Adaptation of HRTFs to Plane Waves with Reduced Modal Order

Benjamin Bernschütz[1,2]
[1] Cologne University of Applied Sciences - Institute of Communications Engineering, Germany
[2] Technical University of Berlin - Audio Communication Group, Germany
benjamin.bernschuetz@fh-koeln.de

Introduction

Modal descriptions of measured or simulated sound fields using spherical harmonics enjoy popularity. The binaural reproduction of the respective datasets is of great interest. Usually a set of head related transfer functions (HRTFs) is involved in the chain, in order to establish the typical binaural cues that can be evaluated by the human auditory system. Measurement and simulation systems often comprise a limited modal resolution only. This entails plane wave descriptions that are basically not compatible to native HRTFs, which inherently involve higher modal orders. The adaptation of the the high-order HRTFs to a low-order wave is discussed in the following, leading to HRTFs with reduced modal order (RHRTFs).

HRTF to Wave Adaptation

Analytic spatial Fourier coefficients $\mathring{P}_{nm}$ corresponding to a plane wave can be obtained using

$$\mathring{P}_{nm}(\omega, r) = 4\,\pi\,\mathrm{i}^n j_n(\tfrac{\omega}{c} r)\,Y_n^m(\Omega_w)^*, \quad (1)$$

where $\mathrm{i} = \sqrt{-1}$ denotes the imaginary unit, $(\cdot)^*$ the complex conjugate, ω the angular frequency, r the radius of an imaginary sphere around the observer's origin, c the speed of sound and $\Omega_w \equiv (\theta_w, \phi_w)$ the solid angle including elevation θ_w and azimuth ϕ_w. Y_n^m denotes the spherical harmonics [1] of order n and mode m, and j_n the spherical Bessel function of the first kind [1]. These two functions solve the angular or the radial portion of the Helmholtz equation in a spherical coordinate system. In practice, similar spatial Fourier coefficients could be delivered from microphone arrays [2], room simulation tools [3] or higher-order ambisonics (HOA) [4] decoders for instance. As the entire sound field in a source-free spherical volume can be expressed in terms of superimposed plane waves [1], the expression of a single plane wave can be generalized for the description of arbitrary complex sound fields. In order to enable binaural playback, directional information is extracted from $\mathring{P}_{nm}$ by applying plane wave decomposition and a subsequent weighting with far-field HRTFs as proposed e.g. in [5] or [6]. The adaptation is conductible in a mathematically closed form using a spherical HRTF set and involving appropriate spherical quadratures [7] serving as a discrete domain base in the spherical wave spectrum. The domain base quadrature with specific nodes Ω_c and node weights β_c is referred to as composite grid in the following. Thus both the plane wave decomposition and the spherical HRTF set need to be defined on identical nodes Ω_c, $c \in \{c, 1, ..., C\text{-}1, C\}$ of a common composite grid. The normalized plane wave decomposition for a node Ω_c yields

$$P_c(\omega, \Omega_c) = \frac{1}{(N_d+1)^2} \sum_{n=0}^{N_d} \sum_{m=-n}^{n} Y_n^m(\Omega_c) \frac{\mathring{P}_{nm}(\omega)}{\mathrm{i}^n j_n(\frac{\omega}{c} r)}, \quad (2)$$

where N_d denotes the maximum modal order of the decomposition that is basically determined by the coefficient's source. In a next step, appropriate HRTFs are multiplied to the decomposed field at each node and finally all nodes are summed. Thereby the HRTF angles need to match the nodes Ω_c of the composite grid which can be achieved by measuring HRTFs for the respective node angles. Alternatively a common high density HRTF set can be used performing a high-order HRTF interpolation in the spherical wave spectrum domain [8] for spatial resampling. This approach is more comfortable, as only one common HRTF set is required for any possible choice of composite grids. Thus in a first step the full HRTF set denoted by $H^{l,r}(\omega, \Omega)$ (l left, r right) is transformed to the spherical wave spectrum [1] using

$$\mathring{H}^{l,r}_{nm}(\omega) = \sum_{g=1}^{G} H^{l,r}(\omega, \Omega_g)\, Y_n^m(\Omega_g)^*\, \beta_g, \quad (3)$$

with Ω_g describing the HRTF angles and β_g the corresponding node weights. Of course, the HRTF set must be captured on a closed and sufficiently dense grid, allowing for high-order spherical harmonic interpolation. In a next step the specific HRTF corresponding to a composite grid node Ω_c is obtained applying the high-order inverse spatial Fourier transform [1] given by

$$\tilde{H}^{l,r}(\omega, \Omega_c) = \sum_{n=0}^{N_i} \sum_{m=-n}^{n} \mathring{H}^{l,r}_{nm}(\omega)\, Y_n^m(\Omega_c). \quad (4)$$

In a last step the decomposed wave field and the resampled HRTF set are merged, yielding a frequency domain binaural signal $Y^{l,r}(\omega)$ using

$$Y^{l,r}(\omega) = \sum_{c=1}^{C} P_c(\omega, \Omega_c)\, \tilde{H}^{l,r}(\omega, \Omega_c)\, \beta_c. \quad (5)$$

As a native spherical HRTF set at high temporal frequencies comprises substantial information in higher modal orders, it is not directly compatible to low-order wave field signals. Generally two systems of different modal resolution can be matched using spatial resampling in the space-frequency domain. The term resampling indicates different options; either upsampling the low-order system or downsampling the high-order system. Depending on the context, both options can be reasonable. For the binaural reproduction of order-reduced plane waves, downsampling the high-order system (thus the HRTFs) to the lower wave order is the appropriate option. If not properly adapted, i.e. by truncating the HRTF wave order in the spherical wave spectrum domain or upsampling the wave order to the HRTF order in the space-frequency domain, the high frequency components get lost and severe low-pass characteristics arise, which can be observed in Figure 1 (intercepted curves). Spatial downsampling of the high density and high-order HRTF set to the low-order wave conserves the high frequency information in a best possible manner for most of the incidence directions, cf. Figure 1. Nevertheless, for some specific incidence directions the result is less satisfying. In this context, the specific choice of the composite grid is of particular importance. First of all, the composite grid should match the order of the plane wave decomposition for minimizing the deviation between RHRTFs and HRTFs, cf. Figure 2. Furthermore the distribution of the nodes influences the properties of the resulting RHRTFs, cf. Figure 2.

