Psychoacoustic evaluation of a fast HRTF measurement system

Jan-Gerrit Richter, Janina Fels
Institute of Technical Acoustics, Medical Acoustics Group
52074 Aachen, Deutschland, Email: jan.richter@akustik.rwth-aachen.de

Introduction

To accurately simulate sound sources from arbitrary directions, a stimuli can be convoluted with the head-related transfer function (HRTF) of the desired direction [1]. The HRTF represents the influence of the listeners head on the sound field. A convolution of a stimuli with a HRTF will reproduce this same effect in a virtual scene. To simulate sound sources in a room, a HRTF with sufficiently high resolution is needed as reflections will arrive at the subjects ears from every directions. To reduce measurement effort so called artificial heads can be constructed which have a HRTF that fits reasonably good on many listeners. The perceived sound direction of virtual sources auralized with dummy-head HRTF will vary in between subjects depending on the geometric properties of the subjects head. The use of individual HRTFs always improves the subjects localization ability of the sources [2] and are thus needed for simulations with high accuracy and in simulations where realism of the simulation is an important factor. To measure individual HRTFs with high resolution, fast measurement techniques need to be applied to reduce measurement time. These techniques often result in a trade-of between measurement accuracy and measurement time.

In this paper a psychoacoustic evaluation of HRTFs measured with a fast HRTF measurement system is presented. Three localization listening experiments are presented that will investigate the influence of such a system on localization accuracy of virtual static sound sources. In the first test, the influence of the pointing device is studied. The second test is used to investigate the influence of the measurement system. The third experiment is used to determine if individual measurements done with a fast system are still beneficial with respect to localization ability even though the measurement might be distorted.

Fast HRTF Measurement System

Traditional HRTF measurement as described by Blauert [1] uses a single loudspeaker that will measure the transfer path to both ear canals. As a reference to the measurement, the measurement is divided by the transferpath to the microphones in the middle of the head, without the head present. This process has to be repeated for every desired direction.

In this paper, a fast measurement system is used that is constructed as an arc around the subject. This arc holds 40 loudspeaker with a diameter of 32mm on different elevation planes [7]. The subject stands on a turntable and is rotated 360°. This setup will allow a gaussian sampling grid with maximum order of 47 with a missing 30° gap at the bottom. Figure 1 shows the measurement arc setup in an semi-anechoic chamber. The 40 loud-

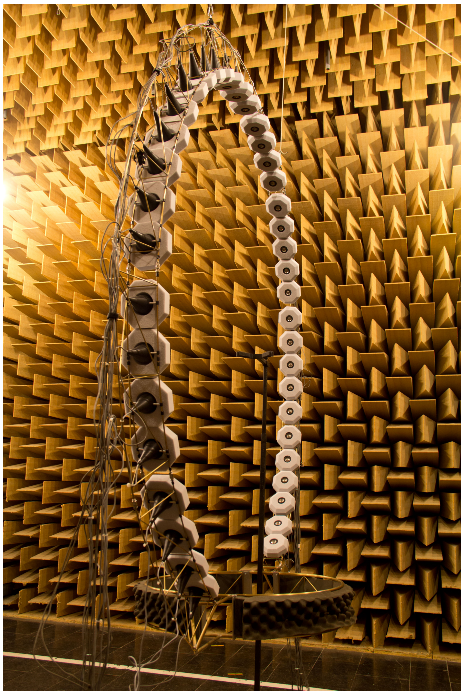


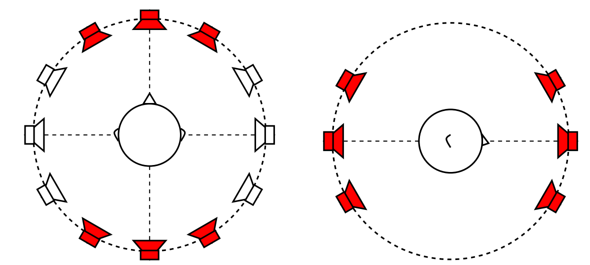
Figure 1: The measurement setup in the semi-anechoic chamber

speaker can measure the HRTFs almost simultaneously with the use of optimized interleaved sweeps [3] which is an extension of the multiple exponential sweep method [4]. The measurement time for 12 azimuth levels that are used in the listening experiment at 40 elevation levels is approximately two minutes.

This setup will introduce several uncertainty factors into the measurement. One factor comes from the setup itself. The added loudspeakers will cause additional reflections and distortion in the sound field. The second factor are uncertainties about the user position, orientation and movement during the measurement. The aim of the listening experiments are to determine if the added uncertainty have a visible influence on the localization accuracy.

