

PAPER

Head-related transfer function interpolation through multivariate polynomial fitting of principal component weights

Lin Wang*, Fuliang Yin† and Zhe Chen‡

School of Electronic and Information Engineering, Dalian University of Technology, 116023 P.R. China

(*Received 4 November 2008, Accepted for publication 25 March 2009*)

Abstract: Head-related transfer function (HRTF) interpolation plays an important role for implementation of 3D sound system because it can not only reduce the number of measurements for HRTFs, but also reduce the data of HRTFs for seamless binaural synthesis. This paper addresses the problem of accurately realizing the interpolation of HRTF for synthesis of virtual auditory space, and proposes a HRTF interpolation method based on principal component analysis. Firstly, the HRTF is decomposed into principal components and corresponding principal component weights, where principal components are direction-independent and principal component weights are direction-dependent; then the directional variation of the principal component weight is multivariate polynomial fitted with a bivariate function of two spatial angulus (azimuth and elevation). Moreover, a sphere-partitioning optimization scheme is employed to improve the approximation precision. Experiment results demonstrate that HRTFs in the entire sphere surface can be interpolated by the proposed method with small distortion, and the proposed method performs better than conventional methods. Therefore the proposed method gives a promising way for HRTF interpolation.

Keywords: Head-related transfer function, Interpolation, Principal component analysis, Multivariate polynomial fitting

PACS number: 43.60.Gk, 43.66.Qp [doi:10.1250/ast.30.395]

1. INTRODUCTION

3D sound technology has great application potentials in multimedia, home entertainment, virtual reality, and human-computer interaction systems. In the study of 3D sound, head-related transfer function (HRTF), which describes the path between the sound source and the ear, plays an important role [1]. Accurate spatial positioning of sound sources in 3D space, to be presented via headphones, can be achieved if we convolve the audio source with a set of HRTFs. HRTFs are usually obtained from their corresponding head-related impulse responses (HRIRs), measured under controlled conditions on a spherical surface of constant radius for a set of positions. However, the measurement only represents the spatial characteristics on discrete positions because of the enormous amounts of time and physical loads of the listener. The need for interpolating the HRTF becomes more important when we address the problem of moving sound sources, or when the perceived direction of the sound changes due to the listener's head movements. It is indispensable to consider how to obtain the HRTFs in arbitrary directions from a limited number of measured HRTFs, so that the switch between different HRTFs is accomplished in a fast and smooth way without creating audible artifacts [2].

The interpolation of HRTF enables us to reduce the number of measurements for new HRTFs, and also reduce the data of HRTFs in auditory virtual systems. Much work has been done to achieve accurate interpolation of HRTF. The interpolation methods may be summarized into two kinds: direct interpolation method and indirect interpolation method. The class of direct interpolation method calculates the target HRTF directly with HRTFs associated with two or more nearest points that circumscribe the desired point. For example, the bilinear method realizes the target position's interpolation with its four nearest measured HRTFs according to their ubiety [1]. In [3], linear interpolation is realized in the form of common acoustics pole/zero (CAPZ) model. And in [4], linear interpolation is realized in the form of inter-positional transfer function (IPTF) model. In [5] a piecewise method is proposed, where the whole frequency band is divided into small

*e-mail: wanglin_2k@sina.com
†e-mail: flyin@dlut.edu.cn
‡e-mail: eeyin@dlut.edu.cn

frequency bands and linear interpolation is implemented in each narrow frequency bands separately. Since the class of direct interpolation method supposes the linear relationship between neighboring positions' HRTFs, it is simple and straightforward. But at some neighboring positions, the relationships among them are not as described by the direct interpolation method. In these cases, the direct interpolation method fails to get good interpolation result.

The class of indirect method estimates the directional variation function of HRTF from the known ones, so that the coordinate of the spatial point may be directly mapped into the corresponding HRTF. For example, the spline interpolation methods [6], the rational interpolation method [7], the plenacoustic function method [8], the spatial frequency response surface (SFRS) method [9], and angular parameterization method [10] all belong to the class of indirect interpolation. Since the indirect interpolation method makes use of the information of all the known HRTFs to calculate the target HRTF, its performance is generally better than the direct method's. On the other hand, the indirect method requires large computation and memory consumption, which is not appropriate for real-time realization. In a word, although much work has been done around the interpolation method, it is still considered as an open problem.

To increase the HRTF interpolation accuracy, an interpolation method based on principal component analysis is proposed in this paper. Differentiating from the methods above which implement HRTF interpolation in time or frequency domain, the proposed method firstly decomposes the HRTF into direction-independent principal components and the direction-dependent principal component weights. Then the directional variation of the principal component weight is multivariate polynomial fitted with a bivariate function of two spatial angulus (azimuth and elevation). Moreover, a sphere-partitioning optimization method is employed to improve the approximation precision. HRTFs in the entire sphere surface can be interpolated by the proposed method with small distortion.

The paper is organized as follows. Two classical interpolation methods, the linear and spline methods are introduced in Section 2. In Section 3, model order reduction is applied to HRTF by principal component analysis and the performance of the conventional interpolation methods are discussed. In Section 4, the multivariate polynomial fitting method is described in detail. Experiments and analysis are carried out in Section 5. At last conclusions are drawn in Section 6.

