Near-field HRTFs

Tobias Lentz

Institute of Technical Acoustics, RWTH Aachen University, 52066 Aachen, Germany,Email: tle@akustik.rwth-aachen.de

Introduction

For creating virtual sound scenes, spatially distributed sources are essential to achieve a high degree of realism. One solution is the binaural synthesis [1, 2]. To provide virtual sources at any position, a HRTF database is required for adding the spatial cues to audio signals. Beside the spatial information, the distance of the source to the listener is very important, especially in a dynamic reproduction system, where the listener has the possibility to change his position and thus the distance to the virtual source. In case of simulating sources located closer to the listener's head (in the following named as proximal region) the simple level-dependent distance correction according to the spherical wave attenuation is not sufficient for a realistic representation. For this reason a listening test was conducted to evaluate, in a first step, the area around the head in which special near-field HRTFs are required. In a second step, the spatial resolution has to be specified to find a good tradeoff between the postulated inaudible artifacts and the amount of data which has to be present in the system.

Listening Test

To evaluate the perceptual perimeter of the proximal region, a simple listening test was performed in the semi-anechoic chamber, examining the ranges where different near-field HRTFs have to be applied. The listeners were asked to compare signals from simulated HRTFs with those from correspondingly measured HRTFs on two criteria, namely the perceived location of the source and any coloration of the signals. In this listening test only the horizontal plane was taken into account. A variation of the elevation angle was omitted as the major differences related to a variation of source distance can be found at lateral angles where the maximum distance between both ears occurs. At any elevation angle outside the horizontal plane, the interaural distance decreases and also the distance-dependent differences [3].

The HRTFs for the listening test were measured at distances between 0.2 m and 1.0 m with a spacing of 0.1 m. Between 1.0 m and 2.0 m, the spacing was chosen to 0.25 m. The azimuth angle was modified in steps of 15 degree between the frontal (0 degree) and rear (180 degree) direction. A preliminary test showed, as expected, that differences become more significant in the lateral area. To reduce the number of trials, the tested angles were limited to 0 degree, 45 degree, 75 degree, and 90 degree. The simulated HRTFs were prepared from far-field HRTFs (measured at a distance of 2 m) with a level correction applied to both channels.

From the comments of all 9 listeners, for each distance and azimuth, a coloration would earn 5 points, while a change of the perceived location earned 10 points. Figure 1 shows the result. The nearer boundary line defines where huge audible differences and a change of the perceived location of the sound were detected. The boundary line being further away from the head is the boundary where any audible difference such as slight coloration was detected. As expected, differences between the simulated distance and the original distance were reported mainly for stimuli in the lateral region. Already at a distance of 1.5 m at 90 degree, slight differences were audible. The first point where a detection of differences occurs defines the boundary line between the proximal and distal region (far-field). Within the proximal region, a near-field representation is necessary by using special HRTFs. The distal distance of 1.5 m used here differs from the value (1.0 m) which is to be found in literature (i.e., [3]). The measured differences between sources at 1.5 m and 1.0 m are not very significant, which is presumably the reason for defining the upper boundary of the proximal region at 1.0 m. But to avoid any audible artifacts at the transition of the two regions, the perception-based boundary of 1.5 m is used here as the basis for the HRTF database design.

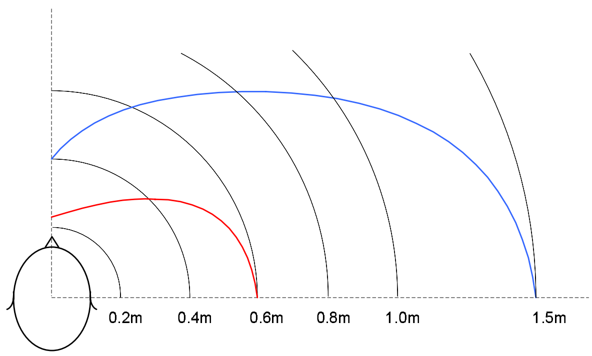


Figure 1: Limits of noticeable differences between near-field and far-field HRTFs.

For the synthesis database, the HRTFs of the ITA artificial head were measured at distances of 0.2 m, 0.3 m, 0.4 m, 0.5 m, 0.75 m, 1.0 m, 1.5 m, and 2.0 m. The spatial resolution is chosen to 1 degree for the azimuth angle and 5 degree for the elevation angle. The closer spacing between measurements at distances near to the head has to be applied to minimize any comb-filter effects and irregularities concerning the perceived distance. Consequently, the generated HRTF database covers all requirements for a near-to-head source imaging system. It should be noted that the system uses the HRTFs of the full sphere because the ITA head has asymmetrical pinnae and head geometry. Non-symmetrical pinnae have positive effects on the externalization of the generated virtual sources [4].

