www.akustik.rwth-aachen.de
DOI: 10.18154/RWTH-2020-11307

Technical Report

A High-Resolution Head-Related Transfer Function Data Set and 3D-Scan of KEMAR

Hark Braren and Janina Fels

Institute and Chair for Hearing Technology and Acoustics, RWTH Aachen University
Corresponding author: jfe@akustik.rwth-aachen.de

Abstract: The KEMAR Head and Torso Simulator is one of the most commonly used artificial head and torso measurement systems and its head-related transfer function (HRTF) is often used in 3D-audio applications. As part of this publication, an HRTF measurement performed in a hemi-anechoic chamber as well as a high resolution 3D-model of the used manikin is provided. This accompanying report details the data acquisition and post-processing steps, that resulted in the provided data sets including a discussion of the shortcomings of the available measurement setup compared to an ideal free-field HRTF.

Keywords: KEMAR, HRTF, anthropometric model, 3D-model

1 Introduction

The human auditory system evaluates the temporal and spectral changes caused by reflections and refraction of incoming sound waves at the head, torso and pinna to identify the location of a sound source [1, 2]. These effects are captured by the head-related transfer function (HRTF), which can in turn be used to create virtual sound sources in 3D virtual auditory environments using binaural techniques [3, 4].
The KEMAR head and torso simulator introduced in 1972 by Knowles [5] and nowadays available from GRAS Sound & Vibration is a widely adopted tool in audio research and engineering applications. Multiple HRTF measurements of KEMAR manikins from different institutions were published over the years [6–8]. This publication aims to extend these with an increased spatial resolution of $1° \times 1°$ and a high-resolution 3D model for simulation applications.
The technical report details the measurement and post-processing procedure that resulted in a high-resolution full sphere measurement of the KEMAR-manikin available from 10.18154/RWTH-2020-11307 and the acquisition of a 3D-model of the head and torso simulator, along with a short discussion of the results.

2 HRTF Measurement

2.1 Procedure

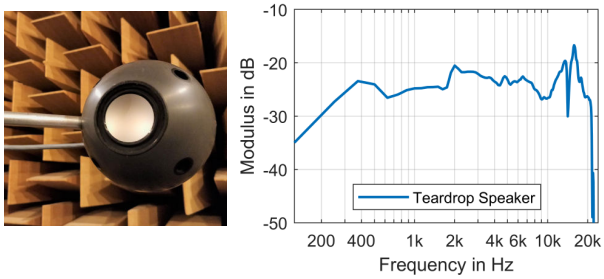
Head-related transfer functions (HRTFs) of an artificial head can be measured by recording the direction-specific transfer function from a loudspeaker to the two microphones placed at the ear canal entrances of the artificial head. In a second step, a reference measurement is taken at the center of where the head used to be using the same loudspeaker. As a free-field environment is assumed during both measurements, the second measurement should be direction-independent and only contain the influence of the microphone and loudspeaker on the measurement result. However, the microphone's placement inside the artificial head results in an acoustic condition that is better described as a pressure field than free-field condition. This can lead to errors at high frequencies which are compensated by using microphones build for these measurement condition, in particular using a free-field compensated microphone during the reference measurement.
An exponential swept sine [9, 10] with a signal length of 65536 samples at 48 kHz was used as the measurement signal. The corresponding impulse responses were then calculated via frequency-domain division as a fast method of de-convolution. The latency of the AD/DA system was measured prior to the acoustic measurements and compensated for.
During each measurement, a single loudspeaker was moved to a position on the equiangular sampling grid using the setup detailed below. There it paused for half a second to let any motion induced ringing die down before the impulse response at that position was measured and the speaker was moved to the next sampling position. Further details on the hardware used to accomplish this, as well as the rest of the measurement hardware is given in the following section.