2. LINEAR AND SPLINE INTERPOLATION METHODS

Two classical interpolation methods, the linear interpolation method and spline interpolation method, are investigated in this section. The linear method is the simplest and straightforward method and the spline method is used widely in various fields.

In linear method, the target (interpolated) HRTF magnitude is computed as a weighted mean of the measured HRTF magnitudes associated with the two nearest points that circumscribe the desired point. This can be expressed as

$$H(\omega) = rH_1(\omega) + (1-r)H_2(\omega), \quad 0 \le r \le 1 \tag{1}$$

where $H(\omega)$ is the target HRTF magnitude, $H_1(\omega)$ and $H_2(\omega)$ are two reference HRTF magnitudes, r is the dividing ratio which relates to the distance of the desired point to the two reference points.

In the case of the spline method, the cubic spline is generally used. It uses the piecewise third-order polynomials to connect the data points which often results in strictly smooth curve. Given a set of data of $N+1$ points, $(\theta_0, H_0(\omega)), (\theta_1, H_1(\omega)), \cdots, (\theta_N, H_N(\omega))$, where $H_i(\omega)$ is the magnitude of the ith HRTF and θ_i is the spatial position of the ith HRTF. The spline method is as follows.

Assuming the second derivatives of the two end points are

$$H_0''(\omega) = H_N''(\omega) = 0 \tag{2}$$

other second derivatives can be evaluated by the equation set

$$\begin{aligned}
&(\theta_i - \theta_{i-1})H_{i-1}''(\omega) + 2(\theta_{i+1} - \theta_{i-1})H_i''(\omega) \\
&\quad + (\theta_{i+1} - \theta_i)H_{i+1}''(\omega) \\
&\quad = \frac{6}{\theta_{i+1} - \theta_i}[H_{i+1}(\omega) - H_i(\omega)] \\
&\quad + \frac{6}{\theta_i - \theta_{i-1}}[H_{i-1}(\omega) - H_i(\omega)], \\
&i = 1, 2, 3, \cdots, N-1
\end{aligned} \tag{3}$$

The cubic function for each interval (θ_{i-1}, θ_i) is

$$\begin{aligned}
H(\omega) = &\frac{H_{i-1}''(\omega)}{6(\theta_i - \theta_{i-1})}(\theta_i - \theta)^3 + \frac{H_i''(\omega)}{6(\theta_i - \theta_{i-1})}(\theta - \theta_{i-1})^3 \\
&+ \left[\frac{H_{i-1}(\omega)}{\theta_i - \theta_{i-1}} - \frac{H_{i-1}''(\omega)(\theta_i - \theta_{i-1})}{6}\right](\theta_i - \theta) \\
&+ \left[\frac{H_i(\omega)}{\theta_i - \theta_{i-1}} - \frac{H_i''(\omega)(\theta_i - \theta_{i-1})}{6}\right](\theta - \theta_{i-1})
\end{aligned} \tag{4}$$

where $H(\omega)$ is the target HRTF magnitude, and $\theta \in (\theta_{i-1}, \theta_i)$.

Comparing two methods above, the linear interpolation method is simple because the information available for interpolation is few, while the spline interpolation method makes use of the information of all the known points and thus is more complicated and computation consuming. Generally, the spline method performs better than the linear method.

3. HRTF INTERPOLATION BASED ON PRINCIPAL COMPONENT ANALYSIS

This section introduces the interpolation method based on principal component analysis, and discusses the interpolation results of conventional interpolation methods. It should be mentioned that the discussion only talks about the magnitude of HRTF and ignores the phase. The reason is that, during the measurement, phase is easy to be contaminated by equipments and circumvents, and the phase we measured is not accurate in the beginning. The information of phase can be recovered by means of the minimum phase characteristics corresponding to magnitude, which can not cause special perception reduction in synthesis [11].

3.1. Principal Component Analysis of HRTF

HRTF interpolation is generally implemented at the frequency domain where the magnitude is composed of a lot of frequency bins. To make the analysis of the interpolation more clearly, we apply principal component analysis (PCA) to HRTF magnitude to reduce the model order [12]. Principal component analysis is a classical model order reduction method. Its key idea is to reduce the dimensionality of a data set while retaining the primary variation in the data. PCA decomposes a set of magnitude spectra into weighted combinations of basis functions. The basis functions are shared by the whole spectrum set, they can be considered the basic spectral shapes from which each spectrum is built. Each spectrum is approximated by a weighted sum of these basis functions, and the weight define the relative contribution of each basis function to the spectrum. Thus each spectrum has its unique weights. The basis function is called principal component and its weight is called principal component weight. Suppose d_k, a vector of dimensionality p, is the kth magnitude spectrum of the data set, it can be represented as a linear combination of basis functions by PCA, this is

$$d_k = \sum_{i=1}^{q} w_{ki} c_i \tag{5}$$

where c_i is the ith principal component, w_{ki} is the ith weight for d_k, and q is the total number of principal components. The number of principal components required to provide an adequate representation of the data depends on the amount of redundancy or correlation presented in the data set. The greater the redundancy, the smaller the number of principal components needed. Generally the number of basis functions is much smaller than the dimensionality of d_k, i.e. $q \ll p$, thus the data size is reduced greatly. The detailed process of principal component analysis is described in [12]. By principal component analysis, a group of HRTFs can be decomposed into the principal components and the corresponding principal component weights. The principal components are shared by the all the HRTFs, they are direction-independent; while the principal component weights remain unique for each HRTF, they are direction-dependent. Based on this, we propose to implement HRTF interpolation via the principal component weight.

