Technical Note

A hybrid compression method for head-related transfer functions

Lin Wang *, Fuliang Yin, Zhe Chen

School of Electronic and Information Engineering, Dalian University of Technology, Dalian, 116023, PR China

ARTICLE INFO

ABSTRACT

The paper proposes a hybrid compression method to resolve the storage problem of a large number of head-related transfer functions (HRTFs). First, each HRTF is approximated by a minimum-phase HRTF and an all pass filter whose group delay equals the interaural time delay (ITD). Second, principal component analysis is applied to the entire HRTF set to derive several basis functions, with a weight vector set defining the contribution of the basis functions to each HRTF. Third, the weight set is vector quantized with the designed codebook. At last, the ITD is curved surface fitted with a cosine series bivariate polynomial. As a result, the HRTF can be reconstructed from the basis functions, codebook indexes, and ITD polynomial coefficients. Simulation results reveal that the proposed method may reduce the data size greatly with similar reconstruction precision comparing with the principal component analysis method.

1. Introduction

A set of transfer functions from different spatial locations to both ears, or head-related transfer functions (HRTFs), are primarily due to interactions of acoustic waves with a listener's head, torso and pinnae [1,2]. It is known that HRTFs convey important cues for human spatial hearing, such as interaural time difference (ITD), interaural level difference (ILD), and spectral cues, where ITD and ILD describe information in the horizon plane while spectral cues help resolving the cone of confusion by giving front–back and elevation perception. HRTFs serve a dominant role for implementation of binaural technology and virtual auditory reality, including 3D sound system. Spatial perception can be made through the convolution of the sound with the appropriate pair of HRTFs. There are two critical aspects to HRTFs: they vary according to sound source position and are unique for each individual. Theoretically, an ideal 3D sound field may be generated with individualized HRTFs, which are measured accurately for each subject separately. But in practice individualized HRTFs are difficult to be obtained because the measurement is time-consuming and requires specific equipments. In practice, generalized or non-individualized HRTFs are used, which may cause large localization errors which typically appear as up–down confusion, front–back confu-

sion and in-head localization [3]. To solve the problems, some spectrum modification methods have been proposed to enhance the perception difference from different sound directions [4,5]. But these methods only improve the localization performance a little because of the diversity of individuals. Recently some methods have been proposed to select a most appropriate HRTF set from a large number of HRTF sets for the listener by subjective listening comparison [6,7]. The subjective selection method requires no extra measurement equipment and thus can resolve the localization error problem to a certain extent when individualized HRTFs are difficult to be acquired. But the HRTF sets of multi-subjects consume a lot of storage space, this impedes the application of the subjective selection method in binaural technology and virtual auditory reality system.

Much research work has been done to reduce the memory consumption of the HRTF. The research may be divided into three kinds. The first kind is to calculate the HRTF by accurately acoustic modelling for the given sound direction [8,9]. This method requires little memory, but its localization performance needs further examination because of the complexity of human body structure. The second kind is pole-zero modelling of the measured HRTF, which represents the HRTF as a low order pole-zero model [10–12]. It can reduce the data size and computation load greatly. The third kind is to represent the HRTFs as weighted combinations of a few basis functions by mathematical decomposition. As an example of the third method, Chen et al. [13] and Bai et al. [14] modeled the external ear as a multisensor broadband beamformer

* Corresponding author. Tel./fax: +86 411 84707049.
E-mail addresses: wanglin_2k@sina.com (L. Wang), flyin@dlut.edu.cn (F. Yin), eeyin@dlut.edu.cn (Z. Chen).

whose weight vector set represented the characteristics of HRTF. Kistler and Wightman applied principal component analysis (PCA) to the logarithms of the HRTF magnitudes and represented them as linear combinations of basis functions [15]. Chen et al. proposed a spatial feature extraction and regularization (SFER) model for the complex valued (magnitude and phase) HRTFs, which represents HRTF as weighted combinations of a set of complex valued eigen transfer functions [16]. It is believed that there are a lot of similarities existing in the HRTFs [15], therefore a HRTF can be expressed as a linear combination of a few basis functions. The compression performance of the third kind of method is the best among the three ones, but the total memory size of it is still huge when there are a large number of HRTFs.

To reduce the memory size of the HRTF database, especially the one composed of multi-subjects' HRTFs, a hybrid compression method for HRTF is proposed. The HRTF is firstly approximated by a minimum-phase HRTF and an all pass filter whose group delay equals the interaural time delay (ITD). Principal component analysis is applied to the minimum-phase HRTF magnitudes to derive a set of basis functions for the entire data set and a weight vector for each HRTF. The weight vector set is vector quantized with the designed vector codebook [17]. The ITD variation with azimuth and elevation is curved surface fitted with a cosine series bivariate polynomial model. Comparing with principal component analysis, the proposed method may reduce the data size greatly with similar reconstruction precision.