Listening Experiment Design

To test the influence of the measurement system on the localization ability, three separate listening experiments where conducted. Localization accuracy has been widely



(a) Virtual source position in horizontal plane **(b)** Virtual source position in elevation planes

Figure 2: Position for virtual sound sources. Red loudspeakers are also tested in shown elevation levels

tested. The location accuracy varys from 1-5° [1] up to 15-25° [8] depending on different factors in the measurement setup.

In the performed listening experiments the subjects are instructed to localize a virtual sound source that is located on 24 points on a sphere around the subject. The subject does not have any additional knowledge about the source origin and can freely point to any point on a full sphere. In the horizontal plane, twelve source positions are placed with 30° distance. The other twelve positions are in 30° and -30° elevation on points in front and in the back of the subject. Figure 2 shows the positions.

The virtual scene is static as it does not consider the listeners head position and movement. As this type of localization task can be very hard a stimuli is selected as pulsed white noise bursts to help localization. The stimuli are played over headphones. To account for individual differences in the headphone fit, the headphone transfer functions are individually measured and equalized [6].

The pointing device potentially has a big influence on the indicated position. Two different pointing devices are compared. The first device uses a gamepad to indicate azimuth and elevation direction. A sphere and a crosshair is displayed on a screen in front of the subject. The second device is a head-mounted display that will display a surrounding sphere to the subject. The subject can point to a direction by looking at it. This also allows pointing to a full sphere. It was shown that such a pointing device allows for more accuracy during localization tests [5].

Test Subjects

Three different subject groups are used to compare localization accuracy with different HRTFs in this paper. Overall, 58 subjects participated in the listening experiments. Subject group one is comprised of 22 subject (11 male, 11 female). This group participated in the comparison between two pointing devices. Subject group two comprises 20 subjects (9 male, 11 female). This group repeated the experiment with a dummy head HRTF from a traditional measurement system. This data was used to compare the same HRTF from two different measurement setups to investigate the influence of the system. The third subject group comprises 16 subjects (12 male, 4 female). From the subjects in this group individual HRTF measurements are taken and the localization test is repeated with these HRTFs. From this data an in-between subject comparison can be made about the influence of individual HRTF measurement on localization accuracy. Table 1 summarises the groups and tests more concise.

Table 1: Visualization of the three subject groups and the four combination tested. Arc and arm are indicators for the measurement system with arc as the fast system. Joystick and Goggles are the two pointing devices used.

	Exp. 1	Exp. 2	Exp. 3
dummy head, arc, joystick	Group 1		
dummy head, arc, goggles	Group 1		
dummy head, arm, joystick		Group 2	
individual, arc, joystick			Group 3

Results

To analyse the error three different error measures are calculated between each stimuli point and the point indicated by the subject. The azimuth and elevation difference measures calculate the difference only in azimuth and elevation planes respectively. To combine both measures, the central angle between the two points is calculated.

Figure 3 shows all three error measures over all subjects and positions, for the comparison between the two pointing devices. Table 2 shows mean and standard deviation values. Although the overall error values are large there is only a small difference between the two systems. A multivariance analysis shows no significant differences between the two systems.

Table 2: Mean and standard deviation values of dummy head listening for two pointing methods.

		Joystick	HMD
Azimuth	Mean	18.28 °	18.35 °
	Std	18.67 °	17.98 °
Elevation	Mean	18.99 °	17.69 °
	Std	16.29 °	15.89 °
Central angle	Mean	27.27 °	26.42 °
	Std	18.35 °	18.10 °

As multiple subjects reported dizziness while using the head mounted display, the joystick pointing method was used for the remaining tests which will also offer the possibility of comparing to real sources in the future.

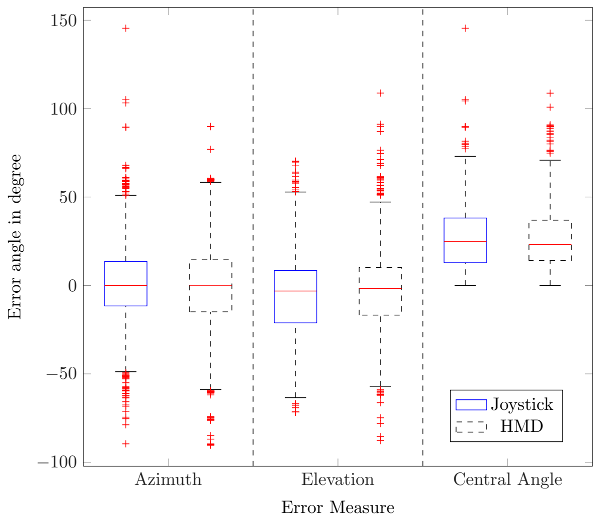


Figure 3: Comparison of two pointing method. All three error measures show differences in localization accuracy for every direction and subject.

