24th INTERNATIONAL CONGRESS ON SOUND AND VIBRATION
23–27 July 2017, London

On the Use of Partial Least Square Regression for Spatial Fitting of Head-Related Transfer Functions

Wang Lei
Northwestern Polytechnical University, School of Marine Science and Technology, Xi'an, China
email: sunny_w@mail.nwpu.edu.cn

Zeng Xiangyang
Northwestern Polytechnical University, School of Marine Science and Technology, Xi'an, China

Yan Liang
Northwestern Polytechnical University, School of Marine Science and Technology, Xi'an, China

The Head-Related Transfer Functions (HRTFs) are spatially continuous while the traditional acoustic measurements are not available to meet this demand. In this paper, we present a spatial fitting method for HRTFs on the basis of a few discrete measurements. We first make the HRTF matrix and the corresponding direction matrix normalized. Then we build an algorithm model between the normalized direction and HRTF matrices with partial least square regression, and determine the number of principle component by cross validation. After that, the regression was implemented on each of their principal components. Finally, the regression equation was recovered according to the inverse process of normalization. We also verified the effectiveness of this method through calculating the similarity between original data and predicted data. The experimental results show that our method is available in spatial fitting of HRTFs.
Keywords: binaural hearing, HRTFs, PLSR, spatial fitting

1. Introduction

The study on binaural hearing has been increasingly noticed, which includes two responses of human hearing system to sound events--subjective attributes and spatial attributes. As for the spatial attributes of sound events such as the distance and location, the head-related transfer functions (HRTFs) is a vital cue. HRTF is the transfer function between the human ears and sound source,and it describes how sound wave is filtered by the pinna, head, and torso of a listener as it propagates from the source to the ear drums in free space[1-3].

As for usual, experimental measurement in limited directions is the most direct way to obtain HRTFs in discrete directions. However, HRTFs are spatially continuous in free space. That means, the measurements are not available to meet the spatial continuity of HRTFs. Thus, making the discrete HRTFs spatially continuous is a significant work in related research. Commonly, HRTFs of unmeasured directions can be inferred from those of the measured directions with interpolating calculating such as adjacent linear interpolation, bi-linear interpolation, trigonometric interpolation, etc. Generally, the interpolation accuracy increases with the descending of sample interval. Meanwhile, the interpolation effect depends on the interpolation methods, frequency range, spatial direction, and the HRTF data.

Previous research efforts attempted to fit HRTFs spatial continuously with different interpolations. On the HRTF database of VALDEMAR artificial head with spatial sample interval of 2°, Christensen[4] analyzed the interpolation effectiveness of HRTF minimum-phase with different elevation angles. The results show that the high frequencies correspond to larger errors. Nishino[5] compared the effects of adjacent linear interpolation and cubic spline interpolation of the minimum-phase of HRTFs in the vertical plane. The results show that cubic spline interpolation obtains better results when the sample interval is lager and homogeneous while the adjacent linear interpolation is more effective when the sample interval is well selected. Wightman [6] studied the bi-linear interpolation of personalized HRTFs through psychological experiments of sound image localization. The outcomes show that the HRTF interpolation in the condition of minimum-phase approximation acquires considerable localization results with which of the actual measurements. On this foundation, Wenzel[7] studied the bi-linear interpolation of non-personalized HRTFs with sound image localization experiments. The results indicate that bi-linear interpolation of HRTFs in both direct and minimum-phase approximation conditions gain less localization errors.

In this paper, we propose a new method to fit HRTF continuous in space on the base of a few discrete measurements. Our results show that the unmeasured HRTFs of spatial direction can be inferred effectively with PLSR algorithm.

2. Database Description

In this paper, the database contains 56 male subjects aged from 25 to 45 with their anthropometric features and the Head-Related Impulse Responses (HRIRs, the denotation of HRTF in time domain) in 1190 directions. All the HRIRs are 800-sample long with the sampling frequency at 44.1 kHz, meanwhile the measurement distance was fixed at 1.5 meter.

3. Partial Least Square Regression and Spatial Continuous HRTFs

3.1 Partial Least Square Regression

Partial Least Square Regression is an expansion statistical method of classical Least Square Regression aiming for finding a linear combination from dependent variables to explain the variation information of independent variables. In this paper, we model the spatial location of azimuth and elevation with HRTF phase and magnitude respectively to construct the continuous model of HRTF in space.

As for spatial points, each one can be described with a coordinate (θ, φ), in which θ stands for the azimuth and φ stands for the elevation. The dependent variable matrix $X_{N\times 2}$ consists of N different spatial direction and the independent variable matrix $M_{N\times L}$ and $P_{N\times L}$ consist of the corresponding magnitudes and phases of HRTFs, where L stands for the length of samples in the matrices.

The first step of PLSR is to normalize the dependent and independent matrices. The normalization process is conducted as the equation below:

$$x_{ij}^{*} = \frac{x_{ij} - m_j}{M_j - m_j} \tag{1}$$

Where x_{ij}^{*} stands for the normalized value of x_{ij}, M_{ij} and m_{ij} stand for the maximum and minimum value of x_{ij} respectively. The normalized matrices are X_0, M_0 and P_0.

