Full Reviewed Paper at ICSA 2019

Presented * by VDT.

How positioning inaccuracies influence the spatial upsampling of sparse head-related transfer function sets

Christoph Pörschmann[1,†], Johannes M. Arend[1,2]

[1] *Institute of Communications Engineering, Technische Hochschule Köln, D-50679 Cologne, Germany*

[2] *Audio Communication Group, TU Berlin, D-10587 Berlin, Germany*

[†] *Corresponding author, E-mail: Christoph.Poerschmann@th-koeln.de*

Abstract

Determining full-spherical individual sets of head-related transfer functions (HRTFs) based on sparse measurements is a prerequisite for various applications in virtual acoustics. To obtain dense sets from sparse measurements, spatial upsampling of sparse HRTF sets in the spatially continuous spherical harmonics (SH) domain can be performed by an inverse SH transform. However, this involves artifacts caused by spatial aliasing and order truncation. In a previous publication we presented the SUpDEq method (Spatial Upsampling by Directional Equalization), which reduces these artifacts by a directional equalization prior to the SH transform. Generally, apart from the spatial resolution of the HRTF set, measurement inaccuracies, for example caused by displacements of the head during the measurement, can influence the spatial upsampling as well. By this direction-depending temporal and spectral deviations are added to the dataset, which in the process of spatial upsampling can cause artifacts comparable to spatial aliasing errors. To reduce the influence of the distance inaccuracies, we present a method for distance error compensation that performs an appropriate distance-shifting of the measured HRTFs. Determining the required values for the shift benefits from the directional equalization performed by SUpDEq and results in time-aligning the directionally equalized HRTFs. We analyze the influence of the angular and distance displacements on spectrum, on interaural cues and on modeled localization performance. While limited angular inaccuracies only have a low impact, already small random distance displacements cause strong impairments, which can be significantly reduced applying the proposed distance error compensation method.

1. Introduction

A spatial presentation of sound sources is a fundamental element of virtual acoustic environments (VAEs). For this, monaural and binaural cues, which are mainly caused by the shape of the pinna and the head, need to be considered. In many headphone-based VAEs, head-related transfer functions (HRTFs) are applied to describe the sound incidence from a source, which is typically in the far-field, to the left and right ear incorporating both, monaural and the binaural cues.

A high number of HRTFs is required to adequately capture these cues for all directions of incidence. Complete sets of HRTFs measured on a spherical grid can be described in the spherical harmonics (SH) domain by a decomposition into spherical base functions of different spatial orders N, where higher orders correspond to a higher spatial resolution [15, 18]. Describing sparse HRTF sets in the SH domain results in a limited order and incorporates an incomplete description of the spatial properties. This results in spatial aliasing and truncation errors. To completely consider these properties, an order $N \geq kr$ with $k = \omega/c$, and r being the head radius is required [7, 14]. Performing a nearly perfect

* Please note that the papers at ICSA can be published by VDT, in print, online and as PDF download.

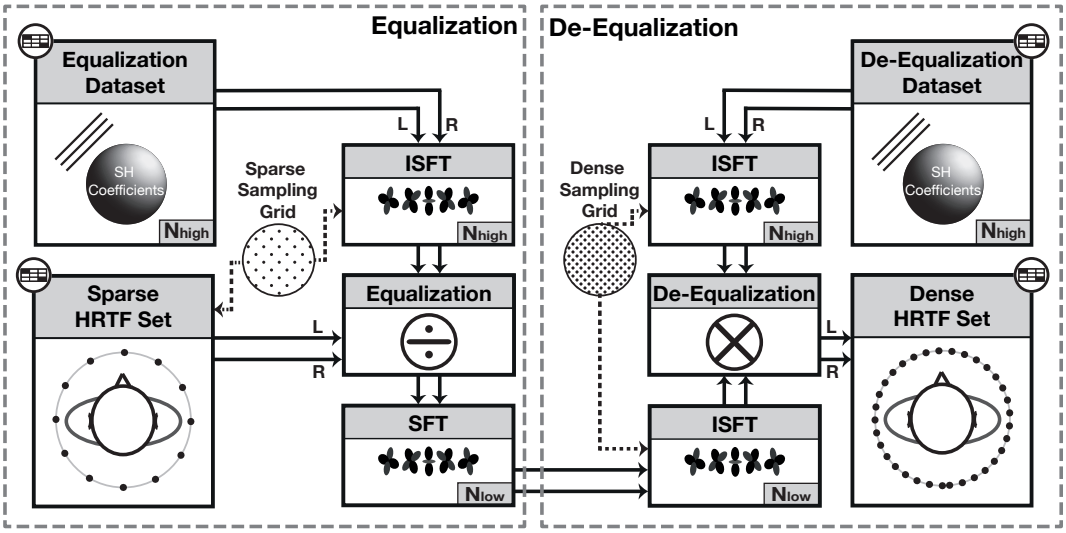


Fig. 1: Block diagram of the SUpDEq method. Left panel: A sparse HRTF set is equalized on the corresponding sparse sampling grid before transformed to the SH domain with $N = N_{low}$. Right panel: The equalized set is de-equalized on a dense sampling grid. If required, the resulting dense HRTF set can again be transformed to the SH domain with $N = N_{high}$.

interpolation for frequencies up to 20 kHz leads to $N = 32$ requiring at least 1089 measured directions when assuming $r = 8.75$ cm and $c = 343$ m/s.

