PROCEEDINGS of the 2nd Symposium: The Acoustics of Ancient Theatres
6-8 July 2022 Verona, Italy

Reinforcement of binaural cues by floor and ceiling reflections

Bernhard U. Seeber[1]

[1]Audio Information Processing, Technical University of Munich, Germany, seeber@tum.de

ABSTRACT

Sound reflections occurring only on floor and ceiling usually have the same azimuthal angle as the direct sound. Their energy thus adds to the direct sound with binaural cues similar to those of the direct sound. The anticipated benefits are of improved source localization and, for a frontal target, of higher binaural coherence of the ear signals. This modelling study investigated these hypotheses using a shoebox room whose specular reflections were modelled with the room acoustic simulation of the Simulated Open Field Environment (SOFE). Adding floor and ceiling reflections decorrelated the binaural ear signals and added late reverberant energy – C50 was higher without a floor or a ceiling. Binaural cues for very low and very high elevation angles converge to zero, i.e. to more centered binaural cues compared to those of a lateral direct sound source. An analysis with mapping these reflections into the horizontal plane did not show consistent evidence for an increase of correlation or coherence due to floor or ceiling reflections.

Keywords: spatial hearing in rooms, speech intelligibility, effects of reverberation

1. INTRODUCTION

Sound reflections in (shoebox) rooms occur – for sources at the same height as the listener – mostly parallel to the main horizontal dimensions of the room, and from floor and ceiling along the direct path from the source to the listener. These floor and ceiling reflections thus come from the same azimuthal angle as the source and could be seen as emphasizing the source's binaural cues. Clapp and Seeber [1] demonstrated that misplaced floor reflections can substantially affect the localization of the source if the direct sound is attenuated, e.g., because the speaker turns away. This was attributed to floor and ceiling reflections emphasizing binaural cues of the source against the diffuse energy in other reflections. The present investigation builds on this idea and questions if floor and ceiling reflections can indeed emphasize the binaural cues of the source. The question arises from the fact that binaural cues become smaller with increasing elevation angle of the source. Figure 1 shows interaural level differences (ILDs; upper panel) and interaural time differences (ITDs; lower panel) for a source at 45° azimuth as a function of source elevation. ILDs and ITDs are largest for horizontal sources (elevation of 0°) and decline if the source comes from below or above the horizontal plane. For sources overhead (elevation of 90°), binaural cues become zero. This is equivalent to the binaural cues of a source in the front, which leads to interaurally correlated ear signals. Hence, floor and ceiling reflections will always have smaller binaural cues by magnitude than the source itself, cues that indicate a horizontal source more toward the front (see right-hand ordinate of Figure 1). The present study investigates the impact of floor and ceiling reflections on interaural parameters relevant for binaural unmasking. Room acoustic simulation is used to create specular reflections of a shoebox room. To test the contribution of floor and ceiling reflections, the binaural room impulse response is manipulated by either omitting floor and ceiling reflections or by mapping them to horizontal sources at the same azimuth.

2. METHODS

2.1 Room Acoustic Simulation

The room acoustic simulation and auralization software of the Simulated Open Field Environment [2] (Matlab code) was used to create binaural room impulse responses (BRIRs). The software uses the mirror-image source method for arbitrary geometries [3] to create a list of image sources. Visible sources are synthesized by spectrally modifying the head-related transfer functions of the KEMAR manikin [4] of the image source location with the wall transfer function and the distance-related attenuation, and placing the resulting impulse response into the BRIR at the time related to the image source distance. Specular reflections up to reflection order 200 were computed, resulting in BRIRs of ca. 12 sec duration. Reflections from order 5 were temporally jittered by up to 5% of their delay to reduce the strict periodicity in the reflection pattern of the rectangular room. No specific diffuse field simulation was used.

In "mapped" conditions, image sources below the horizontal plane (floor reflections) or above the horizontal plane (ceiling reflections) were mapped to the horizontal plane (elevation 0°), while the azimuth angle of the image source was kept the same.

