Development and Validation of a Sound Source for Near-field HRTF Measurements

Yuqing Li, Stephan Preihs, Jürgen Peissig
Institut für Kommunikationstechnik (IKT), Leibniz Universität Hannover, Germany
Email: yuqing.li@ikt.uni-hannover.de

Background

Head-Related Transfer Functions (HRTFs) are widely used for binaural reproduction of spatial sound in virtual auditory displays. HRTFs describe the acoustic effects, i.e. sound reflection, diffraction, and head shadowing, caused by our head, torso, and pinnae, and they inherently contain the key binaural and monaural cues for human listeners to localize spatial sound.

HRTFs are functions of the relative position of the sound source and the listener. When the sound source is far away, the HRTF spectrum variation with the source-receiver distance is neglectable, and the name "far-field HRTF" applies in such circumstances. As the source-listener distance decreases, head shadowing becomes prominent. Usually, when the source-receiver distance is smaller than 1m, some features of HRTFs, especially the low-frequency Interaural Level Difference (ILD), vary with distance [1], and thus have an impact on sound localization. These HRTFs are defined as near-field HRTFs.

The EU-funded project, Virtual Reality Audio for Cyber Environments (VRACE) [2], aims at raising VR audio to a next level and enabling a true world simulation for virtual testing and engineering, with research into physically correct representation of audio events in virtual reality. The acquisition of near-field HRTFs is an important research aspect for the VRACE project, as binaural rendering of nearby sound sources is one of the major project topics.

HRTFs can be measured with a pair of miniature microphones placed at the opening of the ear canals of a dummy head or a human subject. Due to the spatial variation of near-field HRTF spectra, measurements need to be taken at multiple distances as well as directions. With spatial interpolation and extrapolation, it is then possible to represent virtual sound sources at desired positions.

The ideal sound source for measuring near-field HRTFs should be an infinitesimally small, monopole point source, which doesn't exist in real life. Therefore, an alternative must be found. Considering the impact of the sound source's physical and acoustical properties on measurement accuracy, some specific features should be aimed for. The selected sound source should be small enough to avoid multiple reflections between the sound source and the subject's head as the source-receiver distance becomes small. It should be omni-directional so that an equal amount of energy is emitted to all spatial directions. It also should be as full-range as possible, for capturing enough spectral information in the HRTF.

Some sound sources that have been used in near-field HRTF measurements are displayed in Figure 1. They suffer from different drawbacks. In particular, the low-frequency behavior of the probe source and the dodecahedrons are limited by the small size of their drivers. Even for the latest improved design in Figure 1(d), the radiated power below 200Hz is very poor. Another drawback of the dodecahedron loudspeakers is that they get fairly directional at high frequencies due to the interference between the individual drivers.

Therefore, we aim at designing a sound source which is small enough in size but can provide sufficient low-frequency power, for accurately measuring near-field HRTFs.

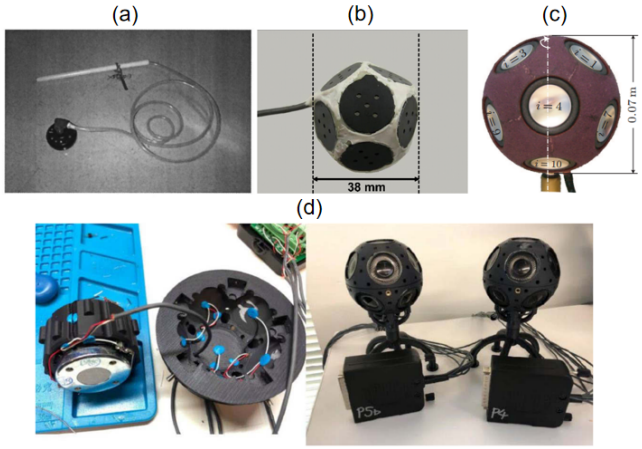

Figure 1: Existing sound sources for near-field HRTF measurement. (a) Probe sound source [1], (b) micro dodecahedron [4], (c) spherical dodecahedron [5], (d) improved spherical dodecahedron [6].

