How wearing headgear affects measured head-related transfer functions

Christoph Pörschmann, Johannes M. Arend, Raphael Gillioz

HOW WEARING HEADGEAR AFFECTS MEASURED HEAD-RELATED TRANSFER FUNCTIONS

Christoph Pörschmann[1] **Johannes M. Arend**[1,2] **Raphael Gillioz**[3]
[1] TH Köln, Institute of Communications Engineering, Cologne, Germany
[2] TU Berlin, Audio Communication Group, Berlin, Germany
[3] TH Köln, Institute of Media and Imaging Technology, Cologne, Germany
Christoph.Poerschmann@th-koeln.de

ABSTRACT

We present spherical high-density measurement data of head-related transfer functions (HRTFs) and analyze the influence of wearing headgear during the measurements. For this we captured datasets from a Neumann KU100 and a HEAD acoustics HMS II.3 dummy head either equipped with a bicycle helmet, a baseball cap, an Oculus Rift head-mounted display, or AKG K1000 headphones. We investigate the influence of different types of headgear in terms of their spectrum and their binaural cues and compare the results to reference measurements of the dummy heads without headgear. Generally, the results show that differences to the reference vary significantly depending on the type of the headgear. The spectral differences to the reference are maximal for the AKG K1000 and smallest for the Oculus Rift and the baseball cap. Analyzing the influence of the incidence directions on the spectral differences we found the strongest deviations for the Oculus Rift and the baseball cap for contralateral sound incidence. For the bicycle helmet, the contralateral directions were also most affected, but shifted upwards in elevation. Finally, for the AKG K1000, which generally has the highest impact on the spectrum of the HRTFs, we observed maximal deviations for sound incidence from behind. Regarding the interaural time differences (ITDs) and interaural level differences (ILDs) the analysis again revealed the highest influences for the AKG K1000. While for the Oculus Rift the ITDs and ILDs were mainly affected for frontal directions, we observed only a very weak influence of the bicycle helmet and the baseball cap. The HRTF sets are available in the SOFA format under a Creative Commons CC BY-SA 4.0 license.

1. INTRODUCTION

The spatial representation of sound sources is an essential element of virtual acoustic environments (VAEs). When determining the sound incidence direction, the human auditory system evaluates monaural and binaural cues, which are caused by the shape of the pinna and the head. While spectral information is the most important cue for elevation of a sound source, we use differences between the signals reaching the left and the right ear for lateral localization. These binaural differences manifest in interaural time differences (ITDs) and interaural level differences (ILDs). In many headphone-based VAEs, head-related transfer functions (HRTFs) are used to describe the sound incidence from a source to the left and right ear, thus integrating both monaural and binaural cues [1]. Various researchers investigated the perceptual influences of individual measurements of HRTFs (e.g. [2, 3]) and found improvements regarding externalization, localization as well as a reduction of front-back confusions when using individualized datasets. Furthermore, specific properties, like for example the torso [4], and probably even headgear [5–7] influence the HRTFs and thus as well localization and other perceptual attributes.

Generally speaking, apart from individualization and head-above-torso movements, in many real-life situations spatial cues are modified by headgear, for example by wearing a baseball cap, a bicycle helmet, or a head-mounted display (HMD), which nowadays is often used in VR applications. However, often a good localization performance is important when wearing such items, e.g. in order to determine approaching vehicles when cycling. Furthermore, when performing psychoacoustic experiments in mixed-reality applications using HMDs, the influence of the HMD on the HRTFs must be considered. Effects of an HTC Vive HMD on localization performance have already been analyzed by Ahrens et al [8]. The authors performed a loudspeaker-based localization experiment presenting conditions with and without an HMD, which showed significant differences in localization due to the HMD. To analyze the influence of headgear for varying directions of incidence, measurements of HRTFs on a dense spherical sampling grid are required. However, HRTF measurements of a dummy head with various headgear are still rare, and to our knowledge only one dataset measured for an HTC Vice on a sparse grid with 64 positions is freely accessible [8].