Different studies analyzed artifacts caused by spatial upsampling of sparsely measured HRTF sets to a dense sampling grid or examined methods to reduce these artifacts (e.g. [5, 7, 9, 19]). In this context, we recently introduced the SUpDEq (Spatial Upsampling by Directional Equalization) method [12], which removes frequency-dependent ITDs and ILDs as well as elevation-dependent spectral features from the HRTFs. For this we apply spectral division (equalization) to the HRTFs with a corresponding equalization function prior to the SH transform. A directional rigid sphere transfer function (STF) can be used here as equalization function, resulting in a significantly reduced spatial order N. After spatial upsampling, a de-equalization by means of a spectral multiplication with the same equalization function is performed to recover a spatially upsampled HRTF set.

Generally speaking, the use of the proposed method is especially advantageous for measuring sets of individual HRTFs which, for example, provide a better localization accuracy in the median plane than non-individual ones [8]. However, measuring such datasets in a simple procedure with a cheap measurement setup, and under non-ideal room-acoustical conditions is a challenging task. In previous papers we already analyzed the suitability of the SUpDEq method for individual HRTF sets [13] and investigated to what extent the use of low-cost loudspeakers in reflective environments affects the HRTF measurements [11].

In this paper we analyze another critical issue. The positioning accuracy of the subject in the array during the HRTF measurement can result in distance and angular errors. These inaccuracies can on the one hand be caused by a constant shift of the listener's center position while measuring. For example, in [4] it is shown that such shifts of the listener position significantly increase the required spatial order of the HRTF set. To compensate for this, methods for recentering the receiver by appropriate postprocessing of the measured dataset have been developed [16]. On the other hand these shifts can be non-systematic and thus be independent of the measured directions, resulting in (nearly) randomly distributed distance and angular inaccuracies. Such deviations in the measured HRTFs might be observed with sequential HRTF measurements. Furthermore, when tracking the listener's position and orientation in such a procedure, the inaccuracies of the tracking device can as well be regarded as being randomly distributed.

To investigate the influence of positioning inaccuracies on spatial upsampling, we performed a study which compares spatially upsampled sparse HRTF sets to a reference sampled on a dense grid. We analyze the influence of angular and distance errors regarding spectral differences, binaural cues and modeled localization performance. Furthermore, we investigate to what extent a method compensating the distance errors can enhance the performance. Thus the result of this study can help to obtain required boundary conditions for performing HRTF measurements.

2. Method

The SUpDEq method has been described and evaluated in detail in [12]. In the following we thus only briefly outline the basic concept. The corresponding block diagram is given in Fig. 1. First, the sparse HRTF set H_{HRTF} measured at S sampling points $\Omega_s = \{(\phi_1, \theta_1), \ldots, (\phi_S, \theta_S)\}$ is equalized direction-dependently with an appropriate equalization dataset H_{EQ}

$$H_{HRTF,EQ}(\omega, \Omega_s) = \frac{H_{HRTF}(\omega, \Omega_s)}{H_{EQ}(\omega, \Omega_s)} \,. \tag{1}$$

While generally different equalization datasets can be applied, in this study a rigid sphere transfer function (STF) is used

which describes an incoming wave on a rigid sphere [18, p. 227]. The radius of the sphere corresponds to the physical dimensions of a human head and an ear position of $\phi = \pm 90°$ and $\theta = 0°$ is considered. The STF can thus be regarded as a simplified HRTF set featuring basic temporal and spectral components but without information on the shape of the outer ears or the fine structure of the head. Thus, by the equalization a time-alignment of the HRTFs is performed and direction-dependent influences of the spherical shape of the head are compensated. The equalization with the STF indeed leads to considerably reduced spatial dependency in $H_{HRTF,EQ}$ and aims at minimizing the required order for the SH transform. As the equalization dataset H_{EQ} can be described analytically, it can be determined at a freely chosen maximal order, typically $N_{high} \geq 35$. The SH coefficients for the equalized sparse HRTF set $H_{HRTF,EQ}$ are obtained by applying the SH transform to the equalized HRTFs up to an appropriate low maximal order N_{low}, which corresponds to the maximal order that can be resolved by Ω_s. Then an upsampled HRTF set $\widehat{H}_{HRTF,EQ}$ is calculated on a dense sampling grid $\Omega_d = \{(\phi_1, \theta_1), \ldots, (\phi_D, \theta_D)\}$, with $D \gg S$ by using the inverse SH transform. Finally, HRTFs are reconstructed by a subsequent de-equalization by means of spectral multiplication with a de-equalization dataset H_{DEQ}

$$\widehat{H}_{HRTF,DEQ}(\omega, \Omega_d) = \widehat{H}_{HRTF,EQ}(\omega, \Omega_d) \cdot H_{DEQ}(\omega, \Omega_d)\,. \quad (2)$$

For de-equalization, again the STF can be used. This last step recovers energies at higher spatial orders that were transformed to lower orders in the equalization. $H_{HRTF} = \widehat{H}_{HRTF,DEQ}$ holds if N_{low} and N_{high} are chosen appropriately. Energy which, after the equalization, still is apparent at high modal orders $N > N_{low}$ is irreversibly mirrored to lower orders $N \leq N_{low}$. Thus we obtain $H_{HRTF} \approx \widehat{H}_{HRTF,DEQ}$.

3. HRTF Datasets

We used HRTFs of a Neumann KU100 dummy head measured on a dense Lebedev grid with 2702 sampling points which can be used for SH processing at a sufficient order of $N = 35$ for the evaluation [6]. The SH representation of the dataset served as the reference in our investigations. From this reference set we generated various sparse HRTF sets which were required as input data for the evaluation. First the sparse HRTF sets varied regarding the accessible spatial order N. These sets were obtained in the same way as described in [12] by spatially subsampling the reference set in the SH domain by means of the inverse SH transform. Furthermore, in order to create datasets incorporating positioning inaccuracies, we randomly varied the distance for each measured direction in a range of $\pm\Delta r_{max}$ to the reference distance of $R = 2\,m$. To perform the distance shifts we used a method which is based on the SUpDEq method [2]. Instead of an incident plane wave (representing a sound source in the far-field), an STF for a spherical wave (point source) at the reference distance of $R = 2\,m$ is used for the equalization and a spherical wave at $R' = R + \Delta r$ for the de-equalization. By this, both the phase and the amplitude are appropriately adapted to the changed distance. To consider angular inaccuracies, we randomly modified the directions for which we determined the HRTFs from the dense HRTF set equally-distributed within a solid angle of $\Delta\phi_{max}$.

