Minimum Representation and Reconstruction of HRTF using Spherical Harmonic Transform

Yuxiang Wang
Hajim School of Engineering, University of Rochester

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.

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.

Method: Spherical Harmonics Transform(SHT)

Spherical Harmonics are the general solutions to Laplace Equations:

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

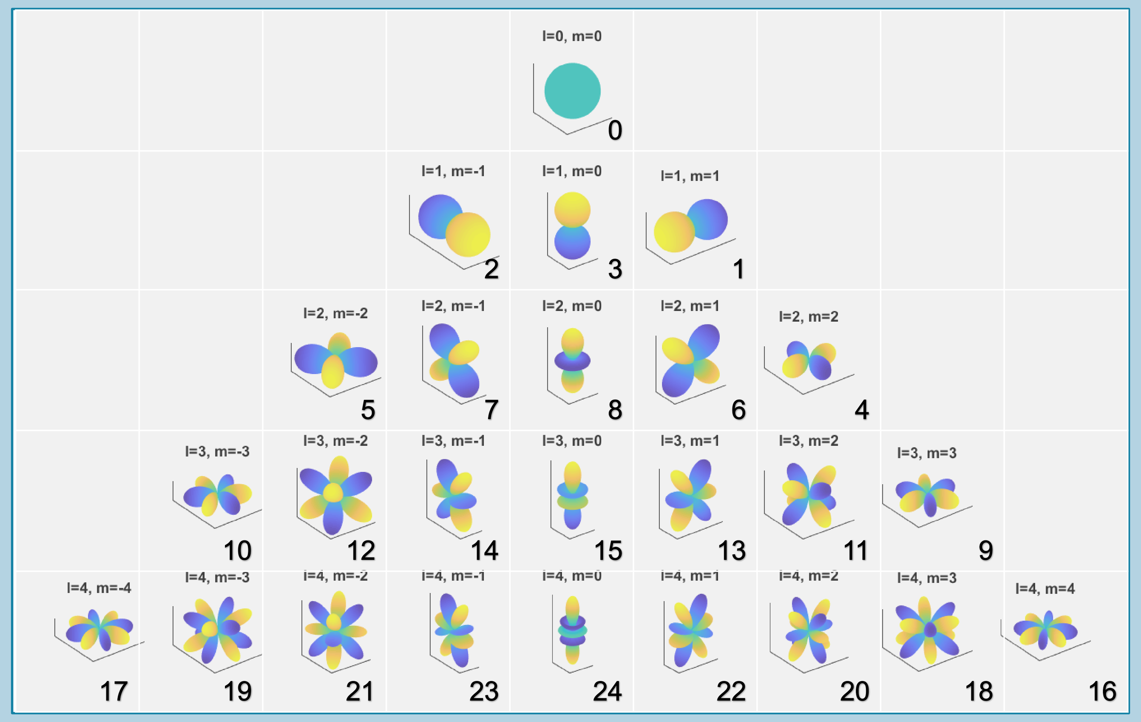
which describes the vibration profile of ideal fluids. The spherical harmonic basis of l-th order and m-th degree is computed as following:

$$Y_\ell^m(\theta,\varphi) = Ne^{im\varphi}P_\ell^m(\cos\theta)$$

Also all the SH bases are orthogonal to each other, which makes them viable bases to expand a 3-dimensional shape.

$$\int_{\theta=0}^{\pi}\int_{\varphi=0}^{2\pi} Y_\ell^m\, Y_{\ell'}^{m'\,*}\, d\Omega = \delta_{\ell\ell'}\,\delta_{mm'},$$

The first few orders of Spherical Harmonics bases are as follows:

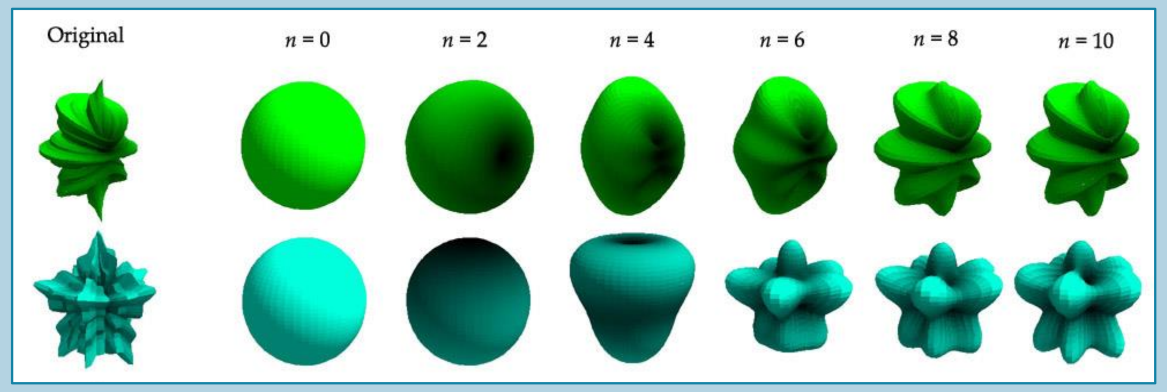


Similar to the Fourier Series, we could use the these bases to expand a certain shape in space, using least square fit approach:

