Journal of the Audio Engineering Society
Vol. 65, No. 6, June 2017 (© 2017)
DOI: https://doi.org/10.17743/jaes.2017.0012



# Perceptual Evaluation of Individualized Binaural Reproduction Using a Virtual Artificial Head

**EUGEN RASUMOW[1], MATTHIAS BLAU[1], SIMON DOCLO,[2]** *AES Associate Member*,
**STEVEN VAN DE PAR,[2]** *AES Member*, **MARTIN HANSEN[1], DIRK PUSCHEL[3], AND VOLKER MELLERT[2]**

[1]*Jade University of Applied Sciences, Oldenburg, Germany (eugen.rasumow@web.de)*
[2]*Carl von Ossietzky University of Oldenburg, Oldenburg, Germany*
[3]*Akustik Technologie Göttingen, Göttingen, Germany*

In binaural recordings, spatial information can be captured by using so-called artificial heads, which are replicas of real human heads with ear microphones and average anthropometric geometries. Mainly because of their non-individual character, such recordings often entail perceptual deficiencies (front-back confusion, internalization, etc.). Alternatively, individually measured head-related transfer functions (HRTFs) can be approximately synthesized using a microphone array in conjunction with a *filter-and-sum* beamformer (referred to as virtual artificial head, VAH). Its main advantage over traditional artificial heads is the possibility to subsequently adapt one recording to individual HRTFs in the reproduction stage by an appropriate modification of the directivity pattern of the VAH. Further advantages of the VAH include its smaller size and weight. The realization of a VAH as a planar microphone array with 24 microphones has been presented in previous studies. In this study binaural reproductions using the VAH, two traditional artificial heads, and individual HRTFs were perceptually evaluated in the horizontal plane with respect to the original free-field presentation. Evaluations were conducted for directions explicitly considered in the optimization of the VAH-filter coefficients but also for intermediate directions, which are assumed to be interpolated implicitly by the VAH. The ratings confirm the validity of the concept of synthesizing HRTFs using a VAH. It is found that the VAH-synthesis enables sufficiently good binaural reproductions that in general yield better perceptual ratings in comparison to traditional artificial heads for explicitly considered directions and approximately equivalent ratings for intermediate directions.

## 0 INTRODUCTION

Spatial information is an important factor for the perception and appraisal of sounds. Spatial information can be introduced into recordings and measurements (to some extent) by using so-called artificial heads, which are reproductions of real human heads with microphones placed in the ear canals (cf., [1] for an extensive review). Alternatively, the direction- and frequency-dependent head-related transfer functions (HRTFs) can be approximately synthesized using a set of spatially distributed microphones with appropriate digital filtering (cf., [2], [3], [4], [5], and [6—8]). Such a device is referred to as a virtual artificial head (VAH). The main advantages of a VAH are the possibility of adjusting the filter coefficients to HRTFs of different listeners (individualization) and to different look directions (orientation), the possibility of employing head tracking in the reproduction stage and a better flexibility and manageability due to the smaller size/weight of the device.

In our implementation of the VAH, the filter coefficients are optimized using measured steering vectors and HRTFs for discrete directions, assuming that the intermediate directions will be interpolated implicitly by the VAH. Alternatively, the optimization could be done in the spherical harmonics domain, where spatial information is represented using model assumptions (cf., [5], [9], and [10]). The optimization using measured steering vectors and HRTFs for discrete directions was chosen in this study since spherical harmonics-based strategies generally need many more HRTF measurements and especially more microphones (cf., [11] and [12]), which has prevented this approach from being implemented in working devices so far.

In order to enable the best possible synthesis of individual HRTFs in our approach, various studies were performed in advance to elaborate an appropriate microphone topology (cf., [6]), an appropriate cost function (cf., [7]), appropriate regularization strategies (cf., [8]), appropriate regularization parameters (cf., [13]), and an appropriate smoothing of HRTFs (cf., [14] and [15]) prior to the synthesis.

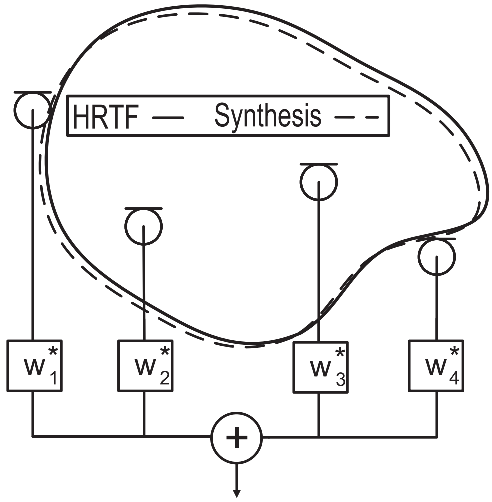


Fig. 1. **Filter-and-sum beamformer.** Schematic diagram of a filter-and-sum beamformer with N=4 microphones and the associated filter coefficients $\mathbf{w} = [w_1 \ldots w_4]^T$. Desired directivity pattern (HRTF) and resulting directivity pattern (Synthesis).

The present study is aimed at perceptually assessing the quality of the VAH-synthesis for explicitly considered directions, but also for intermediate directions, by comparing binaural presentations to free-field presentations. The free-field presentation from (single) loudspeakers in an anechoic room served as the reference condition. Binaural presentations via headphones (with headphone equalization) using the following four test setups were perceptually evaluated in comparison to the reference condition: binaural reproductions using individual HRTFs, individualized synthesis using the VAH, and non-individual presentations using two traditional artificial heads ($DH_1$ and $DH_2$), one of which is commercially available. It is to be expected that binaural reproductions using individual HRTFs should result in the best evaluations in the following experiments. The two traditional artificial heads were chosen in order to classify the ratings of the VAH-synthesis in comparison to the method that presumably is most commonly used for binaural reproductions.

This paper is organized as follows: First, the principle of the VAH and the implementation used in this study are reviewed in Sec. 1. Second, the perceptual evaluation is presented in Secs. 2 and 3. This is then followed by a discussion of the results in Sec. 4 and the conclusions in Sec. 5.

## 1 REVIEW OF THE VIRTUAL ARTIFICIAL HEAD

### 1.1 Principle and Optimization of the Virtual Artificial Head

Using the filter-and-sum method in conjunction with an appropriate microphone array and filter coefficients, it is possible to synthesize desired directivity patterns. The synthesized spatial directivity pattern (see Fig. 1) of the VAH $H(f, \theta)$ for direction $\theta$ and frequency $f$ can be expressed as[1]

$$H(f, \theta) = \mathbf{w}^H(f)\mathbf{d}(f, \theta), \tag{1}$$

with the N× 1 steering vector $\mathbf{d}(f, \theta)$ representing the frequency- and direction-dependent transfer functions between a source and the N microphones and $\mathbf{w}(f)$ the N× 1 vector containing the complex-valued filter coefficients for the N microphones at frequency $f$. Given the motivation to synthesize the directivity patterns of individual HRTFs, the filter coefficients $\mathbf{w}(f)$ for the VAH may be optimized by minimizing a chosen cost function between the desired directivity pattern (individual HRTFs) and the synthesized directivity pattern H. One common choice is to use a narrowband least squares cost function $J_{LS}(\mathbf{w}(f))$, being the sum over P discrete directions of the squared absolute differences between the synthesis $H(f, \theta)$ and the desired $HRTF(f, \theta)$, i.e.,

$$J_{LS}(\mathbf{w}(f)) = \sum_{i=1}^{P} F(f, \theta_i) \cdot \left|\mathbf{w}^H(f)\mathbf{d}(f, \theta_i) - HRTF(f, \theta_i)\right|^2, \tag{2}$$

with $F(f, \theta)$ being the real-valued positive weights assigning more or less importance to certain directions. Minimizing this cost function would lead to the filter coefficients $\mathbf{w}(f)$ for a single frequency $f$. Based on the motivation to mimic the spectral grouping of the human ear, in [8] a least squares cost function incorporating multiple frequency bins was presented as

$$J_v(\mathbf{w}_v(\Omega)) = \sum_{l=1}^{L} J_{LS}(\mathbf{w}(f_l)) = \sum_{l=1}^{L}\sum_{i=1}^{P} F(f_l, \theta_i) \cdot \left|\mathbf{w}^H(f_l)\mathbf{d}(f_l, \theta_i) - HRTF(f_l, \theta_i)\right|^2, \tag{3}$$

with the frequency vector $\Omega = f_1 \ldots f_c \ldots f_L$ and the stacked $N \cdot L \times 1$ filter vector $\mathbf{w}_v(\Omega) = [\mathbf{w}(f_1) \ldots \mathbf{w}(f_c) \ldots \mathbf{w}(f_L)]^T$.

