THE VIKING HRTF DATASET

Simone Spagnol
Dept. of Architecture, Design & Media Technology
Aalborg University
Copenhagen, Denmark
ssp@create.aau.dk

Kristján Bjarki Purkhús, Rúnar Unnthórsson
School of Engineering and Natural Sciences
University of Iceland
Reykjavík, Iceland
runson@hi.is

Sverrir Karl Björnsson
Dept. of Mechanical Engineering
Technical University of Denmark
Kgs. Lyngby, Denmark
s171951@student.dtu.dk

ABSTRACT

This paper describes the Viking HRTF dataset, a collection of head-related transfer functions (HRTFs) measured at the University of Iceland. The dataset includes full-sphere HRTFs measured on a dense spatial grid (1513 positions) with a KEMAR mannequin with 20 different artificial left pinnae attached, one at a time. The artificial pinnae were previously obtained through a custom molding procedure from 20 different lifelike human heads. The analyses of results reported here suggest that the collected acoustical measurements are robust, reproducible, and faithful to reference KEMAR HRTFs, and that material hardness has a negligible impact on the measurements compared to pinna shape. The purpose of the present collection, which is available for free download, is to provide accurate input data for future investigations on the relation between HRTFs and anthropometric data through machine learning techniques or other state-of-the-art methodologies.

1. INTRODUCTION

Binaural sound rendering techniques typically rely on the use of Head-Related Transfer Functions (HRTFs), i.e. filters that capture the acoustic effects of the human head [1]. HRTFs allow accurate simulation of the signal that arrives at the entrance of the ear canal as a function of the spatial location of the sound source (azimuth, elevation, and distance). The ideal rendering in terms of accuracy involves the use of individual HRTFs measured on the listener. By processing a desired monophonic sound signal with a pair of individual HRTFs, one per channel, and by adequately accounting for headphone-induced spectral coloration [2], authentic 3D sound experiences can take place. Virtual sound sources created with individual HRTFs can be localized almost as accurately as real sources and externalized, provided that head movements can be made [3]. Binaural technologies have a wide variety of applications, ranging from personal entertainment [4], through immersive virtual environments [5], to travel aids for the blind [6].

However, obtaining individual HRTF data is only possible with dedicated research facilities and invasive and/or strenuous recording procedures [7]. This is the reason why non-individual HRTFs, acoustically measured on anthropomorphic mannequins or generic human individuals, are often preferred in practice. Several HRTF sets are available online, the most popular being those measured on the KEMAR mannequin [8] or the Neumann KU-100 dummy head [9]. Alternatively, an HRTF set can be taken from one of many public databases of individual measurements (e.g. the CIPIC database [10]); many of these databases were recently unified in a common HRTF format known as Spatially Oriented Format for Acoustics (SOFA). [1]

The drawback with non-individual HRTFs is that they obviously refer to a different anthropometry than the listener's. In particular, the most relevant differences between the HRTFs of two subjects are due to different pinna features (shape, size, and orientation) that give every individual a unique pinna shape [11]. When used for binaural rendering, this HRTF mismatch often results in localization errors such as front/back confusion, wrong perception of elevation, and inside-the-head localization [12].

For the above reasons, studying the relationship between pinna features and HRTFs is an essential step towards understanding of the underlying acoustical mechanisms. It is known, for instance, that the frequency of the first pinna-related peak depends on the dimensions of the concha [13, 14], and that notch frequencies are related to the distance between the ear canal and the most prominent pinna edges [15, 16]. However, previous studies that include applications of anthropometric regression methods to measured HRTF data [17–19] have produced mixed results, highlighting that many of these relations are not fully understood yet. One of the reasons might rely in the HRTF data collected on a human population, with issues related to the relative positioning of the microphones inside the

[1] www.sofaconventions.org

Figure 1. The used HRTF measurement system and room (left); detail of the KEMAR with custom left pinna (right).

ears, head movements during measurement, and non-ideal measurement conditions. Still, even measurements on the same dummy head often result in dissimilar HRTF sets, depending on the used measurement system [9].