Subject group two repeated the same listening experiment with an HRTF from a dummy-head measured with a traditional system. An in-between subject comparison indicates the influence of the measurement system on the localization accuracy. Figure 4 shows all three error measures for the comparison between the measurement setups. Again, no significant differences between the two systems could be found by a multivariance analysis. Any differences caused by the system is most likely too small to register as the overall uncertainty is too large.

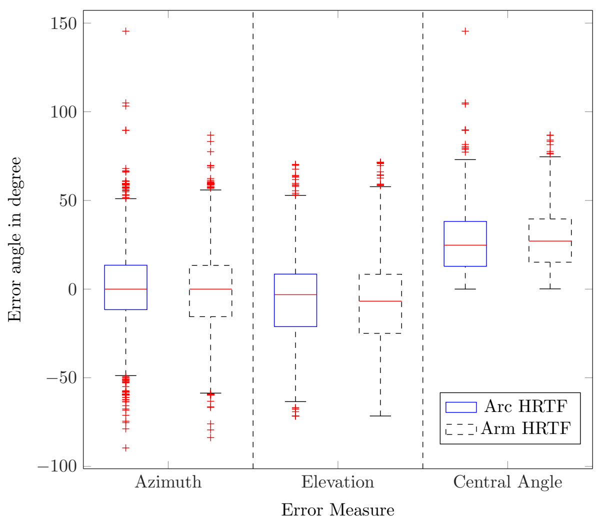


Figure 4: Comparison of two HRTF measurement setups. All three error measures show differences in localization accuracy for every direction and subject.

Subject group 3 participated in the experiment with individually measured HRTFs. The data is compared to results from subject group one with a dummy-head HRTF measured in the fast system. This in-between subject comparison is done to investigate if the added distortions will harm the benefit from individually measured HRTF with respect to localization accuracy. Figure 5 and Table 3 show the mean and standard deviation error for both the individual and the dummy-head HRTF. A clear increase in accuracy is visible in both azimuth and elevation. This is supported by a variance analysis which shows significant differences (Azimuth: $F(1, 34) = 7.40$, $p < 0.05$, $\eta^2 = 0.18$; Elevation: $F(1, 34) = 13.39$, $p < 0.05$, $\eta^2 = 0.29$). This result shows that although the HRTF might be affected from the measurement setup and uncertainties the use of individual HRTF still has benefits in localization experiments.

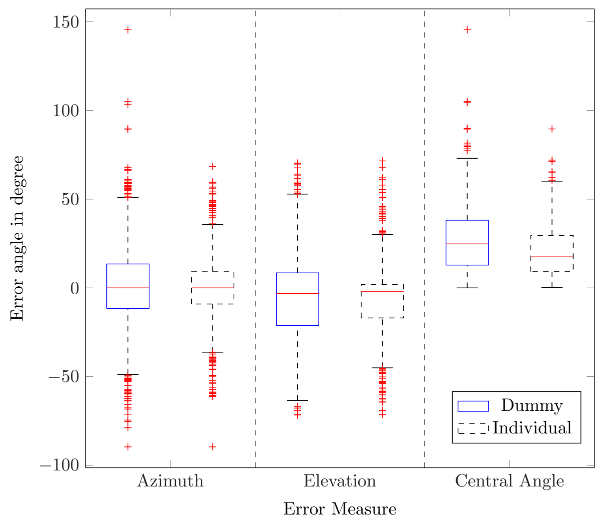


Figure 5: Comparison of individual to dummy head HRTFs. All three error measures show differences in localization accuracy for every direction and subject.

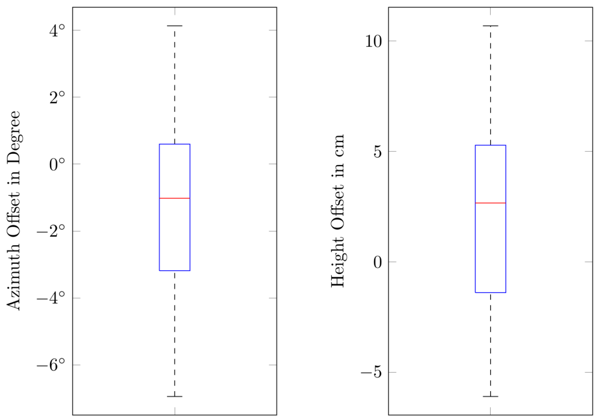
Table 3: Mean and standard deviation values for dummy-head HRTFs and individual HRTFs.

