Perceptual Evaluation of Differences in Anthropometric Parameters

Hark Braren and Janina Fels
Institute of Technical Acoustics, Teaching and Research Area Medical Acoustics, RWTH Aachen University
Email: {hbr, jfe}@akustik.rwth-aachen.de

Introduction

Head-related transfer functions (HRTFs) are known to be highly subject-specific due to varying expressions of anthropometric features. This holds especially true for children, whose torso, head and pinna geometries are continuously changing deeply up into their teenage years [1]. Modern 3D-scanners capture the shape of these features with a high level of accuracy. Combined with well-established simulation methods, e.g. the boundary element method (BEM) it is possible to simulate their effect on the sound field as it is captured by the HRTF. In addition, these 3D-shapes can be manipulated, in order to introduce systematically controlled anthropometric differences to assess their effect on the HRTF and the resulting perceptual differences.

In the present work, a preliminary study to determine the audible influence of manipulating individual geometry parameters on certain perceptual attributes is presented. In this study, the effect of controlled modifications to the pinna geometry of a known artificial head geometry on perceptual parameters is evaluated. Shape parameters of the outer ear are modified and the effect on the HRTF is evaluated a in preliminary listening experiment.

Pinna Parameter Scaling

The baseline used for this study is the ITA Artificial Head [2], shown in Figure 2 in green, whose pinna shape was chosen to work for a broad audience. Pinna parameters can be measured in many different ways. One common approach, which has been widely adopted is described by Algazi et al. and used in the CIPIC database [3]. The parameters used for modification were chosen based on a correlation analysis on the ITA Anthropometric Database [4]. Figure 1 shows the set of pinna parameters with the chosen parameters d3, the cavum concha width, and d6, the pinna width encircled.

To modify the pinna shape, the parameters were increased along their direction whilst keeping the rest of the parameters as constant as possible. The results of an increased pinna width can be seen in Figure 2. The modification strength again was chosen based on a statistical evaluation of the ITA database shown in table 1. The parameters were increased by 1.5 standard deviations in order to see objective differences in the HRTFs, as shown in Figure 3.

The presented HRTFs were simulated using the boundary element method (BEM) in Comsol Multiphysics Version 5.5 [5]. It can be seen, that even an increase of 1.5 standard deviations, in the following indicated by a + sign, the objective differences are still very small. Only an abnormally large increase of a single parameter, shown here for d6 increased to the maximum value found in the used database and indicated by a ++ leads to HRTFs that show a significant variation. The resulting pinna geometry of said large increase can be seen in Figure 4.

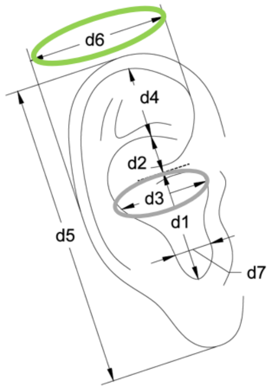

Figure 1: Pinna Dimensions comparable to [3]

Table 1: Statistical evaluation of the pinna parameters

	mean	std	rel. std
d3	18.1 mm	2.6mm	15%
d6	33.3 mm	3.03 mm	9%

Perceptual Attributes

Non-individual HRTFs are known to affect the perception of a given sound event in multiple ways. The spatial audio quality index (SAQI) [6] gives a vocabulary of perceptual qualities with corresponding scale attributes which can be used as a starting point to evaluate the

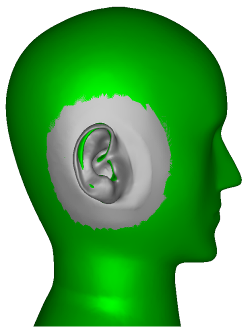

Figure 2: Pinna with increased parameter d6 (grey) and original ear (green).

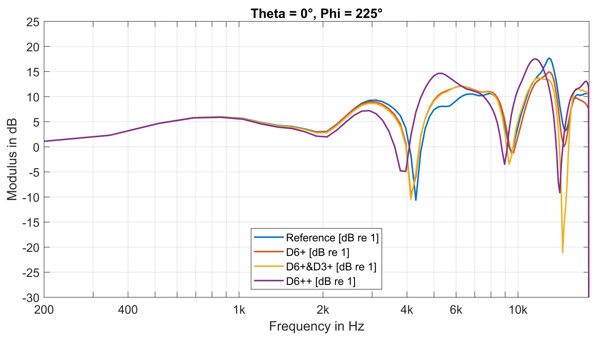

Figure 3: Right-ear-HRTFs resulting from a parameter increase of 1.5 std (+) and to the maximum in the database (++) in the horizontal plane.

