CHALMERS
UNIVERSITY OF TECHNOLOGY

Perceptual Detection Thresholds for Alterations of the Azimuth of Early Room Reflections

Downloaded from: https://research.chalmers.se, 2024-04-26 04:49 UTC

Citation for the original published paper (version of record):
Bederna, F., Müller, L., Ahrens, J. (2023). Perceptual Detection Thresholds for Alterations of the Azimuth of Early Room Reflections. Tagungsband DAGA 2023 - 49. Jahrestagung für Akustik: 216-219

N.B. When citing this work, cite the original published paper.

research.chalmers.se offers the possibility of retrieving research publications produced at Chalmers University of Technology. It covers all kind of research output: articles, dissertations, conference papers, reports etc. since 2004. research.chalmers.se is administrated and maintained by Chalmers Library

(article starts on next page)

Perceptual Detection Thresholds for Alterations of the Azimuth of Early Room Reflections

Felicitas Bederna, Leon Müller, Jens Ahrens
Division of Applied Acoustics, Chalmers University of Technology, 41296 Gothenburg, Sweden

Introduction

Applications like augmented and virtual reality require controlled creation of virtual acoustic spaces. For perfect reproduction of a room not only the direct sound and a few reflections but all reflections need to be rendered and thus not only the pressure room impulse response (RIR) but all incidence angles of the reflections in combination with the RIR, the so-called spatial room impulse response (SRIR) needs to be known either by simulations or measurements of a real room with a microphone array. If the virtual receiver is granted the freedom to move in the room, the matter gets much more complicated since the SRIR is different at every position in the room. Thus the question arises how accurately the spatial information in an SRIR needs to be rendered to create the desired room impression. In the present work, we focus on the early reflections.

There is a vast amount of literature available from which conclusions on the perceptual significance of the early reflections can be drawn. Some high-level conclusions were drawn in the literature on concert halls where it was found that, for example, strong lateral reflections can increase the apparent source width [1]. Similarly, the general perceptual properties of manipulated room responses was investigated in the context of 6-degrees-of-freedom binaural audio reproduction. The relevant literature is summarized in [2], and an example of a practical case study is [3]. Also a number of low-level properties were investigated. In [8, 10], the audibility threshold for manipulation of a single reflection was determined. [4] presents an experiment on the perceptual salience of selected early reflections inside a large set of early reflections, and [5] presents an experiment on the significance of non-horizontal information in room responses.

One conclusion that may be distilled from the literature is that it is not clear at this point what the general thresholds are regarding the required accuracy of the spatial information for achieving a perception that is indistinguishable from the reference. In many cases, the actual room response may differ considerably from the reference while the perceptual properties can still be very close to the reference.

As a complement to especially [8, 10], this work provides a first insight into the perception of room responses in which several reflections were modified compared to a reference. For this, the azimuth of five horizontal early reflections in a selected room response is manipulated in the course of a listening experiment in order to find the median detection threshold, i.e. the angle shift at which the manipulation becomes audible.

Experiment Design

Stimulus creation

In the course of the presented work, it turned out to be challenging to identify a suitable SRIR as identification of reflections in an SRIR is not trivial. Also, alterations of reflections that were not very prominent did many times not result in an audible difference. The room impulse response that was finally chosen for this study was recorded in the Audio Lab at the Division of Applied Acoustics at the Chalmers University of Technology. It is a shoe-box shaped room with a volume of approximately 46 m^2 and a reverberation time of around 0.27 s. The impulse response was analyzed and the directional information in form of the azimuth and elevation angles were determined using the spatial decomposition method (SDM) [6]. The first five reflections in the RIR were very prominent and could be identified as all first order reflections of the room (see upper and middle plot in Figure 1). All subsequent reflections were not clearly identifiable due to ambiguities in the detection criteria. It was therefore decided to investigate alterations of those first five reflections.