2.2 Hardware

The artificial head used in the measurement is a GRAS 45BB-4 KEMAR head and torso with large 55 shore hardness anthropometric ears (KB0090 and KB0091). Compared to other KEMAR manikins, this one does not feature the ear simulators, but has microphones placed directly in the ear canal entrances. The microphones are GRAS 40AO 1/2" pre-polarized micro-

phones on GRAS 26CS CCP pre-amplifiers with GRAS RA0003 adapters and supplied trough GRAS 12AL 1-channel CCP power modules over BNC cables. The microphone signals were digitized using an RME Multiface at 48 kHz sampling rate with an in-house developed measurement front-end (ITA-Robo, see Figure 1), used as a conditioning amplifier. The measurement signals were generated using the ITA-Toolbox [11] available from `www.ita-toolbox.org` and converted to analog signals using the same RME Multiface. The output stage of the ITA-Robo was used as the power amplifier and connected to a special teardrop shaped measurement loudspeaker also developed at the Institute of Technical Acoustics in Aachen as shown in Figure 2a. The speaker has a usable frequency range from 200 Hz to 20 kHz. The on-axis frequency response can be seen in Figure 2b. All microphones were calibrated using the GRAS Type 42AA pistonphone to a sound pressure level of 114 dB .

Figure 1: ITA-Robo measurement frontend: conditioning and power amplifier



(a) Teardrop-speaker **(b)** Frequency response

Figure 2: Teardrop shaped measurement loudspeaker with front facing bass-reflex ports

2.3 Measurement

The measurement took place in the hemi-anechoic chamber at the Institute of Technical Acoustics in Aachen. The room measures 12.6 x 7.6 x 4.5 m with 80 cm thick stone wool absorbers on all four walls and on the ceiling. The tiled floor constitutes an acoustically rigid boundary condition that has to be accounted for when processing the measurements. The room allows for free-field measurement conditions, apart from the reflecting floor, down to approximately 100 Hz.

To measure the impulse responses on the desired sampling grid, two degrees of freedom are required to set the azimuth and elevation angles. This is achieved by mounting the excitation loudspeaker on an arm made from small carbon tubes which can be rotated around its center with a radius of 2.066 m as seen in Figure 3. It rotates in the elevation direction from 0° to 120° zenith angle. The azimuth angle is set by mounting the KEMAR artificial head on a turn table, that allows a rotation around its central mount by 360°. Both directions can be controlled from MATLAB and are set to measure at 1 degree increments, thus creating a spherical cap of the upper hemisphere. At a 1 degree resolution, this results in a total of 43560 measurements for the upper hemisphere. The positions are addressed in order by rotating the elevation angle by 1 deg, then rotating the azimuth angle the whole 360 steps in a clockwise direction before raising the elevation by another degree and rotation, one step at a time, in the opposite direction. To accurately position the artificial head for the measurement, a laser alignment system is used as seen in Figures 3 and 4.

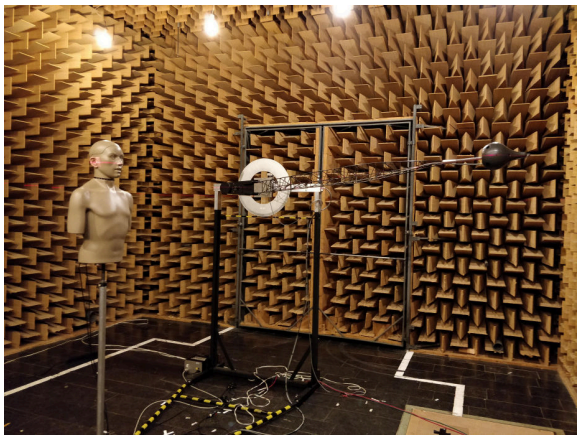


Figure 3: HRTF measurement system located in a hemi-anechoic chamber. The loudspeaker is mounted on an rotating arm covering elevation angles from 0 to 120 degree zenith and the KEMAR mounted on a turntable for azimuth movement

Lower hemisphere

The system cannot measure below the manikin, hence the measurement was split into two hemispheres. To measure the lower hemisphere using the same system, a mount which can support the KEMAR head and torso weighing 11.5 kg upside down without rotational or positional play had to be manufactured. The construction consists of two aluminum tubes (radius 3 cm) in front and behind the manikin, which are connected by two horizontal connecting rods as shown in Figure 4. It is mounted on the turntable to allow azimuth rotation.

