HRTF measurement and HRTF simulation: a comparison of two approaches

Lorenz Kroener[1], Christophe Langrenne[2]

[1] *French-German Research Institute of Saint-Louis, 68300 Saint-Louis, France, Email: lorenz.kroener@isl.eu*

[2] *Conservatoire national des arts et métiers, 75141 Paris, France, Email: christophe.langrenne@lecnam.net*

Introduction

Head related transfer functions (HRTFs) are highly personal functions and are necessary for binaural rendering. The measurement of individual HRTFs is time consuming and a complex procedure, hence for listening tests the subjects' HRTFs are often not available and generic HRTFs have to be used [1]. Different approaches have been proposed to avoid the measurement procedure, but still obtain individualized HRTFs: HRTF fitting, resp. individualization algorithms are performed to tune generic HRTFs to the subjects' individual cues [2]. In other approaches subjects get trained to generic HRTFs [3]. Selecting the HRTF out of a set of HRTFs, with which a subject performs a given tasks the best, is proposed in [4]. In [5] and [6] anthropometric data are measured to estimate individual HRTFs. The simulation of individual HRTFs in [7] is based on magnetic resonance imaging techniques. In this article we obtain individual HRTFs, using a low cost 3D scanner and numerical simulation tools, and compare them to actual measurements.

HRTF Simulation

Preparation of simulation

With a low budged 3D Scanner (3D Systems "Sense") two dummy heads are scanned. Both dummy heads are based on the ISL dummy head "Harry". Harry can be equipped with Type 3.4 ear simulators, also referred to as "Harry3.4" or, as shown in Figure 1, with Type 3.3 ear simulators, also referred to as "Harry3.3". The 3D models of these dummy heads are used for the numerical simulation, which is computing the HRTF for each dummy head based on its 3D model. Prior to the simulation, the 3D meshes are optimized and verified to meet the criteria of the simulation software. This includes the suppression of scanning artifact and the compensation and correction of errors in the meshes. The meshes are smoothed with a shape preserving algorithm in order to control the length of the edges. Finally re-meshing techniques are used to provide a regularly meshed object, but more imported to meet the spatial sampling criteria. The spatial sampling criteria requires, that at least 6 elements per wave length are provided by the mesh. The upper simulation frequency is set to 20 kHz, thus the maximum element size has to be inferior to 2.86 mm. For proper simulation results the limit is set to 2.80 mm. Figure 2 is showing the 3D meshes, as they are ready for simulation, and in particular detailed views of the right ear simulators of Harry3.4 and Harry3.3. The differences in the shapes of the two ear simulators are clearly distinguishable.

Figure 1: Type 3.3 ear simulators mounted on dummy head Harry. The ear simulators can be replaced by Type 3.4 simulators.

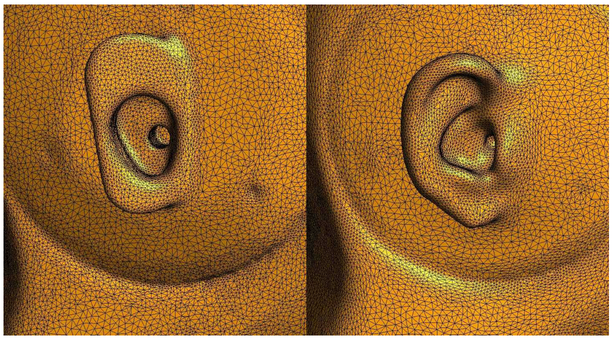

Figure 2: Detailed view of the scanned meshes (verified & optimized) of the dummy head. Left: Type 3.4 ear simulator. Right: Type 3.3 ear simulator.

Simulation process

The simulation software is based on the Multi-Level Fast Multipole Method (MLFMM) [8]. In order to reduce the computing time, the HRTFs are determined using Helmholtz's reciprocity principle. This principle states that the positions of the source and receiver can be exchanged without affecting the results [9]. The simulation is simplified by assuming a rigid body reflection of the acoustic waves at the surfaces of the 3D models and not taking into account the exact impedance values of the different surfaces of the human head, e.g. hair and skin.