In order to get filter coefficients that are robust against small disturbances (in microphone position and sensitivities), some sort of regularization must be used. A common robustness measure is the so-called white noise gain (WNG). Contrary to its common use in the beamforming literature, it could be shown that a *mean* white noise gain averaged over all directions is necessary for a proper performance of the VAH [8].

The constrained optimization problem for the $t$-th frequency band can then be written as

$$\min_{\mathbf{w}_v(\Omega^t)} J_v(\mathbf{w}_v(\Omega^t)) \text{ subject to } WNG_v(\mathbf{w}_v(\Omega^t)) \geq \beta_v, \tag{4}$$

with $\beta_v$ the minimum desired value for $WNG_v$. This constrained optimization problem from Eq. (4) was used in the present study to calculate regularized individual filter coefficients. The equation for the filter coefficients minimizing this constrained optimization problem can be found in reference [8], Eq. (25).

[1] In the following $\mathbf{x}^T$ denotes the transpose and $\mathbf{x}^H$ denotes the Hermitian transpose of $\mathbf{x}$.

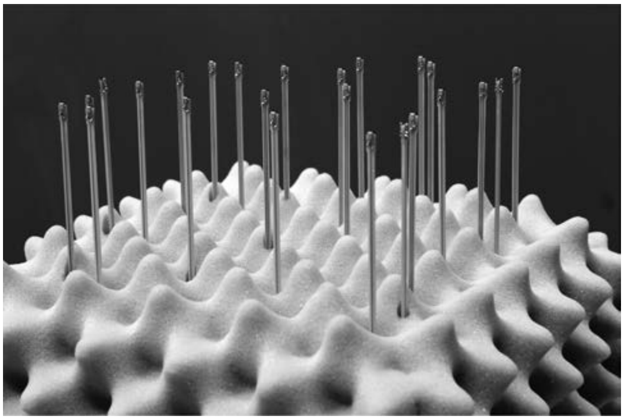

Fig. 2. **Implemented virtual artificial head.** Microphone array with 24 sensors composed of 48 MEMS microphones with a planar microphone topology according to the procedure described in [6].

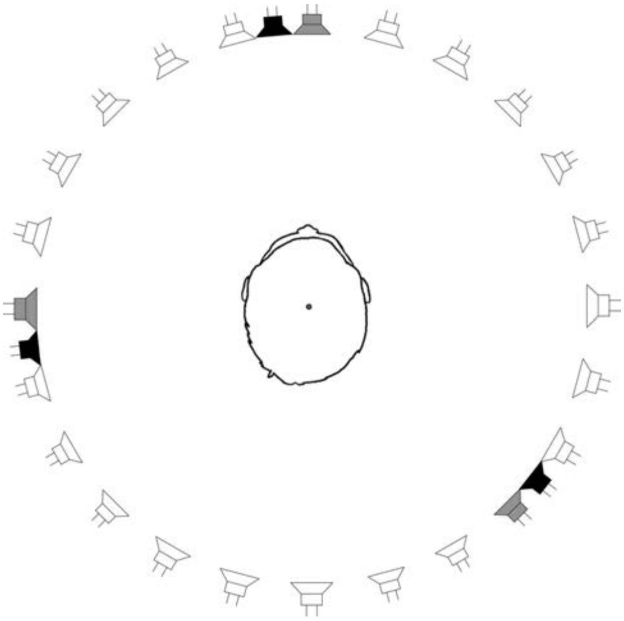

Fig. 3. **Directions considered in the optimization and the evaluation experiment.** Illustration of the directions that are explicitly considered in the optimization of the VAH-filter coefficients (white and gray loudspeaker symbols), the tested directions for three explicitly considered directions (gray) and three intermediate directions (black symbols).

In [8] it is suggested to calculate the filter coefficients in equivalent rectangular bandwidths (ERB, cf., [16]) with L frequency bins. At each frequency of interest, the ERB with frequencies $\Omega$ around its center frequency $f_c$ is considered. This yields a $N \cdot L \times 1$ vector with filter coefficients $\mathbf{w}_v(\Omega)$, where we only consider the filter coefficients at the center frequency $\mathbf{w}(f_c)$ as the solution at that frequency bin. This procedure can be interpreted as a narrowband optimization in the frequency domain that takes neighboring frequencies in ERBs into account for the optimization and regularization of the individual filter coefficients.

### 1.2 Implementation of the Virtual Artificial Head

In order to appropriately synthesize individual HRTFs using the virtual artificial head, several aspects need to be considered.

First, a microphone array with an appropriate microphone topology is required, which is adequate to synthesize various multi-directional desired directivity patterns, namely various individual HRTFs. Based on the findings from [13], in this study a planar microphone array with 24 sensors, each composed of 2 MEMS microphones (Analog Devices ADMP 504 Ultralow Noise Microphone), and a microphone topology according to [6] was used for recording and subsequently for synthesizing the individual binaural reproductions in the horizontal plane, cf., Fig. 2.

Second, the inherent direction-dependent steering vectors of the used microphone array must be known for optimizing the individual VAH-filter coefficients. In this study the steering vectors were measured in an anechoic room in the horizontal plane for 24 equidistantly spaced ($\Delta\theta = 15^\circ$) directions, with $\theta = 0^\circ, 15^\circ, 30^\circ \ldots 345^\circ$. The steering vectors were truncated in the time domain to 320 samples ($\approx$7 ms at a sampling frequency of $f_s = 44100$ Hz) using a tapered Hann-window (window length of 50 samples). Please consider that the steering vectors include the characteristics of the microphone transfer functions and hence need to be measured individually to yield an appropriate equalization associated with the VAH-synthesis.

Third, the desired directivity patterns for the VAH to be synthesized are required. In this study the individual HRTFs of the participating subjects served as the desired directivity patterns. Individual HRTFs were measured for the same 24 equidistantly spaced directions (cf., white and gray symbols in Fig. 3) as the steering vectors in the horizontal plane ($\theta = 0^\circ, 15^\circ, 30^\circ \ldots 345^\circ$) using the blocked ear method (cf., [17]) with the microphones (Knowles FG-23329 miniature electret microphones) embedded in foam earplugs (cf., [18]). All HRIR (head-related impulse responses) were truncated in the time domain to 320 samples ($\approx$7 ms at a sampling frequency of $f_s = 44100$ Hz) with a tapered Hann-window with a descending flank of 50 samples (cf., [14]) following a conservative interpretation of the perceptual limits described in [14]. Based on the findings from [15], the individual HRTFs were smoothed in the frequency- and spatial domains. The phase responses of individual HRTFs were substituted by linear phases (determined by the maximum of the hilbert envelope of the impulse responses) for frequencies $f > 1$ kHz, which enabled the complex smoothing of the HRTFs within constant relative bandwidths of $B_W = \frac{1}{5}$ octaves in the frequency domain. Furthermore, spatial notches of individual directivity patterns (HRTFs for a fixed frequency as a function of direction) were levelled out such that the dynamic range of the directivity patterns across azimuth never exceeded 29 dB at any frequency.