It is worth noting that we chose randomly distributed deviations because they showed in informal pretests the highest impact on the spatial upsampling. Furthermore, such deviations are on the one hand typical, when the subject moves slightly between each of the sequentially measured directions or turns the head not exactly to the target direction. On the other hand measurement errors of a head-tracking device, used to determine the exact subject position and orientation during an HRTF measurement can be regarded as well being randomly distributed.

Accordingly we created datasets considering maximal distance deviations of $\Delta r_{max} = 1\,cm, 2\,cm, 5\,cm$ and maximal angular deviations of $\Delta\phi_{max} = 2°, 5°, 15°$ by spatially downsampling of the reference set for 15 sparse sampling grids – Lebedev grids with 6, 14, 26, 38, 50, 74, 86, 110, 146, 170, 194, 230, 266, 302, and 350 sampling points – equaling (limited) orders of $N = 1 - 15$. For each of these conditions, we generated SH coefficients which we used for the further evaluation. Thus, both order-limited and de-equalized sets were always based on the respective sparse grid.

While the order-limited (OL) datasets were obtained with an SH interpolation without any pre- or postprocessing, we used the Matlab-based implementation of the SUpDEq method as described in [12] to obtain the de-equalized HRTF sets (DEQ). The radius for the rigid sphere model was calculated according to Algazi et al. [1] based on the dimensions of the dummy head, resulting in a radius of $r = 9.19$ cm. Finally the HRTFs of the test grids used in the evaluations were obtained via the inverse SH transform of the order-limited dataset or the de-equalized dataset at the respective positions.

4. Evaluation

4.1. Spectrum

First we analyze the spectral deviations to the reference set as a function of N on a Lebedev grid with $T = 2702$ sampling points as test sampling grid $\Omega_t = \{(\phi_1, \theta_1), \ldots, (\phi_T, \theta_T)\}$. For this the frequency-dependent spectral differences per sampling point were calculated in dB as

$$\Delta g(\omega, \Omega_t) = 20lg \frac{\mid H_{HRTF,REF}(\omega, \Omega_t) \mid}{\mid H_{HRTF,TEST}(\omega, \Omega_t) \mid}, \quad (3)$$

where $H_{HRTF,REF}$ is the left ear HRTF extracted from the reference set and $H_{HRTF,TEST}$ the one extracted from the order-limited or the de-equalized datasets at each sampling point Ω_t. Then the absolute value of $\Delta g(\omega, \Omega_t)$ was averaged across across all sampling points Ω_t to obtain the frequency-dependent measure $\Delta G_f(\omega)$ (in dB)

$$\Delta G_f(\omega) = \frac{1}{n_{\Omega_t}} \sum_{\Omega_t=1}^{n_{\Omega_t}} \mid \Delta g(\omega, \Omega_t) \mid, \quad (4)$$

and across ω and Ω_t, resulting in a single value ΔG (in dB) describing the spectral difference

$$\Delta G = \frac{1}{n_{\Omega_t}} \frac{1}{n_\omega} \sum_{\Omega_t=1}^{n_{\Omega_t}} \sum_{\omega=1}^{n_\omega} \mid \Delta g(\omega, \Omega_t) \mid . \quad (5)$$

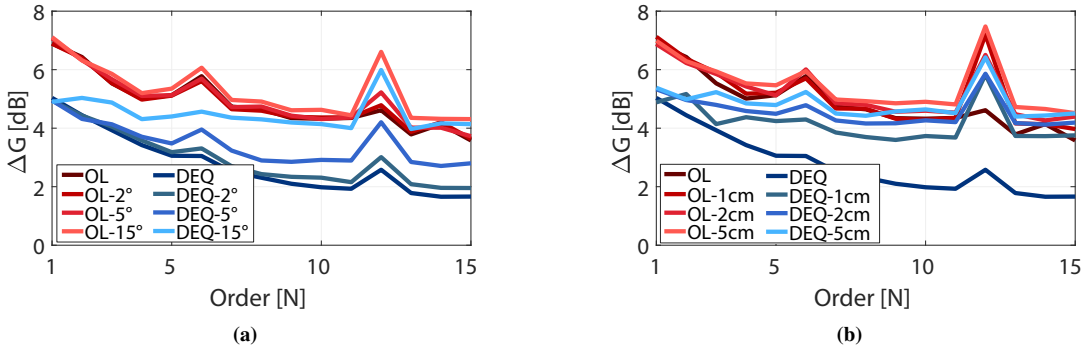


Fig. 2: Mean spectral differences ΔG in dB (left ear) between reference HRTF set ($N = 35$) and the datasets with angular and distances inaccuracies depending on the order N. Red: order-limited datasets (OL), Blue: de-equalized datasets (DEQ). (a) Influence of the angular inaccuracies $\Delta\phi_{max}$, (b) Impact of distance inaccuracies Δr_{max}. The color saturation corresponds to the size of the error.