In this paper we present the design, implementation and preliminary analysis of a novel set of HRTF measurements, that we refer to as the Viking HRTF dataset. Measurements were taken on a KEMAR mannequin with various interchangeable pinna shapes, with the purpose of providing accurate input data for determining the individual HRTF from a 3D representation of the user [20]. We implemented a procedure for casting artificial pinnae from lifelike human heads so as to provide a reasonable sample of pinna shapes for KEMAR from a real human population, having all remaining anthropometric parameters of the head, torso, and opposite pinna fixed. Furthermore, although limited by the available resources, we designed the automated measurement setup to be as accurate and stable as possible, in order to guarantee reproducible measurements.

2. METHODS

2.1 Measurement system

The HRTF measurements were taken automatically with the system pictured in Figure 1. The measurement system consisted of a KEMAR mannequin [2] mounted on a 360° rotating cylindrical stand and a Genelec 8020CPM-6 loudspeaker mounted on an L-shaped rotating arm. The configuration for the mannequin was the 45BB-4 with standard large anthropometric pinnae (35 Shore-OO hardness, GRAS KB5000/01) as reference and half-inch pressure microphones (GRAS 40AO) placed at the ear canal entrances. The two audio channels were fed to an RME Fireface 802 sound card connected to a PC running MATLAB.

The distance between the loudspeaker tweeter and center of the mannequin head was designed to be a constant

[2] http://kemar.us

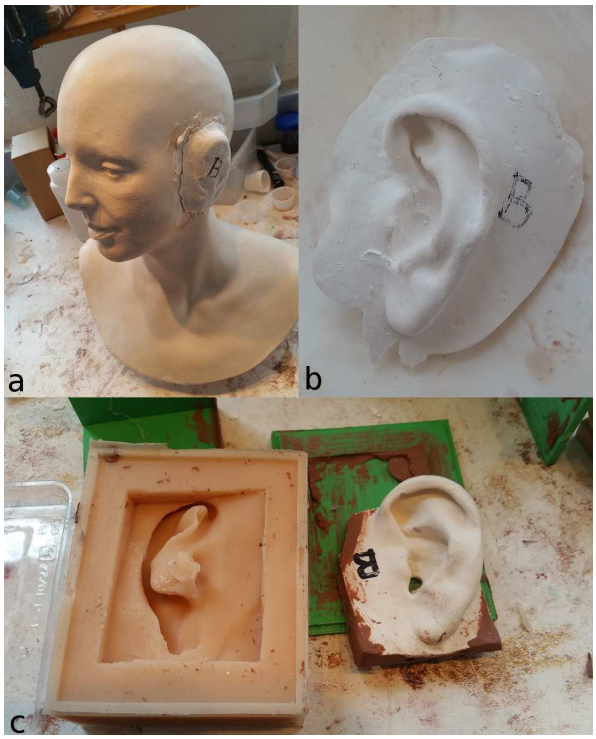

Figure 2. The making of a negative ear mold: (a) 1st negative mold; (b) Jesmonite®replica; (c) 2nd negative mold.

1 m, independently of the orientation of the mannequin and arm. This distance value guarantees the collection of far-field spectral cues with reasonable accuracy [21]. Absolute references for azimuth and elevation were provided through a fixed marker on the bottom of the mannequin (0° azimuth) and a bubble level mounted on the side segment of the arm (0° elevation), respectively. Correct alignment in all three axes could be attained by targeting laser beams to the microphones and tip of the nose of the KEMAR. This experimental choice allows collection of HRTFs on a full-azimuth range with elevations from $-50°$ to $90°$, according to a vertical polar coordinate system.

