

# HRTF compression via principal components analysis and vector quantization

Lin Wang<sup>a)</sup>, Fuliang Yin<sup>b)</sup>, and Zhe Chen

School of Electronics and Information Engineering,  
Dalian University of Technology, China

a) wanglin\_2k@sina.com

b) flyin@dlut.edu.cn

**Abstract:** The computation burden and the huge memory size of the head-related transfer function (HRTF) database challenge the practical application of 3D sound system. This paper therefore proposes a novel method which employs principal components analysis and vector quantization jointly to reduce the size of HRTF set in 3D sound system. Numerical experiment results demonstrate that the proposed method can save the storage space greatly with little localization performance degradation.

**Keywords:** HRTF, data compression, principal components analysis, vector quantization

**Classification:** Science and engineering for electronics

## References

- [1] D. R. Begault, *3D sound for Virtual Reality and Multimedia*, London: Academic Press, 1994.
- [2] J. Huopaniemi, and M. Karjalainen, "Review of digital filter design and implementation methods for 3D sound," *Proc. of the 102th Convention of the Audio Engineering Society*, Preprint 4461, Munich, Germany, 1997.
- [3] A. Kulkarni, and H. S. Colburn, "Infinite-impulse-response models of the head-related transfer function," *J. Acoust. Soc. Am.*, vol. 115, no. 4, April 2004.
- [4] 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. America*, vol. 92, no. 3, pp. 1637–1647, March 1992.
- [5] 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.*, vol. 97, no. 1, pp. 439–452, Jan. 1995.
- [6] V. R. Algazi, R. O. Duda, and D. M. Thompson, "The CIPIC HRTF database," *Proc. of 2001 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, pp. 99–102, New York, Oct. 2001.
- [7] Y. Linde, A. Buzo, and R. M. Gray, "An algorithm for vector quantizer design," *IEEE Trans. Commun.*, vol. COM-28, no. 1, pp. 84–95, Jan. 1980.

- [8] K. K. Paliwal and B. S. Atal, “Efficient vector quantization of LPC parameters at 24 bits/frame,” *IEEE Trans. Speech Audio Process.*, vol. 1, no. 10, pp. 3–14, Jan. 1993.

## 1 Introduction

Head-related transfer function (HRTF) serves the dominant role of implementation for the 3D sound system [1]. But the huge memory space of the HRTF data challenges the practical application of the 3D sound system. Much research work has been done to reduce the size of the HRTF data, such as pole-zero modeling [2, 3], structural or mathematical decomposition [4, 5], and so on, nevertheless the problem is still considered as an open question when there are a large amount of HRTFs. This paper proposes a method employing principal components analysis and vector quantization jointly to reduce the HRTF data size greatly with little localization performance degradation.

## 2 Proposed method

The flow diagram of the proposed method is shown in Fig. 1. First of all, principal components analysis (PCA) is applied to the entire HRTF magnitude set to derive a small set of basis functions and the corresponding principal component weight vectors. Second, the weight vector set is vector quantized, where a weight vector can be represented with a vector index in the codebook. The proposed method is described in detail as follows.

The key idea of principal components analysis is to reduce the dimensionality of a data set while retaining the primary variation in the data [4]. PCA decomposes a set of HRTF magnitude spectrum into basis functions and the corresponding principal component weight vectors. This is expressed as

$$d_k = \sum_{i=1}^q w_{ki} c_i \quad (1)$$

where  $d_k$  is the  $k$ th magnitude spectrum in the set,  $c_i$  is the  $i$ th basis function, and  $w_{ki}$  is the  $i$ th weight for  $d_k$ ,  $q$  is the total number of basis functions. Generally the number of basis functions is much smaller than the dimensionality of  $d_k$ , thus the data size is reduced greatly. The detailed process of principal components analysis can be found in [4].

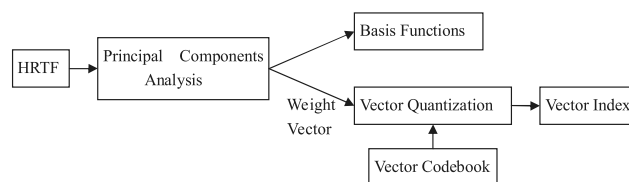


Fig. 1. Workflow of the proposed method

Although PCA reduces the HRTF data size greatly, the weights of each HRTF still consume a lot of memory when there are a large number of HRTFs. Since the variation of the weight as a function of HRTF direction has no regularity, it's difficult to fit it with a mathematical equation precisely. Vector quantization is employed to compress the weight data. 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  $\mathbf{x}$  is represented by the nearest matching vector from the set of  $N$  dimensional vectors  $Y = [y_1, \dots, y_L]$ , where  $Y$  is referred to as the vector codebook,  $L$  is the size of the codebook. It's evident that only the index of the codebook is stored instead of the quantized value. This may conserve the storage space and achieve more compression. In this paper the LBG algorithm is employed for vector quantizer design [7]. The split vector quantization method is also employed to improve the quantization precision [8], where the weight parameters are split into a number of parts and each part is quantized separately using vector quantization.

The basis functions and the vector codebook are shared by the entire HRTF database, 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.

