APPLYING EXTRAPOLATION AND INTERPOLATION METHODS TO MEASURED AND SIMULATED HRTF DATA USING SPHERICAL HARMONIC DECOMPOSITION

Martin Pollow

Institute of Technical Acoustics
RWTH Aachen University
Neustraße 50, 52066 Aachen, Germany
mpo@akustik.rwth-aachen.de

ABSTRACT

Head-Related Transfer Functions (HRTFs) describe a persons or a dummy-heads perception of sound for different angles of incidence. These data are usually simulated or measured on discrete points in space on a fixed distance. Anywhere in between these points the HRTF can either be approximated by one of its neighbors or interpolated from the other measurement points. One way to do so is to approximate the HRTFs as a continuous mode-limited function by spherical harmonic decomposition, allowing to get a function value for any desired direction.

Using the reciprocity principle the incident sound field of the HRTF can be formulated as an exterior radiation problem. With a given set of HRTF data on a fixed distance from the center of the head, this approach allows to extrapolate the range of the HRTF (distance from head to sound source) to larger or smaller values, making it possible to e.g. calculate near-field HRTFs from measurement data obtained on a fixed radius. In this paper, these known extrapolation and interpolation methods will be applied to both a simulated and measured set of HRTF data.

1. INTRODUCTION

Humans have the remarkable ability to localize the angle of sound incidence with high accuracy. Besides interaural level and time differences, there is also variation in the spectral components of the signal due to diffraction and scattering on the upper part of the human body [1]. These directional dependent Head-Related Transfer Functions (HRTFs) can be either measured or simulated, both being a non-trivial and time-consuming task. HRTFs can be determined from individuals or from dummy-heads, the latter being used more often due to their well-known geometry and their ability to be absolutely immobile during extensive measurements. The resulting information is usually a set of discrete points on a sphere around the head, that is used for arbitrary angles of incidence by nearest neighbor search or interpolation, depending on the required processing speed. Beside the directional dependent ability to localize sound, listeners are also able to distinguish between remote and close sources. For lateral incidence the distance sensitivity is higher, showing audible range dependent differences within a radius of 1.5 m [2]. Employing these range dependent differences of auditory impressions enhances the quality of auralization, especially in a dynamic reproduction system where the listener can change the distance to the sound source.

2. STATE OF THE ART

To enhance the level of realism in the virtual reality environment, Lentz measured the HRTFs of a dummy-head on various different radii from 20 cm to 2 m. These sets of data are currently in use in a real-time virtual scene rendering system at RWTH Aachen University. A fast nearest-neighbor search for both angle of incidence and range of the HRTF is used to reach the required performance. Lentz also developed a method for geometric interpolation of the measured HRTFs [3].

Duraiswami et al. show in [4] how to use spherical harmonic decomposition for interpolation and extrapolation of HRTFs. A sufficiently dense spherical sampling grid on a fixed radius allows to decompose the HRTFs into a set of coefficients in terms of spherical harmonic base functions. This data set can be utilized to gain a value of the HRTF for any arbitrary direction (interpolation) and to scale the HRTF to different ranges (extrapolation) in order to get a solution for the HRTFs for all points outside a confined space containing all significant sources (such as scattering sources) that have an impact on the HRTF.

3. CURRENT CONTRIBUTION

In this paper this spherical harmonic decomposition approach is applied to different HRTF data sets. The first is Lentz' measured set of HRTF data of the dummy-head from the Institute of Technical Acoustics in Aachen [3]. The second set is a simulation result using the Boundary-Element-Method to derive the HRTF of a commercial dummy-head. Both sets were evaluated at distances of 2 m, 1 m, 50 cm and 30 cm on a equiangular sampling grid of 5° steps for azimuth and elevation. With this geometry 2522 unique data points on the sphere exist, with a significant higher point density close to the poles.

The measurement data at distances closer than 2 m were obtained on the same dummy-head with a different measurement setup and do not cover the full sphere. Also the BEM simulation for 30 cm also has some missing parts due to intersections of the sphere with the torso of the dummy-head. All missing information was removed before evaluation in spatial domain. The simulation was done only up to 6 kHz as computational constraints did not allow to calculate any higher without loosing accuracy of the used mesh.

As both measurement and simulation data at a radius of 2 m cover the full sphere, these sets were used as reference values to perform interpolation and extrapolation.

4. THEORY

Using reciprocity, HRTFs can be formulated as exterior radiation problems [5]. With a known volume velocity on the opening of the ear channel, the resulting radiation anywhere outside a confined space of all contributing (scattering) sources can be calculated employing the Helmholtz equation and the Sommerfeld radiation condition [4] [6]. This spherical function for a fixed radius can be decomposed into spherical harmonic coefficients that are equal to the decomposed HRTF.

