Investigation on sound localization performance in a virtual acoustic environment designed for hearing aid users

Florian Pausch[1,2], Michael Kohnen[1], Lukas Aspöck[1], Z. Ellen Peng[1,2], Janina Fels[1,2]
[1] *Institute of Technical Acoustics,* [2] *Medical Acoustics Group, 52074 Aachen, Germany,*
Email: florian.pausch@akustik.rwth-aachen.de

Abstract

Sound localization performance in binaural reproduction systems using acoustic crosstalk cancellation filters and loudspeakers relies on head-related transfer functions and has been previously examined in several studies on normal-hearing listeners. For applications on listeners fitted with bilateral hearing aids, the binaural playback setup has to be extended to include a pair of research hearing aids in the playback of simulated binaural signals based on hearing aid-related transfer functions. The fusion of these two playback paths enables to include the remaining hearing capabilities of a listener with hearing loss and to simulate sound perception over hearing aids. In this study, sound localization performance in a system designed for hearing aid users is evaluated with normal-hearing adults.

Introduction

Localization is a crucial performance parameter for systems that mimic real-life acoustic environments by means of virtual acoustic environments (VAE). It is of particular interest how accurate virtual sound sources (VSS) can be localized in comparison to real sound sources (RSS) [1].
In conventional binaural playback systems, which are based on head-related transfer functions (HRTF), an external soundfield is simulated. In such a system, playback is realized over loudspeakers where a binaural input signal is filtered by acoustic crosstalk cancellation (CTC) filters [2] to achieve a sufficiently high channel separation [3, 4].
In order to make VAE accessible for people fitted with bilateral hearing aids (HA), this system is extended by a pair of research hearing aids (RHA) with access to microphone input signals and receiver output signals. This enables not only to measure hearing aid-related transfer functions (HARTF) but also to play back simulated, measurement-based receiver signals to create VSS, as done by Mueller et al. [5].
Yet little is known about the localization performance when combining the two described binaural playback systems. Hence, this study evaluates the system designed for HA users with normal-hearing adults in a dynamic localization experiment. In the current experiment, stimuli are presented by VSS in three conditions: RHA only, loudspeakers only with CTC filters, and combined RHA and loudspeakers with CTC filters. The results are compared to a previous localization experiment where RSS were presented over loudspeakers and VSS were presented binaurally over headphones. Both experiments are evaluated in terms of angular error measures.

Methodology

System description

Figure 1 shows the extended binaural reproduction system which is installed in an acoustically optimized hearing booth [4] with a room volume of approximately $10.5\,\mathrm{m}^3$.
As stimulus, a white noise pulse-train with hard on-/offsets and a total duration of 2.25 s with an intermediate pause of 0.25 s is used (PCM encoded, sampled at $f_s = 44.1\,\mathrm{kHz}$ with 16 bits resolution). The stimulus length is chosen to enable head movements during presentation.
The VSS are generated by convolving the stimulus either with a pair of setup HRTF or HARTF using the real-time auralization software Virtual Acoustics, which has been developed in-house. Both data sets are measured at a radius of $r = 1.86\,\mathrm{m}$ from a dummy head [6] using a measurement arm equipped with a broadband loudspeaker in 1° x 1° spatial resolution covering the full sphere in azimuth and zenith angles. The measured data sets have a filter length of 256 samples. The HARTF data set is measured at the microphone positions of a behind-the-ear receiver-in-the-ear RHA with open dome fitting. Although being equipped with two omnidirectional microphones on each ear side, only the measurements of one microphone pair is used for the generation of VSS.
In the loudspeaker-based binaural reproduction path, the CTC filters are realized as N-CTC filters, with N = 4 loudspeakers, all playing simultaneously with additional loudspeaker free-field equalization filters [2].

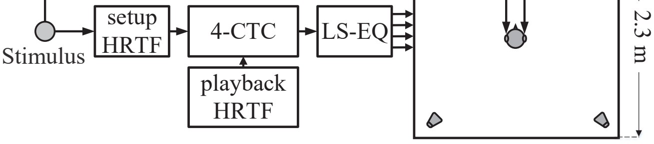

Figure 1: Extended binaural reproduction setup for hearing aid applications consisting of a hearing aid-based and a loudspeaker-based playback path with crosstalk-cancellation filter network and loudspeaker free-field equalization to create virtual sound sources. The system is installed in an acoustically optimized hearing booth.