The first step in the stimulus preparation was quantization of the directions of arrival on a Lebedev sphere grid of 50 points in order to avoid spreading the reflections in multiple directions that can be a consequence of the imperfect direction of arrival estimation. Then, the entire room impulse response was projected onto the horizontal plane by setting the elevation of all incidence angles to 0°. This was assumed to help make the perceptual differences due to azimuth angle alterations as strong as possible. The azimuth angles of the first five reflections were randomly set to 45°, -60°, 135°, -80° and -160° (see lower graph in Figure 1), so that no reflections are lying on the median plane and that the reflections are distributed over all possible directions the horizontal plane. The alterations that were applied to the incidence angles of the reflections were excluded from the quantization process so that the angular shift could be chosen freely. The so created RIR was used as the reference "original" RIR during the listening experiment.

Angle alteration in the adaptive threshold listening test

The subjects' task was to differentiate a test signal that was convolved with the binaural RIR (BRIR) with altered incidence angles of the reflections from the same signal convolved with the reference BRIR. The direct sound was always impinging from straight ahead. The adaptation process started with all five strong reflections

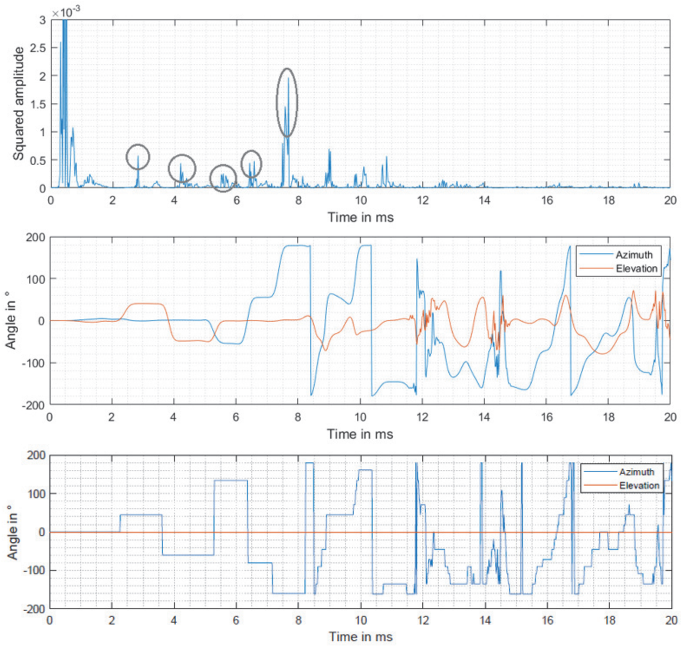

Figure 1: Used RIR in the time domain with marked first reflections (upper graph), the extracted directions of arrival of every sample in the RIR (middle graph) and the adjusted and quantized angles (lower graph).

squashed to the left-hand side relative to the user's nominal look direction, which produced a significant perceptual difference. Cf. the dotted lines in Figure 2.

The comparison was implemented as an ABX comparison and the adaptation process followed the adaptive two-up-one-down method. One of the stimuli A and B was always the reference, and the other stimulus comprised the altered incidence angles. If a participant was able to correctly identify whether the stimulus X was identical to A or to B twice in a row, the angle shift of the reflections in the altered stimulus decreased, and the reflections moved towards the incidence angles in the reference. However, if the participant did not answer correctly, the angle shifts of the reflections increased, and the reflections in the altered stimulus moved farther away from the reference. Furthermore, whenever the participant changed from the "up" movement (wrong answers) to the "down" movement (two correct answers), the step size is halved, so that the step size decreases in the course of the experiment. This increases the resolution of the resulting threshold while it reduces the measurement time as much as possible.

All reflections move with a step size that corresponds to the same percentage of their maximum shift so that they all reach their original positions at the same time. This means that the first reflection, for example, which is with its original angle of 45° closer to 90° than the second reflection with an original angle of 135°, has the same percentage step size as the second one but an accordingly smaller absolute step size. The starting point of the adaptation process was placed between the upper and lower limit, where a clear difference between the RIRs was audible. The specific values of the angles as well as their absolute step sizes are provided in Table 1. Non-head-tracked audio examples of the stimuli are available at[1].

