GUI FOR COMPARING PERCEPTION OF SOUND ADJUSTED BY MEASURED OR MODELED HRTF

F. Rund, D. Štorek, O. Glaser

Faculty of Electrical Engineering

Czech Technical University in Prague, Prague, Czech Republic

Abstract

Using Head Related Transfer Function (HRTF) is a key for creating an effective three dimensional sound. Nowadays, two main methods are used to get a set of HRTF in effective quality – modeling according to basic anthropometrical parameters or direct measurement on particular subject or acoustic manikin. This interface was created for comparing perception of sounds adjusted by modeled or measured HRTF. This project extends our program in [8].

1 Introduction

Creating virtual acoustic spaces, where all sound sources can be positioned in any location is desirable ability for assistive technologies, entertainment, art, or military purposes. Basic access is to use classic stereo systems, when each signal comes to left or right channel with different intensity and / or delay. More advanced system is binaural recording. A system with artificial head with microphones mounted in place of ear channel entrance records the acoustic scene or performance nearby. In this case a separation of left and right ear channels is required, so in practice we have to use headphones. When we then listen to the binaural record, we get almost the same information to our ears like in real situation. A big disadvantage of this system is a need of recording set and the virtual scene can´t be interactive.

Another approach is using HRTF (Head Related Transfer Function) when we filter desired sound with HRTF or left and right channel appurtenant to the specific direction. Equivalent in the time domain is HRIR (Head Related Impulse Response), which is actually the impulse response of path between sound source and ear channel entrance. For left channel it demonstrates Eq. (1).

$$Y_L(\omega) = HRTF_L(\omega) * X(\omega) \qquad \text{resp.} \qquad y_L(t) = HRIR_L(t) * x(t) \tag{1}$$

Main advantage is that we are allowed to make whole acoustic scene only from the monaural recording, so we can (ideally) make the virtual acoustic scene at pleasure only with proper set of HRTFs and chosen bank of sounds. In order to get well perceptible results we have to obtain the set of HRTFs with sufficient dense. One possibility is to measure HRTF in certain points or it can be generated by modeling system.

2 Modelled HRTF

There are many approaches how to model the interaction between propagated sound and main body parts (pinna, head and torso), which affects the final HRTF behavior. Our model is based on system published in [6]. It is based on description of multiple paths, which the sound travels from sound source to listener´s ear. A scheme for left and right ear is depicted on Fig. 1. Both channels work on the same principles.

The simplest case is that sound enters subject´s ear directly with no interaction, but real sources have much more wider directivity, so the same sound strikes different part of body and reflects even to the ear canal entrance. Every reflection in this model is represented by coefficient ρ which determines the ratio between energy of incident and reflected sound, and by coefficient τ representing time delay related to the direct sound.

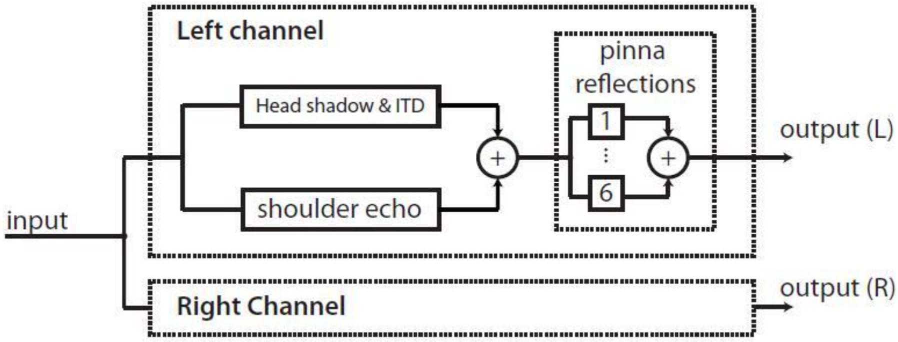

Figure 1: Scheme of HRTF modeling process [3]

Reflection occurs especially on shoulder (shoulder bounce) and pinna, where it is many times multiplied according to unique pinna geometry. When the head shadowing is activated (indirect path between source and ear for contralateral side) even ITD (Interaural Time Difference), which is one of the most important elements for localization in azimuth plane, is set between both channels.

