Customization of HRTF using anthropometric features

Yuxiang Wang & You Zhang
Hajim School of Engineering, University of Rochester

Abstract

An HRTF(Head-related Transfer Function) describes how human ears receive a sound from a certain spatial direction. This transfer function is individual and direction dependent and is vital for virtual acoustic display. However, due to its uniqueness, using generic HRTF for virtual acoustic display may result in compromised results, leading to vague or misplaced acoustic images. Thus, it's meaningful to investigate the customization of HRTF from a subject's physical appearance.
In this project, we find certain relations from features in HRTF to the features of a subject's anthropometric details. And furthermore, by taking a few certain measurements such as taking a photo, we are be able to predict a new subject's HRTF.

Introduction

Head Related Transfer Function (HRTF) is a response that characterizes how an ear receives a sound from a point in space. It contains all the acoustic cues that the auditory system will use to make a localization judgment and is generally an important tool for spatial audio reproduction. However, due to its uniqueness, using generic HRTF for virtual acoustic display may result in compromised results, as one's HRTF may differ significantly from the other.

In this project, we propose a practical approach for connecting the features of subjects' physical appearance, to their HRTFs' certain features. An existing HRTF database is used for our study, and some attempts were made to find such connections. In finalized version, we provide a demo of generating one's HRTF using a picture of his/her ears.

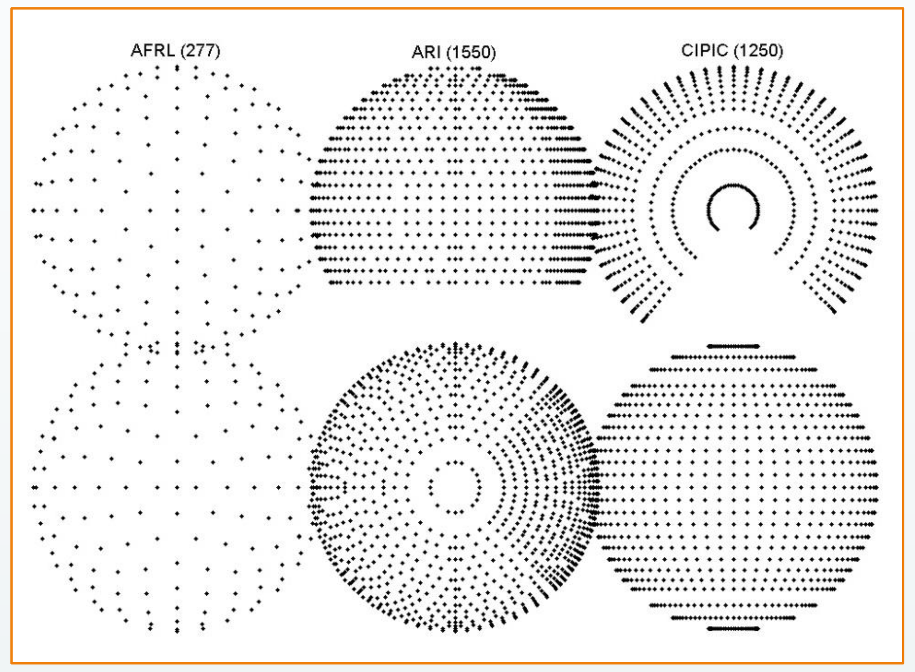
Proposed Methods

First, we choose a proper HRTF database to investigate. We shall choose a database that provides both subjects' HRTF and their corresponding anthropometric features.

Second, we decide the features in the HRTF that is suit- able to formulate relations to.

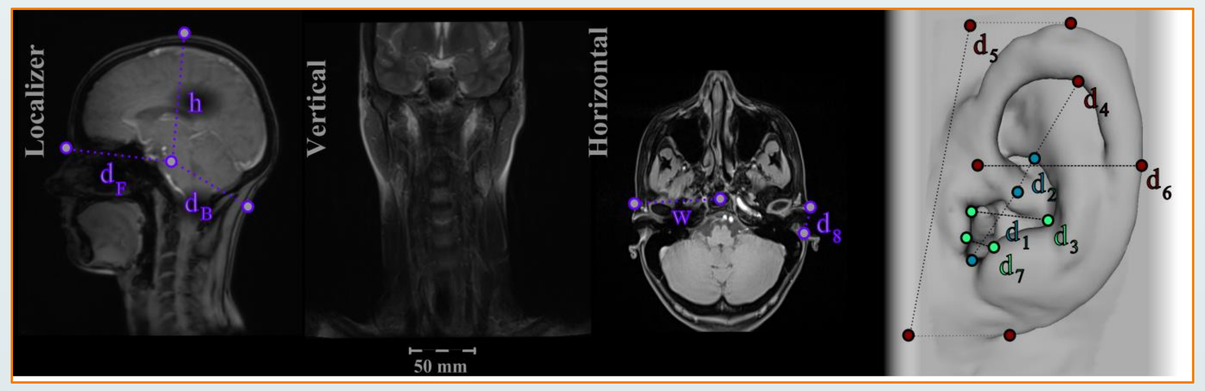
Then, we make attempt to find connections from the subjects' physical features to the features in their HRTF. We try the connections for HRTF at the frontal direction in the beginning, then to move on to the HRTFs at mid plane, then we try to generalize it to the HRTFs at all the directions.