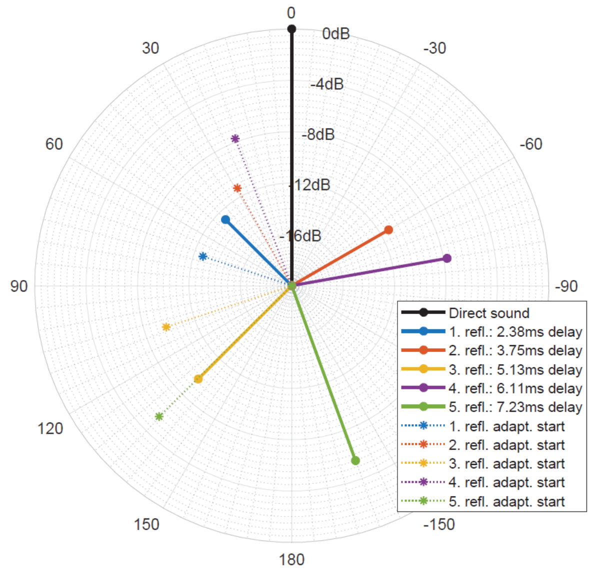

Figure 2: Original angles of incidence of the direct sound and the five first reflections with their relative levels to the direct sound. The dotted lines show the angle of incidence of each reflection at the beginning of the adaptation process.

Table 1: Original and initial azimuth angles and step sizes for the angle shifting for all five first reflections.

Refl.	Orig. angle	Init. angle	Step size
1.	45°	71.74°	0.2647°
2.	-60°	29.12°	0.8824°
3.	135°	108.57°	0.2647°
4.	-80°	21°	1°
5.	-160°	137.64°	0.6471°

Test setup and execution

Prior to the experiment, the reference RIR as well as the RIRs with all possible first reflection shifts were created, quantized, and each digital sample is convolved with a head-related-transfer-function (HRTFs) for the according angle of incidence of that sample to create BRIRs using the implementation of SDM auralization from [7].

The listening test was conducted in a quiet room at Chalmers University of Technology in Gothenburg, Sweden. The participants were placed in front of a computer screen and the signals were presented via headphones with head tracking. The RIRs were tested with both, a speech signal and a drums signal as test signals that were convolved with the BRIRs. This was done using a block-based real-time overlap-add convolution of the BRIRs with the test signals to be able to quickly adjust between the original and the adjusted BRIRs during the comparison of the signals. This way, the participants were able to directly compare the original and adjusted RIR without interruption and without restarting the test

[1]`http://www.ta.chalmers.se/research/audio-technology-group/audio-examples/daga-2023b/`

signal. Each participant started with a short training session, followed by the two main experiments, the ABX threshold test once with the speech and once with the drums signal. Each test was completed after 8 reversals or interrupted in case the participant reached the upper limit of the adaptation range. A more detailed description of the test setup and the generation of the signals can be found in [11].

Participants

The experiment was performed by 20 people, 13 of them experienced and very experienced and 7 without any prior experience with listening tests. The age of the subjects ranged from 23 to 46 years with a mean age of 28.95 years. All subjects reported to have normal hearing.

Results and Discussion

The result of the individual listening tests are two adaption curves per subject, one for the speech and one for the drums signal. An example of such an adaptation curve can be seen in Figure 3, where the given angle shift values on the y-axis refer to the biggest shift of the five individual angle shifts (this is done for simplicity and is the same for all visualizations that only show one angle shift). The angle shift threshold for a test subject and one test signal was determined to be the average of the angle shifts of the last four reversal points of the measurement curve (filled circles in Figure 3). The length of the individual tests as

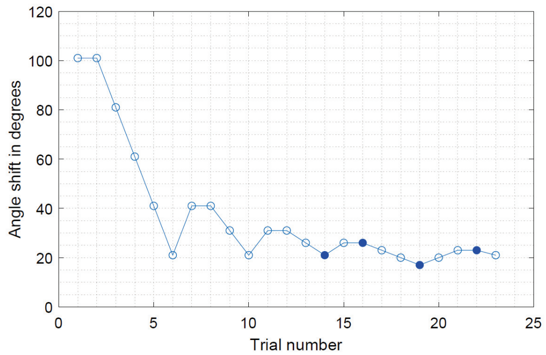

Figure 3: Listening test curve for the second participant and the speech signal.

well as the time needed to complete them varied strongly between the 20 participants. The number of tails until the convergence point was reached differed from 22 up to 44 trails and the time needed was in a range of 4 min for the fastest person up to 30 min for the slowest person. The median angle shifts as well as the 25th and 75th percentiles and outliers for the speech and drums tests are displayed in Figure 4. For simplicity, only the biggest angle shift of the five reflections (the fourth reflection) is plotted in the graph. The corresponding median angle shifts of all five reflections can be seen in Table 2 and are visualized in a polar plot together with the original angles in Figure 5.

