Minimum Representation and Reconstruction of HRTF using Spherical Harmonic Transform

Yuxiang Wang

May 5, 2019

Abstract

In this project, I examined the ideas from the some papers covering the methods of the minimum representation of an Head-Related Transfer Function (HRTF) database, using Spherical Harmonics Transform(SHT) in the frequency domain. By performing experiments of SHT of different orders, I managed to expend an actual HRTF dataset using reduced amount of parameters. The consistency of the original HRTF and the reconstructed HRTF was also examined, and finally a brief demo was created to show the differences of the two.

1 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, not all the information contained by an HRTF is a viable cue to the auditory system , and it's been an interesting topic to find ways to represent an HRTF with minimum parameters. This this project, I followed the methodology proposed by Griffin et al., using Spherical Harmonics Transform(SHT) to expand the HRTF in the frequency domain, and were able to express the original HRTF data in a very reduced scale.

2 Methods: Spherical Harmonics Expansion

2.1 Spherical Harmonics

Spherical Harmonics are the solution to the Laplace equation,

$$\Delta f = \frac{1}{\rho^2}\frac{\partial}{\partial \rho}\left(\rho^2 \frac{\partial f}{\partial \rho}\right) + \frac{1}{\rho^2 \sin\theta}\frac{\partial}{\partial \theta}\left(\sin\frac{\partial f}{\partial \theta}\right) + \frac{1}{\rho^2 \sin^2\theta}\frac{\partial^2 f}{\partial \varphi^2} = 0 \tag{1}$$

which describes the vibration profile of ideal fluids. The spherical harmonic basis of l-th order and m-th degree is computed as following:

$$Y_l^m(\theta,\varphi) = \sqrt{\frac{(2l+1)}{4\pi}\frac{(1-m)!}{(l+m)!}}P_l^m(\cos\theta)\,e^{im\varphi} \tag{2}$$

where $P_l^m(\cos\theta)$ is Associated Legendre Polynomial. The SH base of the first 4 orders are listed as follows:

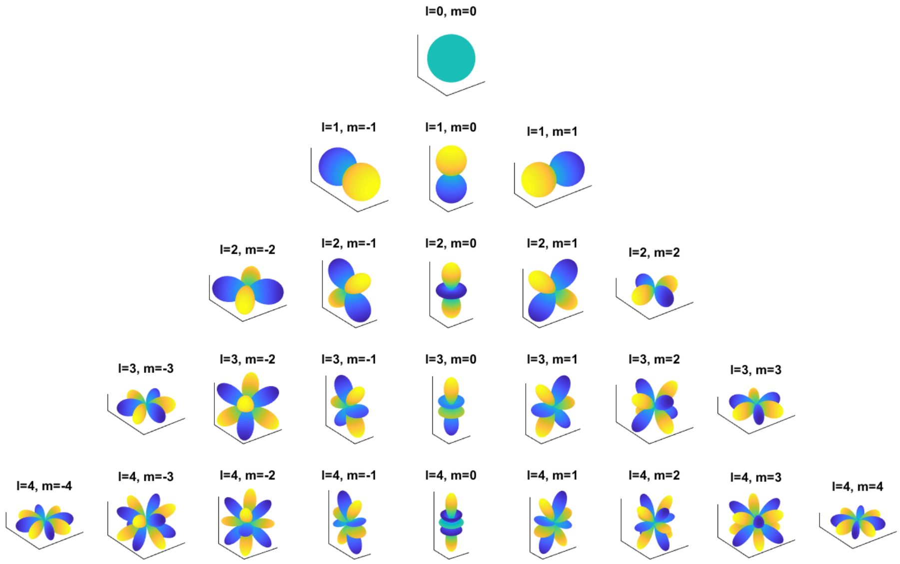

Figure 1: Spherical Harmonics bases up to 4th order

Using all these basis, we are able to expand arbitrary shape that are sampled in spacial locations.For arbitrary grids, the SH coefficients are typically estimated by forming a system of linear equations using the discretized version of S repeated times, one for each spatial location $\{\phi_i, \theta_i\}^S$:

