Controlling the Head Position during individual HRTF Measurements and its Effect on Accuracy

Florian Denk[1,4], Jan Heeren[1,2], Stephan D. Ewert[1], Birger Kollmeier[1], Stephan M.A. Ernst [1,3]
[1] *Carl von Ossietzky Universität Oldenburg, Medizinische Physik and Cluster of Excellence Hearing4all*
[2] *Also at: Kompetenzzentrum HörTech, Oldenburg, Deutschland*
[3] *Now at: Universitätsklinikum Gießen und Marburg GmbH, Klinik für Hals-Nasen-Ohrenheilkunde, Standort Gießen*
[4] *Email: florian.denk@uni-oldenburg.de*

Introduction

The head-related transfer function (HRTF) describes the incident direction and frequency dependent representation of sound at the eardrum of a listener, caused by transmission effects of the specific outer ear, head and torso. It is an individual "fingerprint" that we are used to hear the world through, and when the HRTF of a subject is known, persuading simulation of free-field listening can be achieved by processing a signal with the HRTF prior to presentation over headphones [1].

When measuring HRTFs, one source of error is movement of the subject, particularly the head, causing the effective sound incident direction to dynamically deviate from the desired one. Such movement is inevitable and occurs even if the subjects are instructed to remain still, and do in their subjective impression. Commonly, this source of inaccuracy is addressed by providing a headrest or similar mechanic support, which however cannot avoid such movement completely [2–6]. In this work, a novel method for both recording and interactively controlling the head position based on visual feedback is presented. The approach was successfully implemented and applied in individual HRTF measurements, where repeated rounds of measurements were conducted and a stable subject position was particularly crucial. The results show that excellent control over the head position and orientation is achieved, and the connected residual influence on the obtained HRTFs is in a negligible size.

Headtracking and Visual Feedback for Head Positioning Control

The general approach is to continuously monitor the head position in all degrees of freedom[1] (DOF) using a head tracker, and to graphically display the necessary corrections to restore a reference position to the subject. This allows them to interactively adapt the head position in a way that a target position is achieved and conserved.

The realization of the according Graphical User Interface (GUI) is shown in figure 1. The current head misalignment in each of the six DOFs is visualized to the subject via arrows of variable colouration, two for each dimension, such that the necessary movement to restore the resting position is directly displayed. The colour of the arrows can change gradually between grey (position okay) to deep red (highest deviation). For a better handling and easy adjustment of tolerable deviations, the observed error ε in each dimension is converted to an error score $\mathrm{ER}_{\mathrm{pos}} \in [-1, 1]$ by applying a sigmoidal dependence

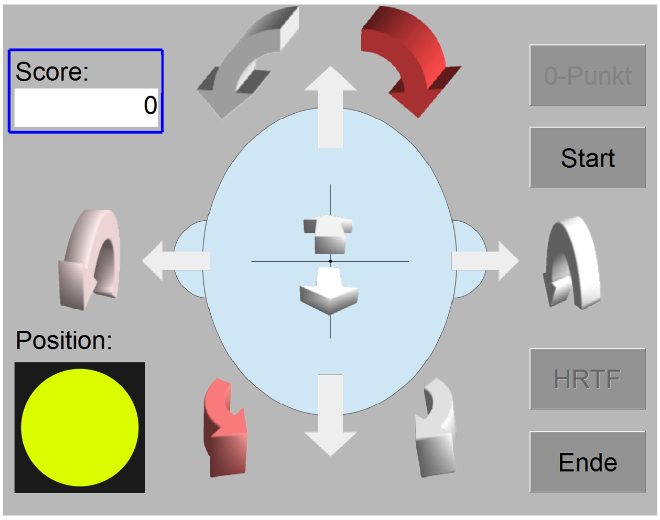

Figure 1: Graphical User Interface (GUI) visualizing the deviation of the current head position from a reference in 6 degrees of freedom. The variable arrow colour guides to subject to realign the head position to a pre-defined reference.

$$\mathrm{ER}_{\mathrm{pos}} = \frac{\mathrm{sign}(\varepsilon)}{1 + \exp\left[-k\left(|\varepsilon| - \tau_{\mathrm{pos}} + \frac{\ln \frac{\mathrm{ER}^{(\tau)}}{1-\mathrm{ER}^{(\tau)}}}{k}\right)\right]}. \quad (1)$$