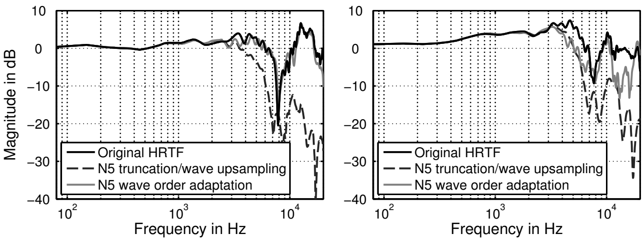

Figure 1: Two exemplary original HRTFs (black) versus their RHRTF counterparts at the wave order N5 for different incidence directions, i.e. approximately 10° (left) and approximately 35° (right) in the horizontal plane. The intercepted curves represent the result without any appropriate adaptation, either truncating the HRTF order or upsampling the decomposed wave. The expected low-pass characteristics is clearly visible. The gray curve is achieved using a Lebedev composite grid that matches the wave order. The results improve considerably for most incidence directions, whereas especially at low orders the RHRTFs still differ considerably from the original HRTFs. The left plot shows an example that achieves very good approximation and the right one an example with considerably higher deviations.

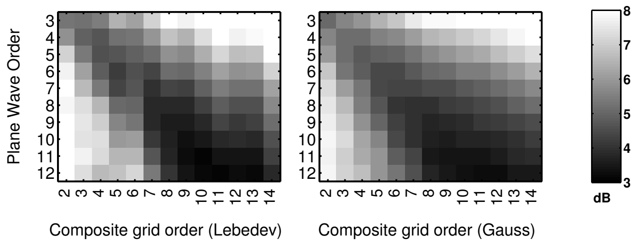

Figure 2: Mean spectral deviation between RHRTF sets of different modal order and the corresponding full resolution HRTF set used as indicator for the reconstruction performance averaged over 974 wave incidence directions involving different composite grid types (Lebedev, Gauss) and orders.

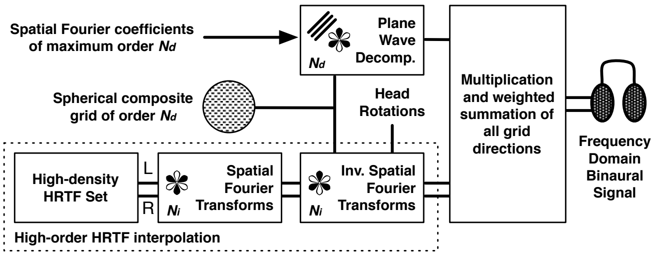

Figure 3: Block Diagram

Conclusions and Outlook

The adaptation of HRTFs to plane waves with reduced modal order was discussed. At rising wave order the resulting RHRTFs converge to the original HRTFs. At low wave orders, spatial downsampling of the HRTF set to the wave order minimizes the differences substantially. Nevertheless the RHRTFs are not identical to the full resolution HRTFs. At that point the question arises, to what extent these differences are perceived by a human listener and which factors mainly influence the perception. These aspects are discussed in a subsequent more detailed publication [9] that is closely related to the present one.

Funding

The research activities are funded by the Federal Ministry of Education and Research in Germany under the support code 17009X11-MARA.

References

[1] E Williams, *Fourier acoustics: Sound radiation and nearfield acoustical holography*, Academic Press, 1999.

[2] J Meyer and G Elko, "A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield," in *Proc. of the ICASSP (IEEE)*, 2002.

[3] M Vorländer, M Pollow, and S Pelzer, "Interfacing spherical harmonics and room simulation algorithms," *JASA*, vol. 131, no. 4, 2012.

[4] J Daniel, R Nicol, and S Moreau, "Further investigations of high order ambisonics and wavefield synthesis for holophonic sound imaging," in *Proceedings of the AES Convention 114*, 2003.

[5] R Duraiswami, D N Zotkin, Z Li, E Grassi, N A Gumerov, and L S Davis, "High order spatial audio capture and its binaural head-tracked playback over headphones with HRTF cues," in *Proc. of the AES Convention 119*, 2005.

[6] S Spors, H Wierstorf, and M Geier, "Comparison of modal versus delay-and-sum beamforming in the context of data-based binaural synthesis," in *Proc. of the AES Convention 132*, 2012.

[7] F Zotter, "Sampling strategies for acoustic holography/holophony on the sphere," in *Proc. of the NAG-DAGA*, 2009.

[8] R Duraiswami, D N Zotkin, and N A Gumerov, "Interpolation and range extrapolation of HRTFs," in *Proc. of the ICASSP (IEEE)*, 2004.

[9] B Bernschütz, A Vázquez-Giner, C Pörschmann, and J Arend, "Binaural reproduction of plane waves with reduced modal order," in *Acta Acustica united with Acustica*, accepted for publication in November 2014.