The system was aligned the same way as in the right-side-up measurement using the 3-axis laser system and placing the ear canal entrances, and thereby the inter-aural axis, in the center of the spherical cap inscribed by the measurement system. The lower hemisphere was only measured down to 99° zenith angle (36000 positions) compared to the of the upper hemisphere. This allows for sufficient overlap for future corrections of any misalignment between the two measurements while reducing the overall measurement time. In the combined

Figure 4: KEMAR in the upside-down mounting rig

(whole sphere) HRTF, the right-side up measurement is used where possible, to limit the artifacts caused by the upside-down rig, which are further discussed in Section 2.5.

Reference measurement

For the reference measurement, the KEMAR was removed from the setup and replaced by a free-field compensated B&K 4190 microphone on a B&K 2669 preamplifier positioned in the center of the interaural axis aligned by means of the same 3-axis laser system used to align the head and torso simulator. The B&K 2669 preamp is supplied with 200 V polarization voltage from a B&K 2610 measuring amplifier. The microphone again was calibrated using the GRAS 42AA pistonphone to 114 dB at 250 Hz. The measurement loudspeaker was positioned directly above the upward facing microphone (at 0° zenith angle) and the reference transfer function was measured.

2.4 Post-processing

In post processing, the 79560 individual measurements (including duplicates for alignment) were time windowed (and cropped) to remove the floor reflection at the lowest measurement positions. A Hann window with 2 ms fall time was chosen as shown in Figure 5.

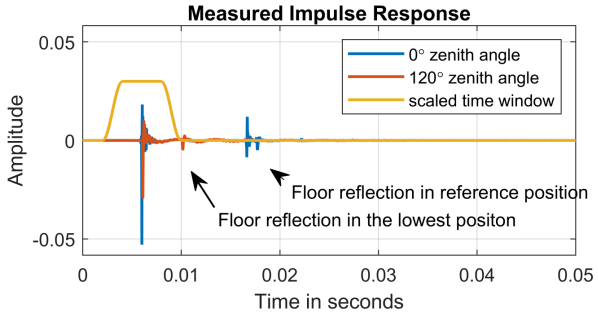


Figure 5: Time-domain measurement with applied time window function

The time-windowed impulse responses at each measurement position were circularly de-convolved by means of regularized spectral division by the reference measurement. Thus, the non-flat frequency response of the loudspeaker, as shown in Figure 2, was compensated for in all measurements.
The low frequency information is affected by the length of the time windows as well as by the modes of the measurement room which are present at frequencies below 100 Hz. To compensate any influence of the room-modes, the frequency information below 100 Hz is interpolated towards unity at 0 Hz using a linear interpolation scheme. This approach is motivated by the fact that wavelengths at these frequencies get large compared to the body and hence the frequency domain information is nearly constant with a linear phase component as described by Xie [12].

Combination of the two halves

Prior to combining the two hemispheres, their azimuth alignment error is evaluated using the interaural time difference (ITD) as a metric. Both hemispheres are expected to have the zero crossing of their ITDs (calculated using the cross-correlation method[13]) in the median plane at a zero-degree azimuth. Thus the individual alignment offset can be calculated by linearly interpolating between the ITD values sampled in 1 degree increments. The offsets are 0 deg and 1.7 deg for the lower and upper hemisphere respectively.

2.5 Results

The resulting HRTF data set contains a full sphere $1° \times 1°$ resolution HRTF data set. Figure 6 shows the resulting linear magnitude directivity pattern of the data set. Zenith angles down to 120° are taken from the right-side up measurement, to reduce the effect of the mounting rig, discussed in the following section. The transition between the measurements of the upper and lower hemisphere is visible in the linear representation as a small step in the directivity, that is present due to slightly varying environmental conditions. However, it vanishes when looking at the data in logarithmic (dB) representation and is expected to be inaudible.