It can easily be seen that the median angle shift of the speech and the drums signal are very similar, even though the percentiles and the outliers vary in a bigger range for the speech signal than for the drums. Since the order of the two test signals was randomized for each participant,

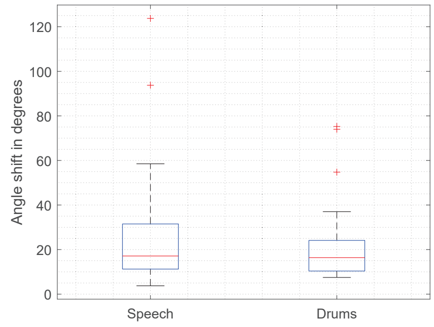

Figure 4: Median and the 25th and 75th percentiles of the thresholds for both the speech and the drum signal. The whiskers extend to the extreme values up to 1.5 times the interquartile range and outliers are marked with a cross.

Table 2: Median angle shifts of all five reflections for the speech and the drum signal at the measured threshold.

Refl.	Shift speech	Shift drums
1.	4.53°	4.33°
2.	15.11°	14.45°
3.	4.53°	4.34°
4.	17.13°	16.38°
5.	11.08°	10.6°

this is not a result of a training effect. Instead, it is assumed that the transients in the drums signal could have helped the participants to hear the reflections more distinctively and thus helped to distinguish the signals more consistently. However, all outliers except for one are caused solely by inexperienced listening test participants.

When analyzing the found thresholds the first step was to determine if all five adjusted reflections would in fact be individually audible in the RIR. This was done by comparing the relative levels of the reflections with the estimated audibility thresholds of the five reflections using [8] (at 10 ms after the direct sound). It was found that all reflections should lie above their estimated thresholds, thus should individually be audible and could all influence the perception. Together with the direct sound, all of the five reflections lie in the range between 1 ms and 50 ms and thus in the area of the precedence effect. According to this law, the localization of the sound should not be influenced by the reflections but they can influence the spatial extend, tone color and the feeling of the center of gravity [9]. This is in accordance with the informal feedback of some subjects who reported the biggest perceptual difference for them was the change of gravity.

The thresholds found in this experiment lie in the range between 4.3° to and 17.1° for the five reflections. These thresholds can be compared to a similar examination in [10], where the authors obtained the maximum angle shift of a single side reflection in a RIR that was not distinguishable from the original non-shifted reflection. The median angle shift found for the reflection originally coming from 70° was 22° and thus similar to the 16.38° and 17.13° median angle shifts found in this work for the fourth reflection with the biggest angle shift. Addition-

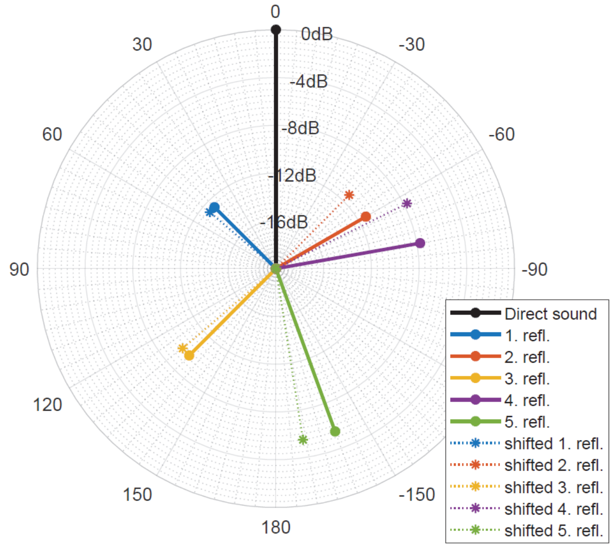

Figure 5: Original angles of incidence of the direct sound and the first five reflections with their relative levels to the direct sound. The dotted lines show the median angle shift thresholds for each reflection for the speech signal.

