A Spherical Far Field HRIR/HRTF Compilation of the Neumann KU 100

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

Introduction

A full sphere far field HRIR/HRTF (head related impulse response and its Fourier transform head related transfer function) [1] compilation of a Neumann KU 100 dummy head is presented. The compilation can be used for different purposes; it is, for instance, very convenient for applications in the field of spherical Fourier acoustics [2]. Different sampling configurations were captured and the compilation on the whole involves around 22.000 post-processed binaural impulse responses. The data is organized and stored in MATLAB® objects, offering different processing features and including gateways to the SOFiA sound field analysis toolbox [3] and to the Sound Scape Renderer [4]. Associated headphone compensation filters for around 20 common headphones were created. Additionally, a method to optimize and extend the low frequency response of measured HRIRs/HRTFs is discussed.

HRIR Measurements

The HRIR measurements were conducted in the anechoic chamber at Cologne University of Applied Sciences. The respective chamber has dimensions of 4.5 x 11.7 x 2.3 m and a lower frequency boundary of around 200 Hz. All surfaces are absorptive. A custom mount was constructed to install the Neumann KU 100 dummy head on the highly accurate VariSphear motion system [5], offering motor position deviations $< 0.01°$. The resulting orientation entailed a transform of the usual VariSphear coordinate system. The impulse response measurements and the motion control were performed using the proprietary VariSphear software. The measurement core of the software has an integrated error detection stage to ensure the validity of all measured impulse responses. An emphasized sine sweep with +20dB low-shelf at 100 Hz of 2^{19} samples at 48kHz temporal sampling rate was used for excitation. The long sweep duration (approx. 11 s) made the measurements robust against background noise and entailed a good overall signal to noise ratio. A RME Fireface UCX including its internal low-noise microphone preamps and AD/DA converters was used as audio interface. The sound source was a Genelec 8260A active 3-way speaker system. The speaker was driven around -3 dB below the internal limiter threshold. No relevant climate changes in the chamber or time-variance of the speaker system due to driver heating were observed during the session. The temperature was constantly tracked and a static microphone was used to detect time variances. The frequency response of the speaker at the physical origin (center of the head) was tuned to be flat within a tolerance of less than ± 1dB for the entire involved spectrum using a digital signal processor. The speaker system was high-passed around 60 Hz in order to reduce the power load. This was feasible as the final low frequency components were generated during the post processing. The physical origin was placed in the vertical acoustic center of the speaker, where the phase response of the single speaker components is perfectly aligned. The head position adjustment was conducted with high accuracy using a cross-grid laser combined with real-time analysis of the phase response differences at the ear channels. The distance between speaker system and center of the head was approximately 3.25 m, which can be considered as far field referring to the speaker dimensions.

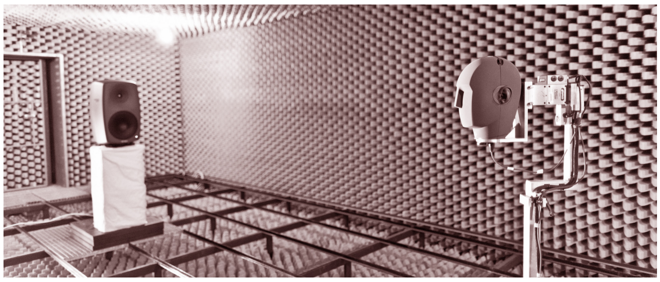

Figure 1: Genelec 8260A speaker and Neumann KU100 dummy head mounted on the VariSphear motion system [5] in the anechoic chamber. [Photo: P. Stade]

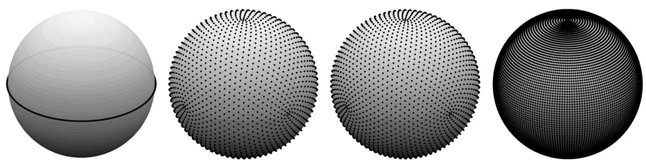

Figure 2: Spatial sampling configurations: Circular 1°, Lebedev 2354, Lebedev 2702 and Gauss-Legendre 2°.