Design and Construction

The primary limitation of decreasing the source size is that the smaller the driver's diaphragm, the less low-frequency power it can radiate. According to [3], the radiated sound power from a driver is proportional to its surface area, the diaphragm displacement, and the wave number. However, small drivers are usually not capable of making large excursion distances.

Therefore, we propose a loudspeaker-tube system as a solution, which is a modified version of the probe source in [1] as shown in Figure 1(a). The sound source in [1] consists of a long tube connected to a small electrodynamic horn driver, so that the relatively small opening of the tube acts like an acoustic point source, and theoretically it is relatively nondirectional even at high frequencies, although the directivity of this sound source was not reported. In this work, we utilize a larger loud-

speaker driver instead for producing more low-frequency sound energy, and we seal it in an enclosure for higher efficiency.

The designed sound source has three main components:

- a driver sealed in a cylindrical enclosure,
- a reduction cone placed on top of the driver to smoothly reduce the diameter of sound propagation path,
- a 3 m-long PVC tube with a diameter of 1.2 cm, sound is emitted at the tube's nozzle.

In our work, the driver's enclosure and the reduction cone are 3D-printed, the material used is PLA.

When pursuing optimal performance of the driver, the back volume of the enclosure needs to be carefully calculated with the aid of some of the driver's Thiele-Small parameters, which are available from the product sheet. Thiele-Small parameters are a set of electromechanical parameters that can be generally used to predict the low-frequency behavior of closed-box loudspeakers [7]. For our design, the key parameter obtained from the calculation is the back volume V_B, which then determines the resonance frequency f_c of the driver-enclosure system. By controlling the overall quality factor Q_{TC} of the designed enclosure, an appropriate V_B can be deduced. The quality factor Q implies the amount of damping. A smaller Q means more damping and thus a higher resonance frequency, and a large Q would make the system underdamped and thus a resonance peak appears. Therefore, a balance must be found. It has been shown that when $Q_{TC} = \frac{1}{\sqrt{2}}$, the loudspeaker produces the flattest response [7]. Therefore, a Q_{TC} of 0.71 was chosen, and the required enclosure volume was deduced according to Equations 1 and 2. The relevant parameters are shown in Table 1.

Thiele-Small Parameters of the Driver	Deduced Parameters of the Closed-box Loudspeaker
f_s: open air resonance	f_c: installation resonance
V_{AS}: equivalent volume	V_B: volume enclosed in the housing
Q_{TS}: overall quality of the driver	Q_{TC}: overall quality of the loudspeaker

Table 1: Design parameters

$$f_c = f_s \sqrt{\frac{V_{AS}}{V_B} + 1} \tag{1}$$

$$Q_{TC} = \frac{f_c}{f_s} Q_{TS} \tag{2}$$

The inner diameter of the reduction cone is reduced exponentially from 10 cm to 1 cm.

Figure 2 shows the dimensions and frequency response of the selected driver. It is full range, with a very low resonance frequency of 75Hz. The CAD model of the designed enclosure and the cone are shown on the left-hand side in Figure 3, with dimensions specified. The upper interior of the enclosure includes a section of smooth radius transition for manufacturing convenience. The cone, the driver, and the enclosure were screwed together (the assembly is referred to as "the loudspeaker part" from now on), and the PVC tube was connected to the small mouth of the reduction cone. On the right in Figure 3 is a photo of the loudspeaker part of the constructed sound source.

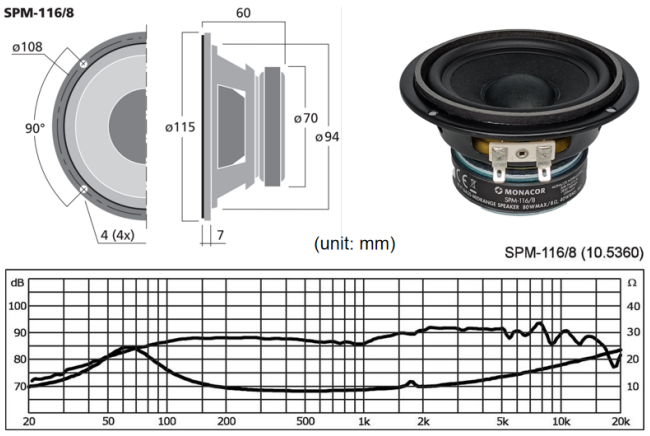

Figure 2: Selected driver [8].