A head-tracking system is included to account for user movements. Any translatory or rotational head movement influences the selection of setup and playback HRTF and/or HARTF as well as the calculation of CTC filters and results in a virtual scene update in real-time. For a correct auralization, the offset of the rigid body, mounted on top of the head, to the center of the interaural axis is corrected individually. A tracker frame rate of 60 Hz is used which leads to a position update approximately every 16.7 ms, neglecting other factors that influence end-to-end latency.

Compared systems

In the current study, five different conditions, tested in two different experiments, are compared:

[LS]	RSS played back over loudspeakers,
[HP]	VSS played back over headphones,
[CTC]	VSS played back over loudspeakers with CTC filters,
[RHA]	VSS played back over RHA, and
[CTCwRHA]	VSS played back over loudspeakers with CTC filters and RHA.

Experiment 1 (E1) is conducted in an anechoic chamber and investigates localization performance in [LS] and [HP] [7]. Experiment 2 (E2) is conducted in a hearing booth assessing [CTC], [RHA] and [CTCwRHA] [8]. To compensate the influence of the headphone frequency response on the localization accuracy in [HP], individual headphone-related transfer functions have been measured and applied as inverse filters prior to playback. The playback levels have been adapted to produce the same binaural loudness in all conditions.

Reproduction delay

In conditions [RHA] and [CTCwRHA] of E2, the RHA signals are time-delayed by $\Delta t \approx 7\,\mathrm{ms}$, relatively to the reproduction in [CTC], to incorporate the latency in a conventional HA [9]. The measured end-to-end latency of both binaural reproduction paths in [CTCwRHA] is shown in Figure 2.

Spectral influence of the ear piece

In the extended binaural reproduction system [CTCwRHA], the RHA with open dome fitting considerably affects the external soundfield for frequencies above $f \approx 1\,\mathrm{kHz}$, as shown in Figure 3. Above 2 kHz, the damping effect reaches its maximum of $-12\,\mathrm{dB}$ at 9 kHz, followed by a damping minimum at 10 kHz, to return to a damping of $-5\,\mathrm{dB}$ at frequencies above.

Procedure

In a sound localization task, the perceived localization of 12 VSS arranged in steps of $\pi/4$ (starting at $\varphi = 0$) in the horizontal plane and four VSS placed in the median plane at $\varphi = 0$ and π with zenith angles $\vartheta = \pi/3$ and $2\pi/3$ at a distance of 2 m is tested [8]. An exocentric pointing method is used where the subject moves a cross-hair to mark the perceived VSS position on a sphere which is displayed on a screen in front of the subject [11].

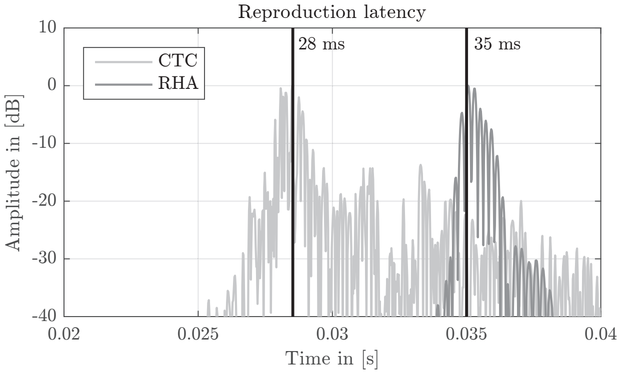

Figure 2: System reproduction latency. Normalized impulse responses of both binaural reproduction systems in [CTCwRHA] measured with a dummy head (left ear's impulse responses) using a RME Fireface UC with 128 samples buffer size at a sampling rate of 44.1 kHz with i/o latency compensation. To obtain absolute latency values, the measured latency has to be corrected by the sound propagation delay due to a virtual sound source distance of 2 m, located at $\varphi = 0^\circ$ in the horizontal plane.

