THE CIPIC HRTF DATABASE

V. R. Algazi, R. O. Duda and D. M. Thompson

CIPIC
U.C. Davis
Davis, CA 95616, USA
algazi@ece.ucdavis.edu

C. Avendano

Creative Advanced Technology Center
1500 Green Hills Road
Scotts Valley, CA 95066
carlosa@atc.creative.com

ABSTRACT

This paper describes a public-domain database of high-spatial-resolution head-related transfer functions measured at the U. C. Davis CIPIC Interface Laboratory. Release 1.0 includes head-related impulse responses for 45 subjects at 25 different azimuths and 50 different elevations (1250 directions) at approximately 5° angular increments. In addition, the database contains anthropometric measurements for each subject. Statistics of anthropometric parameters and correlations between anthropometry and some temporal and spectral features of the HRTFs are reported.

1. INTRODUCTION

Head-related transfer functions (HRTFs) capture the sound localization cues created by the scattering of incident sound waves by the body, and play a central role in spatial audio systems. Most HRTF-based commercial systems convolve the input signal with a single, "standard" head-related impulse response (HRIR), and several studies have employed the public-domain dataset for the KEMAR mannequin [1]. However, it is well known that HRTFs vary significantly from person to person, and that serious perceptual distortions (particularly front/back confusion and elevation errors) can occur when one listens to sounds spatialized with a non-individualized HRTF [2].

Although the determination of individual HRTFs can addressed in a number of ways, most recently by numerical computations based on a detailed geometric mesh of the human body [3, 4], the study of individual variations requires a database of uniformly measured HRTFs. Several laboratories have developed HRTF databases to support their own research (e.g., [5]). However, the only publicly available database is the AUDIS catalog [6], which is limited to 12 subjects measured at approximately 120 positions in space, and cannot be used for commercial purposes.

The CIPIC Interface Laboratory at U.C. Davis has measured HRTFs at high spatial resolution for more than 90 subjects. Release 1.0 — a public-domain subset for 45 subjects (including KEMAR with large and with small pinnae) — is available by downloading from our web site (http:// interface.cipic.ucdavis.edu). In addition to including impulse responses for 1250 directions for each ear of each subject, the database includes a set of anthropometric measurements that can be used for scaling studies. This paper describes the content of the database, and briefly describes the characteristics of the data. Additional technical documentation and MATLAB^{TM} utility programs for inspecting the data are provided with the database files.

2. MEASUREMENTS

Excluding the KEMAR mannequin, the 43 human subjects (27 men and 16 women) were either U.C. Davis students or visitors to the CIPIC Interface Laboratory. All HRTFs were measured with the subject seated at the center of a 1-m radius hoop whose axis was aligned with the subject's interaural axis. The position of the subject's head was not constrained, but the subject could monitor his or her head position [7].[1]

Bose Acoustimass^{TM} loudspeakers (5.8-cm cone diameter) were mounted at various positions along the hoop. A modified Snapshot^{TM} system from Crystal River Engineering generated Golay-code signals. The subject's ear canals were blocked, and Etymōtic Research ER-7C probe microphones were used to pick up the Golay-code signals. The microphone outputs were digitized at 44.1-kHz, 16-bit resolution and processed by Snapshot's `oneshot` function to yield a raw HRIR. A modified Hanning window was applied to the raw HRIR measurements to remove room reflections, and the results were free-field compensated to correct for the spectral characteristics of the transducers.[2] The length of each HRIR is 200 samples, corresponding to a duration of about 4.5 ms.

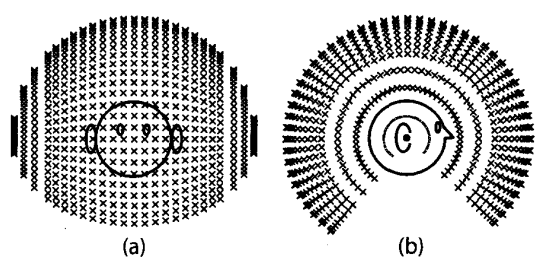

Figure 1: *Locations of data points (a) front (b) side*

Sound source location was specified by the azimuth angle θ and elevation angle ϕ in interaural-polar coordinates.[3] Elevations

[1] Small head motion, detected by abrupt changes in ITD, commonly occurred. Data sets were discarded if significant spectral discontinuities were observed.

[2] The free-field response was measured at the position of the center of the head. The free-field compensation was limited to 15 dB so as to control the duration and the ripples of the compensated response. In listening tests, Møller's procedure was used to compensate for headphones and to re-introduce the missing ear-canal resonance [12].

