Bibliographical reference

Measurement of a Full 3D Set of HRTFs for In-Ear and Hearing Aid Microphones on a Head and Torso Simulator (HATS), Chris Oreinos and Jörg M. Buchholz, Acta Acustica **vol. 99** (Number 5), 2013, pp. 836-844

DOI

https://doi.org/10.3813/AAA.918662

Acta Acustica united with Acustica
Vol. 99 (2013) 836 – 844

DOI 10.3813/AAA.918662

Measurement of a Full 3D Set of HRTFs for In-Ear and Hearing Aid Microphones on a Head and Torso Simulator (HATS)

Chris Oreinos, Jörg M. Buchholz
National Acoustic Laboratories, 16 University Avenue, Macquarie University, Australia
Department of Linguistics, Faculty of Human Sciences, Macquarie University, Australia
The HEARing Cooperative Research Centre, Australia. chris.oreinos@nal.gov.au

Summary
The accurate reproduction of acoustic real-world environments is becoming of increasing importance in hearing device research and development. It is thereby often required to accurately predict the sound pressure at the microphones of a hearing device in a simulated or recorded acoustic environment. For that reason, an extensive set of head-related transfer functions (HRTFs) was measured in free-field with a pair of behind-the-ear (BTE) hearing aids placed on a Head and Torso Simulator (HATS). Transfer functions to the in-ear HATS microphones were also measured. A spherical head model was applied to extend the useable frequency range towards low frequencies. Special care was given to preserve the phase properties of the measurements so that the HRTFs could be widely used in phase-sensitive technical applications, including the evaluation of spatial signal processing algorithms (*i.e.* directional microphones, beamformers) in hearing devices and the evaluation of sound field resynthesis methods. The extended HRTF set can also be used for research in psychoacoustics. It is available for download at: http://www.nal.gov.au/download/HATS_BTE_hrirDatabase.zip.

PACS no. 43.66.Pn, 43.66.Qp, 43.20.Fn, 43.66.Ts

1. Introduction

Head-related transfer functions (HRTFs) contain all the spectral, temporal, and spatial information that is available in a given acoustic environment. The HRTFs in their free-field form [1] are calculated as the ratio of the Fourier transform of the sound pressure at a point in the ear canal of a subject to the sound pressure that would have been measured at the centre of the head, with the subject not being present. They contain the diffraction effect of the head and torso as well as the resonances and scattering effects of the pinna. Given that HRTFs are determined by the detailed size and shape of the head, torso, and in particular ears, they are highly individual. This is demonstrated, for example, in localization experiments, where listeners' performance decreases significantly when non-individual HRTFs are used, in particular along the cones of confusion [2]. However, non-individual HRTFs, as for example measured on a standardized head and torso simulator (HATS), are often used in applications where it is unpractical to measure individual HRTFs, such as virtual auditory spaces and computer games. Non-individual HRTFs can also be considered sufficient when analysing and verifying the general effect of multi-channel loudspeaker playback methods, where listener-specific effects are not of primary interest. In particular in the latter application, the measurement of a full 3D HRTF data set is required. By applying such a 3D HRTF data set, it is then possible to transform the loudspeaker output signals of basically any multi-channel playback system (e.g., using higher-order Ambisonics [3, 4, 5]) into ear signals. Although a large amount of 3D HRTF data sets are publically available (e.g. [6, 7, 8, 9, 10, 11], very limited HRTF data that additionally contain responses to the microphones of hearing aids (HA) fitted to a listener's (or HATS') ears are available. Such HTRF sets are essential for the research and development of hearing devices, the evaluation of hearing aid algorithms, and the verification of multi-channel loudspeaker systems for recreating realistic acoustic scenes (e.g., [12]) aimed at hearing aid testing.

When multi-microphone signal processing (or enhancement) techniques are considered in hearing devices, the HRTFs need to provide accurate phase and amplitude information. Particularly when considering delay-and-subtract directional HA processing [13], even very small phase and amplitude errors between the two HA microphones are critical at low frequencies due to the differentiation and subsequent equalization of the high-pass 6 dB/octave roll-off. Such accuracy was difficult to achieve here at low frequencies, by the measurements alone, for two reasons: (a) because a small measurement loudspea-

Received 27 November 2012,
accepted 1 August 2013.

ker, with poor low-frequency response, had to be used due to weight restrictions of moving equipment and (b) because the used anechoic chamber was not rated anechoic down to the required frequencies. To remedy that, the measured HRTFs were extended towards low frequencies by applying a spherical head model.

Although spherical head models have been previously used for this purpose [14], the applied methods for the combination of the model with the measurements have been evaluated mainly perceptually [15]. The signal processing in hearing devices is very different from the auditory signal processing, and in particular for the case of multi-microphone signal processing techniques, it is much more sensitive to small amplitude and phase variations. Hence, different (or modified) techniques are required for combining measured and modelled HRTFs. The low-frequency extension methods developed and verified in this study are important for any HRTF measurement where the available anechoic chamber has a limited frequency bandwidth or when only semi-anechoic or even reverberant chambers are available and impulse response truncation methods need to be applied to remove the room artefacts from the measured HRTFs.

The present study comprises

- The measurement of an extensive 6-channel HRTF data set for 1784 source locations covering the full 3D space using a Brüel & Kjær type 4128C HATS with two behind-the-ear (BTE) hearing aids, featuring two microphones each, fitted above the left and right pinna.
- The application of a spherical head model to extend the usable frequency range of the measured HRTFs towards low frequencies.
- A method to combine measured and modelled HRTFs that is adequate for multi-microphone hearing aid signal processing techniques. The applicability of this method is verified by considering directivity plots of an example hyper-cardioid microphone that is realised by combining the signals from the two microphones of either BTE hearing aid.

2. Methods

2.1. Measurement setup

HRTFs were measured in the anechoic chamber of the Auditory Neuroscience Laboratory, Department of Physiology, University of Sydney. The chamber is 3.5 m long, 4.6 m wide and 2.4 m high and is rated as anechoic above 300 Hz. It contains a motor controlled semicircular hoop with a radius of 1.2 m, which supports an Audience A3 loudspeaker. The hoop is fully controlled by computer software with an angular precision of less than 0.1° and is calibrated before the start of each set of measurements.