10.58874/SAAT.2022.186

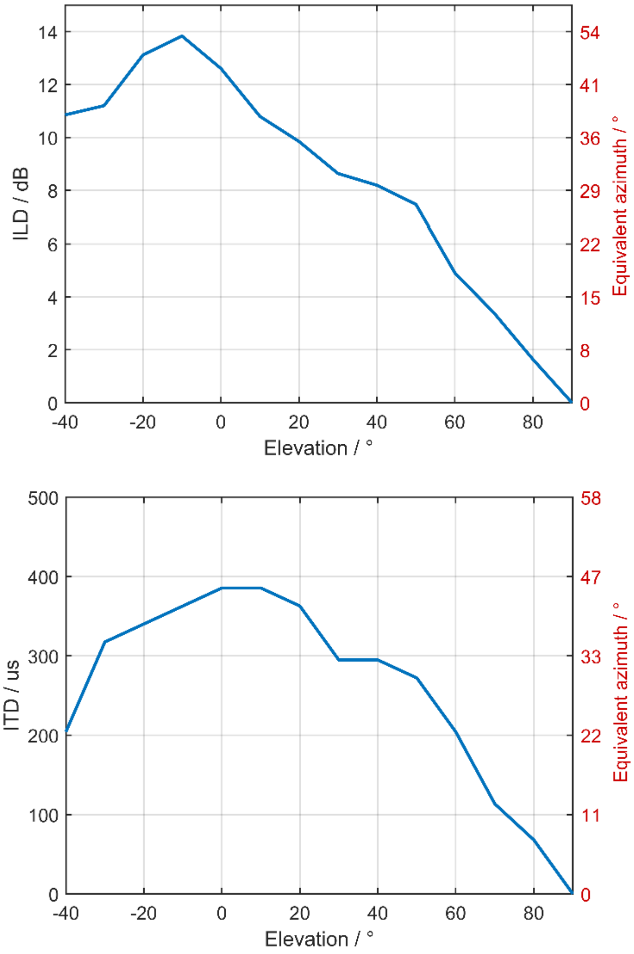

Figure 1 – Interaural level differences (ILD, top panel) and interaural time differences (ITD, bottom panel) as a function of the elevation of a source at 45° azimuth. ITDs and ILDs were extracted from the head-related transfer functions of a KEMAR manikin [4]. The right ordinate presents the azimuth angle of a source in the horizontal plane with the same magnitude of ILDs or ITDs. With increasing absolute source elevation, ITDs and ILDs tend toward zero, representing a horizontal source toward the front, which has interaurally correlated binaural cues.

2.2 Room conditions

BRIRs of an empty rectangular room with dimensions 20 m x12 m x 7 m were simulated. The source was positioned at (4.36 m, 17.64 m, 1.50 m) in 3.7 m distance to and at an angle of 10° left to the frontal direction of the receiver at (5.00 m, 14.00 m, 1.50 m). The directional characteristics of the source were those of a human speaker [5] in the direction of the receiver. Reflection coefficients of vertical walls were those of gypsum board, those of the floor mimicked wood parquet in asphalt on concrete and those of the ceiling stemmed from wood.

Seven room conditions were created to assess the impact of floor and ceiling reflections on binaural parameters:

6 walls, 3D – Simulation of the shoebox room with all 6 walls and “correct” 3D-mapping of the reflections. This room reflects the baseline condition. The top panel of Figure 2 depicts the reflection paths of the reflections in this condition up to reflection order two.

5 walls: No ceiling – The ceiling was omitted in this simulation, resulting in an shoebox room with 5 walls. This room represents an open-air theatre.

4 walls: No floor & ceiling – The floor and ceiling were omitted, resulting in a space made of only side-walls (i.e., perfectly absorbing floor and ceiling). This condition contrasts with the “6 walls, 3D”-condition in that the energy and binaural cues of floor and ceiling reflections are absent. The reflection paths for all reflections up to order two are shown in the lower panel of Figure 2.

6 walls: ceiling mapped horizontally – The above conditions are dominated by the energetic change of omitting floor and ceiling reflections. In order to specifically test the impact of reflection elevation angle on binaural parameters, in this condition reflections coming from above the horizontal plane were mapped to the horizontal plane while their azimuthal angle was kept identical and the number and the energy of all reflections was kept identical to “6 walls, 3D”.

6 walls: floor mapped – Reflections coming from below the horizontal plane were mapped to the horizontal plane while all other reflections were kept identical to the “6 walls, 3D” condition.

