4pPP4

Principal Components Analysis Interpolation of HRTF's Using Locally Chosen Basis Functions

Jacob W. Scarpaci and H. Steven Colburn

Hearing Research Center and Dept. of Biomedical Engineering, Boston University

Boston University Hearing Research Center

Binaural Hearing Lab

Motivation

To evaluate PCA as a means for representing HRTF data sets and compare locally chosen and globally chosen methods of reconstruction.

Principal components analysis (PCA), when applied to an HRTF data set, has advantages in the amount of data stored and how it can be manipulated. We will talk about PCA in the context of spherical spatial interpolation methods, although we will not show data from interpolated HRTFs. It has been shown that proper spatialization can occur with a reduced number of basis functions [1]. We propose a variation of this method that shows improvements in certain aspects of HRTF reconstruction.

Principal Components Analysis

Performed on log Magnitude Directional Transfer function $D(\omega, \theta_k, \phi_k)$. All phase characteristics are dealt with separately.

BF_i	Basis functions where i is the global ranking	W_i	Weight of BF_i at a given position.
H_g	HRTF reconstructed with globally chosen BFs	H_l	HRTF reconstructed with locally chosen BFs
Q_0	Mean function of all the HRTFs in the data set.	N	Number of HRTFs in the data set.
M	Number of BFs used in reconstruction.	S	Number of BFs required to be stored.
θ_k	Azimuth for position k.	ϕ_k	Elevation for position k.

The directional transfer functions ($D(\omega, \theta_k, \phi_k)$) contain all of the directional characteristics of the signal for each of the measured positions.

$$D(\omega, \theta_k, \phi_k) = log_{10}(|H(\omega, \theta_k, \phi_k)|) - Q_0(\omega) \quad (1)$$

Since PCA is based on the common variance in the data set, the first step is to calculate the covariance matrix.

$$\Lambda_{DD} = E\{DD'\} = \frac{1}{N}\sum_{k=1}^{N} D(\omega, \theta_k, \phi_k) D(\omega, \theta_k, \phi_k)' \quad (2)$$

The eigenvectors of the covariance matrix serve as the orthogonal set of basis functions BF_i. The eigenvalues serve as a global ranking of associated basis functions. The weights ($W_i(\theta_k, \phi_k)$) used to reconstruct each position are calculated by projecting the DTFs onto the basis functions.

$$W_i(\theta_k, \phi_k) = BF_i(\omega) D(\omega, \theta_k, \phi_k) \quad (3)$$

One can use M basis functions where $M < N$ to get a reconstruction whose error is dependent on M.

$$log_{10}(|\hat{H}(\omega, \theta_k, \phi_k)|) = Q_0 + \sum_{i=1}^{M} W_i(\theta_k, \phi_k) BF_i(\omega) \quad (4)$$

Globally chosen basis functions: Select the M basis functions that contributed the most to the entire data set by selecting those that have the largest eigenvalues. *This method may not be the most efficient transfer of common variance at a given position.*

Locally chosen basis functions: Select the M basis functions that minimize the error in the reconstruction of a specific HRTF. *A basis function that contributes significantly over the whole data set may not have a significant contribution in the point being reconstructed.* This must be done for each HRTF by selecting the basis functions that have the largest weights (W_i) associated with them in the reconstruction. For spatial interpolation only the weights associated with the measured HRTFs closest to the position are considered.

Overlap of Chosen Basis Functions

Large percentage of positions use locally chosen basis functions that are not in the top M globally ranked basis functions.

- How many of the locally chosen reconstructed HRTFs (H_l) use basis functions that are not in globally chosen reconstructed HRTFs (H_g)?
- How many basis functions must be stored for the H_l reconstructions?

	BFs (M)	H_l contains BFs not in H_g	Stored BFs (S)
CIPIC 2500	3	90.8%	25
	5	93.5%	38
	10	99.6%	60
SOW 1008	3	91.8%	32
	5	96.9%	50
	10	99.9%	68
SDO 288	3	80.9%	16
	5	91.7%	24
	10	99.7%	42

Table 1: Analysis of three data sets; CIPIC [2] with 2500 reconstructions, SOW and SDO [3] with 1008 and 288 reconstructions, respectively.