The measurements of the HRTFs were performed on the 4128C head and torso simulator (HATS) manufactured by Brüel & Kjær. The HATS complies with the ANSI standard for manikins for simulated in-situ airborne acoustic measurements (ANSI S3.36-1985 [R2006]) and the ITU-T

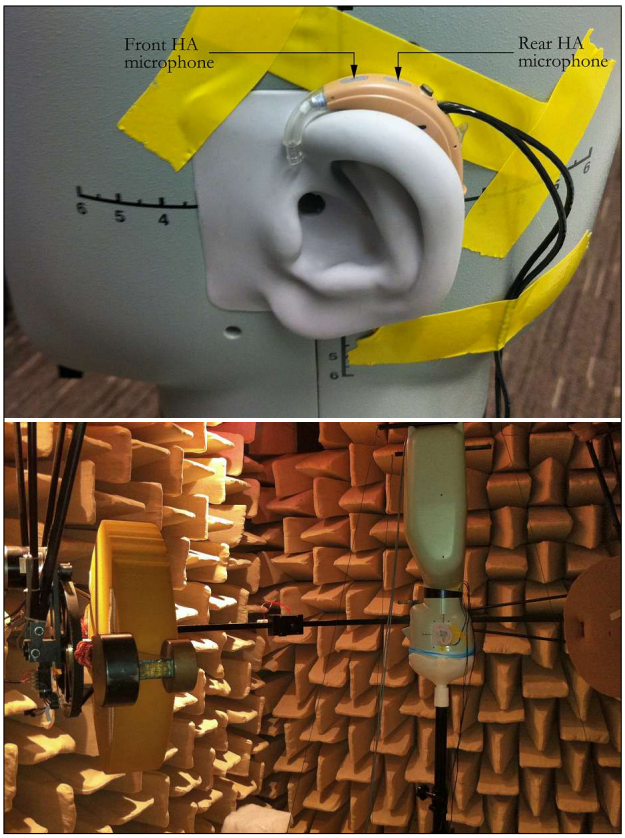

Figure 1. The acoustic mannequin used for the measurements: (a) left ear of the HATS showing the hearing aid satellite with its two microphones; (b) the HATS standing upside down on a tripod and held in place by a purpose-built plastic fitting.

P.58 recommendation for HATS telephonometry. It comes pre-fitted with 4158C/4159C type occluded ear simulators, which were connected to a pair of 2610 Brüel & Kjær measuring amplifiers. For these HRTFs measurements, a pair of –intentionally– slightly asymmetric pinnae (B&K models DZ-9763 and DZ-9764) were used. Two behind-the-ear (BTE) hearing aids were placed and secured, with sticky tape, above the HATS' left and right pinnae (Figure 1a). The hearing aids were provided by Phonak and featured cables connected to their microphones and in-built receiver (not used for this set of measurements). The signals picked up by the microphones were sent via balanced cables to a purpose-built preamplifier designed by Phonak. It must be noted that these hearing aid "satellites" did not perform any signal processing. Finally, the six channels (two in-ear plus two BTE microphones for each side) were fed to an RME Fireface 400 soundcard connected to a PC running MATLAB. Transfer functions were measured using a logarithmic sweep of 1 s duration, sampled at 44.1 kHz with 24-bit resolution.

An important issue that had to be resolved was the fact that the lowest elevation angle where the hoop could be positioned at was $\delta = -55°$. In order to work around this technical restriction the HATS had to be mounted upside down supported by its head, facing the front as in the upright placement, for some of the measurements. For that

Table I. The coordinates of the 1784 points on the sphere where the HRTFs and BTE transfer functions were measured. Resolution of elevation range: 5°. Number of measurements: N

Elevation range	Azimuth resolution	N
$[-40°, -40°]$	5°	1224
$(-60°, -40°) \cup (40°, 60°)$	8°	270
$(-70°, -60°] \cup [60°, 70°)$	10°	144
$(-80°, -70°] \cup [70°, 80°)$	15°	96
$(-90°, -80°] \cup [80°, 90°)$	30°	48
$-90°, 90°$	360°	2

purpose, it had to be placed on a custom-built mount that was shaped around the top of the head using a thermoplastic material (polymorph pellets). Figure 1b shows the setup used for the upside down mounting of the HATS. For those measurements, the source (loudspeaker) coordinates (azimuth θ and elevation δ) were referenced to the upside down HATS according to the transformations: $\delta' = -\delta$ and $\theta' = -\theta$, so that the role of the ears remained consistent. The use of asymmetric pinnae for the left and right ears of the HATS dictated this choice. If both pinnae were exactly the same, another option would have been to transform the coordinates according to $\delta' = -\delta$ and $\theta' = \theta$, while exchanging the roles of left and right side microphones (in-ear and BTE).

In both measurement conditions (standing upright and upside down) the accurate alignment of the HATS was facilitated by using two laser beam pointers mounted on the hoop at an arc distance of $\pi/2$. Correct alignment in all three axes could be attained by targeting the laser beams to the ear canals and the tip of the nose of the HATS. That resulted in effectively aligning the centre of the interaural axis with the centre of the hoop, which is the centre of the coordinate system against which the measurements were referenced. An observation that will prove useful in section 2.2 is that the centre of the interaural axis does not coincide with the centre of the HATS' head.

In order to further verify the correct alignment of the setup, all loudspeaker locations with elevation $-55° \leq \delta \leq -40°$ were measured with the HATS standing both upright and upside down. The corresponding transfer functions were compared during the measurement process and were found to match very well (absolute magnitude error between the upright and upside down HRTF measurements, averaged across all available overlapping locations, $-55° \leq \delta \leq -40°$ and $0° \leq \theta < 360°$, < 1 dB up to at least 4–5 kHz and < 2 dB up to at least 10 kHz), confirming the centred placement of the HATS in both conditions. Table I summarizes all 1784 measurement positions.

