PROCEEDINGS of the
23rd International Congress on Acoustics
9 to 13 September 2019 in Aachen, Germany

Objective differences between individual HRTF datasets of children and adults

Hark BRAREN and Janina FELS

(1)Teaching and Research Area of Medical Acoustics, Institute of Technical Acoustics, RWTH Aachen University, Germany,
{hbr, jfe}@akustik.rwth-aachen.de

Abstract

Subjective evaluations of noise by children and its effect on their learning abilities are a current research topic. Studies in this field often make use of virtual acoustic scenarios to be able to present a controlled acoustic environment in which to conduct the experiments. The presentation of realistic acoustic scenes requires a set of Head-Related Transfer Functions (HRTFs) that in the best case is the subject's individual HRTF, or matches it as closely as possible. However in most cases only a generic HRTF from an artificial head is available. This leads to perceptive deviations in location, coloration and reduces the immersion in the presented scene. Previous studies have shown, that anthropometric parameters, and by inference HRTFs differ especially for younger children compared to adults. In this study, individual high-resolution HRTF datasets as well as head and torso geometries of 26 children (ages 5-9 years) have been measured using a fast HRTF measurement system and a structured light 3D Scanner. The HRTFs are compared to those from adults and an artificial head to objectively evaluate their similarities and differences. The results indicate, that even though anthropometric geometries increase significantly in size with age, both children and adults show similar differences compared to an artificial head. Keywords: children HRTF, binaural hearing

1 INTRODUCTION

The influence of auditory distractions or background noise on the behavior and learning capabilities of children is an ongoing topic of interest. Studies like [1] utilize 3D virtual acoustic scenes, often presented over headphones using binaural technology, to create a controlled acoustic environment. This method of presentation requires a set of head-related transfer functions (HRTFs) which are highly subject specific. Even though other means of sound reproduction, with array based systems like Ambisonics, VBAP or sound field synthesis techniques, that do not rely on HRTF data, exist, headphone-based reproduction methods are often preferred due to the much reduced hardware effort.

While different methods of individualizing HRTFs [2, 3] are the topic of past and ongoing research, using an artificial head measurement remains the HRTF dataset of choice, due to the broad unavailability of HRTF measurement systems. While the accuracy of HRTF reproduction using artificial heads has been evaluated for adults, its applicability to children, with smaller anthropometric dimensions, has not yet seen such an in depth evaluation. It is known from earlier studies by Fels [4, 5], that HRTFs from children differ significantly from adults. In addition, studies by Prodi et al. [6] show differences in the inter-aural cross correlation (IACC) and speech transmission index (STI) when using different artificial heads (children vs. adult) in room acoustic measurements

There are many aspects concerning the quality of a reproduction system using non-individual HRTFs, such as its spectral coloration, apparent source width and distance or localization (in-)accuracy. While the localization accuracy in the horizontal plane is primarily influenced by the inter-aural time difference (ITD) and the inter-aural level differences (ILD) [7], the elevation accuracy dominated to monaural cues, meaning direction and subject specific peaks and notches in the spectrum of a sound event reaching the ear drum. These monaural cues are caused by the subject specific shape of the outer ear, which exhibits characteristic resonances that change in frequency and magnitude depending on the direction of incidence. Baumgartner [8] published a prediction model, based on spectral gradient extraction, that allows an objective prediction of localization accuracy in the sagittal plane when listening with a foreign pair of ears as captured for instance by a artificial head HRTF.

2 DATA ACQUISITION

To obtain children HRTF datasets, 26 children, ages 5 to 9 were invited to the hemi-anechoic chamber at the Institute of Technical Acoustics in Aachen, where the fast HRTF measurement system, evaluated by Richter et. al. [9] was used to measure 5 x 2.5 degree (azimuth x elevation) resolution HRTFs of the children. The measurement system consists of 64 loudspeakers placed at 2.5 degree increments on a circular arc with a radius of 1.2 m from ≈ 0 to 160 degree zenith angle. The arc is anchored at the top, where a silent stepper motor can rotate the arc 360 degrees in azimuth around a subject placed in its center. During the measurement procedure, the children were seated on a variable height chair, such that their head was centered in the sphere created by the rotating measurement-arc with their ear canal entrances located in its vertical center at 2 m high. A variable head rest was used to align the front back and left to right directions. To reduce chances of injury by falling, a seatbelt was attached to the chair which was closed during the measurement. In addition a parent was allowed inside the hemi-anechoic chamber outside the critical radius of the measurement arc. The critical radius is defined by the distance between the lowest speaker and reflecting floor, which causes the first reflection that has to be time windowed in post processing. Any later reflections from equipment and persons inside the measurement room will thus be removed as well. The whole measurement process took approximately 3 minutes using a continuously moving measurement arc that rotates around the measured subject while playing back interleaved sweep signals [10]. To further reduce movement of the childrens' heads during the measurement process, a LCD screen was set up at head height displaying muted TV show clips to give the children a visual target, which they are interested in looking at. The clips were chosen from shows, that are age appropriate and don't contain dialog or other important auditory content so that the lack of sound was not disturbing to the children. Additionally overly humorous content was avoided to prevent disturbances in the measurements due to laughter. The children were instructed to remain quiet and still for the duration of measurement, but that they should indicate any signs of discomfort to their parent or experiment supervisor present in the room.
In addition to the HRTF measurement, a high resolution 3D-scan of the the children's head and upper torso was captured using a structured light scanner (Space Spider by Artec 3D).