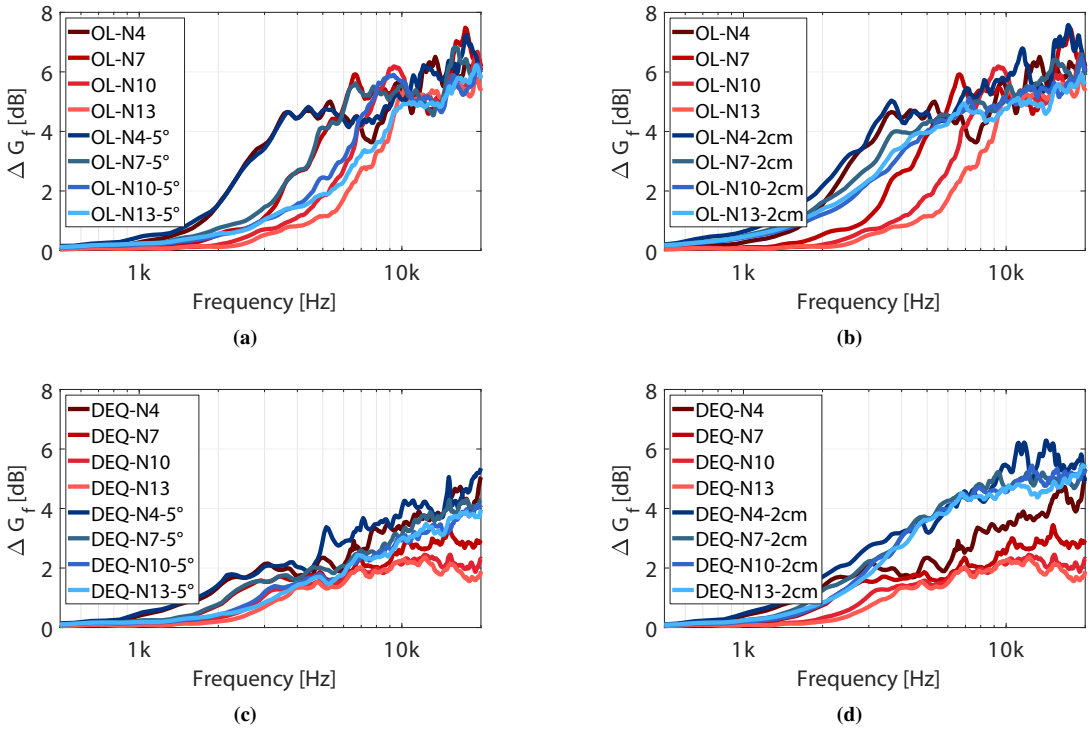


Fig. 3: Spectral differences $\Delta G_f(\omega)$ in dB (left ear) of spatially upsampled datasets (color saturation corresponds to N) to the reference HRTF set ($N = 35$). Red: datasets without positioning inaccuracies, Blue: datasets with angular errors or distance inaccuracies. (a,b) Results for the order-limited sets (OL), (c,d) Results for the de-equalized sets (DEQ). (a,c) Influence of the angular alteration $\Delta\phi_{max} = 5°$, (b,d) Influence of the distance shifts $\Delta r_{max} = 2cm$.

Fig. 2 shows ΔG depending on the order N both for the order-limited datasets and the de-equalized datasets, that means without and with the SUpDEq processing. Data for angular deviations of 2°, 5° and 15° as well as for distance deviations of $1\,cm$, $2\,cm$ and $5\,cm$ is given. Generally, for the order-limited datasets, angular and distance deviations have only a minor influence on ΔG. For de-equalized HRTF sets angular inaccuracies of up to $\Delta\phi_{max} = 5°$ result in a slight increase of ΔG. Only for $\Delta\phi_{max} = 15°$ a strong influence already at low orders N can be observed (Fig. 2 a). On the contrary, distance inaccuracies strongly affect the de-equalized datasets (Fig. 2 b). Already for small Δr_{max}, the spectral differences increase, especially at higher spatial orders N. For example, at $N = 7$ and $\Delta r_{max} = 1\,cm$ the increase is about 1.5 dB.

Fig. 3 shows the spectral differences over frequency. While the influence of angular deviations is minor for $N = 4$ and $N = 7$ it causes an increase for higher orders. However, the deviations are below 1 dB in all cases for frequencies up to 10 kHz. On the contrary, the influences of distance inaccuracies are much larger. Errors of $\Delta r_{max} = 2\,cm$ strongly deteriorate the spectrum both for the de-equalized and the order-limited datasets. Furthermore, the inaccuracies nearly completely outweigh the benefit of the SUpDEq method, especially for higher orders. Thus, when performing HRTF measurements, distance inaccuracies between the sound source and the human head need to be avoided.