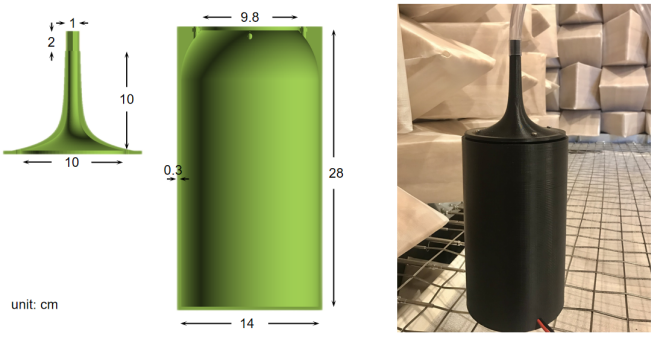

Figure 3: The CAD model and photo of the driver's enclosure and the reduction cone.

Evaluation on the Sound Source

Frequency Response

The free-field frequency response of the developed sound source was measured in the anechoic chamber of IKT, Leibniz Universität Hannover. It was excited with a 5 s log sine-sweep from 20 Hz to 20 kHz, which first fed through and amplified by an SMSL SA-36A Pro Power Amplifier. The distance between the tube nozzle and the microphone was set to 1 m. A 3D-printed sleeve was used to provide support to the end of the tube, which wa fixed by clamps to a tripod arm (Figure 4). To avoid direct sound radiated from the back of the diver to reach the measurement microphone, the loudspeaker part was placed in a storage box, and the gap between the box and the loudspeaker was filled by sound-absorbing materials. To improve signal-to-noise ratio (SNR), measurements were repeated 10 times and an average was taken. The

resulted frequency-independent SNR was 49dB. The Total Harmonic Distortion (THD) up to the 5th harmonic was 2.05%.

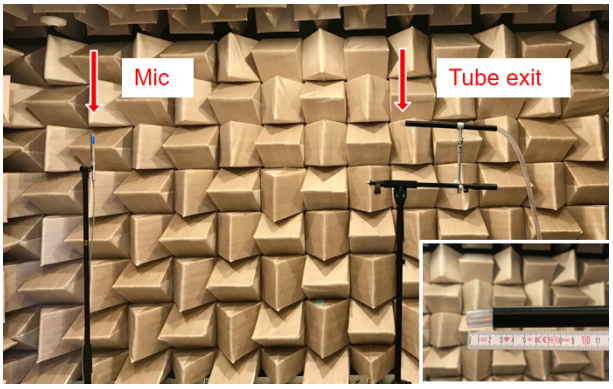

Figure 4: Measurement setup for sound source evaluation.

The measurement results in time and frequency domain are shown in Figure 5. The most prominent feature in the impulse response plot is that there are subsequent peaks after the first one. The intervals between each peak are approximately equal to 18 ms, which corresponds to the time needed for traveling twice from the driver's diaphragm to the opening of the tube, because the subsequent peaks are caused by standing waves in the tube. If desired, this effect can be diminished by inserting some damping material to the end of the tube. These reflections introduce small fluctuations in the frequency response, but generally, energy is concentrated at low to mid frequencies around 500 Hz, and a sufficient SNR is retained down to about 100 Hz. Above 500 Hz, the magnitude drops at an approximately 10 dB/octave rate to about 2 kHz. It then keeps a relatively constant level except for 2 valleys at 3.6 kHz and 7.2 kHz, respectively. For frequencies over 12 kHz, the magnitude response becomes poor.

Directivity

Off-axis responses of the sound source were measured on the horizontal plane every 5° from −90° to 90°, with the aid of a turntable controlled by a MATLAB program. Figure 6 shows the polar magnitude response at different frequencies. The sound source can be regarded as omnidirectional up to 2000 Hz, with a maximum magnitude deviation of about 0.5 dB compared to the on-axis response. For frequencies up to 8000 Hz, the magnitude deviation is within 3 dB. Up to 13 kHz, the sound source has a 3 dB span of 110°.