There, τ_{pos} is the allowed tolerance and k is a parameter allowing to adjust the steepness of the relation to an individual preference. When the current error is equal to the tolerance, the error score $\mathrm{ER}^{(\tau)}$ is displayed, for which 0.8 was found to be an appropriate value. Figure 2 shows examples for the conversion as used in this work. Besides the misalignment in the individual DOFs, an overall position indicator is implemented as a traffic light sign in the lower left corner, visualizing the maximum deviation with respect to the tolerance. Red colour indicates that at least in one DOF the misalignment is intolerable. Aiming to further increase the motivation of the subjects, gamification is included in the task by keeping a score visible to the subjects

[1] x (front/back), y (left/right)and z (up/down) for translational movements, euler angles Yaw $\tilde{\varphi}$, Pitch $\tilde{\vartheta}$ and Roll $\tilde{\zeta}$ for rotations. The term "position" in this work includes all coordinates, i.e. determines both translational and rotational coordinates.

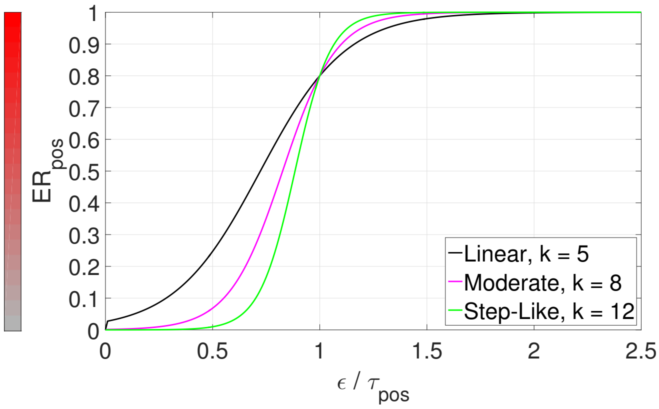

Figure 2: Conversion from observed position error ε (only positive-definite error range displayed) with respect to allowed tolerance τ_{pos} to the error score ER_{pos} according to equation (1). The value of ER_{pos} determines the shade of red of the arrows in figure 1. As in the presented experiments, $\text{ER}^{(\tau)}$ was set to 0.8.

based on the current head position, which is updated during measurements. Additional support is provided by a gazing point in the centre of the GUI, which coincides with the acoustic axis during the measurements.

Individual HRTF Measurements

HRTF measurements were conducted with 16 subjects in the Virtual Reality Lab of Oldenburg University shown in figure 3, which is an anechoic chamber featuring a 98-channel 3D loudspeaker setup, as well as a video system. The configuration allows simultaneous measurement of all incident directions using the Multiple Exponential Sweep Method (MESM, [7]). The individual sweeps covered the frequency range between 50 and 20000 Hz with a length of 4.1 seconds, leading to a duration of 36 seconds for measuring the transfer functions from all loudspeakers. HRTFs to various points in the ear were assessed, and the measurement was repeated for seven earplug styles. In each condition, four repetitions were performed. The overall duration was between 70 and 90 minutes, depending on the subject.
For controlling the head positioning, the method introduced in the previous section was utilized in combination with a headrest. The GUI as shown in figure 1 was presented to the subject via a beamer on a small canvas mounted directly before the frontal speakers using fishing line, and pre-tests verified a negligible acoustic influence of the specialized fabric. The head position data is delivered by a small electro-magnetic head tracker (*Pohlemus Patriot*), which is mounted on the subjects' head using the modified interior of a construction helmet, as shown in figure 3. The bracket provides tight fit on the head, and much care was taken to avoid any contact during the experiment. The sensor's small size reduces the acoustic influence to a minimum, and a sampling rate of 60 Hz with sub-mm and sub-degree resolution facilitate high-accuracy real-time implementation of the interface. The GUI is mirrored to the screen of a PC in the control room

