating to perception of absolute distances based on the floor reflection's influence. Furthermore, there is a clear need for research on distance perception based on different elevated sound source directions. The consistent effect of the floor reflection on distance perception appears to play an important role within familiar environments. Thus, by including a specific floor reflection into VAEs further cues for determining a sound source distance will be provided, whereby the perceived position of a distant virtual sound source could lead to improved stability.

References

[1] Rubak, Per und Lars G. Johansen. „Coloration in Natural and Artificial Room Impulse Responses“. In: AES Conference: 23rd Int. Conf.: Signal Processing in Audio Recording and Reproduction, (2003) 1–19.

[2] Guski, Rainer. „Auditory localization: effects of reflecting surfaces“. In: Perception 19.6, (1990) 819–830.

[3] Bech, Søren. „Timbral aspects of reproduced sound in small rooms. I“. In: Acoustical Society of America. J. 97.3, (1995) 1717–1726.

[4] Bech, Søren. „Timbral aspects of reproduced sound in small rooms. II“. In: Acoustical Society of America. J. 99.6, (1996) 3539–3550.

[5] Gourévitch, Boris und Romain Brette. „The impact of early reflections on binaural cues“. In: The Acoustical Society of America, J. 132.1, (2012) 9–27.

[6] Bernschütz, Benjamin. „A spherical far field HRIR/HRTF compilation of the Neumann KU 100“. In: Proc. 39th German Annual Conf. on Acoustics (DAGA), (2013) 3–6.

[7] Geier, Matthias, Jens Ahrens und Sascha Spors. „The SoundScape Renderer: A Unified Spatial Audio Reproduction Framework for Arbitrary Rendering Methods“. In: AES Convention 124, (2008) 7330.

[8] Vazquez Giner, Arnau. „Scale- Conducting Psychoacoustic Experiments with Dynamic Binaural Synthesis“. In: Proc. of the DAGA, (2015).

[9] Hochberg, Yosef. „A sharper Bonferroni procedure for multiple tests of significance“. In: Biometrika 75.4, (1988) 800–802.