To accurately analyze these influences for all directions of incidence, measurements of HRTFs on a dense

doi:10.25836/sasp.2019.27

(a) (b) (c) (d)

Figure 1: Measurement of the HRTF set in the anechoic chamber at TH Köln for the baseball cap on the HEAD acoustics HMS II.3 (a), the Oculus Rift on the Neumann KU100 (b), the Uvex bicycle helmet on the Neumann KU100 (c), and the AKG K1000 on the HEAD acoustics HMS II.3 (d).

grid need to be performed. Several authors suggested to describe complete sets of HRTFs in spherical harmonics (SH) domain [9, 10]. Here, the HRTF set, measured on a spherical grid, is decomposed into spherical base functions of different spatial orders N, where higher orders correspond to a higher spatial resolution. To completely consider these properties and to avoid spatial aliasing, an order $N \geq kr$ with $k = \omega/c$, and r being the head radius is required [11, 12]. Assuming $r = 8.75\,\text{cm}$ as the average human head radius [13] and $c = 343\,\text{m/s}$ leads to $N = 32$ requiring at least 1089 measured directions. In this case no spatial aliasing occurs for frequencies up to 20 kHz. However, only a few of the datasets mentioned above were measured spherically with an adequate density, none of them investigating the influence of headgear.

This work presents high-density measurement data of HRTF sets from a Neumann KU100 and a HEAD acoustics HMS II.3 dummy head, either equipped with a bicycle helmet, a baseball cap, an Oculus Rift HMD, or a set of extra-aural AKG K1000 headphones captured on a full spherical Lebedev grid with 2702 points. We analyze the datasets in terms of their spectrum and their binaural cues and compare the results to reference measurements of the dummy heads without headgear.

2. HRTF MEASUREMENTS

We performed HRTF measurements of a Neumann KU100 dummy head and a Head acoustics HMS II.3 for four different types of headgear: a baseball cap (Flexfit Snapback), a bicycle helmet (Uvex Cobra RS), an HMD (Oculus Rift, without any kind of headphones) and AKG K1000 extraaural headphones. As loudspeaker we used a Genelec 1029A, which has a flat on-axis frequency response from 50 Hz to 20 kHz (±3 dB). The HRTFs were measured on a Lebedev full spherical grid with 2702 points. We applied the VariSphear measurement system [14] for precise positioning of the dummy head at the spatial sampling positions and for capturing the HRTFs. Besides the motor control and impulse response capture modules, the software provides an automatic error detection which checks every measured impulse response for noticeable variations compared to the previous measurement. This ensures a validity of all obtained impulse responses. The excitation signal for all measurements was an emphasized sine sweep with +20 dB low shelf at 100 Hz (2^{18} samples at 48 kHz sampling rate, length 5.5 s). An RME Babyface audio interface served as AD / DA converter and microphone preamp. For further details on the measurement set-up and procedure please refer to [15, 16].

The measurements were carried out in the anechoic chamber at TH Köln. The room has dimensions of $4.5\,\text{m} \times 11.7\,\text{m} \times 2.3\,\text{m}$ and a lower cut-off frequency of about 200 Hz. Fig. 1 shows the Neumann KU100 and the Head acoustics HMS II.3 with the different types of headgear mounted on the VariSphear device. The height of the loudspeaker and of the dummy head was at 1.25 m and the acoustic center of the loudspeaker was always set to the ear level of the dummy head. All sets of HRTFs were captured at 2 m distance. For all setups, exact alignment of the head was checked for various sampling positions. The distance between the loudspeaker and the entrance of the dummy head's ear canal was for each measurement accurately determined with a laser distance meter. Additionally, we used a Microtech Gefell M296S microphone positioned at the acoustic center of the dummy head to measure omnidirectional impulse responses which we used for the magnitude and phase compensation of the loudspeaker.

In a subsequent postprocessing the raw measurement data were first carefully truncated and windowed. Then we compensated the influence of the loudspeaker by inverse FIR filtering with the measured omnidirectional impulse response. The final length of each HRIR is 128 samples at a sampling rate of 48 kHz. The postprocessing is based on the implementation and description from [15, 16]. A fur-

doi:10.25836/sasp.2019.27

ther step of the postprocessing eliminates low-frequency artefacts resulting from the frequency response of the relatively small loudspeaker, which fails to reproduce low frequencies at adequate sound pressure levels. Additionally this step removes the influence of room modes and reflections arising from the sound field of the anechoic chamber below its cut-off frequency. A well-suited approach is to replace the low-frequency range of the HRTFs by an analytic expression assuming that for low frequencies (e.g. below 200 Hz), pinna and ear canal hardly affect the HRTF and even the spherical shape of the head only has minor influence on the sound field. In this study we use a low-frequency extension in the frequency domain according to [17] and apply a linear cross-fade between the low-frequency component and the raw HRTFs in a crossover frequency range from 200 Hz – 400 Hz. The level is calculated from the mean absolute values, while the phase is linearly extrapolated in the crossover frequency range.