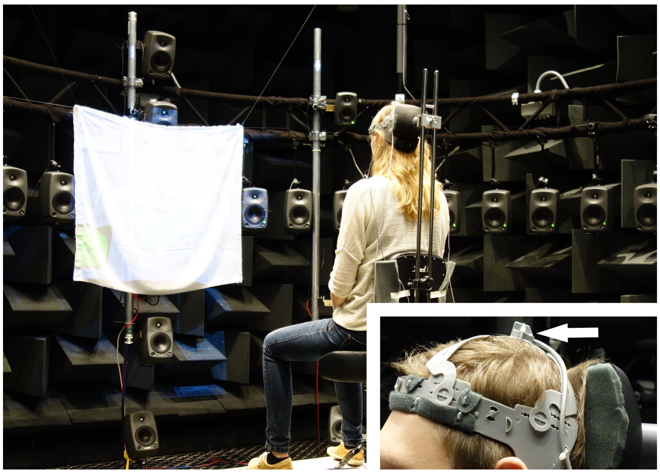

Figure 3: Subject performing the HRTF measurements, up front is the canvas where the GUI is displayed. Inset: *Pohlemus Patriot* positioning sensor (small cube on top of head marked by arrow) attached to a subject's head using the modified interior of a construction helmet.

Table 1: Default Head Positioning Tolerances. For coordinates see footnote[1].

in cm			in °		
x	y	z	$\tilde{\varphi}$	$\tilde{\vartheta}$	$\tilde{\zeta}$
1	1	1	0.6	0.8	0.8

for the experimentator, who can control the parameters such as allowed tolerances, as well as the audio measurement.
After a short period of training, all subjects were able to handle the displayed information, which on the first sight may seem overcharging, and successfully used the interactive control method. The default tolerances shown in table 1 were derived from minimum audible angles in the corresponding dimensions and what was achievable in informal pilot tests. For only a minor number of subjects (2 out of 16), it was necessary to adjust the tolerances to about twice the given values to allow non-frustrating use. Pilot tests showed that using the presented approach, the head position can be held constant even without a headrest. However, in this case the short-term variance as well as the perceived effort were considerably increased, hence a headrest was provided for the current study. On the linux computer the experiment was conducted with, the graphic update rate was about 12/s, achieving a smooth representation of the real-time head position. For the error calculation in each video frame, the median of the latest head position samples (about 5) was used. Also, all head positioning data is stored for further evaluation. The initial positioning of the subjects in the room was performed by means of a pendulum and a laser distance measurement device, and stored internally as the reference the head was aimed to be adjusted to.

Head Positioning Results

For evaluation, the head tracker data obtained during the measurements was segmented into the individual

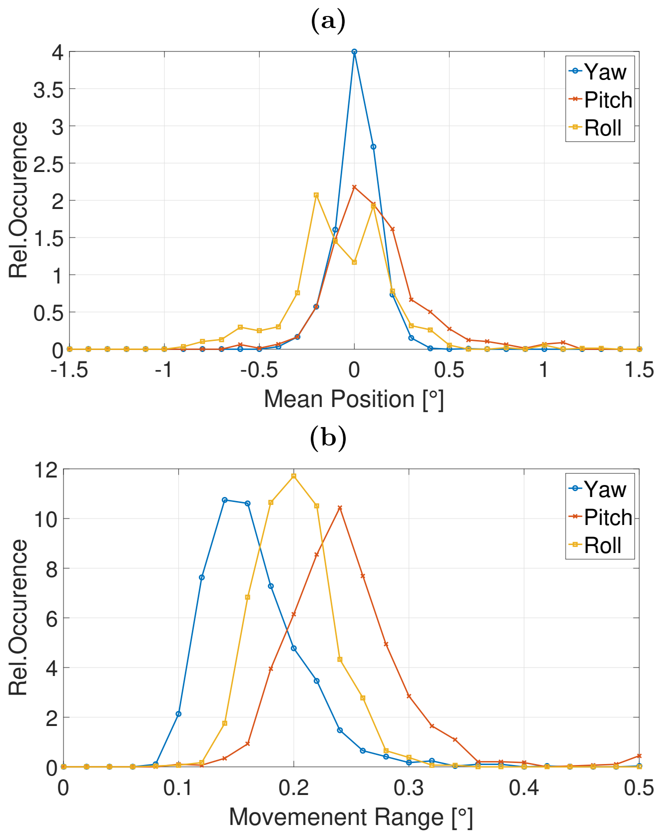

Figure 4: Distributions of head position observed during measurement of each individual HRTF **(a)** achieved mean head position (error against original position) and **(b)** variation of head position during single measurements.