4.1. Order constrained spherical harmonics

The maximal order is to be chosen by observing the condition number of the matrix of spherical harmonic base functions sampled on the point of data. For the used sampling grid of 5° steps a maximum order of $n_{\max} = 35$ was found to create accurate results with a well-invertible matrix Y that contains all sampled spherical harmonic base functions as column vectors. The orthonormal spherical harmonic base functions can be defined for discrete directions as

$$Y_n^m(\theta_i, \phi_i) = \sqrt{\frac{(2n+1)}{4\pi}\frac{(n-m)!}{(n+m)!}} \cdot P_n^m(\cos\theta_i) \cdot e^{jm\phi_i} \quad (1)$$

with n being order and m degree, θ_i and ϕ_i the angles of the ith sample. $P_n^m(x)$ denotes the associated Legendre function.

4.2. Sampling weights

The equiangular sampling used for this computation is a non-regular sampling distribution that has a much higher grid density in vicinity of the poles. To use this grid for the computation these points of the grids have to be weighted according to the area they cover. The weights can be calculated by taking the first row of the matrix resulting from $\sqrt{4\pi}Y^\dagger$, with $Y^\dagger$ being the Moore-Penrose-Inverse (pseudo-inverse) of the base function matrix Y. The weight vector w is purely real and positive for a suitable sampling schema and the sum over its elements equals to 4π, the solid angle over of the full sphere.

4.3. Discrete spherical harmonic transform

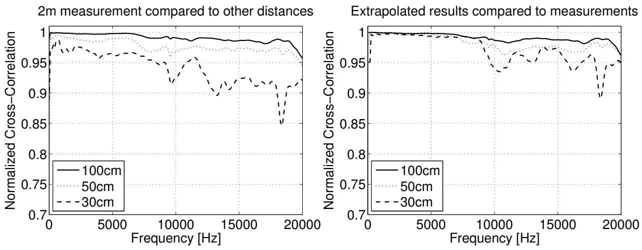
Using these values for the weights allows to perform the discrete spherical harmonic transform with respect to the uneven sampling grid [7]. The vector f denotes the spatially sampled spherical function, the resulting vector $\hat{f}$ its spherical harmonic coefficients:

$$\hat{f} = \left(Y^T \mathrm{diag}\{w\} Y\right)^{-1} Y^T \mathrm{diag}\{w\} \cdot f \quad (2)$$

This forward transformation is an approximation as possible higher orders of the spherical function result in aliasing (cf. [8]).

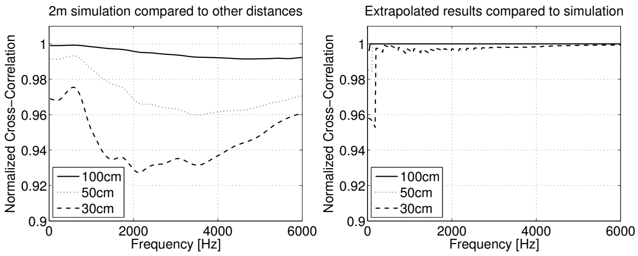
The inverse transform, however, is exact and can be expressed as simple matrix multiplication:

$$f = Y \cdot \hat{f} \quad (3)$$



without extrapolation with extrapolation

Figure 1: Normalized cross-correlation of the measurement in 2 m distance and the measurements at a closer distance



without extrapolation with extrapolation

Figure 2: Normalized cross-correlation of simulation results in 2 m distance and the simulation results at a closer distance

4.4. Performing interpolation and extrapolation

As the spherical harmonic coefficients of the transformed HRTF are independent of the used sampling grid, interpolation can simply be done by using a different sampling for the spherical harmonic matrix Y. As the inverse transform into the spatial domain is exact, the interpolation is perfect for band-limited functions.

Let $\hat{f}_0$ be the spherical harmonic coefficient vector of the measured (or simulated) HRTF of a certain radius r_0 in the valid range that confines all contributing sources. Then the extrapolated HRTF can be calculated for all possible radii r larger than the radius of the confined space by

$$\hat{f}(r) = \frac{h_n(kr)}{h_n(kr_0)} \hat{f}_0 \quad (4)$$

with $h_n(kr)$ being the outbound spherical Hankel function of order n. k denotes the wave number and r is the target radius of the range extrapolation.