This paper is organized as follows. The hybrid compression method for HRTF is developed in Section 2 and described in three aspects: principal component analysis, vector quantization, and curved surface fitting. Experiment results and analysis are given in Section 3 to evaluate the performance of the proposed method. Finally, conclusions are drawn in Section 4.

2. The hybrid compression method for HRTF

The classical compression method, principal component analysis, is well known for its compression efficiency. It decomposes a group of HRTFs into principal components and the corresponding principal component weights, where principal components are shared by the whole group and principal component weights remain unique for each HRTF. Although principal component analysis may reduce the data size greatly, the weight vector of each HRTF still consume much memory. Based on principal component analysis, we employ vector quantization method for the weight vector and curved surface fitting method for the interaural time delay (ITD) data separately to get further compression.

First of all, each HRTF is minimum-phase reconstructed, which results in a minimum-phase HRTF cascaded with an all pass filter whose group delay equals the interaural time delay (ITD).

Second, principal component analysis is applied to the entire minimum-phase HRTF magnitude set. A small set of basis functions are derived from the principal component analysis. Each spectrum can be approximated with a weighted sum of the basis functions, and the weights define the relative contribution of each basis function to the spectrum.

Third, the entire weight set is vector quantized, where a weight vector can be represented as a vector index in the codebook, and thus the storage space is saved.

At last, the ITD variation with direction is curved surface fitted with a cosine series bivariate polynomial for each subject in the database separately. Thus only a few polynomial coefficients need to be stored instead of the ITD set.

The workflow of the proposed method is shown as in Fig. 1. The basis functions and the vector codebook are shared by the entire HRTF database, the ITD polynomial coefficients are calculated for each subject, respectively, and the weight vector index is unique for each HRTF direction. All these parameters have been computed offline and stored in the memory beforehand. When binaural synthesis, the HRTF can be reconstructed with its corresponding parameters read from the memory. Because principal component analysis, vector quantization and curved surface fitting are employed jointly to reduce the size of the HRTF data, we call the new method a hybrid compression method. And the proposed hybrid compression method is described in detail as follows.

2.1. Minimum-phase reconstruction and principal component analysis

Generally, HRTF $H(e^{j\omega})$ can be approximated by a minimum-phase sequence combined with a position-dependent interaural time delay [18]:

$$H(e^{j\omega}) = H_{min}(e^{j\omega})H_{ap}(e^{j\omega}) \tag{1}$$

where H_{min} is the minimum-phase function and H_{ap} is an all pass function with the group delay $\tau = $ ITD, the interaural time difference.

$$H_{ap}(e^{j\omega}) = e^{j\omega\tau} \tag{2}$$

The phase of the minimum-phase function can be estimated from the magnitude function. Thus only the magnitude of the minimum-phase HRTF is processed by principal component analysis (PCA).

The key idea of principal component analysis is to reduce the dimensionality of a data set while retaining the primary variation in the data. It decomposes a set of magnitude spectra into weighted combinations of basis functions. Suppose d_k is the kth magnitude spectrum of the data set, it can be represented as a linear combination of basis functions by PCA. This is described as

$$d_k = \sum_{i=1}^{q} w_{ki}c_i \tag{3}$$

where c_i is the ith basis function, and w_{ki} is the ith weight for d_k, q is the total number of basis functions. The basis functions are shared by the whole spectrum set and is called principal component (PC); the weight define the relative contribution of each basis function to the spectrum and is called principal component weight (PC weight),
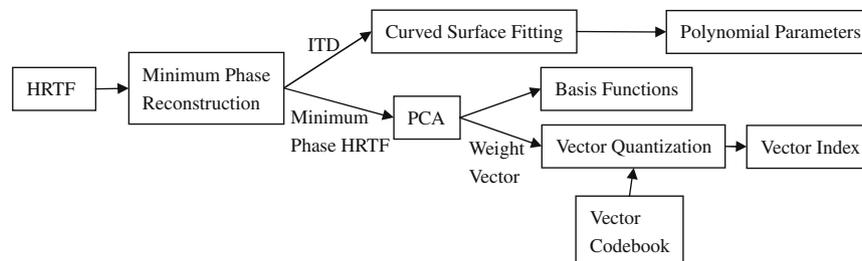

Fig. 1. Workflow diagram of the hybrid compression method for HRTF.

it remains unique for each HRTF. The number of basis functions required to provide an adequate representation of the data depends on the amount of redundancy or correlation presented in the data set. Generally 5–10 PCs are required to account for approximately 90% of the variation in the HRTF magnitude functions. The detailed process of principal component analysis is described in [15].

