Fast Range-Acquisition of Head-Related Impulse Responses

Gerald Enzner

Institut für Kommunikationsakustik, Ruhr-Universität Bochum, 44780 Bochum, Email: gerald.enzner@rub.de

1. Introduction

The head-related impulse response (HRIR) is the core tool in binaural sound technology. In full range, it describes the direction and distance dependent receiver characteristics of the human head and the outer ear. By feeding the HRIRs of both ears with a dry input signal, we can almost ideally mimic the spatial cues of that signal, as if it were reproduced by a sound source in space.

In this contribution, we describe the ingredients of a method for fast and accurate acquisition of HRIRs in a region of interest. For the horizontal plane as a region of particular relevance, we describe the measurement density and the related amount of spatial interpolation and extrapolation. Special attention is directed to a new procedure for fast realization of the measurement part, because it facilitates individualized HRIR acquisition.

2. State-of-the-Art

Previous work on range-acquisition of head-related transfer functions (HRTFs) naturally includes the spatially discrete measurement with certain degrees of radial and angular resolution and geometric interpolation at intermediate locations, e.g., [1]. A holophonic approach to radial extrapolation of discrete HRTF measurements, using the spherical harmonics domain, was presented in [2]. This method was recently verified in [3] by comparing the results to the radial HRTF measurements described in [1]. Another concept makes use of simplified head models in order to apply generic near-field modifications to far-field HRTF measurements [4].

3. Fast Range-Acquisition of HRTFs

In contrast to available concepts, our method tries to overcome the unavoidable errors due to the spatial discretization of HRIR measurements, such as interpolation errors or aliasing. Essentially, we utilize a dynamical measurement setup with which we can perform quasi-continuous acquisition of the HRIR along a desired trajectory. We further exploit reciprocity and consider the entire HRIR as a soundfield emitted by the subject of interest. Then, using the HRIR measurements on a circular trajectory in space, we finally reconstruct the entire HRIR-field in a holophonic sense. In the following, these ingredients of our method are described more precisely.

3.1. Continuous-Azimuth Measurements

Consider the dynamical measurement setup in Fig. 1 with a rotating subject of interest. The fixed loudspeaker at $r_m = 1.5$m distance from the subject emits a broadband probe signal $x(k)$ at sampling time k and we record the binaural signals $y_\ell(k)$ and $y_r(k)$. After a full 360° rotation of the subject, the binaural recording represents an observation of all possible HRIRs on the azimuth-circle.

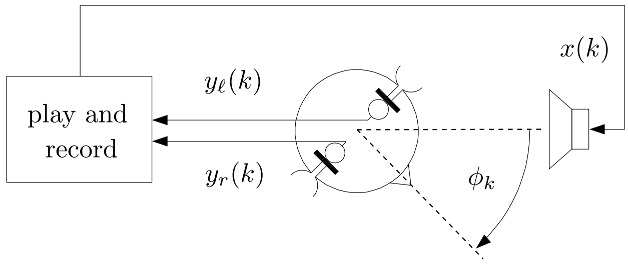

Fig. 1: Rotating, dynamical HRTF measurement setup.

Assuming sufficiently linear and broadband transducers, we can describe the rotating HRIR between the loudspeaker and the subject as a slowly time-varying system with impulse response $h_{\ell,r}(k, \phi_k)$. Therefore, the HRIR formally relates HRIR input and output signals according to a linear convolution (linear regression) model, i.e.,

$$y_{\ell,r}(k) = \sum_{\kappa=0}^{N} x(k-\kappa) h_{\ell,r}(\kappa, \phi_k) + n_{\ell,r}(k) \ , \qquad (1)$$

where N denotes the length (memory) of the HRIRs and $n_{\ell,r}(k)$ is independent observation noise at the recording positions. The quasi-continuous azimuth $\phi_k = \omega_o k T_s$ at time k evolves with constant angular frequency $\omega_o = 2\pi/T_{360}$, where T_{360} is the duration of a 360° revolution and $T_s = 1/f_s$ is the temporal sampling interval.