As suggested in [4] the orders used for computation were limited to $n < kr$ to avoid the explosion of higher order coefficients when extrapolating to close ranges. No regularization was used for data processing in this paper.

5. GRAPHICAL RESULTS

5.1. Correlation plots

Fig. 1 and 2 show the normalized correlations between the 2 m data and the data at closer distances, without and with the use of extrapolation. These values can be used as a measure of how similar the HRTFs are at a closer range in comparison to the reference measurement (simulation) at a distance of 2 m. Note that the correlation plots of measurement and simulation have different axis scales to improve readability.

For both the extrapolated and the original data, the BEM simulation shows a significant higher correlation of the HRTFs of different ranges compared to the measured data. The reason could be the deviations in the different measurement setups used and the high precision of the simulated data. In both cases the extrapolation enhances the correlation, the measurement data shows good results up to approximately 8 kHz. Above that frequency the correlation is small, but still enhanced by extrapolation. The general shape of the correlation curve for the measurement and extrapolation at a range of 30 cm in Fig. 1 is similar, indicating a general deviation in the results during different measurement sessions.

The correlation of the extrapolated HRTF from simulation is perfect for the 1 m and 50 cm range at meaningful frequencies, the 30 cm extrapolation suffers from the lower order truncation and excessive amplification of small values of higher orders. These artifacts are expected to be attenuated using a suitable regularization algorithm.

5.2. Azimuthal HRTF results

In Fig. 3 and Fig. 4 the HRTFs of the horizontal plane are plotted over frequency on a logarithmic scale. The maximum level of the HRTFs was shifted to 0 dB for each range to allow comparison of the function on different radii despite the attenuation over distance. The color scale was chosen to show a dynamic range of 40 dB.

In the left columns the measured (simulated) data at distances of 2 m, 1 m, 50 cm and 30 cm are plotted, whereas the right columns contains the extrapolated data from the measurement (simulation) at the reference distance of 2 m depicted in the top diagram next to the corresponding dummy-head.

The general quality of extrapolation is high with a obvious similarity of the original functions and their extrapolated counterpart. The fine structure at high frequencies of the measurement data differs slightly, which can be explained as we use two measurement data from different measurements. At low frequencies some artifacts occur both in the measured and simulated data due to the frequency dependent upper cut-off limit for the used order.

6. CONCLUSIONS AND OUTLOOK

The methods of data interpolation and range extrapolation as shown in [4] seem to works well for the examined sets of HRTF data. As expected, numerically simulated HRTF can be extrapolated with very high accuracy. At close ranges some artifacts occur due to the band-limitation to avoid blow ups of higher orders of spherical harmonic coefficients. The graphical results of the measurement data and the corresponding extrapolation shows that the used method is capable of reducing the amount of measurement without having to neglect the near-field effects of HRTFs.

However, the deviations for the measurement seem to be much higher than for the simulation results. This might have several reasons: The HRTF for the 2 m range was measured some time before the HRTFs on closer distances. Hereby different excitation signals and measurement setups were used. Furthermore, the measurement data cover a much larger frequency range with good extrapolation results in the lower frequency range. Due to computational constraints it was only possible to derive simulated results at frequencies where the extrapolation is of high quality also for the measured data.

As high orders of spherical harmonic coefficients have a devastating effect when extrapolating their functions to closer ranges, it is important to obtain very clean measurement results and to use sophisticated algorithms for finding appropriate solutions in the spherical harmonic domain. Singular value decomposition or some kind of regularization, can help to further enhance the quality of extrapolation.

The used approach allows to calculate the HRTFs on any point outside a confined area that contains all scattering sound sources that influences the transfer functions. Due to performance issues these algorithms can currently not be implemented directly in the real-time scene rendering system. It is, however, possible to only measure a smaller set of HRTF data and calculate more points in space that can be used for a fast nearest neighbor search to preserve real-time computation. As computation power is expected to rise, implementation of these algorithms in real-time systems is an option for the future to enhance the quality of auralization in virtual reality.

Of course, without subjective criteria it is hard to evaluate the quality interpolation and extrapolation of HRTFs. Listening test are planned, either in the lab or in the virtual reality environment, to get a more accurate comparison of measured and extrapolated HRTFs.

7. ACKNOWLEDGMENT

The author thanks Tobias Lentz and Markus Müller-Trapet for providing the measured and simulated HRTF data used in this paper.