ally, the fourth reflection has an original incidence angle from -80° and is hence comparable to the one from [10] at 70° in terms of the amount of localization blur at this angle (even though left-right mirrored). This comparison allows the possibility that the found angle shift threshold could entirely be caused by the biggest angle shift only while the other reflections with their smaller angle shifts could have been inaudible. That the threshold found in this work is a few degrees lower than in [10] could be due to differences in the test design. However, the smaller angle shift threshold could also indicate that the other reflections or at least one other reflection had an influence on the perceived difference and thus both reflections can be shifted closer to their original angles than it would have been possible for them individually. In order to limit the experiment duration, the angle shift threshold of the fourth reflection was not determined in this study. However, this step, as well as collecting a formal subjective feedback from the participants, could be implemented in a follow-up experiment to better understand the psychoacoustic effects that are caused by shifting different reflections at the same time compared to only shifting the individual reflections on their own.

Conclusion

In this work, the angle shift threshold of five first reflections in a RIR was examined using an adaptive ABX threshold test and speech and drums test signals. The five reflections were shifted simultaneously in order to find the point where the participant is just able to distinguish the shifted from the original RIR. The found median thresholds for the five reflections lie between approximately 4° to 17°. While the chosen experiment design does not reveal whether the reflection with the strongest angular shift or the combination of multiple shifted reflections causes perceivable differences, this study nevertheless showed that a room with reflections that are shifted by less than the values found in this work is not distinguishable from the original one. Given that a room with slightly altered reflections, which can be differentiated from the original room by a direct comparison, can still provide a convincing replica of the original room, the simulation and measurement of RIRs can be significantly simplified by the lower spatial resolution required.

References

[1] D. Griesinger, "The psychoacoustics of apparent source width, spaciousness and envelopment in performance spaces", *Acta Acustica united with Acustica*, vol. 83, no. 4, pp. 721-731, July, 1997.

[2] A. Neidhardt, C. Schneiderwind, F. Klein, "Perceptual Matching of Room Acoustics for Auditory Augmented Reality in Small Rooms - Literature Review and Theoretical Framework", *Trends in Hearing*, vol. 26, 2022.

[3] T. Deppisch, S. V. Amengual Garí, P. Calamia J. Ahrens, "Perceptual Evaluation of Spatial Room Impulse Response Extrapolation by Direct and Residual Subspace Decomposition", AES AVAR, August, 2022.

[4] F. Brinkmann, H. Gamper, N. Raghuvanshi, and I. Tashev, "Towards encoding perceptually salient early reflections for parametric spatial audio rendering", AES AVAR, August, 2020.

[5] L. Müller, J. Ahrens, "Perceptual Differences for Modifications of the Elevation of Early Room Reflections", AES AVAR, August, 2022.

[6] S. Tervo, J. Pätynen, A. Kuusinen, T. Lokki, "Spatial Decomposition Method for Room Impulse Responses", *Journal of the Audio Engineering Society*, vol. 61, no. 1/2, pp. 17-28, January, 2013.

[7] S. V. Amengual Garí, J. M. Ahrend, P. T. Calamia and P. W. Robinson, "Optimizations of the spatial decomposition method for binaural reproduction", *Journal of the Audio Engineering Society*, vol. 68, no. 12, pp. 959-976, January 14, 2021, DOI: 10.17743/jaes.2020.0063.

[8] X. Zhong, W. Guo and J. Wang, "Audible threshold of early reflections with different orientations and delays", *Computers, Materials & Continua*, vol. 61, no. 3, pp. 18-22, 2019, ISSN: 10.32604/sv.2018.03900.

[9] J. Blauert, *The Psychophysics of Human Sound Localization.* The MIT Press, 1996. DOI: 10.7551/mitpress/6391.001.0001.

[10] O. C. Gomes, N. Meyer-Kahlen, W. Lachenmayr and T. Lokki, "Perceptual consequences of direction and level of early reflections in a chamber music hall", 2022.

[11] F. Bederna, "Perceptual detection thresholds for alterations of the azimuth of early room reflections", Master's Thesis, 2022.