Then, let

$$\mathbf{x}(k) = (x(k), x(k-1), \ldots, x(k-N+1))^T \qquad (2)$$

denote the set of most recent HRIR input samples and

$$\mathbf{h}_{\ell,r}(\phi_k) = (h_{\ell,r}(0,\phi_k), \ldots h_{\ell,r}(N-1,\phi_k))^T \qquad (3)$$

the corresponding vector of the HRIR coefficients in (1). Using this notation, we can formulate the normalized least mean-square (NLMS) algorithm [5] for iterative extraction of the rotating HRIRs from the data $y_{\ell,r}(k)$:

$$\widehat{\mathbf{h}}_{\ell,r}(\phi_{k+1}) = \widehat{\mathbf{h}}_{\ell,r}(\phi_k) + \mu_0 \frac{e_{\ell,r}(k)\mathbf{x}(k)}{||\mathbf{x}(k)||_2^2} \qquad (4)$$

$$e_{\ell,r}(k) = y_{\ell,r}(k) - \widehat{\mathbf{h}}_{\ell,r}^T(\phi_k)\mathbf{x}(k) \ . \qquad (5)$$

This gradient-descent approach, which has been tailored for the tracking of time-varying systems, updates the inference $\widehat{\mathbf{h}}_{\ell,r}(\phi_k)$ from ϕ_k to ϕ_{k+1} by explicit minimization of the squared error $e_{\ell,r}^2(k)$. The symbol μ_0 denotes a stepsize factor which can be chosen close to unity in our low-noise environment. It has been verified that the NLMS algorithm with its tracking ability extracts HRIRs from the data and thereby preserves spatial HRIR cues, such as interaural time and level differences between the ears. For a fast measurement with $T_{360} = 20s$, it yields a sufficiently low HRIR noise floor about 50 dB below the direct sound amplitude and, finally, it causes hardly any spatial smoothing of the extracted HRIRs $\widehat{\mathbf{h}}_{\ell,r}(\phi_k)$ [6].

3.2. Reciprocity and HRTF-Field

While the HRIRs $\widehat{\mathbf{h}}_{\ell,r}(\phi_k)$ are in fact measured from the loudspeaker into the direction of the subject, cf. Fig. 1, the same acoustic impulse responses can be used to describe a virtual sound propagation from the measurement points in the ear canals to the respective loudspeaker location in space. This is justified by the principal of reciprocity which applies to soundfields in general [7].

By considering a sound source in the ear canal, we can assign a sound pressure value $p(r, \phi, \omega)$ to each frequency ω and (r, ϕ)-tupel in space. If a unit pulse is then emitted by the source, the resulting sound pressure field $p(r, \phi, \omega)$ in frequency-domain is termed the head-related transfer function (HRTF) field. This field exists independently for left and right ear and is illustrated in Fig. 2 by the spherical wave centered at the left ear position.

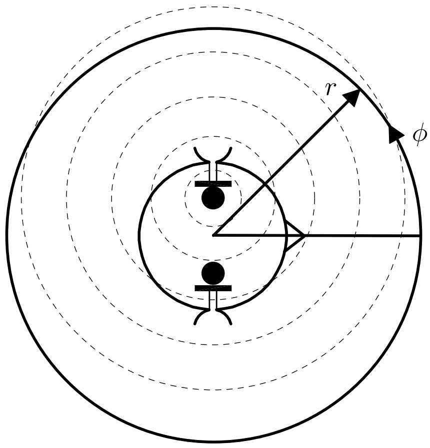

Fig. 2: The left-ear reciprocal HRTF-field (dashed) and the head-centered cylindrical coordinate system (solid).

The HRTF-field $p(r_m, \phi, \omega)$ on the HRIR measurement radius r_m can be obtained by the temporal Fourier transform of the measured HRIRs (left and right ear separately) at the respective azimuth coordinate ϕ,

$$p(r_m, \phi, \omega) = \mathcal{F}\left\{\widehat{h}(k, \phi)\right\} = \sum_{k=0}^{N-1} \widehat{h}(k, \phi) e^{-j\omega T_s k} , \quad (6)$$

where "l, r" indices of HRIRs were dropped for brevity.