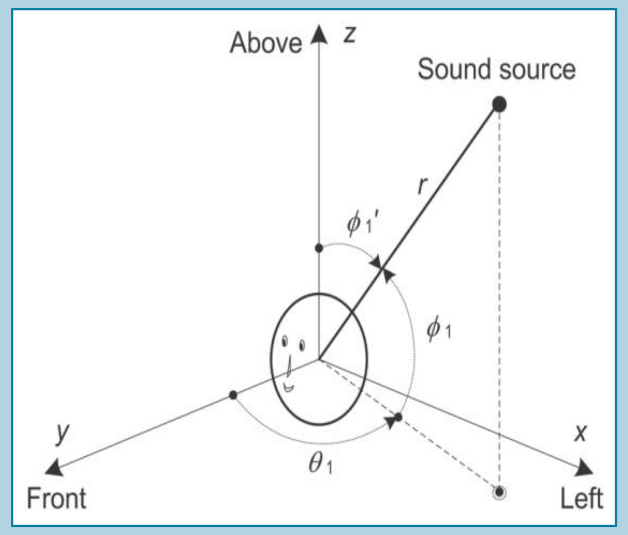
$$f = Yc$$

$$f = \left[f(\phi_1,\theta_1),\dots f(\phi_S,\theta_S)\right]^T$$
$$c = \left[C_{00}, C_{1-1}, C_{10}, C_{11}, \dots C_{PP}\right]^T$$
$$Y = \left[y_{00}, y_{1-1}, y_{10} y_{11}, \dots y_{PP}\right]$$
$$y_{lm} = \left[Y_{lm}(\phi_1,\theta_1),\dots, Y_{lm}(\phi_S,\theta_S)\right]^T$$

$$\hat{c} = \left(Y^T Y\right)^{-1} Y^T f$$

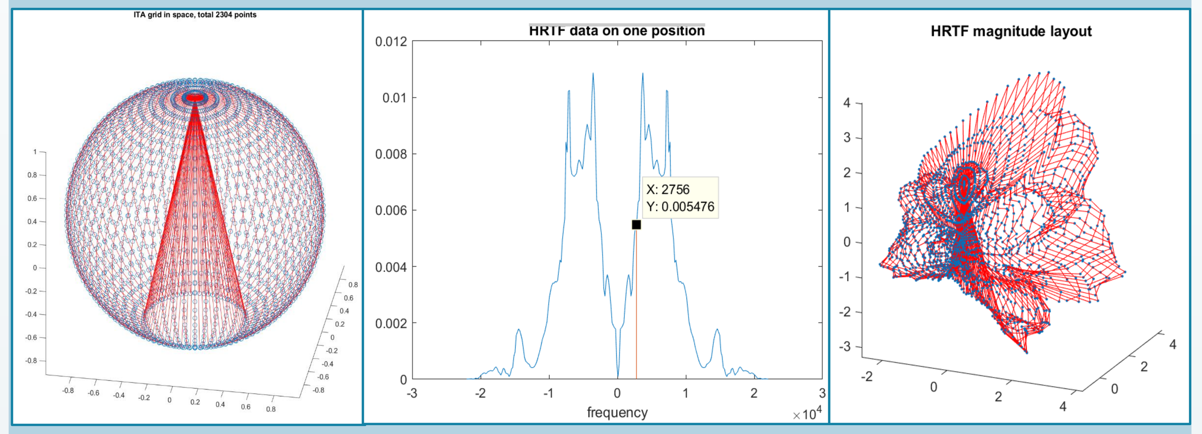


If we consider the human head in spherical coordinate system:



In this case, SHT methods make good physical meaning in the sense of representing the different frequency components of the spatial layout of an HRTF dataset. Thus we could use it to reduce the data scale of HRTF.

