COMPUTED HRIRS AND EARS DATABASE FOR ACOUSTIC RESEARCH

SLIM GHORBAL[1], RENAUD SÉGUIER[1]

(revision 1 - 02/20/2020)

CONTENTS

LIST OF FIGURES

LIST OF TABLES

[1] *IETR/FAST, CentraleSupélec, France*

1 PURPOSE

In the following, The CHEDAR database is described in detail. This database is devolved to the study of the relationships between morphology and binaural listening. It focuses on the influence of the ear.

It is provided under a Creative Commons BY-NC-SA licence, giving you the freedom to redistribute and edit the database for non-commercial purposes if you give credit to the original version and authors. For more information visit http://creativecommons.org/licenses/by-nc-sa/4.0/, and contact the authors for commercial use. If you use this data please reference:

Slim Ghorbal, Xavier Bonjour, and Renaud Séguier. Computed hrirs and ears database for acoustic research. In *Audio Engineering Society Conference: 148th International Conference*, Vienne, Austria, May 2020. Audio Engineering Society.

2 CONTENT

The database is comprised of 1 253 different 3D meshes and their associated computed HRIRs. The meshes are available as binary .ply files and the HRIRs are available as .sofa files. The database also comes with anthropometric measurements for each mesh.

3 DESCRIPTION

3.1 Meshes

All the database meshes derive from the same 3D model of ear, head and torso, whose deformations are controlled by a set of blendshapes. This model (and the meshes created out of it) contains 55 332 vertices and 110 660 facets. Its resolution is finer around the left ear as this is where the virtual microphone is positioned. The edge length is 0.7 mm at the speaker position and grows up to 5 mm, 10 cm away.

Regarding the differences between meshes themselves, choice has been made to focus on the impact of the ear shape only. Thus, the head and torso parameters are left untouched and a set of 7 parameters modifying specifically the ear are used. The table 1 lists them. Their specifications are based on prior knowledge and studies on ear shape influence [1, 2, 3].

Parameter	Nb values	Description
d_3	4	cavum concha width
$d_1 + d_2$	3	cavum concha + cymba height
d_4	3	fossa height
θ_1	3	pinna rotation angle around axis Y
θ_2	3	pinna rotation angle around axis Z
P	2	parabola effect
R	2	crus of helix

Table 1: Parameters used to deform the model.

More in detail, we have retained 5 well-known CIPIC parameters and defined 2 additional ones. These latter are the height of the crus of helix, which can be very prominent for some individuals and referred as parameter R, and a "parabola effect" that leads to protruding ears and named P. The figure 1 illustrates them. Note that P differs from θ_2 as the concha does not move when P changes.

One of the underlying ideas of the database being to better understand the effects of some anthropometric features of the ear, we have chosen to assign a given set of "acceptable values" to each of them and to generate all the possible combinations. By *acceptable values*, we refer to measurements encountered in

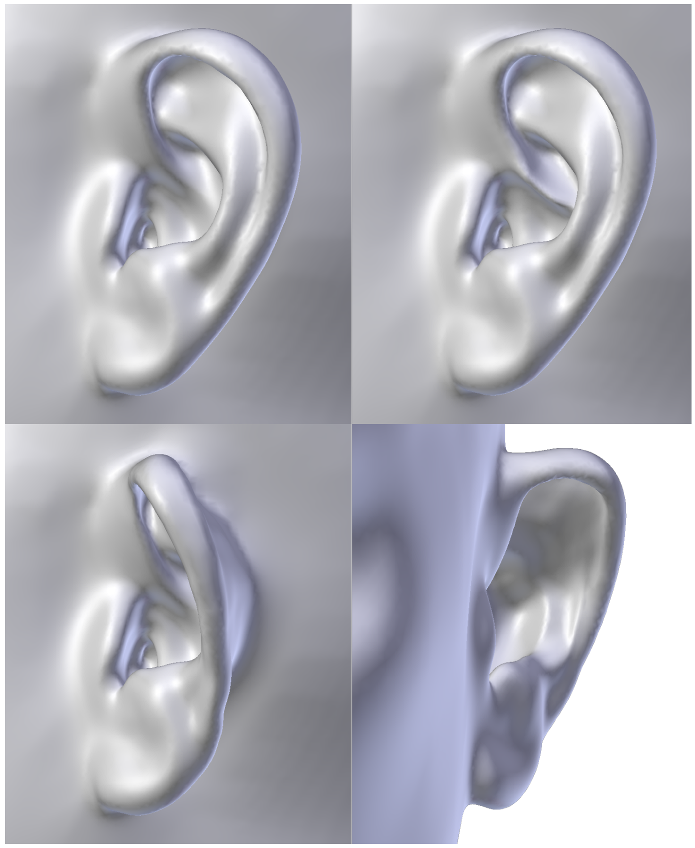

Figure 1: Top left, the base ear untouched. Top right, the base ear deformed by parameter R. Bottom, two views of the base ear deformed by parameter P.

real statistics, such as the one of the CIPIC database. For the parameters P and R, obviously more exotic, the limits have been set based on the personal experience of the team.

