Front. Electr. Electron. Eng. China 2007, 2(4): 388–394
DOI 10.1007/s11460-007-0072-1

RESEARCH ARTICLE

ZHANG Jie

A non-linear spatial hearing model based on bases pursuit algorithm

Abstract In the research on spatial hearing and realization of virtual auditory space, it is important to effectively model the head-related transfer functions (HRTFs) or head-related impulse responses (HRIRs). In our study, we managed to carry out adaptive non-linear approximation in the field of wavelet transformation. The results show that the HRIRs' adaptive non-linear approximation model is a more effective data reduction model, is faster, and is 5 dB on average better than the traditional principal component analysis (PCA) (Karhunen-Loève transform) model based on relative mean square error (MSE) criterion. Furthermore, we also discussed the best bases' choice for the time-frequency representation of HRIRs, and the results show that local cosine bases are more propitious to HRIRs' adaptive approximation than wavelet and wavelet packet base. However, the improved effect of local cosine bases is not distinct. Here, for the sake of modeling the HRIRs more truthfully, we consider choosing optimal time-frequency atoms from redundant dictionary to decompose this kind of signals of HRIRs and achieve better results than all the previous models.

Keywords spatial hearing model, wavelet transformation, non-linear approximation, bases pursuit algorithm

1 Introduction

It is a well-known fact that the human auditory system works well in determining a sound's positions in an acoustic environment. This phenomenon has attracted researchers' attention for a long time, with so many investigations and interpretations present in the field. So far, there are many mathematical models that aim to solve this problem. The most classic and effective solution is the duplex theory, which utilizes interaural time difference (ITD) and interaural intensity difference (IID) to interpret the cause of spatial hearing. By virtue of these results, there have been some tentative applications in auditory display of multimedia systems, for instance, VoiceNotes of MIT Media Lab. However, there are still no fully successful models capable of explaining some of the phenomena associated with spatial hearing, such as the perception of a sound's position at an elevation, the "cone of confusion" and so on.

In the ensuing research, people take cognizance of the cues implied in grotesque while a little inerratic spectrum. Thus, the systemic description of the entire relevant information, head-related transfer function (HRTF), has been introduced to depict them, where ITD can be seen as the delay between two ears and IID as the difference of HRTFs' magnitude. Moreover, people have constructed different mathematical models to simulate this delicate function, and tried to give a felicitous interpretation for this fancy phenomenon [1−4].

As a popular model structure, the principal component analysis (PCA) model has been greatly studied in this field [2−4]. Basically, researchers in Ref. [2] applied PCA to logarithmic magnitudes of HRTFs that were measured for 10 subjects and at 256 positions. In the PCA model, HRTFs for each position were represented as a weighted combination of a set of basis functions in a low-dimensional subspace, which were obtained from the measured logarithmic magnitudes of HRTFs. While the phase was approximated from the magnitude related to a certain position based on minimum phase characteristics. The spatial feature extraction and regularization model put forward in Ref. [3] contains the phase information during the HRTFs' reduction representation. Furthermore, a thin-plate spline using regulation procedures was also introduced to obtain a mathematical representation for sampled spatial positions. As a result, unmeasured positions' HRTFs can be approximately reestablished from the mathematical equation, which effectively surmount the spatially discrete limitation of HRTFs' measurement. However, both these PCA models were used to process complex valued (amplitude and phase) head-related transfer functions, and needed a lot of complex and logarithmic operations. In Ref. [4], Wu applied Karhunen-Loève transformation to HRIRs in time

Received December 8, 2006; accepted March 15, 2007

ZHANG Jie (✉)
The CETC 14th Research Institute, College of Information Science and Engineering, Southeast University, Nanjing 210096, China
E-mail: zjhf1978@yahoo.com.cn

domain to obtain a PCA model, which only required real-valued operations and behaved effectively for real-time implementation of virtual auditory space (VAS).

However, traditional orthogonal bases generally cannot analyze a signal's characteristics in the temporal and spectral domain at one time because of the compact support wavelet functions' capability in time-frequency localization and multi-scale analysis, which have extensive applications in a signal's nonlinear, optimal, and all-sided approximation. Reference [5] indicates that, as to a function $h(t) \in L^2(R)$, there are finite wavelet bases functions to approximate it with arbitrary precision; and the approximation to impulse response of a system using orthogonal wavelet and scale functions have also been studied. Our work just discovers an origin from this, and considers carrying out adaptive non-linear approximation in the field of head-related impulse responses (HRIRs)' wavelet transformation. Recently, there have been some reports of HRTFs' modeling by wavelet transformation. For example, Ref. [6] constructed a group of sparse filters to model HRTF based on the wavelet's multi-scale characteristics, the results of which showed the excellence of filter-bank's model than HRTF's conventional filter design algorithms, Prony, Yule-Walker, and BMT. However, in this paper, on the basis of the distribution characteristic identification of all the HRIRs' data in KEMAR package, our work points out that, because HRIRs cannot well accord with the Gaussian distribution, the PCA reduction model based on Karhunen-Loève transform is not optimal and the non-linear method can work well. As for this, we have carried out detailed data processing experiments to validate our opinion [7].

For this kind of signals just as HRTFs, Batteau et al made some investigations on the HRIRs in time domain [8,9]. They found that the impulse response could be piled up from echoes reflected on outer ears

$$h(t) = \delta(t) + a_1\delta(t-\tau_1) + a_2\delta(t-\tau_2) \tag{1}$$

where a_1 and a_2 are the reflection coefficients; τ_1 and τ_2 are the time delay of reflected signals. Furthermore, when the elevation angle moves towards lower latitudes, the time delays usually increase. Hiranaka and Yamasaki further validated the above phenomena, and found that, all the reflection delays are less than 350 μs for humans through experimentation. Moreover, they also discovered that, when sound source is located in front of the human body, there are at least two reflection waves; at the backside, there is only one; while on top of the head, there are scarcely any reflection components [10]. Wright also confirmed the relationship between HRTFs' characteristics and delay of sound's reflections [11].

All these reveal the facts that, as for the kind of HRIRs' signals, the best bases choice for time-frequency representation of HRIRs is also necessary and important for the efficient modeling of HRIRs. Some simulation results have shown that local cosine bases are more propitious to HRIRs' adaptive approximation than wavelet and wavelet packet base in our experiments. However, the improved effect of local cosine bases is not distinct. Here, for the sake of modeling the HRIRs more truthfully, we consider choosing optimal time-frequency atoms from redundant dictionary [12] to decompose this kind of signals of HRIRs and achieve rather better results than all the previous models.

Synthetically, the paper is organized as follows. In Sect. 2, the algorithm of HRIRs' adaptive non-linear approximation based on bases pursuit algorithm is introduced. Then the contrasting results with PCA's, wavelet's, wavelet packet's, and local cosine packet's model for HRIRs are presented in Sect. 3. After that, several next-step directions are brought forward for future study.

2 Principle of best bases' selection for HRIR's spectro-temporal representation

To optimize non-linear approximation for HRIRs' signal, we can adaptively choose some bases according to the signals' characteristics. The essential scheme is that, first, we construct a concave cost function, such as entropy or l^p function. Then the best bases are selected from a dictionary of bases, such as wavelet packet bases, local cosine bases or even a redundant dictionary for pursuit algorithm by minimizing the cost function. In the following contents, the approximation models of HRIRs are compared under different basis dictionary of PCA, wavelet, wavelet packet, local cosine bases, and a redundant dictionary for pursuit algorithm to evaluate their performance.

Suppose a group of compact support wavelet set in the space $L^2(R)$ is $B = \left[\left\{\phi_{J,k}\right\}_{0\leqslant k\leqslant 2^{-J}}, \left\{\psi_{j,k}\right\}_{-\infty<j\leqslant J,\ 0\leqslant k<2^{-J}}\right]$, then the optimal nonlinear approximation of signal $\boldsymbol{h} \in L^2(R)$ using M wavelets is

$$h_M = \sum_{(j,k)\in I_M} \left\langle \boldsymbol{h}, \boldsymbol{\psi}_{j,k} \right\rangle \boldsymbol{\psi}_{j,k} \tag{2}$$

Here, I_M is the subscript set of the M biggest wavelet coefficients $\left|\left\langle \boldsymbol{h}, \boldsymbol{\psi}_{j,k} \right\rangle\right|$, which correspond to the most remarkable pertinence between signal $\boldsymbol{h}$ and wavelets function, and can be regarded as the primary characteristics of signal $\boldsymbol{h}$. Then, the minimum error of non-linear approximation is

$$\varepsilon(M) = \left\|\boldsymbol{h} - \boldsymbol{h}_M\right\|^2 = \sum_{(j,k)\notin I_M} \left\langle \boldsymbol{h}, \boldsymbol{\psi}_{j,k} \right\rangle \boldsymbol{\psi}_{j,k} \tag{3}$$

where $h(l) = \left\langle \boldsymbol{h}, \boldsymbol{\psi}_{j_l,k_l} \right\rangle$ is the l coefficient with descending rank, i.e. $h(l) \geqslant h(l+1)$. It is also indicated that in Ref. [12], when the ranked wavelet coefficients have rapid attenuation, $|h(l)| = O(l^{-s})$ if and only if $\varepsilon(M) = O(M^{-2s})$, the nonlinear approximation can produce fewer errors and certainly excels linear approximation. The detailed conclusion can consult the theory of Besov space [12]. In practice, the Mallat algorithm can be used in the multi-resolution representation for signal $\boldsymbol{h}$'s decomposition and reconstruction, which also ensures the wavelet adaptive model's superiority in processing efficiency.

Equally, we can also realize the selection of wavelet coefficients by employing the following threshold function

$$\text{threshold}(x)=\begin{cases}x, & |x|\geqslant T\\ 0, & |x|<T\end{cases} \tag{4}$$

to retain the primary characteristics of HRIR's signal $\boldsymbol{h}$, where T satisfies the following relationship: $h_{j,k}^{r}(M+1)<T\leqslant h_{j,k}^{r}(M)$, $h_{j,k}^{r}=\left\langle \boldsymbol{h},\boldsymbol{\psi}_{j,k}\right\rangle$, the sorted wavelet coefficients [12].

For convenient narration, we first suppose a dictionary $\boldsymbol{D}=\bigcup_{\lambda\in\Lambda}\boldsymbol{B}^{\lambda}$ composed of a union of orthonormal bases in a signal space of finite dimension N, where each orthonormal basis is a family of N vectors $\boldsymbol{B}^{\lambda}=\{\boldsymbol{g}_m^{\lambda}\}_{1\leqslant m\leqslant N}$. Wavelet packets and local cosine bases are just the examples of this kind of signal spaces.

We want to optimize the non-linear approximation of signal $\boldsymbol{h}\in L^2(R)$ by choosing a best basis in $\boldsymbol{D}$; and the resulting best non-linear approximation of $\boldsymbol{h}$ is

$$\boldsymbol{h}_M^{\lambda}=\sum_{m\in I_M^{\lambda}}\left\langle \boldsymbol{h},\boldsymbol{g}_m^{\lambda}\right\rangle \boldsymbol{g}_m^{\lambda}\sqrt{a^2+b^2} \tag{5}$$

where $\boldsymbol{I}_M^{\lambda}$ is the subscript set of the M biggest wavelet coefficients $\left|\left\langle \boldsymbol{h},\boldsymbol{g}_m^{\lambda}\right\rangle\right|$.

For HRIR signal of length N, from wavelet packet or local cosine bases' dictionary, choosing best bases is achieved by minimizing the following cost function

$$C(\boldsymbol{h},\boldsymbol{B}^{\lambda})=\sum_{m=0}^{N-1}\Phi\left(\frac{\left|\left\langle \boldsymbol{h},\boldsymbol{g}_m^{\lambda}\right\rangle\right|^2}{\|\boldsymbol{h}\|^2}\right) \tag{6}$$

where Φ can be chosen as the entropy function

$$\Phi(x)=-x\log_{\mathrm{e}}x \tag{7}$$

During practical realization, dynamic programming algorithm based on the dictionary structure of wavelet packet or local cosine bases can be used in searching for the optimal bases [12,13].

Now, we give some analyses and explanation of wavelet packet and local cosine bases' decomposition characteristics for HRIRs' multi-resolution modeling. First, a demonstrative division of the time-frequency plane for wavelet packet basis is given in Fig. 1. From the demonstration, we can see that wavelet packet basis divides the frequency axis in separate intervals of varying sizes and the best packet basis can thus be interpreted as the optimal frequency segmentation.

If a signal has similar time-frequency energy structure, located at different times and in the same frequency interval, the best wavelet packet bases can achieve good approximation for this kind of signals; but for the condition that the signal is translated, the wavelet packet coefficients are usually modified greatly and the resulting minimization of the cost function may give a very different basis, that is, the best wavelet packet bases are not translation invariants, and do not accord with the extraction of a signal's characteristics and corresponding pattern recognition for the signal.

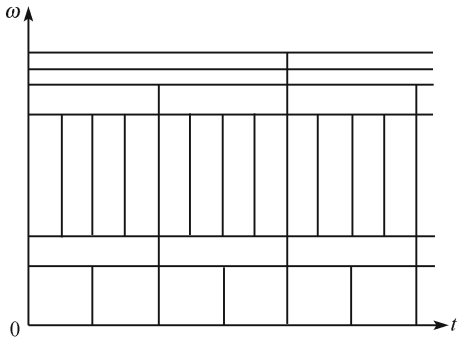


Fig. 1 Demonstrative division of the time-frequency plane for wavelet packet basis [12]

Second, orthonormal bases of space $L^2(R)$ can also be constructed dividing the time axis. For example, local cosine bases are just designed by multiplying smooth widows $g_p(t)$ with cosine functions $\cos(\omega t+\phi)$ at successive finite intervals $[t_p, t_{p+1}]$. Figure 2 gives a demonstrative division of the time-frequency plane for local cosine bases. From this figure, we can see that local cosine bases divide the time axis into intervals of different sizes, whose best resulting local bases adapt to the time segmentation of a signal's time-frequency structure. Therefore, a best local cosine basis is well adapted to approximate signals whose spectro-temporal characteristics fluctuate in time. However, it does not include very different structures of time and frequency spreading at any time. As a result of different characteristics for these two kinds of bases, they respectively tend to approximate different kinds of signals. In the following experimental results, these two kinds of bases' approximation performance for HRIRs' data is also given for an elaborate illustration of their applicability in adaptive non-linear modeling of spatial hearing research.

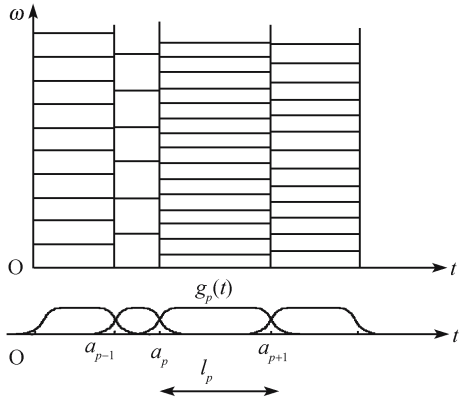


Fig. 2 Demonstrative division of the time-frequency plane for local cosine basis [12]

After that, we are going to give the bases pursuit algorithm, which is optimal for a signal's approximation, but not always orthogonal. In a signal space of dimensions N, let $\boldsymbol{D}=\{g_p\}_{0\leqslant p<P}$ be a redundant dictionary of P vectors, where

$P > N$, then an approximation h_M of $\boldsymbol{h}$ can be expressed as a linear combination of selected M vectors from the above dictionary $\boldsymbol{D}$. In general, the chosen vectors are not necessarily orthogonal, which provides much more freedom than the circumstance under orthogonal restriction.

As for a signal of length N, in the bases' dictionary $\boldsymbol{D}$, the optimal basis $\boldsymbol{B} = \{g_{p_m}\}_{0 \leqslant m < N}$ is selected accordingly to minimize the following cost function

$$C(\boldsymbol{h}, \boldsymbol{D}) = \sum_{m=0}^{N-1} \Phi\left(\frac{\left| \left\langle \boldsymbol{h}, \boldsymbol{g}_m^{\lambda} \right\rangle \right|^2}{\|\boldsymbol{h}\|^2} \right) \tag{8}$$

where the concave function Φ is set as

$$\Phi(u) = u^{1/2} \tag{9}$$

for bases pursuit algorithm. As a result, the cost function C can be rewritten in the next formula

$$C(\boldsymbol{h}, \boldsymbol{B}) = \frac{1}{\|\boldsymbol{h}\|} \sum_{m=0}^{N-1} \left| a(p_m) \right| \tag{10}$$

The optimal solutions related to minimization of cost function C can be acquired by linear programming [12,13].

Furthermore, it is well known that PCA is the optimal reduction representation of a group data under linear condition [2−4]. If $\boldsymbol{H}(n)$ is the Stochastic vector composed of all the HRIRs of KEMAR package, the HRIR $h_l(n)$ on a certain position can be expressed as [4]

$$h_l(n) = \boldsymbol{Q}\boldsymbol{w} + \boldsymbol{h}_{\text{av}} = \sum_{i=1}^{N} w_i(n, \theta_l, \varphi_l)\boldsymbol{q}_i + \boldsymbol{h}_{\text{av}} + \varepsilon_l \tag{11}$$

where $\boldsymbol{Q} = [\boldsymbol{q}_1, \boldsymbol{q}_2, ..., \boldsymbol{q}_N]$, satisfying the normalized condition: $\boldsymbol{Q}\boldsymbol{Q}^{\text{T}} = \boldsymbol{Q}^{\text{T}}\boldsymbol{Q} = \boldsymbol{I}$, is composed of the eigenvectors $\boldsymbol{q}_i$ of $\boldsymbol{H}(n)$'s autocovariance matrix $\boldsymbol{R}_h$; $\boldsymbol{R}_h = \frac{1}{P}\sum_{l=1}^{P}(\boldsymbol{h}_l - \boldsymbol{h}_{\text{av}})(\boldsymbol{h}_l - \boldsymbol{h}_{\text{av}})^{\text{T}}$, $\boldsymbol{h}_{\text{av}} = \frac{1}{P}\sum_{l=1}^{P}\boldsymbol{h}_l$ and $\boldsymbol{w}_i = \boldsymbol{Q}^{\text{T}}(\boldsymbol{h}_{\text{i}} - \boldsymbol{h}_{\text{av}})$. P is all the measured HRIRs' number of KEMAR package; and then ε_l is the approximation error of $h_l(n)$'s PCA model with N ranks.

Such results can be made clear in the ensuing geometrical explanation: the HRIR connected to each direction is a realization of Stochastic vector $\boldsymbol{H}$ of HRIRs, bases vectors $\boldsymbol{q}_i$ of Karhunen-Loève transformation present the principal axes of HRIRs data distribution, and the biggest eigenvalues correspond to the directions with dense distribution of the data. Consequently, HRIRs' reconstruction of projection onto these directions can result in minimum average errors. However, the presupposition is that, if the Stochastic vector $\boldsymbol{H}$ of HRIRs has a Gaussian distribution, its probability density is uniform along ellipsoids whose axes are proportional to the eigenvalue σ_i in the direction of vector $\boldsymbol{q}_i$. However, if $\boldsymbol{H}$ is not a Gaussian process, a non-linear approximation may be much more precise than the linear approximation, and the Karhunen-Loève transformation basis is no longer optimal. Here, we perform a Lilliefors test for goodness of fit to a normal distribution of HRIRs. As shown in Fig. 3, it is observable that most of the HRIRs' samples do not satisfy the hypothesis of normal distribution very well (The result of the hypothesis test is a Boolean value that is 0 when you do not reject the null hypothesis and 1 when you do reject that hypothesis.). Therefore, the PCA model based on Karhunen-Loève transformation cannot optimally approximate non-Gaussian HRIRs' data.

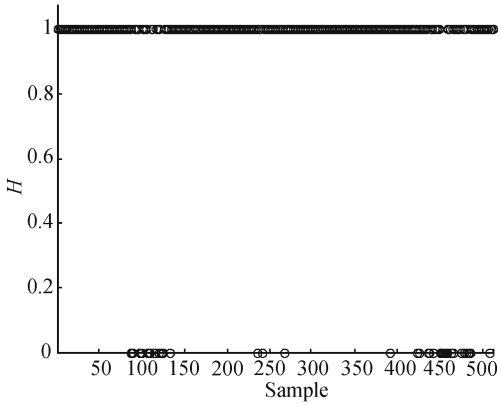


Fig. 3 Lilliefors hypothesis testing for goodness of fit to a normal distribution of HRIRs

In the following experiments, we apply wavelet multiresolution analysis algorithm to decompose HRIRs signal under different scales, and then set an appropriate threshold to approximate the HRIRs nonlinearly, whose results show better performance than those of HRIRs' PCA model.

3 Simulation results

Here, the data of KEMAR's HRTFs, offered by MIT Media Lab [14], are used in our simulation work. The measurement of the data was taken with a speaker on the discrete positions every 10° in elevation, and 5°–30° unequally in azimuth. Moreover, the measured data may be contaminated by deficient factors, and thus it is necessary to get rid of the contamination before further processing. For some other meticulous and important processing, the reader is referred to Ref. [14]. In view of all the 710 measurement positions of KEMAR, we first process the HRIRs data according to PCA model in the time domain, where 18 principal components, accounting for 98% of the variance in the original HRIRs, are selected as the basis vectors to approximate the measured HRIRs. The resulting PCA reduction model has a total of 21 996 values (including the principal components and weighted coefficients corresponding to all the measured spatial directions). Figure 4(a) is a demonstration of the HRIRs' PCA model.

In our work, we will emphasize the results of PCA, wavelet, wavelet packet, local cosine, and non-orthogonal bases resulting from bases pursuit algorithm to model HRTFs in adaptive non-linear approach.

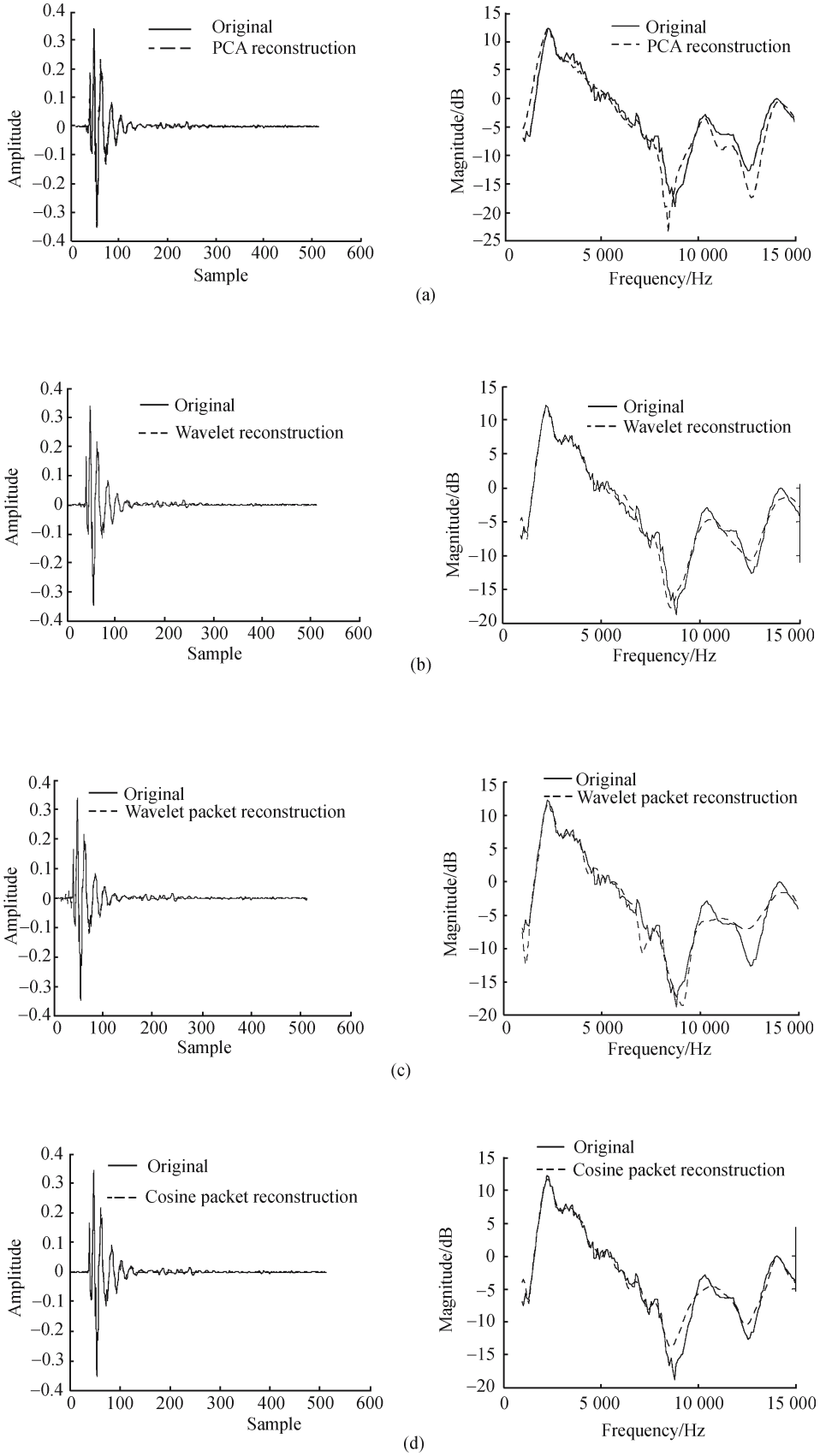
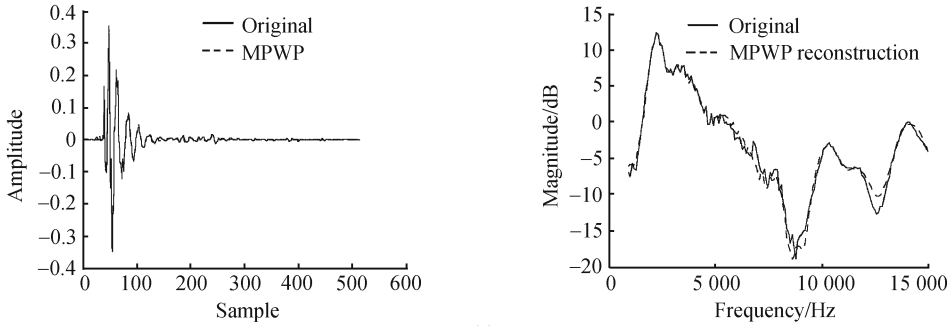


Fig. 4 A comparison demonstration of the HRIRs' reduction models based on different bases (a) PCA; (b) wavelet; (c) wavelet packet; (d) local cosine bases; (e) pursuit algorithm



(e)

Fig. 4 (*Continued*)

Figure 5 gives a typical comparison of performance among wavelet, wavelet packet, local cosine bases, and non-orthogonal bases grounded on pursuit algorithm, where we can see the pursuit algorithm's superiority over the other kinds of bases.

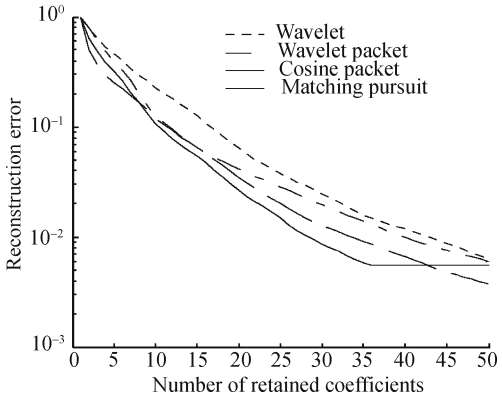


Fig. 5 An error comparison demonstration of the HRIRs' reduction models based on different bases

In addition, in Fig. 4, a demonstration of the HRIRs' reduction models based on different bases (Elevation = 0°, Azimuth = 0°) is given, where the thresholds for coefficients' selection are chosen as 3% of the HRIRs' Euclid norm. From the computational results for this position's HRIR, we can see that, the approximation error of the PCA model is 0.242 5; wavelet basis (dB10) is 0.118 5; wavelet packet is 0.156 5; local cosine basis is 0.111 1; while pursuit algorithm is 0.074 4, which excel the foregoing models' results.

As for all the measurement positions whose results are given in Figs. 6 and 7, under the circumstance with approximately equal amount of the remaining parameters of the PCA model, their corresponding average errors of models are respectively −13.1 dB, −18.1 dB, −18.2 dB, −18.9 dB, and −21 dB, shown in Figs. 6 and 7. (The PCA model preserves 21 996 parameters; the non-linear reduction model for wavelet basis holds 23 223 parameters' values, where we select Daubechies10 wavelet; the model of wavelet packet basis has 23 374 values; the local cosine model retains 21 952 parameters; and the pursuit algorithm reserves 26 913 coefficients.) We can see the superiority of the pursuit algorithm for HRIRs' modeling. Of course, the improved approximate error is achieved by a little redundancy of processing time. On a notebook PC with Intel Celeron M 1.5 GHz and 224 M memory, PCA needs 10 s, while the other models (wavelet, wavelet packet, local cosine, and pursuit algorithm) correspondingly need 4.3, 9.8, 28.8, and 91 s. However, fortunately, the processing time is all within the bounds of virtual auditory system's implementation [1,15].

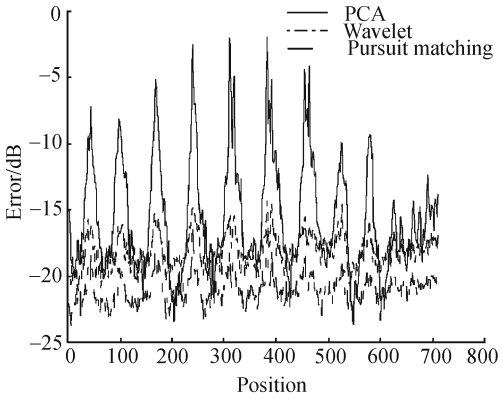


Fig. 6 Comparison 1 of reduction models on all the positions

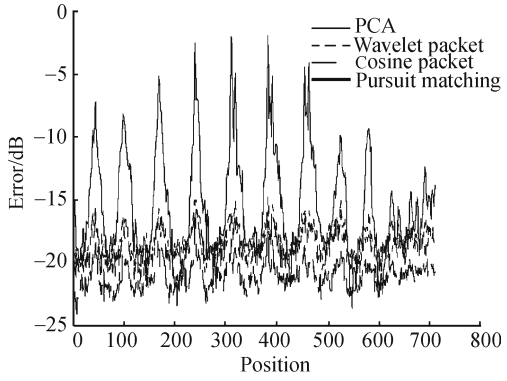


Fig. 7 Comparison 2 of reduction models on all the positions

Here, we use the following error formulation [4]

$$\text{Error} = \frac{\| \boldsymbol{h} - \hat{\boldsymbol{h}} \|^2}{\| \boldsymbol{h} \|^2} \tag{12}$$

where $\boldsymbol{h}$ is the measured HRIR, and $\hat{\boldsymbol{h}}$ is the reduction model's HRIR.

4 Conclusion

As seen from the results of different models, the adaptive model based on bases' pursuit algorithm for HRIRs' approximation achieves better effects than those models using PCA, wavelet, wavelet packet, and local cosine bases. This shows the adaptive method's validity based on non-orthogonal bases in HRIRs' modeling. As for future work, listening tests will be performed to evaluate subjective performance of these models.

Acknowledgements This work was supported by the National Basic Research of China (No. 2002CB312102).

References

1. Blauert J P. Spatial Hearing. Revised ed. Cambridge: MA, MIT, 1997
2. Kistler D J, Wightman F L. A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction. Journal of the Acoustical Society of America, 1992, 91(3): 1 637−1 647
3. Chen J, Van Veen B D, Hecox K E. A spatial feature extraction and regularization model for the head-related transfer function. Journal of the Acoustical Society of America, 1995, 97(1): 439−452
4. Wu Z Y, Chan F H Y, Lam F K, et al. A time domain binaural model based on spatial feature extraction for the head-related transfer function. Journal of the Acoustical Society of America, 1997, 102(4): 2 211−2 218
5. Zhang Qinghua. Using wavelet network in nonparametric estimation. IEEE Transactions on Neural Network, 1997, 8(2): 227−236
6. Torres J, Petraglia M, Tenenbaum R. Low-order modeling of head-related transfer functions using wavelet transforms. In: Proceedings of ISCAS, Vancouver, Canada. IEEE Press, 2004, 513−516
7. Zhang Jie, Wu Zhenyang, Ma Hao. A smoothing method of head-related transfer functions based on reconstruction from wavelet transform modulus maxima. Journal of Electronics and Information Technology, 2007, 29(2): 473–477 (in Chinese)
8. Wu Zhenyang, Wang Weipin. Denoising of HRTF (Head Related Transfer Function) based on singularity of wavelet transformation. Acta Biophysica Sinica, 1997, 13(3): 473−478 (in Chinese)
9. Batteau D W. The role of the pinna in human localization. In: Proceedings of Royal Society London, 1967, 168b: 158−180
10. Hiranaka Y, Yamasaki H. Envelope representations of pinna impulse responses relating to three-dimensional localization of sound sources. Journal of the Acoustical Society of America, 1993, 73(1): 291−296
11. Hebrank J, Wright D. Spectral cues used in the localization of sound sources on the median plane. Journal of the Acoustical Society of America, 1974, 56(6): 1 829−1 834
12. Mallat S. A Wavelet Tour of Signal Processing. Boston: Academic Press, 1997
13. Donoho D, Duncan M, Huo Xiaoming, et al. The Wave-lab Package. Stanford University, 1999
14. Gardner B, Martin K. HRTF measurements of a KEMAR dummy-head microphone. Technical Report. MIT Media Lab Perceptual Computing, Cambridge: MA,1994, 5
15. Zotkin D N, Duraiswami R, Davis L S. Creation of virtual auditory spaces. In: Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing. IEEE Press, 2002, 2 113−2 116