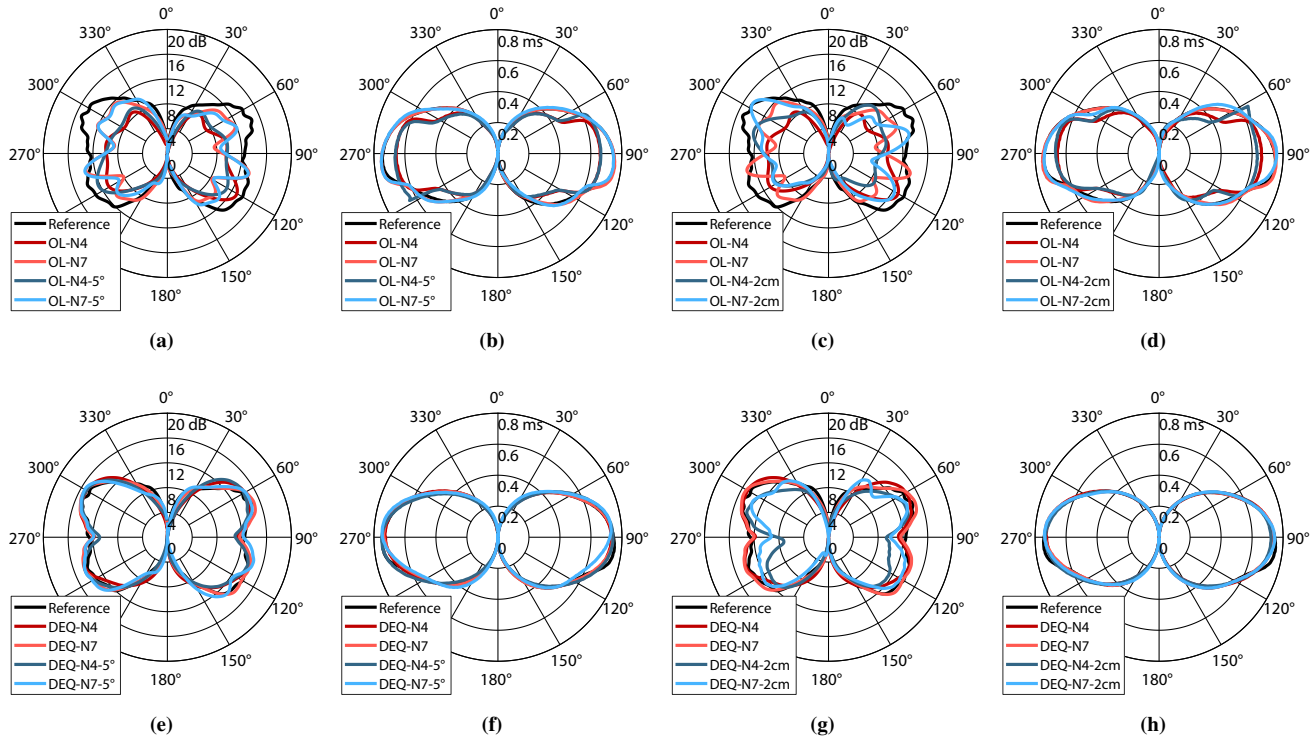


Fig. 4: ILDs and ITDs in the horizontal plane. Black: Reference HRTF set, Red: Datasets without positioning inaccuracies, Blue: Datasets with altered angles $\Delta\phi_{max} = 5°$ or distances $\Delta r_{max} = 2\,cm$. In the upper line (a–d) the results for the order-limited datasets (OL) are shown, in the lower one (e–h) the ones for the de-equalized sets (DEQ). The angle represents the azimuth ϕ of the sound source. The radius describes the magnitude of the level differences (in dB) or time differences (in ms). The left two rows (a,b,e,f) shows the results for the angular deviations and the right rows (c,d,g,h) the ones for the distance inaccuracies.

4.2. Binaural cues

In Fig. 4 the interaural level differences (ILDs) and the interaural time differencs (ITDs) are shown. For the order-limited sets (Fig. 4 (a–d)), differences of the ITDs and ILDs to the reference vary most depending on the spatial order N. The maximal deviations from the reference are up to 4 dB for the order-limited sets at $N = 7$ and more than 8 dB at $N = 4$. For the lateral sound incidence variations in ITDs of up to 0.1 ms at $N = 4$ occur. While the influence of the angular inaccuracies is minor, the distance error affects the ILDs, but however does not generally lead to a strong increase of the error. As shown in Fig. 4 (e–h) for the de-equalized datasets (DEQ), the ITDs are only slightly affected by the positioning inaccuracies. On the contrary, the considered distance inaccuracies of 2 cm strongly influence the ILDs and lead for lateral sound incidence to deviations of 4 dB at $N = 4$ and of about 2 dB at $N = 7$.

4.3. Localization performance

To analyze the impact of the distance and angular inaccuracies on localization performance in the median sagittal plane, we used the model from Baumgartner et al. [3]. The model compares the spectral structure of a reference HRTF set to a set of test HRTFs and calculates a probabilistic estimate of the perceived sound source location. Based on this estimate, the polar RMS error is determined which describes the expected angular error between the actual and perceived source positions. Additionally, it determines the quadrant error rate specifying the rate of front-back or up-down confusions. To estimate the localization performance in the horizontal plane, we used the model from May et al. [10] which weighs the frequency-dependent binaural cues (ILDs, ITDs) to estimate the azimuthal position of a sound source based on a trained Gaussian mixture model. A lateral error can be calculated by comparing the intended and the estimated source position. We used the Auditory Modeling Toolbox (AMT) [17] for these calculations. The procedure for determining the errors has been described in detail in [12] and is in the following briefly outlined. We used a test sampling grid Ω_t with $\phi = \{0°, 180°\}$ and $-30° \leq \theta \leq 90°$ in steps of 1° to estimate median plane localization performance and assumed a median listener sensitivity of $S = 0.76$ (in accordance with Baumgartner et al. [3]). For the horizontal plane localization performance, we applied a test sampling grid with $\phi = \pm 90°$ in steps of 5°. We calculated the absolute polar error difference (PE in degree)

$$\Delta PE = \mid PE_{REF} - PE_{TEST} \mid, \tag{6}$$

the absolute quadrant error difference (QE in percent)

$$\Delta QE = \mid QE_{REF} - QE_{TEST} \mid, \tag{7}$$

as well as the absolute lateral error difference (LE in degree)

$$\Delta LE = \frac{1}{T} \sum_{t=1}^{T} \mid LE_{REF}(\Omega_t) - LE_{TEST}(\Omega_t) \mid, \tag{8}$$