2.2. Vector quantization of the weight vector

The weights derived from principal component analysis represents the contribution of the basis functions to the spectrum. Although principal component analysis reduces the HRTF data size greatly, the total memory size of weight vectors is still huge when there are a large number of HRTFs. As the variation of the weight as a function of HRTF direction has no evident regularity, it is difficult to fit it with a mathematical equation precisely. In this paper a vector quantization method is proposed to reduce the size of the entire weight vector set.

Vector quantization is an efficient compression method, in which the basic idea is to represent the set of scalars as a single vector and quantize them jointly in the vector space. In vector quantization, the N dimensional vector $x = [x_1, \ldots, x_N]$ is represented by the nearest matching vector from the set of N dimensional vectors $Y = [y_1, \ldots, y_L]$, where Y is referred to as the vector codebook, L is the size of the codebook, and y_i is the ith N dimensional code vector in the codebook. The matching metric is to find a code vector y_{index} in the codebook with the smallest distortion:

$$\text{index} \stackrel{def}{=} \min_i \{\|x - y_i\|^2\}, \quad i = 1, \ldots, L \tag{4}$$

We say that x is quantized as y_{index}, and y_{index} is the quantized value of x. Only the index of the codebook is sent instead of the quantized value. This can conserves space and achieves more compression.

The split vector quantization method is used to improve the quantization precision, where the weight parameters are split into a number of parts and each part is quantized separately using vector quantization [19,20]. The experiment in Section 3 will illustrate the effectiveness of the split vector quantization method.

2.3. Curved surface fitting of ITD

ITD, the group delay of the all pass filter H_{ap} in (1), is unique for each HRTF. Curved surface fitting is employed to decrease the size of the ITD data [21]. The variation of ITD as a function of azimuth and elevation is fitted with the cosine series bivariate polynomial. The cosine series bivariate polynomial is defined as

$$z(x, y) = \sum_{j=0}^{Q} \sum_{v=0}^{Q-j} \beta_{jv} \cos(jx) \cdot \cos(vy) \tag{5}$$

where z is the fitted ITD, x and y are azimuth and elevation normalized to $[0, \pi]$, respectively, Q is the polynomial order and β_{jv} is the polynomial coefficient. Denote Eq. (5) as a matrix form

$$z = cb \tag{6}$$

where $b = [\beta_{0,0}, \ldots, \beta_{0,Q}, \beta_{1,0}, \ldots, \beta_{1,Q-1}, \ldots, \beta_{Q,0}]^T$ is the coefficient vector which consists of $\{\beta_{jv}\}$ arrayed in ascending order of j and v, respectively, and $c = [1, \ldots, \cos y, \cos x, \ldots, \cos x \cos(Q-1)y, \ldots, \cos Qx]$ according to the elements in b. The size of b is $L = \frac{1}{2}(Q+1)(Q+2)$.

Suppose there are N ITD variables, we can get

$$Z = Cb \tag{7}$$

where $Z = [z_1, \ldots, z_N]^T$, $C = [c_1, \ldots, c_N]^T$ and b is the coefficient vector. The least-squares solution to Eq. (7) is given by

$$b = (Z^T Z)^{-1} Z^T C \tag{8}$$

For each subject, the polynomial coefficient vector b is stored instead of ITDs of all directions. This saves the storage space greatly.

3. Experiments and analysis

In this section we carry out experiments for the proposed method in three aspects separately: principal component analysis, vector quantization, and surface fitting. The CIPIC HRTF database (from CIPIC Interface Laboratory, University of California) is used as the experimental data.

3.1. CIPIC HRTF database

A lot of research projects of institutions and universities have collected some libraries of HRTF measurements in their anechoic chamber. One of the famous libraries is called the CIPIC database, which is available on the Internet [22]. The HRTF measurements were made in the anechoic chamber of U. C. Davis CIPIC Interface Laboratory. The database includes head-related impulse responses for 45 subjects at 25 different azimuths ($-80°$–$80°$) and 50 different elevations ($-45°$–$231°$) with approximately $5°$ angular increments (1250 positions totally). The impulse response of each HRTF measurement is 200 taps long at the sample rate 44.1 kHz. It is noticed that the HRTFs of the left and right ears are measured separately for the subject. Since it is hypothesized that the left and right ears of normal human being are symmetric, only the measured HRTFs of a single ear are enough for 3D sound processing. Here we take the HRTFs of left ears as the experiment data, and the storage of the HRTFs of all 45 subjects (a total of 56,250 HRTFs) requires a large amount of memory.