Several different sampling configurations were captured. First of all, two circular horizontal turns with an angle resolution of 1° were run. The first circle was recorded using a simple thin microphone stand on the rotating ground-plane and the second circle using the complete 3D robot arm and rotation mount that was also used for all subsequent spherical grids. Two different equidistant spherical Lebedev grids (2354 and 2702 nodes) were recorded. These are convenient for typical applications in spherical acoustics. The grids were simulated in advance using the SOFiA sound field analysis toolbox [3]. Both of the presented Lebedev grids enable a stable transform into the spherical wave spectrum domain (spherical Fourier transform based on surface spherical harmonics) [2] with low contribution of spatial aliasing on the entire audio spectrum up to >20 kHz on the KU100 head radius. Finally a full sphere equiangular 2° Gauss quadrature with 16020 nodes was captured. There are no grid discontinuities as all grid nodes could be measured. This is particularly relevant in spherical acoustics. In addition to the HRIRs, single impulse responses were captured at the physical origin using an omnidirectional Microtech Gefell M296S microphone.

miro Data Format

A proprietary data type was developed in order to store, organize and access the datasets in a comfortable way. It is a simple object based MATLAB® data type called *miro* (measured impulse response object); *miro* integrates three stages: It combines the storage of raw impulse responses, the access to around 50 meta information properties and a method set for the access to the impulse response as well as for the treatment and conversion of the signals. It offers interfaces to the SOFiA sound field analysis toolbox [3] for spherical harmonics processing and to the Sound Scape Renderer [4] for dynamic binaural synthesis.

Headphone and Diffuse Field Filters

Even headphones and dummy heads explicitly advertising free field equalization [6] do usually still not meet the requirements for a highly natural and color-free reproduction that is needed for an immersive binaural listening experience. Hence the frequency response of the headphones, respectively the full transducer chain, must be explicitly compensated by application of appropriate compensation filters [7]. In order to make this highly specific compensation filters accessible to a broad range of different users, around 20 different common headphone models were analyzed to create specifically adapted compensation filters. To create the filters, each headphone was put on the KU100 head and replaced 12 times in order to capture stable and representative transfer functions. Additionally, a common free field compensation filter has been derived from spherical magnitude averaging, compare Figure 3. This filter compensates the dummy head itself and does not refer to a specific headphone. The compensation filters were computed based on a semi-automatic log-spline inversion algorithm. The proprietary *miro* data type enables a direct inline processing of the headphone and free field compensation filters.

The low frequency range of HRIRs

In a system for binaural synthesis, source signals are convolved with HRIRs. Hence all magnitude and phase properties of the HRIRs are directly imprinted on the resulting audio signal. Besides the desired magnitude and phase properties that are inherent to binaural techniques [1] [8], additional undesired parasitic magnitude and phase changes may arise which originate from the process of capturing the HRIRs. Especially at the low frequency end, below 100 Hz for instance, it is difficult to capture HRIRs without undesired artifacts due to several reasons. Small studio monitors that are typically used for HRIR measurements, mostly cannot reproduce frequencies e.g. below 60 Hz with sufficient sound pressure level. This lack of low frequencies cannot be compensated using filters or equalizers. As a consequence, the measured HRIRs cannot transmit low frequencies either. A possible solution is to use a bigger and more powerful speaker system, which otherwise brings along an increased cabinet sizes and larger spatial distributions of the single drivers. Especially for near-field measurements this is not desirable. A less obvious problem of speaker systems is the typical surge of group delay towards low frequencies. The group delay rises due to different mechanical reasons, bass-reflex constructions and filter networks [9]. Besides of the question at which dimensions group delay distortions lead to any noticeable impairment of the audio signal [10], there is a simple and very practical problem with it: Group delay differences directly come along with a spread of the time domain signal. Thus the HRIRs need more filter taps to transmit the full audio spectrum and demand for more computational power. While the problems concerning speaker systems are widely solvable with some effort, the anechoic chamber can bring up much more fundamental problems due to finite room dimensions and the limited length of absorption wedges. Below certain boundary frequency, reflections and room modes arise. As the low frequencies do not contribute significantly to localization [8], the respective reflections do usually not directly affect the binaural hearing. But reflections and room modes entail an overlay of the direct sound with reflected waves at the transducers. Depending on the phase relation this can lead to an amplification or to cancellation of the respective frequency. This has dramatic influences on the frequency response of the system. Furthermore room modes bring along a very excessive surge of group delay, as the decay of energy is slow. If these issues are not considered and the effects are eliminated, the measured HRIRs do not transmit low frequencies properly.