3 OBJECTIVE DIFFERENCES

3.1 Horizontal Plane

Auditory localization tasks in the horizontal plane mainly rely on the binaural cues, namely the ILD and ITD [7]. The ITD in particular is highly dependent on the circumference of the head, which can grow by up to 10 cm from age 5 to adulthood [12]. The maximum ITD averaged across the 26 child datasets is 615 μs whilst the average adult ITD according to Fels [4] is in the range of 800 μs, with the ITA artificial head[1] [13] splitting the difference with an maximum ITD of 723 μs.
The just noticeable differences in ITDs has been found by Aussal et al. [14] to be 16μs while Bomhardt [15] reported results in the range of 30 to 80 μs, depending on the angle of incidence with regard to the observed sound event.
This means that while there are significant and noticeable difference between children and adults, when listening through the ears of an artificial head, they both exhibit similar difference magnitudes, be it with an inverted sign. This is due to the fact, that the anthropometric dimensions of a artificial head are generally smaller than the typical European average, with the ITA artificial residing on the larger side compared to other representatives such as the KEMAR or B&K Type 4128 head and torso simulator.

[1] ITA artificial head dataset available at: http://www.akustik.rwth-aachen.de/go/id/pein/lidx/1

3.2 Elevation localization

The localization performance in the sagittal plane is evaluated using the probabilistic localization model by Baumgartner et al. [8], available as part of the Auditory Modeling Toolbox [16]. The subject specific parameter apart from the HRTF data, such as the listener-specific sensitivity, were left at default values for the prediction. As a target HRTF, the ITA artificial head HRTF was used. The adult datasets evaluated in the same way are available from the ITA individual HRTF database [2] [17].

Figures 1 and 2 show examples of prediction results in the sagittal plane for a child and adult participant respectively. Figure 3 shows a typical result of miss-matched HRTFs, predicting a low localization performance. While some children show similar results to the ones shown in Figure 1, many exhibit significantly worse performance. The resulting polar error rates range from 18% to 60% with quadrant error rates between 10% and 44%, with averages of 40% and 18% respectively. This shows, that while the model predicts a good localization performance for some children, for many, the localization performance is predicted suffer when using a artificial head HRTF. These resulting error magnitudes are comparable to those of adults using non-individualized HRTFs reported by Baumgartner [8].

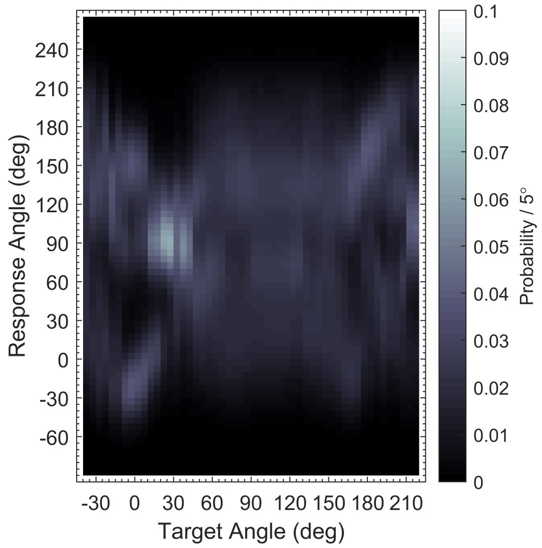

Figure 1. Predicted response likelihood in the sagittal plane for child 6 when using the ITA artificial head HRTF

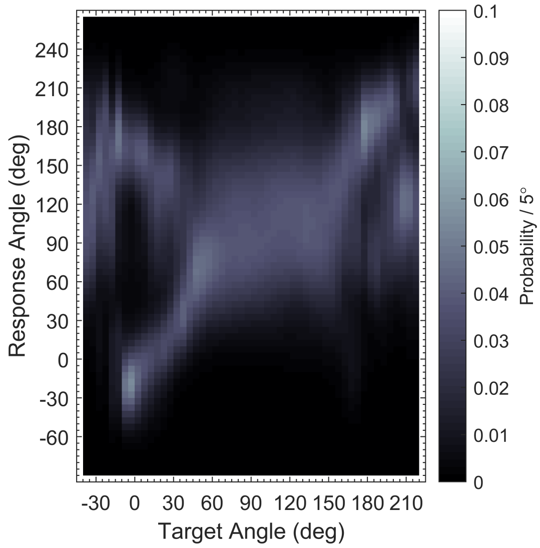

Figure 2. Predicted response likelihood in the sagittal plane for adult 25 when using the ITA artificial head HRTF