HRTFs are typically normalized by the free-field response measured with the same loudspeaker and microphones as used in the HRTF measurements, but with the microphones placed in the origin of the applied coordinate system (here the centre of the hoop) and the head (HATS) removed [1]. Unfortunately, due to time restrictions on the use of the facility, the loudspeaker response couldn't be measured with the in-ear and the BTE microphones in a free-space configuration. Instead, the loudspeaker magnitude response was equalized, with a 512-tap FIR filter, using a measurement provided by our collaborators, and the magnitude responses of the BTE microphones were equalized, with a 64-tap FIR filter, after being measured inside a hearing aid test box. The frequency response of the HATS ear simulators (4158C and 4159C) were already calibrated up to 20 kHz, on purchase of the product, so no further correction was necessary. The gain necessary to compensate for the in-ear microphones/amplifiers sensitivity was estimated by comparing the measured $(0°, 0°)$ HRTF to the calibration chart provided by B&K, for the specific HATS that was used, as well as by comparing the measured HRTFs to the ITU-T P.58 recommendation ("Head and torso simulator for telephonometry") values for HRTFs at 0°, 90°, 180°, 270° on the horizontal plane. The sensitivity of the four BTE microphones/amplifiers was compensated for by minimizing the RMS error between their individual magnitude responses, averaged over all source directions, and the corresponding averaged in-ear microphone responses in the frequency range of 300–600 Hz.

2.2. Spherical head model realisation

In this work, a spherical head model was applied to extend the useable frequency range of the HRTFs, *i.e.* to extrapolate information to lower frequencies, while preserving the sensitive phase information of the HA microphones. This model approximates the human head by a rigid sphere and as such, it does not take into account the acoustic resonances and scattering by the pinnae, at high frequencies, or the shadowing effects of the torso. Although it is an idealised model, it has been suggested that it can be used to significantly improve the low frequency accuracy of measured HRTFs [15, 16, 17]. Using the Fourier-Bessel series expansion of the wave equation solution [18] for the incident (interior problem) and scattered (exterior problem) sound field created by a single point source positioned at $(r_s, \theta_s, \delta_s)$ and imposing the boundary condition of zero total radial velocity on the rigid sphere of radius $r = a$, the pressure on its surface can be computed as [19]

$$p(r = a, \theta, \delta) = \sum_{m=0}^{\infty} \frac{\mathrm{i}^{-1}}{(ka)^2 h_m^{-'}(ka)} \frac{h_m^-(kr_s)}{h_0^-(kr_s)} (2m+1) P_m(\cos\gamma), \quad (1)$$

where h_m^- is the mth-order spherical Hankel function of the second kind, $h_m^{-'}$ its derivative and γ is the angle between the vector of the point source and that of the observation point. P_m is the mth-degree Legendre polynomial defined as

$$P_m(x) = \frac{1}{2^m m!} \frac{\mathrm{d}^m}{\mathrm{d}x^m} (x^2 - 1)^m. \quad (2)$$

Equation (1) assumes waves normalized so that they have unit amplitude and zero phase at the origin, having an implied $\mathrm{e}^{+\mathrm{i}\omega t}$ time dependence.

Table II. Spherical head model parameters.

Parameter	Value
Head radius	0.1 m
Ear canal entrance locations	L: (100°, −11°)
	R: (−100°, −11°)
Front BTE microphone locations	L: (104°, 10°)
	R: (−104°, 10°)
Rear BTE microphone locations	L: (109°, 10°)
	R: (−109°, 10°)
"Point" source distance	1.2 m

The spherical head model response was computed using the parameters shown in Table II. Given the non-spherical nature of the HATS' head as well as the existence of the torso, which was disregarded by the applied spherical head model, the parameters of TABLE II were adjusted in a more heuristic, rather than rigorous, way. The head-related values were chosen by comparison to their locations on the HATS while iteratively minimizing the RMS error of the modelled and measured ILDs and ITDs. The sound source distance was taken to be exactly the radius of the hoop where the speaker was mounted. The resulting mean absolute errors, averaged across all 1784 measurement points, were $\overline{|\mathrm{ILD}^{\mathrm{dB}}_{\mathrm{meas}} - \mathrm{ILD}^{\mathrm{dB}}_{\mathrm{mod}}|} < 0.9\,\mathrm{dB}$ and $\overline{|\mathrm{ITD}_{\mathrm{meas}} - \mathrm{ITD}_{\mathrm{mod}}|} < 46\,\mu\mathrm{s}$, where the ILDs were estimated at the region 300–600 Hz. A formal optimisation was impractical due to the computational complex nature of the spherical head model.

It should be mentioned that alternative approaches exist to the spherical head model, e.g. the snow-man model described by Algazi *et al.* [14], which are expected to provide a more accurate representation of the low frequency response of the HRTFs. However, the spherical head model already resulted in very small mean absolute ITD and ILD errors and thus, such complex models were not further considered here.

In order to maximize the computational efficiency, equation (1) was computed using a set of recursive formulas as described by Duda *et al.* [16] instead of being directly computed via the analytic expression. Given that the origin of the coordinate system used in the above spherical head model is defined in the centre of the spherical head and thus is different from the one used in the HRTF measurements (see section 2.1), the source positions (radius r'_s, azimuth θ'_s, elevation δ'_s) to be used for the spherical head model computation had to be derived from the source positions of the HRTF data set according to

$$\begin{aligned} \theta'_s &= \arctan\left(\frac{y}{x + x_{ic}}\right), \\ \delta'_s &= \arctan\left(\frac{z + z_{ic}}{(x + x_{ic})^2 + y^2}\right), \\ r'_s &= \sqrt{(x + x_{ic})^2 + y^2 + (z + z_{ic})^2}, \end{aligned} \tag{3}$$

where $[x, y, z] = [r_s \cos\theta_s \cos\delta_s, r_s \sin\theta_s \cos\delta_s, r_s \sin\delta_s]$ is the position vector of a given source, referenced to the centre of the hoop and $[x_{ic}, 0, z_{ic}]$ is the position vector of the centre of the hoop (centre of the interaural axis) referenced to the centre of the HATS head. This change of coordinate systems is necessary so that, for example, the angle γ between the vector of the left ear (100°, −11°), as referenced to the HATS' centre system, and the leftmost source point, whose coordinates expressed in the system centred at the midpoint of the interaural axis are (90°, 0°), is correctly calculated to be zero.

