Comparative Study of Optimal Design Strategies of Reverberators

Mingsain R. Bai[*] and Huiwen Chen

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

Abstract: Artificial reverberator is a key element in spatial audio reproduction. This paper compares various optimal design strategies of artificial reverberators for room response simulation. From the comparison, it is hoped that guidelines for design of natural-sounding reverberators can be found. Infinite impulse response (IIR) filters such as allpass filters or comb filters are used as building blocks for the design. Early reflections of room responses are modeled by three different approaches: the traditional image method, the exponential truncation method, and network approximation method. On the other hand, comb/nested allpass filter networks are exploited for modeling late reverberations, with parameters predetermined by a genetic algorithm (GA)-based procedure. Up/down-sampling and spectral band replication (SBR) techniques are also employed to save memory storage and recover the lost high frequency portion of the reverberation. Subjective listening tests were conducted to compare the proposed artificial reverberator modules. The test results revealed that the reverberator in which early reflections are modeled by network approximation and late reverberations are modeled by comb/nested allpass filter network is superior in delivering natural-sounding reverberation among all methods.

1. INTRODUCTION

Reverberation is an acoustical phenomenon as a result of infinite number of reflections from the boundary in an enclosure. Reverberation embodies the characteristics of the listening environment and is hence an important ingredient of natural-sounding audio signals. Artificial reverberations are often employed to enrich dry recordings for practical reproduction. Reverberation can generally be divided into two generic portions of early reflections and late reverberation, as shown in Fig. (**1**). In typical room impulse responses, early reflections that may last for as much as 80 ms, depending on room geometry as well as the relative positions of the source and the listener, whereas the late reverberation is comprised of dense reflections that exhibit little structure but exponentially decreasing envelop in the time domain. These two parts of room responses carry important acoustical cues for human perception of the listening environment such as the size of room and the absorptivity of the boundary.
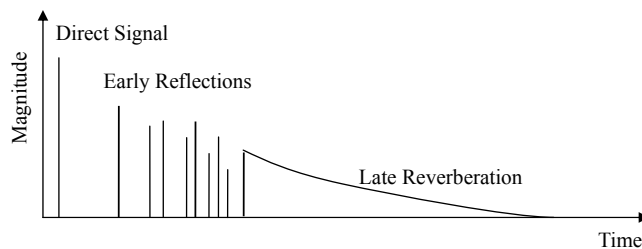
Fig. (1). A typical room impulse response.

A straightforward approach of realizing reverberation is direct filtering using the impulse response, or the Finite Impulse Response (FIR) filter, measured in a room. However,

this is usually deemed impractical for real-time applications due to its excessive computation complexity. Instead of the brute-force approach, artificial reverberators are commonly used at the expense of some timbral quality and naturalness. A simple approach of artificial reverberator is called the image method [1] that models the reflections from the room boundary using the sound waves emanated from a finite number of image sources. Instead of pursuing exact models that are prohibitively complex, it is often more practical to find alternative models capable of capturing salient psychoacoustical attributes of natural reverberation within reasonable cost of implementation. Schroeder [2] proposed an Infinite Impulse Response (IIR)-based filter network to simulate room reverberations. In his pioneering work, he introduced recursive comb filters and allpass filters as filter building blocks to design the reverberators. Later, Gerzon [3] generalized the single-channel Schroeder reverberator to a multichannel reverberator by replacing the delay line with a 'unitary' network [4]. Stautner and Puckette [5] introduced feedback delay networks (FDN's) for synthesizing reverberation. This structure consists of a set of delay lines connected in a feedback loop through a 'feedback matrix'. The FDN's generalize the recursive feedback comb filter by replacing the single delay line by a diagonal matrix of delay lines of different lengths, and by replacing the feedback gain by the matrix $\mathbf{G} = \mathbf{UD}$, where $\mathbf{U}$ is any unitary matrix and $\mathbf{D}$ is a diagonal matrix having all elements less than $1-\varepsilon$ in magnitude ($\varepsilon > 0$ determines the stability margin). Jot [6] suggested a FDN reverberator in which a lowpass filter is placed in series with each delay line for simulating the absorptive losses. Bai and Bai [7] proposed a filter network reverberator consisting of ten parallel comb filters cascaded with nested allpass filters. Schroeder's comb/allpass filters network, the FDN, and Bai's comb/nested allpass filters network are IIR filters well suited to simulating the late reverberations of room responses. As for the early reflections that are rela-

*Address correspondence to this author at the Department of Mechanical Engineering, National Chiao-Tung University, 1001 Ta-Hsueh Road Hsin-Chu 300, Taiwan; E-mail: msbai@mail.nctu.edu.tw

tively sparse and structured, tapped delay lines (TDL) in FIR filter form are generally appropriate for implementation [9].

This paper aims to develop a systematic procedure for the design of natural-sounding reverberators with low computational cost. Three methods, the conventional image method, the exponential truncation method and the network approximation method, are employed to model the early reflections of room responses. Among these three methods, the image method and the exponential truncation method are categorized as FIR-based approaches, whereas the network approximation method is categorized as an IIR-based approach. On the other hand, the late reverberation is always modeled by comb/nested allpass filter networks in the paper. A common problem with this kind of reverberator is that inadequate choice of filter parameters could lead to artifacts such as metallic and ringing noises. It is then desirable to have a systematic and efficient way in search of optimal filter parameters. To combat this problem, a search procedure extended from the previous paper [7] that makes use of the Genetic algorithms (GA) is presented to minimize the effort of trial and error in designing the artificial reverberators. However, the present work differs itself from Ref. [7] in that the structure of comb/nested allpass filter networks is modified. As a simplification, a single first-order IIR lowpass filter is placed after the nested allpass filter in substitution of the absorptive filters in comb filters used in [7]. Up/down-sampling technique is exploited to ease the requirement of memory storage and processing speed. Under a lower sample rate, much fewer delays are needed in the networks at the cost of a reduced bandwidth. To address the problem, the spectral band replication (SBR) [8] technique is employed to recover the lost high frequency signals.