Finally we transformed the dataset to the spherical harmonics (SH) domain and stored it in form of SH coefficients [9, 10]. This allows for a calculation of the HRTFs at any direction by means of the inverse SH transform.

Even though the measurements were conducted with great care, there are several influencing factors which should be considered. The influence of the robot arm of the VariSphear system on sound radiation to the ear is hard to quantify and depends very much on frequency and incidence direction. Please refer to [15, 16] for a more detailed analysis of these influences. However, for our comparisons of the measured datasets to a reference without headgear these influences are quite irrelevant as they occur in the same way in all measured HRTF sets. In the context of our study, differences which are induced by small variations, e.g. by non exact placement of the loudspeaker and the dummy head are more important. We minimized these inaccuracies by exactly positioning and calibrating the device before each measurement session.

3. RESULTS

3.1 Spectrum

Fig. 2 shows the magnitude responses for frontal and contralateral sound incidence both for the reference and the different headgear. Below 1 kHz the differences are quite small because the wavelength is larger than the geometric structures of the headgear. Generally, the results show that the deviations to the reference are minor for the baseball cap and the Oculus Rift. However, for the baseball cap we observed differences of more than 10 dB at frequencies between 5 kHz and 8 kHz for contalateral sound incidence. The bicycle helmet has nearly no influence for frontal sound incidence, but contralaterally already below 3 kHz strong comb-filter effects become apparent. For all headgear tested, the differences to the reference are maximal for the AKG K1000. Even though the headphones do not directly cover the ears, they strongly influence sound incidence. Accodingly, for frontal sound incidence already at 2 kHz the magnitude response is reduced by 10 dB.

To analyze the spectral deviations to a reference set, we calculated the averaged values across all 2702 measured directions $\Omega\{(\phi_1, \theta_1), \ldots, (\phi_T, \theta_T)\}$:

$$\Delta G_f(\omega) = \frac{1}{n_\Omega} \sum_{\Omega=1}^{n_\Omega} |20lg \frac{|\mathit{HRTF}_{\mathit{REF}}(\omega, \Omega)|}{|\mathit{HRTF}_{\mathit{TEST}}(\omega, \Omega)|}|, \quad (1)$$

with ω describing the temporal frequency, $\mathit{HRTF}_{\mathit{REF}}$ the reference HRTF set and $\mathit{HRTF}_{\mathit{TEST}}$ the HRTF set of the tested headgear.

Fig. 3 illustrates the frequency-dependent spectral differences $\Delta G_f(\omega)$ between the reference without headgear and the different headgear both for the KU100 and the HMS II.3. Below 1 kHz the deviations are mostly in the range of 1 dB or below. Only for the AKG K1000 on the HMS II.3 dummy head they exceed 2 dB. Generally, we observed the lowest differences to the reference for the baseball cap and the Oculus Rift which are in the range of 2 dB for frequencies up to 10 kHz. The deviations are higher for the bicycle helmet, reaching 4 dB already at frequencies below 10 kHz. Finally the highest differences of all tested headgear occur for the AKG K1000. Already at 2 kHz the values exceed 4 dB and reach 6 dB at about 10 kHz.

In a next step we analyzed the spatial distribution of the differences and calculated the directional deviation across all frequencies as

$$\Delta G_{sp}(\Omega) = \frac{1}{n_\omega} \sum_{\omega=1}^{n_\omega} | \, 20lg \frac{| \, \mathit{HRTF}_{\mathit{REF}}(\omega, \Omega) \, |}{| \, \mathit{HRTF}_{\mathit{TEST}}(\omega, \Omega) \, |} \, |, \quad (2)$$

In this case the sampling grid Ω_t was full spherical, in steps of 1° for azimuth and elevation. The results are very similar for both tested dummy heads. Fig. 4 shows the results for the Neumann KU100. For the baseball cap (a) and for the Oculus Rift (c) the spectral differences are mainly located at contralateral directions. Here, the sound incidence which is dominated by diffraction around the the head is strongly affected. For the different types of headgear we found maximal deviations $\Delta G_{sp,max}$ in a range of $8\,dB$ to $10\,dB$. For the baseball cap the maximal deviations are located at $\phi = 272^\circ$ and $\theta = 13^\circ$ and for the Oculus Rift at $\phi = 260^\circ$ and $\theta = 10^\circ$. As the bicycle helmet mainly covers the top of the head we observed maximal spectral differences shifted upwards at $\phi = 273^\circ$ and $\theta = 59^\circ$. For the AKG K1000 the plot shows distinct spectral differences spread over the entire angular range with large areas of high spectral deviations. We observed the highest deviations at $\phi = 220^\circ$ and $\theta = -9^\circ$.