[3] Readers who are more familiar with vertical-polar coordinates should be warned that interaural-polar azimuth is limited to the range from $-90°$ to $+90°$. Points that are in back of the subject are found at 180° elevation.

Meta Platforms, Inc.
Exhibit 1046
Authorized licensed use limited to: New York University. Downloaded on October 25,2022 at 02:04:45 UTC from IEEE Xplore. Restrictions apply. Page 001

were uniformly sampled in $360/64 = 5.625°$ steps from $-45°$ to $+230.625°$. To obtain roughly uniform density on the sphere, azimuths were sampled at $-80°, -65°, -55°$, from $-45°$ to $45°$ in steps of $5°$, at $55°, 65°$, and $80°$. This leads to spatial sampling at 1250 points, as illustrated in Fig. 1.

3. ANTHROPOMETRY

Although the exact HRTF is complicated, its general behavior can be estimated from fairly simple geometric models of the torso, head and pinnae [8, 9, 10]. These models can be individualized to particular listeners if appropriate anthropometric measurements are available [11]. However, specifying a general set of well-defined and relevant measurements is problematic. The problem is particularly difficult for the pinna, where small variations can produce large changes in the HRTF. Anthropometric measurements, even if imperfect, enable the investigation of possible correspondences or correlations between physical dimensions and HRTF features.

The choice of anthropometry relevant to understanding or estimating HRTFs lead us to follow an approach proposed by Genuit [8], and to define a set of 27 anthropometric measurements — 17 for the head and torso (Fig. 2) and 10 for the pinna (Fig. 3).[4]

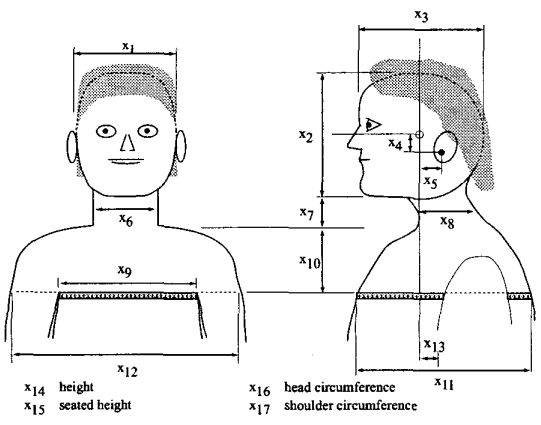

Figure 2: *Head and torso measurements*

The range of variation for the individuals in the CIPIC database can be measured by some statistics for the anthropometric measurements. In general, histograms of the individual measurements indicate a basically normal distribution of values. The means and standard deviations for the anthropometric parameters are listed in Table 1.[5] Here distances are measured in cm and angles in degrees, and the percentage variation is $2\sigma/\mu$ in percent. For example, the mean head width was 14.49 cm, and, assuming a normal distribution, 95% of the cases were within ± 13 % of the mean. Excluding the offset measurements x_4, x_5 and x_{13}, for which percentage deviation is not meaningful, we see that the average percentage deviation is $\pm 26\%$. Thus, there is considerable variation in the sizes and shapes of the subjects in the database.

[4] In general, a particular measurement was included if (a) it was deemed to have a significant influence on the HRTF, and (b) it could be reliably and reasonably easily measured. In addition we recorded each subject's weight, age and sex as possibly relevant.

[5] Although the number of subjects in the database is too small to be representative of the general population, our results are in general agreement with published values for adults [13].