One-sided pressure spectra were calculated with the spherical head model for each of the 6 microphone locations (see Table II) at 4097 ($n_{fft} = 8192$) equidistant frequency points between 0 Hz and $f_s/2$. The corresponding real-valued head-related impulse responses (HRIRs) were then derived after first "mirroring" the resulting spectra to negative frequencies by applying the complex conjugate transformation: $H(-\mathrm{i}\omega) = H^*(\mathrm{i}\omega)$ and then applying an inverse (discrete) Fourier transform ($n_{fft} = 8192$). Note that the n_{fft} was chosen long enough to avoid time-aliasing after the successive time shifting and filtering of the impulses.

2.3. Combination of the head model with the measured data

A primary question that needed to be addressed was how to combine the measured HRTFs with the transfer functions derived from the spherical head model at low frequencies. Algazi *et al.* [14] have proposed cross-fading the magnitudes of the model and the data (linearly in dB) while either keeping the entire phase response of the model, or re-inserting the estimated time of arrival of the data impulses to the minimum-phase inverse-DFT of the combined (cross-faded) magnitude responses. The latter method, called minimum-phase reconstruction [15], has been perceptually verified for sound localization. However, the measured transfer functions of this HRTF set will be used as inputs to hearing aids whose processing is very different from human auditory processing. In particular, multi-microphone signal enhancement techniques (e.g., directional microphones and adaptive beamformers) are highly sensitive to very small phase and amplitude variations. Hence, the HRTFs must preserve as accurately as possible phase and amplitude relationships, especially between the individual BTE hearing aid microphone pairs. In this regard the method proposed by Algazi *et al.* [14] did not provide satisfactory results and needed to be modified.

The combination method pursued in this study consists of time aligning the model's impulse responses to the corresponding measured impulse responses and then combining them using "crossover" type low-pass and high-pass filters with cut-off frequencies in the regions of 400–500 Hz. First, the time of arrival (TOA) for all 1784 measured and modelled HRIRs was estimated separately for the in-ear and BTE microphones according to the method proposed by Nam *et al.* [20] as the time instant of the maximum peak of the absolute of the cross-correlation function between a given HRIR and its minimum-phase ver-

sion,

$$\hat{\tau} = \underset{\tau}{\operatorname{argmax}} \left\{ \left| \sum_n h[n-\tau] h_{mp}[n] \right| \right\}. \quad (4)$$

Since the BTE hearing aid microphones were only spaced apart by $d_{\mathrm{mic}} = 0.009\,\mathrm{m}$, their maximum TOA difference across all source locations was about 26 µs. Considering that the applied sampling frequency was $f_s = 44.1\,\mathrm{kHz}$ and thus, the resolution of the applied TOA estimation method was only $T = 1/f_s \approx 22.7\,\mu\mathrm{s}$, the HRIRs had to be up-sampled (a factor of 100 was chosen) before the cross-correlation operation of equation (4) was performed.

An alignment delay was then calculated from the estimated TOA of the modelled HRIRs, $\hat{\tau}_{\mathrm{mod}}$, and the corresponding measured HRIRs, $\hat{\tau}_{\mathrm{meas}}$, given by

$$\Delta = \frac{[\hat{\tau}_{\mathrm{mod}} - \hat{\tau}_{\mathrm{meas}}]}{100}. \quad (5)$$

To preserve even the smallest phase differences between each BTE microphone pair, the average alignment delay for each microphone pair was always applied to both microphones. In this way any (small) TOA estimation error only affected the absolute timing of the BTE microphone pair signals, but the inter-microphone phase differences provided by the spherical head model were preserved. Since the combination of the measured and modelled HRTFs was performed in the frequency domain, the derived alignment delays Δ were transformed into complex gain factors,

$$g\Delta(\mathrm{i}f) = \mathrm{e}^{-\mathrm{i}2\pi f \Delta / f_s}, \quad (6)$$

which were then multiplied with the corresponding modelled HRTFs. No further amplitude adjustments were required than those already described in section 2.1.

Inspired by the design of loudspeaker crossovers, 256-tap long linear-phase FIR filters were used to realise the crossover between measured and time-aligned modelled HRTFs that resembled the magnitude spectrum of a complementary pair of asymmetric Linkwitz-Riley filters. An 8th order low-pass filter with a cut-off frequency of 500 Hz and a 4th order high-pass filter with a cut-off frequency of 400 Hz were chosen. The 4th order HP filter slope combined with the roll-off of the measurements effectively resulting in a higher order HP response, almost complimentary to the 8th order LP filter. The different slopes and cut-off frequencies of the filters was chosen to: (1) apply a small gain-peak of 1–2 dB around 300–500 Hz and in turn to provide a smoother transition between measured and modelled HRTFs and (2) ensure good agreement of the combined HRTF responses with data from the ITU-T P.58 recommendation and the HATS' own calibration chart. This filter design approach in the frequency domain controlled the steepness of the slopes of the cross-over filters and thereby limited the temporal ringing of the filters in the time domain. Choosing a linear-phase filter design minimized any phase (or comb-filter) effects that could otherwise be introduced by adding the low-pass and high-pass HRTF components. The filtering was done in the frequency domain, as mentioned above, and an impulse response was computed taking the inverse DFT (as in section 2.2) of the combined modelled/measured conjugate-mirrored transfer function.

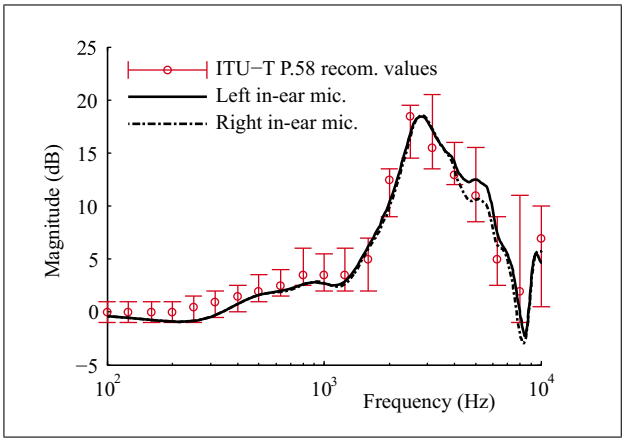

Figure 2. Processed (0°, 0°) HRTF magnitude response, extended to low frequencies using the spherical head model, compared against the values of the ITU-T P.58 recommendation.