3.2 Binaural cues

Next we compared the ILDs and ITDs of the different headgear to the reference without headgear. For this purpose, we extracted HRTFs in the horizontal plane ($\theta = 0^\circ$) with an angular spacing of $\phi = 1^\circ$ from the datasets. The broadband ILDs were then calculated as the ratio between the energy of the left and right ear HRIR. The ITDs were calculated from the HRIRs by applying the threshold onset method with ten-times oversampling for more precise onset detection.

doi:10.25836/sasp.2019.27

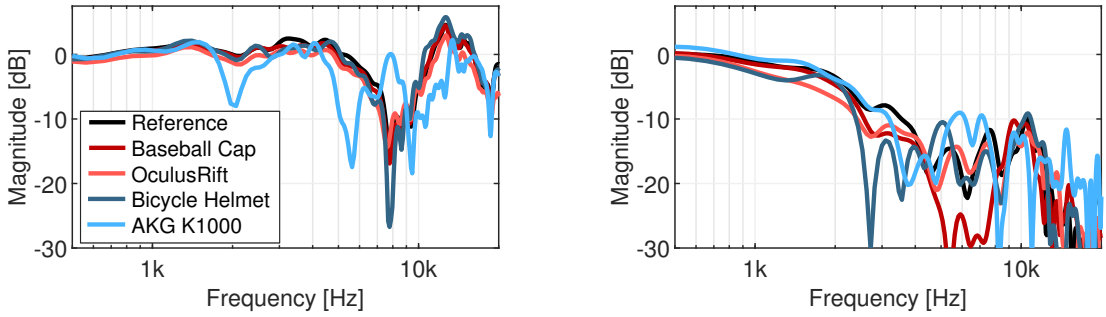

Figure 2: Left ear magnitude, extracted from the sets of the reference (black) and the different headgear on a Neumann KU100. In red the results for the baseball cap and the bicycle helmet are shown, in blue the results for the Oculus Rift and for the AKG K1000. (a) Front direction ($\phi = 0°, \theta = 0°$). (b) Contralateral direction ($\phi = 270°, \theta = 0°$).

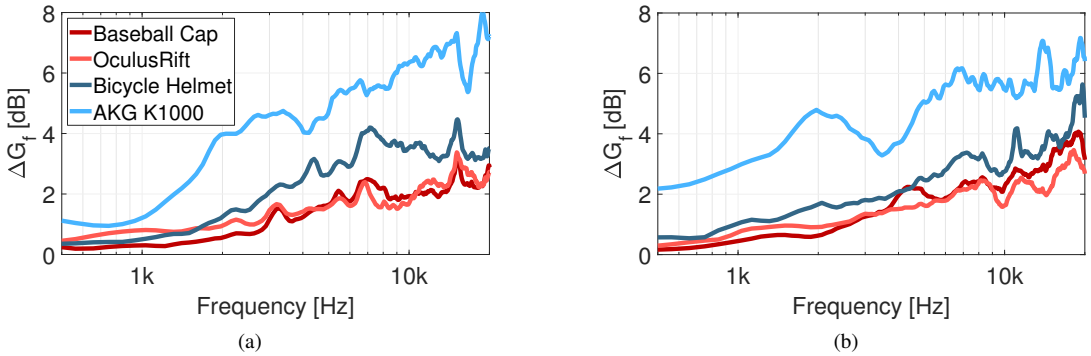

Figure 3: Spectral differences $\Delta G_f(\omega)$ in dB (left ear) between reference HRTF set and the different headgear. In red the results for the baseball cap and the bicycle helmet are shown, in blue the results for the Oculus Rift and for the AKG K1000. (a) Spectral differences for Neumann KU100, (b) for HEAD acoustics HMS II.3.