Each mesh is composed of around 150 000 elements, satisfying the spatial sampling criteria at 20 kHz. The HRTFs are simulated at 1600 frequencies, using logarithmic incrementing step sizes, ranging from 20 Hz to 20 kHz. The MLFMM is implemented in Fortran using OpenMP and the simulation software is designed as a parallel processes. The simulation is launched at the LMSSC laboratory at the CNAM in Paris, on a computing machine, providing 16 dual core CPUs.

HRTF Measurement

The HRTF measurements are conducted in a fully anechoic chamber, e.g. the chamber's floor, ceiling and walls are entirely equipped with triangular prism shaped acoustic absorbers. The chamber is a room-in-room construction to reduce the transmission of structure-borne noise and environmental noise into this chamber. The chamber‘s dimensions are 5.00 m × 6.40 m × 2.60 m (W×L×H). The absorbers inside the chamber have a total height of 0.40 m, reducing the chamber's dimension by 0.80 m, leading to a usable acoustical room of 4.20 m × 5.60 m × 1.80 m (W×L×H) and a volume of 42.34 m^3. The chamber's acoustical properties are verified prior to the HRTF measurements. The reverberation time RT60 is determined in octave bands between 63 Hz and 8 kHz. Table 1 summarizes the results of the RT60 measurement, which can be concluded to provide good measurement conditions.

Table 1: RT60 of the anechoic chamber.

Frequency [Hz]	63	125	250	500
RT60 [s]	0.24	0.20	0.17	0.06
Frequency [Hz]	1k	2k	4k	8k
RT60 [s]	0.06	0.05	0.05	0.05

A spectrum analyzer (Stanford Research SR780) is used as signal generator and for the calculation of the transfer functions. The test tone, a logarithmic sinusoidal sweep ranging from 20 Hz to 20 kHz, generated by the SR780, is sent through a power amplifier (DaytonAudio MA1260) to a loudspeaker (JBL Control 1 Pro). The acoustic signal, captured by the receiver, is passed via a signal conditioner (B&K Type 5935L) to the input of the spectrum analyzer. The speaker is positioned on the chamber's longitudinal axis at half chamber's height and at a distance of 1.30 m from the chamber's back wall, facing the receiver position. The receiver position is defined on the chamber's longitudinal axis at half chamber's height and at a distance of 3.00 m to the speaker's position.

Initially, the transfer function of the measurement chain is obtained. Therefore a reference microphone (B&K Type 4192) and a microphone amplifier (B&K Type 2669) are mounted at the receiver position and the transfer function of the measurement setup is determined.

In the following, the transfer functions of the left and right ear of Harry3.4 and Harry3.3 are measured, one after the other. The previously used reference microphone is replaced by the dummy head, which is equipped with a left and right ear canal. The left, resp. right ear simulators are connected to the entrance of the left, resp. right ear canals. At the end of each ear canal a microphone (B&K Type 4192) and a microphone amplifier (B&K Type 2670) are capturing the incoming sound. The positions of the microphones in the dummy head correspond to the positions of ear drums in the human head. The dummy head is mounted on a turntable and aligned such that the midpoint of the dummy head's interaural axis, also called center of head (CoH), is placed at the receiver position. The transfer functions are measured in the horizontal plane (elevation angle $\theta = 0°$) at 16 angles of incidence (azimuth angle $\phi = 0°, 22.5°, 45°, \ldots, 337.5°$). The underlying right hand coordinate system is centered in the CoH, orientating its x-axis to the facial side of the dummy head.

The transfer functions of the dummy head are divided by the transfer function of the measurement chain to remove the characteristics of the measurement chain. Furthermore all non-directional cues, e.g. the ear canal resonance, are equalized by removing the mean value over all angles. This leads to the directional transfer functions (DTFs) [10], which are used as HRTFs in the following.

To reduce the noise floor and to increase the SNR it is assumed that the left and right HRTF are symmetrical to each other. Hence the average between the left HRTF and the mirrored right HRTF is calculated to obtain a less noisy HRTF of the left ear.