3.2. Interpolation of the Principal Component Weights

The idea of interpolation on the principal component weight has ever been proposed in [13,14]. In [13], a two-dimensional spline method is used for interpolation, which is computation consuming. In [14], a Fourier series expansion fitting of the principal component weight is proposed for interpolation in the horizontal plan. To simplify the discussion, here we only discuss interpolation from two points by linear and spline method.

For linear method, the interpolation of frequency magnitude is equivalent to the interpolation of the corresponding principal component weights. The proof is as follows.

By principal component analysis, the two reference HRTF magnitude can be expressed as $H_1(\omega) = \sum_i w_{1i} c_i$, $H_2(\omega) = \sum_i w_{2i} c_i$, where $H_1(\omega)$ and $H_2(\omega)$ are two reference HRTF magnitudes, c_i is the ith principal component, w_{1i} and w_{2i} are the corresponding ith principal component weight of H_1 and H_2 respectively. The target HRTF magnitude is

$$\begin{aligned} H(\omega) &= rH_1(\omega) + (1-r)H_2(\omega) \\ &= r\sum_i w_{1i}c_i + (1-r)\sum_i w_{2i}c_i \\ &= \sum_i [rw_{1i} + (1-r)w_{2i}]c_i \end{aligned} \tag{6}$$

It can be seen from (6) that the interpolation of magnitude is equivalent to the interpolation of the corresponding principal component weights.

Take the KEMAR HRTF as the experiment data [15], and apply principal component analysis to all HRTF magnitudes in the database (710 positions totally). Take the principal component weight-1 as an example, the weight-1 in the horizontal plane is shown in Fig. 1. It can be seen from Fig. 1 that the curve of weight-1 in the horizontal plane is not a monetary straight line; it is composed of some turning curves. If the target point and its neighboring points lie on one straight line, the target HRTF can be calculated using linear equation with little distortion. For example, points D, E, F are on one straight line, point E can be interpolated linearly from D and F accurately. On the other hand, if the target point and its neighboring points lie on different turning curves, linear interpolation may cause large error. For example, points A and C lie on different

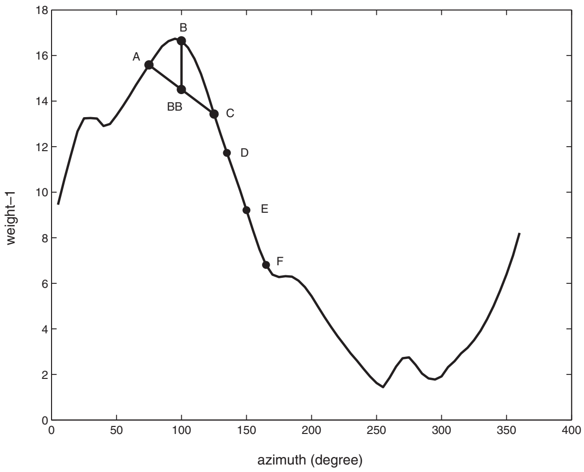


Fig. 1 Weight-1 in the horizontal plane.

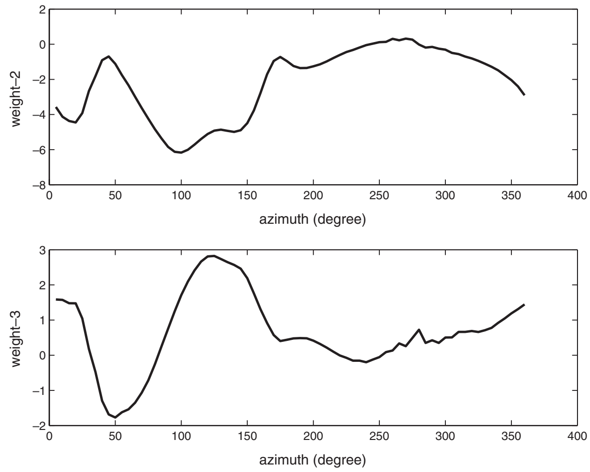


Fig. 2 Weight-2 and weight-3 in the horizontal plane.

turning curves, and B lie at the turning point of two turning curves, BB is interpolated from A and C by linear method. It is obvious that the interpolated point BB deviates from the target point B greatly, and the error of linear interpolation is very large.

The curves of weight-2 and weight-3 are displayed in Fig. 2. It is shown in Fig. 2 that the variation trends of weight-2 and weight-3 are similar to weight-1's. Both of them are composed of many turning curves. Thus from the aspect of interpolation of the weights, the linear method is not appropriate for interpolation because large errors may occur at turning points of weight curves.

Similarly, the spline method can not settle the problem that large errors will occur at turning points of weight curves either. The interpolation accuracy at the turning points affects the global interpolation performance greatly.