Finally, the measured steering vectors, the individual HRTFs, and an adequate value for the minimum desired white noise gain $\beta_v$ may be used to optimize the individual VAH-filter coefficients according to Sec. 1.1. In this study

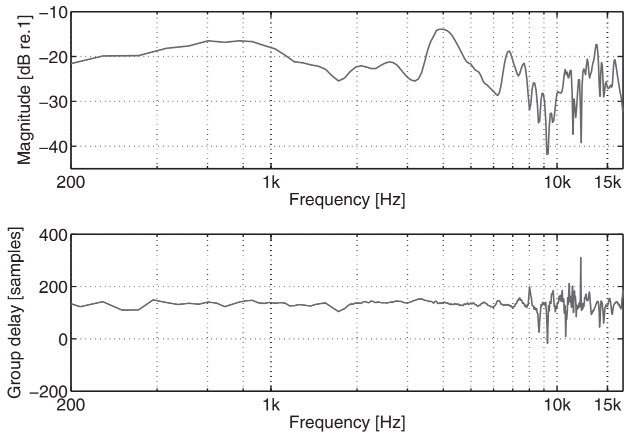


Fig. 4. **Filter coefficients for the virtual artificial head.** Exemplary VAH-filter coefficients of the VAH depicted as magnitude in dB (top) and group delay (bottom) in samples as functions of frequency.

a $\beta_v = 2$ dB was chosen. This particular value for $\beta_v$ was chosen based on previous preliminary tests. In order to give a first impression of the VAH-filter coefficients, exemplary VAH-filter coefficients are shown in Fig. 4 as a function of frequency.

It is worth noting that in this study the filter coefficients were optimized using the measured steering vectors (and HRTFs) for 24 equidistantly spaced directions only. *No* steering vectors were measured for intermediate directions (e.g., the tested intermediate directions $\theta = 7.5°$, 97.5°, and 232.5°). Instead, it is assumed that the directivity patterns for the intermediate directions are interpolated implicitly by the VAH. To test this hypothesis, the test signal (cf., Sec. 2.5) was recorded with the VAH in an anechoic room and hence synthesized for three explicitly considered loudspeaker directions ($\theta = 0°, 90°$, and 225°, cf., gray loudspeaker symbols in Fig. 3) and for three intermediate loudspeaker directions ($\theta = 7.5°, 97.5°$ & 232.5°, cf., black loudspeaker symbols in Fig. 3). Note that the latter three directions exhibit the largest possible angular deviation of 7.5° from the explicitly considered directions within the horizontal plane and hence are assumed to yield the most salient (possibly negative) perceptual effects due to the interpolation of the VAH.

## 2 METHODS

### 2.1 Experimental Design

In addition to the VAH, we considered individually measured HRTFs and two traditional artificial heads as further setups to be tested. Hence, the binaural headphone reproductions of four[2] setups were evaluated in the experiment

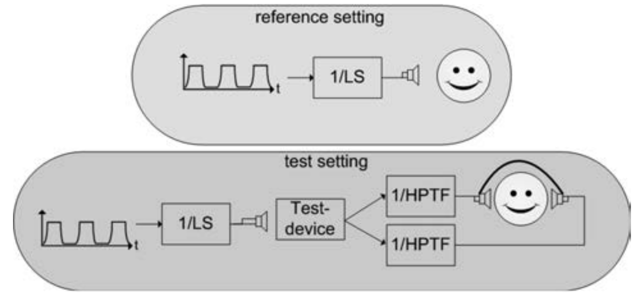


Fig. 5. **Test paradigm.** Paradigm with the test signal associated with one of the four setups (Test device) played back via headphone in the test setting and via spectrally-equalized loudspeakers (1/LS) in the reference setting.

with reference to the free-field playback in an anechoic room. The setups to be evaluated were

HRTF – Individual binaural reproduction, where the test signal was filtered with individual HRTFs and equalized using individual HPTFs.
VAH – Individualized synthesis using the VAH including individually equalized HPTFs.
$DH_1$ – Binaural reproduction using the artificial head $DH_1$ and equalized HPTFs measured with $DH_1$
$DH_2$ – Binaural reproduction using the artificial head $DH_2$ and equalized HPTFs measured with $DH_2$

These four setups (test setting, cf., Fig. 5), presented over headphones, were evaluated by the subjects with reference to a presentation via loudspeakers (reference setting, cf., Fig. 5). The used loudspeakers were custom-built in order to be sufficiently small and efficient for such evaluation experiments. They were composed of Omnes Audio BB3.AL broadband drivers and Kemo M033N 18W power amplifiers. To eliminate its influence, the transfer functions of each loudspeaker were equalized in magnitude and phase before the implementation of the experiment. The subjects could freely switch between the test setting (presented in a hidden, randomized order) and the reference setting by pushing a toggle-button attached to the headphones (cf., white arrow in Fig. 6). Within each condition, the test signal was initially played via headphones (test setting) until the subject pushed the toggle-button and put off the headphones for listening to the reference setting over loudspeakers. The signals were played in an infinite loop and subjects were able to pause playback using the stop button on the graphical user interface (GUI, cf., Fig. 7). As proposed in [25], it was possible to sort the setups according to the entered ratings and hence to further compare the setups among each other by pushing the sort button.

A total of six normal hearing subjects (five male, one female) participated in the evaluation experiment. Five of the subjects were members of the scientific staff of the Institut für Hörtechnik und Audiologie. Three of them are among the authors of this study. One subject was a student who was paid for the participation. All subjects had extensive experience with this kind of psychoacoustical evaluation

[2] Originally, the evaluation was conducted including the evaluation of *three* traditional artificial heads, rather than two, whereas one of the used artificial heads (in hindsight) turned out to have damaged microphone capsules. Hence, the ratings for this third traditional artificial head were removed from the further discussion.

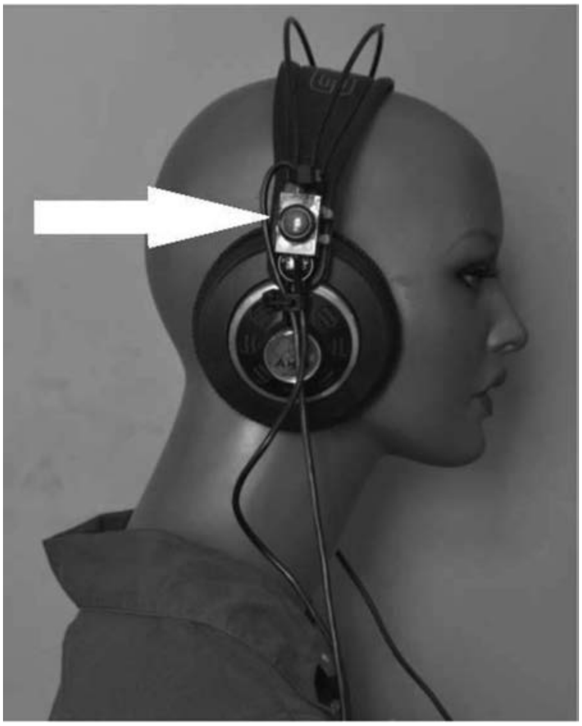

Fig. 6. **Headphones.** Headphones seated on the $DH_2$ with a toggle button (cf., white arrow) that allows to switch between the test and the reference setting.