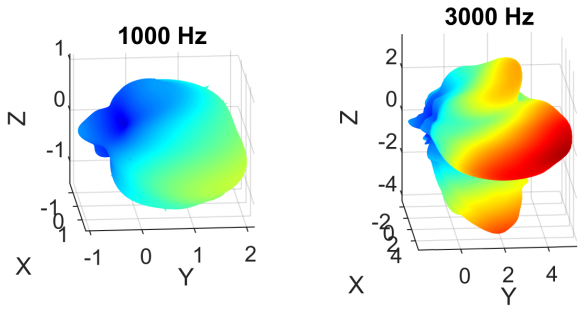


Figure 6: 3D-surface representation of the linear magnitude of the resulting HRTF at different frequencies.

Effect of the upside down mounting rig

Elevation angles on the lower hemisphere below the transition at 120° zenith angles are affected by reflections and diffraction at the mounting rig, see Figure 4. To illustrate the effect, the horizontal plane from both the upside-down and right-side up measurements are shown in Figure 7. It can be seen that the rig affects the frequency domain HRTF starting at around 3kHz with a diffraction pattern symmetric to its location at 0° and 180° azimuth angle. The effect is especially dominant in areas of low energy in the HRTF namely the contralateral side.

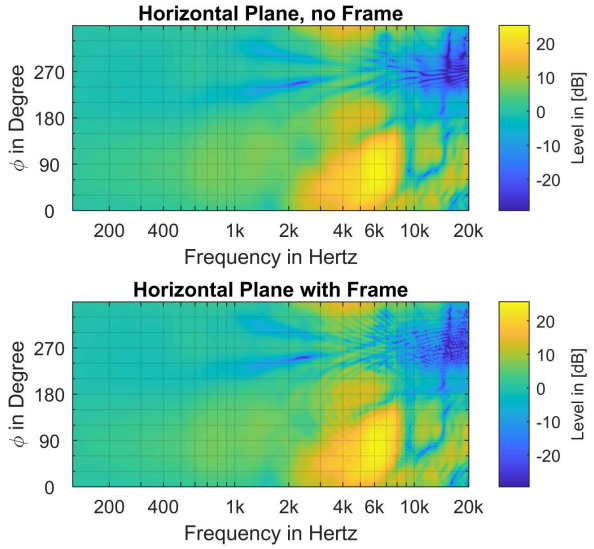


Figure 7: Horizontal plane of the measurement with the KEMAR mounted right-side up (top) and upside-down (bottom) demonstrating the effect of the upside-down mounting rig (see Figure 4).

A closer look at the effect in the frequency domain in two directions is presented in Figure 8. On the ipsilateral side, indicated in red, the diffraction and reflection pattern manifests as a low magnitude comb filter that is barely visible where the dotted line of the measurement (with the upside-down mounting rig) is visible behind the solid line (without it). On the contralateral side, indicated in blue, the effect seen as peaks and valleys of the dotted line is more prominent.
Even though it is unfortunate to have the diffraction pattern superimposed on the HRTF of the artificial head, it is unavoidable using the available measurement system.

3 3D-Data

In addition to the acoustic measurement data, a high resolution 3D-Scan of the KEMAR manikin as shown in Figure 9 was captured using an Artec Space Spider [14] structured light 3D-scanner. The scanner has a maximum resolution of 0.3 mm, however, the scans