4. MULTIVARIATE POLYNOMIAL FITTING METHOD

Based on the analysis in Section 3, we find that large interpolation errors generally occur at the turning points of weight curves. Excellent interpolation result will not be obtained until the weight curves are represented precisely. A multivariate polynomial fitting method is presented to approximate the weight curve with a bivariate function of two spatial variables azimuth and elevation. The nth weight of HRTF at position (azimuth, elevation), $w_n(\theta,\phi)$, is expressed as

$$w_n(\theta,\phi) = \sum_{p=0}^{P}\sum_{q=0}^{Q} \theta^p \phi^q c_n^{p,q} \tag{7}$$

where θ and ϕ are azimuth and elevation normalized into [0, 1] respectively, $c_n^{p,q}$ is the polynomial coefficient, P and Q are polynomial order of azimuth and elevation respectively.

Suppose $F_n(\theta,\phi)$ is the target weight corresponding to $w_n(\theta,\phi)$, the total modeling square error is

$$\begin{aligned} J &= \sum_{k=1}^{K}\sum_{l=1}^{L} |F_n(\theta_k,\phi_l) - w_n(\theta_k,\phi_l)|^2 \\ &= \sum_{k=1}^{K}\sum_{l=1}^{L} \left| F_n(\theta_k,\phi_l) - \sum_{p=0}^{P}\sum_{q=0}^{Q} \theta_k^p \phi_l^q c_n^{p,q} \right|^2 \end{aligned} \tag{8}$$

where L is the number of azimuth, K is the number of elevation, and the total number of HRTFs involved in computation is $M = LK$.

Solving the minimization problem of J about the coefficient $c_n^{p,q}$, the least squares solution of $c_n^{p,q}$ may be obtained by

$$C_n = (X_n^{\mathrm{T}} X_n)^{-1} X_n b_n \tag{9}$$

where C_n is the coefficient vector of length $(P+1)(Q+1)$

$$C_n = [c_n^{0,0}, \cdots, c_n^{0,Q}, c_n^{1,0}, \cdots, c_n^{1,Q}, \cdots, c_n^{P,0}, \cdots, c_n^{P,Q}]^{\mathrm{T}} \tag{10}$$

and b_n is a vector of length M, X is a matrix of size $M \times (P+1)(Q+1)$

$$\begin{aligned} b_n = [&F_n(\theta_1,\phi_1), \cdots, F_n(\theta_1,\phi_K), F_n(\theta_2,\phi_1), \cdots, \\ &F_n(\theta_L,\phi_1), \cdots, F_n(\theta_L,\phi_K)]^{\mathrm{T}} \end{aligned} \tag{11}$$

$$X = \begin{bmatrix} \theta_1^0\phi_1^0 & \cdots & \theta_1^P\phi_1^Q \\ \cdots & \cdots & \cdots \\ \theta_L^0\phi_K^0 & \cdots & \theta_L^P\phi_K^Q \end{bmatrix} \tag{12}$$

Given the desired accuracy, the required polynomial order P and Q are related to the total number of HRTFs involved in the computation. Applying the solution over the entire reference sphere may leads to a very high polynomial order, which may increase computational cost greatly. A sphere-partitioning parameter optimization method is employed to decrease the polynomial order, which partitions the sphere into several subareas [10]. And polynomial fitting is implemented in each subarea separately. There are many partitioning schemes, such as along

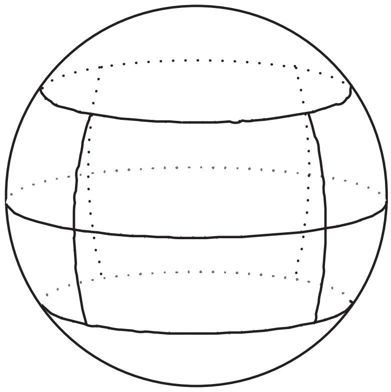


Fig. 3 Partition scheme by constant-area regions.

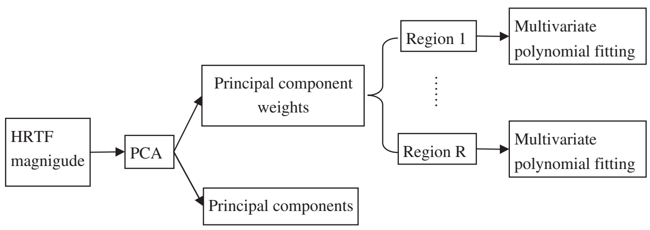


Fig. 4 Flow chart of the proposed method.

the meridians, along the horizontal strips, and by constant-area regions. A typical partitioning scheme by constant-area regions is shown in Fig. 3. Suppose the sphere is divided into R subregions, the total number of required parameters is $R(P+1)(Q+1)$. For a desired accuracy, small regions may lead to low model order, while the total parameter number in the entire sphere may be large. On the other hand, large regions needs fewer parameters in the entire sphere, while the model order may be high. Thus a careful balance should be considered between the polynomial order and the number of regions.

Based on the analysis above, the whole flow chart of the proposed method is shown in Fig. 4. The polynomial coefficients have been computed offline and stored in the memory beforehand. Given the spatial position, the target HRTF principal component weight can be calculated from (7) with the corresponding polynomial coefficients read from the memory. The reconstruction process consumes little computation.

5. EXPERIMENT RESULTS AND ANALYSIS

To evaluate the performance of the proposed method, we firstly compare the interpolation results of the linear, spline, and proposed method in the horizontal plane, then give the interpolation result of the proposed method over the entire sphere region. The experiments above are conducted using one dataset: KEMAR HRTF. To examine the generality of the proposed method, its interpolation performance are investigated for various HRTF measurements.