The second step is to construct the internal and external relationships between dependent variables and independent variables, and implement the regression of X_0 and M_0 (X_0 and P_0) on the first component:

$$\begin{aligned} X_0 &= t_1 p_1^{'} + X_1 \\ M_0 &= t_1 r_1^{'} + M_1 \end{aligned} \tag{2}$$

Where t_1 represents the first component extracted from X_0, p_1 and r_1 represent the regression coefficients, M_1 and X_1 represent the residual of X_0 and M_0 respectively.

Next, replace X_0 and M_0 with M_1 and X_1 and repeat the second step for a times until the accuracy of regression equation meets the requirement, in which a is determined by the cross validation expressed below. Then, the estimation of M_0 can be denoted as:

$$\hat{M}_0 = \sum_{i=1}^{a} t_i r_i' \tag{3}$$

Finally, according to the inverse process of normalization, the regression equation can be restored.

$$M = f(X) \tag{4}$$

The specific calculation process see [8] in detail.

The PLSR model between spatial direction and HRTF phases is the same way as before.

3.2 Cross Validation

Reject the i*th* (*i=1,2,...,n*) observed data each time and construct the PLSR model with the left $n-1$ data. Meanwhile, fit the regression model after $h(h<r)$ components' being extracted, and then put the i*th* rejected data into the regression equation to obtain the predicted value of $\hat{b}_{ij}(h)$ of m_j $(j=1,2,...,p)$ on i*th* observed data. For each $i(i=1,2,\cdots n)$, repeat the validation. Thus, the prediction error square sum of the j*th* variable with *h* extracted components can be expressed as:

$$PRESS_j(h) = \sum_{i=1}^{m} (b_{ij} - \hat{b}_{ij}(h))^2, j = 1,2,\cdots,n \tag{5}$$

The predicted error square sum of $M = [m_1, m_2, \cdots, m_n]^T$ can be wrote as:

$$PRESS(h) = \sum_{i=1}^{M} PRESS_j(h) \tag{6}$$

Besides, fit *h* regression equations with all the samples. Herein, denote the predicted value of the i*th* sample as $\hat{b}_{ij}(h)$, further, the error square sum of h_i can be denoted as

$$SS_j(h) = \sum_{i=1}^{m} (b_{ij} - \hat{b}_{ij}(h))^2 \tag{7}$$

and the error square sum of M would be:

$$SS(h) = \sum_{j=1}^{P} SS_j(h) \tag{8}$$

The number l of component extracted corresponds to the least value of $PRESS(h)$. The effectiveness of cross validation is defined as:

$$Q_h^2 = 1 - PRESS(h)/SS(h-1) \tag{9}$$

In this case, in the end of each modeling, calculate the effectiveness of cross validation until $Q_h^2 < 1 - 0.95^2$. That means the model contains h components has met the accuracy demand.

3.3 Spatial-continuous HRTFs

According to the models above, the unknown HRTFs can be synthesized with the corresponding magnitudes and phases:

$$H(n) = M * [\cos(P) + i * \sin(P)] \tag{10}$$

4. Experiments and Error Evaluation

4.1 Experiments and Results

We selected one subject with integrated measured HRTF from the database to model the PLSR relationship between the spatial location and HRTF magnitude and phase. The range of spatial location is of the azimuth from 0° to 360° and the elevation from -30° to 30° with the integration of 10°. Meanwhile, we left out some spatial location with measured HRTFs (case in point, we have measured the HRTF of location (0°, 45°), and we made this direction as the test data) to be the verification group using for testing the accuracy of the model.

Figure 1 shows the predicted and measured phases of direction (0°, 45°) and (0°, 135°), figure 2 shows the corresponding magnitudes.