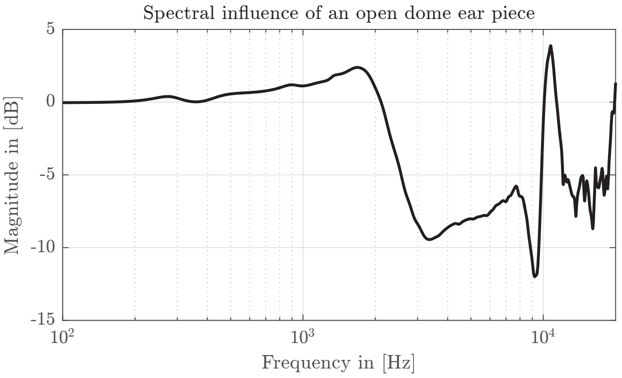

Figure 3: Influence of an open dome ear piece on the perceived magnitude spectrum. The plot shows the magnitude difference in transfer functions between occluded and unoccluded ear, measured from a Neumann KH 120 A placed at $\varphi = \pi/4$ in the horizontal plane to an ear simulator acc. IEC 60318-4 (2010-01) built into a dummy head [10].

A training session is conducted before each test condition in which ten VSS positions are tested, the presented VSS position being additionally marked with a red cross. In the test session, all VSS positions are queried counterbalanced using a Latin Square design and are tested three times each, which leads to a total of 180 trials per subject for all conditions. Subjects were allowed to repeat each stimulus two times per trial.

Subjects

Nine female and six male, normal-hearing adults at the age of 24 ± 5.4 (mean ± std) participated in both experiments in two different sessions on two different days. All subjects received a representation allowance for their participation.

Results

Error measures

To evaluate the localization performance, three different angular error measures are used: the azimuth error ϵ_φ, the zenith error ϵ_ϑ, and the overall angular error ϵ_γ. Each error measure shows the angular deviation between the presented and the perceived VSS position with respect to the particular angular component or a vector error [11].

Evaluation

The tested null hypothesis assumes that the reproduction systems used in E2 allow the same localization accuracy as the reproduction systems used in E1, i.e. the error angles in E2 do not significantly differ from the error angles in E1 with respect to their median values.
A Shapiro-Wilk test shows that the measured data does not come from a normal distribution, therefore, a Friedman test is applied to review the data for significances, with given Kendall's W. For post-hoc analyses, a Wilcoxon signed rank test with Bonferroni correction is applied.[1] In the following, all differences are referred to [HP], where the indices of the Deltas describe the compared conditions, e.g. Δ_{24} symbolizes the comparison between the median error angles in [HP] and [RHA]. To investigate the effect of [CTC] on localization performance in [CTCwRHA], the comparison Δ_{45} is drawn additionally.

Azimuth error

A Friedman test shows a significant influence of the reproduction systems on median azimuth error ϵ_φ, $\chi^2(4) = 140.84, p < .001, W = .07$. In Figure 4, ϵ_φ is plotted over all conditions. Planned comparisons show no significant differences for Δ_{23}, $Z = -.146, p = .884, r = .004$, and Δ_{25}, $Z = -2.567, p = .01, r = .08$. A significant difference exists for Δ_{12}, $Z = -5.884, p < .001, r = .18$, Δ_{24}, $Z = -7.808, p < .001, r = .08$, and Δ_{45}, $Z = -5.904, p < .001, r = .18$.

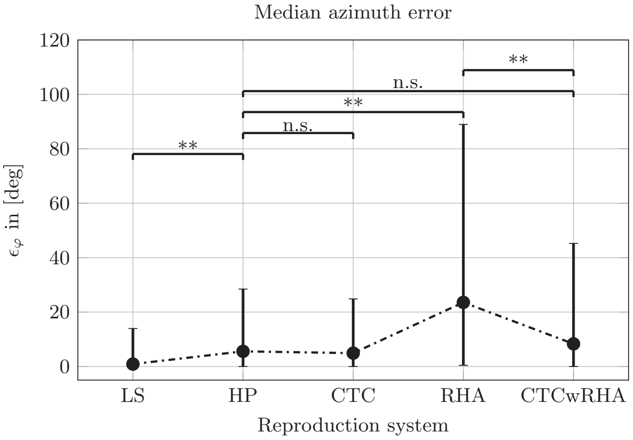

Figure 4: Median azimuth error ϵ_φ with interquartile range, plotted over all tested reproduction systems.