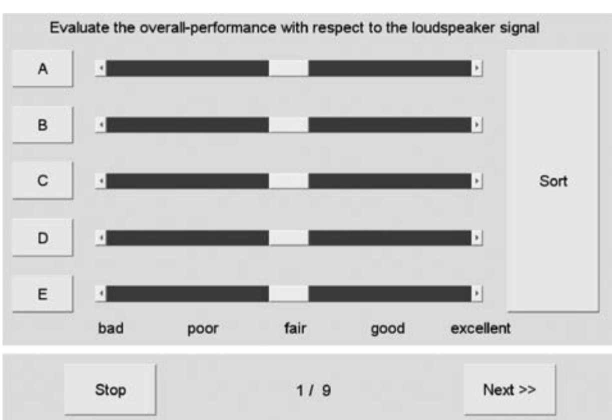


Fig. 7. **Graphical user interface**. GUI for evaluating the different setups with respect to loudspeaker presentation for each direction and perceptual quality.

experiment and participated voluntarily. As familiarization, all subjects completed at least one run of the evaluation experiment. The subjects were instructed to evaluate the perceptual qualities of the test settings regarding *localization*, *spectral coloration*, and *overall performance* in three separate (subsequent) experiments with respect to the reference setting (test signal played via loudspeakers) on an English category scale ranging between *bad*, *poor*, *fair*, *good*, and *excellent* (cf., Fig. 7). Each condition (unique combination of direction and tested setup) was evaluated three times in a randomized order. One run of a test lasted approximately 30 minutes. All subjects were encouraged to imagine the position of the sound source outside of the head (also when listening to the test signals over headphones during the familiarization phase), which was assumed to enhance the immersive reconstruction of a spatial scenario.

Subjects gave their written informed consent to participating in the study. All experimental procedures were approved by the Carl-von-Ossietzky-University of Oldenburg ethics committee.

### 2.2 Head-Related Transfer Functions (HRTFs)

In addition to the HRTFs for the 24 equidistantly spaced directions that were used for computing the filter coefficients of the VAH, HRTFs were also measured for three intermediate directions in the horizontal plane ($\theta = 7.5°$, $97.5°$, and $232.5°$, cf., black symbols in Fig. 3). Analogously to the previous procedure, the HRIR for the three intermediate directions were truncated in the time domain (320 samples) and smoothed in the frequency domain ($B_W = \frac{1}{5}$ octave bands) after substituting the measured phases with a linear phase for frequencies $f > 1$ kHz.

The individual HRTFs from the six directions were used to generate individual binaural reproductions by filtering the test signal (cf., Sec. 2.5) with the associated HRTFs and subsequent headphone equalization (cf., Sec. 2.4).

### 2.3 Traditional Artificial Heads

Two traditional artificial heads (referred to as $DH_1$, $DH_2$) were chosen to be evaluated in this experiment in addition to the individualized VAH-synthesis and individual binaural HRTF reproductions. To this end, the test signal arriving from loudspeakers positioned at the six chosen directions in the horizontal plane ($\theta = 0°$, $7.5°$, $90°$, $97.5°$, $225°$, and $232.5°$, cf., black and gray symbols in Fig. 3) was recorded with the artificial heads at the same position in the anechoic room that was also used for the steering vector- and HRTF measurements. Artificial head $DH_1$ was the commercially available HMSII.2 by Head Acoustics, whereas artificial head $DH_2$ was a custom-made device made from a mannequin (cf., Fig. 6) with built-in microphones into faithful copies of human ears. In general, the two chosen traditional artificial heads are based on two different design principles: The microphones of $DH_1$ (HMSII.2 by Head Acoustics) are placed at the blocked entrance of the ear canal, which is a similar position compared to the measurement of HRTFs using the blocked ear method. Further, $DH_1$ exhibits rather schematically designed pinnae without mimicking distinctive details of the outer ear. On the contrary, the microphones of the artificial head $DH_2$ are placed at the ends of approximately 30 mm long ear canals. Moreover, $DH_2$ exhibits rather detailed replicas of the outer ear. The positioning of the microphones considerably influences the spectral characteristics of the sound arriving at the microphones. In particular, a microphone position at the end of the ear canal will yield an additional notch in the headphone transfer function (HPTF), which in turn may become problematic in the HPTF equalization, cf., Sec. 2.4.

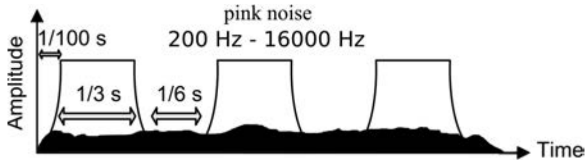


Fig. 8. **Test stimulus**. Temporal properties of the noise bursts with an additional stationary white noise of 40 dB SPL.

## 2.4 Headphone Transfer Functions (HPTFs)

The binaural test signals were presented via a D/A-converter (ADI-8 DS, RME Audio) and headphones (K-240 Studio, AKG Acoustics) to the subjects. Individual HPTFs were measured for each subject immediately after measuring the HRTFs with the microphones in the blocked-ear canal still left in place. It is well known from literature that the particular HPTF varies considerably with the individual placement/fit of the headphone (cf., [21], [22]) and may lead to audible artifacts (cf., [23]), when equalizing HPTFs that, for instance, exhibit narrowband spectral notches. Therefore, subjects were instructed to reposition the headphone 10 times to various realistic positions, prior to the 10 measurements of the HPTFs, which successively yielded 10 different, yet realistic individual HPTFs. That individual HPTF resulting in the smallest dynamic range for frequencies 300 Hz $\leq f \leq$ 16000 Hz was inverted according to the method given in [24] with the regularization parameter of $\alpha_{\text{inv}} = 0.01$. The regularized inversion was carried out after adjusting the root mean square-level of the individual HRIR in the time domain to –30 dB re 1.

The described procedure was likewise applied to measure and invert the HPTFs on subjects and on traditional artificial heads using the built-in microphones. It is worth noting that due to the differing acoustical paths between the headphone and the microphones of the artificial heads, the HPTFs and consequently their equalization vary considerably with the used artificial heads. Therefore, the described method to achieve robust equalization filters for the HPTFs (tenfold repetition and regularized inversion according to [24]) was also chosen for the two artificial heads to ensure that the quality of the various equalization filters was as constant as possible across all the different devices/setups.

## 2.5 Procedure and Stimuli

Burst sequences of pink noise with a spectral content of 200 Hz $\leq f \leq$ 16000 Hz were chosen as the test signal for the evaluation experiment. Each noise burst lasted $\frac{1}{3}$ s (333, $\overline{3}$ ms) with $\frac{1}{100}$ s (10 ms) onset- and offset ramps (raised cosine) followed by silence of $\frac{1}{6}$ s (166, $\overline{6}$ ms cf., Fig. 8). This test stimulus was intended to facilitate the evaluation of spectral but also temporal aspects. The test signal was recorded with the VAH and with the traditional artificial heads in an anechoic room using spectrally-equalized loudspeakers placed in the six chosen directions in the horizontal plane. The binaural reproductions and VAH-synthesis presented via headphone were calibrated individually to have an overall level of 75 dB SPL measured in the artificial ear coupler (G.R.A.S. type 43AA) for the ipsilateral side at $\theta = 90°$.