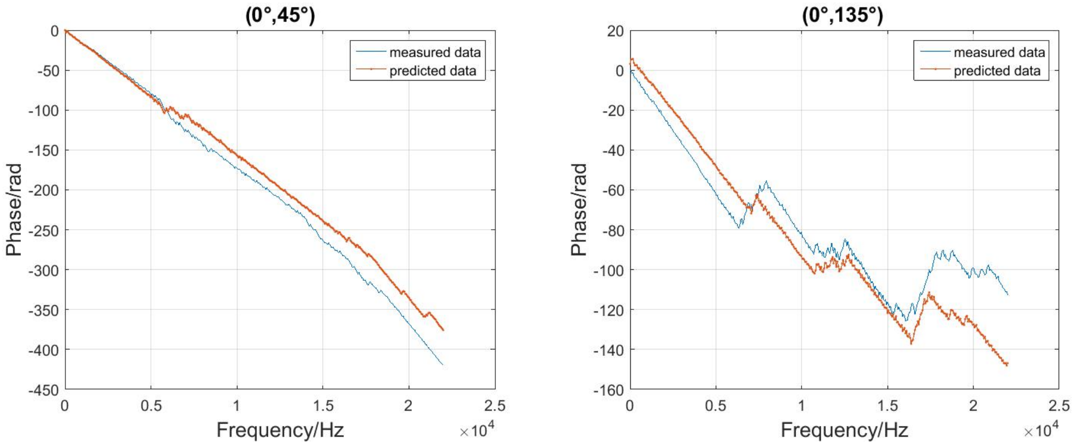

Fig 1 Measured and predicted data of HRTF phases

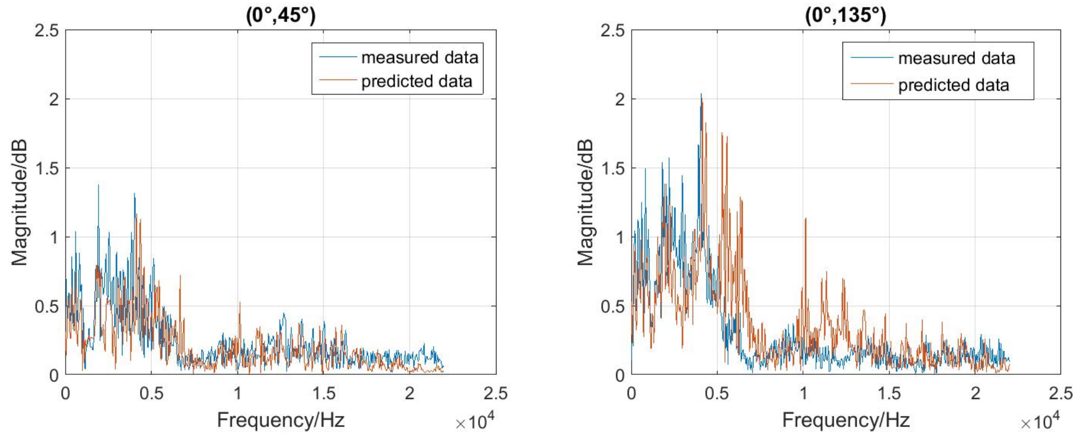

Fig 2 Measured and predicted data of HRTF magnitudes

4.2 Error Evaluation

We test the PLSR models with some measured HRTFs by calculating the similarity of phases and magnitudes predicted and measured. The formula for calculating the similarity is:

$$\cos(x, y) = \frac{x^{'} y}{\|x\| \bullet \|y\|} \tag{11}$$

where $x'y$ denotes dot-product of measured data x and predicted data y, $\|x\|$ and $\|y\|$ denote the norm of x and y respectively. A larger value of $\cos(x, y)$ represents a less similarity of x and y.

Table 1 gives the similarity of magnitudes and phases in different spatial directions.

Table 1 Similarity of original data and predicted data

	magnitude		phase	
	left	right	left	right
(0°, 45°)	0.811	0.833	0.893	0.897
(0°, 135°)	0.796	0.767	0.826	0.882
(0°, 225°)	0.762	0.789	0.832	0.787
(0°, 315°)	0.790	0.818	0.855	0.839

The results show that, the predicted phases and magnitudes are similar to the measured ones. Comparing the similarity of magnitudes and phases, apparently, the latter one has a larger similarity than the former, which indicates that the PLSR model between the spatial direction and HRTF phases obtains a higher accuracy.

5. Conclusion

In this paper, we proposed a new method to construct a spatial-continuous model of HRTFs. We studied two partial least square regression relationships between spatial directions and HRTF magnitudes, HRTF phases with some measured HRTFs in different direction. Then the unmeasured HRTFs in space can be synthesized through the predicted magnitudes and phases. Our experiments suggest that the regression models are effective, which provides a guide that the HRTFs can be inferred from a small amount of measurements by some statistical fitting method.

Acknowledgements

This paper was supported by the National Natural Science Foundation of China (No. 11374241 and No. 11404265).

REFERENCES

1 Salvador, C. D., Sakamoto, S., Trevino, J. A., et al, Accuracy of head-related transfer function synthesized with spherical microphone arrays, *Proceedings of meetings on Acoustics*, **19**(1), 055085(2013).
2 Huang, Q., Fang, Y., Interpolation of head-related transfer functions using spherical fourier expansion, *Chinese 1. Electron*., **26** (4), 571-576, (2009).
3 Algazi, Y. R., Duda, R. O.,Thomposon, D. M., and Avendano, C., The CIPIC HRTF Database, *Proc. IEEE Workshop Applicat. Signal Process. Audio and Acoust.*, New Paltz, NY, 99-102 (2001).
4 Christensen, F., Mϕller, H., Interpolating between head-related transfer functions measured with low directional resolution, *AES 107th Convention*, New York,USA, 5253(2000).
5 Nishino, T., Mase, S. And Kajita, S., Interpolating HRTF for auditory virtual reality, *The Journal of the Acoustical Society of America*, **100**(4), 2602(1996).
6 Wightman, F. L., Kister, D. J., et al, The dominant role of low-frequency interaural time difference in sound localization, *The Journal of theAcoustical Society of America*, **91**(3), 1648-1661(1992).
7 Wenzel, E. M., Arruda, M., Kistler, D. J., et al, Localization using nonindividualized head-related transfer functions, *The Journal of the Acoustical Society of America*, **94**(1), 111-123(1993).
8 Wang, H. W., *Partial least square regression--method and application*, National Defense Industry Press, Beijing(1999).