Available online at www.sciencedirect.com
ScienceDirect
Journal of Sound and Vibration 321 (2009) 1090–1108
JOURNAL OF SOUND AND VIBRATION
www.elsevier.com/locate/jsvi

Multirate synthesis of reverberators using subband filtering

Mingsian R. Bai*, Kwuen-Yieng Ou, Pingshun Zeung

Department of Mechanical Engineering, National Chiao-Tung University, 1001 Ta-Hsueh Road, Hsin-Chu 300, Taiwan

Received 15 January 2008; received in revised form 5 June 2008; accepted 1 October 2008
Handling Editor: L.G. Tham
Available online 18 November 2008

Abstract

As an important element in three-dimensional audio reproduction, reverberators serve to enhance the sense of spaciousness by synthesizing the sound reflections and reverberations pertaining to a particular listening environment. However, direct implementation of room responses generally proves impractical due to the extreme complexity of system dynamics. To overcome the difficulties encountered in long convolutions, an efficient multirate signal processing technique is proposed in this paper on the basis of subband filtering scheme. In this method, the frequency dependent property of reverberation time is incorporated into the design procedure. Further simplification follows from the synthesized infinite impulse response (IIR) implementation for each subband. Specifically, the regular IIR filter, the cascaded comb filters and nested allpass filters, and the cascaded comb filter and an FIR filter are developed to implement the subband filters. Genetic algorithms (GA) are employed to optimize the parameters of reverberators. Subjective listening tests are carried out to assess the performance of the proposed techniques. The results demonstrate that the implemented multirate reverberators are capable of delivering natural sensation of room reverberations, as compared to direct measurement.

1. Introduction

Reverberation plays a vital role in 3D audio reproduction in that it creates a realistic sensation of diffusiveness of sound field and spaciousness of the acoustic environment. For instance, if appropriate reverberation is not incorporated when a headset is used as a means for reproduction, the virtual sound would be perceived as if it emanates from inside the head. However, direct implementation of a reverberator is usually not a trivial task because of its long impulse response of a real room recorded at the sampling rate, say, 44.1 kHz. This approach is too computationally intensive to warrant a real-time implementation. One solution to this problem is the use of high speed signal processing hardware. For example, Lake technology [1] developed fast convolution engines using multiprocessor architecture. On the other hand, software approaches offer less expensive alternatives. The commonly used method for this long convolution process is to utilize artificial reverberators comprised of comb and allpass filters [2–4]. However, this method suffers from a disadvantage in that it takes much experience to adjust the parameters of the comb and allpass network before a natural sounding reverberator (without a metallic artifact) can be reached. An alternative approach is the

*Corresponding author.
E-mail address: msbai@mail.nctu.edu.tw (M.R. Bai).

0022-460X/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jsv.2008.10.003

fast convolution technique such as block processing methods and their variants using fast Fourier transform (FFT). Gardner devised a hybrid convolution algorithm that combines direct convolutions and FFT, without introducing input–output delay [5]. Along the same line of software approach, this paper proposes an efficient technique to deal with long convolutions involved in reverberators. This method differs from the FFT-based approaches in that it seeks to realize the required filtering in the subband domain.

Subband processing is widespread in the area of multimedia data compression and perceptual audio coding [6]. In particular, the subband coding in MPEG-1 layer I and II is implemented with a pseudo-QMF (quadrature mirror filter) structure. This paper is motivated by the similar notion. We seek to enhance the filtering efficiency in light of the parallelism of filter bank structure. Computationally intensive filtering is carried out in subbands, with the frequency dependent reverberation time taken into account.

A review of the previous work on subband filtering is given as follows. A filter bank convolver with comparable computational complexity was proposed by Vaidyanathan [7]. Vetterli attempted to simplify subband filtering by converting a length-KN convolution into M length-K convolutions down sampled by N [8]. However, this simplification is still insufficient for the present reverberation problem, where long filters taps is needed in each subband. This difficulty was discussed by using a matrix representation of subband filtering [9]. Although the aliasing problem can be eliminated by using a general filter bank, the computation remains intensive due to the cross terms of the subband matrix. It was also pointed out in Ref. [10] that aliasing problems will arise in the transition band and result in unpleasant ringing sound if the cross terms are ignored. To alleviate the problems of previous method, a subband filtering technique based on cosine modulated pseudo QMF with high stopband attenuation [11,12] is suggested in this paper. Instead of using the FIR structure in Ref. [9], IIR structures are employed to implement the subband filters. Specifically, the regular IIR filter, the cascaded comb filters and nested allpass filters, and the cascaded comb filter and an FIR filter. Properties of reverberation are exploited to simplify the filter design. To evaluate the performance of the proposed reverberators, subjective listening tests are conducted in terms of spaciousness, clarity, naturalness, and richness. The results reveal that the proposed subband techniques are capable of rendering natural reverberation with little penalty on processing efficiency.

2. Filtering by cosine modulated filter bank

There are three advantages in implementing reverberators under the subband structure. The first advantage is that the parallelism inherent in the filter bank is exploited to divide a long filter into short filters to represent different bands of the frequency response. Only thus could the dynamics in each subband be approximated by a low-order filter. This approach is apparently simpler than directly approximating the total band. The second advantage is that frequency dependent acoustic properties pertaining to reverberation is accounted for. The third advantage is that the proposed method can be embedded to existing subband-based systems which are widespread in the area of multimedia data compression and perceptual audio coding. The idea is depicted in Fig. 1, whereby we seek a subband implementation in Fig. 1(b) that produces a filtered output approximately

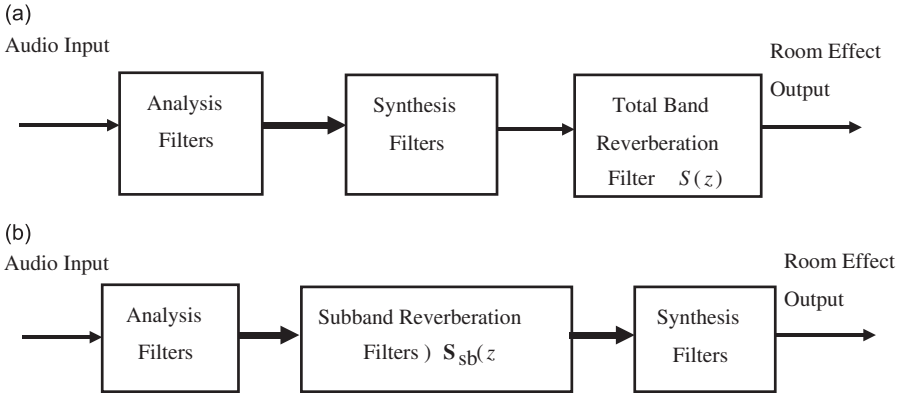

Fig. 1. Schematic diagrams of reverberation filters. (a) Total band filtering and (b) subband filtering.

identical to that from the system in Fig. 1(a). It has been demonstrated in Ref. [9] that an M-channel subband filter bank can be expressed as

$$S_{\mathrm{sb},nk}(z) = \left[z^{L-1}H_n(z)S(z)F_k(z)\right]_{\downarrow M}, \quad 0 \leqslant n, \;\; k \leqslant M-1 \tag{1}$$

where $H_n(z)$ is the nth *analysis* filter, $F_k(z)$ is the kth *synthesis* filter, $S(z)$ is the total band reverberation filter, L is the length of the analysis and synthesis filters, and $\downarrow M$ represents *decimation* by M. The concept of the general subband filtering is shown in Fig. 2(a). From Eq. (1), it is understood that the filtered output in the subband n is the sum of the convolutions of the kth subband decimated signal, filtered with the decimated product $H_n(z)S(z)F_k(z)$. The decimated subband signals contain the aliased components that can be canceled after synthesis, using the cross-term filtering ($K \neq n$). However, ideal cancellation can only be achieved at the cost of intensive computations. Conversely, if the cross terms are neglected, aliasing in the transition band will result in unpleasant ringing problems [10]. To avoid this dilemma, we choose to design the subband filter bank with very high stopband attenuation and cross terms neglected, as shown in Fig. 2(b). In Fig. 3, the frequency domain illustrating the idea of subband filtering is shown with an example of a three-channel filter bank. In this figure, $\omega = \Omega T$, $\omega' = \Omega T' = M\Omega T$; T and T' are the sampling periods before and after down sampling; Ω is the analog frequency; ω and ω' are the digital frequencies before and after down sampling. The parallelism stems from the frequency division of processing, as is evident in the evolution of signals at each stage.