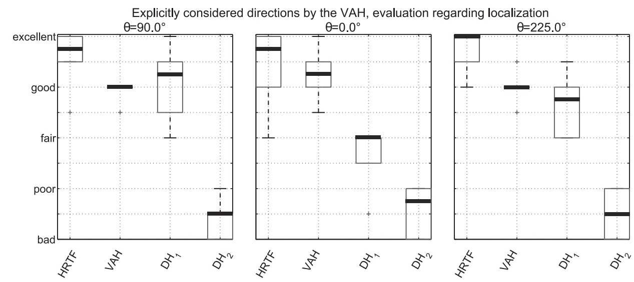


Fig. 9. **Explicitly considered directions, evaluation regarding localization**. Aggregated ratings (y-axis) for all six subjects regarding the perception of *localization* for the four tested setups on the x-axis and three explicitly considered directions (left, middle and right panel).

During preliminary tests, it became evident that the different test setups exhibit characteristically different sensor noise floors (in level and in spectral coloration), which could potentially be recognized by the subjects in the evaluation of the various setups. The sensor noise was primarily audible with the VAH. Such sensor noise is known to occur when using microphone arrays with a large number of microphones and when synthesizing distinctive desired directivity patterns [13]. To avoid an undesired detection of the setups due to the associated sensor noise, an additional stationary white noise signal of 40 dB SPL was added diotically to the noise bursts prior to the presentation (cf., dark area in Fig. 8). Because of the relatively low level of the additional white noise, it is reasonable to assume no relevant interference between the two noise signals.

## 3 RESULTS

The results of the performed experiments are presented as boxplots (cf., Fig. 9–14), where the central mark (black horizontal line) is the median and the edges of the box are the 25th and 75th percentiles. Outlying data points are plotted individually as + symbols.

The following description is divided into the results for explicitly considered directions (cf., Sec. 3.1) and intermediate directions (cf., Sec. 3.2) of the VAH.

## 3.1 Explicitly Considered Directions

The aggregated results over all subjects with regard to the perceived *localization* are illustrated in Fig. 9. Consistent with the expectations, the best ratings were obtained for the HRTF-setup followed by the individualized synthesis using the VAH. The median of the ratings of $DH_1$ is slightly better than the median of the ratings for the VAH-setup at $\theta = 90°$. However, the variance of the ratings across subjects is generally slightly lower for the individualized setups (HRTF and VAH, except for $\theta = 0°$), while the ratings

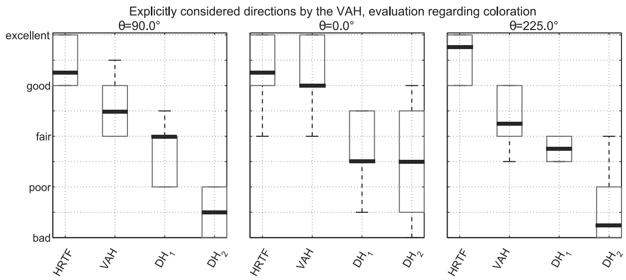


Fig. 10. **Explicitly considered directions, evaluation regarding spectral coloration**. Aggregated ratings (y-axis) for all six subjects regarding the perception of *coloration* for the four tested setups on the x-axis and three explicitly considered directions (left, middle and right panel).

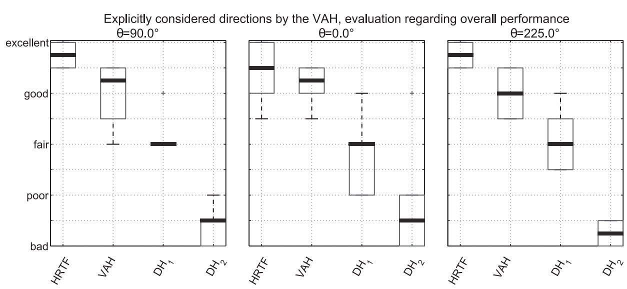


Fig. 11. **Explicitly considered directions, evaluation regarding overall performance**. Aggregated results for all six subjects regarding the *overall performance* for the four tested setups and three explicitly considered directions (left, middle and right panel).

vary slightly more across the subjects for the traditional artificial heads. Generally, the ratings for the traditional artificial heads are at or below the HRTF- and VAH-ratings, with the best ratings for the HRTF-setup and the worst ratings for $DH_2$.

The aggregated ratings regarding *spectral coloration* are illustrated in Fig. 10. Again, for the frontal direction the best ratings were obtained for the HRTF- and the VAH-setup, and the rating of the VAH-synthesis is approximately equivalent to the HRTF-setup. For the lateral direction $\theta = 90°$, the ratings for the VAH-setup vary between *fair* and *excellent*, being slightly above the ratings for $DH_1$ and clearly above $DH_2$. The median of the ratings associated with the VAH-setup at $\theta = 225°$(approximately varying between *fair* and *good*) is slightly above the ratings for $DH_1$ and clearly above the ratings for $DH_2$.

The aggregated ratings with regard to the *overall performance* are illustrated in Fig. 11. More clearly than for the ratings regarding *localization* and *spectral coloration*, the HRTF- and VAH-setup were evaluated considerably better than traditional artificial heads. For all three explicitly considered directions, the median ratings of the VAH-setup are *good* or better, whereas the median *overall performance* ratings of the traditional artificial heads are *fair* ($DH_1$) or worse ($DH_2$).

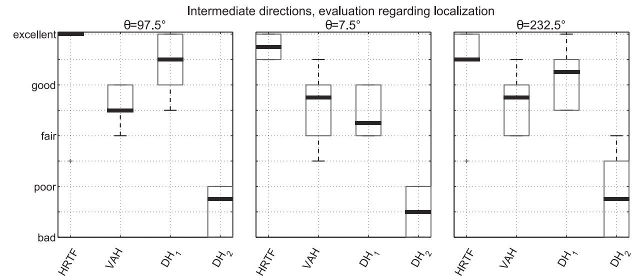


Fig. 12. **Intermediate directions, evaluation regarding localization**. Aggregated ratings (y-axis) for all six subjects regarding the perception of *localization* for the four tested setups on the x-axis and three intermediate directions ($\theta = 7.5°, 97.5°$ & $232.5°$).

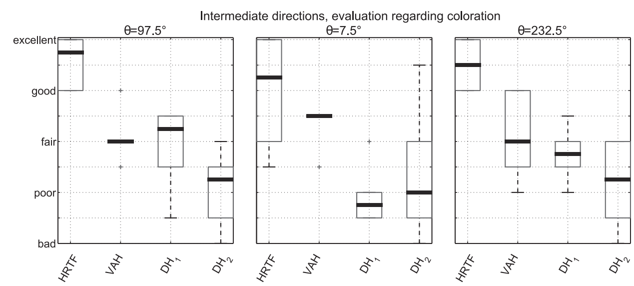


Fig. 13. **Intermediate directions, evaluation regarding spectral coloration**. Aggregated results for all six subjects regarding the perception of *spectral coloration* for the four tested setups and three intermediate directions.

## 3.2 Intermediate Directions

The aggregated ratings with regard to *localization* for the three intermediate directions are illustrated in Fig. 12. The HRTF-setup resulted in the best (median) ratings, followed by $DH_1$ ($\theta = 97.5°$ and $\theta = 232.5°$) and the VAH-setup ($\theta = 7.5°$). In comparison to the ratings for explicitly considered directions, the (median) *localization* ratings for intermediate directions for the VAH-setup drop slightly between *fair* and *good* and exhibit a slightly larger variance across the subjects. The (median) *localization* ratings associated with $DH_1$, unalteredly vary between *fair* and *excellent* (cf., Fig. 9). Also the *localization* ratings associated with $DH_2$ are similar to those for explicitly considered directions, with the median ratings ranging approximately between *bad* and *poor*.