The final impulse responses of the combined HRTFs were then truncated to 256 samples using a one-sided Tukey window with a tapered-to-constant-section-ratio of 1/5. Special care was taken in the truncation process to maintain the very short pre-ringing that was introduced by the linear phase crossover filtering process. Finally, the impulse responses were saved separately to MATLAB files (.mat extension), with every file having the impulse responses of all six microphones relevant to a given source position. The final HRTF data set is publically available at http://www.nal.gov.au/download/HATS_BTE_hrirDatabase.zip.

3. Results

Figure 2 shows the derived (0°, 0°) HRTF response (measurement combined with the spherical head model), along with the corresponding ITU-T P.58 recommendation's range of values, highlighting the compliance with this standard. The HRTFs at 0°, 90°, 180°, 270° on the horizontal plane were also compared to the relevant ITU-T P.58 values (not shown here) and were similarly found to comply.

The magnitude spectra of the measured, modelled, and combined HRTFs are plotted in Figure 3 for an example source direction of (60°, −30°). Responses for the left ear are shown in the left figures and for the right ear in the right figures. The top figures refer to the front microphone of the BTE hearing aids and the bottom figures refer to the in-ear microphones. In the top panels of Figure 3 it can be seen that the overall trend of the modelled HRTFs matches with the measured HRTF of the BTE microphone. In contrast, the measured in-ear responses deviate significantly from the model at higher frequencies. This deviation is

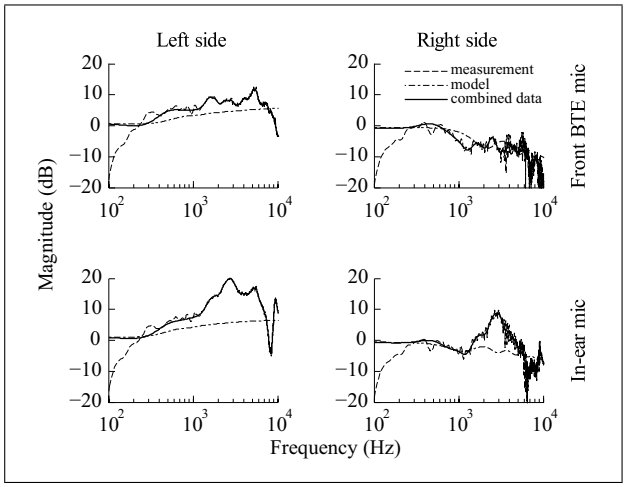

Figure 3. Spectrum magnitude of HRTF measurements (dashed curve), spherical head model (dash-dotted curve) and final combined and windowed responses (thick solid curve) relevant to the source direction . The left and right panels show left and right ear microphone responses, respectively. The top curves refer to the front BTE microphones and the bottom curves to the in-ear microphones.

mainly due to the ear canal resonances, the most prominent of which is at around 2.5 kHz, as well as the effect of the pinnae. At low frequencies the application of the spherical head model increases the usable bandwidth and removes the low frequency dips and peaks, which are due to the measurement chamber (which is not anechoic below 300 Hz) room modes.

When comparing the details contained in the modelled and measured spectra, in particular at frequencies above about 1 kHz, clear differences can be observed. These differences can be even more pronounced at other source positions, especially at the contralateral side where more extensive rippling occurs. However, these discrepancies are expected because the shape of the head of the HATS is not strictly spherical and the presence of the torso and pinnae has not been considered in the spherical head model. Since the model is only applied to extend the usable frequency range below about 300 Hz, the mentioned discrepancies are not significant and can be neglected.

Figure 4 shows the HRTF magnitudes of the left in-ear and front HA microphones for all available source positions on the horizontal plane. The contour plots of the in-ear responses can be qualitatively compared to similar plots found in literature [11, 7, 17, 21]. However, a quantitative one-to-one comparison is not meaningful unless one is comparing measurements from the same mannequin, fitted with the same pinnae and ear canal and simulator and driven by a similar measuring loudspeaker at the same distance. Comparing Figure 4a and 4b, it is evident that the ear-canal resonances at ∼2.5 kHz and ∼5 kHz are eliminated in the BTE responses. Additionally, some of the high frequency fine structure appears smoothed and attenuated due to the absence of pinna reflections.

As a next step to validate the combined HRTFs, the Interaural Level Differences (ILDs) versus frequency are

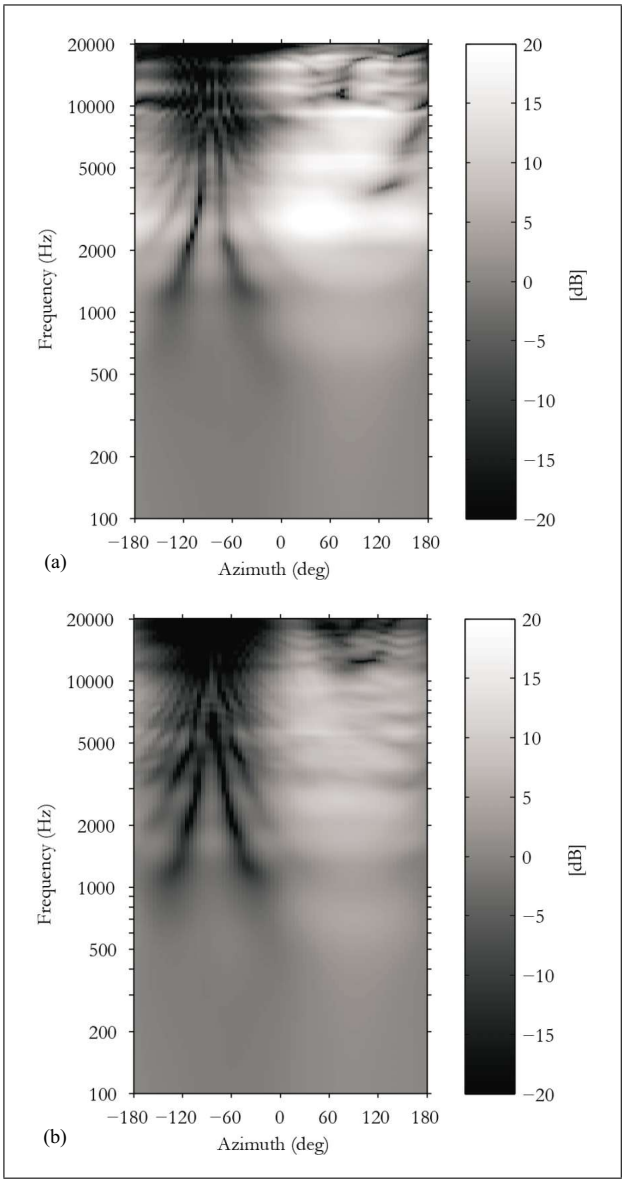

Figure 4. Contour plots of the magnitude of HRTFs to the: (a) in-ear and (b) hearing aid (front) microphone of HATS' left side. The abscissa of the plots depicts the source direction on the horizontal plane, with azimuth $\theta \in [-180°, 180°]$.