Zenith error

A Friedman test shows a significant influence of the reproduction systems on median zenith error ϵ_ϑ, $\chi^2(4) = 248.44, p < .001, W = .12$. In Figure 5, ϵ_ϑ is plotted over all conditions. Planned comparisons show no significant differences for Δ_{45}, $Z = -.841, p = .4, r = .03$. A significant difference exists for Δ_{12}, $Z = -4.060, p < .001, r = .12$, Δ_{23}, $Z = -7.730, p < .001, r = .24$, Δ_{24}, $Z = -7.263, p < .001, r = .22$, and Δ_{25}, $Z = -8.371, p < .001, r = .25$.

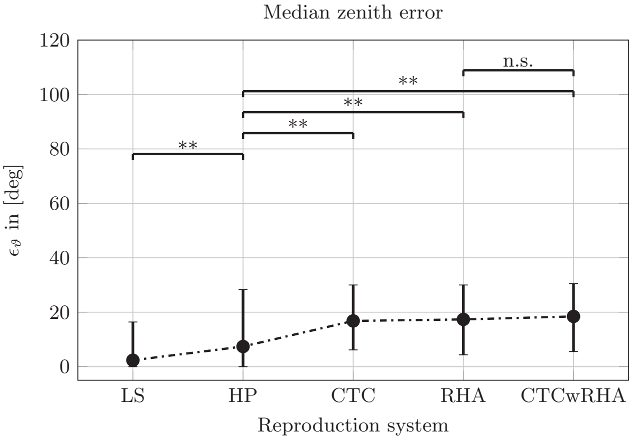

Figure 5: Median zenith error ϵ_ϑ with interquartile range, plotted over all tested reproduction systems.

Overall angular error

A Friedman test shows a significant influence of the reproduction systems on median overall angular error ϵ_γ, $\chi^2(4) = 302.29, p < .001, W = .14$. In Figure 6, ϵ_γ is plotted over all conditions. Planned comparisons show significant differences for Δ_{12}, $Z = -7.453, p < .001, r = .23$, and Δ_{23}, $Z = -3.380, p = .001, r = .10$, Δ_{24}, $Z = -9.453, p < .001, r = .29$, Δ_{25}, $Z = -6.078, p < .001, r = .18$, and Δ_{45} $Z = -3.989, p < .001, r = .12$.

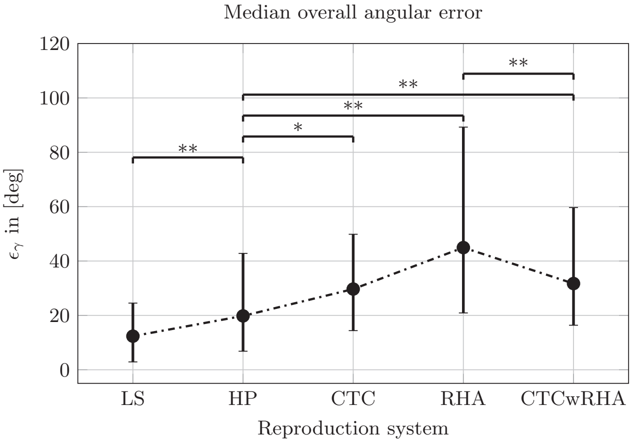

Figure 6: Median overall angular error ϵ_γ with interquartile range, plotted over all tested reproduction systems.

---

[1] This results in a corrected $p < .05/5 < .01$. Non-significant differences with $p > .01$ are marked with (n.s.), asterisks denote significant differences with $p < .01$ (∗) and with $p < .001$ (∗∗).

Discussion

The results of all planned comparisons for the binaural reproduction systems in E2, given [HP] as reference, are shown in Table 1. For [CTC] and [CTCwRHA], no significant difference was found in terms of azimuth error ϵ_φ, at BONFERRONI-corrected $p < .01$. All other comparisons were significantly different.