This way we obtain a specific combination of τ and ρ coefficients for all directions. So the final one channel signal consists of sum of the direct sound and it´s attenuated and delayed copies. This process makes unique filtration (HRTF) because of positive and negative frequency interferences caused by its phase shift. Detailed explication of model work, equations, calculation and basic geometry can be found in [1], [3], [6].

3 Measured HRTF

It is not easy to measure a set of HRTF with sufficient quality and density without proper equipment. We decided to use the data from CIPIC HRTF database [4] free available on web. This database includes 1250 channel pairs of measuring for over 40 subjects, which is quite enough for our purposes. These data are stored as Head Related Impulse response in standard CD quality of sample frequency at 44.1 kHz and 16 bits per sample. Each HRIR consists of 200 samples.

4 GUI Manual

After running the m-file ‘localization_start.m’ the preamble window is opened, shown on Fig. 2. As mentioned in chapter 2, ITD is important element for azimuth localization cues and it is derived from subject´s head diameter. User is firstly asked for input own value of head diameter in order to concert written and perceived directions in modeled GUI. Default value is a population average diameter from [5]

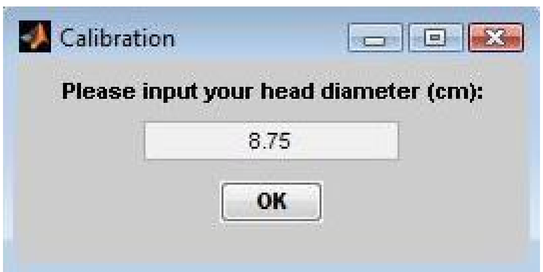

Figure 2: Preamble window for both

After confirmation, the first window offering GUI for modeled HRTF starts. Main objects for creating HRTF for desired direction are two sliders controlling azimuth and elevation. Both values of θ and ϕ are displayed on text field, and it is also possible to enter concrete number in cooperation with slider shifting.

In lower section there is a scheme similar to Fig. 1, which allows user to switch on/off an influence of two main parts of the model - head (shadowing, ITD) or pinna (specific filtration for median plane). In case of switching off the selected part turns red as shown on Fig. 3. This example demonstrates that disabling pinna does not allow to change elevation perception (slider disappears).

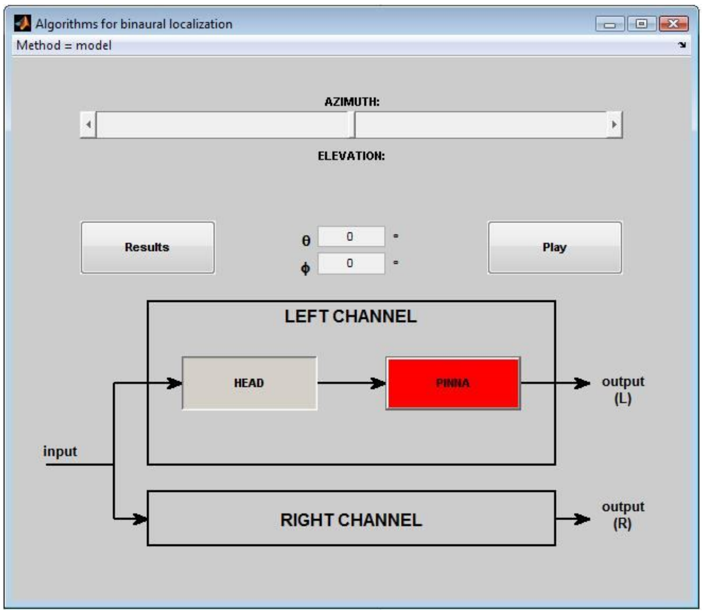

Figure 3: Illustration of the GUI *modeled* HRIR

Modeled and measured method can be switched in upper left corner. When set of measured HRTF is chosen, the lower scheme turns into HRIR length setting, which can be set from 1 to 200 samples. User can also select subject whose HRTF set will be used. GUI of measured variant is presented on Fig. 4.