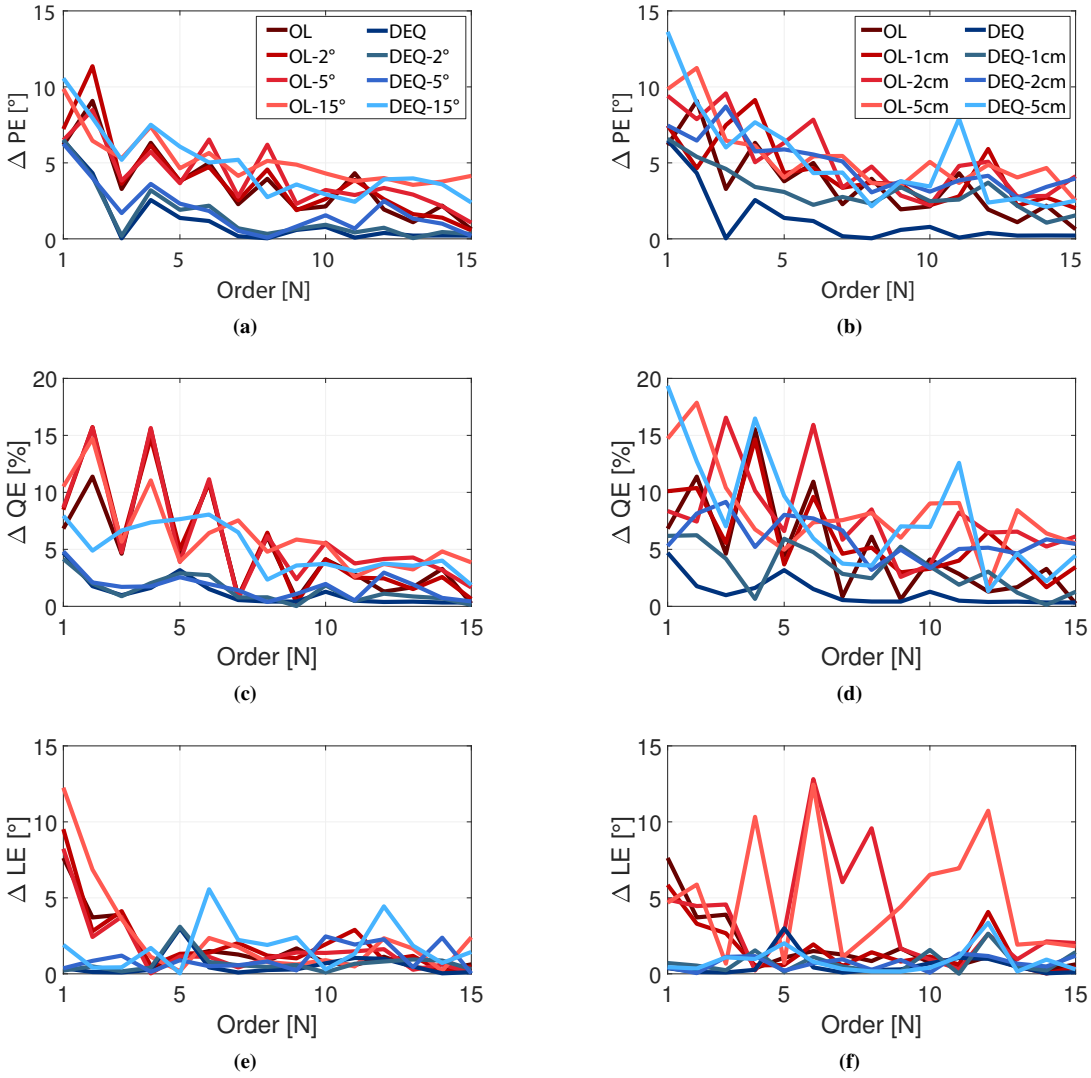


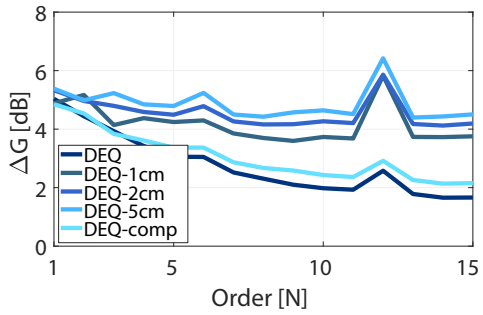
Fig. 5: Influence of distance and angular error on modeled localization performance. (a,b) Absolute polar error difference ΔPE, (c,d) Quadrant error difference ΔQE, (e,f) Lateral error difference ΔLE over SH order N. Red: Results for the order-limited datasets (OL) , Blue: De-equalized datasets (DEQ). Left row: (a,c,e) Influence of the angular inaccuracies $\Delta\phi_{max}$, Right row: (b,d,f) Impact of distance inaccuracies Δr_{max}. The color saturation corresponds to the size of the angular respectively the distance error.

for each order N with the subscripts *REF* for the reference set and *TEST* for the tested HRTF set.

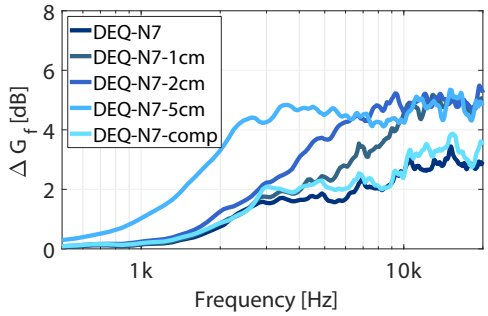
As illustrated in Fig. 5 (a, c, e) angular inaccuracies only slightly affect localization performance of the order-limited datasets. For the de-equalized datasets the localization performance is strongly affected only for angular inaccuracies of $\Delta\phi_{max} = 15°$. This is completely different for the distance inaccuracies which are shown in Fig. 5 (b, d, f). In the median sagittal plane already for $\Delta r_{max} = 2\,cm$, both the polar error difference ΔPE and the quadrant error difference ΔQE are strongly increased. For some of the data, the localization errors are even stronger for the de-equalized datasets than for the order-limited datasets. Thus as already analyzed based on the spectral differences (see Sec. 4.1), distance inaccuracies have a severe impact on the spatial upsampling. However, the increase of the lateral error difference ΔLE is quite small at least for the de-equalized datasets.