Table 1 The azimuth intervals of KEMAR's measurement.

Elevation (°)	Number	Interval (°)
−20–20	72	5.00
−30 and 30	60	6.00
−40 and 40	56	6.43
50	45	8.00
60	36	10.00
70	24	15.00
80	12	30.00
90	1	/

5.1. KEMAR HRTF

A lot of research projects of institutions and universities have collected libraries of HRTF measurements in their anechoic chambers. One of the famous libraries is called KEMAR database, which is available on the Internet [15]. The HRTF measurements were made in the anechoic chamber of MIT Media Laboratory. In the database, the HRIRs were measured for elevations −40°–90° in 10° steps, and azimuths 0°–360° for the steps listed in Table 1 (710 positions altogether). It is apparent that the spatial sampling is quite discrete. Therefore, interpolation is requisite for practical application. The impulse response of each HRTF measurement is 512 taps long at the sample rate 44.1 kHz. Before the experiment, some preprocessing operation, such as windowing, equalization, and so on, have been applied to the original data, and the final HRIR data is 128 taps long [15]. In Section 5.2. and Section 5.3., the experiments are carried out using KEMAR HRTF.

5.2. Experiments of Interpolation in the Horizontal Plane

Two interpolation methods (linear and spline) and the proposed method are considered in this experiment. In principal component analysis, 10 principal components are chosen to account for about 92% of the variation in the HRTF magnitude functions. It has been proved in [12] that the localization performance degraded little when above 90% of variation is retained. Interpolation of the principal component weights is performed in the horizontal plane using the three methods respectively. In linear and spline methods, the reference points are selected at angles 5° to 365° with an interval of 20°, i.e. 19 reference points are used. In the proposed method, the horizontal plane is divided into three subregions, i.e. 0°–120°, 120°–240°, and 240°–360°, the polynomial order is set as $P = 7$, $Q = 0$. The total parameter number of the proposed method is 24, which is a little larger than those of the linear and spline methods. To compare the interpolation precision of the weight, we define the relative error between the original weight and the interpolated weight as

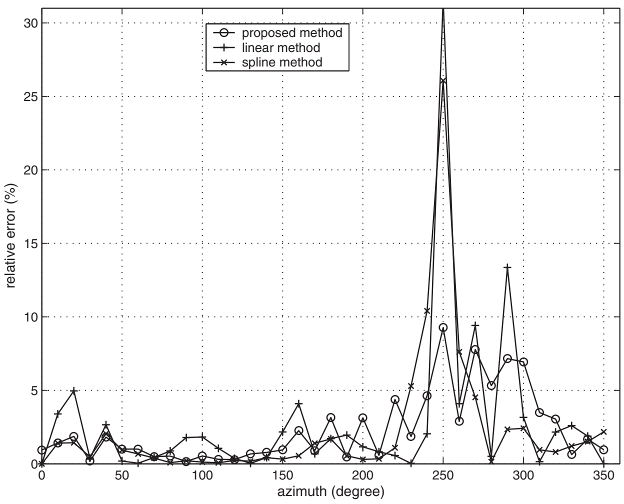


Fig. 5 Relative interpolation errors of weight-1 in the horizontal plane for the linear method, spline method, and proposed method.

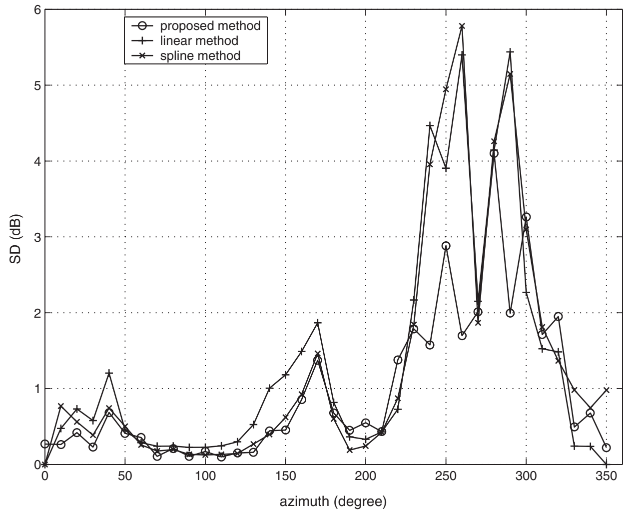


Fig. 6 Global spectral distortions in the horizontal plane for the linear method, spline method, and proposed method.

$$e = \left|\frac{w_o - w_r}{w_o}\right| \times 100\% \qquad (13)$$

where w_o is the original weight and w_r is the interpolated weight. Take weight-1 as an example, its relative error as a function of azimuth is shown in Fig. 5. It can be seen that the performance of the linear method is the worst among the three ones. The interpolation precision of the spline method is close to that of the proposed method except in areas around azimuth 250°, where the error of the proposed method (9.3%) is much less than the one of the spline method (26%). Similar cases occur in the interpolations of other weights too. Thus from the aspect of weight interpolation, the performance of the proposed method is the best among the three ones, which verifies the analysis of Section 3.