The aggregated ratings with regard to *spectral coloration* for the intermediate directions are illustrated in Fig. 13. Again, the best median ratings were obtained for the HRTF-setup, however, with a rather large inter-subject variance for $\theta = 7.5°$. This effect is also visible for $\theta = 0°$ (cf., Fig. 10) and may presumably be explained by the difficulty of some subjects to externalize binaural reproductions for frontal directions (cf., [20]). A missing externalization, however, may possibly lead to a perceived spectral deviation and hence result in stronger inter-subject variances. The next-best spectral coloration ratings were obtained for the VAH-setup ($\theta = 7.5°$and $\theta = 232.5°$) and $DH_1$ ($\theta = 97.5°$), with the median ratings at *fair* or slightly above. The worst (median) *spectral coloration* ratings were obtained for $DH_2$

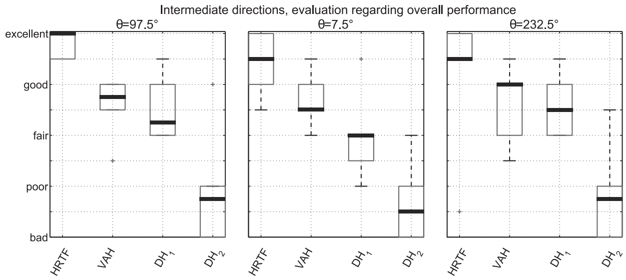


Fig. 14. **Intermediate directions, evaluation regarding overall performance**. Aggregated results for all six subjects regarding the *overall performance* for the four tested setups and three intermediate directions.

($\theta = 97.5°$and $\theta = 232.5°$) and $DH_1$ ($\theta = 7.5°$), approximately ranging around *poor*.

The aggregated ratings with regard to the *overall performance* for the intermediate directions are illustrated in Fig. 14. Similarly to the ratings for explicitly considered directions (cf., Fig. 11), the overall performance ratings obtained for the HRTF-setup are again followed by the ratings obtained for the VAH-setup, with the median ratings ranging between *fair* and *good*. The next-best median ratings were obtained for $DH_1$, also ranging between *fair* and *good*, yet constantly below the VAH-setup. The worst (median) *overall performance* ratings were obtained for $DH_2$ and are constantly below *poor*.

Interestingly, the aggregated ratings with regard to *spectral coloration* are quite comparable for both traditional artificial heads $DH_1$ and $DH_2$, whereas $DH_1$ is generally evaluated better with regard to *localization* (cf., Figs. 9 and 12) and *overall performance* (cf., Figs. 11 and 14). This may probably be explained by the greater difficulty of some subjects to externalize the binaural reproductions associated with $DH_2$.

## 4 DISCUSSION

In line with the expectations, the HRTF-setup was constantly evaluated best across the tested setups in each condition, indicating its validity and the perceptual benefits of individual binaural reproductions. The next-best ratings were obtained with the VAH-setup and the traditional artificial head $DH_1$. The characteristics of the gathered ratings will be analyzed from a statistical point of view in the next paragraphs.

In general, the experimental data for explicitly considered directions (cf., Figs. 9–11) and for intermediate directions (cf., Figs. 12–14) show a large variance across the tested setups and across subjects. For the present study it is of primary interest to investigate whether the perceptual ratings vary significantly across the tested setups. A normality test (Lilliefors-test) indicated that the gathered experimental data cannot be assumed to be normally distributed, which may primarily be attributed to the low number of subjects. This is why non-parametric statistical tests were used in the following data analysis.

The Friedman-test was used to assess the effect of the tested setups on the ratings in each condition (unique combination of direction and evaluated aspect). As often used in literature, we chose to consider a result of the Friedman-test as statistically significant if the associated p-value is smaller than 0.05. In order to control the familywise error rate (probability of at least one type I error), the critical p-value for significance is usually adapted according to the number of the implemented tests using the Bonferroni correction. Since the Friedman-test was implemented for 18 different conditions (3 evaluated aspects $\times$ 6 directions), only p-values smaller than $\frac{0.05}{18}$ are considered to indicate a significant effect of the tested setups and are indicated as bold numbers in Table 1.

The p-values in Table 1 indicate that the tested setups seem to have a significant effect on the ratings for a total of 10 conditions, which is in quite good agreement with the visual impression of the aggregated data in Figs. 9–14. In many conditions, this effect is presumably due to the mainly large difference in ratings between the HRTF-setup and $DH_2$, which is, however, not the effect of primary interest in this study. In contrast, one main aspect of interest in this study is to investigate the suitability of the VAH-setup and hence to investigate the difference in ratings between the VAH-setup and the setup associated with the best ratings when using a traditional artificial head (best of $DH_1$ or $DH_2$). In general, $DH_1$ was determined (by visual inspection of the aggregated data in Figs. 9–14) as the traditional artificial head associated with the best ratings. The Wilcoxon signed-rank test was then used as a post-hoc test within each of the 10 conditions to assess the difference in the mean ranks between the ratings for the VAH-setup and $DH_1$. Note that each of the 10 conditions (bold numbers in Table 1) was analyzed separately. No correction of the significance level (0.05) was applied. The Wilcoxon signed-rank test revealed three conditions indicating significant differences in the mean ranks between the VAH-setup and $DH_1$: The results indicate a significant difference in the mean ranks of the ratings with regard to *localization* between the VAH-setup and $DH_1$ for $\theta = 0°$and $\theta = 97.5°$($p = \frac{1}{32} = 0.03125$). Evidently from the aggregated results in Figs. 9 and 12, the ratings for the VAH-setup are significantly better compared to the ratings associated with $DH_1$ for $\theta = 0°$ and significantly worse for $\theta = 97.5°$. This fact paradigmatically emphasizes the benefits of the VAH-setup for explicitly considered directions but also indicates partial disadvantages of the VAH-setup for some intermediate directions compared to traditional artificial heads. Regarding the ratings of the *overall performance*, the Wilcoxon signed-rank test indicates significant differences in mean ranks for $\theta = 225°$ (between the VAH setup and $DH_1$, with $p = \frac{1}{32}$). Here the ratings obtained with the VAH-setup are clearly above the ratings obtained with $DH_1$, indicating significantly better ratings of the VAH-setup. No significant differences in the mean ranks could be found for the ratings regarding *spectral coloration* between the VAH-setup and $DH_1$.

In sum, the ratings emphasize the importance of individualization for binaural reproduction regarding all of the

Table 1. p-values resulting from the Friedman test. Values associated with a significant effect of the tested setups on the ratings ($p \leq \frac{0.05}{18}$) are indicated as bold numbers.

| p-values | $\theta = 0°$ | $\theta = 7.5°$ | $\theta = 90°$ | $\theta = 97.5°$ | $\theta = 225°$ | $\theta = 232.5°$ |
|---|---|---|---|---|---|---|
| *Localization* | **0.0012** | **0.0009** | 0.0037 | **0.0018** | **0.0009** | 0.0031 |
| *Coloration* | 0.0789 | 0.0098 | **0.0005** | **0.0016** | **0.0011** | 0.0087 |
| *Overall* | 0.0072 | **0.0012** | **0.0005** | 0.0042 | **0.0004** | 0.0382 |

tested aspects (*localization*, *spectral coloration*, and *overall performance*). The median ratings obtained for the individualized VAH-synthesis range mainly around *good* for explicitly considered directions and between *fair* and *good* for intermediate directions. In general, the worst results obtained for the VAH-setup were gathered for the evaluation regarding *spectral coloration*, which is also known from previous studies (cf., [26]). For some conditions (primarily for intermediate directions) the *localization* is evaluated better with the traditional artificial head $DH_1$, whereas for the majority of conditions the ratings obtained for the VAH-setup are better compared to the tested traditional artificial heads.