$$\boldsymbol{f} = \boldsymbol{Y}\boldsymbol{c} \tag{3}$$

$$\begin{aligned} \boldsymbol{f} &= [f(\phi_1, \theta_1), \dots f(\phi_S, \theta_S)]^T \\ \boldsymbol{c} &= [C_{00}, C_{1-1}, C_{10}, C_{11}, \dots C_{PP}]^T \\ \boldsymbol{Y} &= [\boldsymbol{y}_{00}, \boldsymbol{y}_{1-1}, \boldsymbol{y}_{10}\boldsymbol{y}_{11}, \dots \boldsymbol{y}_{PP}] \\ & and \\ \boldsymbol{y}_{lm} &= [Y_{lm}(\phi_1, \theta_1), \dots, Y_{lm}(\phi_S, \theta_S)]^T \end{aligned} \tag{4}$$

2.2 Physical meaning of performing SHT to HRTF

If we consider the human head in spherical coordinate system, and due to reciprocity principle, the HRTF could be regarded as the vibration pattern of a speaker mounted on the entrance of the ear canal.

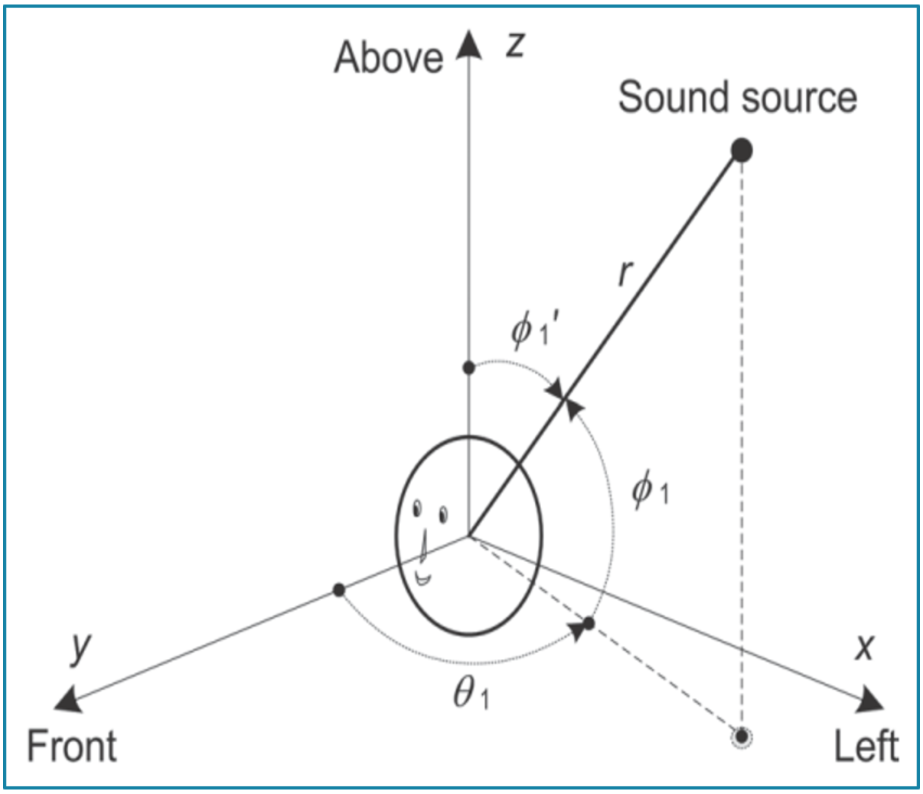

Figure 2: Human head in spherical coord system

For each HRIR measuring grid location, we compute the HRTF in frequency domain, then for each frequency bin in the HRTF, we pick this certain frequency and plot the magnitude layout of each corresponding source location, assigning the magnitudes as the distance to the origin. Right now, one single 3-d shape has the physical meaning of the magnitude layout of a single frequency sound source at the ear canal entrance, in spherical coord system.

In such way, we can expand such 3-dimensional shape(representing frequency-dependent magnitude spatial layout) using SHT, and use the resulting coefficients to reconstruct the original HRTF.