Three subjective listening tests were conducted for assessing the performance of the proposed reverberators. The first test compared the field measurements with the comb/nested allpass filter networks in modeling late reverberations. The second test compared three methods in modeling early reflections. The third test investigated the reverberators with and without the up/down sampling and SBR processing. Apart from the subjective tests, computational cost required by each reverberator was also compared in this paper.

2. SYNTHESIS OF REVERBERATION

2.1. Characteristics of Room Responses

Natural room responses contain highly complex dynamics resulting from not only the direct sound but the infinite number of reflections from the room boundary. Room responses are characterized by a number of acoustical properties such as reverberation time, impulse response and frequency response, which are dependent of boundary absorption, room geometry and volume of the room, relative positions of the source and receiver, etc. A number of characteristics of the room response are relevant in the design of reverberators. First, echo density is defined in the time domain as the number of echoes per second in the room response [10]:

$$E_d = \frac{4\pi c^3}{V} t^2 , \qquad (1)$$

where t is the time variable, c is the speed of sound and V is the volume of room. Note that echo density is proportional to the squares of the time. The second characteristic is the modal density defined in the frequency domain as the number of normal modes per Hertz. Modal density can be estimated using the following formula [10]:

$$M_d = \frac{dN_f}{df} \approx \frac{4\pi V}{c^3} f^2 , \qquad (2)$$

where N_f is the number of normal modes below the frequency f. Hence, the modal density of a room grows in proportion to the squares of frequency. Third, the reverberation time T_{60} is the time required for the sound pressure level (SPL) to decay by 60 dB after a steady-state source is switched off. Reverberation time can be estimated by the Sabine's formula [11]:

$$T_{60} = \frac{0.163 \cdot V}{\sum_i a_i S_i} , \qquad (3)$$

where V is the volume of the room, S_i and a_i are the surface area and the associated absorption coefficient of the boundary element i. Reverberation time is proportional to the volume of the room and inversely proportional to the wall absorptivity and the interior surface area of the room. Because most materials become more absorptive at high frequencies, reverberation time generally decreases with the frequency. A method for estimating reverberation time from the room impulse response was proposed by Schroeder [12]. His method estimates the reverberation time for which the level of the energy decay curve (EDC) decays by 60 dB.

$$EDC(t) = \frac{\int_t^\infty h^2(\tau)d\tau}{\int_0^\infty h^2(\tau)d\tau} , \qquad (4)$$

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

2.2. Modeling Late Reverberations

In the section, comb filters and allpass filters are used as filter building blocks to construct natural-sounding reverberation networks. As mentioned previously, this is a realistic approach to reconcile the reverberation quality and computation loading. Fig. (**2**) illustrates the artificial reverberator proposed by Schroeder [2], which consists of four parallel comb filters and two serial allpass filters. The comb filters were modified by Moorer [13], as shown in Fig. (**3**), considering the fact that the reverberation time usually decreases with frequency. The modified comb filter is given by the following transfer function:

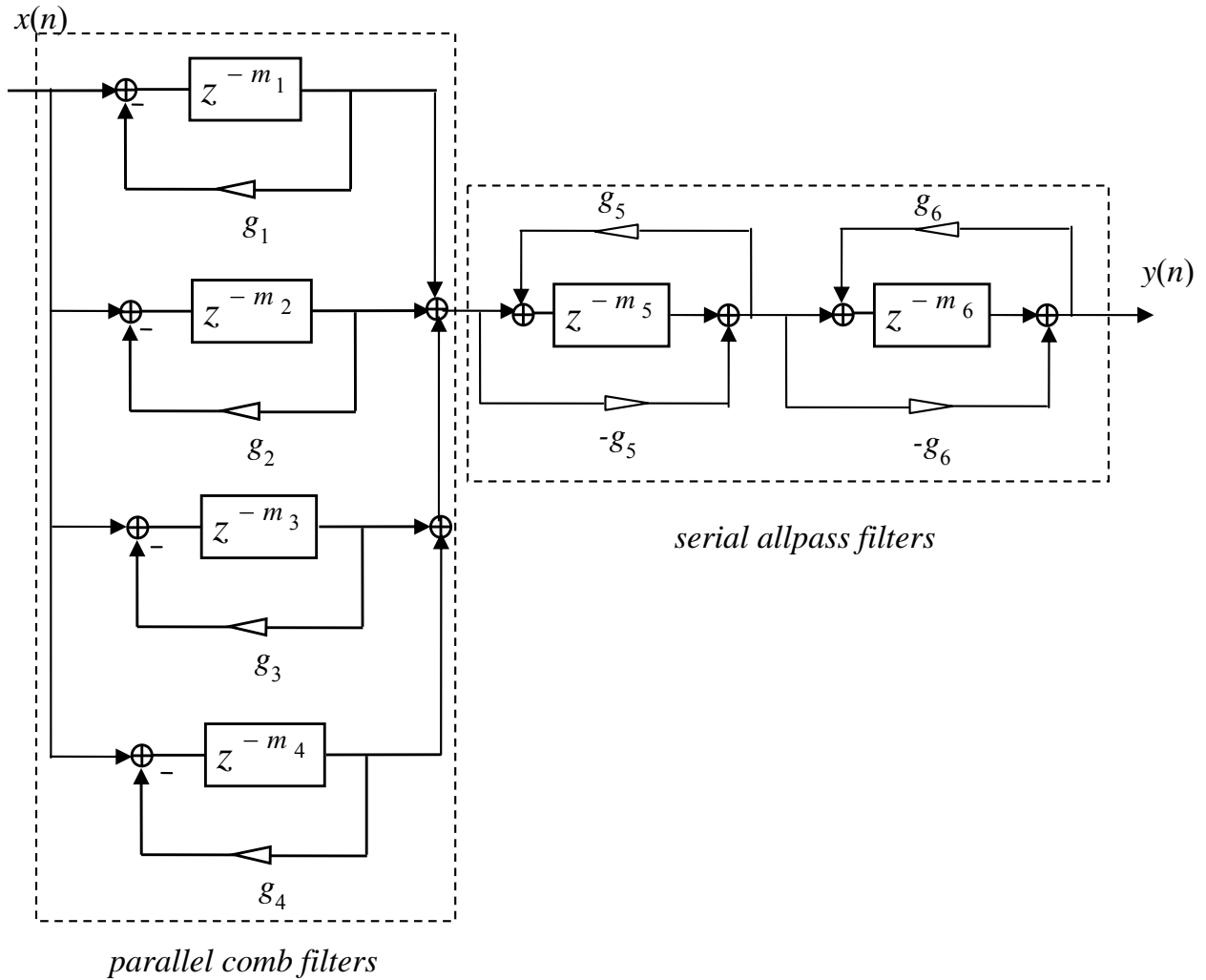$$H(z) = \frac{z^{-m}}{1 - gH_L(z)z^{-m}} , \qquad (5)$$