3.2. Experiments on principal component analysis of HRTF

Principal component analysis is applied to the log-magnitude functions of all the HRTFs (a total of 56,250 HRTFs) to derive the basis functions, or principal components (PC). All the 200 log-magnitudes in the full frequency region are included in the analysis. The basis functions are the q eigenvectors with the q largest eigenvalues of the magnitude covariance matrix, and the contribution of each basis function to data variation is proportional to its eigenvalue. Fig. 2 depicts the percentage variance of HRTF repre-
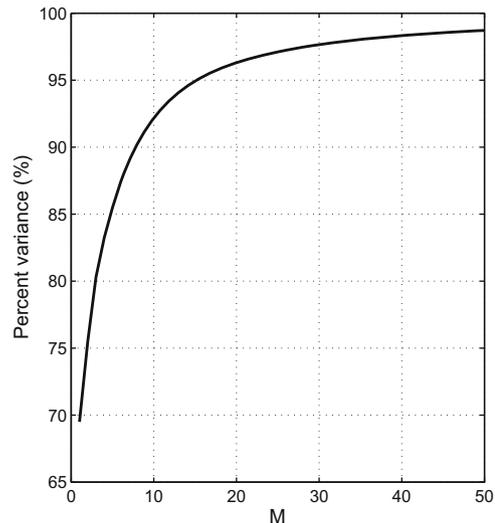

Fig. 2. Percent variance of HRTF represented by first M basis functions.

sented by the first M eigenvalues as a function of M, which is given by

$$\text{var}(M) = \frac{\sum_{i=1}^{M} \lambda_i}{\sum_{i=1}^{N} \lambda_i} \times 100\% \tag{9}$$

where the percentage variance $\text{var}(\cdot)$ is a function of M, λ_i is the eigenvalue corresponding to the ith basis function, and $N = 200$ is the total number of eigenvalues. Fig. 2 shows that the first few eigenvalues represent most of the variations contained in the HRTFs. We choose 15 PCs to account for about 95% of the variation in the HRTF magnitude. Fig. 3 depicts the percentage contributions of the first 10 PCs to the variation of the data set, which is given by

$$\text{con}_i = \frac{\lambda_i}{\sum_{j=1}^{N} \lambda_j} \times 100\% \tag{10}$$

where con_i is the percentage contribution of the ith PC, and λ_i is the ith eigenvalue and $N = 200$ is the total number of eigenvalues. Fig. 4 depicts the spectrum of PC 1–5, respectively. It is shown that the first basis function and its weight (which we designate as PC 1) capture the majority of common variation presented in the data, and the remaining basis functions and weights (PC 2, PC 3, etc.) reflect decreasing common variation and increasing unique variation. To display the variation of the weights, put PC 1 weight of all the 56,250 HRTFs into a big array, where the position of HRTFs are arranged according to the ascending order of subject, elevation, and azimuth, respectively. And the same operation is applied to other weights separately. The PC1–5 weights across all the HRTFs are shown in Fig. 5 with the horizon axes representing the relative position of the weight in the vector, and the vertical axes representing the weight value. The distribution exhibit no evident regularity except PC 1 weight. Vector quantization is employed to compress the weight set.

3.3. Experiments on vector quantization of the weight set

The entire weight set derived from principal component analysis, which consists of 56,250 vectors of dimensionality 15, is used for both training and testing. Split vector quantization for weight vector is employed. We split the weight vector into 5 parts: PC 1, PC 2–3, PC 4–6, PC 7–10, and PC 11–15 according to the contribution of each basis function to data variation. A 256-level vector quantizer is designed separately for each part using the data in
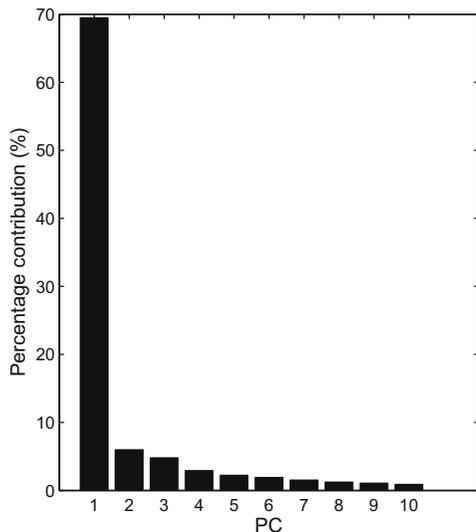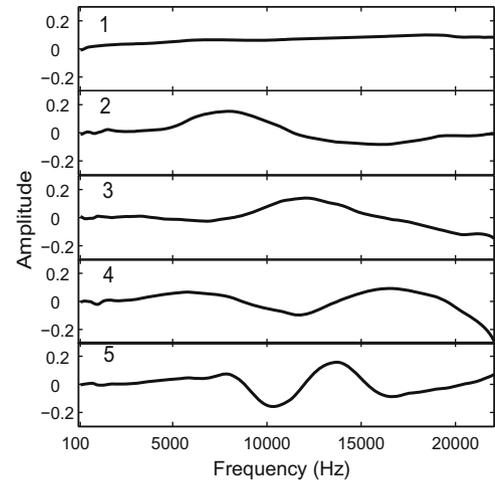