Distance Interpolation

Some further considerations are necessary if a dynamic solution is desired. A variation of the distance induces an interpolation between two HRTFs at different measured distances to realize an arbitrary virtual distance of the source. Beyond the proximal region, the ILD caused by lateral sound incidence from a source in the far-field is dominated by the head shadowing effect not by the distance difference between the ears. The level difference between both ears for lateral sound incidence becomes important with a decreasing distance to the head and can not be neglected. Due to the distance dependency of the level difference, the interpolation between distances is more complex in this case. The time delay or distance difference respectively, of each ear related to the frontal direction as a function of azimuth and elevation is an useful value for considerations about distance interpolation (see Figure 2). The time difference related to the frontal direction will be entitled further as FTD. The summation of the FTD absolute value of both ears is identical to the ITD. In contrast to the ITD which is rotational symmetric to the point ($\varphi = 180$ degree, $t_{ITD} = 0$ s), the FTD shows a strong asymmetry which reflects the indirection of the sound waves to the contralateral ear. Furthermore the FTD calculated by using the spherical model is a good approximation to the measured values.

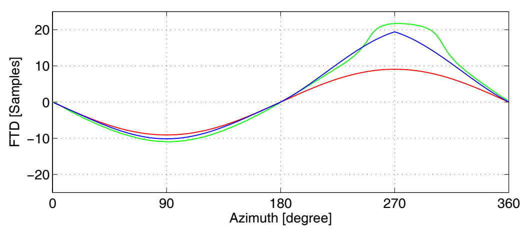
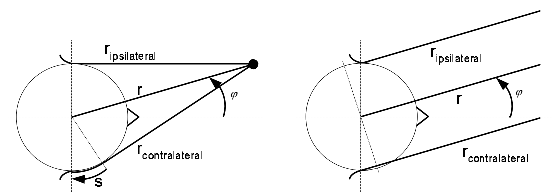


Figure 2: Time difference of the ear signals related to the frontal direction of sound incidence as a function of the head angle. The distance between source and the center of the head is 0.4 m. Measurement = green curve, exact model = blue curve, and simple model = red curve.



(a) Near-field (b) Far-field

Figure 3: Spherical models.

$$r_c = \varphi_1 r_H + \varphi_2 r_H + \sqrt{r^2 - r_H^2} \qquad (1)$$

$$r_i = \begin{cases} \varphi_1 r_H + \varphi_2 r_H + \sqrt{r^2 - r_H^2} & \varphi \geq a \\ \sqrt{(r - r_H \sin(\varphi_2))^2 + (r_H \cos(\varphi_2))^2} & \varphi < a \end{cases} \qquad (2)$$

$$a = \frac{\varphi_1}{\cos(\vartheta)} \quad ; \quad \varphi_1 = \arcsin(r_H/r) \quad ; \quad \varphi_2 = \varphi\cos(\vartheta)$$

Figure 4 shows the level at the ears as a function of distance for different angles of sound incidence. The diameter of the head is chosen to 18 cm. The level is calculated using two different spherical models (see Figure 3). Equation 1 and 2 represent the exact solution for any distance with respect to the assumption of a spherical head. This is delineated by the the solid curves in Figure 4. It can be seen, that the exact model fits well to the measured values (labeled with $*$ and **o**) at all distances, whereas the approximation gives acceptable results at distances of 2 m and beyond at least at the contralateral ear. All distances between the measured values can be interpolated using the exact spherical model, but due to the distance dependency of the FTD (see Figure 2) the interpolation has to be done separately for magnitude and phase to avoid audible artifacts.

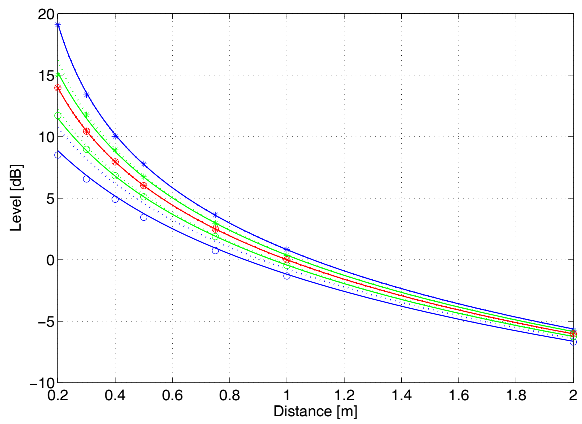


Figure 4: Level at the ears as a function of distance for different angles of sound incidence; $\varphi = 0$ (red, center), $\varphi = 45$ (green) and $\varphi = 90$ degree (blue, outer curve). The level is calculated using the exact (bold curves) and the simplified (dotted curves) spherical model. The diameter of the head is chosen to 18 cm. Points marked with $*$ and **o** represent the real level of the measured HRTFs at 100 Hz.

References

[1] Lentz, T.: *Dynamic Crosstalk Cancellation for Binaural Synthesis in Virtual Reality Environments*, Journal of the Audio Engineering Society 54 (2006), 283–294

[2] Lentz, T., Schröder, D., Vorländer, M., Assenmacher, *Virtual Reality System with Integrated Sound Field Simulation and Reproduction*, EURASIP Journal on Advances in Signal Processing, Special Issue on Spatial Sound and Virtual Acoustics, (2007)

[3] Brungart,D.S., Durlach N.I., Rabinowitz, W.M.: *Auditory localization of nearby sources. II. Localization of a broadband source*, Journal of the Acoustical Society of America 106 (1999), 1956–1968

[4] Brookes, T., Treble, C.: *The Effect of Non-Symmetrical Left/Right Recording Pinnae on the Perceived Externalisation of Binaural Recordings.*, Proceedings of the 118th Audio Engineering Society Convention