3 SH expansion experiments

Using the similar methods of the two papers, I did some practice using SHT to expand a known shape as well as an HRTF data base

3.1 SHT to a spacial shape

I calculated the SHT coefficients of a sphere-like object, started with lower order:

Then I moved with higher orders, the results shows that SH reconstructed data converges quite consistant to the original shape:

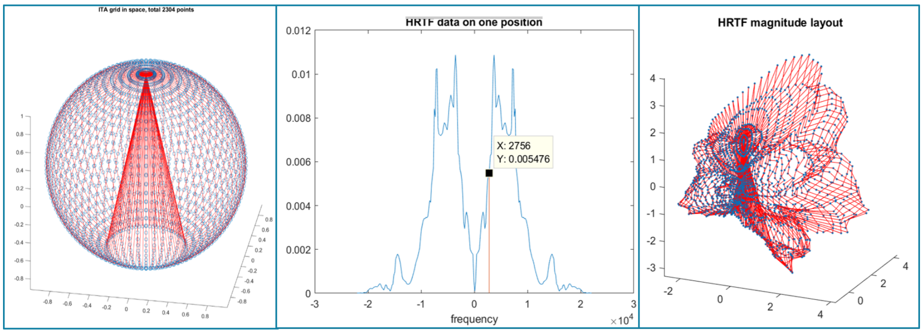

Figure 3: HRTF magnitude layout of one single freqeuency

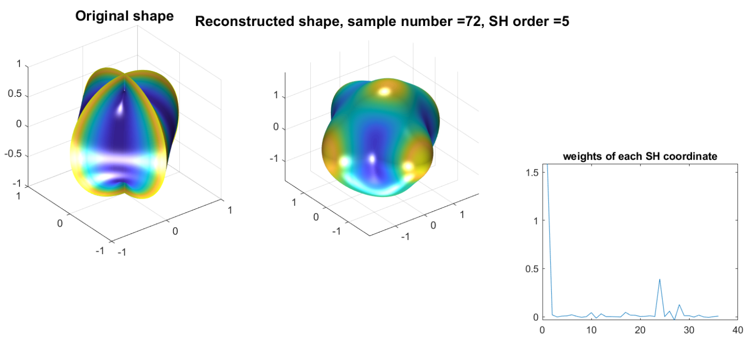

Figure 4: SHT of 5th order

3.2 SHT of an HRTF database

In this actual practice, I followed Griffin et al.'s work, and did experiments on expanding and HRTF at one certain frequency magnitude of 2756 Hz. I started with 5th order SH, and as I moved with higher order(up to 30), the representation is converging nicely to the original data:

3.3 Reconstruction from SHT to orginal HRTF database

Using the same weight matrix C, by performing a simple multiplication we could reconstruct the HRTF at arbitrary spatial angle. Below are the results for performing the reconstruction with different SH orders:

Also a brief demo was created to illustrate the audio externalization difference of performing SHT reconstruction at different SHT orders.