Figure 4: Increased d6 to the maximum of the ITA Anthropometric Database.

influence of the geometric modifications. In addition, sound examples for the various attributes are provided which can be used to help guide non-expert listeners in a listening experiment where such attributes are to be evaluated.
For the evaluation in the present work, the following parameters were evaluated:

- Horizontal localization
- Elevation localization
- Distance
- Externalization
- Comb-filter effects
- Coloration
- Overall difference

Experimental Design

A short listening experiment to evaluate the ability to determine perceptual differences between the original and different modified HRTFs was conducted. A pink-noise burst train consisting of three bust, each 300 ms in length, with 200 ms pauses in between bursts was used.
The experiment design chosen to evaluate the objectively small differences in the HRTFs caused by modifying individual parameters was a modified MUSHRA [7] design. The design uses a hidden reference and low-pass-filtered anchor stimulus as per the MUSHRA standard, however the scale has been adapted to better accommodate the different perceptual parameters as described in section. As an example, the coloration can be both brighter and darker, therefore a symmetric scale with a zero coloration value as the starting central value.
Three directions of the modified and reference HRTFs, namely the frontal direction $(\theta, \varphi = 0°, 0°)$, right side direction $(0°, 270°)$ and an elevated position $(75°, 65°)$, were chosen and convolved with the noise train to create the stimuli. The stimuli were presented via static headphone presentation. The headphones used were Sennheiser HD 650 whose headphone transfer functions (HpTFs) were measured for each subject and compensated according to Masiero [8]. Prior to the experiment, an audiometrie was conducted to ensure each subject was of normal hearing ability up to 16 kHz. Three male subjects took part in this preliminary evaluation of whether the described modifications were actually audible.

Results

The results of the described test are presented below. They are to be taken with a grain of salt and do only indicate trends to improve an actual listening test design due to the limited number of participants. The general consent of the participants was, that there were three HRTFs that were almost indistinguishable from the reference namely d3+, d6+ and the hidden reference. This also shows in the results.

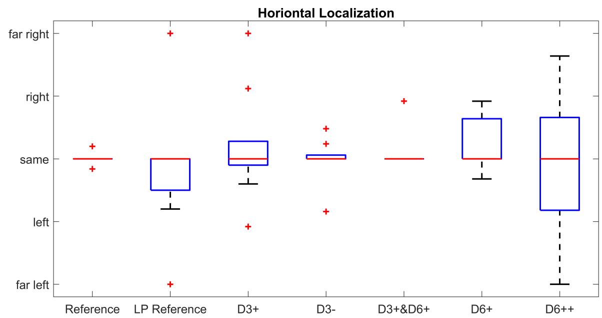

Figure 5: Perceived difference in horizontal localization.

Figure 5 shows the results for horizontal localization averaged across the three tested directions. No notable difference was observed, not even for the low-pass filtered anchor and abnormally increased d6++ parameter. Here only a notable increase in the spread of the answers was observed, with no clear direction.

The results for other perceptual attributes were a bit clearer. With regard to a perceived coloration compared to the known reference, the low-pass filtered anchor was correctly identified, so was the hidden reference (denoted as reference in the plot). The increased cavum concha d3+ shows an increase in brightness as does the abnormally stretched pinna width d6++. However the variance among the three subjects is too large to give a definitive result.

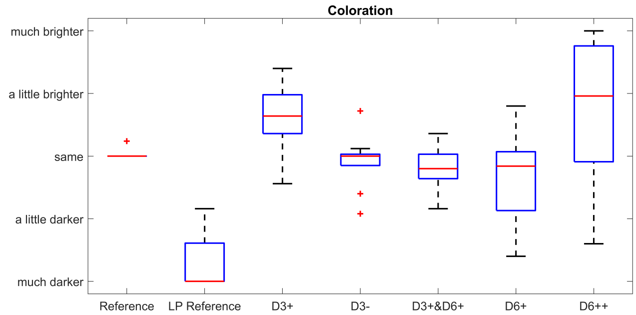

Figure 6: Perceived coloration difference.

Conclusion

As shown in the trends presented in the previous section, the differences caused by modifying only individual parameters of an HRTF are barely audible, thus an evaluation should include a larger spread of the selected parameters. Also head geometries should be considered, when having a comparison to a similar experiment with child subjects in mind. Fels et al. [1] showed that these parameters change at even at older ages compared to the pinna dimension. In addition they affect the important binaural cues, namely the interaural time and level differences.

References

[1] Fels, J., Buthmann, P., & Vorländer, M. (2004). Head-related transfer functions of children. Acta Acustica United with Acustica, 90(5), 918–927.

[2] Schmitz, A. (1995). Ein neues digitales Kunstkopfmesssystem. Acustica : International Journal on Acoustics, 81(4), 416–420.

[3] Algazi, V. R., Duda, R. O., Thompson, D. M., & Avendano, C. (2001). The CIPIC HRTF database. Proceedings of the 2001 IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics, 99–102. https://doi.org/10.1109/ASPAA.2001.969552

[4] Bomhardt, R., Klein, M. de la F., & Fels, J. (2016). A high-resolution head-related transfer function and three-dimensional ear model database. 172nd Meeting of the Acoustical Society of America, 172.

[5] Comsol Multiphysics, Version 5.5, `www.comsol.com`

[6] Lindau, A., Erbes, V., Lepa, S., Maempel, H. J., Brinkman, F., & Weinzierl, S. (2014). A spatial audio quality inventory (SAQI). Acta Acustica United with Acustica, 100(5), 984–994. https://doi.org/10.3813/AAA.918778

[7] Recommendation ITU-R BS.1534-1 (01/2003) Method for the subjective assessment of intermediate quality level of coding systems. International Telecommunication Union, Radiocommunication Assembly

[8] Masiero, B., & Fels, J. (2011). Perceptually robust headphone equalization for binaural reproduction. Audio Engineering Society 130th Convention, 1–7. https://doi.org/10.13140/2.1.1598.6882