HRTF Individualization: Investigations on the Influence of Adapted Binaural Parameters on the Localization

Ramona Bomhardt[1], Marcia Lins[1], Janina Fels[1]
[1] *Institut für Technische Akustik, Medizinische Akustik, 52074 Aachen, Germany,*
Email: {rbo, mal, jfe}@akustik.rwth-aachen.de

Why Do We Need Individualization?

When head-related transfer functions (HRTFs) of artificial heads are used, localization errors are reported very often. This is likely because the interaural level and time differences (ILD and ITD) as well as the spectral cues often fail to coincide with those of the individual listeners. To minimize these differences, the ITD and ILD can be adjusted for the used HRTF dataset. The ITD adjustment can be applied by the model of Bomhardt et Al. [1] and the ILD adjustment by frequency scaling between 1 and 4 kHz. Therefore Lins et Al. investigate the relation between the depth of the head as well as the offset of the ear and the ILD [2]. Based on boundary element method studies on an ellipsoid with human ears, it is found that the large offset of the ears will shift the extrema of the ILD towards the back of the head meanwhile a large depth of the head will rise the maxima of the ILD. Besides this, a large offset will change the hight of the maxima as well but in contrast to the head depth, this effect is less significant.

ILD scaling

The ILD is dominated by two maxima whose position will be shifted towards lateral positions and whose amplitude will be increased for higher frequencies. Due to the fact that the same applies for larger heads, the frequency vector is scaled in the range between the limiting frequencies $f_{lower} = 1$ and $f_{upper} = 4$ kHz:

$$f_{Bézier}(t) = (f_{lower} - f_{upper})\, s \cdot t^2 + (1+s) \cdot (f_{upper} - f_{lower}) \cdot t + f_{lower} \quad (1)$$

For this purpose a quadratic Bézier curve $f_{Bézier}(t)$ in

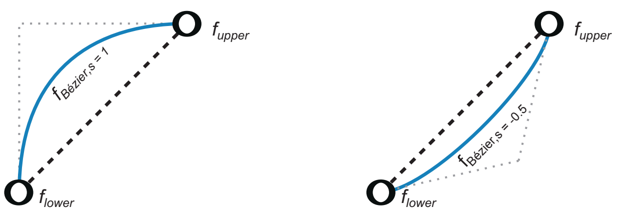

Figure 1: The original frequency vector (dashed line) is linearly spaced from f_{lower} to f_{upper}. The blue curves are the resulting scaled frequency vectors for $s = 1$ (left) and $s = -0.5$ (right).

dependency of f_{lower}, f_{upper} and a scaling factor $s \in [-1 \quad 1]$ is used for scaling [3]. Hereby, the parameter

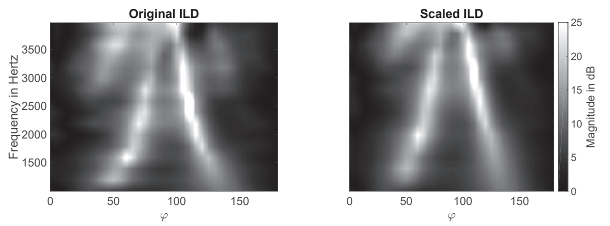

Figure 2: Both images show an individually measured ILD dependent on the frequency and the direction. Meanwhile the left image shows the original ILD, the right image shows the scaled ILD with a scaling factor $s = 0.75$.

$t \in [0 \quad 1]$ is the position on the Bézier curve between f_{lower} and f_{upper} (see figure 1). If the scaled frequency vector $f_{Bézier}$ is plotted over the original linear frequency vector, the curve will be deflected to higher frequencies for a positive and to lower frequencies for negative scaling factor s. The middle point is located at

$$f_{middle} = \frac{1}{2} f_{upper}(s+1) + f_{lower}(1-s) \quad (2)$$

and can be substituted in the formula:

$$f_{Bézier}(t) = (f_{lower} - 2f_{middle} + f_{upper}) \cdot t^2 + 2(f_{middle} - f_{lower}) \cdot t + f_{lower}. \quad (3)$$

The resulting scaled ILD shows in figure 2, in comparison to the original one, that the maxima are wider spaced and their magnitudes are lower. This corresponds to a smaller head. Consequently, it can be assumed that this scaling will lead to a shift of the perceived direction towards the front for head-on presented sources. The monaural localization cues, which are related to resonances of the pinna, are not affected by the scaling. The same applies to lower frequencies where the ITD cues play a major role for localization.

Localization Experiment

The localization performance is evaluated by the proximal pointing method [4] where the subject points with a stick in the direction of a virtual sound source in respect to the center of the head. The stimuli are binaurally played back over headphones using individually measured HRTF and averaged headphone transfer functions. Five directions in the horizontal plane at $\varphi = -40°$, $-50°$, $-60°$, $-70°$ and $-80°$ are tested eight times per direction for five scaling factors $s = [-1, -0.5, 0, 0.51]$ with pulsed pink noise.