3.3. Radial HRTF Projection

For the exterior configuration in Fig. 2, which basically assumes a source- and scatter-free region outside the head, the general solution to the acoustic wave equation can be written as a modal series in polar coordinates [8],

$$p^*(r, \phi, \omega) = \sum_{n=-\infty}^{\infty} A_n(\omega) H_n^{(1)}(kr) e^{jn\phi} , \quad (7)$$

where $H_n^{(1)}(x)$ is the complex Hankel function of the first kind, $k = \omega/c$ the acoustic wavenumber, c the speed of sound, n the modal index, and A_n the independent set of modal coefficients. Notice that the soundfield in (7) is independent of the height of the horizontal plane.

For a given radius $r = r_m$ and frequency ω, the soundfield according to (7) represents a complex Fourier series with 2π-periodicity in ϕ. The coefficients $A_n(\omega)$ can thus be obtained by complex Fourier analysis of the sound pressure $p(r_m, \phi, \omega)$ along the azimuth circle and subsequent normalization by the Hankel function at radius $r = r_m$:

$$A_n(\omega) H_n^{(1)}(kr_m) = \frac{1}{2\pi} \int_0^{2\pi} p^*(r_m, \phi, \omega) e^{-jn\phi} \, d\phi \; . \quad (8)$$

Knowing $A_n(\omega)$, we can perform holophonic reconstruction of the entire HRTF-field $p(r, \phi, \omega)$ outside the head radius using (7). Fig. 3 depicts an HRTF reconstruction example, based on continuous-azimuth measurements.

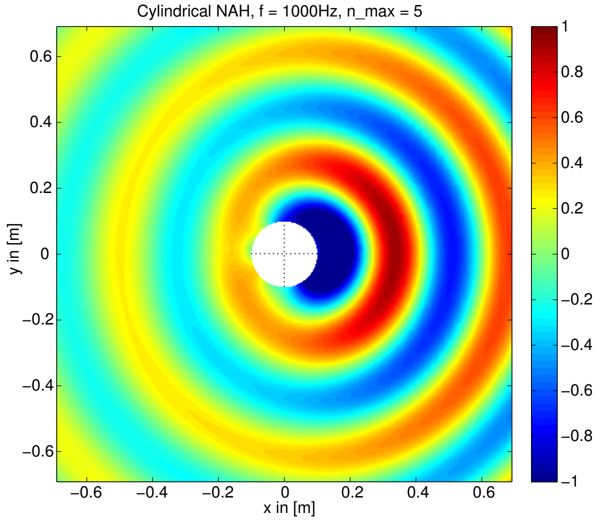

Fig. 3: HRTF-field (right); nose in y-direction; $w = 2\pi \cdot 1$ kHz.

The HRIR on a desired contour $r \neq r_m$ can be finally obtained by inverse Fourier transform $\mathcal{F}^{-1}$ of the extrapolated sound pressure $p(r \neq r_m, \phi, \omega)$ according to (6). The particular advantage of this holophonic HRIR reconstruction based on our continuous-azimuth measurements, cf. Sec. 3.1., lies in the opportunity of quasi-continuous evaluation of the integral in (8), thus avoiding possible errors due to soundfield discretization.

References

[1] Lentz, T.: Near-Field HRTFs, DAGA 2007

[2] Duraiswami, R. et al.: Interpolation and Range Extrapolation of HRTFs, ICASSP 2004

[3] Pollow, M.: Applying Extrapolation and Interpolation Methods to Measured and Simulated HRTF Data, 2nd Intl. Symp. on Ambisonics (Paris) 2010

[4] Romblom, D., Cook B.: Near-Field Compensation for HRTF Processing, 125th AES Convention, Oct. 2008

[5] Haykin, S.: Adaptive Filter Theory, Prentice-Hall

[6] Enzner, G.: Analysis and Optimal Control of LMS-type Adaptive Filtering for Continuous-Azimuth Acquisition of HRIRs, ICASSP 2008

[7] Blauert, J.: Acoustics for Engineers, Springer, 2008

[8] Williams, E.G.: Fourier Acoustics, Acad. Press, 1999