Computation and Storage Costs

Amount of storage (numbers) needed for 5-basis function (M=5) reconstructions of the entire CIPIC data set. For H_g reconstructions S=M.

$$\frac{\text{measured}}{\text{reconstructed}} = \frac{\overbrace{2500}^{filters} \times \overbrace{200}^{pts/filter}}{\underbrace{200 \times S}_{BFs\ stored} + \underbrace{2500 \times S}_{weight\ vector}}$$

measured	global (S=5)	local (S=38)
500,000	13,500	102,600

For the CIPIC data set H_l reconstructions still has a storage advantage over the measured data set as long as $S \leq 185$.

Computational load for averaging 4 spatial points in real time. This does not take into account the computational load to create the basis functions, only to use them.

Measured: $\overbrace{200 \times 4\ multiplies}^{scale\ filters} + \overbrace{200 \times 4\ adds}^{sum\ filters}$

Global: $\overbrace{S \times 4\ multiplies + S \times 4\ adds}^{create\ weights} + \overbrace{200 \times M\ multiplies}^{scale\ BFs} + \overbrace{200 \times M\ adds}^{sum\ BFs}$

Local: $L_G + \overbrace{S \times M}^{sort\ weights}$ (Depends on algorithm)

measured	global(S=5,M=5)	local(S=38,M=5)
800/800	1020/1020	1152/1152/190

Table 2: Amount of operations required (multiplies/adds/sort). Example shows CIPIC database for M=5

Depending on real-time constraints, difference in computational loads may be inconsequential.

Reconstructions

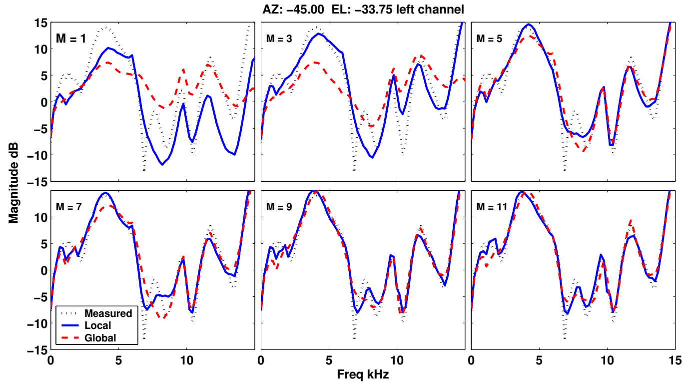

Figure 1: Sample reconstructions using different numbers of basis functions. The right channel of position (-45,-33.75) is shown.

- Frequencies weighted equally in the creation of the BFs.
- *RMS error may not capture errors in spectral features*
- Although overall RMS error will be reduced when adding BFs, the error at a given frequency may go up

Numerical Error Analysis

There are positions that show several dB of improvement using locally chosen basis functions.

We calculate the RMS error between the reconstructions (r) and the measured (m) functions as follows where reconstruction can be either global (g) or local (l). **No frequency weighting.**

$$E_r = \sqrt{\frac{1}{N}\sum (20 * log_{10}(H_m) - 20 * log_{10}(H_r))^2} \quad (5)$$

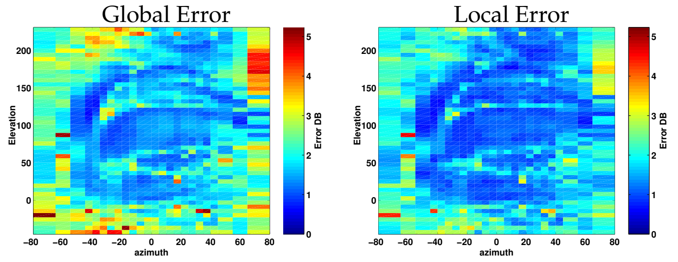

Figure 2: RMS error of reconstructed functions vs measured functions. Shown for M=5 as a function of azimuth and elevation.

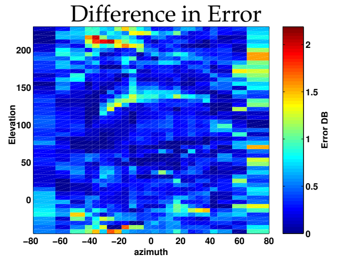

Figure 3: Difference of RMS errors ($E_l - E_g$) shown for M=5 as a function of azimuth and elevation.

Psychophysics

- Non-frozen Gaussian noise lowpass filtered at 14 kHz.
- Measured phase used for both H_m and H_r.
- 4I2AFC Discrimination task between measured and reconstructed filters.

H_m H_r H_m H_m or H_m H_m H_r H_m

- 4 trained subjects 100 trials at each position for different values of M .
- Position (0,0) chosen due to its small RMS error, small error difference, and its common use as a standard position for psychophysics.
- Position (-45,-33.75) chosen due to its large global error with relatively small local error at small values of M.

Results

- Correlation between psychophysical results and RMS error not apparent.
- As number of BFs used in reconstruction increases the ability of the subject to discriminate between the measured and reconstructed filters drops towards chance.
- *Discriminability reaches chance quicker for the locally chosen reconstructions.*

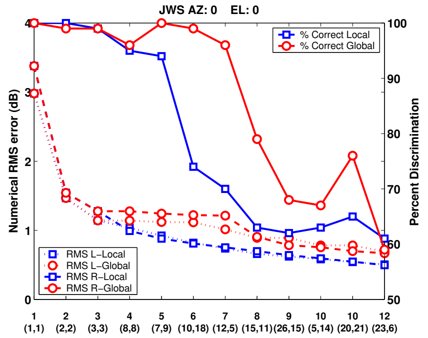

Figure 4: Solid Curves: Discrimination performance for local (blue) or global (red) reconstructions as a function of M. Chance performance indicates better HRTF representation. Dashed and Dotted Curves: Local (blue) or global (red) RMS error for right (dashed) and left (dotted) channels as a function of M. Numbers below graph indicate the global rank of the basis functions that were added to the local reconstruction for left and right HRTFs.

- *Subjects show a psychophysical advantage for using locally chosen BFs for reconstructions at both spatial positions.*
- Performance advantage changes with position.
- Number of BFs needed in the reconstruction is dependent on spatial position.

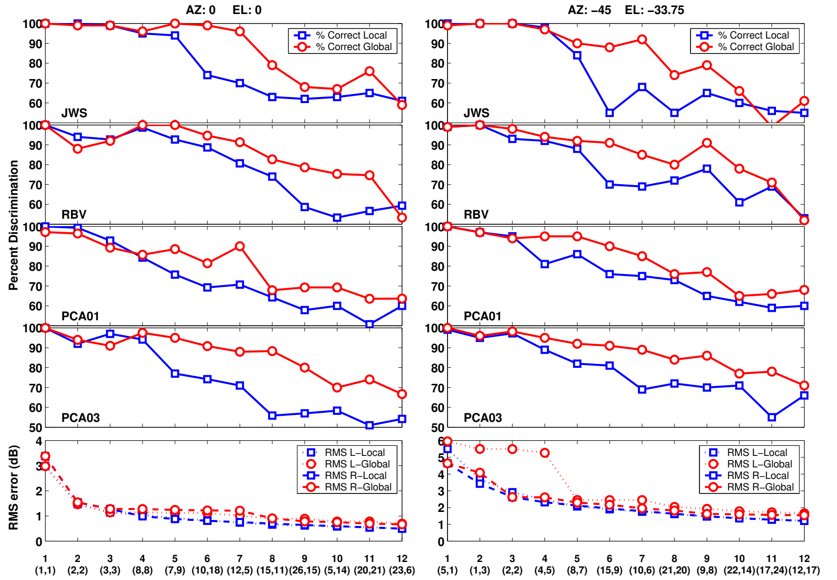

Figure 5: Individual subject performance and RMS error for two positions at different values of M.

Summary and Conclusions

- There are spatial positions where the locally chosen method gives better results numerically and psychophysically.
- The relationship between numerical error and psychophysical performance is not apparent.
- Changes in location cause changes in numerical error and psychophysical performance.

Future Work

- Continue discrimination task at additional positions.
- Psychophysical evaluation using localization task.
- Create BFs with frequency-weighted cost functions.
- Run PCA analysis on smoothed HRTF.
- Split analysis into ipsi, contra, and medial data sets.

References

[1] D. J. Kistler and F. L. Wightman. A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction. *J. Acoust. Soc. Am.*, 91(3), March 1992.

[2] V. R. Algazi, R. O. Duda, D. M. Thompson, and C. Avendano. The CIPIC HRTF database. In *Proc. 2001 ieee workshop on applications of signal processing to audio and electroacoustics*, pages 99–102, 2001.

[3] A. Kulkarni, S. K. Isabelle, and H. S. Colburn. Sensitivity of human subjects to head-related transfer-function phase spectra. *J. Acoust. Soc. Am.*, 105(5), May 1999.

Acknowledgments

Supported by NIDC(DC00100). Special thanks to Reggie Villacorta and Ian McNulty for their contributions.