6 walls: floor & ceiling mapped – In this condition, both floor and ceiling reflections were mapped into the horizontal plane, i.e. all reflections were within the horizontal plane while the number of reflections and their energy was identical to the “6 walls, 3D” condition.

2.3 Room acoustic and binaural parameters

The ***direct-to-reverberant ratio (DRR)*** was computed as the ratio of the direct sound energy and the reverberant energy expressed in dB and averaged across both ears. To compute the reverberation-only BRIR, the BRIR of the direct sound was subtracted from the overall BRIR of the respective condition.

C50 expresses the ratio of the energy arriving within the first 50 ms (here: up to sample 2205 of the BRIR at a sampling frequency of 44.100 Hz) of the room impulse response and the energy arriving after 50 ms.

The ***interaural cross-correlation (IACC)*** is the normalized correlation coefficient of the two ear signals in the BRIR.

The ***coherence*** was computed as the maximum of the interaural cross-correlation function evaluated over a delay of ±1 ms. The coherence will compensate for the decorrelation caused by the azimuthal source offset of 10° and expresses the delay-independent similarity of the ear signals. Note that both IACC and coherence were computed from broadband signals.

The ***coherence ratio*** reflects the idea of useful source information arriving within the first 50 ms and detrimental cues arriving later (c.f. C50). Binaural unmasking is driven by a contrast in correlation. The ratio of the coherence of the first 50 ms of the BRIR and the coherence of the BRIR after 50 ms is computed.

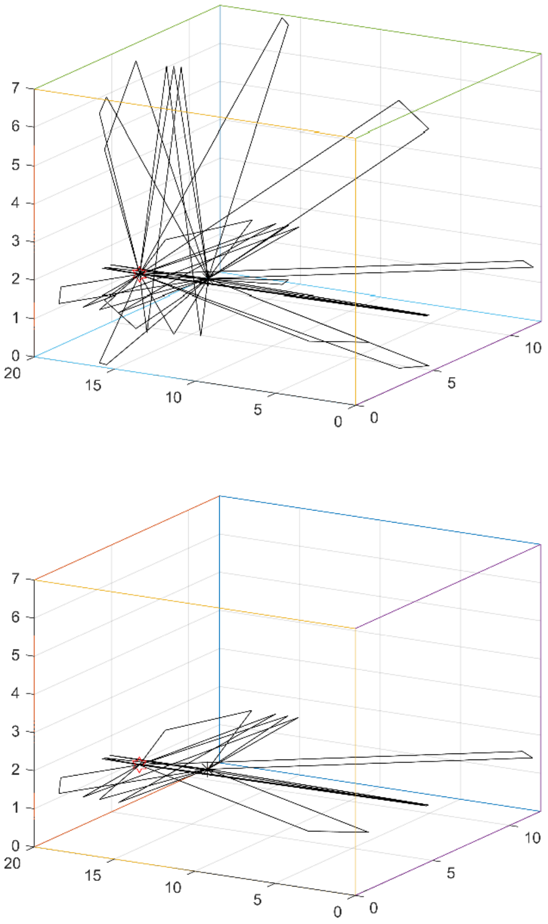

Figure 2 – Reflection paths up to the second reflection order in a simulated rectangular room with 6 walls (top panel) and in the same room without floor and ceiling (bottom panel). The source is at 350° azimuth relative to the receiver and depicted in red; the receiver is marked by a black asterisk.

3. RESULTS

Table 1 shows room acoustic parameters of the 7 room conditions. As one would expect, the DRR improves when omitting floor and ceiling reflections, indicating that an open-air arena should be beneficial in binaural terms. Mapping reflections to the horizontal plane does not affect the DRR since it is an energetic measure and the change was mostly binaural (though there is also a spectral change). C50 increases strongly when there is no floor or no ceiling. While some floor reflections arrive early, the majority of floor and ceiling reflections arrive in the late, “detrimental” part of the BRIR such that omitting those leads to an increase in C50.

Table 1 – Room acoustic parameters.

Room condition	DRR / dB	C50 / dB
6 walls, 3D	-4.3	2.4
5 walls: No ceiling	-1.4	8.1
5 walls: No floor	1.9	7.1
4 walls: No floor & ceiling	3.7	8.2
6 walls: ceiling mapped horizontally	-4.3	2.6
6 walls: floor mapped	-4.3	2.8
6 walls: floor & ceiling mapped horizontally	-4.6	2.6

Table 2 – Binaural parameters.