Results

HRTF

The left ear HRTFs with an improved SNR are considered in the following. Figure 3 and Figure 4 are showing the magnitude of the measured HRTFs, while Figure 5 and Figure 6 are showing the magnitude of the simulated HRTFs. The abscissa denotes the frequency in Hertz, the ordinate denotes the azimuth angle in degree. The magnitude of the HRTF is color coded and its unit is given in Decibel.

By visual inspection, differences between the measurement and simulation can be identified below 100 Hz. There, peaks and notches occur in the measured HRTFs which are due to the low SNR of the test tone in this frequency range. The loudspeaker transmits too few energy below 100 Hz to provide a sufficient strength of the signal. Further, the resonance of the measurement chain around 150 Hz is not perfectly compensated by the reference measurement, leading to peaks and notches around 150 Hz. Common characteristics of all four HRTFs are first, the general shape of the magnitude between 1 kHz and 10 kHz. Second, the increase of the magnitude at 270° (contralateral ear) over the entire frequency range, which is explained by the summation of the two waves diffracted around the head. And last, the maximum of the magnitude at 45° between 3 kHz and 4 kHz, due to the ear canal resonance. The magnitudes of the measured and the simulated HRTFs match quite well, when comparing the same dummy head configurations.

RMSE

The root-mean square error (RMSE) between two HRTFs, denoted by $\mathrm{HRTF_I}$ and $\mathrm{HRTF_{II}}$, is obtained by calculating for each azimuth angle the RMSE over all frequencies of the ratio between $\mathrm{HRTF_I}$ and $\mathrm{HRTF_{II}}$. Then calculating the RMS value over the azimuth angles, c.f. equation 1.

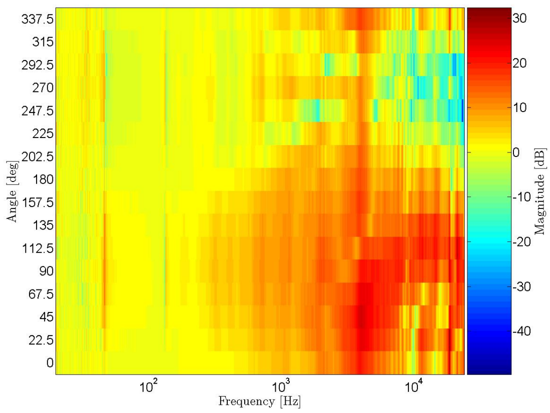

Figure 3: Measured HRTF of dummy head Harry3.4.

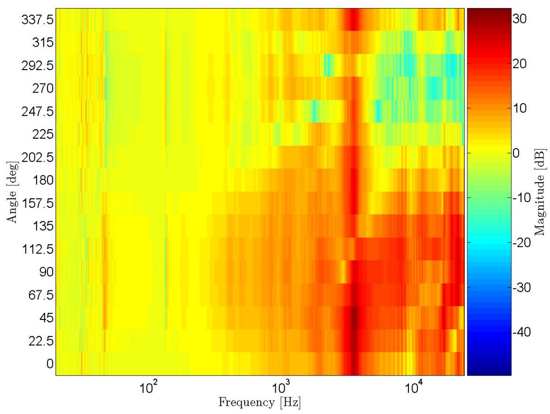

Figure 4: Measured HRTF of dummy head Harry3.3.

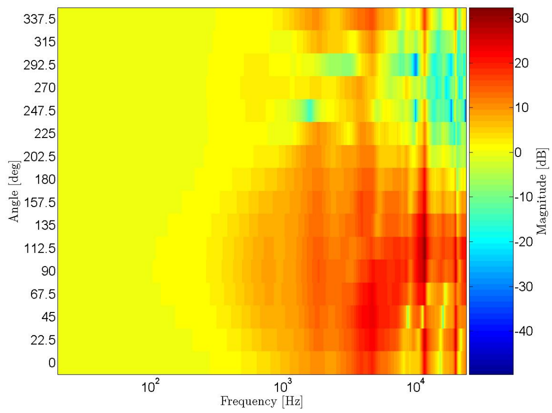

Figure 5: Simulated HRTF of dummy head Harry3.4.