To evaluate the global performance of the three interpolation methods, we define the global spectral distortion (SD) in dB as

$$SD = \sqrt{\frac{1}{K}\sum_{k=0}^{K-1}[20\log_{10}|H(\omega_k)| - 20\log_{10}|\hat{H}(\omega_k)|]^2} \qquad (14)$$

where $H(\omega_k)$ and $\hat{H}(\omega_k)$ are the original and interpolated HRTF magnitude respectively, and K is the total number of frequency bins in $[0, \pi]$. The smaller SD is, the more accurate the interpolation is. The spectral distortion in the horizontal plane of the three methods is shown in Fig. 6. It can be seen from Fig. 6 that in areas around 240°–300° where large errors occurs, the SD curve of the proposed method is almost always below the other two methods. In other areas, the interpolation errors of the three methods are all small, and their SD curves are very close. The average SD of the linear, spline and proposed method across the horizontal plane are 1.32 dB, 1.30 dB, and 0.98 dB respectively, where the proposed method has the lowest average SD among the three methods. Since the proposed method has a more gently-varying SD curve, and a lower average SD, it is superior to the other two methods. The interpolation performance of the linear method is similar to the spline method.

5.3. Experiments of Interpolation over the Sphere

The interpolation in the horizontal plane has been investigated, but the spatial information reconstruction needs not only the information from the horizontal plane, but also the information from the median plane. Thus the interpolation performance in the entire sphere area using the proposed method is examined. In the experiment, different polynomial orders are tested, where $P = 3$ to 8 and $Q = 3$ to 4. The detailed sphere partitioning scheme is shown in Table 2, where the sphere is partitioned into 10 subregions. The average spectral distortion of the proposed method at different polynomial orders over the entire sphere area is shown in Table 3. It can be seen from Table 3 that the interpolation performance increases with the polynomial order P and Q. A phenomenon is observed that given a constant parameter number, i.e. $(P+1)(Q+1)$, the performance of the condition with higher Q is better than the condition with lower Q. For example, SD at $P = 7$, $Q = 3$ is 0.99, while SD at $P = 5$, $Q = 4$ is 0.89. Thus the performance seems to be more affected by the elevation order Q than the azimuth order P. The authors think it is due to a greater variation of HRTF in elevation than in azimuth. However this is subject to further investigation because HRTF is subject-dependent. Take the interpolation precision and parameter number both into consideration, we set $P = 6$, $Q = 4$. In this

Table 2 Partitioning scheme for KEMAR HRTF.

Region	Elevation (°)	Azimuth (°)
1–2	40–80	0–180, 180–360
3–6	0–40	0–90, 90–180, 180–270, 270–360
7–10	−40–0	0–90, 90–180, 180–270, 270–360

Table 3 Spectral distortion at different polynomial order.

P	Q	Average SD (dB)
3	3	1.52
4	3	1.30
5	3	1.17
6	3	1.06
7	3	0.99
8	3	0.92
3	4	1.34
4	4	1.06
5	4	0.89
6	4	0.71
7	4	0.62
8	4	0.51

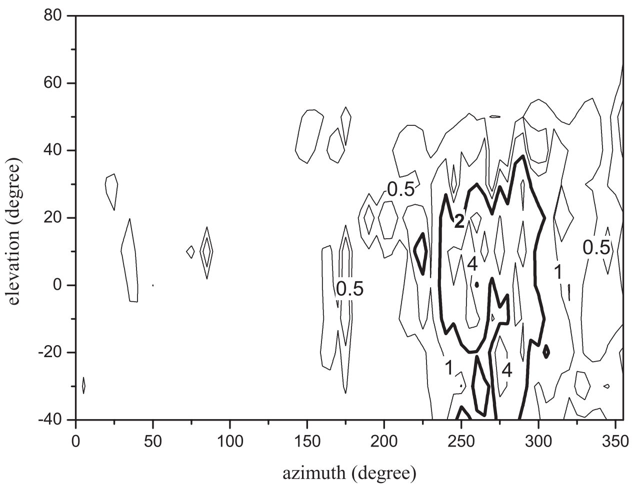


Fig. 7 Contour plot of spectral distortion over the entire sphere area for the proposed method (KEMAR HRTF).

condition, the average SD $= 0.71$ dB, the percentage of SD > 2 dB is 9.6%, and the maximum SD is 6.7 dB. It has been shown in [16] that listeners can detect HRTF magnitude errors larger than approximately 2 dB. The contour plot of the spectral distortion over the entire sphere region is shown in Fig. 7. On the contour line, the number is the spectral distortion between the original and interpolated HRTF magnitude in dB. The contour line of SD $= 2$ dB is highlighted in thick line. The error (SD > 2 dB) mainly occurs in region of elevation $-30°$–$30°$ and azimuth 230°–320°. And the errors in other regions are very small. In error regions, large errors occur because the HRTFs in this area have a great variation, and the