8. REFERENCES

[1] J. Blauert, *Spatial Hearing. The Psychophysics of Human Sound Localization*, The MIT Press, Cambridge, MA, USA, 1983

[2] T. Lentz, *Near-field HRTFs*, Fortschritte der Akustik, Stuttgart, Germany, 2007

[3] T. Lentz, *Binaural Technology for Virtual Reality*, Dissertation RWTH Aachen University, Germany, 2007

[4] R. Duraiswami, D.N. Zotkin, N.A. Gumerov, *Interpolation and range extrapolation of HRTFs*, IEEE International Conference on Acoustics, 2004

[5] J. Fels, *From Children to Adults: How Binaural Cues and Ear Canal Impedances Grow*, Dissertation RWTH Aachen University, Germany, 2008

[6] E.G. Williams, *Fourier acoustics: sound radiation and nearfield acoustical holography*, Washington D.C., USA, 1999

[7] F. Zotter, *Analysis and Synthesis of Sound-Radiation with Spherical Arrays*, Dissertation KU Graz, Austria, 2009

[8] B. Rafaely et al, *Spatial Aliasing in Spherical Microphone Arrays*, IEEE Transactions On Signal Processing, 2007

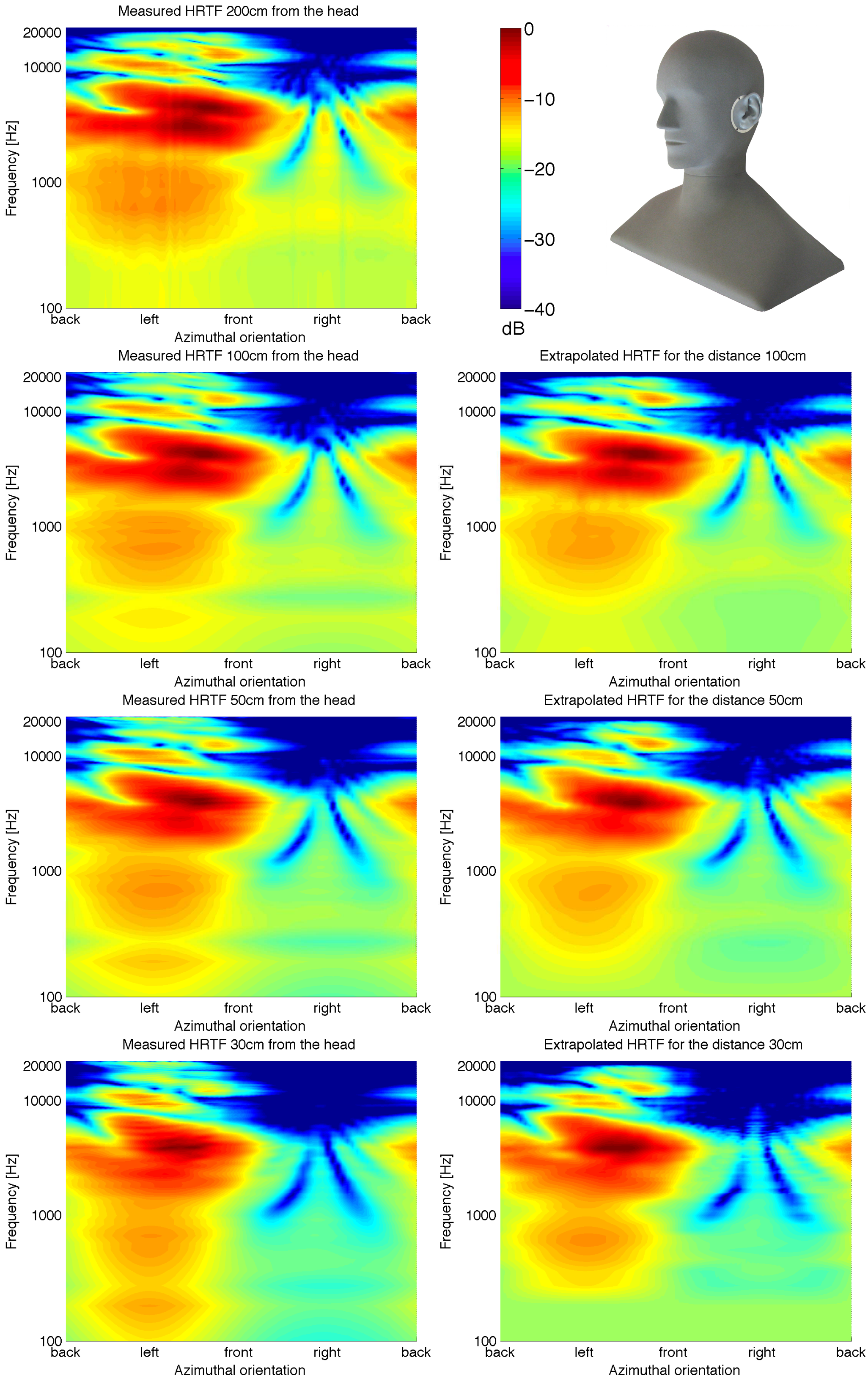


Figure 3: Measurement of a dummy-heads left-ear HRTF (left column) and the corresponding extrapolation results (right column) from the measured data at 2 m (top left plot)