		Dummy-Head	Individual
Azimuth	Mean	18.28 °	12.86 °
	Std	18.67 °	14.42 °
Elevation	Mean	18.99 °	13.93 °
	Std	16.29 °	13.94 °
Central angle	Mean	27.27 °	20.62 °
	Std	18.35 °	14.98 °

Subject Positioning

One potentially negative influence on the individual localization results can be the subject positioning during measurement. Although the subject is positioned in the arc using a three axis laser, uncertainties remain about the influence of the position durin the measurement. A non ideal subject position will influence the measured HRTF in direction and distance. As the directivity of the measurement system is not uniform, even frequency dependent influences are possible.

To analyse the subject positioning inside the arc, two different parameters are calculated. The first parameter



(a) Estimated azimuth offset in degree calculated over 8 subjects.

(b) Estimated height offset in cm calculated over 8 subjects.

Figure 6: Estimated user position uncertainties during fast HRTF measurement.

is the height offset of the subject inside the measurement arc. As each measurement comprises 40 measurements around the subject, this parameter can be calculated using time of arrival information from the 40 impulse responses for each ear. The second parameter describes the offset of the subjects head in azimuth direction. As this parameter is not easily fitted from one measurement, a fit using the Interaural Level Difference is used. Over one azimuth rotation, the ITD approximately will follow a sinusoid form. An approximation of a sinusoid gives information about the approximate offset of the subject during the measurement.

Figure 6 shows both height and azimuth offset calculated for 8 subjects from Group 3. The Figure shows a relatively large deviation in both parameters. The height varies between -5 and 10 cm while the offset varies between -6° and 4°. The data from each individual offset with both parameters are then used to correct the measured HRTF positions. This is done to achieve more realistic error measures as each subject will be presented with stimuli from slightly different directions. A multi-variance analysis again shows no significant differences between accuracy with this correction and uncorrected positions as the achieved improvement is marginal. The effect may again be to small to be of influence on global localization accuracy.

Conclusion

In this paper three localization listening experiment are presented that analyse the influence of a fast HRTF measurement system on global localization performance. The fast measurement system introduces distortions and uncertainties about the subject into the measurement. These influences are studied in detail. In the first listening experiment a suitable pointing device for the tests is selected. No significant differences could be found between a gamepad and a head-mounted display. The gamepad is selected for the other tests for practical reasons. The second test compares the influence of the measurement setup itself as both HRTFs stem from the same dummy head measured with a fast and a traditional setup. Again, no significant differences have been found. The last test compares localization ability from dummy head HRTFs to individual HRTFs from a fast measurement system. Here a significant improvement in accuracy has been found which signifies that potentially negative influences from the measurement system are not too large to mask the improved localization accuracy from individual HRTFs.

Outlook

To further analyze the influence of the fast measurement system on the quality of the measured HRTF several other listening experiments are needed. The localization has to be studied in more detail, including dynamic scenes which will allow small head movements as they should increase the localization accuracy. A further step should be the auralization of moving sources that should show differences in the influence of the system more clearly. Other aspects of the perception of sound, for example coloration of the sound, have to be evaluated as well.

References

[1] J. Blauert. *Spatial hearing: The psychophysics of human sound localization.* MIT Press Cambridge, 2nd edition, 1997.

[2] A. Bronkhorst. Localization of real and virtual sound sources. *The Journal of the Acoustical Society of America*, 98(November 1995):2542–2553, 1995.

[3] P. Dietrich, B. Masiero, and M. Vorländer. On the optimization of the multiple exponential sweep method. *J. Audio Eng. Soc*, 61(3):113–124, 2013.

[4] P. Majdak, P. Balazs, and B. Laback. Multiple exponential sweep method for fast measurement of head-related transfer functions. *J. Audio Eng. Soc*, 55(7/8):623—-637, 2007.

[5] P. Majdak and B. Laback. The accuracy of localizing virtual sound sources: Effects of pointing method and visual environment. *Audio Engineering Society Convention 124*, 2008.

[6] B. Masiero and J. Fels. Perceptually robust headphone equalization for binaural reproduction. *Audio Engineering Society Convention 130*, pages 1–7, 2011.

[7] M. Pollow, B. Masiero, P. Dietrich, J. Fels, and M. Vorländer. Fast measurement system for spatially continuous individual HRTFs. In *4th Int. Symposium on Ambisonics and Spherical Acoustics*, University of York, UK, 2012.

[8] F. L. Wightman and D. J. Kistler. Headphone simulation of free-field listening. II: Psychophysical validation. *The Journal of the Acoustical Society of America*, 85(2):868—-878, Feb. 1989.