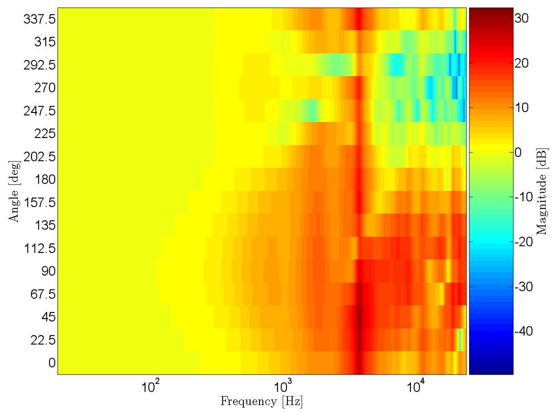

Figure 6: Simulated HRTF of dummy head Harry3.3.

$$\epsilon_{\mathrm{RMS}} = \sqrt{\frac{1}{N_\phi}\sum_\phi\left(\sqrt{\frac{1}{N_f}\sum_f\left(\frac{HRTF_I(f,\phi)}{HRTF_{II}(f,\phi)}\right)^2}\right)^2} \quad [\mathrm{dB}] \tag{1}$$

The RMSEs between each of the four HRTFs are shown for the entire frequency range in Table 2. The RMSE of two consecutive measurements of the same head configuration (Harry3.4m, resp. Harry3.3m), equals 1.66 dB, resp. 1.86 dB, which is contrary to what one would expect, namely 0 dB. Even though the measurement setup was kept unchanged between those measurements, measurement uncertainties lead to these non-zero values. The numerical simulations are time invariant processes and each simulation run leads to equal results, hence the RMSE between two consecutive simulations of the same configuration (Harry3.4s, resp. Harry 3.3s) is 0.00 dB. The pairs measurement – simulation for each head configuration Harry3.4m – Harry3.4s, resp. Harry3.3m – Harry3.3s lead to RMSEs (3.31 dB, resp.3.11 dB) which are up to 0.93 dB lower than the RMSE between the measurements of two different head configurations (4.05 dB). Hence the HRTFs of the simulations equal to the HRTFs of the corresponding measurements more than the HRTFs of the measurements of the different dummy heads equal to each other. In contrast, the RMSE between the two simulations (Harry3.4s – Harry3.3s; 3.29 dB) is 0.02 dB lower than the RMSE between the two measurements (Harry3.4m – Harry3.3m; 3.31 dB). This might be explained by the measurement uncertainty.

Table 2: RMSE values in decibel for the entire frequency range. (s): simulated; (m) measured.

Harry	3.4m	3.4s	3.3m	3.3s
3.4m	1.66	3.31	4.04	3.61
3.4s	–	0.00	4.38	3.29
3.3m	–	–	1.86	3.11
3.3s	–	–	–	0.00

Recalling that the simulation is simplified regarding the acoustic impedance of the model. For evaluation purpose we now split the entire frequency range into a lower range ($f \leq 6$ kHz) and an upper range ($f > 6$ kHz). For both ranges the RMSE values are calculated as done before. The previously discussed characteristics of the entire frequency range reappear in the lower resp. upper frequency range, c.f. Table 3, resp. Table 4. The main differences between Table 2 to Table 4 are the absolute values of the RMSEs. All nonzero elements of Table 3 are smaller than those of Table 2 and all nonzero elements of Table 4 are larger than those of Table 2. Hence the RMSEs in the lower frequency range is lower than the RMSE in the upper frequency range. The overall number of 1600 frequencies is split into 1285 frequencies below 6 kHz and 315 frequencies above 6 kHz. The upper frequency range with a fewer number of frequencies contribute a higher error than the lower frequency range, which comprises the issue of the low SNR issue below 100 Hz. This difference can be explained by the assumption made on the acoustic impedance.