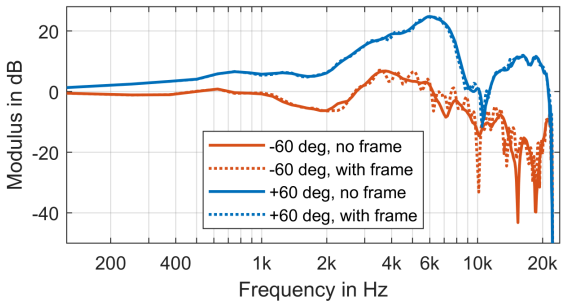
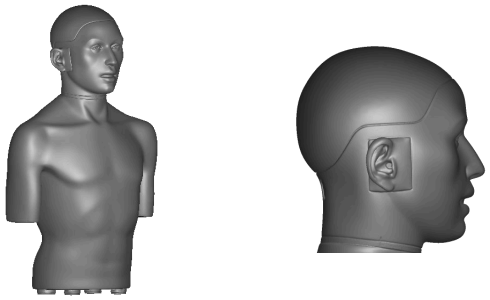


Figure 8: The effect of the upside-down mounting rig (see Figure 4) manifested as a superimposed frequency domain comb filter.

used for the model had a mean resolution of 0.5 mm. This resolution is more than sufficient to have an accurate representation of the acoustic effects of the structure in the audible frequency range up to 20 kHz and beyond according to the 6-element per wavelength criterion typically applied in engineering applications such as a boundary element method (BEM) simulation. The resolution of the available model has then been adaptively reduced using the maximum deviation criterion to allow for an accurate representation of the geometry while at the same time reducing the resulting file size. This results in high vertex densities of the triangulated mesh in areas with complex geometries, such as the pinnae whilst showing a reduced density in simpler areas, such as the more or less flat torso region. However, it does not take into consideration the mesh density needed for simulations.



(a) Whole body **(b)** Detailed view of the right ear.

Figure 9: 3D-Scan of the KEMAR Manikin.

It is thus advised to re-mesh and simplify the mesh by removing unnecessary details for simulation purposes in order to reduce calculation times. A compromise has to be made between computation time and accuracy of the geometry, especially in complex areas such as the pinna.

3.1 Acquisition

The 3D-mesh was crated in Artec Studio 13 by capturing the manikin from different angles in a number of surface scans. This is especially needed to accurately

digitize the complex geometry of the pinnae. The individual scans were aligned based on visual and geometric features to form a complete model of the manikin. More than thirty scans were combined in total to form the final mesh. A more detailed description of the scanning procedure can be found in [15], which should be referred to for reference.

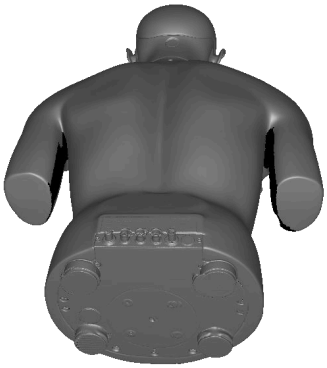


Figure 10: Problematic areas under the arms that needed to be filled in post processing indicated in black.

Most of the body was easy to scan and did not pose any problems. The areas under the arm pits however were hard to capture due to the method of the 3D-scanner and its physical dimensions. The problematic areas, shown in Figure 10, could not be captured with the structured light scanning system but instead were filled in using hole filling algorithms in Autodesk Meshmixer [16] (Version 3.5.474) in post-processing.
The mesh of the final model consists of approximately 150k vertices with an average edge length of 1.9 mm. It is provided as an *.stl* surface mesh under the DOI of this publication alongside the measured HRTF data.

3.2 Simulation

The boundary element method (BEM) is a proven tool to simulate HRTFs from 3D models [17, 18]. Prior to using the provided 3D-model in a boundary element simulation, some of the fine-structure in the model such as the seam at the neck where the head can be rotated relative to the torso, or the feet and connection ports at the bottom of the model where removed and smoothed in order to further reduce mesh vertices. The model was then re-meshed resulting in a mesh containing approximately 85k vertices, valid up to 20 kHz. The HRTF was simulated using the *mesh2hrtf* (Mesh2HRTF, H. Ziegelwanger, ARI, OEAW (`mesh2hrtf.sourceforge.net`)) [19, 20] boundary element solver utilizing the fast multi-pole method (FMM).