Fig. 5 shows the calculated ILDs and ITDs of the reference HRTF set and the different headgear. As depicted in Fig. 5 (a,c), for the baseball cap and the bicycle helmet the ILDs are similar to the reference without headgear. Only marginal deviations can be observed here, mainly at lateral directions. For the Oculus Rift stronger deviations occur and for the AKG K1000 the ILDs are altered by up to 4 dB. As can be seen in Fig. 5 (b,d), the ITDs of the measured HRTF sets are generally in good agreement with the ITDs of the reference. While only small deviations exist for the baseball cap, the bicycle helmet and the HMD, the ITDs are significantly reduced laterally for the AKG K1000 measured on the Neumann KU 100.

4. DISCUSSION

In the previous section we have shown that headgear significantly affects measured HRTFs. Now we discuss how this relates to some other factors influencing HRTF measurements. In this context Brinkmann et al. [4] analyzed deviations due to different head-above-torso rotations. The results showed direction-dependent spectral deviations comparable to the ones of the baseball cap or the Oculus Rift. Furthermore, the authors performed a listening experiment showing the audibility of these deviations.

In the context of HRTF measurements, spectrum, temporal structure and interaural differences are as well affected by spatial upsampling. For sparse HRTF sets, which are often used for individual measurements, a subsequent spatial upsampling needs to be performed, e.g. by applying a spatial Fourier transform of the data in the spherical harmonics domain and resampling the HRTFs by an inverse transformation on a dense sampling grid [9, 10]. However, this results in spatial aliasing and truncation errors. The influence of these so-called sparsity errors on the upsampling of HRTFs has for example been analyzed in [18, 19]. Depending on the spatial resolution of the sparse HRTF set, the contributions of the baseball cap, the Oculus Rift and even the bicycle helmet are in the same range or even lower than the sparsity errors. For example, for spatial upsampling by spherical harmonic interpolation at a spatial order of $N = 13$, resulting in 266 measurements on a Lebedev grid, the mean spectral deviations are still above 4 dB at frequencies of 10 kHz [19]. Even when performing an improved interpolation method according to Pörschmann et al. [19], the averaged spectral deviations on this grid are still in the same range as for the baseball cap or the Oculus Rift. Comparing the results of our study to investigations on spatial upsampling of sparse HRTF sets reveals, that a specific focus must be put on the spatial upsampling when