Table 1: Summary of planned comparisons in Experiment 2 with headphone-based binaural reproduction system [HP] as reference.

Error	Condition		
	CTC	DHA	CTCwRHA
ϵ_φ	n.s.	**	n.s.
ϵ_ϑ	**	**	**
ϵ_γ	*	**	**

The significant drop in localization performance in [DHA] has been expected and can be traced back to the fact that the simulation of the VSS is based on a pair of HARTF measured by one omnidirectional RHA microphone behind the respective ear. Due to the position of the RHA microphones, the binaural cues, interaural time and level differences, differ from conventional HRTF cues and especially pinna-related cues are missing in HARTF [10]. Additionally, the frequency range of the HA receiver signals is limited and considerably spectrally modified because of the the open dome fitting and the receiver transfer function, respectively [10]. These deteriorations add up and lead to worse localization performance, although being significantly improved in [CTCwRHA] with respect to ϵ_φ and ϵ_γ, cf. Figures 4 and 6. However, as the RHA are mainly utilized to improve audibility in upcoming experiments, this localization degradation is not too problematical.

Conclusion

An experiment to assess the localization performance in an extended binaural reproduction system designed for hearing aid users was conducted with normal-hearing adults. Compared to a headphone-based binaural reproduction of virtual sound sources, the loudspeaker-based binaural reproduction with acoustic crosstalk-cancellation filters, without and with additional research hearing aids, showed a comparable localization performance with respect to azimuth error. Although zenith and overall angular localization errors were found to be significantly larger, the presented reproduction approach will result in acceptable to good localization results and is thus applicable for auralization purposes.

Further investigations have to be carried out involving hearing aid users to get more practical results. In this context, the comparison between sound localization in the presented reproduction system and the localization of real sound sources with their own hearing aids is especially interesting.

Acknowledgements

The authors would like to thank Jan Gerrit Richter and Suliang Wang for providing the data of their localization experiment. This work received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no. ITN FP7-607139: Improving Children's Auditory Rehabilitation (iCARE).

References

[1] Bronkhorst, A. W.: Localization of real and virtual sound sources. The Journal of the Acoustical Society of America (1995), 98, 2542-2553.

[2] Masiero, B.: Individualized Binaural Technology. PhD Thesis (2012), Institute of Technical Acoustics, RWTH Aachen University.

[3] Parodi, Y. L., Rubak, P.: A Subjective Evaluation of the Minimum Channel Separation for Reproducing Binaural Signals over Loudspeakers. The Journal of the Audio Engineering Society 59 (2011), 487-497.

[4] Pausch, F. et al.: Evaluation of binaural sound reproduction systems with focus on perceptual quality. In Fortschritte der Akustik: DAGA 2015, 41. Jahrestagung für Akustik, Nuremberg.

[5] Mueller, M. F. et al.: Localization of virtual sound sources with bilateral hearing aids in realistic acoustical scenes. The Journal of the Acoustical Society of America (2012), 131, 4732-4742.

[6] Schmitz, A.: Ein neues digitales Kunstkopfmeßsystem. Acta Acustica united with Acustica 81.4 (1995), 416-420.

[7] Wang, S.: Investigation of the influence of small head movement on the localization ability of human listeners. Master Thesis (2015), Institute of Technical Acoustics, RWTH Aachen University.

[8] Stockmann, A.: Investigation of realistic acoustic scene simulation for people with hearing impairment. Master Thesis (2015), Institute of Technical Acoustics, RWTH Aachen University.

[9] Stone, M. A. et al.: Tolerable Hearing Aid Delays. V. Estimation of Limits for Open Canal Fittings. Ear Hear (2008), 29(4):601-17.

[10] Tumbrägel, J.: Hearing aid signal processing for a real-time auralization system. Master Thesis (2015), Institute of Technical Acoustics, RWTH Aachen University.

[11] Friedrich, A.: Psychoacoustical validation of fast individual HRTF measurements. Master Thesis (2014), Institute of Technical Acoustics, RWTH Aachen University.

[12] Rosenthal, R.: Meta-analytic procedures for social research. CA: Sage, Second Edition, Newbury Park, 1991.