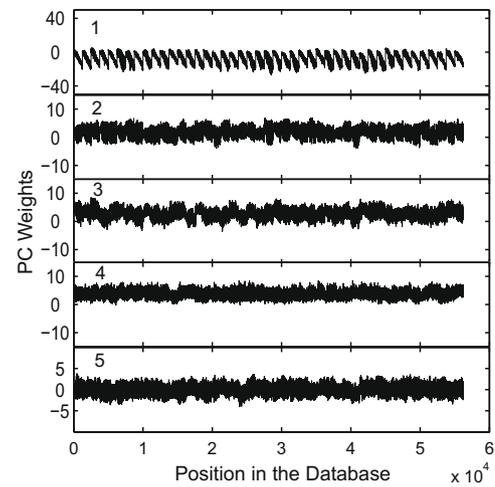**Fig. 4.** The first five basis functions.

Fig. 5. Weights of each PC.

the training set. Thus each vector of 15 parameters can be represented with the codebook index of 40 bits. A classical vector quantization design algorithm (often termed LBG-algorithm originated by Linde, Buzo, and Gray [19]), is used here with the Euclidean distance measure for designing these vector quantizers. The quantization results at 40 bits/vector are computed from the test data. To evaluate the quantization performance of the vector quantization method, we employ the spectral distortion (SD) measure in speech coding. The spectral distortion for the ith weight vector, SD_i, is defined (in dB) as follows.

$$\text{SD}_i^2 = \frac{1}{\pi} \int_0^{\pi} [20\log_{10}|P_i(\omega)| - 20\log_{10}|\widehat{P}_i(\omega)|]^2 d\omega \tag{11}$$

where $P_i(\omega)$ and $\widehat{P}_i(\omega)$ are the unquantized and quantized HRTF magnitude, respectively. Table 1 shows the quantization performance of the split vector quantizer. Little quantization distortion

Fig. 3. Percentage contribution of each basis function.

Table 1
Spectral distortion (SD) performance of the split vector quantizer.

Bits used	Avg. SD (in dB)	Outliers (in %)	
		2–4 dB	>4 dB
40	0.91	0.32	0

may be noticed with an average SD across all HRTFs of about 0.91 dB, about 0.32% outliers in the range 2–4 dB, and no outlier with SD greater than 4 dB.

To evaluate the reconstruction precision of the hybrid compression method employing both PCA and vector quantization, we define the percent mean square error (PMSE) of the reconstructed HRTF relative to the original one as

$$e_i = \frac{\|h_i - \hat{h}_i\|^2}{\|h_i\|^2}\% \tag{12}$$

where h_i and $\hat{h}_i$ are the ith original and reconstructed HRTF log-magnitude, respectively, and $\|\cdot\|$ denotes 2-norm operation. For convenience, we call the PCA reconstructed HRTF the unquantized HRTF, and call the PCA and vector quantization reconstructed HRTF the quantized HRTF. Table 2 presents the average PMSE across all HRTFs, where e_{o_u} denotes the error of the unquantized HRTF relative to the measured one, e_{o_q} denotes the error of the quantized HRTF relative to the measured one, and e_{u_q} denotes the error of the quantized HRTF relative to the unquantized one. The relative error of quantized HRTF to measured HRTF (4.13%) is only a little larger than that of the unquantized HRTF to measured HRTF (3.62%), which demonstrates that the error of the proposed method is nearly identical to that of the PCA method. It also reveals that about 95% of the variation is retained by PCA. On the other hand, the small relative error (0.75%) of the quantized HRTF to the unquantized HRTF also demonstrates the transparent quantization quality of the quantization method.