Rotation of the mannequin and arm was managed by two independent high-torque step motors (JVL MST001A, 1.2 Nm), controlled through two digital step drives (Geckodrive G213V) in full-step mode (200 steps per revolution). In order to increase the torque and angular resolution of the system, two $100:1$ gearboxes were installed between the first motor and the arm and between the second motor and the mannequin. As a result, the minimum rotation angle of both the arm and mannequin was $0.018°$. Furthermore, in order to reduce the needed torque on the arm, a 22 kg counterweight was applied to the shorter appendix so to balance the weight of the loudspeaker and the longer appendix of the arm. Communication between the step drives and PC was also managed in MATLAB.

2.2 Pinna casting

In order to provide a sufficient sample of pinna shapes for the HRTF measurements, we set up and applied a custom

ISBN 978-84-09-08518-7

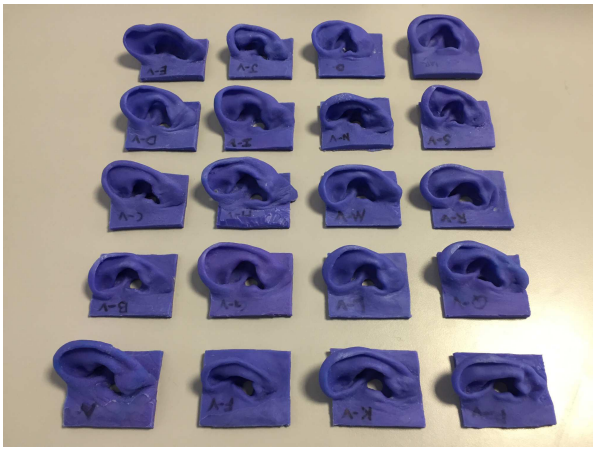

Figure 3. The 20 left pinna replicas obtained through the casting procedure.

procedure to cast silicone replicas of pinnae from dummy heads to fit the KEMAR as follows.

Step 1: First negative mold. We applied two layers of silicone mix to the ear, first a thin one with a paintbrush and then a thick one ($1/2$ cup) with a spatula, letting each layer dry for 1 day before the next step. In order to reduce leakage onto vertical surfaces, we added a few drops of thixotropic additive to the silicone for the thick layer. Then, we applied wet plaster of Paris strips over the negative silicone mold (Figure 2a) in order to efficiently peel it from the dummy head and to give it support so it does not deform in Step 2.

Step 2: Jesmonite® replica. We poured a quick mix of $1/2$ cup of water and two teaspoons of Jesmonite® into the negative mold obtained in Step 1 placed above a vibration table (in order to avoid formation of air bubbles), and let it harden for $1-2$ days. Then, we removed the Jesmonite® ear (Figure 2b) and occasionally filled gaps or deformations in it with clay. Finally, we drilled a hole in the back of the concha and sanded the ear base with a belt sander to accurately fit the size of the KEMAR pinna slot.

Step 3: Second negative mold. We placed the Jesmonite® ear replica into a custom-made plastic box with open top and secured it to the bottom with clay. Then, we poured 2 cups of silicone mix inside the plastic box, placed on the vibration table, and let it dry for 1 day. We finally removed the plastic box and Jesmonite® ear (Figure 2c).

Step 4: Final pinna replica. We first poured car wax into the inner parts of the negative mold obtained in Step 3 as a release agent, and let it dry for half a day. Then we poured $1/3$ cup of a different silicone mix from the one used in Steps 1 and 3 (25 Shore-A hardness) into the negative mold, and let it dry for 1 day. We finally removed the silicone pinna replica and cut the excess parts on the base with a knife.

As pictured in Figure 3, we applied the procedure to a series of left ears of 20 different subjects. These include the KEMAR with standard large anthropometric pinnae and 19 different lifelike dummy heads made out of plaster, borrowed from the Saga Museum in Reykjavík. The heads, manufactured between 2001 and 2003 for artistic purposes,[3] reproduce with high fidelity the anthropometric traits of 19 Icelandic humans (7 female), aged between 7 and 77 at the time of manufacturing.

Table 1. HRTF measurement positions.