A side by side comparison of the horizontal plane HRTF is depicted in Figure 11. In general, the HRTFs are in good agreement with slight differences in the depth of

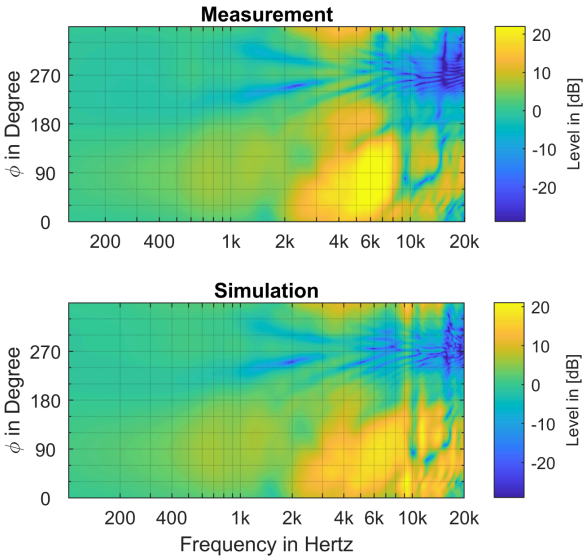


Figure 11: Left ear horizontal plane HRTF from measurement and BEM simulation.

notches (blue lines in the plot) and their center frequencies. This becomes even more obvious when looking at the frequency domain representation of a single direction, as shown in Figure 12. These differences can be explained on the one hand by a mismatch in temperature and thus the speed of sound between the measurement and simulation, which leads to a small shift of the resonance frequencies in the simulation.

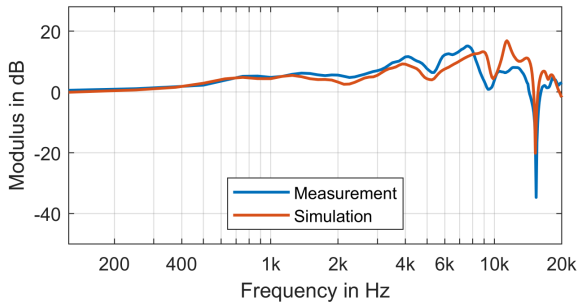


Figure 12: Comparison between measured and simulated HRTF of the left ear at $\varphi = 150°$ in the horizontal plane.

It should also be mentioned, that the high-frequency content of the simulation appears more crisp, less noisy compared to the measurement. This is to be expected from the non-ideal real-world measurement conditions which include equipment noise and small positional inaccuracies, which primarily affect high frequency information.

Acknowledgements

The authors want to thank Jens Ahrens and the acoustics group at Chalmers University of Technology, Gothenburg, Sweden for providing the KEMAR head and torso measurement system.

Supplementary Material

The data sets described in this technical report can be downloaded from 10.18154/RWTH-2020-11307 in both *.ita* format, compatible with the ITA-Toolbox for Matlab (ita-toolbox.org) and *.sofa* [21] format (see sofaconventions.org for suitable APIs).
The provided files include the HRTF as described in this report (*KEMAR_HRTF.ita/.sofa*), alongside the raw impulse response measurements (*KEMAR_HRTF_unprocessed.ita/.sofa*), to which only a time window as described in Section 2.4 and Figure 5 was applied. The free-field Loudspeaker response, time windowed in the same manor, is also made available (*Loudspeaker_Freefield_Response.wav*).

References

[1] J. Blauert, *Spatial Hearing: The Psychophysics of Human Sound Localization*. Cambridge: MIT Press, 1997.

[2] J. Fels and M. Vorländer, "Anthropometric Parameters Influencing Head-Related Transfer Functions," *Acta Acustica united with Acustica*, no. 2, pp. 331–342, DOI: 10.3813/AAA.918156.

[3] H. Møller, M. F. Sørensen, D. Hammershøi, and C. B. Jensen, "Head-related transfer functions of human subjects," *Journal of the Audio Engineering Society*, vol. 43, no. 5, pp. 300–321, 1995.

[4] B. Xie, *Head-related transfer function and virtual auditory display*. J. Ross Publishing, 2013.