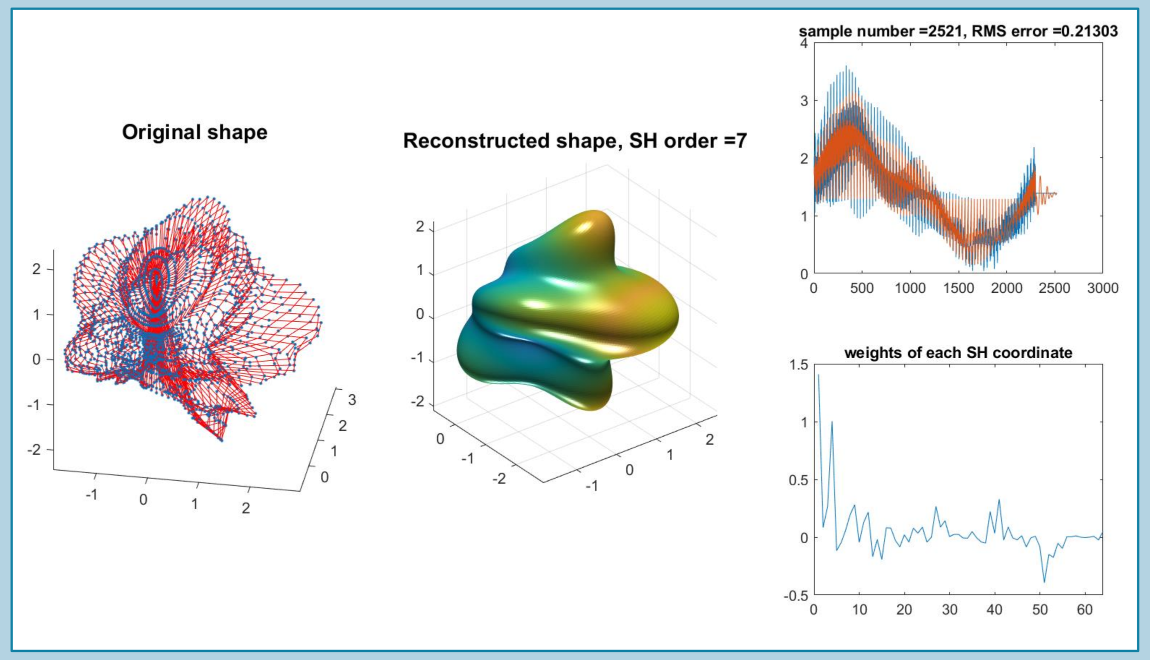
For each HRIR measuring grid location, we compute the HRTF in frequency domain, then for each frequency bin in the HRTF, we pick this certain frequency and plot the magnitude layout of each corresponding source location, assigning the magnitudes as the distance to the origin.

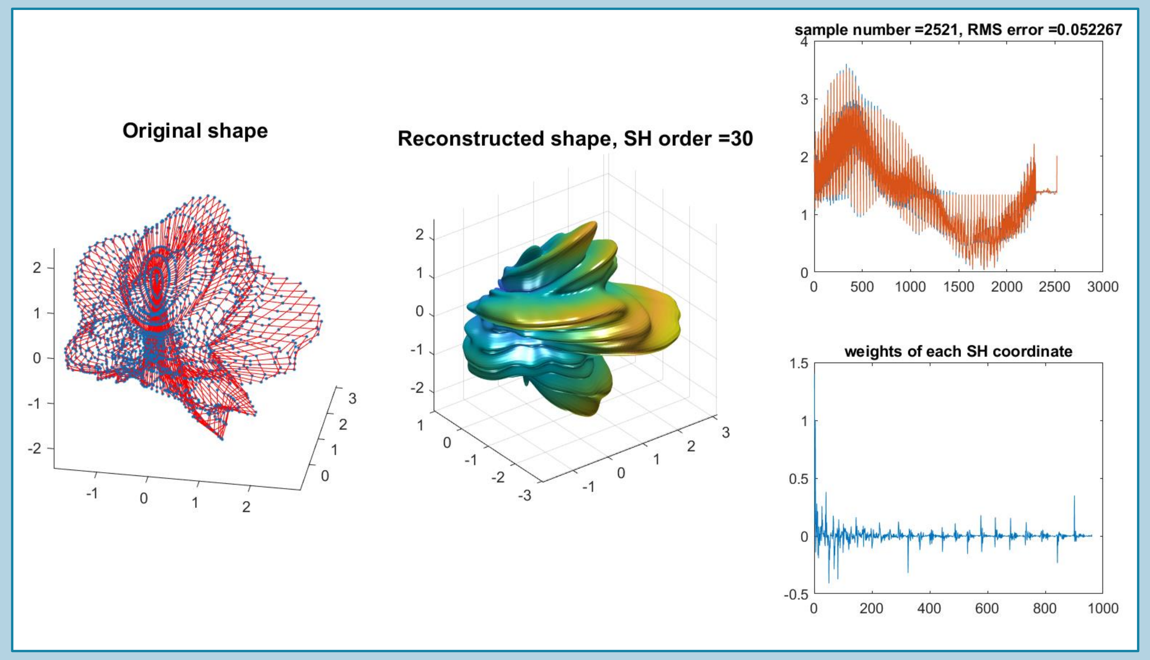


In such way, we can expand such 3-dimensional shape(representing frequency-dependent magnitude spatial layout) using SHT, and use the resulting coefficients to reconstruct the original HRTF.