plotted in Figure 5 for sources located on the horizontal plane. ILDs were calculated as the magnitude-difference between the left and right ear's HRTFs. A number of observations can be made from these plots. First, due to the different pinnae used on the measured HATS the ILD for $\theta = 0°$ is slightly different from zero at high frequencies. A second observation is that for positive azimuths close to 90° (source on-axis with left ear) a positive ILD, which takes substantial values at high frequencies, is observed, while an almost symmetrical picture up to about 5 kHz is seen for negative azimuths (above ∼5 kHz the used pinnae cause deviations to the symmetric nature of the ILDs). The theoretical bright spots at ±90° are also confirmed on that plot (*i.e.*, a source at 90° creates a bright spot at 180° around the perimeter of the sphere, that is at −90°,

where the contralateral ear lies). As mentioned above for the HRTF magnitude case, it is not straightforward to compare the ILD contours from another data set, unless it is ensured that all geometrical, structural and acoustical properties of the setup are the same. The ILDs measured at the BTE-microphones (not shown here) look very similar to the in-ear ILDs of Figure 5 at low frequencies, but above about 5 kHz are significantly smaller due to absence of the pinna and ear canal effects.

The broadband ITD of the measured HRTFs as well as the modelled HRTFs before and after TOA alignment (see section 2.3) are shown in Figure 6 for the in-ear microphones as a function of source azimuth. The ITDs were calculated as the TOA difference between the left and right ear microphones using equation (4). The ITDs for the measurement (solid line) as well as the TOA-aligned model (dashed line) match very well, confirming the accuracy of the TOA-alignment procedure described in section 2.3. Moreover, these ITDs exhibit the typical ITD shape measured in humans as seen for example in [20] or [22]. The ITDs for the modelled HRTFs before TOA-alignment (dashed-dotted line) show a slightly less "peaky" behaviour than the ITDs for the measured HRTFs, but otherwise provide a very good fit and thus confirm the applicability of the parameters given in Table II for the calculation of the spherical head model responses. As expected, the ITDs estimated from the HA microphones (not shown here) look almost the same as those of Figure 6.

In order to verify the accurate representation of phase (or timing) of the combined HRTFs, in particular between the front and back microphones of the BTE hearing aids, 3D directivity plots of a hyper-cardioid and an omnidirectional microphone were realised and shown in Figure 7 for two example frequencies, 300 Hz and 1600 Hz. The radial distance of each directivity plot surface node represents the magnitude of the directional microphone output response (in dB) due to a point-source located at that node's azimuth and elevation at a distance of $r = 1.2$ m (distance between loudspeaker and HATS' head centre). The hyper-cardioid output was generated by applying delay and subtraction beamforming [13] to the HRTFs of the front and back microphones of the left BTE hearing aid. Such an approach for creating directional microphones is highly sensitive to phase and amplitude errors, in particular at low frequencies. If the phase difference between the microphones is slightly disrupted (as was the case in the measured HRTFs at low frequencies), the directivity plots get severely distorted. The bottom polar plots of Figure 7 show the hyper-cardioid directivity attained at low frequencies (300 Hz) to be smooth and resembling that at higher frequencies (1600 Hz). Thus the low frequency region where the model dominates the combined response produces a directional result that resembles that at higher frequencies where the measured data dominate.

Note that the presented directivity patterns differ from the well-known ideal hyper-cardioid polar plot [13] due to the existence of the head. The head shadowing effectively creates a higher-order directional behaviour, as compared

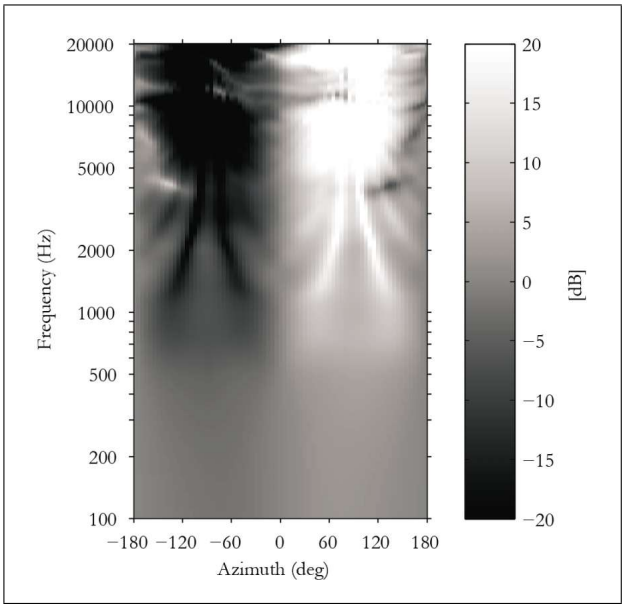

Figure 5. Interaural level differences between left and right in-ear microphones ($H_L - H_R$), versus frequency, relevant to positions lying on the horizontal plane with azimuth $\theta \in [-180°, 180°]$.