Fig. (2). Schroeder's artificial reverberator consisting of four parallel comb filters and two series allpass filters.

where $H_L(z) = \dfrac{1}{1 - az^{-1}}$ being an one-pole lowpass filter to simulate high frequency absorption. The parameters, a and g, must satisfy the stability condition, $\dfrac{g}{1-a} < 1$. Gardner [14] proposed artificial reverberators that make use of a "nested" allpass filter. Fig. (**4**) shows the block diagram of a two-layered nested allpass filter and its impulse response function. The transfer function of a two-layered nested allpass filter is given by

$$H(z) = \frac{-g_2 + g_1 g_2 z^{-m_1} - g_1 z^{-m_2} + z^{-(m_1+m_2)}}{1 - g_1 z^{-m_1} + g_1 g_2 z^{-m_2} - g_2 z^{-(m_1+m_2)}}, \qquad (6)$$

where the parameters g_i and m_i ($i = 1, 2$) are the gains and the delays of the nested allpass filter. There are two noteworthy features of the nested allpass filters. First, the same pattern of response never repeats in the time domain. Second, the echo density increases with time, which is closed to the late reverberation of a real room response. Bai and Bai [7] suggested artificial reverberators with a network comprising ten parallel modified comb filters and a three-layered nested allpass filter. This network is slightly modified in this paper into

another structure that involves three comb filters cascaded with a three-layered nested allpass filter, as shown in Fig. (**5**). In the network, three parallel comb filters (instead of ten filters used in the network in [7]) serve to increase both modal density and echo density, whereas the three-layered nested allpass filter serves to further increase echo density of reverberation. In order to simulate the high-frequency absorption effec, a first-order IIR lowpass filter is placed after the three-layered nested allpass filter, as given by the following transfer function:

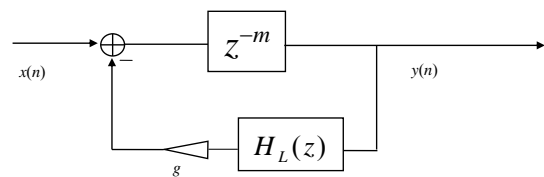$$H(z) = \frac{b(1 + z^{-1})}{1 + az^{-1}}. \qquad (7)$$

Fig. (3). Moorer's modified comb filter with a first-order lowpass filter $H_L(z)$ inserted in the feedback loop.
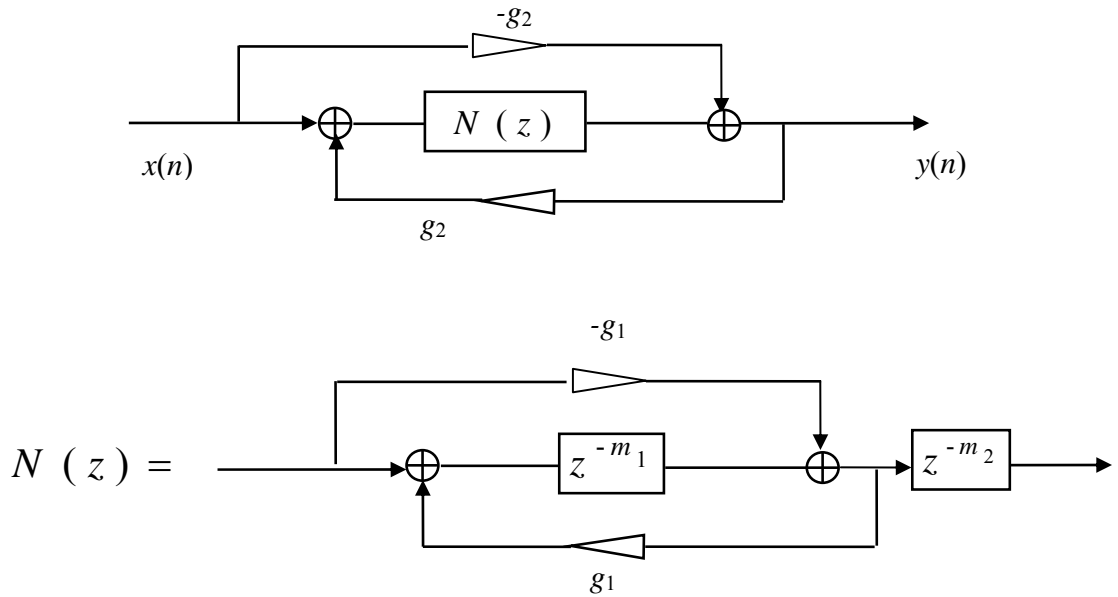
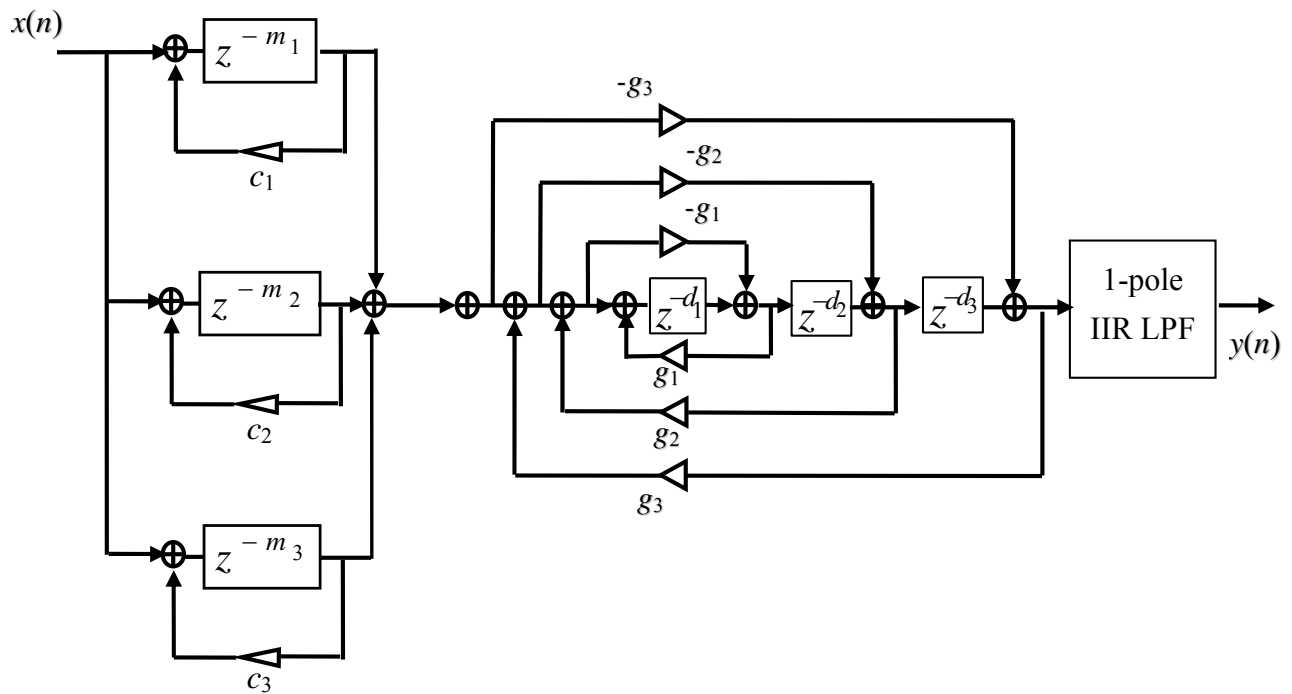**Fig. (4).** The block diagram of a two-layered nested allpass filter.