Elevations [deg]	[-45,45]	[50,70]	[75,85]	90
Step [deg]	5	15	45	360
No. of azimuths	72	24	8	1

2.3 Measurement procedure

The measurements took place on several days outside office hours (4PM to 12PM) in a silent office room with reflecting walls, floor and ceiling inside the Tæknigarður building of the University of Iceland (see left panel of Figure 1). Since we did not record the free-field response, the measurements include the response of the loudspeaker and microphones as well as some effects of the room (later minimized by windowing as explained in the next subsection). We used the logarithmic sweep method [22] to record the single acoustical responses. The input sweep signal, whose level was kept constant throughout the whole measurement schedule, spanned frequencies between 20 Hz and 20 kHz in 1 s, at a sampling rate $f_s = 48$ kHz. The average SPL level at 1 kHz for a frontal stimulus as collected through a Class 1 sound level meter at the center of the reference system was 82 dB.

Sound source location was specified through the azimuth angle θ and elevation angle ϕ in vertical-polar coordinates. Elevations were uniformly sampled in $5°$ steps from $-45°$ to $90°$, while azimuths were sampled in different increments as shown in Table 1 in order to obtain roughly uniform density towards the upper pole of the sphere. The total number of spatial positions per measurement was 1513. At each measurement session, and for each elevation, the mannequin was consecutively rotated of the corresponding angular step and sweep responses were acquired at each azimuth angle. After completion of all azimuth angles for the current elevation, the mannequin was rotated back to its starting position and the arm moved down to the next elevation angle. In order to maintain the HRTF measurements as silent as possible, 0.5-ms pauses were introduced between all motor commands and the previous/next recording. The duration of one single measurement session was approximately 105 minutes.

Twenty-three measurement sessions (3 control measurements + 20 test measurements) were scheduled in total. A standard large anthropometric pinna (35 Shore-OO hardness) was installed on the right channel of the KEMAR throughout the whole measurement schedule, while the left pinna changed at every measurement session. In the first and second control sessions ($CS1$ and $CS2$) we installed the corresponding original left KEMAR pinna and its variant with different hardness (55 Shore-OO), respectively. These sessions were introduced in order to check for dif-

[3] https://sagamuseum.is/sagadesign/

ISBN 978-84-09-08518-7

ferences in materials and to assess the fidelity of our KEMAR pinna replica to the original ones. The following 20 sessions were devoted to the 20 custom-made pinnae, including the KEMAR pinna replica (right panel of Figure 1) and the 19 human pinnae (alphabetically labeled A to S, see Figure 3). The third and last control session ($CS3$) was again a measurement with both original 35 Shore-OO hardness pinnae, and was introduced in order to control for reproducibility of our measurements. Between each two measurement sessions, the HRTF measurement system was manually calibrated to the correct starting position by tuning each motor to its corresponding absolute reference.

2.4 Post-processing

A post-processing script was written to recover the HRTF from the corresponding raw sweep response, based on the following sequential steps. First, we compute the cross-correlation function $\Psi[n]$ between the raw sweep response $y[m]$ and the input sweep signal $s[m]$ in order to find the starting point of the sweep response. In particular, we find the maximum correlation value Ψ_M in $\Psi[n]$ and extract the lag n_i of the first sample in $\Psi[n]$ such that $\Psi[n] > 0.5\Psi_M$. Then, the starting point of $y[m]$ corresponds to the lag n_j with maximum correlation value within a 21-sample interval centered in $\Psi(n_i)$. This trick was introduced so as to avoid directly picking Ψ_M as starting point, that due to the echoic measurement conditions could correspond to a wall reflection, especially in the case of contralateral HRTFs.