doi:10.25836/sasp.2019.27

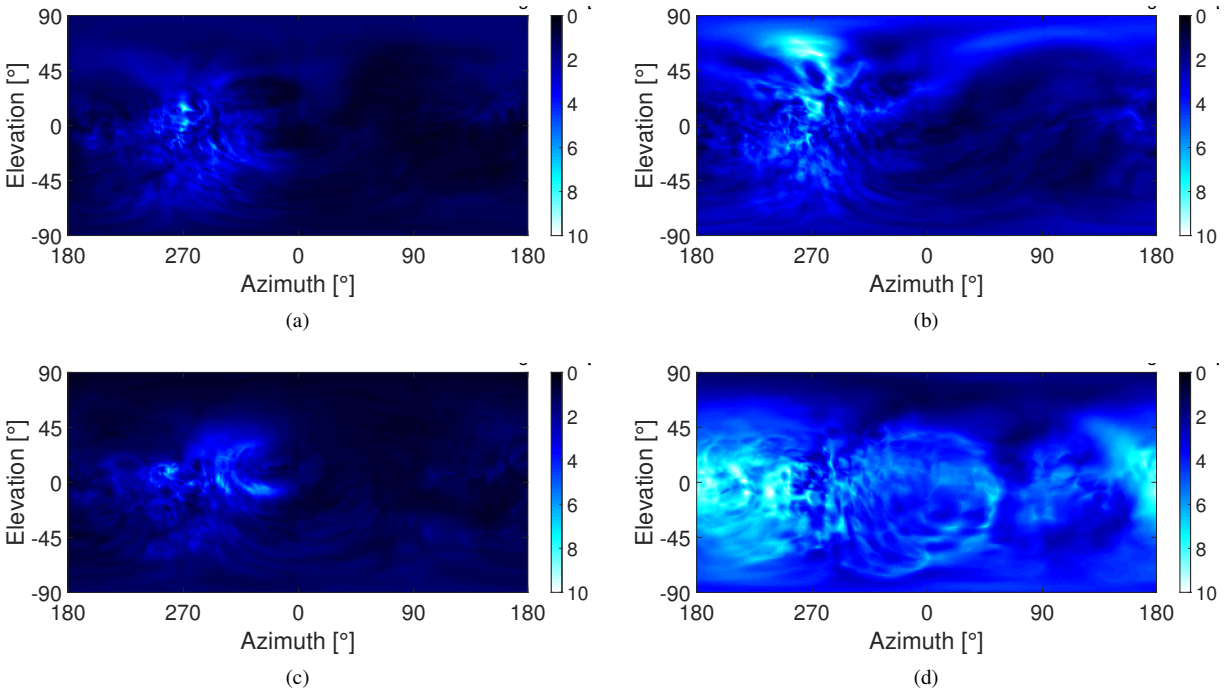

Figure 4: Spectral differences $\Delta G_{sp}(\Omega_t)$ per sampling point and $f \leq 10\,\text{kHz}$ for the different types of headgear and the KU100: baseball cap (a) bicycle helmet (b) , Oculus Rift (c) and the AKG K1000 (d).

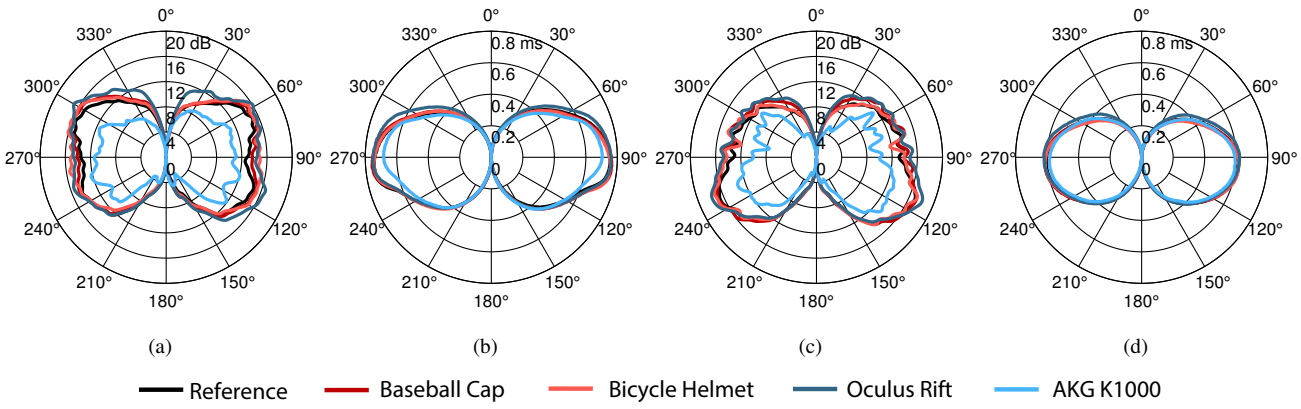

Figure 5: ILDs (a,c), and ITDs (b,d) in the horizontal plane for the reference HRTF set (black) and for the different headgear. The angle represents the azimuth ϕ of the sound incidence. The radius describes the magnitude of the level differences (in dB) or time differences (in ms). Results for the Neumann KU100 (a,b). Results for the HEAD acoustics HMS II.3 (c,d).

using sparse sets of low spatial order, more than on the influence of headgear.

5. CONCLUSION

In this paper we have described a series of measurements of spherical HRTF sets with two different dummy heads equipped with various headgear. We analyzed their influence on the spectrum and on the binaural cues. The results show that differences to the reference without headgear vary significantly depending on the type of the headgear. Regarding the ITDs and ILDs, the analysis revealed the highest influences for the AKG K1000. While for the Oculus Rift HMD, the ITDs and ILDs are affected strongest for frontal directions, generally only a very weak influence of the bicycle helmet and the baseball cap on ITDs and ILDs was observed. This suggests that localization in the horizontal plane is hardly affected by the headgear. The spectral differences to the reference are maximal for the AKG K1000, lowest for the Oculus Rift and the baseball cap. Furthermore, we analyzed for which incidence directions the spectrum is influenced most by the headgear. For the Oculus Rift and the baseball cap, the strongest deviations were found for contralateral sound incidence. For the bicycle helmet, the directions mostly affected are as well contralateral, but slightly shifted upwards in elevation. Finally,

doi:10.25836/sasp.2019.27

the AKG K1000 headphones generally have the highest impact on the measured HRTFs, which becomes maximal for sound incidence from behind.