Fig. (5). The structure of the reverberator consisting of three comb filters and three-layered nested allpass filter cascaded with a first-order lowpass filter.

This represents a simplification to the reverberator previously suggested by Bai and Bai [7], where each comb filter requires an absorbent filter to account for high frequency absorption.

2.3. Modeling Early Reflections

Three methods are presented in this section to model the early reflections of room responses.

- *The image method*. In this method, a cubic room model is employed to construct image sources. The impulse response of the room is calculated by summing the impulses from all image sources. Experience suggests that sixth order of reflections produced by 376 image sources generally suffices to approximate the early reflections.

- *The exponential truncation method*. This method aims to simplify the early reflections by retaining only the

major portion of the impulse response of the room. To be specific, the following exponential threshold is utilized to truncate the measured impulse response:

$$T(n) = c_1 r^{-n}, \quad 0 < r < 1, \tag{8}$$

where $T(n)$ is the threshold at the time instant n. In the expression, parameter c_1 determines the initial value of $T(n)$, while c_2 determines the decay rate of $T(n)$. The impulse response at the instant n will be retained if its amplitude is greater than $T(n)$. Otherwise, the value will be reset to zero. An example of the use of the exponential threshold method is shown in Fig. (**6**). Fig. (**6a**) shows the impulse response of Bethel Church, where the dotted line represents the threshold curve with $c_1 = 0.05$ and $r = 0.92$. After applying the threshold, the 3528-tapped room response is truncated to only 86 coefficients, as shown in Fig. (**6b**).

- *Filter network approximation method*. In this method, three parallel comb filters and a three-layered nested allpass filter are employed to approximate the early reflections of room responses. The relative positions of source and receiver and room geometry are irrelevant in this approach. Parameters of the filter network are obtained by matching the frequency response functions of the early reflections between the measured response and the artificial network. The matching procedure is carried out by using the genetic algorithm which will be introduced in the following section.

3. OPTIMAL DESIGN OF ARTIFICIAL REVERBERATORS

To find adequate parameters of a network that gives natural sounding reverberation with no artifacts is a tedious and time consuming task. In the section, the genetic algorithm (GA) will be presented for optimizing the parameters of the comb/nested allpass filter network in an automatic and efficient manner. Six types of rooms including a living room, a medium room, a theater, a church, a gym and a great hall are investigated.

The GA procedure begins with encoding all parameters into binary strings called the *chromosomes* [15]. The resolution of a parameter is dependent on the number of bits per string in the search domain. The search objective is termed the *fitness* function. A chromosome with high fitness has higher probability to survive the natural selection and to reproduce offspring in the next generation. *Crossover* enables exchanging genes in the chromosomes *via* probabilistic decision in the mating pool. *Mutation* is crucial for preventing the premature convergence problem as the genes become increasingly homogeneous. The mutation point is randomly chosen. Mutation is then carried out by alternating the gene from zero to one, and vice versa.
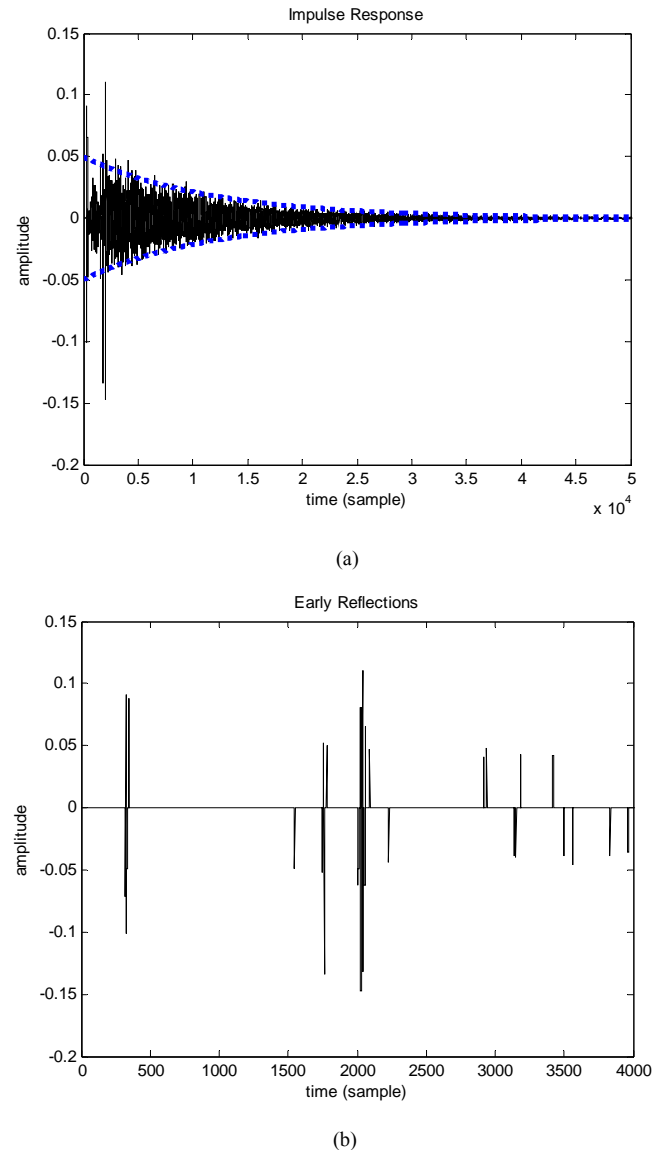