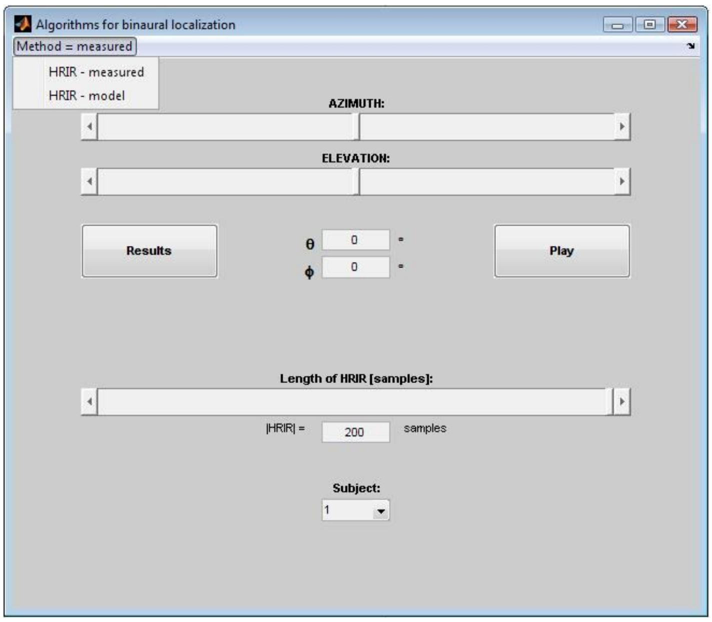

Figure 4: Illustration of the GUI using *measured* HRIR

Sound adjusted by HRIR for selected direction can be played by “Play” push button. For both variants is used a wideband white Gaussian noise stimulus, which is one of the most suitable for localization experiments [4].

The “results” button opens a new window with depiction of selected parameters. First figure shows estimated position in 2D projection of spherical system of coordinates. It is almost the same like a view of subject´s back of the head. The second figure shows HRIR for left and right ear – delay and attenuation for contralateral channel is well noticeable. Two bottom pictures show input noise spectrum and spectrum for left and right ear channel after application of HRTF.