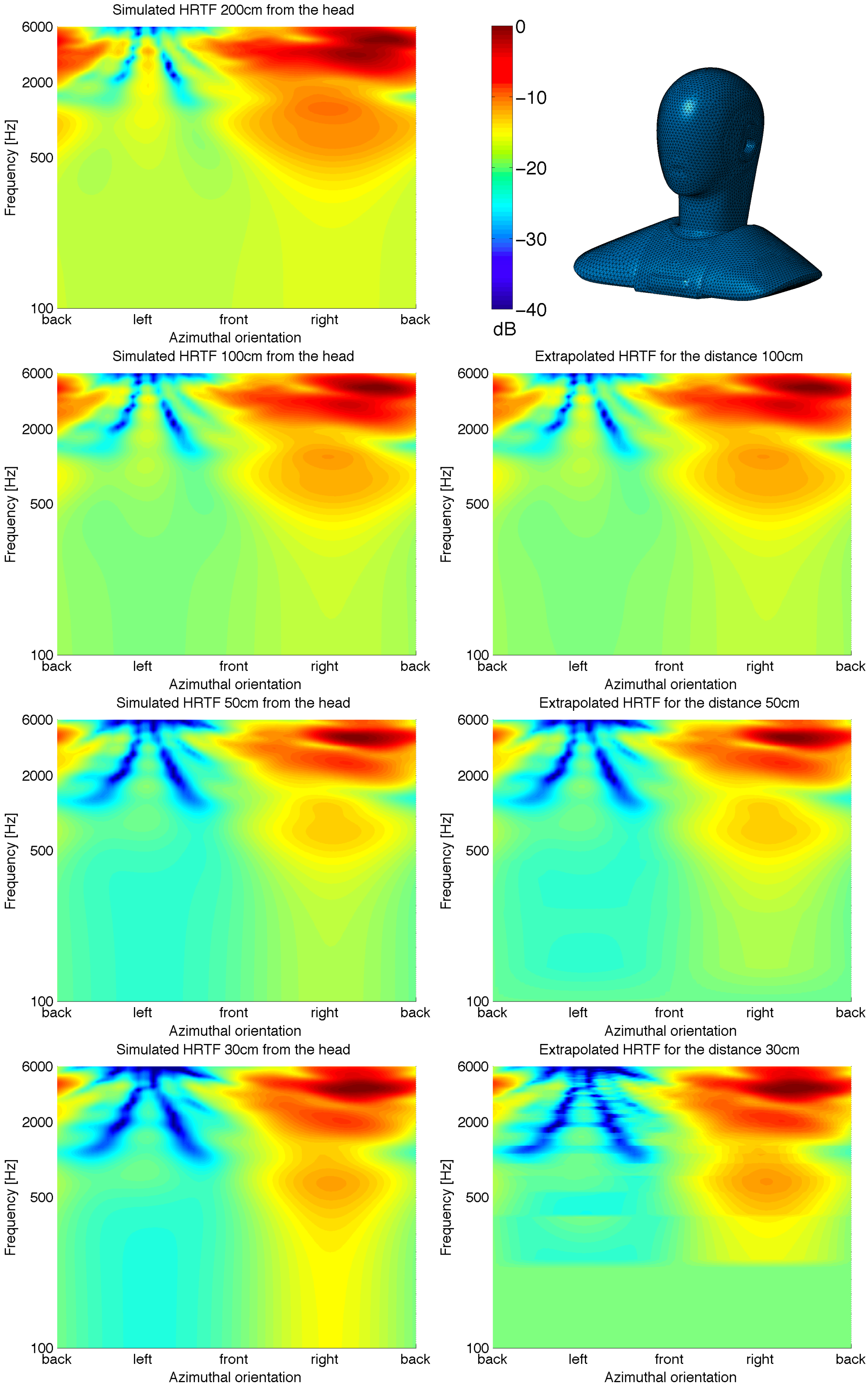


Figure 4: BEM simulation of a dummy-heads right-ear HRTF (left column) and the corresponding extrapolation results (right column) from the simulation at 2 m (top left plot)