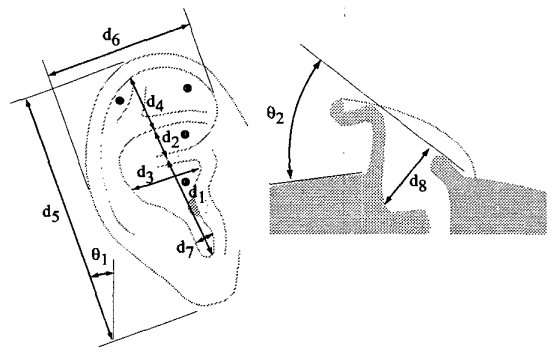

Figure 3: *Pinna measurements*

Var	Measurement	μ	σ	%
x_1	head width	14.49	0.95	13
x_2	head height	21.46	1.24	12
x_3	head depth	19.96	1.29	13
x_4	pinna offset down	3.03	0.66	43
x_5	pinna offset back	0.46	0.59	254
x_6	neck width	11.68	1.11	19
x_7	neck height	6.26	1.69	54
x_8	neck depth	10.52	1.22	23
x_9	torso top width	31.50	3.19	20
x_{10}	torso top height	13.42	1.85	28
x_{11}	torso top depth	23.84	2.95	25
x_{12}	shoulder width	45.90	3.78	16
x_{13}	head offset forward	3.03	2.29	151
x_{14}	height	172.43	11.61	13
x_{15}	seated height	88.83	5.53	12
x_{16}	head circumference	57.33	2.47	9
x_{17}	shoulder circumference	109.43	10.30	19
d_1	cavum concha height	1.91	0.18	19
d_2	cymba concha height	0.68	0.12	35
d_3	cavum concha width	1.58	0.28	35
d_4	fossa height	1.51	0.33	44
d_5	pinna height	6.41	0.51	16
d_6	pinna width	2.92	0.27	18
d_7	intertragal incisure width	0.53	0.14	51
d_8	cavum concha depth	1.02	0.16	32
θ_1	pinna rotation angle	24.01	6.59	55
θ_2	pinna flare angle	28.53	6.70	47

Table 1. Anthropometric statistics, $\% = 100(2\,\sigma/\mu)$

Correlations between these measurements may be of interest, since one might conjecture that a subject with a large head would also have large pinnae. Indeed, this is the basic assumption behind Middlebrooks's procedure for scaling HRTFs to account for changes in body size [14]. In general, there are statistically significant but weak correlations between most pairs of measurements.[6] Scatterplots and correlation coefficients ρ for four interesting examples are shown in Fig. 4.

We focus on the important but difficult to measure pinna di-

[6] For 45 subjects, any magnitude of more than .28 is statistically significant at the 95% confidence level.

Meta Platforms, Inc.
Exhibit 1046
Page 002
Authorized licensed use limited to: New York University. Downloaded on October 25,2022 at 02:04:45 UTC from IEEE Xplore. Restrictions apply.

mensions. Fig. 4a shows that there is a fairly good correlation between pinna height and cavum concha height ($\rho = 0.45$). There is also some correlation between head depth and cavum concha width (Fig. 4b, $\rho = 0.33$). Interestingly, there is not much correlation between these two concha dimensions (Fig. 4c, $\rho = 0.25$). Perhaps more surprising, there is little correlation between head height and pinna height (Fig. 4d, $\rho = 0.16$), and about the same correlation between head height and cavum concha height $\rho = 0.17$. In general, there appears to be relatively little correlation between the sizes of large and small anatomical features, and accurate estimation of pinna dimensions from head and torso measurements is problematic.

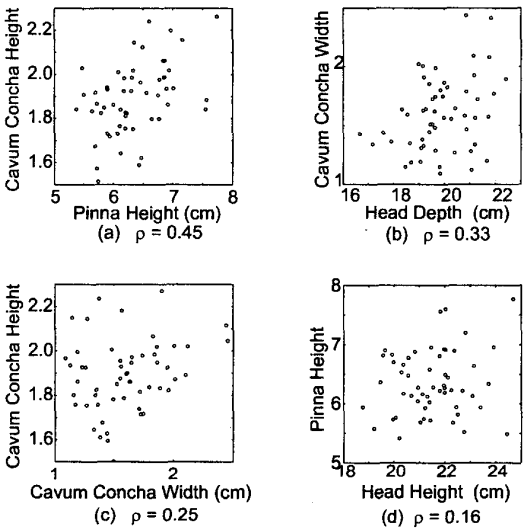

Figure 4: *Selected scatterplots*

4. HRTF VARIATION

One of the advantages of measuring HRTF data at high spatial resolution is that the data can be represented as an image. Fig. 5a shows such an image representation of the impulse response for KEMAR's right ear. Each column in this image is one impulse response at a particular azimuth,[7] with brightness coding the strength of the response. The variation of arrival time with azimuth is clearly seen in the roughly sinusoidal shape of the top envelope of the response. The weakening of the response as the azimuth approaches the opposite side of the head shows the effect of head shadow.

Fig. 5b shows the spectrum for this same case. Here each column is the magnitude of the HRTF in dB, after the power spectrum was smoothed by a constant-Q filter ($Q = 8$). The generally darker appearance of the right half of the image shows the effect of head shadow. The strong response on the ipsilateral side around 5 kHz corresponds to the quarter-wavelength depth resonance identified by Shaw [9], and the weak response around 9 kHz is the so-called "pinna notch." In addition, other interesting but more difficult to explain azimuth-dependent spectral features can be seen.

[7] Of course, interaural-polar azimuth must lie between $-90°$ and $+90°$. For convenience, we show the conventional vertical-polar azimuth in this figure.