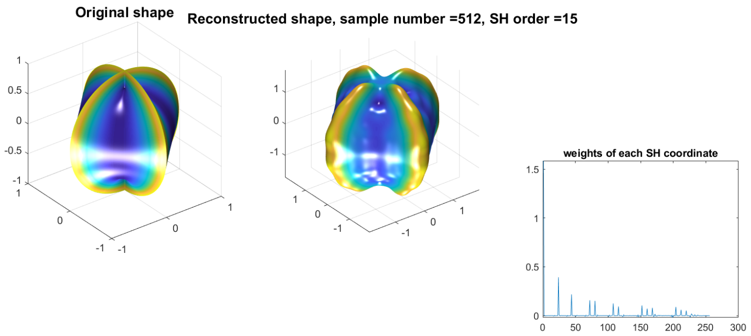

Figure 5: SHT of 15th order

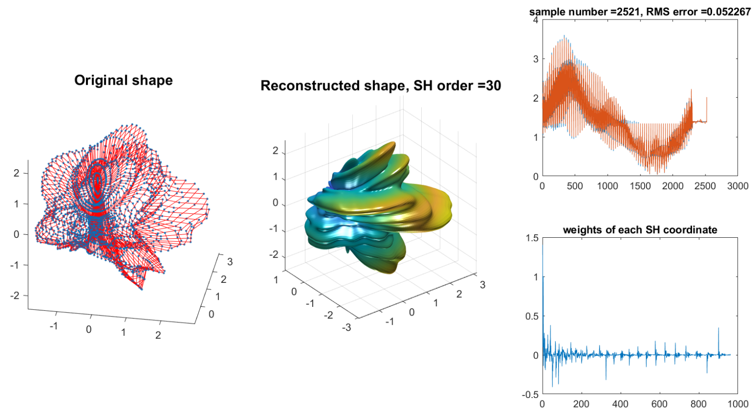

Figure 6: SHT of HRTF at 30th order

4 Discussion

Spherical Harmonics Transform is a useful tool for representing a large scale data that is defined on multiple spatial locations. The minimum parameter required for a perfect localization task could be as low as 5th order SHR, according to the latter paper. In this case, the orginal HRTF data scale is:

$$2304 * 256 * 2$$

Using SHT at 5th order, the total data scale is:

$$36 * 129 * 2$$

which only takes less than 1% of the original data.

Still, it requires more work into how human perceptual system adapt all these information encoded in HRTF.

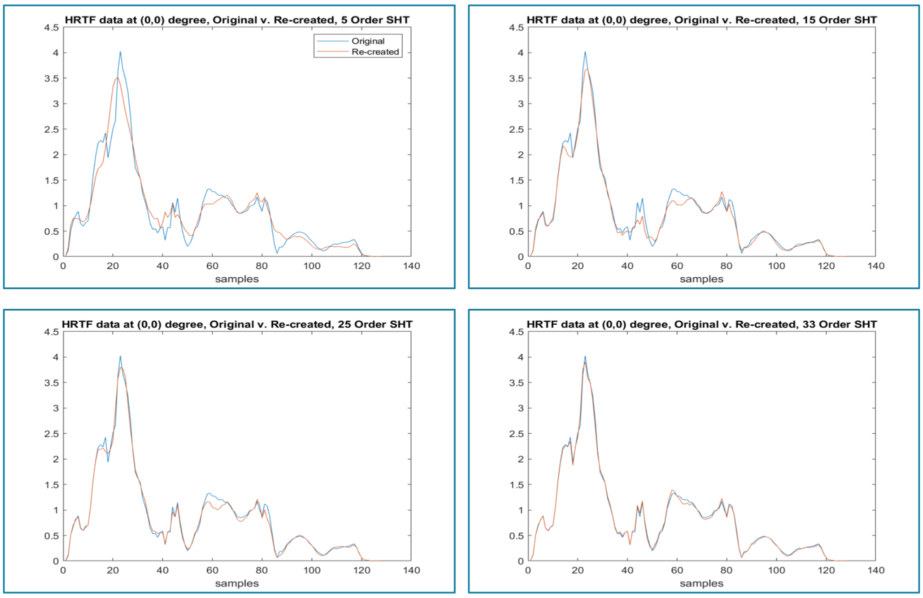

Figure 7: HRTF reconstruction at different orders

References

[1] Evans, M. J., Angus, J. A., & Tew, A. I. (1998). Analyzing head-related transfer function measurements using surface spherical harmonics. *The Journal of the Acoustical Society of America*, 104(4), 2400-2411.

[2] A. Kulkarni and H. S. Colburn, "Role of spectral detail in sound-source localization," *Nature*, vol. 396, pp. 747–749, 1998

[3] Romigh, G., Brungart, D., Stern, R., & Simpson, B. (2015). Efficient Real Spherical Harmonic Representation of Head-Related Transfer Functions. *IEEE Journal of Selected Topics in Signal Processing*, 9(5), 921–930.

[4] M. Aussal, F. Alouges, and B. Katz, "A study of spherical harmonics interpolation for HRTF exchange," *in Proc. Meetings Acoust.*, 2013..

[5] D. N. Zotkin, R. Duraiswami, and N. A. Gumerov, "Regularized HRTF fitting using spherical harmonics," *in Proc. IEEE Workshop Applicat. Signal Process. Audio Acoust. (WASPAA'09)*, 2009, pp. 257–260..

[6] J. Breebaart, F. Nater, and A. Kohlrausch, "Spectral and spatial param- eter resolution requirements for parametric, filter-bank-based HRTF processing," *J. Aud. Eng. Soc., vol. 58, no. 3, pp. 126–140*, 2010..