Then, according to the logarithmic sweep method, we perform inverse filtering on the measured sweeps in order to obtain the corresponding impulse responses. The inverse reference spectrum $S^{-1}[n]$ of the excitation signal is first computed and then low-passed and high-passed with second-order digital Butterworth filters to compensate for the original zero sound pressure level below 20 Hz and above 20 kHz in the sweep signal. Then, the impulse response is obtained as

$$h[n] = \Re(\mathcal{F}^{-1}(\mathcal{F}(y[n]) * S^{-1}[n])), \qquad (1)$$

where $\mathcal{F}$ and $\mathcal{F}^{-1}$ are the DFT and inverse DFT functions, respectively.

Subsequently, a 128-sample Hann window is applied to each impulse response $h[n]$ with the aim of removing unwanted early and late reflections occurring later than approximately 2.5 ms from the onset on the measurement equipment and inside the room, yielding $h_w[n]$. This value guarantees an appropriate windowing even for those measurement points where the loudspeaker is close to the floor or the ceiling. The HRTF is then simply calculated as the magnitude response of the DFT of $h_w[n]$.

3. RESULTS

In this section we show some preliminary results as an assessment of the trustworthiness of the collected HRTF dataset. Figure 4 shows an example plot of HRTF magnitudes on the median plane for a custom-made pinna (set J). Well-known effects can be recognized here, such as

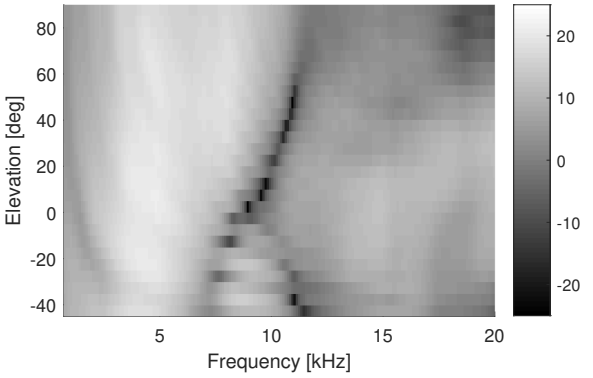

Figure 4. Median-plane HRTF magnitudes of set J, left channel.

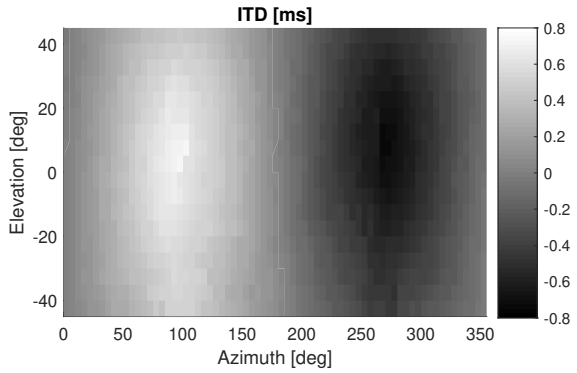

Figure 5. ITD measurements of set J (left channel: custom made pinna, right channel: KEMAR 35 Shore-OO pinna).

the shoulder reflection ridge between 1 and 2 kHz, the omnidirectional resonance around 4 kHz, and the elevation-dependent pattern of peaks and notches [1], with the higher number of notches at lower elevations [23, 24].

We also computed interaural time differences (ITD) as the offset between the starting points of the left and right channel for each sweep response on all sets. Figure 5 shows ITDs for the same example set. In accordance with previous literature [25], ITD values range between zero (median plane) and ± 0.8 ms for lateral directions. We also report in general noticeable asymmetries between left and right channels, resulting in ITD maxima and minima at two different elevation angles. This is due to the presence of two different pinnae in most measurements (KEMAR on the right channel and custom-made on the left one).