Fig. 6 compares the measured, quantized, and unquantized HRTF magnitudes of subject "028" at three typical locations. The solid lines represent measured HRTFs, the dashed lines represent quantized HRTFs, and the dotted lines represent unquantized HRTFs. The azimuth and elevation of each HRTF is noted by symbols "AZ=" and "EL=", respectively. It can be seen from Fig. 6 that the magnitude curves of the quantized HRTFs nearly coincide with those of the unquantized ones, and both of them can capture the primary variations in the measured HRTF magnitude.

The split vector quantization can obtain transparent quantization quality, and thus the magnitude error between the proposed method and the PCA method is unnoticeable. The magnitude distortion between the original data and the above two methods is mainly determined by the number of the retained basis functions, where 15 basis functions may obtain about 95% of the variation of the magnitude and more reconstruction precision may be obtained by increasing the number of the retained basis functions.

3.4. Experiments on surface fitting of ITD

The ITD variation with azimuth and elevation is surface fitted for each subject using least-squares fitting method. The cosine series bivariate polynomials of order 3–7 are used for fitting, respectively. To evaluate the fitting performance, we define the mean square error (MSE) of fitting as

$$E = \frac{1}{N} \sum_{i=1}^{N} (z_i - \hat{z}_i)^2 \tag{13}$$

where N is the ITD number of each subject (1250 for each subject), z_i and $\hat{z}_i$ are the ith original and fitted ITD, respectively. The average MSE across 45 subjects is calculated for the fitting polynomial of order 3–7, respectively. The result is presented in Table 3. It is shown

Table 2
Average PMSE between original and reconstructed HRTF.

	avg. PMSE (%)
e_{o_u}	3.62
e_{o_q}	4.13
e_{u_q}	0.75

Table 3
Average fitting mean square errors as functions of polynomial order.

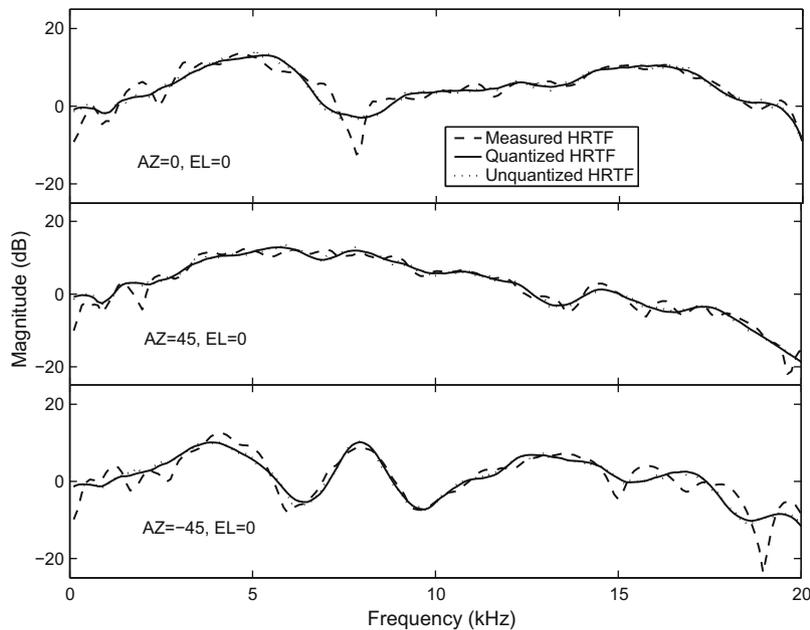| Polynomial order	Parameter number	avg. MSE
3	10	1.82
4	15	1.37
5	21	1.24
6	28	0.53
7	36	0.50

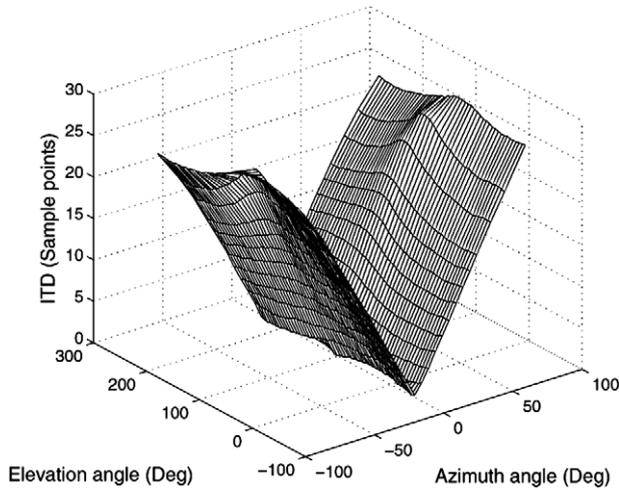**Fig. 6.** Comparison between measured and quantized HRTFs at three locations.