5. Distance error compensation

A major outcome of the present study is that already small (random) inaccuracies in the distance between sound source and listener strongly influence the spatial upsampling, especially when applying the SUpDEq method. As such positioning inaccuracies might not be systematic but somehow randomly distributed over the different measured directions they cannot be compensated using the recentering methods proposed e.g. by Richter et al. [16]. In the following we describe and examine the so-called DEC (Distance Error Compensation) method which reduces the influence of distance inaccuracies. The method is to some extent comparable to the approach from Ziegelwanger and Majdak [20] splitting up the HRTF in a direction-dependent part representing the influence of the sphere and direction-independent part. However, our implementation benefits from the directional equalization which is part of the SUpDEq method. As explained in Sec. 2, the spectral equalization (Eq. 1) removes direction-dependent



(a)

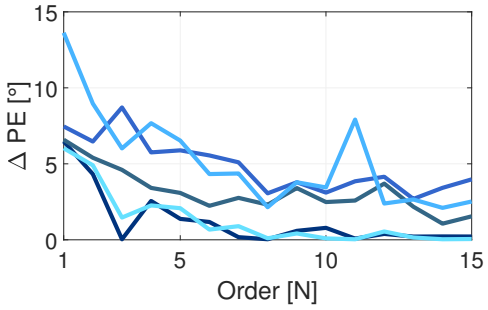


(b)

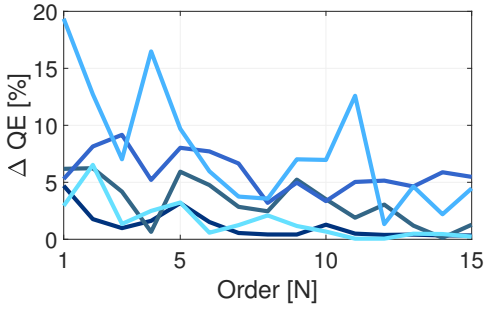
Fig. 6: Influence of the compensation on the upsampled HRTF sets on the mean spectral differences to the reference HRTF set ($N = 35$). (a) ΔG depending on the order N. (b) $\Delta G_f(\omega)$ over frequency for $N = 7$. In addition to the compensated HRTF set (comp) the values for the measurement without positioning inaccuracies and for different distance inaccuracies Δr_{max} are given (color saturation).

spectral and temporal components from the measured HRTFs. By this head-related differences are removed in $H_{HRTF,EQ}$ and ideally all peaks of the equalized head-related impulse responses (HRIR, the time-domain equivalent of an HRTF) are time-aligned. However, deviations due to the positioning inaccuracies remain after the equalization and thus the onset differences between the different equalized HRIRs directly relate to the distance errors. Thus, a simple onset-detection of the spatially equalized equalized HRIRs is applied to estimate the distance errors and to determine the required distance shift Δr. To apply the distance shift we use the directional equalization and subsequent de-equalization at different distances as already described in Sec. 3. For each measured HRTF a point source STF at the distance of $R_{error} = R + \Delta r$ is used for the equalization and a point source STF at the reference distance of $R = 2\,m$ for the de-equalization. After performing the DEC, the peaks of all equalized HRIRs are time-aligned. We implemented the DEC as a separate preprocessing step performed on the sparse HRTF set (see Fig. 1). A ten-times oversampling was applied for more precise onset detection and subsample accuracy. To be robust against noise we determined the onset based on -1 dB related to the maximal value of the equalized HRIRs.

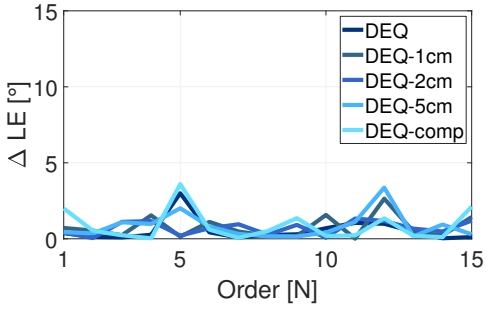
The result of the compensation is given in Fig. 6 showing that the spectral differences are significantly reduced compared to the different distance inaccuracies Δr_{max} and are only slightly higher than for the dataset not comprising any positioning inaccuracies. The remaining differences are



(a)



(b)



(c)

Fig. 7: Influence of the compensation on the localization errors for the de-equalized datasets (DEQ). Absolute polar error difference ΔPE (a), quadrant error difference ΔQE (b), and lateral error difference ΔLE (c) over SH order N for the compensated dataset (comp), the measurement without positioning inaccuracies and for varying distance inaccuracies Δr_{max} (color saturation).

mainly caused by the non-perfect directional equalization. Due to differences between the measured HRTF set and the equalization dataset some differences in the temporal structure remain, which are unintentionally included in the compensation. In Fig. 7 the impact of the distance error compensation on the localization performance is shown. The polar error difference ΔPE (Fig. 7(a)), the quadrant error difference ΔQE (Fig. 7(b)), and the lateral error difference ΔLE (Fig. 7(c)) of the compensated datasets are as well very close to the sets without positioning inaccuracies. Thus an appropriate compensation of the distance inaccuracies as proposed in this paper can minimize localization errors of upsampled HRTF datasets.