In order to evaluate the robustness and reproducibility of our measurements, we calculated the mean spectral distortion between each pair of HRTF sets H_a and H_b as [26]

$$SD(a,b) = \frac{1}{N_\alpha} \sum_i \sqrt{\frac{1}{N_f} \sum_k \left(20 \log_{10} \frac{|H_a(\alpha_i, f_k)|}{|H_b(\alpha_i, f_k)|} \right)^2}, \qquad (2)$$

where α_i is an available spatial position, f_k is an available frequency, N_α is the number of common spatial positions, and N_i is the number of frequencies in the con-

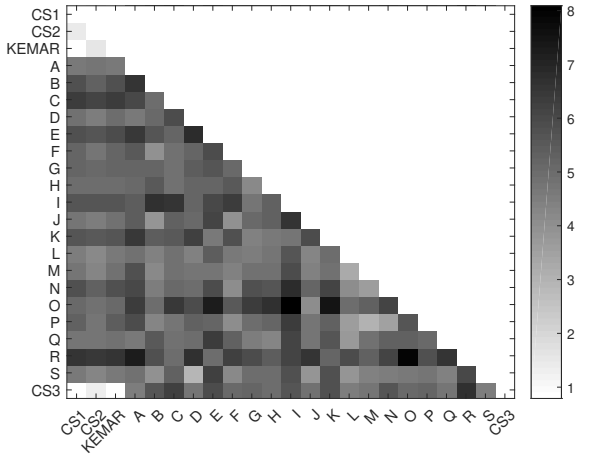

Figure 6. Mean spectral distortion [dB] between all pairs of HRTF sets (left channels only). For simplicity, values above the diagonal line are not repeated.

sidered range. For the following computations, in order to best capture the effect of the pinna, we limit this range between 3 kHz and 10 kHz. While the right channel is just affected by measurement noise, as the low mean spectral distortion (mean 0.26 dB, std 0.07 dB) demonstrates, the left channel results show considerable differences between different pinnae. As displayed in Figure 6, mean spectral distortion between each pair of custom-made human pinnae is never less than 2.93 dB (mean 5.25 dB, std 0.87 dB).

On the other hand, the KEMAR pinna replica scores a lower mean distortion with respect to the reference set measured in $CS1$ (0.84 dB), which is interestingly lower than that of the original KEMAR 55 Shore-OO pinna measured in $CS2$ (1.37 dB). This result suggests the accuracy of the pinna casting procedure and the negligible impact of the silicone hardness. In other words, although differences in hardness may slightly influence the spectral similarity of HRTFs, even the most minimal difference in pinna shape seems more influential. Furthermore, results of $CS3$ indicate that although repositioning the KEMAR pinna on the left channel introduces some additional distortion (0.79 dB), this does not exceed twice the distortion due to measurement noise (right channel, 0.42 dB).

Finally, as an assessment of the fidelity of our KEMAR measurements, in Figure 7 we show the current measurements of the KEMAR with the reference pinna (35 Shore-OO hardness, measured in control session 1) on the horizontal plane and compare them to previous high-quality measurements of the KEMAR mannequin at 1-m distance by Brungart and Rabinowitz (top right panel of Figure 5 in [27]). If we exclude the low-frequency region, where differences are clearer because of the different ear-canal configuration of the KEMAR in the two measurements, we can recognize the same salient spectral features above 2 kHz, corresponding to the five reference points A-E (spectral peaks and notches). This qualitative result evidences the effectiveness of our measurement setup in conveying high-quality reference KEMAR measurements.

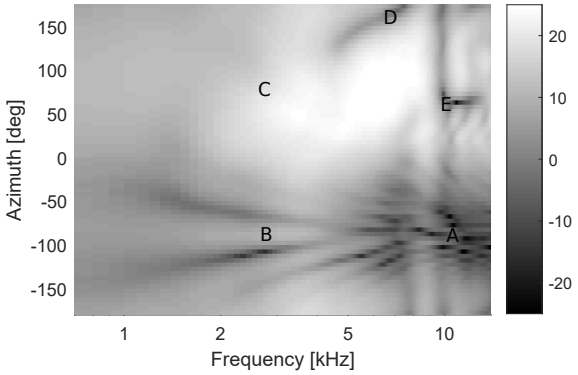

Figure 7. Reference horizontal-plane KEMAR measurements, left channel.