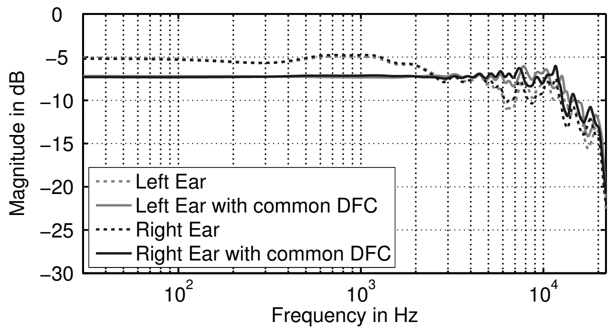

Figure 3: Averaged and weighted magnitude responses over the full sphere for the left and the right ear with and without the common diffuse field compensation filter (common DFC).

Adaptive low frequency extension (LFE)

In order to avoid negative influences on the low frequencies of the measured HRIRs, the entire low frequency range is replaced by an analytic extension. A similar approach has been presented in [11]. At high frequencies the head as a rigid body evokes shading and scattering of the sound field and the pinna and ear canal work as filters [1] [8]. Whereas at frequencies below e.g. 400 Hz, pinna and ear canal filters do barely influence the signal and the head itself does only have minor influence on the sound field. Analytic simulations of a plane wave

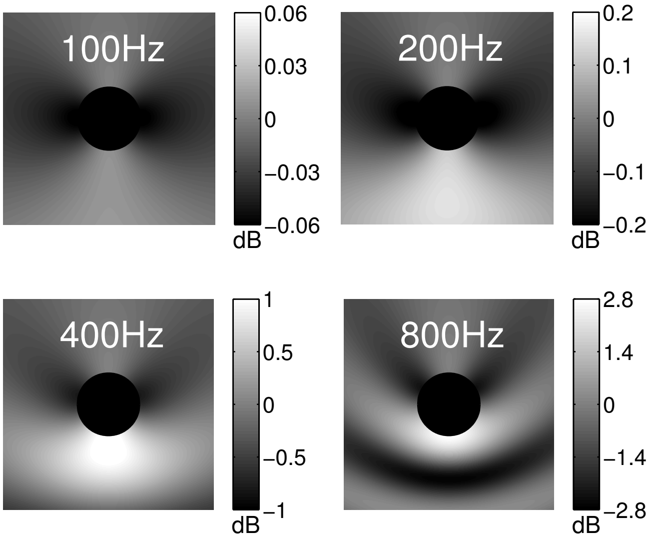

Figure 4: Analytic simulation of the pressure magnitude at different frequencies for a plane wave impact with unit gain from south to a rigid sphere with a typical diameter of 17.5 cm serving as a simplified model of the dummy head. Below the LFE crossover frequency of 200 Hz, the magnitude deviations on the sphere surface are already less than ±0.2 dB.

impact on a rigid sphere as a simplified model of the dummy head (Figure 4) indicate that there are only neglectable differences in the pressure magnitude around the sphere below 200 Hz. Thus it is feasible to replace the low frequency band using simplified analytic descriptions including a flat frequency response. An adaptive low

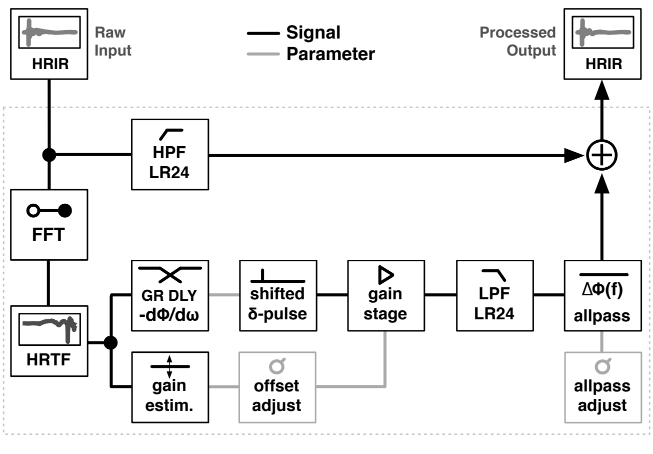

Figure 5: Basic structure of the adaptive LFE algorithm.