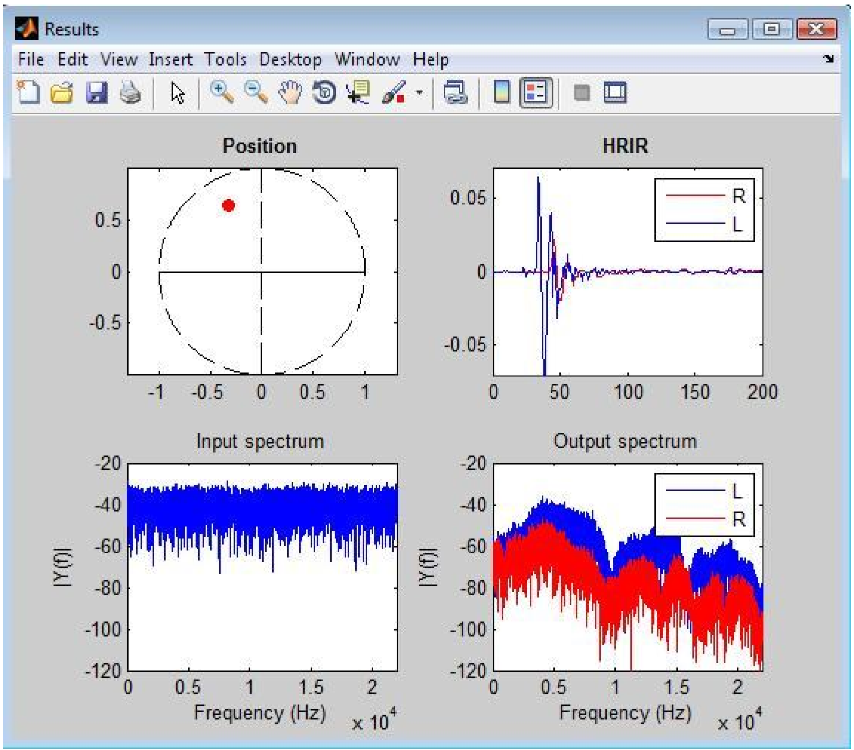

Figure 5: Parameters of adjusted signal shown on demand

5 Results

Graphical user interface for comparing two main methods of obtaining HRTF (measuring and modeling) was created. It allows to the user to verify the final perception of virtually positioned sound for angles $\theta \in < -90°, +90° >$ in azimuth plane and $\phi \in < -90°, +90° >$ in median plane. WGN was used as default sound and for selected directions is available to be played after HRTF application. Parameters of HRIR resp. HRTF can be shown by pressing “Results” button.

This work is an extension of programs published in [8]. We expect the program to be exceeded for real time application to verify the perception of moving sources adjusted by HRTF.

Acknowledgements

The project "GUI For Comparing Perception of Sound Adjusted by Measured or Modeled HRTF" was supported by the Grant Agency of the Czech Technical University in Prague, grant No. SGS10/082/OHK3/1T/13. Presented algorithms use data from CIPIC HRTF database.

References

[1] Duda, R., "*Modeling Head Related Transfer Function,*" Proc. 27th Asilionar Conference of Signals, Systems & Computers, pp. 457-461, Asilomar, CA, 1993

[2] Algazi, R., Divenyi, P., Duda, R., "*Subject Dependent Transfer Functions in Spatial Hearing,*" in Proc. IEEE ICASSP, pp. 877-880, 0-7803-3694-1/97, 1997

[3] Brown, P., Duda, R., "*A Structural Model for Binaural Sound Synthesis,*" IEEE Trans. on Speech and Audio Processing, vol. 6, pp. 476-488, Sept. 1998

[4] Algazi, R., Aveando, C., Thompson, D., Avendano, C., "*CIPIC HRTF Database,*" in Proc. IEEE WASPAA´01, New Platz, NY, 2001, http://interface.cipic.ucdavis.edu/.

[5] Algazi, R., Aveando, C., Duda, R., "*Estimation of a Spherical-Head Model from Anthropometry,*" in J. Aud. Eng. Soc., vol. 49, no.6, pp. 472-478, June 2001

[6] Brown, P., Duda, R., "*An Efficient HRTF Model for 3-D Sound,*" University of Maryland, San Jose State University, 2002

[7] Algazi, R., Duda, R., Duraiswami, R., Gumerov, N., Tang, Z., "*Approximating the Head-Realted Transfer Function Using Simple Geometric Models of the Head and Torso,*" J. Acoust. Soc. Am., vol. 112, pp. 2053-2064, Nov. 2002

[8] Rund, F., Glaser, O., Štorek, D., „*GUI pro demonstraci principů binaurální lokalizace zdrojů zvuku,*" In Technical Computing Prague 2009 [CD-ROM], Praha: Humusoft, 2009, s. 1-5. ISBN 978-80-7080-733-0.

Ing. František Rund, Ph.D.
Department of Radioelectronics, FEE, CTU in Prague, Technická 2, 166 27 Praha 6, Czech Republic
email: xrund@fel.cvut.cz

Ing. Dominik Štorek
Department of Radioelectronics, FEE, CTU in Prague, Technická 2, 166 27 Praha 6, Czech Republic,
email: storedom@fel.cvut.cz