sweep playback times, yielding a course of head positions for each measured incident direction. Figure 4(a) a shows the histogram of the mean head orientation in each sweep, after selection of the trial with the best head position for each incident direction and in-ear measurement point. The range is below $\pm$ 0.5° for the overwhelming majority of samples, i.e. well-bounded by the allowed tolerances and particularly below the worst-case minimum audible angles in the corresponding dimensions. Also, the position error range is only marginally larger than the observed position variation inside one transfer function measurement lasting 4.1 s, expressed by the double standard deviation of head position in each DOF and shown in figure 4(b). This observation may be interpreted such that the control of the mean position is bounded by (inevitable) short-term head movements, and may thus be close to the optimum of what is achievable.

Any misorientation in the individual DOFs results in an effective sound source position that is different from the one intended. This displacement can be expressed as the (shortest possible) angular displacement between both when projecting the locations on a sphere around the head. Figure 5 shows the result for the sequence of all measured HRTF incident directions over the course of the experiment, pooled over the subjects. In the median over the subjects, the angular error is around 0.3°, and even the largest deviations hardly exceed 0.5°. Also, an excellent stability over the course of the measurement is

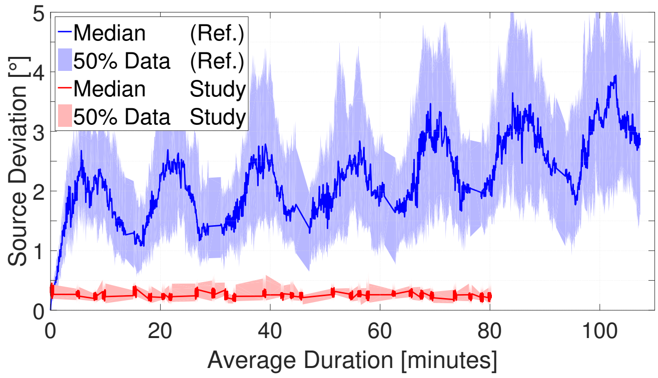

Figure 5: Angular displacement between target and effective sound incident directions, showing the median and 50 % data range (i.e., between 25% and 75% quantiles) of individual measurements over the course of the experiment in minutes. Comparison of a comparable experiment where only a headrest was used (from [5], blue lines) with the current study (red lines).

observed. For comparison, the according data from a similar study [5] is shown in the same figure, where the overall duration was 90-120 minutes and 45 subjects were included. There, the misalignment in the first measurement was removed from all trials, eluding effects of incorrect initial position of the subject. While the head position was recorded for later correction, no visual feedback was provided and only a mechanic support analogous to the present study was used - this represents the standard procedure in such measurements. As compared to this state-of-the art reference, the observed source position deviation is decreased by a factor of about 5-10 when using the proposed visual feedback, and also the variance across subjects is considerably reduced. Another observation is an improved time-stability: Whereas virtually no dependence is observed with the visual feedback method, a systematic deterioration is obvious in the reference case in the very beginning, as well as towards the end.

It is also worth noting that the subjective reactions to the interactive interface were very positive. The fact that minor head movements can be reliably corrected gives the possibility to allow them in pauses, which was a great relief to the participants. Another common feedback was that it is good to have something to do during the acoustic measurements besides listening to sweeps.

Effect of Positioning Errors on HRTF

Finally, the effect of residual head position differences on the measured HRTF is assessed. For a better comparability to other literature, only HRTFs from the presented measurements measured at the blocked ear canal entrance are considered. The evaluation is performed by considering the deviation in HRTFs observed between the 4 trials, and comparing it to the corresponding difference in head position. Thereby, for each incident direction only the combination of trials with the most different head position was considered. As a perceptually relevant difference metric, the impulse responses were processed