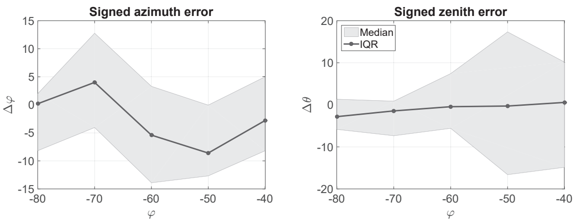

Figure 3: The signed localization error for the azimuth angle $\Delta\varphi$ (left figure) and the zenith angle $\Delta\theta$ (right figure) of the proximal pointing is shown dependent on the five tested horizontal directions φ. The errors are pictured by the median (dark line) and IQR (gray area) of all subjects.

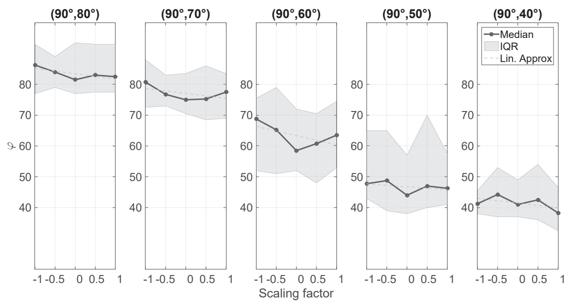

Figure 4: The effect of the scaling is shown for the tested five horizontal directions on the right side of the listener. Meanwhile the scaling factor is given on the x-axis, on the y-axis the pointed azimuth direction is shown by the median (dark line) and interquartile range (gray area) of all subjects. To demonstrate the general tendency, a linear approximated function (dashed line) is displayed as well.

In total, 19 male and 11 female subjects at the age of 26 ± 3 (mean ± std.) participated.

Evaluation of the Pointing Method

For the evaluation of the azimuth error $\Delta\varphi$, the front-back confusions in the horizontal plane, which are caused by the binaural static scene reproduction without room acoustics, are corrected by mirroring the subjects responses from the back to the front. The results in figure 3 show especially for the signed azimuth error $\Delta\varphi$ that deviations from the original position can be detected, although the subjects and loudspeakers are carefully aligned. Nevertheless, the decrease for frontal and lateral directions of the unsigned azimuth error $|\Delta\varphi|$ as well as the deviations of the signed azimuth error $\Delta\varphi$ are in line with the results of Bahu et Al. [4].

Meanwhile the interquartile range (IQR) for the unsigned azimuth error $\Delta\varphi$ is approximately 15° across all tested directions, the IQR of the signed zenith error $\Delta\theta$ will rise for frontal directions. However, the median of the signed zenith error $\Delta\theta$ rises only weakly for sources which are further ahead.

Influence of the Scaling on the Localization

The results of the listening experiment show the expected tendency that a smaller head $s > 0$ leads to a horizontal shift of the perceived sound source position φ towards the front and accordingly a smaller head $s < 0$ leads to a shift towards $-90°$ in figure 4. Due to the scaling, the perceived direction can be moved by 5°. However, the results point out that this effect is particularly pronounced at the directions $\varphi = -70°$ and $-60°$. Nevertheless, in some cases such as direction $\varphi = -70°$ with a scaling factor $s = 1$, the perceived sound source location is not always shifted towards the front for increasing scaling factors $s > 0$. As already detected in the observation of the localization error, the deviations of the subject's responses differ especially for $\varphi = -60°$ and $-50°$ for all tested scaling factors. Additionally, it can be observed that the scaling does not affect the front-back confusions and the zenith angle.

Conclusion

The Bézier curve scaling of the frequency vector allows the scaling of a limiting frequency range meanwhile the rest of the spectrum remains the same. This fact allows frequency dependent modifications. Because of the fact that the ILD is important for the localization for frequencies above 1.5 kHz [5] and above 4 kHz the ILD is influenced by the resonance of the cavum conchae, this range is scaled in the listening test. Although, this is a very limited range, the listening experiment shows that the scaling can be used for the correction of a mismatch between the subject and the used HRTF dataset. In other words, the scaling effects a shift of lateral sources.

References

[1] Bomhardt, R.; Fels, J. (2014): Analytical interaural time difference model for the individualization of arbitrary Head-Related Impulse Responses. Audio Engineering Society Convention 137. USA, Los Angeles, 9-12.10.2014. Audio Engineering Society. 137. Edition. USA, New York

[2] Lins, M.; Bomhardt, R.; Fels, J. (2016): Individualisierung der HRTF (1): Ein Ellipsoidmodell zur Anpassung von interauralen Pegeldifferenzen. DAGA 2016, 42. Jahrestagung für Akustik. Germany, Aachen, 14.-17.03.2016. Deutsche Gesellschaft für Akustik e.V.

[3] Middlebrooks, J. C. (1999): Virtual localization improved by scaling non-individualized external-ear transfer functions in frequency. In: The Journal of the Acoustical Society of America 106, S. 1493.

[4] Bahu, H.; Carpentier, T. Noisternig, M.; Warusfel, O. (2016): Comparison of Different Egocentric Pointing Methods for 3D Sound Localization Experiments. In: Acta Acustica united with Acustica 102 (1), S. 107?-118.

[5] Wightman, Frederic L.; Kistler, Doris J. (1992): The dominant role of low-frequency interaural time differences in sound localization. In: The Journal of the Acoustical Society of America 91, S. 1648.