## 5 CONCLUSION

In conclusion, the gathered ratings emphasize the importance of individualization for binaural reproductions (see also [27] regarding localization and [28] regarding coloration) and confirm the validity of the concept to synthesize HRTFs in the horizontal plane using the proposed VAH. The results show that individual HRTFs in conjunction with individually equalized HPTFs result in the best subjective appraisals (see also [29]). The ratings obtained for the VAH-setup indicate a high level of acceptance among the subjects. Especially the VAH-synthesis for explicitly considered directions yielded mainly *good* to *excellent* median ratings. The ratings for intermediate directions seem to slightly worsen compared to the explicitly considered directions, however, still performing approximately at the same level (or slightly better regarding the *overall performance*) as binaural reproductions using the tested traditional artificial heads.

The performance of the VAH for intermediate directions may presumably be enhanced by considering a finer directional grid within the optimization (e.g., by using interpolated steering vectors and HRTFs) and/or when using a higher WNG-constraint. Moreover, it is reasonable to assume that the microphone topology of the used microphone array may have a significant effect on the directivity pattern, also for intermediate directions. In further research we will investigate the influence of the microphone topology on the performance of the interpolation for intermediate directions.

## 6 ACKNOWLEDGMENTS

This work was partially funded by Bundesministerium für Bildung und Forschung under grant no. 17080X10, the Cluster of Excellence 1077 Hearing4All of the German Research Foundation, and Akustik Technologie Göttingen.

## 7 REFERENCES

[1] S. Paul "Binaural Recording Technology: A Historical Review and Possible Future Developments," *Acta Acustica united with Acustica,* vol. 95, no. 5, pp. 767–788 (2009). DOI: https://doi.org/10.3813/AAA.918208

[2] J. Chen, B. D. V. Veen, and K. E. Hecox "External Ear Transfer Function Modeling: A Beamforming Approach," *J. Acoust. Soc. Am.,* vol. 92, no. 4, pp. 1933–1944 (1992). DOI: https://dx.doi.org/10.1121/1.405241

[3] N. Tohtuyeva and V. Mellert "Approximation of Dummy-Head Recording Technique by a Multimicrophone Arrangement," *J. Acoust. Soc. Am.,* vol. 105, no. 2, pp. 1101–1101 (1999). DOI: https://dx.doi.org/10.1121/1.425158

[4] Y. Kahana, P. A. Nelson, O. Kirkeby, and H. Hamada "A Multiple Microphone Recording Technique for the Generation of Virtual Acoustic Images," *J. Acoust. Soc. Am.,* vol. 105, no. 3, pp. 1503–1516 (1999). DOI: https://dx.doi.org/10.1121/1.426690

[5] J. Atkins "Robust Beamforming and Steering of Arbitrary Beam Patterns Using Spherical Arrays," *Applications of Signal Processing to Audio and Acoustics (WASPAA), 2011 IEEE Workshop on* (2011), pp. 237–240. DOI: https://dx.doi.org/10.1109/ASPAA.2011.6082332

[6] E. Rasumow, M. Blau, M. Hansen, S. Doclo, S. van de Par, V. Mellert, and D Paschal "Robustness of Virtual Artificial Head Topologies with Respect to Microphone Positioning Errors," *Proc. Forum Acusticum, Aalborg*. Aalborg (2011), pp. 2251–2256.

[7] E. Rasumow, M. Blau, S. Doclo, M. Hansen, S. van de Par, D. Paschal, and V. Mellert "Least Squares versus Non-Linear Cost Functions for a Virtual Artificial Head," *J. Acoust. Soc. Am.,* vol. 133, no. 5, pp. 3525–3525 (2013). DOI: https://dx.doi.org/10.1121/1.4800595

[8] E. Rasumow, M. Hansen, S. van de Par, D. Püschel, V. Mellert, S. Doclo, and M. Blau "Regularization Approaches for Synthesizing Multi-Directional Directivity Patterns with Microphone Arrays," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 24, no. 2, pp. 215–225 (2016). DOI: https://dx.doi.org/10.1109/TASLP.2015.2504874

[9] J. Atkins "Spatial Acoustic Signal Processing for Immersive Communication," Ph.D. thesis, Johns Hopkins University, Baltimore, MD (2011).

[10] D. N. Zotkin, R. Duraiswami, and N.A. Gumerov "Regularized HRTF Fitting Using Spherical Harmonics," *WASPAA* (IEEE, 2009). p. 257–260. DOI: https://dx.doi.org/10.1109/ASPAA.2009.5346521

[11] C. D. S. Castaneda, S. Sakamoto, J. A. T. Lopez, J. Li, Y. Yan, and Y. Suzuki "Accuracy of Head-Related Transfer Functions Synthesized with Spherical Microphone Arrays," *Proceedings of the 21st International Congress on Acoustics*, Montreal (Canada), vol. 19 (ASA, 2013). DOI: https://dx.doi.org/10.1121/1.4800833

[12] S. Sakamoto, S. Hongo, T. Okamoto, Y. Iwaya, and Y. Suzuki "Improvement of Accuracy of Three-Dimensional Sound Space Synthesized by Real-Time 'SENZI,' a Sound Space Information Acquisition System Using Spherical Array with Numerous Microphones," *Proceedings of the 21st International Congress on Acoustics*, Montreal (Canada) (2013). DOI: https://dx.doi.org/10.1121/1.4801090

[13] E. Rasumow, M. Blau, M. Hansen, S. Doclo, S. van de Par, V. Mellert, and D. Paschal "The Impact of the White Noise Gain (WNG) of a Virtual Artificial Head on the Appraisal of Binaural Sound Reproduction," *Proc. of the EAA Joint Symposium on Auralization and Ambisonics*, Berlin, Germany, 3–5 April 2014, p. 174–180. DOI: https://dx.doi.org/10.14279/depositonce-29

[14] E. Rasumow, M. Blau, M. Hansen, S. Doclo, S. van de Par, D. Paschal, and V. Mellert "Smoothing Head-Related Transfer Functions for a Virtual Artificial Head," *Acoustics 2012*. Nantes, France (2012), p. 1019–1024.

[15] E. Rasumow, M. Blau, M. Hansen, S. van de Par, S. Doclo, V. Mellert, and D. Paschal "Smoothing Individual Head-Related-Transfer Functions in the Frequency and Spatial Domains," *J. Acoust. Soc. Am.,* vol. 135, no. 4, pp. 2012–2025 (2014 Apr.). DOI: https://dx.doi.org/10.1121/1.4867372

[16] B. C. J. Moore and B. R. Glasberg "Suggested Formulae for Calculating Auditory Filter Bandwidths and Excitation Patterns," *J. Acoust. Soc. Am.,* vol. 74, no. 3, pp. 750–753 (1983). DOI: https://dx.doi.org/10.1121/1.389861

[17] D. Hammershøi and H. Møller "Sound Transmission To and Within the Human Ear Canal," *J. Acoust. Soc. Am.,* vol. 100, no. 1, pp. 408–427 (1996). DOI: https://dx.doi.org/10.1121/1.415856

[18] S. Raufer, E. Rasumow, and M. Blau "HRTF- Measurements with Earmolds and Conventional Ear Plugs—A Comparison," *Proceedings of AIA-DAGA 2013*, Merano (Italy) (2013), pp. 1320–1321.