[5] M. D. Burkhard and R. M. Sachs, "Anthropometric manikin for acoustic research," *Journal of the Acoustical Society of America*, vol. 58, no. 1, pp. 214–222, 1975. DOI: 10.1121/1.380648.

[6] W. G. Gardner and K. D. Martin, "HRTF measurements of a KEMAR," *Journal of the Acoustical Society of America*, vol. 97, no. 6, pp. 3907–3908, 1995. DOI: 10.1121/1.412407.

[7] V. R. Algazi, R. Duda, D. Thompson, and C. Avendano, "The CIPIC HRTF database," *Proceedings of the 2001 IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics (Cat. No.01TH8575)*, no. October, pp. 99–102, 2001. DOI: 10.1109/ASPAA.2001.969552.

[8] H. Wierstorf, M. Geier, A. Raake, and S. Spors, *A Free Database of Head-Related Impulse Response Measurements in the Horizontal Plane with Multiple Distances*, Jun. 2016. DOI: 10.5281/zenodo.55418.

[9] A. Farina, "Simultaneous measurement of impulse response and distortion with a swept-sine technique," *Proc. AES 108th conv, Paris, France*, no. I, pp. 1–15, 2000. DOI: 10.1109/ASPAA.1999.810884.

[10] S. Müller and P. Massarani, "Transfer-Function Measurement with Sweeps," *Journal of the Audio Engineering Society*, vol. 49, no. 6, pp. 443–471, 2001.

[11] M. Berzborn, R. Bomhardt, J. Klein, J.-G. Richter, and M. Vorländer, "The ITA-Toolbox: An Open Source MATLAB Toolbox for Acoustic Measurements and Signal Processing," 43th Annual German Congress on Acoustics, Kiel (Germany), 6 Mar 2017 - 9 Mar 2017, Mar. 6, 2017.

[12] B. Xie, "On the low frequency characters of head-related transfer function," *Shengxue Xuebao/Acta Acustica*, 2008.

[13] B. F. G. Katz and M. Noisternig, "A comparative study of interaural time delay estimation methods.," *The Journal of the Acoustical Society of America*, vol. 135, no. 6, pp. 3530–3540, 2014. DOI: 10.1121/1.4875714.

[14] Artec Group, *Artec Space Spider*. [Online]. Available: https://www.artec3d.com/portable-3d-scanners/artec-spider.

[15] H. Braren and J. Fels, "A High-Resolution Individual 3D Adult Head and Torso Model for HRTF Simulation and Validation: 3D Data," Medical Acoustics Group, Institute of Technical Acoustics, RWTH Aachen, 2020. DOI: 10.18154/RWTH-2020-06760.

[16] Autodesk, *Meshmixer*. [Online]. Available: http://www.meshmixer.com/.

[17] B. F. G. Katz, "Boundary element method calculation of individual head-related transfer function. II. Impedance effects and comparisons to real measurements," *The Journal of the Acoustical Society of America*, vol. 110, no. 5, pp. 2449–2455, 2001. DOI: 10.1121/1.1412441.

[18] ——, "Boundary element method calculation of individual head-related transfer function. I. Rigid model calculation," *The Journal of the Acoustical Society of America*, vol. 110, no. 5, pp. 2440–2448, 2001. DOI: 10.1121/1.1412440.

[19] H. Ziegelwanger, W. Kreuzer, and P. Majdak, "Mesh2hrtf: Open-source software package for the numerical calculation of head-related transfer functions," 22nd International Congress on Sound and Vibration, 2015.

[20] H. Ziegelwanger, P. Majdak, and W. Kreuzer, "Numerical calculation of listener-specific head-related transfer functions and sound localization: Microphone model and mesh discretization," *The Journal of the Acoustical Society of America*, pp. 208–222, 138 2015.

[21] P. Majdak and M. Noisternig, "Aes69-2015: Aes standard for file exchange-spatial acoustic data file format," in *Audio Engineering Society*, 2015.