(a)

(b)

Fig. (6). An example for the exponential truncation method. (**a**) The room impulse response of Bethel Church. The dotted line represents the threshold. (**b**) The early reflection of Bethel Church extracted by using the exponential truncation method.

To model the late reverberation using the parallel comb filters and the nested allpass filter, the parameters to be optimized are the delays and the gain of the comb filter, the delays and the gain of the allpass filter, and the coefficient of the first-order IIR lowpass filter. However, instead of searching these fourteen parameters by GA all at one time, the optimization procedure is divided into four steps according to different purposes of the filters. The flow chart of the present optimization procedure is shown in Fig. (**7**), where the coefficients to be optimized in the first step are three delays and three gains, d_i and g_i, $i = 1, 2, 3$, of the three-layered nested allpass filter. The upper and lower limits for the delays are selected to be 1000 and 50, respectively. The upper and lower limits for the gains are selected to be 1 and 0.1, respectively. The objective of optimization in this step is to achieve
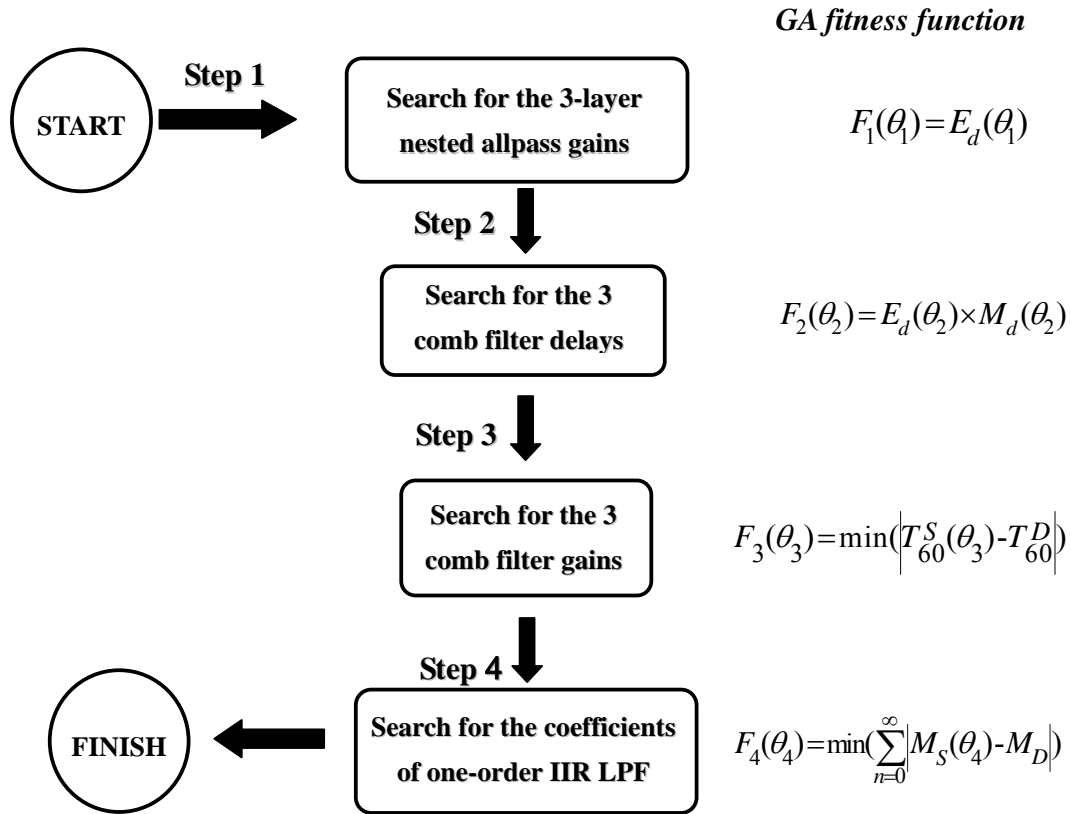
GA fitness function

Fig. (7). The flowchart of the GA optimization procedure.

a response with high echo density E_d. Accordingly, the fitness function $F_1(\theta_1)$ is defined as:

$$F_1(\theta_1) = E_d(\theta_1), \tag{9}$$

where $\theta_1 = [d_1 \ d_2 \ d_3 \ g_1 \ g_2 \ g_3]$ is the chromosome. This is different from the fitness function chosen in [7] in which an energy term is included.