The design strategy of subband filter bank employed in this paper is the cosine modulated pseudo QMF [11,12]. In this method, an FIR filter must be selected as the prototype. Using this prototype, an M-channel maximally decimated filter bank (number of subbands = up/down sampling factor) is generated with the aid

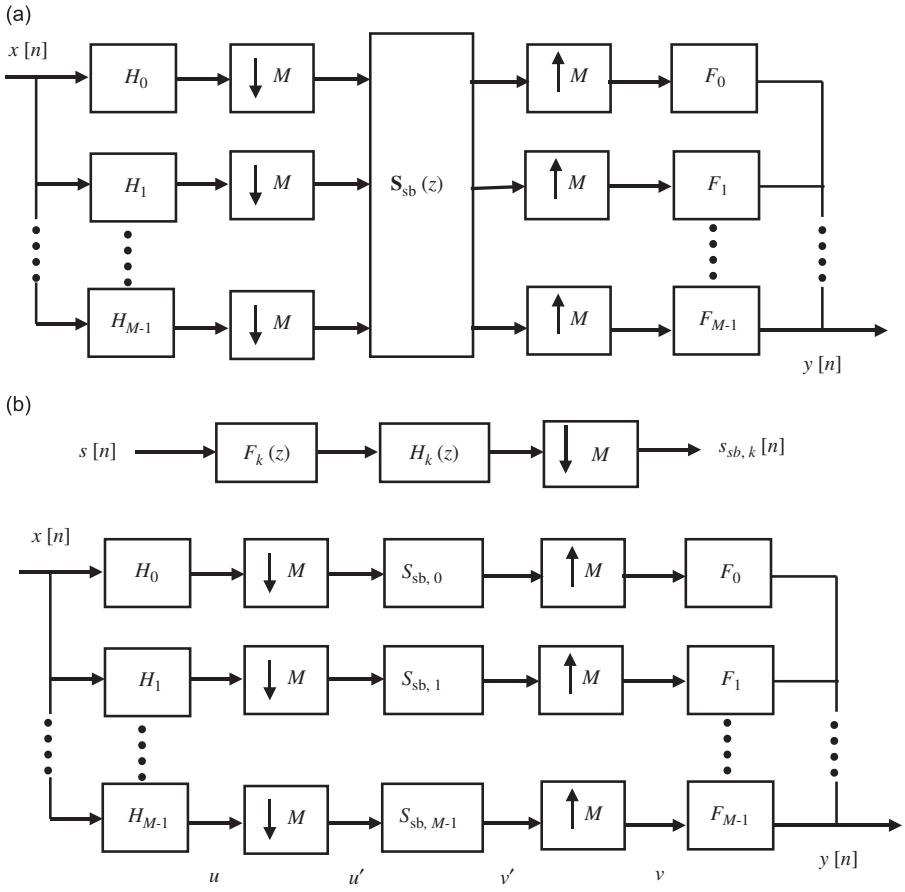

Fig. 2. Subband filtering using a M-channel maximally decimated filter bank. H's and F's are analysis filters and synthesis filters, respectively. (a) Representation of subband filtering by a filter matrix $\mathbf{S}_{\mathrm{sb}}(z)$ and (b) subband filtering with cross-terms neglected.

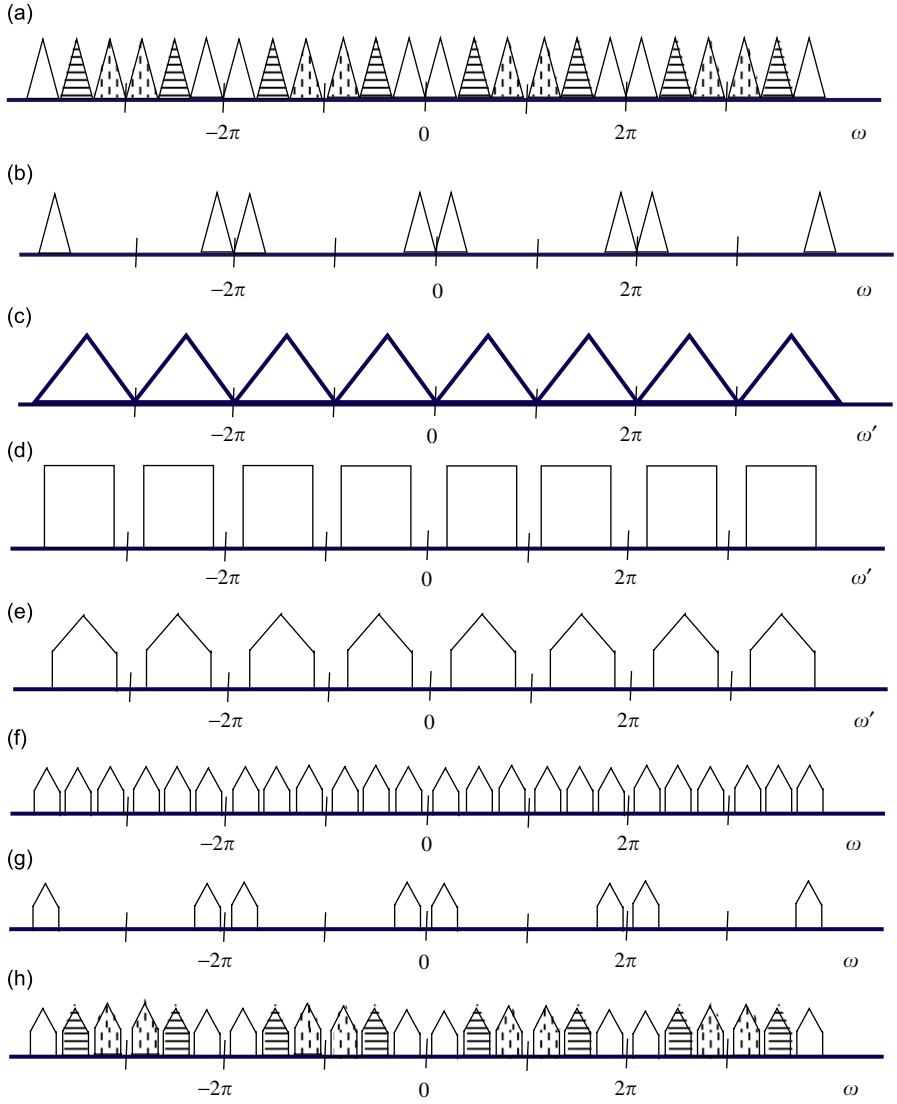

Fig. 3. Frequency-domain illustration of subband filtering for a three-channel filter bank. (a) The spectrum of the input signal. The signal of each subband is indicated with different patterns, (b) the signal filtered by the analysis filter of the first subband, (c) the filtered first subband signal after down sampling, (d) the first subband reverberation filter, (e) the first subband signal filtered by the subband reverberation filter, (f) the first subband signal after up sampling, (g) the first subband signal filtered by the synthesis filter and (h) the synthesized signal at the output end. The signal of each subband is again indicated with different patterns.

of cosine modulation. The maximum attenuation that can be attained by a perfectly reconstructing (PR) cosine modulated filter bank is about 40 dB. Nevertheless, this PR filter bank would still present an undesirable ringing problem. To alleviate this problem, the PR condition is relaxed in the FIR filter design to gain more stopband attenuation. From our experience, as much as 60 dB attenuation is required for acceptable reproduction of reverberations.

Based on the method in Ref. [13], the following analysis and synthesis filter banks represented by $h_k(z)$ and $f_k(z)$, respectively, are employed to minimize phase distortion and aliasing.

$$h_k(n) = 2p_0(n)\cos\left[\frac{\pi}{M}(k+0.5)(n-\frac{N}{2})+\theta_k\right] \tag{2}$$

$$f_k(n) = h_k(N-n) \tag{3}$$

where $\theta_k = (-1)^k \pi/4, 0 \leqslant k \leqslant M-1$ and $p_0(n), n = 1, 2, \ldots, N$ are the coefficients of the prototype FIR filter. The remaining problem is how to minimize the amplitude distortion. The distortion function $T(z)$ for the filter bank is given by [13]

$$T(z) = \frac{1}{M} \sum_{k=0}^{M-1} F_k(z) H_k(z) \tag{4}$$

Z transform of Eq. (3) leads to $F_k(z) = z^{-N} \tilde{H}(z)$, where $\tilde{H}(z)$ is the paraconjugate of $H(z)$. The distortion function can thus be written in frequency domain as

$$T(e^{j\omega}) = \frac{1}{M} e^{-j\omega N} \sum_{k=0}^{M-1} \left| H_k(e^{j\omega}) \right|^2 \tag{5}$$

A filter $G(z)$ is called a Nyquist (M) filter if the following condition is met:

$$g(Mn) = \begin{cases} c, & n = 0 \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

where $g(n)$ is the impulse response of $G(z)$ and c is a constant. In frequency domain

$$\sum_{k=0}^{M-1} H(e^{j(\omega - 2\pi k/N)}) = Mc \tag{7}$$

Eqs. (5) and (7) indicate that, if $|H_k(e^{j\omega})|^2$ is a Nyquist (M) filter, or equivalently, $|P_0(e^{j\omega})|^2$ is a Nyquist $(2M)$ filter, the magnitude of $T(z)$ will be flat.