3.2 HRIRs

The *mesh2hrtf* [4, 5] software and its implementation of the FM-BEM have been used for the acoustic simulations.

The directions at which HRTFs are computed are defined as a set of 4 evaluation grids. Their respective radii are 2 m, 1 m, 50 cm and 20 cm. The three largest are full sphere comprising 2 522 directions (see figure 2). The smallest one is a subset of 1 801 directions chosen not to intersect with the meshes.

The reciprocity principle being applied, all the grid points are sound receivers. The sound emitter, i.e. the speaker, is located at the end of the left ear auditory canal, whose length is about 5 mm. It consists of a set of 55 adjacent facets. A rigid impedance condition has been imposed on all other facets. Simulations are run for frequencies ranging from 100 Hz to 16 kHz by 100 Hz steps. Right ear data are then obtained by symmetry with respect to the median plane. The resulting HRTFs are then diffuse-field equalised (see figure 3 for an example).

Before packaging the simulation output as .sofa files, a few post-processing steps are applied. First, the 0 Hz spectrum is copied from the 100 Hz data. A first set of HRIRs, sampled at 96 kHz, is calculated and gives access to the time of arrival (*TOA*). Then, the original phase is removed from the DTFs and replaced by the minimal phase deduced from the spectrum amplitude. Afterwards, the TOA is re-injected and a

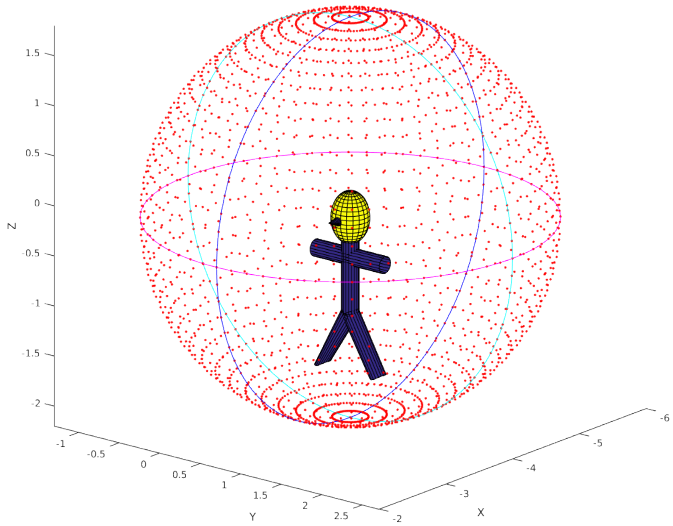

Figure 2: The 2 m evaluation grid. The red dots represent the microphones. The azimuthal plane contains 72 of them.

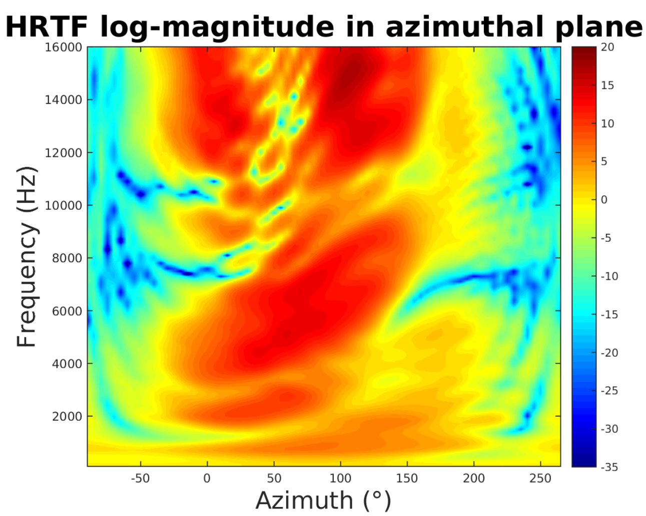

Figure 3: An example of resulting DTF in the azimuthal plane.

second set of HRIRs, sampled at 48 kHz, is calculated. Finally, a volume normalisation and a 4^{th} order Butterworth filter are applied.