Room condition	IACC	Coherence	Coh ratio
6 walls, 3D	0.06	0.42	12.9
5 walls: No ceiling	0.03	0.62	5.0
5 walls: No floor	0.06	0.59	5.7
4 walls: No floor & ceiling	0.07	0.67	8.4
6 walls: ceiling mapped horizontally	0.03	0.45	14.7
6 walls: floor mapped	0.06	0.46	6.9
6 walls: floor & ceiling mapped	0.02	0.45	5.7

Table 2 shows results in terms of binaural parameters. For all room conditions, the IACC is close to zero, indicating highly decorrelated ear signals. The coherence, however, is considerably higher, suggesting that the 10° azimuthal offset of the source results in interaurally decorrelated, but still largely coherent ear signals. The anticipated effect of floor and ceiling reflections leading to higher correlation cannot be observed – it is more that their absence in the late reverberant tail leads to an increase in coherence.

The mapping of reflections from elevated angles to the horizontal plane increases coherence somewhat, in agreement with the hypothesis. The ratio of early and late coherence was another attempt to investigate the potential benefit of floor and ceiling reflections for emphasizing the source (correlation). However, the ratio increases when the ceiling reflections are mapped to the horizontal plane. A similar measure based on the correlation values decreases (from 2.22 for the 6 walls, 3D condition to 1.09 for the condition with horizontally mapped ceiling reflections), which could be supporting the hypothesis. An issue of useful-to-detrimental ratio analyses is that individual reflections might fall into or just out of the useful window, thereby changing computed parameters considerably [6]. Thus, these values appear too unreliable to draw formal conclusions from the present analysis – a more substantial modelling study using more rooms and a time and frequency dependent analysis of spatial unmasking would be needed to gain a better understanding.

4. CONCLUSIONS

This study investigated the question if the smaller binaural cues for sound reflections below and above the horizontal plane increase binaural correlation of a frontal target stimulus, which might be beneficial for localization and spatial unmasking. Analyses using manipulated reflections of a shoebox type room could not find consistent evidence – if a ceiling exists, the additional reflections decorrelate the target sound and they add late reverberant energy, while an open-air theatre configuration has increased C50. If floor reflections are mapped into the horizontal plane, the coherence slightly increases in favour of the hypothesis, but the ratio between coherence in the first 50 ms of the binaural room impulse response and coherence in the late part decreases, suggesting that the binaural contrast between both parts is reduced. A more detailed analysis using different rooms and additional temporal measures is needed to shed more light on the question.

ACKNOWLEDGEMENTS

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Projektnummer 352015383 – SFB 1330, Project C5. rtSOFE development is supported by the Bernstein Center for Computational Neuroscience, BMBF 01 GQ 1004B.

5. REFERENCES

[1] S. Clapp and B. U. Seeber, "Sound Localization in Partially Updated Room Auralizations," in *Fortschritte der Akustik - DAGA '16*, M. Vorländer and J. Fels, Eds., 2016: Dt. Ges. f. Akustik, pp. 558-560.

[2] B. U. Seeber, S. Kerber, and E. R. Hafter, "A System to Simulate and Reproduce Audio-Visual Environments for Spatial Hearing Research," *Hear Res,* vol. 260, no. 1-2, pp. 1-10, Feb 2010, doi: 10.1016/j.heares.2009.11.004.

[3] J. Borish, "Extension to the image model to arbitrary polyhedra," *J Acoust Soc Am,* vol. 75, no. 6, pp. 1827-1836, 1984.

[4] B. Gardner and K. Martin, "HRTF Measurements of a KEMAR Dummy-Head Microphone" MIT Media Lab, Technical Report 280, 1994.

[5] J. L. Flanagan, "Analog measurements of sound radiation from the mouth," *J. Acoust. Soc. Am.,* vol. 32, no. 12, pp. 1613-1620, 1960.

[6] J. Rennies, A. Warzybok, T. Brand, and B. Kollmeier, "Measurement and Prediction of Binaural-Temporal Integration of Speech Reflections," *Trends in Hearing,* vol. 23, Jun 2019, doi: 10.1177/2331216519854267.