In this QMF design, the Kaiser window is used as the FIR prototype [11,12,14]. Given the specifications of transition bandwidth Δf and stopband attenuation A_s, the parameter β and the filter order N can be determined according to

$$\beta = \begin{cases} 0.1102(A_s - 8.7) & \text{if } A_s > 50 \\ 0.5842(A_s - 21)^{0.4} + 0.07886(A_s - 21) & \text{if } 21 < A_s < 50 \\ 0 & \text{if } A_s < 21 \end{cases} \tag{8}$$

and

$$N \approx \frac{A_s - 7.95}{14.36\,\Delta f} \tag{9}$$

For example, the frequency responses of the FIR prototype and the analysis filters for an 8-channel filter bank design obtained using cosine modulation are shown in Figs. 4(a) and (b), respectively.

An optimization procedure is employed here to make $P_0(z)\tilde{P}_0(z)$ an approximate Nyquist $(2M)$ filter, as posed by the following min–max problem [14]:

$$\min_{\omega_c} \max_{n \neq 0} \left| p_0(n) * p_0(-n) \right|_{\downarrow 2M} \tag{10}$$

where $*$ denotes the convolution operator. Because this is a convex problem, optimal cutoff frequency can always be found [14]. After obtaining the optimal prototype filter, the remainder of the analysis and synthesis filters are generated, according to Eqs. (2) and (3), respectively. The filter bank can be easily implemented with techniques such as polyphase structure or discrete cosine transform (DCT) [13].

3. Structures of subband reverberation filter

Although subband filtering enhances computation efficiency, the loading remains high for FIR implementation of the desired filters. Hence, three IIR structures are proposed to approximate the subband filters such that the computational loading can be further reduced.

The first structure is the regular IIR filter. Partitioning the desired room response into subbands enables the identification of the subband filter with smaller number of modes than the total band response. Following the

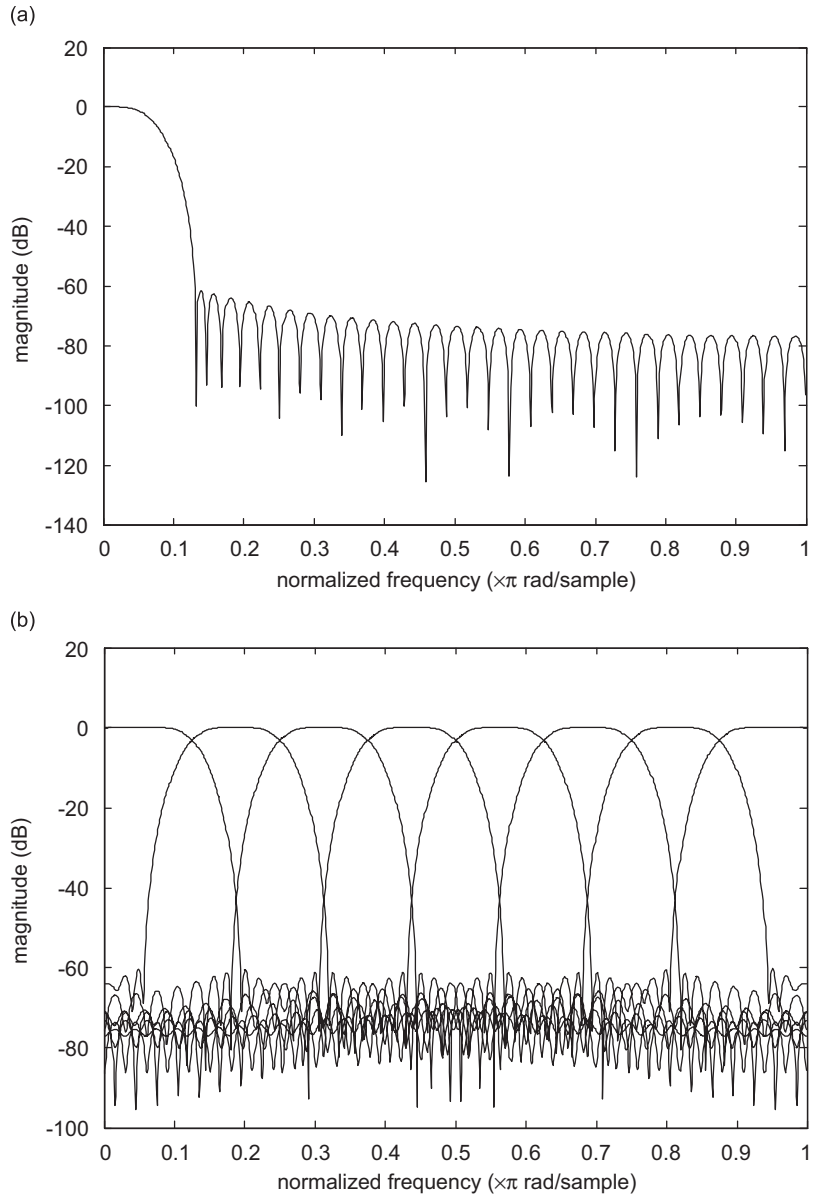

Fig. 4. Magnitude of the frequency response of an 8-channel filter bank. (a) The FIR prototype resulting from the Kaiser window design with transition bandwidth $\Delta f = 0.0625$ and stopband attenuation $A_s = 60\,\mathrm{dB}$ and (b) the combined response of the analysis filter bank obtained using cosine modulation.

process of Fig. 2(b), the room impulse response is filtered by the corresponding analysis filter and synthesis filter, and decimated to yield the subband impulse response. By transforming this response to the frequency domain, the subband filter model can be identified by using the Matlab instruction *invfreqz* [15]. This command finds a discrete-time transfer function to match a given frequency response based on an equation error criterion. Numerator and denominator coefficients **a** and **b** are determined by the following equation:

$$\min_{\mathbf{b},\mathbf{a}} \sum_{k=1}^{n} \omega t(k) \left| h(k) - \frac{B(\omega(k))}{A(\omega(k))} \right|^2 \tag{11}$$

where ωt is the vector of the weighting factors, $A(\omega)$ and $B(\omega)$ are the Fourier transforms of the polynomials **a** and **b**, respectively, ω is the frequency, and n is the number of frequency points. The damped Gauss–Newton method is used for iterative search the optimal solution [16,17].

For an example of a rectangular room, the analytical solution of the room response can be expressed as the following series expansion [18,19]:

$$p(x) \simeq \sum_{n=0}^{\infty} \frac{c^2 \psi_n(x)}{\omega_n^2 - \omega^2 + 2\mathrm{j}\zeta_n \omega_n \omega} \int_V \psi(x_0) Q(x_0)\, \mathrm{d}V(x_0) \tag{12}$$

where ψ_n is the eigenfunction, Q is the source,

$$\zeta_n = \frac{cD_{nn}}{2\omega_n} \tag{13}$$

$$D_{nn} = \frac{1}{V} \int_s \psi_n^2 \beta\, \mathrm{d}S \approx \frac{1}{V} \int_s \psi_n^2 \chi\, \mathrm{d}S \tag{14}$$

where S and V are the surface area and the volume of the rectangular room, respectively.

In addition,

$$\beta(x) \triangleq \frac{\rho_0 c}{z} = \chi + \mathrm{j}\sigma \tag{15}$$

where z is the *specific acoustic impedance*, ρ_0 is the air density, c is the sound speed, χ is the *normalized specific acoustic conductance* and σ is the *normalized specific acoustic susceptance*. The parameter σ slightly shifts the modal frequencies (very small in general), whereas χ damps the resonant modes. If the admittance is small, only D_{nn} has significant effects on the overall damping of the field.

Reverberation time T_{60}, an important property of the room response, is defined as the time required for the sound pressure level to attenuate by 60 dB. T_{60} is in general frequency dependent on the absorptivity of the boundary and medium. T_{60} can be estimated on the basis of the energy decay curve (EDC) [20].

$$\mathrm{EDC}(t) = \frac{\int_t^{\infty} h^2(\tau)\, \mathrm{d}\tau}{\int_0^{\infty} h^2(\tau)\, \mathrm{d}\tau} \tag{16}$$

where $h(\tau)$ is the subband impulse response.

The second structure of subband filter contains parallel comb filters cascaded with nested allpass filters, as shown in Fig. 5. Unlike Schoroeder's reverberator [21] that has four parallel comb filters, cascaded with two serial allpass filters, the nested allpass filter suggested by Gardner [22] is adopted here. One feature of the nested allpass filters is that no periodic response will arise in the time domain and the echo density increases with time, which is similar to the response of a realistic room. As compared to a regular IIR filter, it is straightforward to determine parameters such as dense echo density, modal density, and reverberation time from the second structure. The stability condition is also easily satisfied with the filter gains retrained to be less than unity.

The transfer function of the comb filter used in the second structure takes the form

$$C(z) = z^{-m_i}/(1 - g_i z^{-m_i}) \tag{17}$$

where g_i and m_i are the gain and delay of the ith comb filter. The only parameter to be determined for the parallel comb filters is the gain g_1 of the first comb filter. The remaining gain and delay parameters of comb filters are calculated according to

$$\frac{20 \log 10(g_i)}{m_i T} = \frac{-60}{T_{60}}, \quad i = 1, 2, \ldots, p \tag{18}$$

where $m_1{:}m_p = 1{:}1.9$ and p is the number of comb filters and T is the sampling period. Thus, the gain and delay parameters of the nested allpass filters together with g_1 of the first comb filter will be optimized, using the GA, as will be detailed in the next section.

The third structure of subband filter is composed of an FIR filter cascaded with a comb filter. In what follows, the design procedure is illustrated for one of the subband filters. The FIR filter is intended for

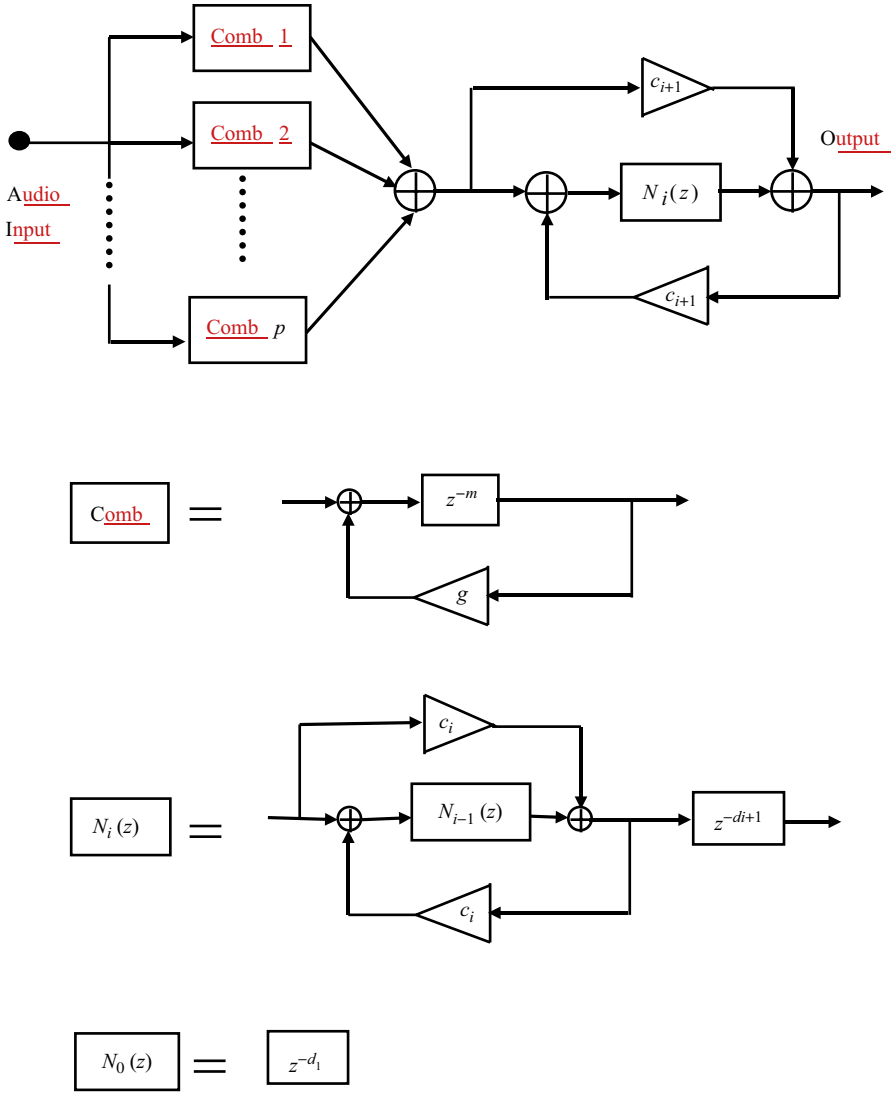

Fig. 5. Block diagrams of the second structure composed of parallel comb filters and nested allpass filters.

emulating the early reflections and is written as

$$H_E(z) = a_0 + a_1 z^{-1} + \cdots + a_M z^{-M} \tag{19}$$

where a_i, $i = 1, 2, \ldots, M$ are the filter coefficients and M is the filter order. The comb filter serves to approximate the late reverberation that has an exponentially decay envelope whose decay rate is dictated by the reverberation time. The comb filter has the transfer function, $H_C(z) = 1/(1-gz^{-N})$, where g and N are the gain and delay parameters, respectively. Next, the cascaded transfer function $H(z) = H_E(z)H_C(z)$ is matched to the subband filter.

$$H(z) = \frac{a_0 + a_1 z^{-1} + \cdots + a_M z^{-M}}{1 - gz^{-N}} = \sum_{n=0}^{\infty} h_s(n) z^{-n} \tag{20}$$

Multiplying both sides of Eq. (20) by the denominator $(1-gz^{-N})$, truncating the subband impulse response $h_s(n)$ to K samples, and comparing the coefficients of both sides lead to the following set of equations [23]:

$$\begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_M \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} h_0 & 0 & 0 & \cdots & 0 \\ h_1 & h_0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ h_M & h_{M-1} & h_{M-2} & \cdots & h_{M-N} \\ h_{M+1} & h_M & h_{M-1} & \cdots & h_{M-N+1} \\ \vdots & \vdots & \vdots & & \vdots \\ h_K & h_{K-1} & h_{K-2} & \cdots & h_{K-N} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ -g \end{bmatrix} \tag{21}$$

where $h_s(n)$ is written as h_n for simplicity. The value of N is generally chosen to be the length of the repetitive pattern occurring in the subband impulse response. The gain g is determined from Eq. (19). If $M = N$, the coefficients $a_k = h_k$ for $k = 0, 1, \ldots, M-1$ and $a_M = h_M - gh_0$.

4. Optimization by the genetic algorithm

In this section, genetic algorithm (GA) is employed to find the optimal parameters of the preceding second subband filter structure, without resorting to much experience. GA is a multi-starting-point algorithm and is less susceptible to local optima problem, making it well suited for reverberator optimization.

In the GA procedure, the design variables are concatenated into the so-called *chromosome*

$$\theta = [g_1 \;\; c_1 \;\; c_2 \;\; \cdots \;\; c_q \;\; d_1 \;\; d_2 \;\; \cdots d_q]^{\mathrm{T}}, \quad \theta \in \Theta \tag{22}$$

where g_1 is the gain of the first comb filter, c_i and d_i $(i = 1,2,\ldots,q)$ are the gain and delay of the nested allpass filter, respectively, q is the number of loops in the nested allpass filters, and Θ is the parameter space. There are $1+2q$ genes per parameter vector.

The chromosomes encoded as binary numbers are generated randomly to form the population. The objective is to find the chromosome with which the EDC of the synthesized reverberator best approximates the desired one. The cost function for match, $\eta(\theta)$, is defined as

$$\eta(\theta) = \max_t(|E(t) - \hat{E}(t)|) \tag{23}$$

where $E(t)$ is the desired subband EDC and $\hat{E}(t)$ is the synthesized EDC. This cost function is then related to a fitness function through the following linear mapping [24,25], also depicted in Fig. 6:

$$F(\theta) = a\eta(\theta) + b \tag{24}$$

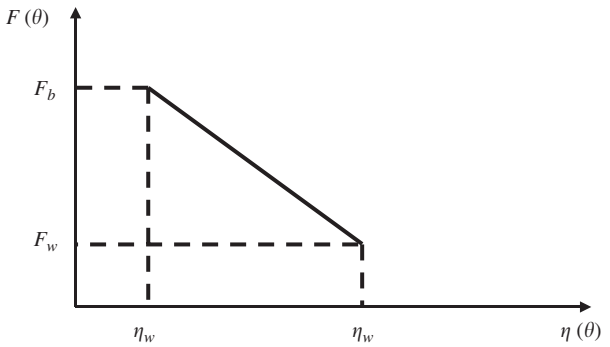

Fig. 6. The relationship between the cost function and the fitness function in GA.

where

$$a = \frac{F_b - F_w}{\eta_b - \eta_w} \tag{25}$$

$$b = F_b - a\eta_b \tag{26}$$

η_b and η_w are the best and worst values of the cost function, respectively.

The first GA procedure, *reproduction*, decides which chromosome in current generation will survive according to the probability

$$p_k = \frac{F_k(\theta)}{\sum_{k=1}^{N_p} F_k(\theta)} \tag{27}$$

where N_p is the population size. Based on p_k, the cumulative distribution function (CDF) is calculated. With reference to the CDF, a random number between 0 and 1 is then generated to decide which chromosome should enter the mating pool. This operation is repeated N_p times to obtain a new population.

The second GA operation, *crossover*, exchanges information between the two randomly chosen chromosomes in the mating pool. Another random number between 0 and 1 is generated. If this number is less than the prespecified crossover probability P_c, the following crossover operation will take place:

parents 1: xxxxxxxx
parents 2: yyyyyyyy
↑ splice point

becomes

child 1: xxxyyyyy
child 2: yyyxxxxx

after crossover.

The third operation, *mutation*, is necessary for preventing the population from premature convergence. This operation is initialized by specifying a mutation probability P_m. Then, a "masking" string of the length of a chromosome is constructed for each chromosome by randomly generating numbers between 0 and 1. If the number is less than the mutation probability, the corresponding digit in the masking string is set to 1, otherwise, 0. Mutation is carried out by the "exclusive-OR" operation between the chromosome and the associated masking string. For example,

maskingpattern : 01000100
chromosome : 10110110 before mutation
chromosome : 11110010 after mutation

The above GA steps should be repeated before the optimal solution can be reached. The overall flow chart of GA optimization is summarized in Fig. 7. According to the elitism policy, the best two chromosomes are preserved during the reproduction. They are passed directly to next generation without crossover and mutation.

5. Experimental investigations

A binaural frequency and impulse responses of a concert hall measured by the Physikalisch-Technische Bundesanstalt [26] is shown in Figs. 8(a) and (b), respectively. The sampling rate is chosen to be 44.1 kHz. The room response has high gain at low frequencies and significant roll-off at high frequencies due to air absorption. We began with the implementation of the first subband structure mentioned in Section 3 to match this measured room response. In this structure, 32 channels were included in the filter bank. All subband filters were identified from frequency domain with the Matlab command *invfreqz*. IIR filters ranging from 60 to 72 orders were used to fit the frequency responses of the first 8 subbands, while 30 orders were used for the rest of the subbands because the low frequency components seem to dominate the room response. For example, the

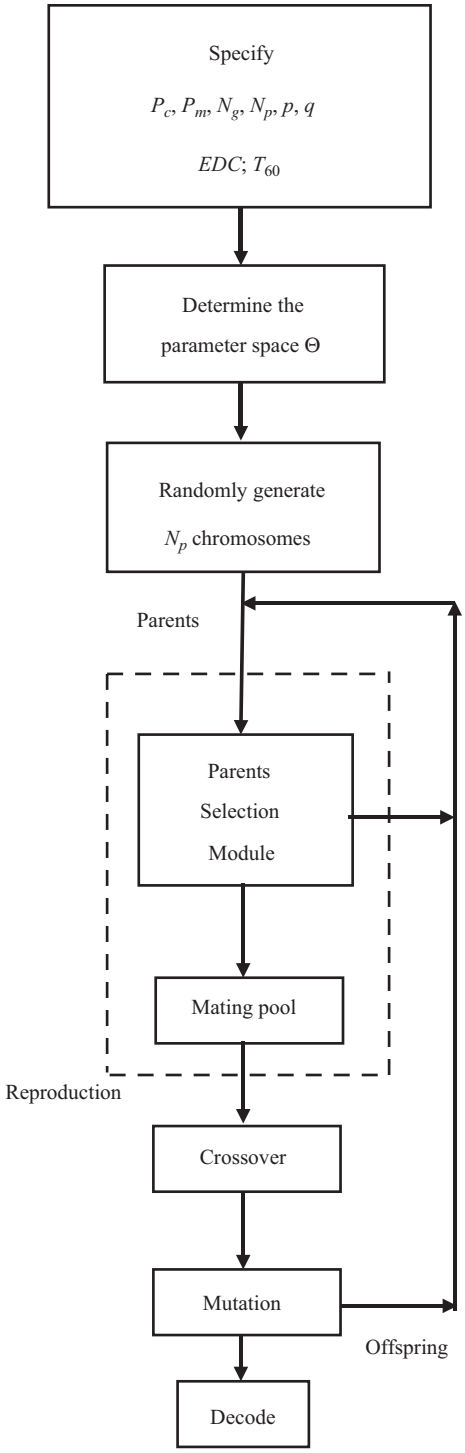

Fig. 7. Flow chart of the GA optimization procedure.

identified frequency responses of the first and the thirteenth subband filters are shown in Figs. 9 (a) and (b), respectively. The resonant peaks are still quite many even if 32 subbands have been used, making it difficult to model the complete dynamics in the system. The total order of IIR filters is 1234, in comparison with the original measured 50 000-tapped FIR filter.

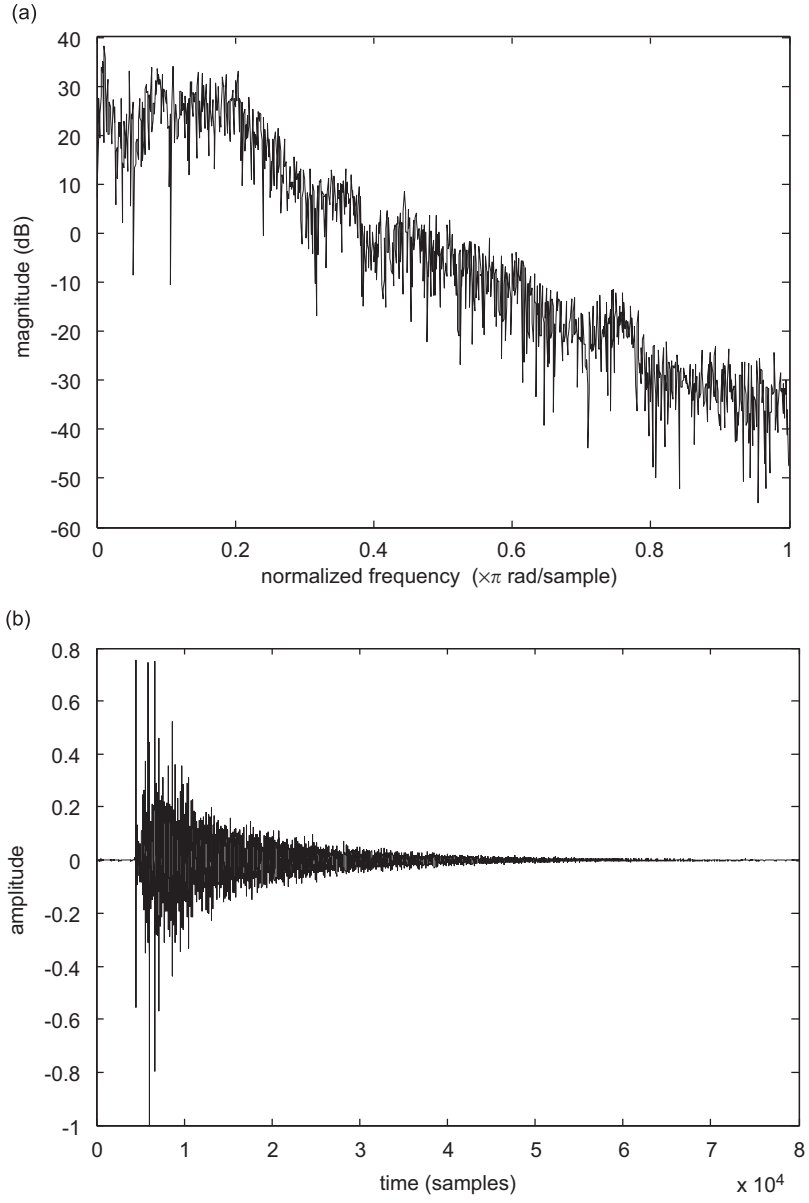

Fig. 8. Room response of a concert hall. (a) The frequency response and (b) the impulse response.

Based on the EDC of the subband impulse response in Fig. 10, the reverberation time $T_{0.30}$ is first estimated and then multiplied by two to get $T_{0.60}$. From the estimated reverberation time, Eqs. (16) and (18) are used to determine the gain and delay parameters of the comb filters. Next, the preceding GA optimization procedure was conducted to find the best parameters for the second structure to fit the EDC of the measured room response. The parameters used in GA were set to be: population size = 100, number of generations = 100, crossover probability = 0.8, mutation probability = 0.01, and four comb filters cascaded with a six-looped nested allpass filter. The 16-channel filter bank proved appropriate for this problem. Fig. 11 compares the optimized EDC, the impulse response, and the frequency response with the desired response for the first subband. The optimal parameters found by the GA procedure of this case are presented in Table 1. It is observed that the optimized system seems to have captured the general time-domain features. Even though the

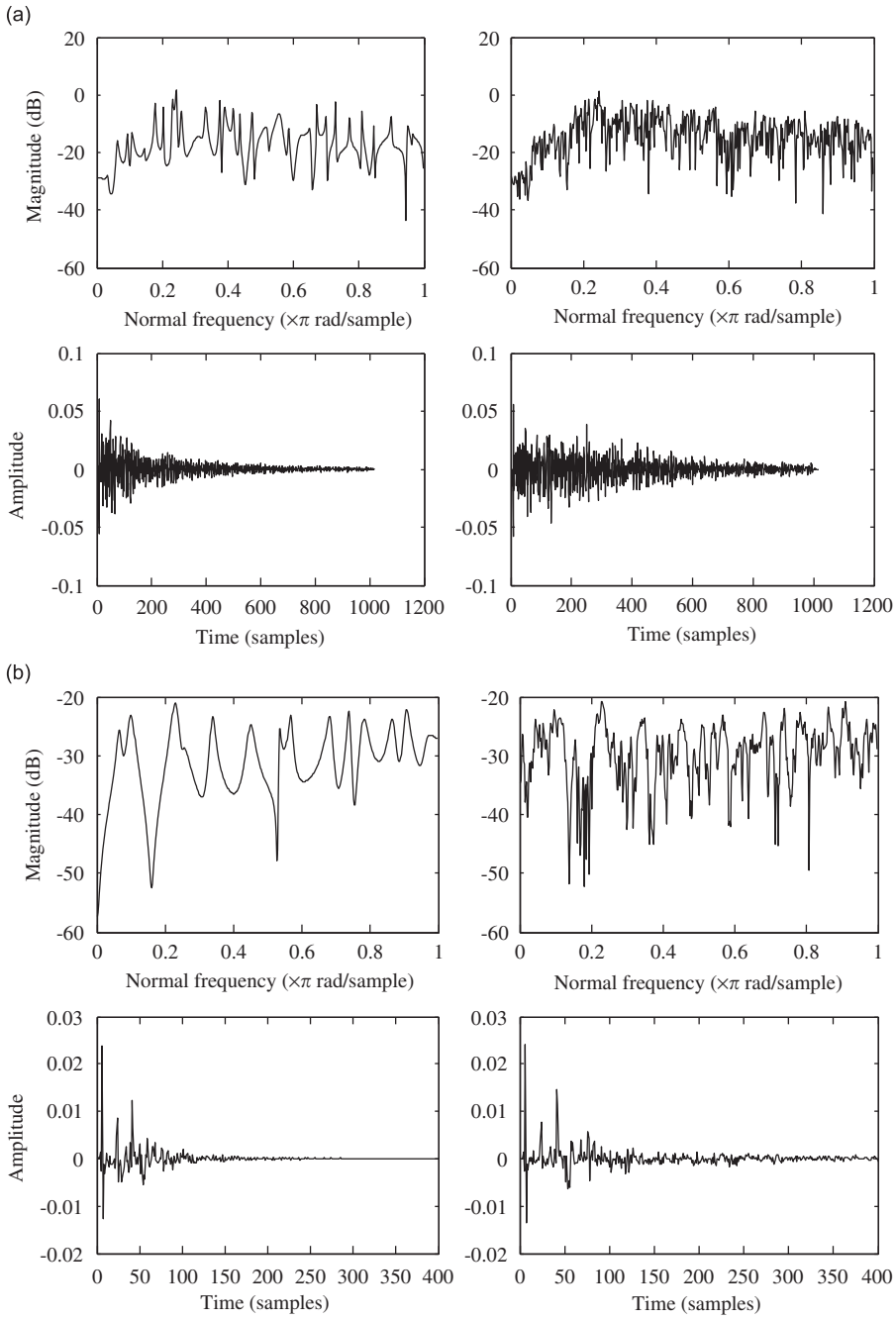

Fig. 9. Comparison between the synthesized response and the measured response for the first subband structure. (a) Frequency and time responses for the first subband and (b) frequency and time responses for the thirteenth subband. (Left side: synthesized response; right side: measured response.)

frequency responses did not match well, the comb filters did not present notable coloration problem, as commonly encountered in comb and allpass reverberators. The computation efficiency is dramatically enhanced in that the subband filter with 3125 taps reduced to 4 comb filters cascaded with a six-looped nested allpass filter.

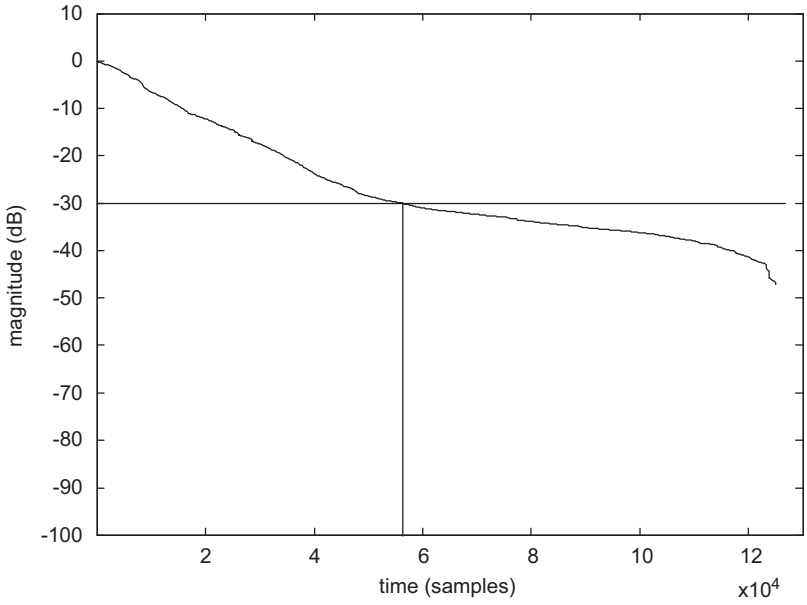

Fig. 10. EDC of the concert hall. T_{30} is also indicated in the figure.

The implementation of the third subband structure is considered next. First, the interval of the repetitive pattern in a subband impulse response must be estimated for the delay term of the comb filter. The gain of the comb filter is determined by Eq. (18). The FIR filter coefficients are then calculated via letting $M = N$ in Eq. (21). Experience indicates that eight subbands are appropriate for this problem. The time domain and frequency domain first subband responses of the synthesized and the measured responses are compared in Fig. 12. The agreement is reasonably consistent and no significant coloration is introduced by the comb filter. Furthermore, the structure embedded in the early portion of the impulse response was preserved well by the FIR filter. The parameters of the eight subbands including the gains and delays of the comb filters are summarized in Table 2. Since the energy at high frequency decays more rapidly than that at low frequency, the corresponding gains of high frequency subbands are smaller.

In order to compare the above-mentioned three subband structures, computational loading and subjective performance are considered. First, the computational loading of the three structures is compared in Table 3. According to EDC the reverberation times (T_{60}) of the reverberators implemented using three structures are calculated and shown in Table 4. It is seen that the filtering efficiency of the second structure is lowest among all. Subjective listening tests involving 10 human subjects were then performed to evaluate the rendering performance of the three approaches. A headphone was used as the means for reproduction. Subjective indices, spaciousness, clarity, naturalness, and richness are adopted in the tests. Each index is assessed in terms of five discrete levels: 1 = very poor, 2 = poor, 3 = fair, 4 = good and 5 = perfect. The meaning of these subjective indices was fully explained to the participants prior to the listening test. A symphony music and a female speech are used as the audio inputs. Each subject was asked to compare the audio signals synthesized by the measured room response with those synthesized by proposed methods. Each index is divided into the above-mentioned five discrete levels to indicate how close each subject perceives between the synthesized signal and the measured room response. The test results for the reverberators using three subband structures are illustrated with error-bar plots of Figs. 13 (a) and (b). Several observations can be derived from these experimental results. The spaciousness resulting from the 32-subband regular IIR filters is somewhat lower than the other two structures, even though the common problem with headphone of in-head rendering has been eliminated. Irrespective of the methods used, a trade-off is evident between the two indices, spaciousness and clarity. Three methods exhibit similar behavior for both types of audio input. The total scores defined as

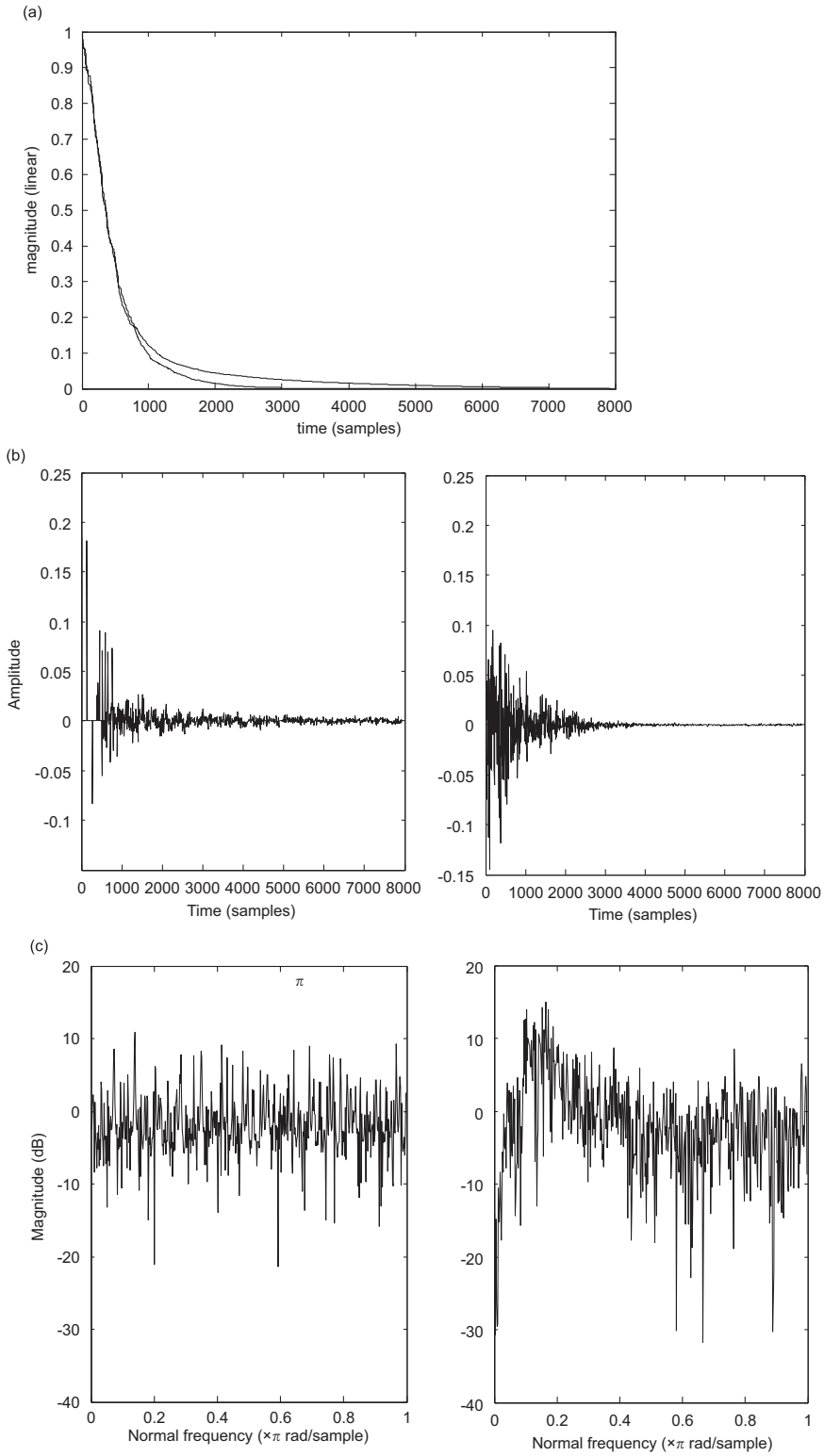

Fig. 11. Implementation results of the second structure for the first subband. (a) Comparison of the EDC resulting from the synthesized subband filter and the measured filter, (dashed line: synthesized response; solid line: measured response), (b) comparison of the impulse responses resulting from the synthesized subband filter and the measured filter, (left side: synthesized response; right side: measured response) and (c) comparison of the frequency responses resulting from the synthesized subband filter and the measured filter (left side: synthesized response; right side: measured response).

Table 1
Coefficients of the comb and nested allpass filters in the first subband for the second structure.

Filter	Comb filter				Nested allpass filter					
Gain	0.8924	0.8632	0.8349	0.8059	0.3771	0.5535	0.8078	0.7866	0.8388	0.5739
Delay	58	75	92	110	173	148	231	96	116	135

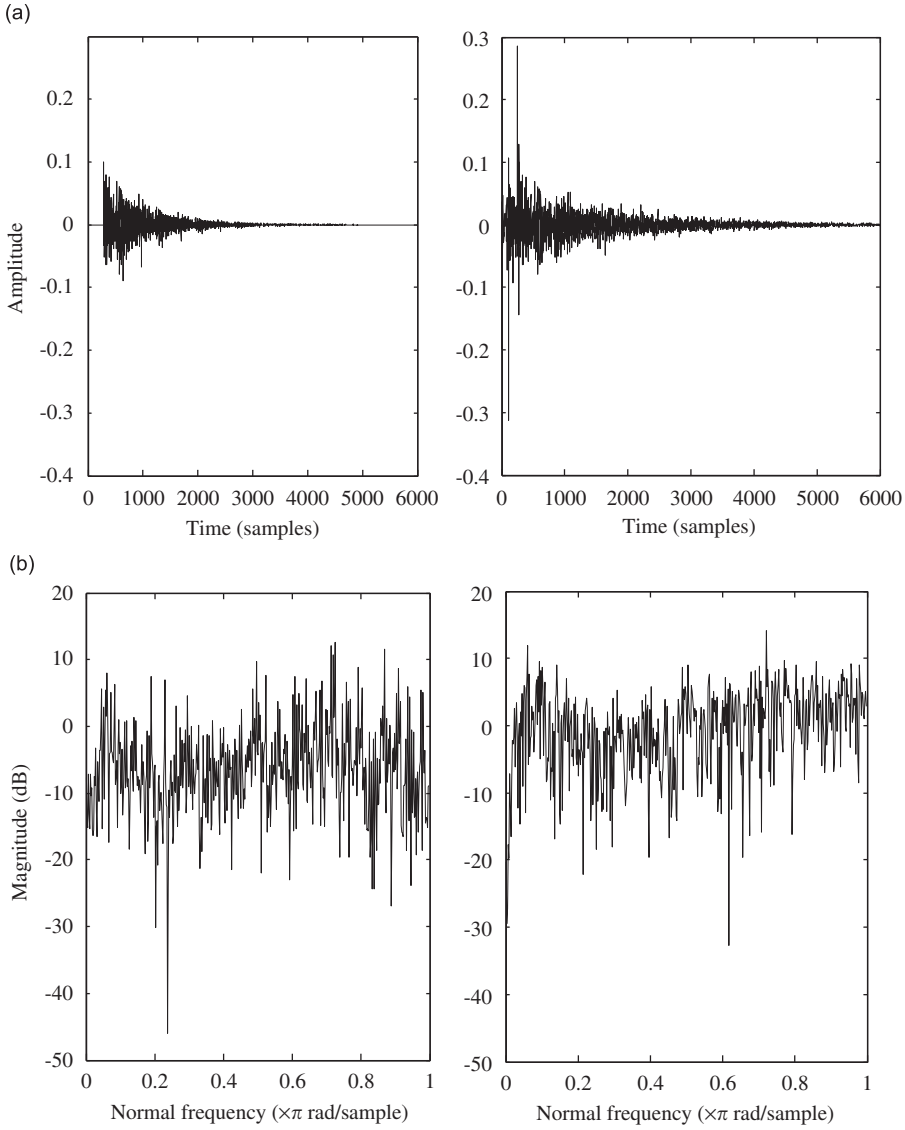

Fig. 12. Implementation results of the third structure for the first subband. (a) Comparison of the impulse responses resulting from the synthesized subband filter and the measured filter and (b) comparison of the frequency responses resulting from the synthesized subband filter and the measured filter (left side: synthesized response; right side: measured response).

the sum about the four indices and both audio inputs are 30.0, 28.5, and 30.2 for the first, the second, and the third structures, respectively. This listening test indicates that the third structure performs best in the subjective evaluation at the expense of computational cost. Even though the performance of the second

Table 2
Coefficients of the comb filters for the third structure.

Band	Band 1	Band 2	Band 3	Band 4	Band 5	Band 6	Band 7	Band 8
Gain	0.6363	0.6315	0.7916	0.6781	0.4778	0.5547	0.2391	0.255
Delay	360	295	124	138	190	115	122	91

Table 3
Comparison of computational loading in the three structures.

Model	Model 1	Model 2	Model 3
Number of subbands	32	16	8
Order of prototype filter	231	115	57
Total taps of FIR filter	None	None	1435
Total number of comb filters	None	64	10
Total number of allpass filters	None	96	None
Total order of IIR filter	1234	None	None
Design complexity	Median	High	Low
Computation complexity	High	Low	Median

Table 4
Reverberation times of reverberators implemented using the three structures.

T_{60}	Model		
	Model 1	Model 2	Model 3
Measured	0.0226	0.0680	0.1360
Synthesized	0.0226	0.1864	0.1133

structure was somewhat poorer in clarity and naturalness than the other two models, significant computational saving and superior spaciousness has been achieved. This may be attributed to the fact that the filter parameters of the second structure are optimized by using the GA procedure.

6. Conclusions

Efficient realization of natural reverberation based on a measured room response has been developed in this paper. The design procedures of the reverberator in conjunction with three subband filtering schemes are described in detail. IIR structures are chosen to implement these subband filters to further enhance the efficiency of subband filtering. These IIR structures involve the regular IIR filter, the cascaded comb filters and nested allpass filters with coefficients optimized via GA, and the cascaded comb filter and an FIR filter. Cosine modulated pseudo QMF is used to implement the subband reverberation filters because of its high stopband attenuation that is necessary for avoiding ringing problems. The proposed methods fully exploit parallelism provided by the maximally decimated filter bank structures and the frequency dependent properties of reverberations.

The computational loading and a subjective rendering performance were compared for the three structures. From the results, the proposed techniques prove effective in delivering quality reproduction of room reverberations. In particular, the third structure performs best in the subjective evaluation, suggesting that this

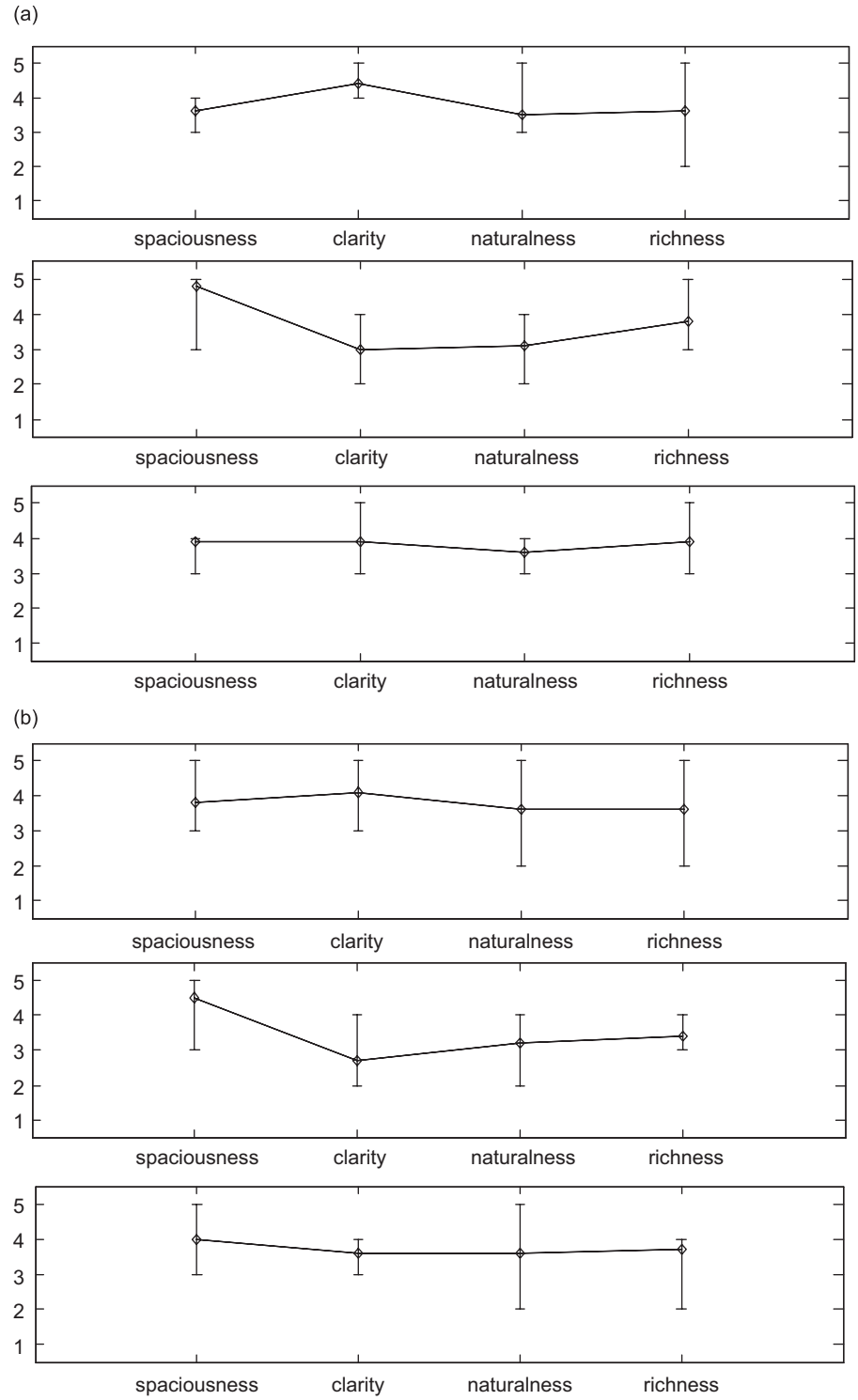

Fig. 13. Results of subjective listening test for reverberation effects. (a) Symphony music and (b) female speech. The upper, middle, and lower diagrams represent the results of the first, the second, and the third structures, respectively. The scale on the vertical axis: 1 = very poor, 2 = poor, 3 = fair, 4 = good, and 5 = excellent. The total scores defined as the sum about the four indices and both audio inputs are 30.0, 28.5, and 30.2 for the first, the second, and the third structures, respectively.

structure captures the salient features of practical room acoustics at the expense of computational cost. While the performance of the second structure was somewhat poorer in clarity and naturalness than the other two models, it achieves significant computational saving and exhibits comparable spaciousness as the measured room response.

Acknowledgments

The work was supported by the National Science Council in Taiwan, Republic of China, under the Project number NSC 87-2212-E009-022. Special thanks also go to Dr. Yuan-Pei Lin in the Department of Control Engineering of National Chiao-Tung University for her helpful suggestions on the design of filter banks.

References

[1] Lake Technology Ltd. 〈http://www.lake.com.au〉.
[2] J.M. Jot, An analysis/synthesis approach to real-time artificial reverberation, *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Vol. 2, 1992, pp. 221–224.
[3] J.A. Moore, The synthesis of complex audio spectra by means of discrete summation formulae, *Journal of the Audio Engineering Society* 24 (1976) 717–727.
[4] J.A. Moore, About the reverberation business, *Computer Music Journal* 3 (1979) 3255–3264.
[5] W.G. Gardner, Efficient convolution without input–output delay, *Journal of the Audio Engineering Society* 43 (1995) 127–136.
[6] M. Bosi, Filter banks in perceptual audio coding, *AES 17th Conference on High-quality Audio Coding*, 2–5, September, Florence, Italy, 1999.
[7] P.P. Vaidyanathan (Fellow, IEEE), Orthonormal and biorthonormal filter banks as convolvers, and convolutional coding gain, *IEEE Transactions on Signal Processing* 41 (6) (1993) 2110–2130.
[8] M. Vetterli (Member, IEEE), Running FIR and IIR filtering using multirate filter banks, *IEEE Transactions on Acoustics, Speech and Signal Processing* 36 (5) (1988) 730–738.
[9] C.A. Lanciani, R.W. Schafer, Subband-domain filtering of MPEG audio signals, *IEEE International Conference on Acoustics, Speech and Signal Processing*, 1999.
[10] J.R. VandeKieft, Computational Improvements to Linear Convolution with Multi-rate Filtering Methods, Master Thesis, University of Miami, Florida, 1998.
[11] N.R. Rayavarapu, N.R. Prakash (Member, IEEE), A computationally efficient design for prototype filters of an M-channel cosine modulated filter bank, *Proceedings of World Academy of Science, Engineering and Technology* 13 (2006) (ISSN 1307–6884).
[12] F. Cruz-Roldán (Member, IEEE), P. Amo-López (Associate Member, IEEE), S. Maldonado-Bascón (Associate Member, IEEE), S.S. Lawson (Senior Member, IEEE), An efficient and simple method for designing prototype filters for cosine-modulated pseudo-QMF banks, *IEEE Signal Processing Letters* 9 (1) (2002) 29–31.
[13] P.P. Vaidyanathan, *Multirate Systems and Filter Banks*, PTR Prentice-Hall, Englewood Cliffs, NJ, 1993.
[14] Y.-P. Lin (Member, IEEE), P.P. Vaidyanathan (Fellow, IEEE), A kaiser window approach for the design of prototype filters of cosine modulated filterbanks, *IEEE Signal Processing Letters* 5 (6) (1998) 132–134.
[15] A. Grace, A.J. Laub, J.N. Little, C.M. Thompson, *Matlab Control System Toolbox*, The Math Works, Inc., Natick, MA, 1999.
[16] E.C. Levi, Complex-curve fitting, *IRE Transactions on Automatic Control* 4 (1959) 37–44.
[17] J.E. Dennis Jr., R.B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice-Hall, Englewood Cliffs, NJ, 1983.
[18] M.R. Bai, *Engineering Acoustics*, Chuan-Hua Scietific Publisher, 2004.
[19] M.R. Bai, K.-Y. Ou, Synthesis of room responses using virtual source representation with application in reverberator design, 2005.
[20] M.R. Schroeder, New method of measuring reverberation time, *Journal of the Audio Engineering Society of America* 37 (1965) 409–412.
[21] M.R. Schroeder, Natural sounding artificial reverberation, *Journal of the Audio Engineering Society* 10 (1962) 219–233.
[22] W.G. Gardner, Reverberation algorithm, in: M. Kahrs, K. Brandenburg (Eds.), *Applications of Signal Processing to Audio and Acoustics*, Academic Publishers, Norwell, MA, 1998.
[23] U. Zolzer, *Digital Audio Signal Processing*, Wiley, New York, 1997.
[24] B.S. Chen, Y.M. Cheng, A structure-specified H_∞ optimal control design for practical applications: a genetic application, *IEEE Transactions on Control Systems Technology* 6 (6) (1998).
[25] T. Kozek, T. Roska, L.O. Chua, Genetic algorithm for CNN template learning, *IEEE Transactions on Circuits and Systems* 40 (1993) 392–402.
[26] Physikalisch-Technische Bundesanstalt, Project 1.401: simulation of room acoustics 〈http://www.ptb.de/en/org/1/14/1401/_index.htm〉.