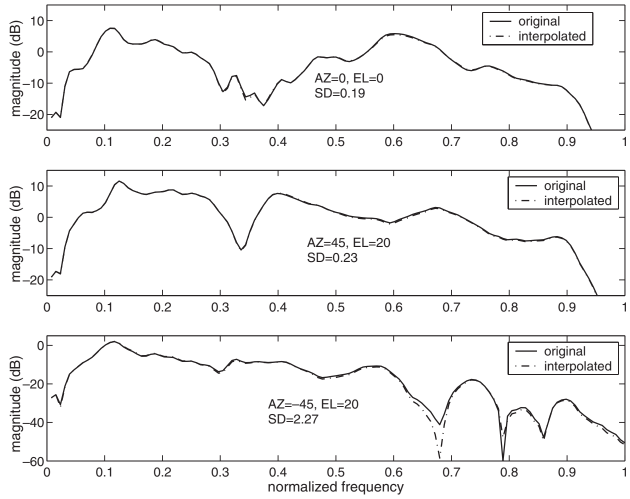


Fig. 8 Original versus interpolated HRTF magnitude.

polynomial order is not high enough to catch the variation precisely. If we increase the polynomial order in this area, the interpolation performance can be further improved. However a balance should be considered because higher polynomial order may increase the computation and memory cost. Figure 8 compares the original HRTFs with those of their interpolated counterparts at three positions. The solid lines represent original HRTFs, the dashed lines represent interpolated HRTFs. The azimuth and elevation of each HRTF is noted by symbols "AZ=" and "EL=" respectively. The symbol "SD" denotes the spectral distortion between the original and interpolated HRTF. Figure 8 shows that the original functions are well approximated by the proposed method.

5.4. Experiments of Interpolation with CIPIC HRTF

The experiments above are carried out using only one dataset, KEMAR HRTF, however HRTF is subject-dependent, thus we apply the proposed method to other measured HRTF datasets to examine its generality. The CIPIC HRTF is another famous HRTF database available on Internet. The HRTF measurements were made in the anechoic chamber of U. C. Davis CIPIC Interface Laboratory. It provides head-related impulse responses for 45 subjects at 1,250 positions around each subject [17]. The impulse response of each HRTF measurement is 200 taps long at the sample rate 44.1 kHz. Interpolation performances in the entire sphere area using the proposed method are investigated for three subjects from the CIPIC database: "003," "058," and "133." The partitioning scheme is shown in Table 4, where the sphere is partitioned into 12 subregions. The polynomial order is set as $P = 7$, $Q = 5$. The interpolation result is given in Table 5. It can be seen from Table 5 that the interpolation results are similar to the one on KEMAR HRTF in Section 5.3, where the average SD $= 0.71$ dB, the percentage of SD > 2 dB is

Table 4 Partitioning scheme for CIPIC HRTF.

Region	Elevation (°)	Azimuth (°)
1–4	40–80	0–90, 90–180, 180–270, 270–360
5–8	0–40	0–90, 90–180, 180–270, 270–360
9–12	−40–0	0–90, 90–180, 180–270, 270–360

Table 5 Interpolation results for CIPIC HRTF.

Subject	Average SD (dB)	Percentage (SD > 2 dB)
003	0.83	9.5%
058	0.55	1.7%
133	0.71	4.7%

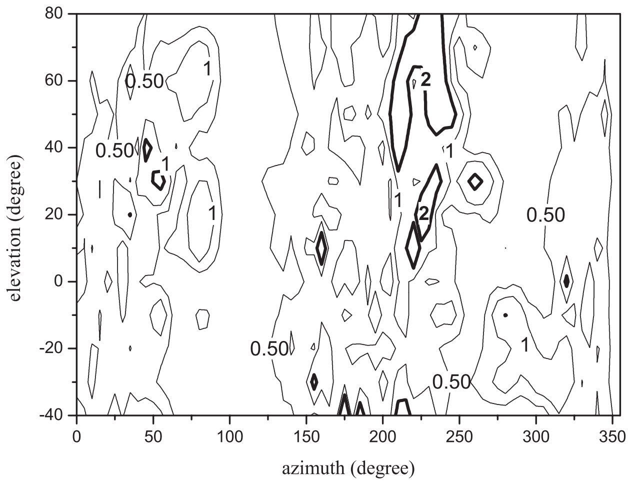


Fig. 9 Contour plot of spectral distortion over the entire sphere area for the proposed method (CIPIC HRTF of subject "133").

9.6%. For subject "133," the contour plot of the spectral distortion over the entire sphere region is shown in Fig. 9. On the contour line, the number is the spectral distortion between the original and interpolated HRTF magnitude in dB. The contour line of SD = 2 dB is highlighted in thick line. It can be seen from Fig. 9 that large errors (SD > 2 dB) also mainly occur in some small areas. From the experiments in Section 5.4, we can conclude that the proposed method is valid for various HRTF measurements.

6. CONCLUSION

In this paper, the problem of efficiently realizing HRTF interpolation is investigated. After model order reduction by principal component analysis which decomposes HRTF into a small set of principal components and the corresponding principal component weights, the performances of conventional interpolation methods are discussed from the aspect of interpolation of the weights. It is found that large interpolation errors generally occur at the turning points of the weight curves. A multivariate polynomial fitting method is therefore proposed to approximate the weight curve precisely with a bivariate polynomial of spatial variables. A sphere-partitioning optimization scheme is employed to improve the modeling accuracy. The proposed method performs better than conventional methods in the horizontal plane, and the interpolation result in the entire sphere surface is excellent. The generality of the proposed method is also verified by experiments on various HRTF measurements. The whole HRTF interpolation process of the proposed method requires little computation. Therefore, with the proposed method, it is possible for a set of HRTFs to be represented efficiently.