4 SUMMARY AND OUTLOOK

In order to investigate objective differences in head-related transfer functions of children and adults, HRTF datasets of 26 first and second grade children were measured using the continuous fast measurement method. As objective measures, ITD differences, as a main binaural cue for localizing sound sources in the horizontal plane, and predicted localization performance in the sagittal plane according to the model by Baumgartner et al. were evaluated. An artificial head was chosen as a reference HRTF, as it is often used instead of the individually measured HRTF in listening tests. In terms of ITD, the differences compared to the artificial head are about the same in magnitude, with children ITDs being smaller and adult ITDs being larger according to their

[2]ITA individual HRTF database available at: http://www.akustik.rwth-aachen.de/go/id/lsly/lidx/1

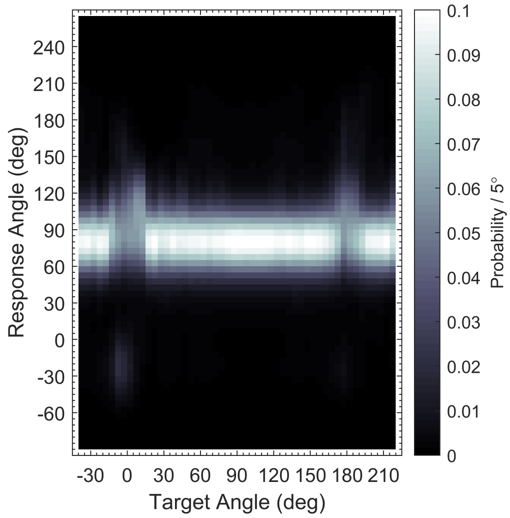

Figure 3. Predicted response likelihood in the sagittal plane for child 5 as an example of poor localization performance using the artificial head HRTF

respective head sizes. The sagittal plane localization performance also yielded similar error magnitudes for the children using an artificial head compared to adults using non-individualized HRTFs. These results suggest that children can hear worse when listening through the ears of an artificial head. However, there are still issues concerning the accuracy of the HRTF datasets, which may skew the results.

It is unlikely that children will sit still during a 3 minute measurement. Thus, even though measures were taken to reduce movement, the HRTF data still contains position uncertainties due to increased head movements compared to measurements of adult subjects. One way to reduce these artifacts is to compensate for the tracked head movement in post processing, which can however result in additional artifacts of its own from the new post-processing overhead. A different approach is to use the 3D scans acquired from the same children to calculate an HRTF dataset using the boundary element method and have them undergo a similar analysis in future works.

REFERENCES

[1] Loh, K.; Fels, J. Speech perception of children and adults in a dual-task paradigm under noisy conditions. Proceedings of the 48th International Congress on Noise Control Engineering, INTER-NOISE 2019 MADRID, 2019

[2] Xu, S.; Li, Z.; Salvendy, G. Individualization of head-related transfer function for three-dimensional virtual auditory display : A review, 12th International Conference on Human-Computer Interaction, (2007), 397–407.

[3] Bomhardt, R., Lins, M.; Fels, J. Analytical ellipsoidal model of interaural time differences for the individualization of head-related impulse responses. Journal of the Audio Engineering Society, 64(11), 2016, 882-894.

[4] Fels, J.; Buthmann, P.; Vorländer, M. Head-Related Transfer Functions of Children, Acta Acustica united with Acustica, 90(5), 2004, 918–927.

[5] Fels, J.; Vorländer, M. Anthropometric parameters influencing head-related transfer functions, Acta Acustica united with Acustica, 95(2), 2009, 331–342.

[6] Prodi, N.; Farnetani, A.; Smynova, Y., FELS, J. Investigating classroom acoustics by means of advanced reproduction techniques, Proceedings of the 122th Audio Engineering Society Convention, Vienna (Austria), 2007.

[7] Blauert, J. Spatial Hearing: The Psychophysics of Human Sound Localization. Cambridge: MIT Press, 1997.

[8] Baumgartner, R.; Majdak, P.; Laback, B. Modeling sound-source localization in sagittal planes for human listeners, The Journal of the Acoustical Society of America, 136(2), 2014, 791–802.

[9] Richter, J.-G.; Behler, G.; Fels, J. Evaluation of a fast HRTF measurement system, AES 140th Convention, Paris, 2016.

[10] Richter, J.-G. Fast measurement of individual head-related transfer functions, Logos Verlag, 2019.

[11] Wightman, F. L.; Kistler, D. J. The dominant role of low-frequency interaural time differences in sound localization. JASA, 91(3), 1992.

[12] Fels, J. From children to adults: How binaural cues and ear canal impedances grow, Logos-Verlag, 2008.

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

[14] Aussal, M.; Alouges, F.; Katz, B. HRTF interpolation and ITD personalization for binaural synthesis using spherical harmonics, Audio Engineering Society Conference: UK 25th Conference: Spatial Audio in Today's 3D World, 2012.

[15] Bomhardt, R. Anthropometric individualization of head-related transfer functions, PhD thesis, RWTH Aachen, 2017.

[16] Søndergaard, P.; Majdak, P. The auditory modeling toolbox, The Technology of Binaural Listening, 2013, 33–56.

[17] Bomhardt, R.; de la Fuente Klein, M.; Fels, J. A high-resolution head-related transfer function and three-dimensional ear model database, Proceedings of Meetings on Acoustics, 29(1), 2016.