Table 3: RMSE values in decibel for the lower frequency range ($f \leq 6$ kHz). (s): simulated; (m) measured.

Harry	3.4m	3.4s	3.3m	3.3s
3.4m	1.46	2.11	2.68	2.52
3.4s	–	0.00	3.21	1.91
3.3m	–	–	0.81	2.38
3.3s	–	–	–	0.00

Table 4: RMSE values in decibel for the upper frequency range ($f > 6$ kHz). (s): simulated; (m) measured.

Harry	3.4m	3.4s	3.3m	3.3s
3.4m	2.22	5.89	6.92	5.93
3.4s	–	0.00	7.32	5.94
3.3m	–	–	3.60	5.25
3.3s	–	–	–	0.00

Acquisition time

During the measurement the acquisition time for one sine sweep is 178 s. For one dummy head, it takes 98 min to obtain the HRTF of the left and right ear ($2 \cdot 178$ s/angle) at 16 angles, counting 12 s for rotating the turntable between two azimuth angles. To obtain a proper 3D mesh without any severe artifacts, the 3D scan of one dummy head takes about 15 min. The time for optimization depends on the quality of the mesh which differs largely from mesh to mesh. The numerical simulation takes 180 min for each set of HRTF. Thanks to the reciprocity principle the entire acoustic field around the dummy head is simulated at once. The time for obtaining the HRTF by the simulation approach is nearly constant, independent of the number of required incidence angles, while for the measurement approach it is linear dependent to the number of required incident angles.

Conclusion

We used dummy heads to compare simulated HRTFs and measured HRTFs of two different anthropometric data sets. The HRTFs were evaluated against each other using an objective method by RMSE calculation. We showed, that the impedance of the object has to be considered during the simulation process. Otherwise higher deviation between the simulated HRTF and the measured HRTF appear in the upper frequency range. Furthermore, in terms of acquisition times the measurement procedure is preferable for a small number of incidence angles, while the simulation is preferable for a large number of incidence angles.

References

[1] Frantisek, R. and Saturka. F.: Alternatives to HRTF measurement. 35th International Conference on Telecommunications and Signal Processing (TSP), IEEE, 2012.

[2] Wenzel, E. M. and Arruda, M. and Kistler, D. J. and Wightman, F. L.: Localization using non-individualized head-related transfer functions. J. Acoust. Soc. Am. 94 (1993), 111–123

[3] Blum, A and Katz, B. F. G. and Warusfel, O.: Eliciting adaptation to non-individual HRTF spectral cues with multi-modal training. Proceedings of Joint Meeting of the German and the French Acoustical Societies (CFA/DAGA'04), Strasboug, France, 2004

[4] Seeber, B. and Fastl, H.: Subjective selection of non-individual head-related transfer functions. International Conference on Auditory Display, Boston, (2003), 259-262

[5] Zotkin, D. Y. N., et al.: HRTF personalization using anthropometric measurements. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, IEEE, 2003.

[6] Torres-Gallegos, E. A., and Orduna-Bustamante, F. and Arámbula-Cosío, F.: Personalization of head-related transfer functions (HRTF) based on automatic photo-anthropometry and inference from a database. Applied Acoustics 97 (2015), 84-95.

[7] Mokhtari, P., et al.: Computer simulation of HRTFs for personalization of 3D audio. Second International Symposium on Universal Communication, IEEE, 2008.

[8] Gumerov, N. A. and Duraiswami, R.: Fast Multipole Methods for the Helmholtz Equation in Three Dimensions. Elsevier, Oxford, 2004.

[9] Gumerov, N. A. and O'Donovan, A. E. and Duraiswami, R. and Zotkin, D. N.: Computation of the head-related transfer function via the fast multipole accelerated boundary element method and its spherical harmonic representation. J. Acoust. Soc. Am. 127(1) (2010), 370-386

[10] Larcher, V.: Techniques de spatialisation des sons pour la réalité virtuelle. Dissertation, Paris 6, 2001.