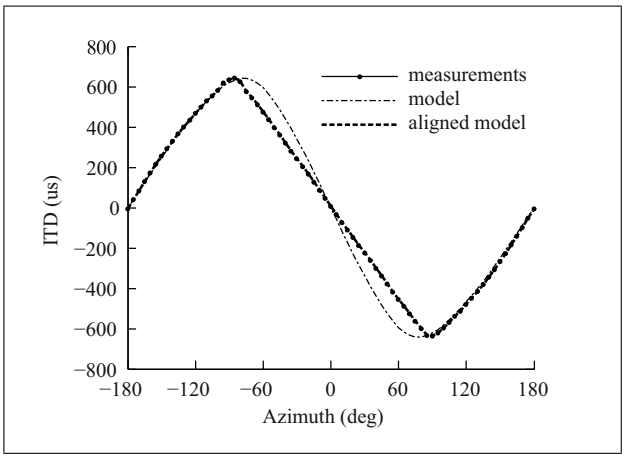

Figure 6. Interaural time differences between left and right in-ear microphones relevant to positions lying on the horizontal plane with azimuth $\theta \in [-180°, 180°]$. The ITD estimates of the measurements and the model, before and after time alignment, are plotted.

to the first order ideal free-field patterns. In order to illustrate the head shadow effect on the microphone directionality, the directivity patterns of an omnidirectional BTE microphone on the head (*i.e.* only considering the front microphone) are shown in the top figures of Figure 7 at 300 Hz and 1600 Hz. They are very similar to the directivity patterns shown by Kates [13], further demonstrating the validity of the measured HRTFs.

Lastly, Figure 8 shows the SNR, averaged across all six microphone channels (which gave very similar values) and all measurement positions (solid line). A noise estimate was formed using the last 256 samples of the 2048-sample long unprocessed measurement data. A signal estimate (in fact signal + noise) was formed using samples (100:999)

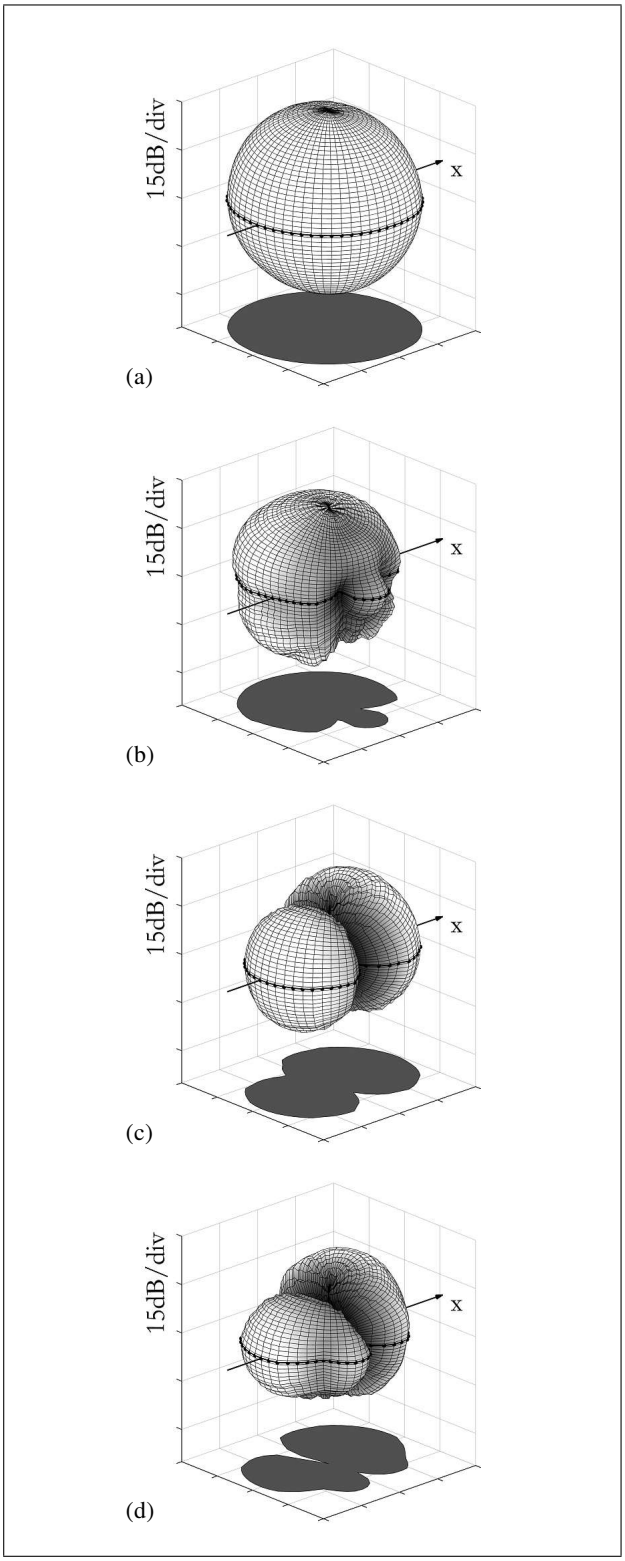

Figure 7. Directivity patterns of the left BTE microphone pair (worn on the HATS) for an omnidirectional configuration (*i.e.*, using only the signals of the front BTE microphone) at (a) 300 Hz, (b) 1600 Hz and a hyper-cardioid configuration at (c) 300 Hz, (d) 1600 Hz. The 2D shadow plots illustrate the directivity plots of the 3D patterns on the horizontal plane.

from the measurement data. The estimated SNR does not strictly depict the true microphone – amplifier – A/D SNR

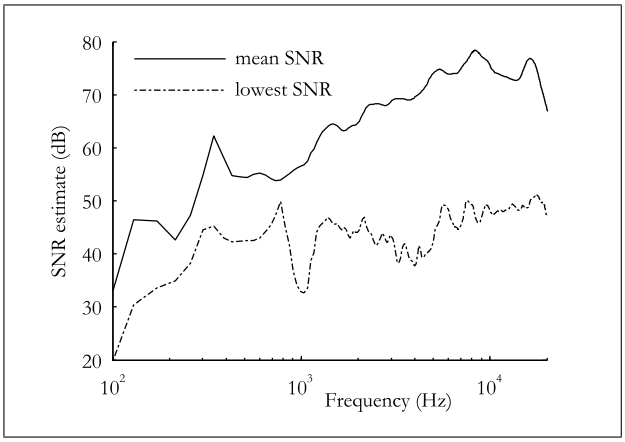

Figure 8. The frequency-dependent mean SNR of all microphone channels averaged across all positions (solid curve) along with the SNR of the measurement with the lowest *broadband* SNR (dash-dotted curve).

but describes the effective "SNR" after averaging and de-convolving the recorded sine sweeps. The frequency-dependent SNR of the measurement with the worst *broadband* SNR, occurring at the position (30°, −85°), is also plotted as a worst-case scenario. It can be seen that the SNR at low frequencies (below 300–400 Hz) is significantly lower than the values at mid and high frequencies. Especially if the 6 dB/octave roll-off of a delay-and-subtract directional HA microphone is considered, the output SNR becomes negative at low frequencies. This confirms again the necessity of applying the spherical head model for a "phase-preserving" extrapolation of the magnitude to low frequencies.