frequency extension (LFE) algorithm has been employed that generates the low frequency component and matches gain and phase, see Figure 5. Each HRIR and each ear is processed independently. At certain crossover frequency (200Hz) the original HRIR is high-passed (24 dB/oct Linkwitz-Riley) and a matched low frequency path is added. To generate the LFE path, the HRIR is transformed into a HRTF using the Fourier transform. The group delay of the original HRTF is evaluated around the crossover frequency and a shifted δ-pulse is generated accordingly. The required gain for the pulse is estimated by an analysis of the original HRTF gain around the crossover frequency on the one hand and either evaluation of single selected stable bins at lower frequencies or the simulation of a sphere model on the other hand. This is done as besides residual shading effects, even in the far field the different ear distances to the source bring along certain attenuation of the sound pressure level. This effect obviously becomes much more important for near-field measurements. In a next

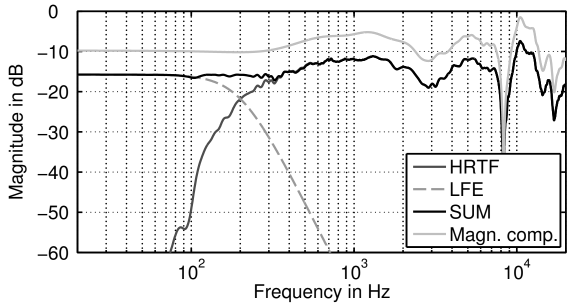

Figure 6: Post-processed HRTF split into the original high-passed HRTF path and the adaptive LFE path. Additionally the final magnitude compensated HRTF is depicted.

step the impulse is lowpass filtered. Finally an allpass filter is employed in order to match the phase slope of the original high-passed HRTF path and the LFE path around the crossover frequency. Out of this range, the LR24 crossover filters attenuate the signal sufficiently. The resulting phase drop of the generated signal is much lower than the phase drop of the original measured HRIR below the crossover frequency. This implies a lower group delay, as depicted in Figure 7. Hence the resulting HRIRs have optimized properties concerning frequency response and group delay. In order to eliminate some remaining minor reflection contributions from the floor framework and the VariSphear base plate, the back section of the HRIRs was slightly smoothed using floating average filters. The HRIRs at that point could directly be employed without further processing besides head and tail windowing. The windowing is performed in the processing core within the *miro* data type.

Magnitude and Phase Compensation

In a final optimization step the remaining group delay differences $\Delta\tau_{gr}$ were removed and the minor frequency response deviations (around ±1 dB) of the speaker system were compensated. For that purpose a specific FIR (finite impulse response) filter was designed and applied. FIR filters generally admit to design magnitude and phase responses independently. The respective filter was derived from the post-processed (LFE) center impulse response by appropriate magnitude and phase inversions and windowing in the time domain similar to [12]. The inherent non-causal portions of the FIR compensation filter clear away in the processed signal and the ensuing overall latency can be easily removed. The influence of the compensation filter on the magnitude response of an exemplary HRTF can be observed in Figure 6. The phase compensation has a direct impact on the group delay that is depicted in Figure 7. At this point all substantial magnitude and phase influences of the speaker and the chamber are eliminated and the final HRIRs are contracted to the tightest possible time domain signal. The HRIRs were reduced to 128 taps while maintaining the full spectral bandwidth. (Whereas already around 30 taps are covering the sonic runtime difference between the ears only.) Besides an improved signal integrity due to the phase compensation, the reduced number of filter taps directly entails decreased computational demands for the convolution.