4. CONCLUSIONS

In this paper we presented a dataset of full-sphere HRTFs of KEMAR with 20 different left pinnae obtained from molds of lifelike human heads. The measurement setup and procedure was described in detail, along with the post-processing operations designed to extract polished HRTFs from measured responses. Our preliminary results suggest the accuracy, variety, and reproducibility of the collected data. On the other hand, the hardness finding warrants quantified replications with a measure of central tendency and dispersion. These results will hopefully encourage investigations on the relation between HRTFs and anthropometric data through machine learning techniques or other state-of-the-art methodologies [28]. Full HRTF and ITD data, together with detailed scans of the used pinnae, are available for free download at the dataset web page [4] that will constantly be updated with new data and information.

Acknowledgments

The authors would like to thank Ernst Backman, owner of the Saga Museum in Reykjavík, for his invaluable assistance with the ear casting process and the access to his workshop and tools. This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements No 643636 and No 797850.

5. REFERENCES

[1] C. I. Cheng and G. H. Wakefield, "Introduction to head-related transfer functions (HRTFs): Representations of HRTFs in time, frequency, and space," *J. Audio Eng. Soc.*, vol. 49, no. 4, pp. 231–249, April 2001.

[2] Z. Schärer and A. Lindau, "Evaluation of equalization methods for binaural signals," in *Proc. 126th Conv. Audio Eng. Soc.*, Munich, Germany, May 2009.

[3] A. W. Bronkhorst, "Localization of real and virtual sound sources," *J. Acoust. Soc. Am.*, vol. 98, no. 5, pp. 2542–2553, November 1995.

[4] `itsadive.create.aau.dk/index.php/viking-hrtf/`

[4] D. S. Brungart, "Near-field virtual audio displays," *Presence*, vol. 11, no. 1, pp. 93–106, February 2002.

[5] S. Spagnol, R. Hoffmann, F. Avanzini, and A. Kristjánsson, "Effects of stimulus order on auditory distance discrimination of virtual nearby sound sources," *J. Acoust. Soc. Am.*, vol. 141, no. 4, pp. EL375–EL380, April 2017.

[6] S. Spagnol, G. Wersényi, M. Bujacz, O. Balan, M. Herrera Martínez, A. Moldoveanu, and R. Unnthórsson, "Current use and future perspectives of spatial audio technologies in electronic travel aids," *Wireless Comm. Mob. Comput.*, vol. 2018, p. 17 pp., March 2018.

[7] B. Xie, *Head-Related Transfer Function and Virtual Auditory Display*, 2nd ed. Plantation, FL, USA: J.Ross Publishing, June 2013.

[8] M. D. Burkhard and R. M. Sachs, "Anthropometric manikin for acoustic research," *J. Acoust. Soc. Am.*, vol. 58, no. 1, pp. 214–222, July 1975.

[9] A. Andreopoulou, D. R. Begault, and B. F. G. Katz, "Inter-laboratory round robin HRTF measurement comparison," *IEEE J. Select. Topics Signal Process.*, vol. 9, no. 5, pp. 895–906, August 2015.

[10] V. R. Algazi, R. O. Duda, D. M. Thompson, and C. Avendano, "The CIPIC HRTF database," in *Proc. IEEE Work. Appl. Signal Process., Audio, Acoust.*, New Paltz, New York, USA, October 2001, pp. 1–4.

[11] A. Abaza, A. Ross, C. Hebert, M. A. F. Harrison, and M. S. Nixon, "A survey on ear biometrics," *ACM Trans. Embedded Computing Systems*, vol. 9, no. 4, pp. 39:1–39:33, March 2010.

[12] H. Møller, M. F. Sørensen, C. B. Jensen, and D. Hammershøi, "Binaural technique: Do we need individual recordings?" *J. Audio Eng. Soc.*, vol. 44, no. 6, pp. 451–469, June 1996.