[19] F. L. Wightman and D. J. Kistler "Headphone Simulation of Free-Field Listening. II: Psychophysical Validation," *J. Acoust. Soc. Am.,* vol. 85, no. 2, pp. 868–878 (1989). DOI: https://dx.doi.org/10.1121/1.397558

[20] S. M. Kim and W. Choi "On the Externalization of Virtual Sound Images in Headphone Reproduction: A Wiener Filter Approach," *J. Acoust. Soc. Am.,*vol. 117, no. 6, pp. 3657–3665 (2005). DOI: https://dx.doi.org/10.1121/1.1921548

[21] B. S. Masiero, *Individualized Binaural Technology: Measurement, Equalization and Subjective Evaluation* (Institute of Technical Acoustics RWTH Aachen University, Logos Verlag, 2012).

[22] F Volk "Inter- and Intra-Individual Variability in the Blocked Auditory Canal Transfer Functions of Three Circus-Aural Headphones," *J. Audio Eng. Soc.,*vol. 62, pp. 315–323 (2014 May). DOI: https://doi.org/10.17743/jaes.2014.0021

[23] M. Paquier, V. Kohl, and B. Jantzen "Effects of Headphone Transfer Function Scattering on Sound Perception," *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA), 2011* (2011), pp. 181–184. DOI: https://doi.org/10.1109/ASPAA.2011.6082317

[24] O. Kirkeby and P. A. Nelson "Digital Filter Design for Inversion Problems in Sound Reproduction," *J. Audio Eng. Soc.,* vol. 47, pp. 583–595 (1999 Jul./Aug.).

[25] P. Y. Michaud, S. Meunier, P. Herzog, M. Lavender, and G. D. d'Aubigny "Perceptual Evaluation of Dissimilarity between Auditory Stimuli: An Alternative to the Paired Comparison," *Acta Acustica united with Acustica*, vol. 99, no. 5, pp. 806–815 (2013). DOI: https://doi.org/10.3813/AAA.918658

[26] E. Rasumow, M. Blau, S. Doclo, M. Hansen, S. van de Par, D. Paschal and V. Mellert "Individualized Binaural Reproduction Using a Virtual Artificial Head," *Fortschritte der Akustik - DAGA 2014*. Oldenburg, Germany (2014), pp. 26–27.

[27] H. Møller, M. F. Sorensen, C. B. Jensen and D. Hammershøi "Binaural Technique: Do We Need Individual Recordings?" *J. Audio Eng. Soc.,*vol. 44, pp. 451–469 (1996 June).

[28] F. Völk and H. Fastl "Locating the Missing 6 dB by Loudness Calibration of Binaural Synthesis," presented at the *131st Convention of the Audio Engineering Society* (2011 Oct.), convention paper 8488.

[29] F. Volk "Interrelations of Virtual Acoustics and Hearing Research by the Example of Binaural Synthesis," Ph.D. thesis, TU Munich (2013).

## THE AUTHORS

Eugen Rasumow

Matthias Blau

Simon Doclo

Steven van de Par

Martin Hansen

Dirk Püschel

Volker Mellert

Eugen Rasumow received the Dipl.-Ing. (FH) degree in hearing technology and audiology from the Oldenburg University of Applied Science, Germany, in 2007 and the M.Sc. degree in hearing technology and audiology from the University of Oldenburg, Germany, in 2009. In 2015 he received the Ph.D. degree from the University of Oldenburg, Germany. His research interests include psychoacoustics, virtual acoustics, and binaural signal processing.

•

Matthias Blau studied electrical engineering at TU Dresden (Germany) and Purdue University (USA). In 1999 he received a Ph.D. in electrical engineering from TU Dresden. Since 2003 he is Professor of Electroacoustics at Jade Hochschule in Oldenburg, Germany. His current research interests include the acoustics of outer and middle ears, virtual acoustics, implantable transducers, room acoustics, and acoustical measurements. Dr. Blau is member of the German Acoustical Society, the Acoustical Society of America, and the Marie Curie Fellowship Association.

•

Simon Doclo (S'95-M'03-SM'13) received the M.Sc. degree in electrical engineering and the Ph.D. degree in applied sciences from the Katholieke Universiteit Leuven, Belgium, in 1997 and 2003. From 2003 to 2007 he was a Postdoctoral Fellow with the Research Foundation Flanders at the Electrical Engineering Department (Katholieke Universiteit Leuven) and the Adaptive Systems Laboratory (McMaster University, Canada). From 2007 to 2009 he was a Principal Scientist with NXP Semiconductors at the Sound and Acoustics Group in Leuven, Belgium. Since 2009 he is a full professor at the University of Oldenburg, Germany, and scientific advisor for the project group Hearing, Speech and Audio Technology of the Fraunhofer Institute for Digital Media Technology. His research activities center around signal processing for acoustical applications, more specifically microphone array processing, active noise control, acoustic sensor networks, and hearing aid processing. Prof. Doclo received the Master Thesis Award of the Royal Flemish Society of Engineers in 1997 (with Erik De Clippel), the Best Student Paper Award at the International Workshop on Acoustic Echo and Noise Control in 2001, the EURASIP Signal Processing Best Paper Award in 2003 (with Marc Moonen), and the IEEE Signal Processing Society 2008 Best Paper Award (with Jingdong Chen, Jacob Benesty, Arden Huang).

He was member of the IEEE Signal Processing Society Technical Committee on Audio and Acoustic Signal Processing (2008–2013) and Technical Program Chair for the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA) in 2013. Prof. Doclo has served as guest editor for several special issues (*IEEE Signal Processing Magazine, Elsevier Signal Processing*) and is associate editor for *IEEE/ACM Transactions on Audio, Speech and Language Processing* and *EURASIP Journal on Advances in Signal Processing*.

•

Steven van de Par studied physics at the Eindhoven University of Technology, Eindhoven, The Netherlands, and received the Ph.D. degree from the Eindhoven University of Technology, in 1998, on a topic related to binaural hearing. As a Postdoctoral Researcher at the Eindhoven University of Technology, he studied auditory visual interaction and was a Guest Researcher at the University of Connecticut Health Center. In early 2000, he joined Philips Research, Eindhoven, to do applied research in digital signal processing and acoustics. His main fields of expertise are auditory and multisensory perception, low-bit-rate audio coding, and music information retrieval. He has published various papers on binaural auditory perception, auditory visual synchrony perception, audio coding, and music information retrieval (MIR)-related topics. Since April 2010 he has held a professor position in acoustics at the University of Oldenburg, Oldenburg, Germany.

•

Martin Hansen studied physics in Göttingen and did his Ph.D. in Oldenburg about psychoacoustical modeling of low bit-rate coded speech transmission quality. After working in R&D in a hearing aid company in Copenhagen for several years he became professor for medical acoustics at the Jade University for Applied Sciences in Oldenburg. His research interests include psychoacoustics and technical audiology.

•

Dirk Püschel was born in 1960. He studied physics from 1980–1984 and received his Ph.D. degree in 1988. His research was in the field of psychoacoustic and room acoustics at III. Institute of Physics in Goettingen until 1993. He was the founder of Akustikbüro Goettingen in 1992 and Akustik Technologie Goettingen in 1993. He is the owner, Director of Research, and Managing Director of Soundtec GmbH.

•

Volker Mellert was head of the Acoustics Group at the Institute for Physics of Oldenburg University and retired in 2009. He received his Ph.D. in physics in Göttingen 1972 and became a Professor for Applied Physics in Oldenburg in 1976. Acoustic wave propagation and human perception (psychoacoustics) are his main scientific interests. He was President of the German Association for Acoustics (DEGA) and of the European Association for Acoustics (EAA). He served as board member of the International Commission of Acoustics and as Editor-in-Chief for *ACUSTICA*. He received the German Medal of Merit for his contributions in acoustics and the Helmholtz-Medal of the DEGA. He is Fellow of the Acoustical Society of America.