ACKNOWLEDGMENT

This work is supported by the National Natural Science Foundation of China (60772161, 60372082) and the Specialized Research Fund for the Doctoral Program of Higher Education of China (200801410015).

REFERENCES

[1] D. R. Begault, *3D Sound for Virtual Reality and Multimedia* (Academic Press, London, 1994).

[2] E. M. Wenzel and S. H. Foster, "Perceptual consequences of interpolating head-related transferfunctions during spatial synthesis," *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, pp. 102–105 (1993).

[3] K. Watanabe, S. Takane and Y. Suzuki, "Interpolation of head-related transfer functions based on the common-acoustical-pole and residue model," *Acoust. Sci. & Tech.*, **24**, 335–337 (2003).

[4] F. P. Freeland, L. W. P. Biscainho and P. S. R. Diniz, "Efficient HRTF interpolation in 3D moving sound," *AES 22nd Int. Conf. on Virtual, Synthetic and Entertainment Audio*, pp. 106–114 (2002).

[5] J. Zhang and Z. Y. Wu, "A piecewise interpolation method based on log-least square error criterion for HRTF," *IEEE 7th Workshop on Multimedia Signal Processing*, pp. 1–4 (2005).

[6] T. Nishino, S. Kajita, K. Takeda and F. Itakura, "Interpolating head related transfer functions in the median plane," *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, pp. 167–170 (1999).

[7] F. Keyrouz and K. Diepold, "A rational HRTF interpolation approach for fast synthesis of moving sound," *12th Digital Signal Processing Workshop and 4th Signal Processing Education Workshop*, pp. 222–226 (2006).

[8] T. Ajdler, L. Sbaiz and M. Vetterli, "Plenacoustic function on the circle with application to HRTF interpolation," *IEEE Int. Conf. Acoust. Speech Signal Process.*, pp. 273–276 (2005).

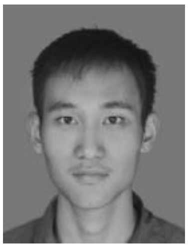
[9] C. I. Cheng and G. H. Wakefield, "Spatial frequency response surfaces: an alternative visualizationtool for head-related transfer functions (HRTFs)," *IEEE Int. Conf. Acoust. Speech Signal Process.*, pp. 961–964 (1999).

[10] F. P. Freeland, L. W. P. Biscainho and P. S. R. Diniz, "HRTF interpolation through direct angular parameterization," *IEEE Int. Symp. on Circuits and Systems*, pp. 1823–1826 (2007).

[11] A. Kulkarni and H. S. Colburn, "Efficient finite-impulse-response filter models of the head-related transfer function," *J. Acoust. Soc. Am.*, **97**, 3278–3278 (1995).

[12] 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**,

1637–1647 (1992).
[13] J. Chen, B. D. Van Veen and K. E. Hecox, "A spatial feature extraction and regularization model for the head related transfer function," *J. Acoust. Soc. Am.*, **97**, 439–452 (1995).
[14] T. Nishino, S. Mase, S. Kajita, K. Takeda and F. Itakura, "Interpolating HRTF for auditory virtual reality," *J. Acoust. Soc. Am.*, **100**, 2602–2602 (1996).
[15] W. G. Gardner and K. D. Martin, "HRTF measurements of a KEMAR," *J. Acoust. Soc. Am.*, **97**, 3907–3908 (1995).
[16] P. R. Runkle, M. A. Blommer and G. H. Wakefield, "A comparison of head related transfer function interpolation methods," *IEEE ASSP Workshop on Applications of Signal Processing to Audio and Acoustics*, pp. 88–91 (1995).
[17] V. R. Algazi, R. O. Duda and D. M. Thompson, "The CIPIC HRTF database," *Proc. 2001 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, pp. 99–102 (2001).



Lin Wang was born in Anhui, China, in 1981. He received the B.S. degree in electronic engineering from Tianjin University, China in 2003. He is pursuing the Ph.D. degree at the School of Electronic and Information Engineering, Dalian University of Technology, Dalian, China. He is currently a visiting fellow with the Institute for Microstructural Sciences, National Research Council Canada, from 2008 to 2009. His research interests include video and audio compression, blind source separation, 3D audio processing.

Fuliang Yin was born in Fushun city, Liaoning province, China, in 1962. He received the B.S. degree in electronic engineering and the M.S. degree in communications and electronic systems from Dalian University of Technology (DUT), Dalian, China, in 1984 and 1987, respectively. He joined the Department of Electronic Engineering, DUT, as a Lecture in 1987 and became an Associate Professor in 1991. He has been a Professor at DUT since 1994, and the Dean of the School of Electronic and Information Engineering of DUT since 2000. His research interests include digital signal processing, speech processing, image processing, broadband wireless communication, and integrated circuit design.

Zhe Chen was born in Heilongjiang province, China, in 1975. He received the B.S. degree, the M.S. degree in signal and information processing, and the Ph.D. degree in signal and information processing from Dalian University of Technology (DUT), Dalian, China, in 1996, 1999 and 2003 respectively. He joined the Department of Electronic Engineering, DUT, as a Lecture in 2002, and became an Associate Professor in 2006. His research interests include stochastic signal processing, speech processing, and broadband wireless communication.