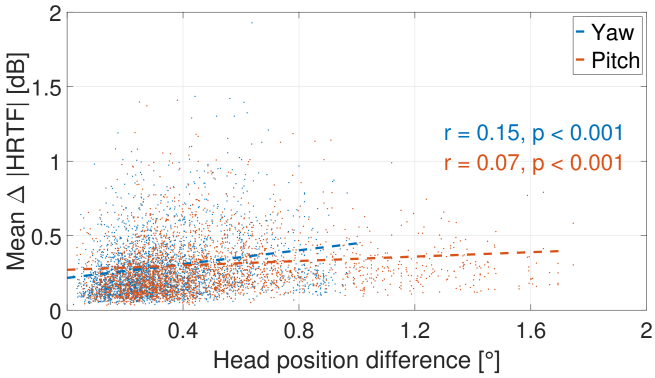

Figure 6: Mean HRTF Error between two trials in dependence of difference in head positions, with linear fits and correlation coefficients.

by a 4-th order gammatone filterbank [8] with half-ERB spacing. Then, the spectral amplitude difference in dB was averaged over the rms-output of filters with centre frequencies between 1 and 12 kHz. This mean-HRTF error in dependence on the head orientation difference in the yaw domain as well as the pitch domain is shown in figure 6, where rare (ca. 10/3000) outliers exceeding 2 dB have been excluded.

The mean HRTF deviation shows significant correlations with both the yaw and the pitch difference, however this dependence is not very strong. It should be kept in mind that the head position difference between trials is very small, and for differences $< 0.3°$ the inevitable movement during one sweep is in the same order of magnitude. It can thus be stated, that in terms of the head position, the observation is in most trials in the range of those inevitable inaccuracies. Although the dependence of the HRTF magnitude error depends significantly on the head position, the resulting deviations are below 0.5 dB for the vast majority of observations. Altogether, the influence of head position deviations observed here on the HRTF error can be termed negligibly small - which will particularly hold after selection of the trials with the best head position, as it can be performed in the presented study.

Conclusion

We presented a novel method for controlling the head position of human subjects based on an interactive graphical display, and successfully applied it in individual HRTF measurements. A clear benefit was demonstrated in comparison to baseline results where only a headrest is provided (as is common practice). The head position was stabilized to an accuracy of about 0.5°, which mainly appears to be bounded by small involuntary head movements of the subjects on a short time scale. The connected errors in the measured HRTF were usually below 0.5 dB and can probably be termed negligible, however a significant dependence of the error on the head orientation verifies the need for accurate control methods like the proposed one.

Acknowledgement

The work was funded by DFG FOR1732 *Research Unit Individualised Hearing Acoustics*. We thank Thibaut Carpentier and Olivier Warusfel from IRCAM for kindly providing their head positioning data measured in [5].

References

[1] Henrik Møller. “Fundamentals of Binaural Technology”. In: *Applied Acoustics* 36.3/4 (1992), pp. 171–218.

[2] Tatsuya Hirahara et al. “Head movement during head-related transfer function measurements”. In: *Acoustical Science and Technology* 31.2 (2010), pp. 165–171.

[3] G. Wersényi and J. Wilson. “Evaluation of the Range of Head Movements in Measurements and Recordings with Human Subjects using Head-Tracking Sensors”. In: *Acta Technica Jaurinensis* 8.3 (July 2015), p. 218.

[4] G. König and W. Sussmann. “Zum Richtungshören in der median-sagittal-ebene”. In: *European Archives of Oto-Rhino-Laryngology* 167.2 (1955), pp. 303–307.

[5] Thibaut Carpentier et al. “Measurement of a head-related transfer function database with high spatial resolution”. In: *7th Forum Acusticum (EAA)*. 2014.

[6] V. Ralph Algazi, Carlos Avendano, and Dennis Thompson. “Dependence of subject and measurement position in binaural signal acquisition”. In: *Journal of the Audio Engineering Society* 47.11 (1999), pp. 937–947.

[7] Piotr Majdak, Peter Balazs, and Bernhard Laback. “Multiple exponential sweep method for fast measurement of head-related transfer functions”. In: *Journal of the Audio Engineering Society* 55.7/8 (2007), pp. 623–637.

[8] V. Hohmann. “Frequency analysis and synthesis using a Gammatone filterbank”. In: *Acta Acustica united with Acustica* 88.3 (May 2002), pp. 433–442.