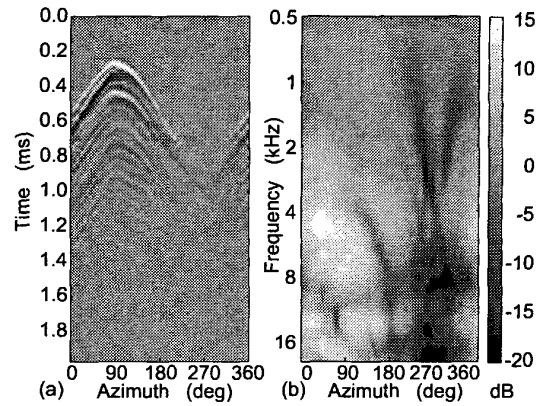

Figure 5: *Horizontal plane (a) HRIR (b) HRTF*

These images give some idea of the HRTF variability for a single subject. It is more difficult to characterize the range of HRTF variability between subjects. However, two simple measures — the maximum interaural time difference $\mathrm{ITD}_{\max}$ and the pinna-notch frequency f_{pn} — are simple, perceptually relevant parameters that characterize the variability that exists.

For the subjects in the database, $\mathrm{ITD}_{\max}$ is approximately normally distributed, with $\mu = 646\,\mu$sec and $\sigma = 33\,\mu$sec, which corresponds to a $\pm 10.3\%$ variation. Not surprisingly, $\mathrm{ITD}_{\max}$ is strongly correlated with head size (see Fig. 6), and it can be estimated quite accurately using simple linear regression. The best single predictor is the head width, with a correlation coefficient of $\rho = 0.78$ between the estimated and the actual ITD. The best pair are the head width and the head depth, for which $\rho = 0.87$. For a more detailed presentation of the estimation of ITD from anthropometry, see [11].

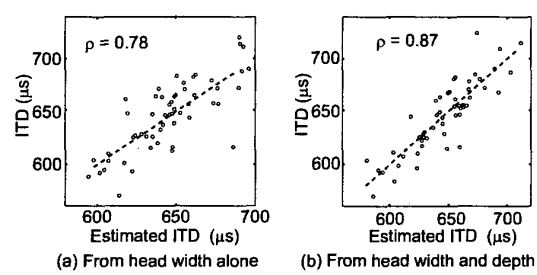

Figure 6: *Scatterplots for estimation of the ITD*

In the frequency-domain, most HRTFs exhibit the prominent depth resonance around 3 to 4 kHz, followed by the pinna "notch" [9]. Fig. 7 shows the HRTF magnitudes for $\theta = \phi = 0°$ for a set of 54 subjects.[8] The pinna notches are indicated by the black dots, and the graphs are sorted by the pinna-notch frequency f_{pn}.

[8] Some subjects used to evaluate statistics have not been included in the database release.

Meta Platforms, Inc.
Exhibit 1046
Authorized licensed use limited to: New York University. Downloaded on October 25,2022 at 02:04:45 UTC from IEEE Xplore. Restrictions apply. Page 003

Statistically, f_{pn} is approximately normally distributed, with $\mu = 7600$ Hz and $\sigma = 1050$ Hz, which corresponds to a rather large $\pm 28\%$ variation. As expected, f_{pn} is correlated with the pinna measurements, but the relationship is not strong, and linear regression is not as successful in estimating f_{pn} from anthropometry. The best single predictor of f_{pn} is the cavum concha height ($\rho = 0.33$). Somewhat surprisingly, the best pair of predictors are the two angles θ_1 and θ_2 ($\rho = 0.42$), and the best triple adds to these the fossa height ($\rho = 0.51$). These results reflect the fact that the scattering of incident waves by the pinna is a complex process related to detailed features, and that accurate estimation of f_{pn} may well require additional concha parameters not included in our measurements. However, simple regression analysis does help identify the most significant of the measured parameters or indicates the need for additional measurements. It our view that the effective customization of HRTFs will requires a deeper understanding of the perceptually important characteristics of the HRTF and of their dependence on detailed pinna features.[9]