Validation

To validate the capability of the developed sound source for near-field HRTF acquisition, measurements were carried out on the G.R.A.S KEMAR with a pair of Sound Professionals in-ear binaural microphones, for 2 source distances, 50 cm and 1 m, at 4 different azimuthal directions, 0°, 30°, 60°, and 90°.

Expected features appear in the results as shown in Figure 7 and 8, for example, the increase in magnitude at the ipsilateral ear for a lateral source compared to a frontal

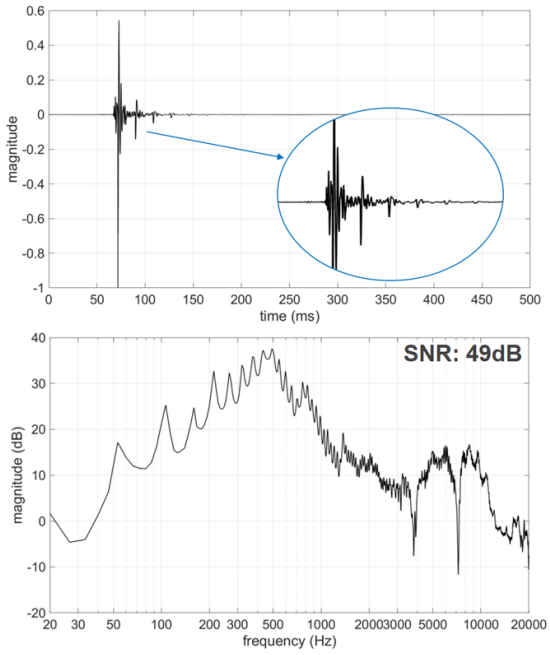

Figure 5: Impulse response and frequency response.

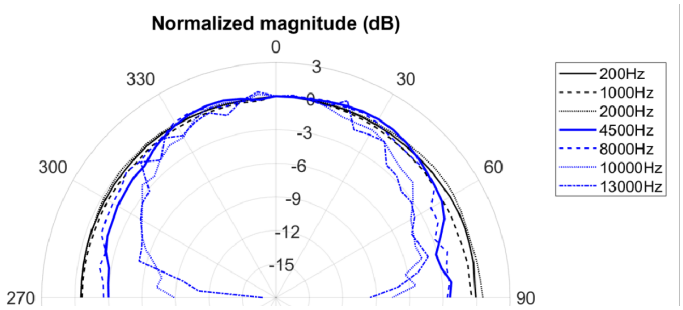

Figure 6: Polar magnitude plot.

source, and the decrease in magnitude for the contralateral ear at the same time. Particularly, the change in magnitude at the contralateral ear with the azimuth angle is slightly larger when the source is at 50 cm than 1 m, as a result of enhanced head shadowing for a nearby sound source.

ILDs were calculated by subtracting the HRTF magnitude at the contralateral ear from that at the ipsilateral ear. Figure 9 shows that for a lateral source, when the sound source is closer, the ILD is larger, especially at low frequencies. This is a characteristic property for near-field HRTFs as mentioned before.

Conclusion

This work introduces the design and construction with the aid of 3D-printing of a point sound source, to be used in near-field HRTF measurements. The low-frequency behavior and directivity of the developed sound source are particularly evaluated. Preliminary measurement results validate the developed sound source to be used for near-field HRTF acquisition.

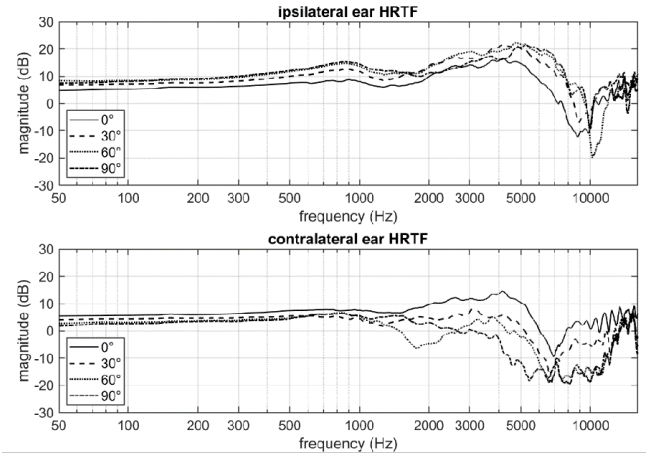

Figure 7: HRTFs measured at 50 cm.