6. Conclusion

In this paper we evaluated the influence of positioning inaccuracies of measured sparse HRTF sets on spatial upsampling.

For this we analyzed the impact of angular and distance errors on spectral cues, binaural cues and modeled localization performance for different spatial orders. The results can be summarized as follows. First, distance inaccuracies have much stronger impact than angular inaccuracies. Second, the effect of distance inaccuracies are much higher for the de-equalized sets than for the order-limited HRTF sets. Third, the influence of these inaccuracies becomes stronger with an increasing spatial resolution of the sparse HRTF set. Fourth, an appropriate distance error compensation (DEC), which applies a distance shift based on a time-alignment of the equalized HRIRs, can nearly completely eliminate the influence of distance inaccuracies. This way, distance inaccuracies in the HRTF sets have only minor impact on the spectral deviations and on the binaural cues. Examining modeled localization performance showed, that after the DEC nearly no influence of the investigated distance inaccuracies on localization performance remains.

Thus results show that the SUpDEq method can be used with simple and non-optimal measurement equipment incorporating angular and distance inaccuracies. Of course these findings need to be validated perceptually and based on measured individual datasets. To further validate the applicability of the method, a demonstration system needs to be set up allowing to measure spherical sparse HRTF sets.

The research presented in this paper has been funded by the German Federal Ministry of Education and Research. Support Code: BMBF 03FH014IX5-NarDasS. A Matlab-based implementation of the SUpDEq method is available on https://github.com/AudioGroupCologne/SUpDEq.

7. References

[1] V. Algazi, C. Avendano, and R. O. Duda. Estimation of a Spherical-Head Model from Anthropometry. *J. Audio Eng. Soc.*, 49(6):472 – 479, 2001.

[2] J. M. Arend and C. Pörschmann. Synthesis of Near-Field HRTFs by Directional Equalization of Far-Field Datasets. In *Proceedings of the 45th DAGA*, pages 1454–1457, 2019.

[3] R. Baumgartner, P. Majdak, and B. Laback. Modeling sound-source localization in sagittal planes for human listeners. *J. Acous. Soc. Am.*, 136(2):791–802, 2014.

[4] I. Ben Hagai, M. Pollow, M. Vorländer, and B. Rafaely. Acoustic centering of sources measured by surrounding spherical microphone arrays. *J. Acous. Soc. Am*, 130(4):2003–2015, 2011.

[5] Z. Ben-Hur, F. Brinkmann, J. Sheaffer, S. Weinzierl, and B. Rafaely. Spectral equalization in binaural signals represented by order-truncated spherical harmonics. *J. Acous. Soc. Am.*, 141(6):4087–4096, 2017.

[6] B. Bernschütz. A Spherical Far Field HRIR / HRTF Compilation of the Neumann KU 100. In *Proceedings of the 39th DAGA*, pages 592–595, 2013.

[7] B. Bernschütz, A. Vázquez Giner, C. Pörschmann, and J. M. Arend. Binaural reproduction of plane waves with reduced modal order. *Acta Acustica united with Acustica*, 100(5):972–983, 2014.

[8] J. Blauert. *Spatial Hearing - The Psychophysics of Human Sound Localization*. MIT Press, Cambridge, MA, revised edition, 1996.

[9] F. Brinkmann and S. Weinzierl. Comparison of head-related transfer functions pre-processing techniques for spherical harmonics decomposition. In *Proceedings of the AES Conference on Audio for Virtual and Augmented Reality*, pages 1–10, 2018.

[10] T. May, S. Van De Par, and A. Kohlrausch. A probabilistic model for robust localization based on a binaural auditory front-end. *IEEE Trans. Audio, Speech, Lang. Process.*, 19(1):1–13, 2011.

[11] C. Pörschmann and J. M. Arend. Obtaining Dense HRTF Sets from Sparse Measurements in Reverberant Environments. In *Proceedings of the AES Conference on Immersive and Interactive Audio*, 2019.

[12] C. Pörschmann, J. M. Arend, and F. Brinkmann. Directional Equalization of Sparse Head-Related Transfer Function Sets for Spatial Upsampling. *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, 27(6):1060 – 1071, 2019.

[13] C. Pörschmann, J. M. Arend, and F. Brinkmann. Spatial upsampling of individual sparse head-related transfer function sets by directional equalization. In *Proceedings of the 23rd International Congress on Acoustics*, 2019.

[14] B. Rafaely. Analysis and Design of Spherical Microphone Arrays. *IEEE Trans. on Speech and Audio Proc.*, 13(1):135–143, 2005.

[15] B. Rafaely. *Fundamentals of Spherical Array Processing*. Springer-Verlag, Berlin Heidelberg, 2015.

[16] J. G. Richter, M. Pollow, F. Wefers, and J. Fels. Spherical harmonics based hrtf datasets: Implementation and evaluation for real-time auralization. *Acta Acustica united with Acustica*, 100(4):667–675, 2014.

[17] P. Søndergaard and P. Majdak. The Auditory Modeling Toolbox. In J. Blauert, editor, *The Technology of Binaural Listening*, pages 33–56. Springer-Verlag, Berlin Heidelberg, 2013.

[18] E. G. Williams. *Fourier Acoustics - Sound Radiation and Nearfield Acoustical Holography*. Academic Press, London, UK, 1999.

[19] M. Zaunschirm, C. Schoerkhuber, and R. Hoeldrich. Binaural rendering of Ambisonic signals by HRIR time alignment and a diffuseness constraint. *J. Acous. Soc. Am.*, 143(6):3616 – 3627, 2018.

[20] H. Ziegelwanger and P. Majdak. Modeling the direction-continuous time-of-arrival in head-related transfer functions. *J. Acous. Soc. Am.*, 135(3):1278–1293, 2014.