3.3 Anthropometric measurements

Taking advantage of the fact that all the meshes share the same vertices and facets numerations, we have selected a set of facets acting as landmarks for automatic anthropometric measurements. Most of our 3D parameters are derived from the CIPIC 2D parameters (except P and R which are totally new) and this new dimension forces us to interpret their initial definitions. For instance, the full concha height, which is given in the CIPIC world by summing the cavum concha and the cymba concha heights ($d_1 + d_2$), no longer satisfies this equation and needs to be measured separately (see parameter $d_1 d_2$ and figure 4).

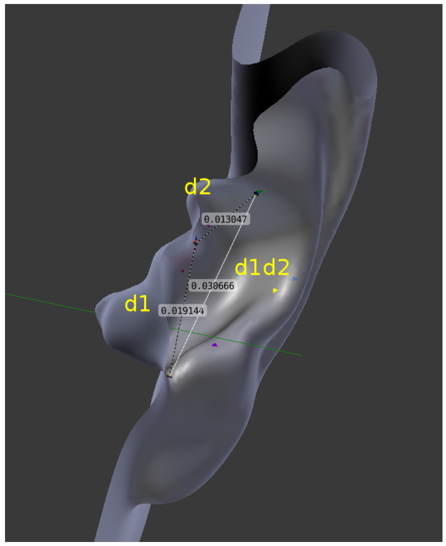

Figure 4: Definition of d_1d_2 versus definitions of d_1 and d_2.

To each landmark is assigned an ID encoding its role in the measurement process and constructed as follow:

{param}_{axis}_{index}

Here *param* refers to the measured parameter name. All the facets sharing the same *param* value are used to measure it.

If *param* refers to a distance, only 2 facets should have an ID starting with it (ex: ***d3**_a* and ***d3**_b* - see figure 5) and we measure the euclidian norm of the vector joining their centroids. If *axis* is specified (ex: *x4_**Z**_a* and *x4_**Z**_b*), we project this vector on *axis* before measuring it. Distances are given in cm.

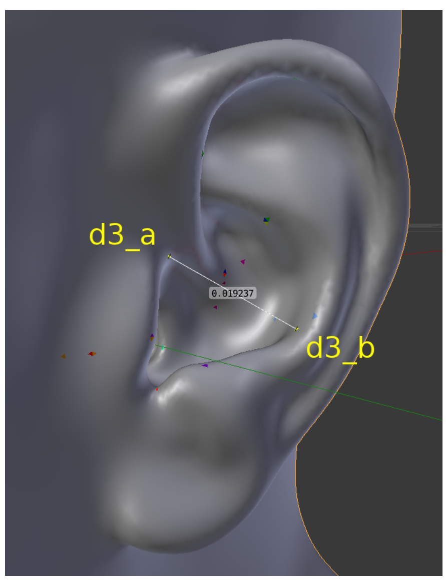

Figure 5: Measure of d_3.

If *param* refers to an angle, 3 facets should have an ID starting with it (ex: ***R**_a*, ***R**_b* and ***R**_c*) and we measure the angle $(\overrightarrow{ba}, \overrightarrow{bc})$ (if we denote a, b and c the facets centroids). If *axis* is specified (ex: *t1_**Y**_a*, *t1_**Y**_b* and *t1_**Y**_c*), we operate in the plane orthogonal to *axis*. Angles are given in degrees. t_1 and t_2 stand for θ_1 and θ_2.

REFERENCES

[1] V Ralph Algazi, Richard O Duda, Dennis M Thompson, and Carlos Avendano. The cipic hrtf database. In *Applications of Signal Processing to Audio and Acoustics, 2001 IEEE Workshop on the*, pages 99–102. IEEE, 2001.

[2] Janina Fels and Michael Vorländer. Anthropometric parameters influencing head-related transfer functions. *Acta Acustica united with Acustica*, 95(2):331–342, 2009.

[3] Slim Ghorbal, Théo Auclair, Catherine Soladié, and Renaud Séguier. Pinna morphological parameters influencing hrtf sets. In *Proceedings of the International Conference on Digital Audio Effects (DAFx), Edinburgh, UK*, pages 353–359, September 2017.

[4] Harald Ziegelwanger, Wolfgang Kreuzer, and Piotr Majdak. Mesh2hrtf: Open-source software package for the numerical calculation of head-related transfer functions. In *22st International Congress on Sound and Vibration*, 2015.

[5] Harald Ziegelwanger, Piotr Majdak, and Wolfgang 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*, 138(1):208–222, 2015.