In the second step, the coefficients to be optimized are three delays of the parallel comb filters m_i ($i = 1, 2, 3$). The upper limit of delay is set to be 3840 samples, or 80 ms at 48k Hz sampling rate. The lower limit is 480 samples, or 10 ms, for a medium size room. The purpose here is to find the best chromosome to achieve the maximum echo density and modal density simultaneously. The fitness function $F_2(\theta_2)$ is defined as

$$F_2(\theta_2) = E_d(\theta_2) \times M_d(\theta_2), \tag{10}$$

where $\theta_2 = [m_1 \ m_2 \ m_3]$ and M_d is the modal density. To trade-off between echo density and modal density, a multiplication is used in Eq. (10) because modal density generally decreases as echo density increases for the comb filter. In the third step, the design parameters are the gains of comb filters c_i, $i = 1, 2, 3$. The fitness function is the difference of reverberation times (T_{60}) between the desired room and the synthesized room:

$$F_3(\theta_3) = \min\left(\left|T_{60}^S(\theta_3) - T_{60}^D\right|\right), \tag{11}$$

where $\theta_3 = [c_1 \ c_2 \ c_3]$, T_{60}^S and T_{60}^D denote the reverberation time estimated by the EDC curves of the synthesized room and the desired room, respectively. In the last step, the design parameters are the coefficients, a and b, of the first-order IIR lowpass filter. The fitness function in this step is chosen to be the difference of magnitude of frequency response functions between the synthesized room and the desired room:

$$F_4(\theta_4) = \min\left(\sum_{n=0}^{\infty}\left|M_S(\theta_4) - M_D\right|\right), \tag{12}$$

where $\theta_4 = [a \ b]$, M_S and M_D are the magnitude of frequency response functions of the synthesized room and the desired room, respectively. There are forty populations formed randomly for each parameter. Each population contains eight chromosomes. The crossover rate and mutation rate are 0.85 and 0.008, respectively. After executing each step for 100 generations using the aforementioned GA procedure, optimal parameters of the filter network are obtained, as summarized in Table **1**. The reverberation times of each room of the measured and the synthesized responses are listed in Table **2**. There is only slight difference in the reverberation time between the synthesized and the measured room responses. Fig. (**8a-f**) illustrate the measured and the synthesized impulse response functions of the rooms. It can be seen that the synthesized impulse responses match quite well to the measured impulse responses. Fig. (**9a-f**) show the corresponding frequency responses of rooms between the

measured and the synthesized responses. The synthesized frequency responses also match quite well to the measured frequency responses.

Table 1.   The Optimized Parameters of Comb Filters and Nested Allpass Filters

Filter Number	1	2	3
Comb delay (m_i)	1784	1712	1482
Allpass delay (d_i)	808	228	539
Allpass gain (g_i)	0.5922	0.4623	0.4746

Table 2.   The Reverberation Time (Seconds) of Six Types of Rooms

T_{60}(s) / Room Type	Living Room	Medium Room	Theater	Church	Gymnasium	Great Hall
Measured	0.2539	0.6735	1.2028	1.6236	2.4112	2.6325
Synthesized	0.2488	0.674	1.1134	1.5847	2.4112	2.7438

The computation cost of the reverberators presented in this paper and the direct convolution using measured response for a large room are compared in Table **3**. The module with early reflections modeled by the filter network approximation method and late reverberation modeled by the comb/nested allpass filters network method requires the lowest computational cost, while the direct convolution using measured response is the most computationally demanding.

Table 3.   The Computation Cost of Reverberator Modules

Room Mode: Great Hall			
Reverberator Module	**APU**	**MPU**	**Memory (Byte)**
Direct convolution	168527	168528	337056
Im/Network	76	78	26674
Exp/Network	58	60	22456
Fi/Network	54	56	23582

APU-Additions per unit time; MPU-Multiplications per unit time; Direct convolution-Direct convolution using the measured room response; Im/Network-Full reverberator with image method modeling early reflections and comb/nested allpass filters network modeling late reverberation; Exp/Network-Full reverberator with exponential truncation modeling early reflections and comb/nested allpass filters network modeling late reverberation; Fi/Network-Full reverberator with filter network approximation modeling early reflections and comb/nested allpass filters network modeling late reverberation.

4. EXPERIMENTAL INVESTIGATIONS

In order to assess the perceptual performance of the artificial reverberators, a series of subjective listening tests are conducted according to the double-blind triple-stimulus with hidden reference method suggested in the standard ITU-R BS. 1116-1 [16] and the multi-stimulus test with hidden reference and anchor (MUSHRA) suggested in the standard ITU-R BS. 1534 [17]. There are fifteen experienced subjects participating in these subjective tests.

- *Experiment-1*. The first listening test is conducted according to the method of the double-blind triple-stimulus with hidden reference for assessing the late reverberation simulators. The measured room impulse response with the direct impulse and the early reflections (before 80ms) removed served as the reference of the late reverberation. The hidden reference and the synthesized late reverberation are randomly assigned to be Cases 1 and 2. The grading scale is set to be 1~5, as recommended in ITU-R BS. 1284 [18]. Six types of rooms including living room, medium room, theater, church, gymnasium and great hall were investigated in this test. A headset is used for sound rendering. The sound pressure level is measured using the pink noise as the input at the point 10cm away from the loudspeaker of the headset. The gain of the headset was adjusted to ensure the loudness to be equal for all reproduced signals. Three subjective indices including *overall quality*, *spatial quality* and *timbral quality* were employed in this listening test. Every subject participating in the test was instructed with the definitions of the preceding subjective indices and the procedure before the listening test. The listeners can switch the cases randomly during the test. The results of the listening test are shown in Figs. (**10a-g**). The vertical bars denote 0.95 confidence intervals. It can be observed from the test results that the grades of overall quality range from 3 to 5. This indicates that the subjective difference between the measured and the synthesized room responses are not statistically significant.