There're quite a few existing HRTF database and their measuring grids:



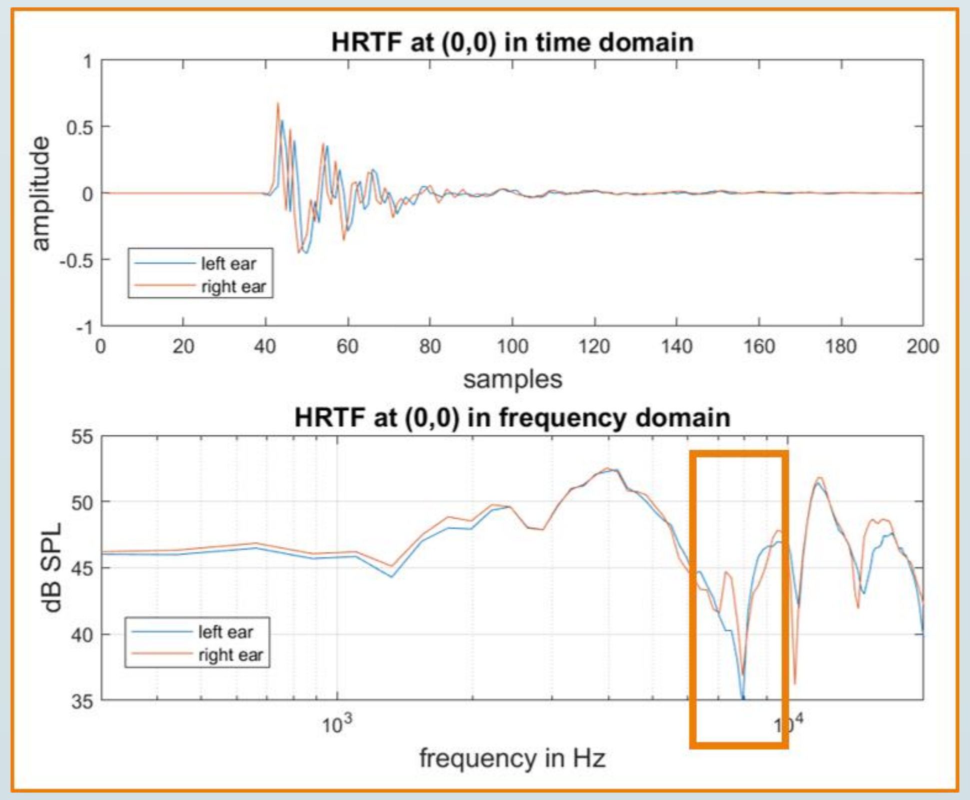
Typical HRTF databases grid layout

For the purpose of this project, we choose CIPIC database(on the right) [1], as it uses interpolar coordinate system, which is more than suitable for studying the connections between physical features to acoustic features, as it panned the source locations vertically. The CIPIC database also provide their corresponding physical measuring data, of the subjects' head and torso, which uses the following measuring standard:



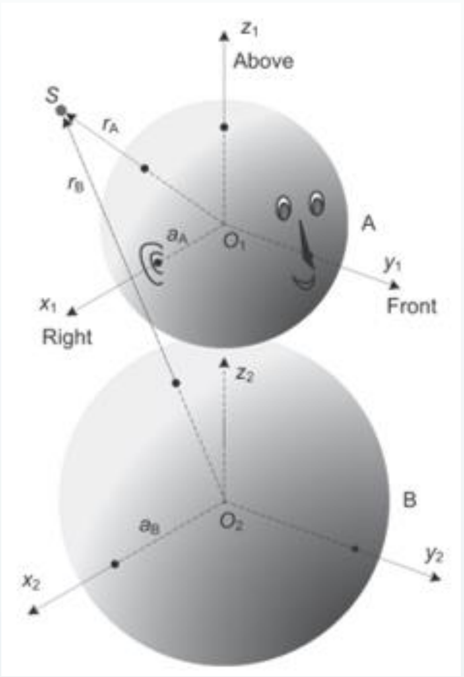
Anthropometric measuring standards

In this project, we choose to use the pinna notches as the features of interest:



Pinna notch existing in HRTF frequency domain

For lower frequency part of the HRTF, we perform the simulation with a physics model called Snowman Model:



Snowman model for analytical calculation of HRTF

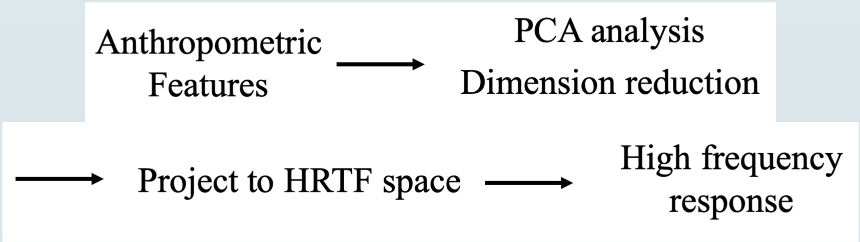
By solving the Helmholtz equation to identify pressure generated by a sinusoidal point source, we can write the sound pressure at the location of the ears in the form of summation of the Spherical Harmonic Expansions:

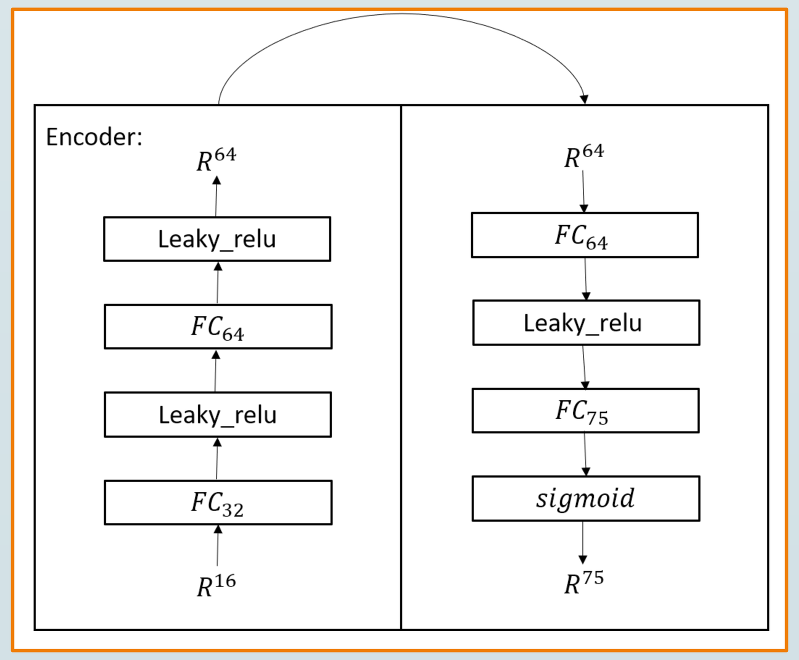
$$\nabla^2 P(\boldsymbol{r}, \boldsymbol{r}', f) + k^2 P(\boldsymbol{r}, \boldsymbol{r}', f) = -jk\rho_0 c Q_0 \delta(\boldsymbol{r} - \boldsymbol{r}')$$

$$P_A(\boldsymbol{r}, \boldsymbol{r}', f) = P_A(\boldsymbol{r}_A, \boldsymbol{r}'_A, f) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} A_{lm}^{A} h_l(kr_A) Y_{lm}(\Omega_A)$$

$$P_B(\boldsymbol{r}, \boldsymbol{r}', f) = P_B(\boldsymbol{r}_B, \boldsymbol{r}'_B, f) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l'} A_{lm}^{B} h_l(kr_B) Y_{lm'}(\Omega_B)$$