Results

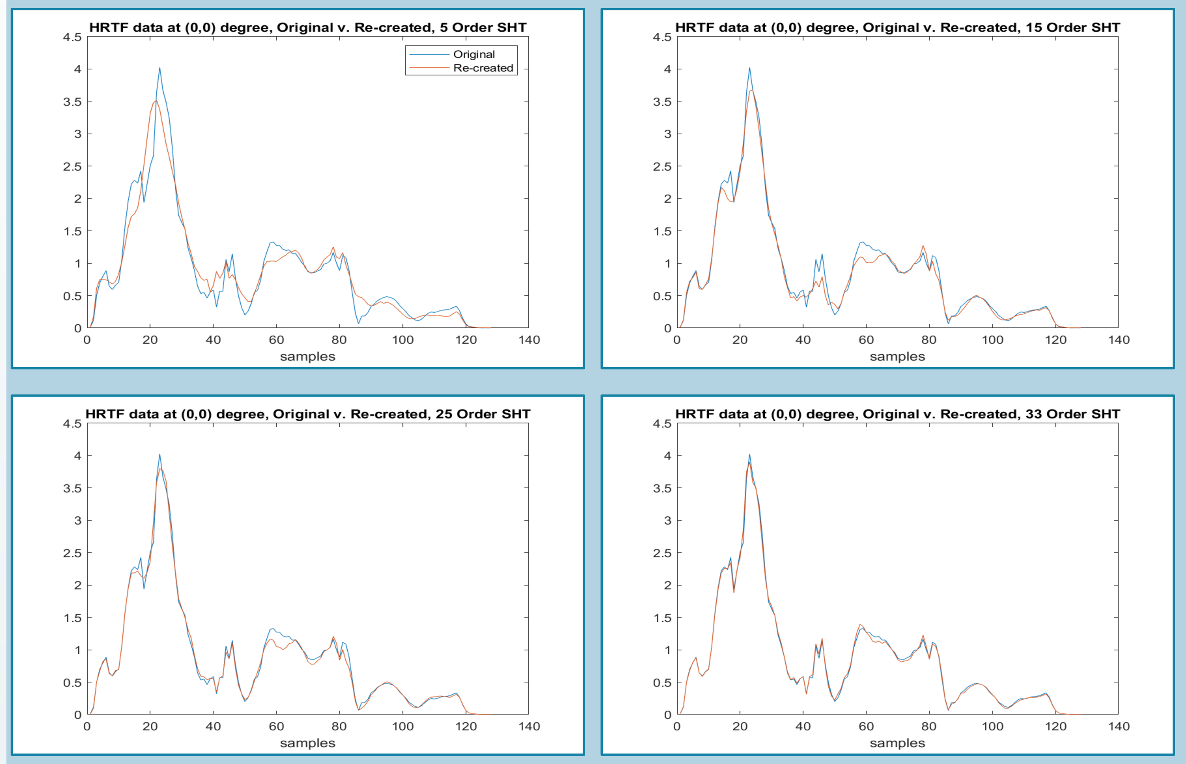
By applying this method, we performed SHT to the original HRTF database in frequency domain, at different orders:





As order increases from low to high, the reconstructed HRTF shows lower RMS errors.

Further more, it's meaningful to examine the reconstructed HRTF after performing SHT methods to each frequency bin.



In order to represent the original HRTF in whole, the data scale required are:

- **Original HRTF data scale:**
  $2304 * 256 * 2$
- **L-th order SHT data scale:**
  $(L+1)^2 * (128+1) * 2$
- **For L = 5, the total data scale is**
  $36 * 129 * 2$

In literature, Griffin et al. found that using SHT up to order L=5, human subjects could perform perfect localization tasks. The SHTed data scale is less than 1% of the original HRTF.

Conclusion

Spherical Harmonics Transform(SHT) in frequency domain makes an efficient way to represent an HRTF database in minimum parameters. We could achieve almost identical reconstruction of the original data using far less data scale. Still, more perceptual tasks remains to be performed to evaluate the efficiency of this method.

References

- 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.
- 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.P. Guillon and R. Nicol, “Head-related transfer function reconstruction from sparse measurements considering a priori knowledge from data- base analysis: A pattern recognition approach,” in Proc. 125th Conv. Audio Eng. Soc., 2008.. Bazinga!
- A. Kulkarni and H. S. Colburn, “Role of spectral detail in sound-source localization,” Nature, vol. 396, pp. 747–749, 1998
- 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.