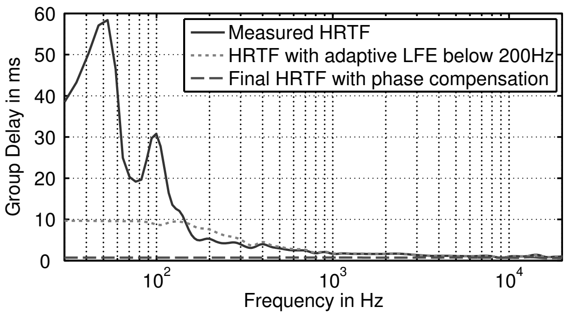

Figure 7: Group delay τ_{gr} of a measured and a post-processed HRTF. The measured HRTF shows a strong surge of τ_{gr} according to the speaker system and additional room resonances in the anechoic chamber at low frequencies (50, 100 Hz). Besides a little increase of τ_{gr} around the crossover frequency due to the filters, the processed HRTF has a considerably reduced $\Delta\tau_{gr}$. The final phase compensated HRTF has a nearly constant τ_{gr} and hence $\Delta\tau_{gr} \approx 0$ leading to a perfectly tight time domain signal.

Conclusion

A free spherical far field HRIR/HRTF compilation of the Neumann KU100 dummy head including several headphone compensation filters is presented. The compilation is useful for applications in spherical acoustics as different dedicated sampling grids are included. It is suitable for high quality audio applications and music production due to the excellent signal quality and the wide frequency response. Owing to the phase compensation, the final HRIRs could be reduced to 128 filter taps while maintaining the full bandwidth. This saves computational power and makes the HRIRs applicable e.g. in mobile devices, where processing power is a highly important issue.

License and Access

Acknowledgements

The author likes to thank Philipp Stade for supporting the headphone measurements and for taking the pictures. Thanks to the MUSIC STORE professional GmbH in Cologne for the hospitality and for providing the headphones to create several compensation filters. Thanks to Markus Haßler from the WDR radio production department for providing the speaker system. The author particularly thanks Christoph Pörschmann, professor for acoustics and audio signal processing at Cologne University of Applied Sciences and head of the research group for his valuable contributions and ideas, and for making all of the presented work possible.

Funding

The research activities are funded by the Federal Ministry of Education and Research in Germany. Support Code: 17009X11 MARA. We appreciate the support.

References

[1] Møller H., "Fundamentals of binaural technology," In: Applied Acoustics 36, pp 171-218, 1992.

[2] William E.G., "Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography," Academic Press, San Diego, 1999.

[3] Bernschütz B., Pörschmann C., Spors S. and Weinzierl S., "SOFiA Sound Field Analysis Toolbox," In: Proceedings of the International Conference on Spatial Audio (ICSA), pp 7-15, 2011.

[4] Geier M., Ahrens J. and Spors S., "The Sound Scape Renderer: A Unified Spatial Audio Reproduction Framework for Arbitrary Rendering Methods," In: Proceedings of the 124th Convention of the AES, Convention Paper 7330, 2008.

[5] Bernschütz B., Pörschmann C., Spors S. and Weinzierl S., "Entwurf und Aufbau eines sphärischen Mikrofonarrays für Forschungsanwendungen in Raumakustik und Virtual Audio," In: Forschschritte der Akustik (DAGA), pp 717-718, 2010.

[6] Theile G., "On the Standardization of the Frequency Response of High-Quality Studio Headphones," In: Journal of the AES, Volume 34(12), pp. 956- 969, 1986.

[7] Schärer Z. and Lindau A., "Evaluation of Equalization Methods for Binaural Signals," In: Proceedings of the 126th Convention of the AES, Preprint No. 7721, 2009.

[8] Blauert J., "Spatial Hearing," The MIT Press, Cambridge/MA, 1997.

[9] Goertz A., "Lautsprecher," In: Handbuch der Audiotechnik, Weinzierl S. (ed.), pp 421-490, 2008.

[10] Blauert J. and Laws P., "Group Delay Distortions in Electroacoustical Systems," In: Journal of the Acoustical Society of America, Volume 63, Number 5, pp 1478-1483, 1978.

[11] Xie B., "On the low frequency characteristics of head-related transfer function," In: Chinese Journal of Acoustics 28(2), pp 1-13, 2009.

[12] Müller S., "Digitale Signalverarbeitung für Lautsprecher," Doctoral Dissertation, RWTH Aachen University, 1999.