On the Implementation of Range Extrapolation of Head-Related Impulse Responses by Virtual Sound Field Synthesis using Integer Delays

Sascha Spors and Hagen Wierstorf

Quality and Usability Lab, Deutsche Telekom Laboratories, Technische Universität Berlin, Germany.

Email: {Sascha.Spors,Hagen.Wierstorf}@tu-berlin.de

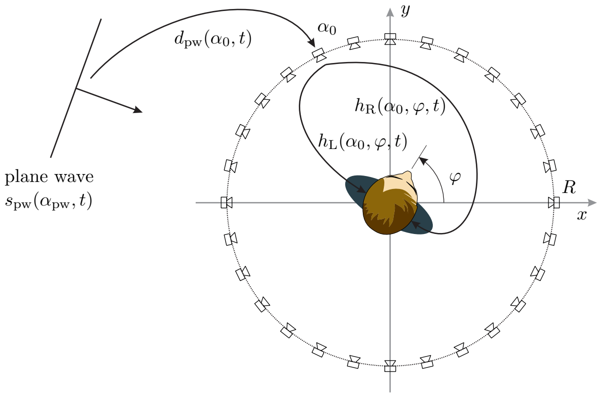


Figure 1: Principle of computing far-field HRIRs by range extrapolation on the basis of virtual WFS.

Introduction

Recently an approach to range extrapolation of head-related impulses responses (HRIRs) has been proposed [1]. It is based on interpreting measured HRIRs as a virtual secondary source distribution which is driven according to a virtual source at the desired distance. In principle, any sound field synthesis (SFS) approach can be used. However, Wave Field Synthesis (WFS) is favorable since it allows for a numerical stable and efficient implementation. The proposed technique requires prefilterings of the HRIR dataset, and summing up weighted and delayed versions of the HRIRs. In WFS the delay operations are typically quantized towards multiples of the sampling interval [2]. For range extrapolation, quantized delays pose limits with respect to the achievable physical accuracy. These limits and potential solutions are discussed in the paper, as well as the perceptual consequences in the context of range extrapolation of HRIRs. The investigations are limited to horizontal plane HRIRs.

Range Extrapolation of HRIRs

WFS aims at the synthesis of a desired sound field within an extended area using an ensemble of individually driven loudspeakers (secondary sources). HRIRs represent the impulse responses from an acoustic source to the ears. Datasets of HRIRs are typically measured for a fixed distance R of the source and varying incidence angle α_0. Such datasets can be interpreted as virtual circular loudspeaker array. Refer to Figure 1 for an illustration.

The far-field extrapolated HRIRs are given by the impulse response from the virtual source to the left/right ear [1]

$$\bar{h}_{\{\mathrm{L,R}\}}(\varphi, t) = \int_0^{2\pi} d_{\mathrm{pw}}(\alpha_0, t) * h_{\{\mathrm{L,R}\}}(\alpha_0, \varphi, R, t)\, R\, d\alpha_0, \quad (1)$$

were $d_{\mathrm{pw}}(\alpha_0, t)$ denotes the driving signal and $h_{\{\mathrm{L,R}\}}(\alpha_0, \varphi, R, t)$ the left/right HRIRs for a given head orientation φ, respectively. For the synthesis of far-field HRIRs the desired virtual source constitutes a Dirac shaped plane wave. The driving function is given by specializing [3, eq.(27)] to the given geometry

$$d_{\mathrm{pw}}(\alpha_0, t) = a(\alpha_0) w(\alpha_0) \left(f_{\mathrm{pre}}(t) * \delta(t - \tau(\alpha_0)) \right), \quad (2)$$

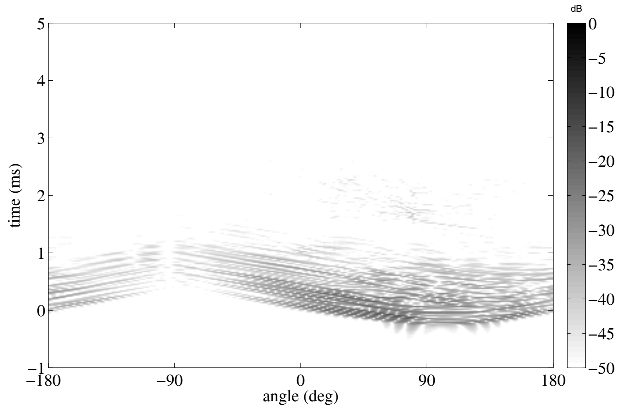
were $a(\alpha_0)$ denotes a window function that selects the active secondary sources for a given incidence angle α_{pw} of the virtual plane wave, $w(\alpha_0)$ a weighting factor, $f_{\mathrm{pre}}(t)$ the pre-equalization filter of WFS and $\tau(\alpha_0) = \frac{R}{c} \cos(\alpha_0 - \alpha_{\mathrm{pw}})$ a delay. Introducing (2) into (1) reveals that the plane wave HRIRs can be computed efficiently by pre-filtering the HRIRs, and summing up weighted and delayed versions of the pre-filtered HRIRs.