- *Experiment-2*. The second listening test is conducted to assess the three early reflection simulators presented previously. The MUSHRA method is adopted to process the data in this test. The synthesized late reverberation part in experiment-1 plus the direct sound is chosen to be the known reference. A mono signal obtained by averaging the two-channel stereo signals and filtering with a 128-tap FIR lowpass filter of 3 kHz cut-off was employed as the anchor. The known reference along with three early reflection simulators are compared. In this test, the three types of rooms including the living room, the theater and the great hall are chosen to represent the small room, the medium room and the large room, respectively. The grading scale ranges from -3 to +3. Three subjective indices including *total preference*, *sense of presence* and *timbral quality* were employed in the test. The score of the known reference is designated to be
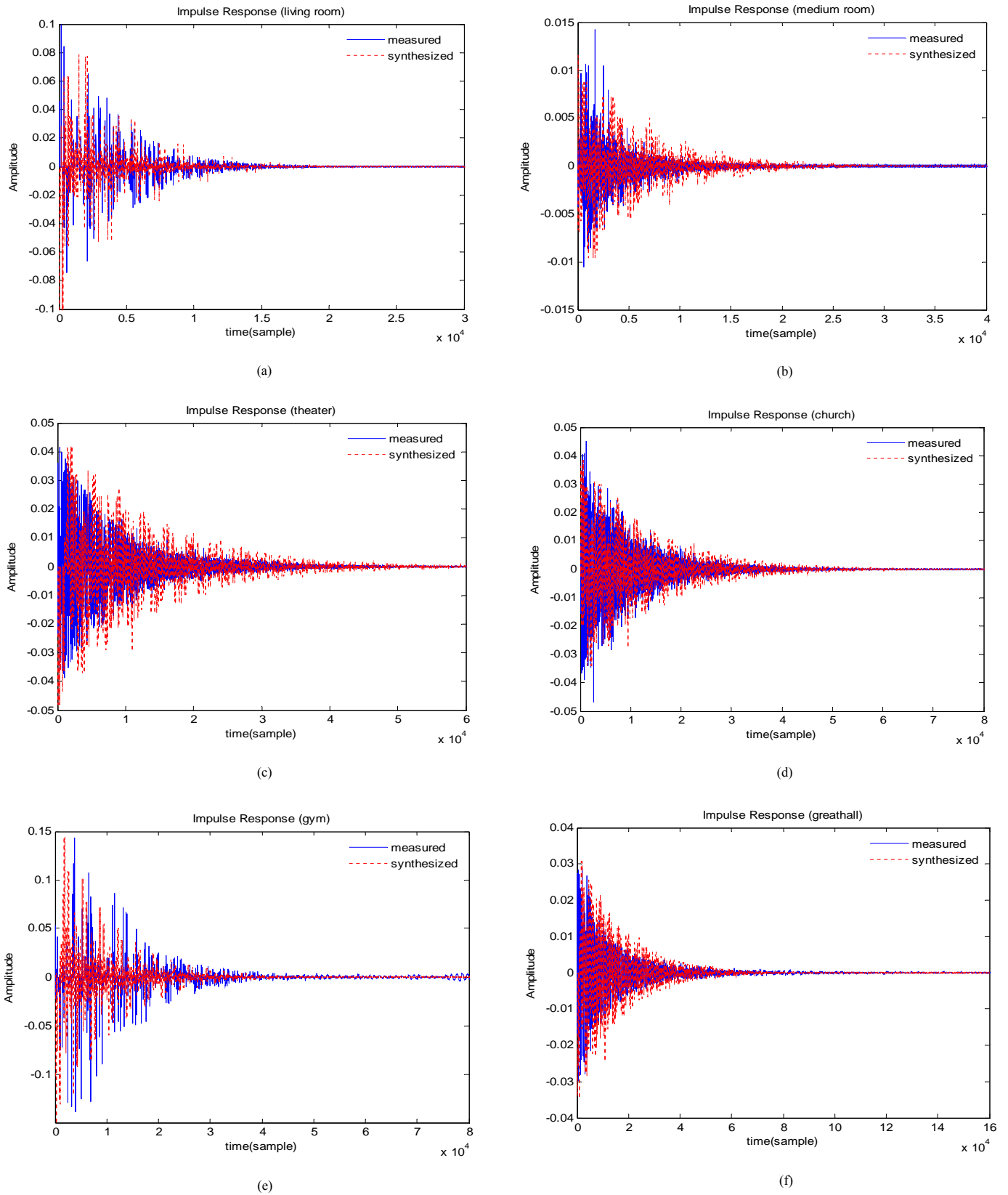
Fig. (8). The measured and the synthesized impulse responses of six types of rooms. (**a**) Living room. (**b**) Medium room. (**c**) Theater. (**d**) Church. (**e**) Gymnasium. (**f**) Great hall.
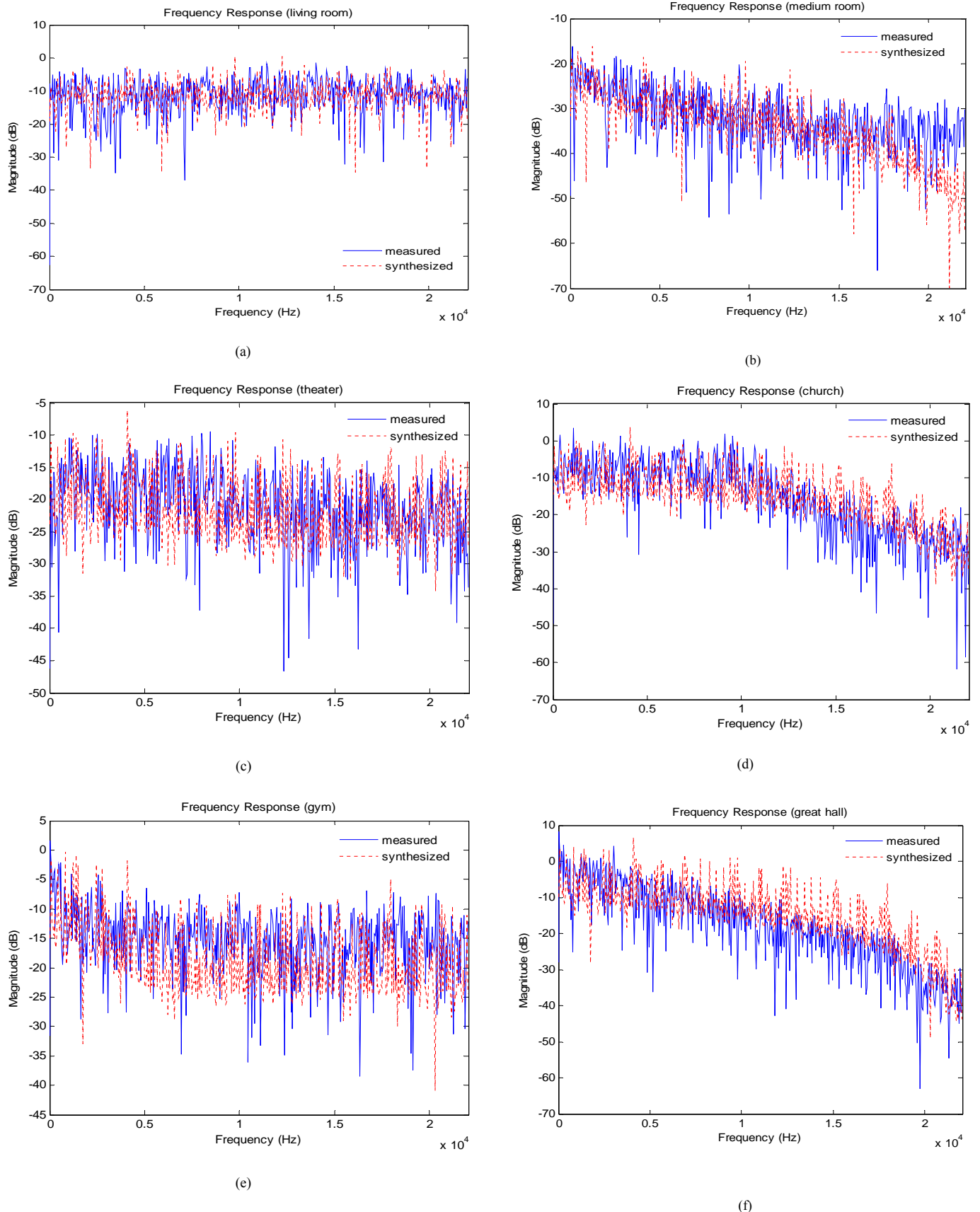
Fig. (9). The measured and the synthesized frequency responses of six types of rooms. (**a**) Living room. (**b**) Medium room. (**c**) Theater. (**d**) Church. (**e**) Gym. (**f**) Great hall.
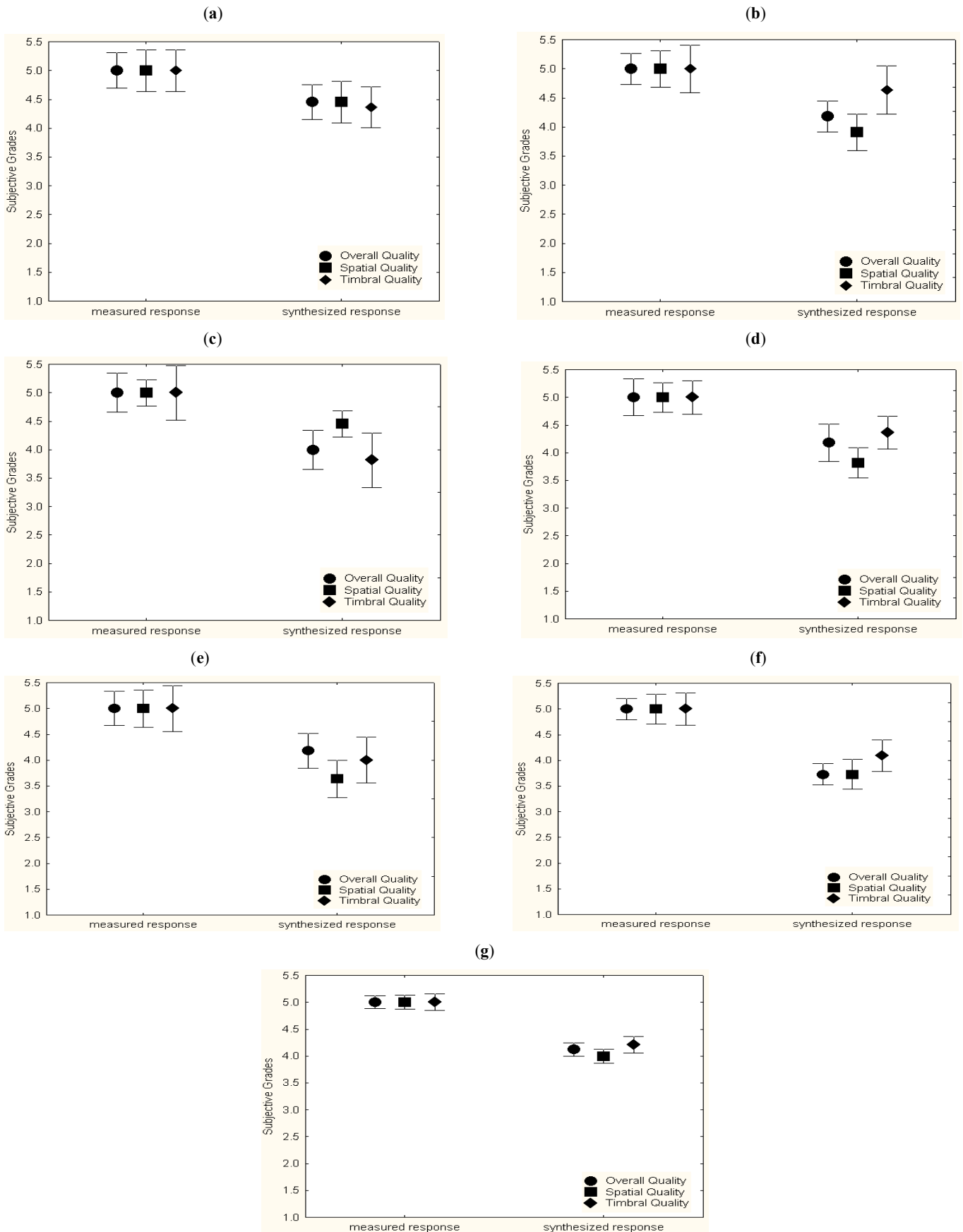
Fig. (10). The listening test results of Experiment-1. (**a**) