

# A method to evaluate the individual differences of head-related transfer functions based on spatial principal component analysis

Bosun Xie, Yang Liu

Acoustic Lab., Physics Dept., South China University of Technology, Guangzhou, 510641, P.R.China

## Abstract

Head-related transfer functions (HRTFs) are functions of sound source position and frequency. They are also individual-dependent. A direct way to evaluate the differences in HRTFs among different subjects is to compare the HRTFs for each pair of subjects at all sound source positions. But this method is complicated. The present work proposed a simplified method to evaluate the individual differences in HRTFs based on spatial principal component analysis (SPCA). Firstly, HRTFs were decomposed as a weighted combination of common spatial basis functions (CSBFs), which are only direction-dependent; while the weights are frequency and individual-dependent. Then, the individual differences of HRTFs were evaluated by calculating the cross-correlation of the weights for each pair of subjects. Moreover, the HRTFs of 52 subjects were employed to validate the proposed method.

## Introduction

In the far field source distance ( $r > 1.0$  m), head-related transfer functions (HRTFs) are functions of source position (including azimuth  $\theta$  and elevation  $\phi$ ) and frequency  $f$ . They are also individual-dependent. One of the important applications of HRTFs is virtual auditory display (VAD)[1]. It is well known that the subjective performance of a VAD varies depending on the similarity between the HRTFs used for synthesizing binaural signals and those of the listener [2]. It is thus necessary to evaluate the difference or similarity in HRTFs among different subjects. A direct but complicated approach is to compare the HRTFs for each pair of subjects at all sound source positions. In the present work, a simplified method to evaluate the individual differences of HRTFs based on spatial principal component analysis (SPCA) was proposed.

## Methods of Analysis

The SPCA is different from the traditional principal component analysis (PCA) in that it uses the common spatial basis functions instead of the common spectral functions to decompose the HRTFs [3][4]. By using the SPCA, the HRTF of a given ear  $e$  (left or right) of subject  $s$  at direction  $(\theta, \phi)$  is expanded as a linear combination of spatial basis functions:

$$H(\theta, \phi, f, s, e) = \sum_q d_q(f, s, e) W_q(\theta, \phi) + H_{av}(\theta, \phi) \quad (1)$$

Where the common spatial basis functions (CSBFs)  $W_q(\theta, \phi)$  and mean spatial function  $H_{av}(\theta, \phi)$  are only direction-dependent. Since HRTFs are roughly left-right symmetric, the HRTFs for one ear can be handled as that for another ear after the azimuthal reflection, which results in the HRTFs at

both ears sharing the same set of  $W_q(\theta, \phi)$  and  $H_{av}(\theta, \phi)$  with frequency, ear and individual dependent weights  $d_q(f, s, e)$ . Accordingly, the similarity of the HRTFs between two subjects can be evaluated by calculating the cross-correlation of the weights in SPCA expansions for the two subjects.

Assume that there are  $2SM$  HRTF magnitude spectra from the two ears of  $S$  subjects and  $M$  source directions for each subject. Each HRTF magnitude spectrum is represented by its values at  $N$  discrete frequencies. Subtracting the mean (across the frequency, subjects and two ears) from each magnitude spectrum, the data form a  $(2N S) \times M$  zero mean matrix  $\mathbf{H}$ . And then an  $M \times M$  covariance matrix  $\mathbf{R}$  can be constructed by  $\mathbf{R} = \mathbf{H}^T \mathbf{H}$ , where the superscript "T" denotes the transpose of the matrix. Then the common spatial basis vectors (CSBVs), which represent the values of CSBFs at  $M$  discrete directions, can be obtained from the eigenvectors  $\mathbf{W}_q$  of matrix  $\mathbf{R}$  for the preceding  $Q$  largest and positive eigenvalues  $\lambda_q$ :

$$\mathbf{R} \mathbf{W}_q = \lambda_q \mathbf{W}_q \quad (2)$$

And then the weights for each subject  $s$  and ear  $e$  at  $N$  discrete frequency are evaluated by:

$$d_q(f_n, s, e) = \sum_{m=0}^{M-1} H_{\Delta}(\theta_m, \phi_m, f_n, s, e) W_q(\theta_m, \phi_m) \quad (3)$$

The  $H_{\Delta}(\theta_m, \phi_m, f_n, s, e)$  and  $W_q(\theta_m, \phi_m)$  are elements from matrix  $\mathbf{H}$  and vector  $\mathbf{W}_q$ , respectively.

The contribution of the  $q$ th CSBVs to the percentage of variance of energy is proportional to the eigenvalues  $\lambda_q$ . And the cumulative percentage variance of energy represented by preceding  $Q$  CSBVs is evaluated by:

$$\text{var} = \left( \sum_{q=1}^Q \lambda_q / \sum_{q=1}^M \lambda_q \right) \times 100\% \quad (4)$$

The cross-correlation between the  $q$ th weight of subjects  $i$  and  $j$  at a given ear  $e$  is

$$c_{ij}(q, e) = \frac{\sum_{n=0}^{N-1} [d_q(f_n, i, e) - \bar{d}_q(i, e)][d_q(f_n, j, e) - \bar{d}_q(j, e)]}{\sqrt{\sum_{n=0}^{N-1} [d_q(f_n, i, e) - \bar{d}_q(i, e)]^2 \sum_{n=0}^{N-1} [d_q(f_n, j, e) - \bar{d}_q(j, e)]^2}} \quad (5)$$

The  $\bar{d}_q$  is the mean of weight  $d_q(f_n, s, e)$  across  $N$  discrete frequencies. And the similarity or difference of HRTF magnitudes between subjects  $i$  and  $j$  at a given ear is evaluated by the similarity coefficient

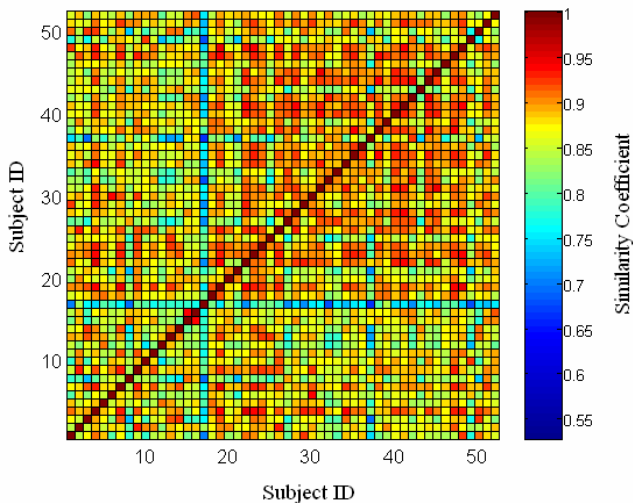
$$C_{ij}(e) = \sum_{q=1}^Q \lambda_q c_{ij}(q, e) / \sum_{q=1}^Q \lambda_q \quad (6)$$

By definition,  $0 \leq |C_{ij}| \leq 1$ , and a larger  $C_{ij}$  means a better similarity.

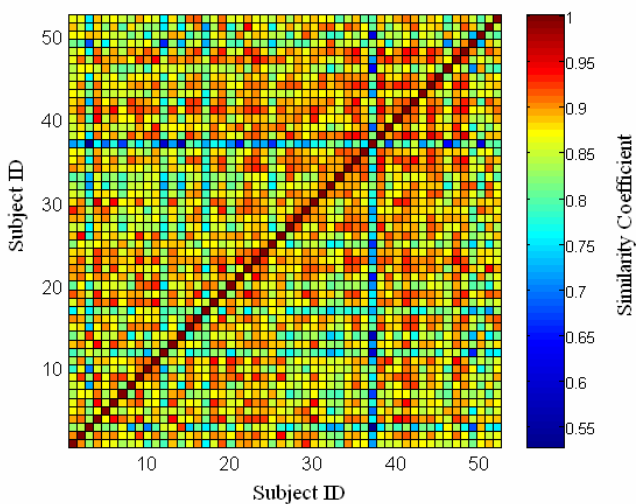
## Results and Discussion

The HRTF data used contain 52 Chinese subjects (26 male and 26 female) with 493 directions for each subject at the sample frequency of 44.1 kHz [5]. The original HRTF magnitudes are firstly smoothed by a moving frequency window with 1 ERB width, resulting in HRTF magnitudes at  $N = 42$  discrete frequencies which represent a frequency range from 0 to 14 kHz with an uniform interval.

The calculation from Eq.(4) shows that the cumulative percentage variance of energy represented by preceding  $Q = 15$  CSBVs reaches 95.9 %.



(a) The right ear



(b) The left ear

**Figure.1:** The similarity coefficient

Figure 1(a) demonstrates the similarity coefficient  $C_{ij}$  for the HRTF magnitudes at right ear between each subjects pair. The similarity varies across the subject pair. Between subjects 4 and 22, the  $C_{ij}$  reaches a maximum of 0.954.

While between subjects 17 and 37, the  $C_{ij}$  exhibits a minimum of 0.667. The mean  $C_{ij}$  across all the subject pairs is 0.867. It is interesting that the HRTF magnitudes of subject 17 is distinctly different from that of the other 51 subjects, and all the  $C_{ij}$  between subject 17 and other subjects are less than 0.842, with a mean value 0.761 across other 51 subjects.

Figure 1(b) demonstrates the results for the left ear. It is roughly similar to that of the right ear. Between subjects 38 and 48, the  $C_{ij}$  reaches a maximum of 0.953. While between subjects 3 and 37, the  $C_{ij}$  exhibits a minimum of 0.528. The mean  $C_{ij}$  across all the subject pairs is 0.861. The HRTF magnitudes of subject 37 is distinctly different from that of the other 51 subjects, and all the  $C_{ij}$  between subject 37 and other subjects are less than 0.848, with a mean value 0.744 across other 51 subjects. In addition, the difference in the details between Figure 1(a) and Figure 1 (b) arises from the asymmetry of HRTF magnitudes between the right and left ears, especially at the high frequency above the 5 kHz [6].

## Conclusion

The proposed method is valid in evaluating the similarity or individual difference of HRTFs. The future works include testifying the results by psychoacoustic experiments and classifying data by multidimensional scaling analysis [7].

## Acknowledgements

The authors acknowledge the support of National Nature Science Fund of China Grant No.10774049 and State Key Lab of Subtropical Building Science, South China University of Technology. The author also thanks Dr. Xiaoli Zhong for her help.

## Reference

- [1]. Xie, B.S.: Head-Related Transfer Function and Virtual Auditory (in Chinese). National Defense Industry Press, Beijing, 2008.
- [2]. Wenzel E.M., et.al: Localization using nonindividualized head-related transfer functions. J. Acoust. Soc. Am. 94(1993), 111-123.
- [3]. Kistler D.J, Wightman F.L.: A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction. J. Acoust. Soc. Am. 91 (1992), 1637-1647.
- [4]. Xie B.S., Spatial principal components analysis on head-related transfer functions and individualized customization from few measurements (abstract). J.Acoust.Aoc.Am. 126(2009), 2242.
- [5]. Xie B.S., et.al: Head-related transfer function database and analysis. Science in China Series G 50(2007), 267-280.
- [6]. Zhong X.L., Xie B.S.: Spatial symmetry of head-related transfer function. Chinese Journal of Acoustics 26 (2007), 73-84.
- [7]. Wightman F., Kistler D.J.: Multidimensional scaling analysis of head-related transfer function, Proceeding of IEEE 1993 WASPA, 98 – 101, 1993.