[13] E. A. G. Shaw and R. Teranishi, "Sound pressure generated in an external-ear replica and real human ears by a nearby point source," *J. Acoust. Soc. Am.*, vol. 44, no. 1, pp. 240–249, 1968.

[14] P. Mokhtari, H. Takemoto, R. Nishimura, and H. Kato, "Frequency and amplitude estimation of the first peak of head-related transfer functions from individual pinna anthropometry," *J. Acoust. Soc. Am.*, vol. 137, no. 2, pp. 690–701, February 2015.

[15] V. C. Raykar, R. Duraiswami, and B. Yegnanarayana, "Extracting the frequencies of the pinna spectral notches in measured head related impulse responses," *J. Acoust. Soc. Am.*, vol. 118, no. 1, pp. 364–374, July 2005.

[16] S. Spagnol, M. Geronazzo, and F. Avanzini, "Fitting pinna-related transfer functions to anthropometry for binaural sound rendering," in *Proc. IEEE Int. Work. Multi. Signal Process. (MMSP'10)*, Saint-Malo, France, October 2010, pp. 194–199.

[17] P. Bilinski, J. Ahrens, M. R. P. Thomas, I. J. Tashev, and J. C. Platt, "HRTF magnitude synthesis via sparse representation of anthropometric features," in *Proc. 39th IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP 2014)*, Firenze, Italy, May 2014, pp. 4501–4505.

[18] F. Grijalva, L. Martini, D. Florencio, and S. Goldenstein, "A manifold learning approach for personalizing HRTFs from anthropometric features," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 24, no. 3, pp. 559–570, March 2016.

[19] S. Spagnol and F. Avanzini, "Frequency estimation of the first pinna notch in head-related transfer functions with a linear anthropometric model," in *Proc. 18th Int. Conf. Digital Audio Effects (DAFx-15)*, Trondheim, Norway, December 2015, pp. 231–236.

[20] S. Spagnol, M. Geronazzo, D. Rocchesso, and F. Avanzini, "Synthetic individual binaural audio delivery by pinna image processing," *Int. J. Pervasive Comput. Comm.*, vol. 10, no. 3, pp. 239–254, July 2014.

[21] S. Spagnol, "On distance dependence of pinna spectral patterns in head-related transfer functions," *J. Acoust. Soc. Am.*, vol. 137, no. 1, pp. EL58–EL64, January 2015.

[22] S. Müller and P. Massarani, "Transfer-function measurement with sweeps," *J. Audio Eng. Soc.*, vol. 49, no. 6, pp. 443–471, June 2001.

[23] V. R. Algazi, C. Avendano, and R. O. Duda, "Elevation localization and head-related transfer function analysis at low frequencies," *J. Acoust. Soc. Am.*, vol. 109, no. 3, pp. 1110–1122, March 2001.

[24] S. Spagnol, M. Hiipakka, and V. Pulkki, "A single-azimuth pinna-related transfer function database," in *Proc. 14th Int. Conf. Digital Audio Effects (DAFx-11)*, Paris, France, September 2011, pp. 209–212.

[25] G. F. Kuhn, "Model for the interaural time differences in the azimuthal plane," *J. Acoust. Soc. Am.*, vol. 62, no. 1, pp. 157–167, July 1977.

[26] S. Spagnol, E. Tavazzi, and F. Avanzini, "Distance rendering and perception of nearby virtual sound sources with a near-field filter model," *Appl. Acoust.*, vol. 115, pp. 61–73, January 2017.

[27] D. S. Brungart and W. M. Rabinowitz, "Auditory localization of nearby sources. Head-related transfer functions," *J. Acoust. Soc. Am.*, vol. 106, no. 3, pp. 1465–1479, September 1999.

[28] M. Geronazzo, S. Spagnol, and F. Avanzini, "A modular framework for the analysis and synthesis of head-related transfer functions," in *Proc. 134th Conv. Audio Eng. Soc.*, no. 8882, Rome, Italy, May 2013.