Practical Aspects

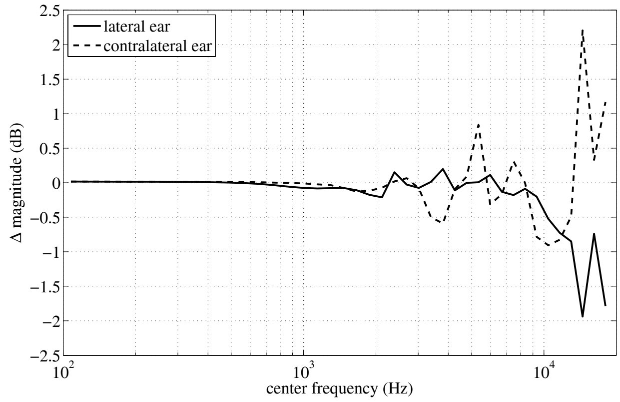
The extrapolation of HRIRs using virtual WFS is subject to a number of practical aspects.

Equation (1) assumes the availability of spatially continuous HRIRs. In practice, HRIRs are spatially sampled which may lead to spatial sampling artifacts in the extrapolated HRIRs. The considered scenario of a circular distribution of virtual secondary sources constitutes 2.5-dimensional synthesis. The resulting amplitude deviations for WFS (and other techniques) are well researched [3]. For the computation of far-field HRIRs this undesired amplitude decay has to be compensated for. Otherwise the interaural level differences (ILDs) will not be correct for the synthesized HRIRs. The artifacts following from sampling and 2.5-dimensional synthesis are not considered in this paper.

For computational efficiency, the delay operation in (2) is quantized to multiples of the temporal sampling interval. This quantization has shown to be inaudible for static virtual sources synthesized by typical WFS setups [2]. Range extrapolation of HRIRs may involve very densely distributed virtual secondary source distributions. Here, temporal quantization of the delay operation may lead to numerical artifacts. These emerge from the fact that multiple secondary sources are delayed by the same value. The consequences of delay quantization are investigated in this paper.



(a) HRIRs (left-ear)



(b) weighted frequency response

Figure 2: Differences between far-field BRIRs calculated using continuous and quantized delay operations.

Results

A custom dataset of horizontal plane HRIRs [4] measured at a source distance of $R = 0.5\,\mathrm{m}$ with angular resolution of one degree is used in order to avoid spatial sampling artifacts. A dataset of left/right far-field HRIRs was calculated by numerical evaluation of (1) and (2). Two scenarios are considered in order to investigate the effect of quantizing the delay $\tau(\alpha_0)$ in (2): (i) no quantization and (ii) quantization towards multiples of the sampling interval $T = 1/f_\mathrm{S}$ with $f_\mathrm{S} = 44.1\,\mathrm{kHz}$. The former is approximated by a fractional delay filter, which was designed as least squared error finite impulse response [5] of length $N_\mathrm{fd} = 30$. Figure 2a shows the absolute difference between the HRIRs calculated with continuous/quantized delays. Note the color scale denotes level in dB. A difference of up to -15 dB can be observed. However, the HRIRs itself and their differences provide only limited insight into the perceptual relevance of the observed differences. In order to investigate further on this, we calculated the interaural level differences (ILDs) and interaural time differences (ITDs) for both HRIR datasets. No relevant differences were observed, indicating that localization is not affected by quantized delays. Figure 2b shows the weighted magnitude frequency response for the lateral and contralateral ear for a plane wave impinging onto the left ear. The response was processed by an auditory filter bank composed of gammatone filters with 1 ERB wide filters. For frequencies above 10 kHz differences up to 2 dB can be observed. These differences may lead to audible coloration.

An informal listening experiment was carried out in order to investigate the audability of the observed differences. Stimuli were produced by convolving the calculated HRIRs with different source material for a variety of different angles. As source material speech and castanets were used. Direct comparison of the stimuli pairs revealed no perceptual difference between continuous and quantized delays. Listening examples are available at `http://audio.qu.tu-berlin.de/?p=835`.

The instrumental results found by our investigations indicate that fractional delay filters are required for the implementation of HRIR extrapolation. However, results from informal listening revealed that no perceptual differences between both methods can be observed. Hence, quantized delays seem to be suitable for far-field extrapolation of HRIRs using virtual WFS. This allows for an computationally efficient implementation of this method. Note, the derived results may not hold for the computation of near-field HRIRs using extrapolation techniques.

Reproducible Research

The numerical simulations in this paper are based on the Sound Field Synthesis Toolbox [6]. The additional scripts used in this paper, as well as the slides from the talk (German) and the listening examples are available at `http://audio.qu.tu-berlin.de/?p=835`.

Acknowledgments

This work was conducted within the research group '*Simulation and Evaluation of Acoustical Environments (SEACEN)*' supported by the grant FOR 1557 of the Deutsche Forschungsgemeinschaft (DFG).

References

[1] S. Spors and J. Ahrens. Generation of far-field head-related transfer functions using sound field synthesis. In *German Annual Conference on Acoustics (DAGA)*, March 2011.

[2] J. Ahrens, M. Geier, and S. Spors. Perceptual comparison of wave field synthesis implemented with integer and fractional delays. In *German Annual Conference on Acoustics (DAGA)*, March 2010.

[3] S. Spors, R. Rabenstein, and J. Ahrens. The theory of wave field synthesis revisited. In *124th AES Convention*. Audio Engineering Society (AES), May 2008.

[4] H. Wierstorf, M. Geier, A. Raake, and S. Spors. A free database of head related impulse response measurements in the horizontal plane with multiple distances. In *130th AES Convention*. Audio Engineering Society (AES), May 2011.

[5] T.I. Laakso, V. Välimäki, M. Karjalainen, and U.K. Laine. Splitting the unit delay. *IEEE Signal Processing Magazine*, 13(1):30–60, January 1996.

[6] H. Wierstorf and S. Spors. Sound field synthesis toolbox. In *132th AES Convention*. Audio Engineering Society (AES), April 2012.