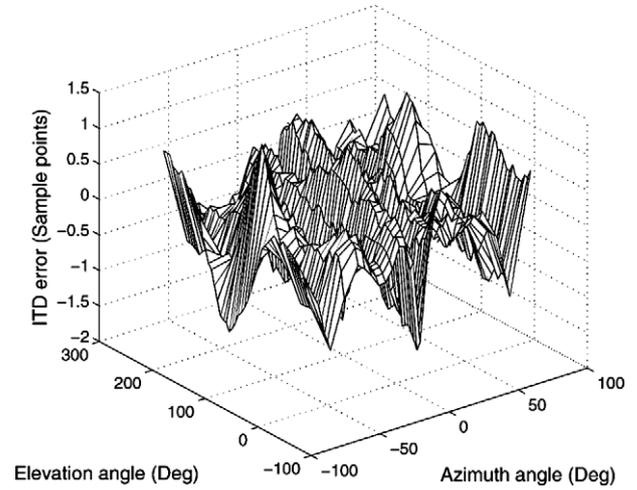**Fig. 7.** Original ITD variation as a function of azimuth and elevation.

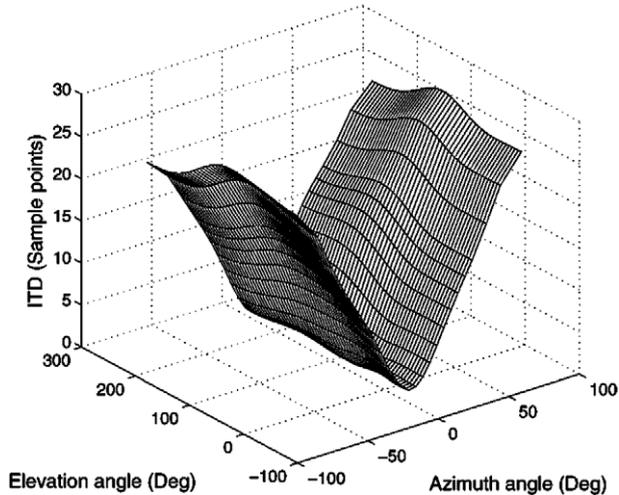**Fig. 9.** Fitted error variation as a function of azimuth and elevation.

Fig. 8. Fitted ITD variation as a function of azimuth and elevation.

Table 4
Memory consumption of the proposed method for CIPIC HRTF database.

Data class	Parameter number	Memory (bytes)
Polynomial coef.	45×28	5040
Basis functions	15×200	12,000
Vector indexes	45×1250	281,250
Code book	256×15	15,360
Total	/	313,650

3.6. Comparison of memory consumption

To indicate the compression efficiency of the proposed method, we compare the memory consumption between the original data and the compressed data. Suppose the storage of a float-point number requires 4 bytes. The storage of the original CIPIC requires $56,250 \times 200 \times 4 = 4.5 \times 10^7$ bytes, which consumes an extraordinary large memory space. The principal component analysis method requires about 3.612×10^6 bytes for 15 basis functions, 56,250 weight vectors and 56,250 ITDs. The memory consumption of the data processed by the proposed method is shown in Table 4, in aspects of polynomial coefficients, basis functions, vector indexes, and vector codebook, respectively. It is shown in Table 4 that the total memory consumption of the data processed by the proposed method is about 3.137×10^5 bytes, which is about 0.7% of the original memory consumption, and 8.7% of that of principal component analysis. Thus about 99.3% of storage space is saved compared to the original one, and about 91.3% of the storage space is saved compared to the principal component analysis method, which is a remarkable compression result.

4. Conclusion

A hybrid compression method is proposed in the paper to reduce the large memory consumption of the HRTF database, which consists of HRTF sets of multi-subjects. This method employs, principal component analysis, vector quantization, and curved surface fitting jointly to compress the HRTF data. Principal component analysis decomposes the minimum-phase reconstructed HRTF magnitudes into a set of basis functions and a weight vector set. Vector quantization and surface fitting are applied to the weight vector set and the ITD data set, respectively, for further compression. All the parameters have been computed offline and stored in the memory beforehand. When binaural synthesis, the HRTF

in Table 3 that the fitting performance increases with the polynomial order, and the increase becomes slowly at order 6. Taking the performance and parameter number both into consideration, we take 6 as the polynomial fitting order. Figs. 7–9, respectively, display the original ITD variation, fitted ITD variation, and fitted residual variation as the function of azimuth and elevation for subject "028". It can be seen from Figs. 7–9 that the ITD variation can be surface fitted with cosine series bivariate polynomial of order 6 with high precision, where the maximum absolute fitted error does not exceed 1.5 sample points. Thus the error of the curving fitting method is unnoticeable.