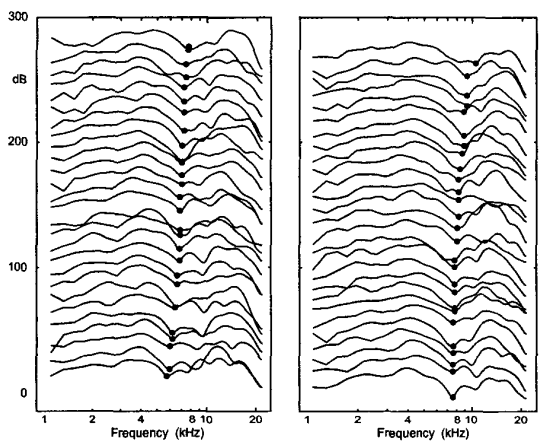

Figure 7: *HRTF magnitudes for* $\theta = \phi = 0°$

5. CONCLUSIONS

High-spatial-resolution HRTF measurements clarify the physical sources of HRTF behavior. A uniform database of HRTFs enables the study of person-to-person differences and of the relation of temporal and spectral characteristics of the HRTF to the anthropometric data. We hope that public availability of the CIPIC HRTF database, augmented with anthropometric measurements, will facilitate further research in the understanding, modeling, and use of individualized HRTFs.

6. ACKNOWLEDGMENTS

This work was supported by the National Science Foundation under grants IRI-96-19339 and ITR-00-86075, and by the University of California DiMI program, with additional support from Aureal Semiconductor, Creative Advanced Technology Center, Hewlett-Packard, and Interval Research Corporation. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation or our other sponsors.

[9] This area of research would benefit from systematic parametric studies using geometric models and computational techniques [3, 4].

7. REFERENCES

[1] Gardner, W. G. and Martin, K. D., "HRTF Measurements of a KEMAR," J. Acoust. Soc. Amer., Vol. 97, 3907-3908, 1995. See also http:// www.sound.media.mit.edu/KEMAR.html.

[2] Wenzel, E. M., Arruda, M., Kistler, D. J. and Wightman, F. L., "Localization Using Non-individualized Head-Related Transfer Functions," J. Acoust. Soc. Amer., Vol. 94, 111-123, 1993.

[3] Katz, B. F. G., "Measurement and Calculation of Individual Head-Related Transfer Functions Using a Boundary Element Model Including the Measurement and Effect of Skin and Hair Impedance," PhD dissertation, Graduate Program in Acoustics, Pennsylvania State University, 1998.

[4] Kahana, Y., Nelson, P. A., Petyt, M. and Choi, S., "Numerical Modeling of the Transfer Functions of a Dummy-head and of the External Ear," Proc. AES 16th Int. Conf. Spatial Sound Reproduction, Rovaniemi, Finland, 1999.

[5] Møller, H., Sørensen, M. F., Hammershøi, D. and Jensen, C. B., "Head-Related Transfer Functions of Human Subjects," J. Aud. Eng. Soc., Vol. 43, 300-321, 1995.

[6] Blauert, J. et. al, "The AUDIS Catalog of Human HRTFs," Proc. 16th ICA, Seattle, WA, 2901-2902, June, 1998. See http://www.eaa-fenestra.de/ documenta/ publications/09-de2.htm

[7] Algazi, V. R., Avendano, C., and Thompson, D., "Dependence of Subject and Measurement Position in Binaural Signal Acquisition," J. Audio Eng. Soc., Vol. 47, 937-947, 1999.

[8] Genuit, K., "Ein Modell zur Beschreibung von Außenohrübertragungseigenschaften", doctoral dissertation, Rheinisch-Westfälischen Technichen Hochschule Aachen, Germany, 1984.

[9] Shaw, E. A. G., "Acoustical Features of the Human External Ear," Gilkey, R. H. and Anderson, T. R., eds., Binaural and Spatial Hearing in Real and Virtual Environments, 25-47, Lawrence Earlbaum Associates, Mahwah, NJ, 1997.

[10] Algazi, V. R., Avendano, C. and Duda, R. O., "Elevation Localization and Head-Related Transfer Function Analysis at Low Frequencies," J. Acoust. Soc. Amer., Vol. 109, 1110-1122, March, 2001.

[11] Algazi, V. R., Avendano, C. and Duda, R. O., "Estimation of a Spherical-Head Model from Anthropometry," accepted for publication in J. Aud. Eng. Soc., 2001.

[12] Møller, H., "Fundamentals of Binaural Technology," Applied Acoustics, Vol. 36, 171-218 (1992).

[13] Tilley, A. R. and H. Dreyfuss Associates, The Measure of Man and Woman, The Whitney Library of Design, NY, 1993.

[14] Middlebrooks, J. C.,"Individual Differences in External-Ear Transfer Functions Reduced by Scaling in Frequency," J. Acoust. Soc. Amer., Vol. 106, 1480-1492, 1999.

Meta Platforms, Inc.
Exhibit 1046
Authorized licensed use limited to: New York University. Downloaded on October 25,2022 at 02:04:45 UTC from IEEE Xplore. Restrictions apply. Page 004