The results of this study are relevant for applications where headgear is worn and localization or other aspects of spatial hearing are considered. This could be the case in mixed-reality applications where natural sound sources are presented while the listener is wearing an HMD, or when investigating localization performance in certain situations, e.g. in sports activities where headgear is used. Of course, our findings need to be verified for individually measured HRTF sets and be validated in a subsequent perceptual evaluation. However, it was the primary intention of this study to provide freely available HRTF sets which are well-suited for auralization purposes and which allow to further investigate the influence of headgear on auditory perception. The HRTF sets are available in the SOFA format under a Creative Commons CC BY-SA 4.0 license and can be downloaded at: http://audiogroup.web.th-koeln.de/headgear.html. The research presented in this paper has been funded by the German Federal Ministry of Education and Research, support Code: BMBF 03FH014IX5-NarDasS.

6. REFERENCES

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

[2] E. M. Wenzel and S. H. Foster, "Perceptual consequences of interpolating head-related transfer functions during spatial synthesis," in *Proc. of IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, pp. 102–105, 1993.

[3] D. R. Begault, E. M. Wenzel, and M. R. Anderson, "Direct comparison of the impact of head tracking, reverberation, and individualized head-related transfer functions on the spatial perception of a virtual speech source.," *Journal of the Audio Engineering Society*, vol. 49, no. 10, pp. 904–916, 2001.

[4] F. Brinkmann, R. Roden, A. Lindau, and S. Weinzierl, "Audibility and Interpolation of Head-Above-Torso Orientation in Binaural Technology," *IEEE Journal on Selected Topics in Signal Processing*, vol. 9, no. 5, pp. 931–942, 2015.

[5] G. Wersényi and A. Illényi, "Differences in Dummy-head HRTFs caused by the Acoustical Environment Near the Head," *Electronic Journal of Technical Acoustics*, vol. 1, pp. 1—-15, 2005.

[6] G. Wersényi and J. Répás, "Comparison of HRTFs from a Dummy-Head Equipped with Hair, Cap and Glasses in a Virtual Audio Listening Task over Equalized Headphones," in *Proc. of the 142nd AES Convention, Berlin*, 2017.

[7] R. Gupta, R. Ranjan, J. He, and W.-s. Gan, "Investigation of effect of VR/AR headgear on Head related transfer functions for natural listening," in *Proc. of the AES Conference on Audio for Virtual and Augmendet Reality*, no. August, 2018.

[8] A. Ahrens, K. D. Lund, M. Marschall, T. Dau, and H. S. Group, "Sound source localization with varying amount of visual information in virtual reality," *PLoS ONE*, vol. 14, pp. 1–19, 2019.

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

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

[11] B. Rafaely, "Analysis and Design of Spherical Microphone Arrays," *IEEE Transaction on Speech and Audio Processing*, vol. 13, no. 1, pp. 135–143, 2005.

[12] 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*, vol. 100, no. 5, pp. 972–983, 2014.

[13] R. V. L. Hartley and T. C. Fry, "The Binaural Location of Pure Tones," *Physical Review*, vol. 18, no. 6, pp. 431–442, 1921.

[14] B. Bernschütz, C. Pörschmann, S. Spors, and S. Weinzierl, "Entwurf und Aufbau eines variablen sphärischen Mikrofonarrays für Forschungsanwendungen in Raumakustik und Virtual Audio," in *Proc. of the 36th DAGA*, pp. 717–718, 2010.

[15] B. Bernschütz, "A Spherical Far Field HRIR / HRTF Compilation of the Neumann KU 100," in *Proc. of the 39th DAGA*, pp. 592–595, 2013.

[16] J. M. Arend, A. Neidhardt, C. Pörschmann, P. Christoph, and C. Pörschmann, "Measurement and Perceptual Evaluation of a Spherical Near-Field HRTF Set," in *Proc. of the 29th Tonmeistertagung - VDT International Convention*, pp. 52–55, 2016.

[17] B. Xie, "On the low frequency characteristics of head-related transfer function," *Chinese J. Acoust.*, vol. 28, pp. 1–13, 2009.

[18] F. Brinkmann and S. Weinzierl, "Comparison of head-related transfer functions pre-processing techniques for spherical harmonics decomposition," in *Proc. of the AES International Conference on Audio for Virtual and Augmented Reality*, pp. 1–10, 2018.

[19] 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.*, vol. 27, no. 6, pp. 1060 – 1071, 2019.

 doi:10.25836/sasp.2019.27