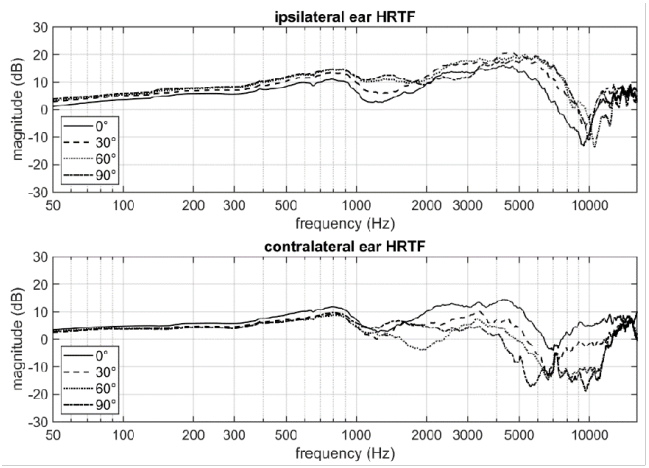

Figure 8: HRTFs measured at 100 cm.

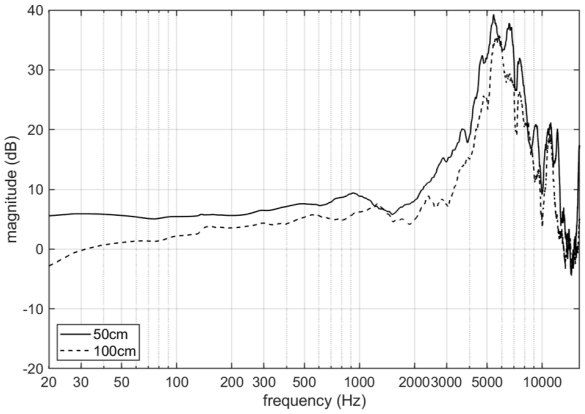

Figure 9: ILDs at 90°.

References

[1] D. Brungart and W. Rabinowitz, "Auditory localization of nearby sources. Head-related transfer functions", The Journal of the Acoustical Society of America, vol. 106, no. 3, pp. 1465-1479, 1999. Available: 10.1121/1.427180.

[2] VRACE, URL:`https://vrace-etn.eu/`

[3] R.H. Schmidt, "Low Frequency Sound Generation by Loudspeaker Drivers", RMS Acoustics & Mechatronics and Grimm Audio, Netherlands, 2017 [Online]. Available: https://www.grimmaudio.com/wp-content/uploads/RMS-white-paper-1-Sound-generation.pdf.

[4] S. Hosoe, T. Nishino, K. Itou and K. Takeda, "Development of Micro-Dodecahedral Loudspeaker for Measuring Head-Related Transfer Functions in The Proximal region," 2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings, 2006, pp. V-V, doi: 10.1109/ICASSP.2006.1661279.

[5] G. Yu, B. Xie and D. Rao, "Directivity of Spherical Polyhedron Sound Source Used in Near-Field HRTF Measurements", Chinese Physics Letters, vol. 27, no. 12, pp. 124302 1-4, 2010. Available: 10.1088/0256-307x/27/12/124302.

[6] B. Chojnacki, S. Terry Cho, and R. Mehra, "Full Range Omnidirectional Sound Source for Near-Field Head-Related Transfer-Functions Measurement," J. Audio Eng. Soc., vol. 69, no. 5, pp. 323-339, (2021 May.). doi: https://doi.org/10.17743/jaes.2021.0004.

[7] R. H. Small, "Closed-Box Loudspeaker Systems Part I: Analysis", J. Audio Eng. Soc., Vol. 20, No. 10, pp. 798–808, 1972.

[8] MONACOR SPM-116/8, URL:`https://www.monacor.de/produkte/components/lautsprechertechnik/hi-fi-mitteltoener-/spm-116-8/`