It should be mentioned that PCA only models the magnitude components of the HRTFs, and 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 [4].

### 3 Experiment and analysis

The proposed method may reduce the memory consumption greatly especially for a large number of HRTFs. In this section experiments on the proposed method are carried out using the CIPIC HRTF database [6], which includes head-related impulse responses for 45 subjects (each measured at 1250 spatial positions). The impulse response of each HRTF measurement is 200 taps long at the sample rate 44.1 kHz. The storage of the whole database (a total of 56250 HRTFs) requires a lot of memory. PCA and the proposed method are applied to the HRTF database respectively and the compression results are compared.

Principal components analysis is applied to the log-magnitude functions of all the 56250 HRTFs. We choose 10 basis functions to account for about 92% of the variation in the HRTF magnitude functions. It has been proved in [4] that the HRTF localization performance degraded little when above 90% of variation is retained. The weight set, which consists of  $45 \times 1250$  vectors of dimensionality 10, is used for both vector quantizer training and testing. The weight vector is split into four parts according to the contribution of each basis function to data variation:  $w_1$ ,  $w_2 - w_3$ ,  $w_4 - w_6$ , and  $w_7 - w_{10}$ . A

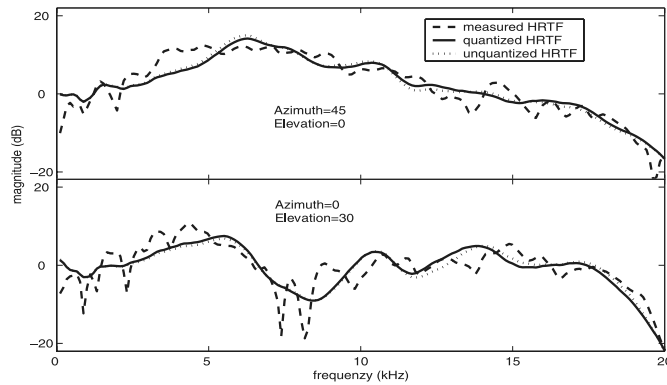


Fig. 2. Comparison between measured and quantized HRTFs at four locations

256-level vector quantizer is designed separately for each of these four parts. Thus each vector of 10 parameters can be represented with the codebook index whose size is 32 bits.

To evaluate the reconstruction precision of the proposed method, we define the percent mean square error (PMSE) of the reconstructed HRTF relative to the original one as

$$e_i = \frac{\|h_i - \hat{h}_i\|^2}{\|h_i\|^2} \times 100\% \quad (2)$$

where  $h_i$  and  $\hat{h}_i$  are the  $i$ th original and reconstructed HRTF log-magnitude respectively, and  $\|\cdot\|$  denotes 2-norm operation. The average PMSE across all the 56250 HRTFs is calculated for the PCA method and the proposed method respectively. The error of the PCA method, i.e. the average PMSE of the unquantized HRTF (PCA reconstructed) relative to the measured one, is 6.71%. The error of the proposed method, i.e. the average PMSE of the quantized HRTF (PCA and vector quantization reconstructed) relative to the measured one, is 7.23%. And the error of vector quantization, i.e. the average PMSE of the quantized HRTF relative to the unquantized one is 0.61%. It can be said that the error of vector quantization is unnoticeable, and the error of the proposed method is nearly the same as that of the PCA method. Thus the HRTF data may be compressed with the proposed method with little localization performance degradation.

Fig. 2 compares the measured, quantized, and unquantized HRTF magnitudes of subject “028” at two typical positions. It can be seen from Fig. 2 that the magnitude curves of the quantized HRTFs nearly coincide with those of the unquantized ones, and both of them can approximate the curves of the measured ones precisely. This also verifies the effectiveness of the proposed method.

Suppose the storage of a float-point number requires 4 bytes. The storage of the original CIPIC HRTFs requires  $56250 \times 200 \times 4 = 4.5 \times 10^7$  bytes, which consumes an extraordinary large memory space. The principal components analysis method requires about  $2.26 \times 10^6$  bytes for 10 basis functions and 56250 weight vectors. The memory cost of the proposed method is summa-

**Table I.** Memory consumption of the proposed method

Data class	Parameter number	Memory (bytes)
Basis functions	$10 \times 200$	8000
Vector indexes	56250	225000
Code Book	$256 \times 10$	10240
Total	/	243240

rized in Table I from aspects of basis functions, vector indexes, and codebook respectively. The total memory consumption of the proposed method is about  $2.43 \times 10^5$  bytes, which is only about 0.54% of the original one, and 10.7% of that of principal components analysis. Thus about 89.3% of the storage space is saved compared to the principal components analysis method, which is a remarkable improvement.

#### 4 Conclusion

Principal components analysis reduces the HRTF data size greatly, but the principal component weights of each HRTF still consume much memory when there are a large number of HRTFs. The proposed method represents the weight as the codebook index to conserve storage space and achieve more compression. Numerical experiment results demonstrate that the proposed method can reduce the memory consumption remarkably with little localization performance degradation.

#### 5 Acknowledge

This work is supported by the National Natural Science Foundation of China under Grant No. 60772161 and 60372082.