3.5. Subjective evaluation

The reconstruction precisions of the proposed method and the PCA method are nearly identical, thus the two methods have similar sound localization performances and subjective evaluation results. The sound localization performance is mainly related to the number of the retained basis functions. The subjective evaluation is not discussed in this paper. Researchers who are interested in it may refer to [15] where the subjective evaluation is carried out for the PCA method.

can be reconstructed with its corresponding parameters read from the memory. The HRTF reconstruction process does not need much computation. Simulation results demonstrate that the proposed method can obtain similar HRTF reconstruction precision with only 8.7% of the memory space required by the principal component analysis method. This is a remarkable compression result, and thus the proposed method is potential for 3D sound processing. In addition, although this method is developed primitively for HRTF data sets of multi-subjects, it can also be applied to the HRTF data of a single subject with similar compression result.

Acknowledgements

References

[1] Begault DR. 3D sound for virtual reality and multimedia. London: Academic Press; 1994.
[2] Cheng CI, Wakefield GH. Introduction to head-related transfer function (HRTFs): representations of HRTF in time, frequency and space. J Audio Eng Soc 2001;49(4):231–49.
[3] Wenzel EM, Arruda M. Localization using nonindividualized head-related transfer functions. J Acoust Soc Am 1993;94(1):111–23.
[4] Tan CJ, Gan WS. User-defined spectral manipulation of HRTF for improved localization in 3D sound systems. Electron Lett 1998;34(25):2387–9. Dec..
[5] Park MH, Choi SI, Kim SH. Improvement of front–back sound localization characteristics in headphone-based 3D sound generation. In: The 7th international conference on advanced communication technology, Phoenix Park, South Korea; 2005. p. 373–85.
[6] Seeber B, Fastl H. Subjective selection of non-individual head-related transfer functions. In: Proceedings of the 2003 international conference on auditory display, Boston, MA, USA; 2003. p. 373–85.
[7] Iwaya Y. Individualization of head-related transfer functions with tournament-style listening test: listening with other's ear. Acoust Sci Technol 2006;27(6): 340–3.
[8] Algazi VR. Approximating the head-related transfer function using simple geometric models of the head and torso. J Acoust Soc Am 2002;112(5): 2053–64.
[9] Xiao T, Liu QH. Finite difference computation of head-related function of human hearing. J Acoust Soc Am 2003;113(5):2434–41. May.
[10] Huopaniemi J, Karjalainen M. Review of digital filter design and implementation methods for 3D sound. In: Proceedings of the 102th convention of the audio engineering society, Preprint 4461, Munich, Germany; 1997.
[11] Kulkarni A, Colburn HS. Infinite-impulse-response models of the head-related transfer function. J Acoust Soc Am 2004;115(4).
[12] Brown CP, Duda RO. A structural model for binaural sound synthesis. IEEE Trans Speech Audio Process 1998;6(5):476–88.
[13] Chen J, Van Veen BD, Hecox KE. External ear transfer function modeling: a beamforming approach. J Acoust Soc Am 1992;92(4):1933–44.
[14] Bai MR, Ou KY. Head-related transfer function (HRTF) synthesis based on a three-dimensional array model and singular value decomposition. J Sound Vib 2005;281:1093–115.
[15] Kistler DJ, Wightman FL. A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction. J Acoust Soc Am 1992;92(3):1637–47.
[16] Chen J, Van Veen BD, Hecox KE. A spatial feature extraction and regularization model for the head related transfer function. J Acoust Soc Am 1995;97(1):439–52.
[17] Wang L, Yin F, Chen Z. HRTF compression via principal components analysis and vector quantization. IEICE Electron Exp 2008;5(9):321–5.
[18] Kulkarni A, Isabelle SK, Colburn HS. On the minimum-phase approximation of head-related transfer functions. In: Proceedings of IEEE ASSP workshop on applications of signal processing to audio and acoustics, Mohonk Mountain House, New Paltz, New York; October 1995.
[19] Linde Y, Buzo A, Gray RM. An algorithm for vector quantizer design. IEEE Trans Commun COM 1980;28(1):84–95.
[20] Paliwal KK, Atal BS. Efficient vector quantization of LPC parameters at 24 bits/frame. IEEE Trans Speech Audio Process 1993;1(10):3–14.
[21] Ahn SJ. Least squares orthogonal distance fitting of curves and surfaces in space. Berlin: Springer; 2004.
[22] Algazi VR, Duda RO, Thompson DM. The CIPIC HRTF database. In: Proceedings of 2001 IEEE workshop on applications of signal processing to audio and acoustics, New York; October 2001. p. 99–102.