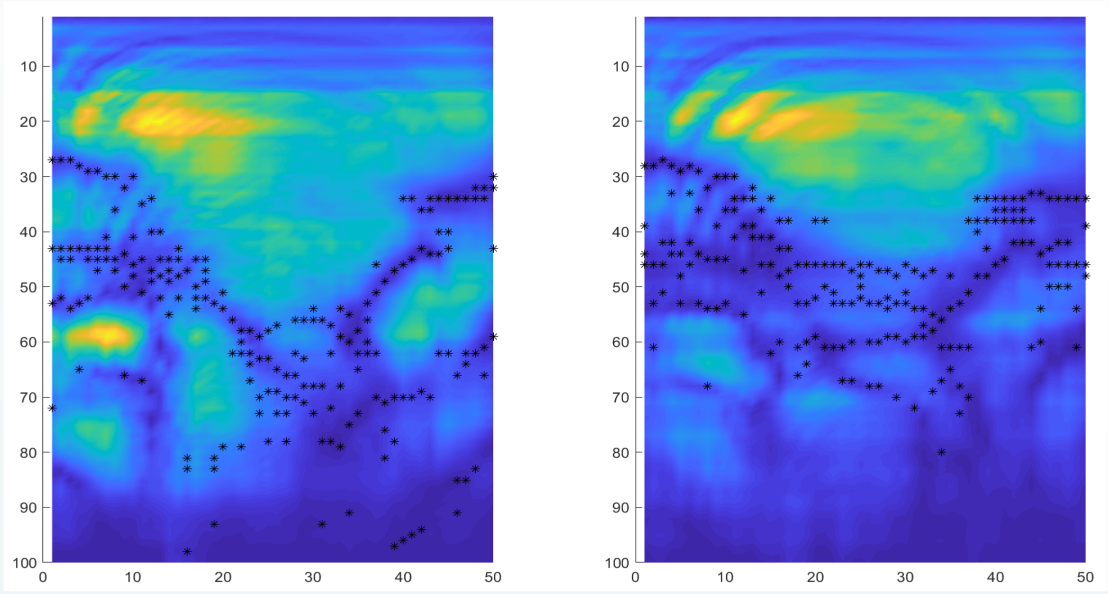
To learn the mapping from anthropometric features to HRTF features in higher frequency, we proposed the following approaches:





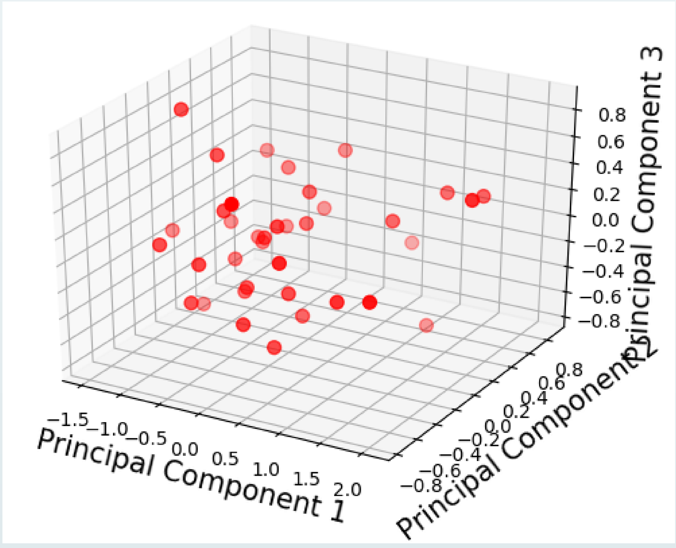
Results

We extracted the pinna notches of a subject's HRTF in the medium plane using the peak detection methods, and used the results for later steps



Pinna notch extraction results of a subject's HRTF in the medium plane

Also, we performed PCA to visualize the principal components of entire CIPIC database, showing a lower dimension representation of all the subject's head and torso measures:



Conclusion

In this project, we made our attempts to build the connections form the subject's anthropometric data to the features in their corresponding HRTF in the frequency domain. This is a comprehensive research problem that combines the backgrounds of both acoustics and machine learning. We are able to extract some information in the frontal direction at this phase, and further investigations in all the directions remains as an interesting future task.

References

[1] V Ralph Algazi, Richard O Duda, Ramani Duraiswami, Nail A Gumerov, and Zhihui Tang. Approximating the head-related transfer function using simple geometric models of the head and torso. The Journal of the Acoustical Society of America, 112(5):2053–2064, 2002.

[2] Pierre Guillon, Rozenn Nicol, and Laurent Simon. Head-related transfer functions reconstruction from sparse measurements considering a priori knowledge from database analysis: A pattern recognition approach. In Audio Engineering Society Convention 125. Audio Engineering Society, 2008.

[3] Abhijit Kulkarni and H Steven Colburn. Role of spectral detail in sound-source localization. Nature, 396(6713):747, 1998.

[4] Griffin D Romigh, Douglas S Brungart, Richard M Stern, and Brian D Simpson. Efficient real spherical harmonic representation of head-related transfer functions. IEEE Journal of Selected Topics in Signal Processing, 9(5):921–930, 2015.