4. Summary and conclusions

The procedure for the measurement of an extensive set of HRTFs from 1748 directions covering the whole sphere to the 2 in-ear and 4 BTE hearing aid microphones on a HATS was presented. A spherical head model was applied to extend the usable frequency range of the measured HRTFs towards low frequencies. A method was also proposed to combine the measurements with the model responses while maintaining the very sensitive amplitude and phase information carried by the HRTFs, which is particularly critical between the front and back hearing aid microphones. The applicability of this approach as well as the validity of the entire HRTF data set, especially when used for hearing aid applications, was confirmed by analyzing magnitude spectra of the combined HRTFs, ITDs, ILDs and directivity plots of hyper-cardioid and omnidirectional microphones.

The developed HRTF data set covering the full sphere at a high resolution can be used in a wide range of applications ranging from recreating virtual auditory spaces to research and development with hearing devices, evaluation of hearing aid algorithms, including psychoacoustic aspects, and verification of multi-channel loudspeaker systems for recreating realistic acoustic scenes. The ex-

istence of measurements from both in-ear and BTE microphones permits realistic simulations of aided scenarios where sound reaches the hearing aids of the wearer but at the same time leaks through to the eardrum (this "acoustic path" will be affected by the fitting of the hearing aid, which may include ear-molds with different vent sizes).

Acknowledgement

The authors acknowledge the financial support of the HEARing CRC, established and supported under the Cooperative Research Centres Program – an initiative of the Australian Government. In addition, we would like to thank Johahn Leung, Heather C. Kelly, and Prof. Simon Carlile for letting us use the Auditory Neuroscience Laboratory, Department of Physiology, University of Sydney, as well as two anonymous reviewers for their helpful suggestions.

References

[1] J. Blauert: Spatial hearing: The psychophysics of human sound localization. MIT, Cambridge, MA, 1983.

[2] E. M. Wenzel, M. Arruda, D. J. Kistler, F. L. Wightman: Localization using nonindividualized head-related transfer functions. J. Acoust. Soc. Am. **94** (1993) 111–123.

[3] S. Moreau, J. Daniel, S. Bertet: 3D sound field recording with higher order ambisonics. Objective measurements and validation of spherical microphone. AES 120th Convention, Paris, France, 2006.

[4] D. B. Ward, T. D. Abhayapala: Reproduction of a plane-wave sound field using an array of loudspeakers. IEEE Trans. Speech Audio Process. **9** (2001) 697–707.

[5] N. Epain, P. Guillon, A. Kan, R. Kosobrodov, D. Sun, C. Jin, A. van Schaik: Objective evaluation of a three-dimensional sound field reproduction system. Proceedings of 20th International Conference on Acoustics, ICA, Sydney, Australia, 2010.

[6] W. G. Gardner, K. D. Martin: HRTF measurements of a KEMAR. J. Acoust. Soc. Am. **97** (1995) 3907–3908.

[7] V. R. Algazi, R. O. Duda, D. M. Thompson, C. Avendano: The CIPIC HRTF database. Proc. IEEE Workshop on Applications on Signal Processing to Audio and Acoustics, 2001, 99–102.

[8] E. Grassi, J. Tulsi, S. Shamma: Measurement of head-related transfer functions based on the empirical transfer function estimate. Proc. International Conference on Auditory Display, 2003, 119–122.

[9] T. Nakado, T. Nishino, K. Takeda: Head-related transfer function measurement in sagittal and frontal coordinates. Acoust. Sci. & Tech. **29** (2008) 335–337.

[10] H. Kayser, S. D. Ewert, J. Anemüller, T. Rohdenburg, V. Hohmann, B. Kollmeier: Database of multichannel in-ear and behind-the-ear head-related and binaural room impulse responses. EURASIP, J. on Advances Signal Processing (2009).

[11] B. P. Bovbjerg, F. Christensen, P. Minnaar, X. Chen: Measuring the head-related transfer functions of an artificial head with a high directional resolution. AES 109th Convention, Los Angeles, USA, 2000.

[12] M. Pauli, S. Favrot, J. M. Buchholz: Improving hearing aids through listening tests in a virtual sound environment. Hearing Journal **63** (2010) 40–44.

[13] J. M. Kates: Digital hearing aids. Plural Publishing, San Diego, CA, 2008.

[14] V. R. Algazi, R. O. Duda, D. M. Thompson: The use of head-and-torso models for improved spatial sound synthesis. AES 113rd Convention, Los Angeles, USA, 2002.

[15] D. J. Kistler, F. L. Wightman: A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction. J. Acoust. Soc. Am. **91** (1992) 1637–1647.

[16] R. O. Duda, W. L. Martens: Range dependence of the response of a spherical head model. J. Acoust. Soc. Am. **104** (1998) 3048–3058.

[17] D. S. Brungart, W. M. Rabinowitz: Auditory localization of nearby sources. Head-related transfer functions. J. Acoust. Soc. Am. **106** (1999) 1465–1479.

[18] E. G. Williams: Fourier acoustics. Academic, San Diego, CA, 1999.

[19] J. Daniel: Spatial sound encoding including near field effect: Introducing distance coding filters and a viable, new ambisonics format. AES 23rd International Conference, Copenhagen, Denmark, 2003.

[20] J. Nam, J. S. Abel, J. O. Smith: A method for estimating interaural time difference for binaural synthesis. AES 125th Convention, San Francisco, USA, 2008.

[21] R. Duraiswami, D. N. Zotkin, N. A. Gumerov: Interpolation and range extrapolation of HRTFs. Proceedings of the ICASSP '04, 2004.

[22] S. Busson, R. Nicol, B. F. G. Katz: Subjective investigations of the interaural time difference in the horizontal plane. AES 188th Convention, Barcelona, Spain, 2005.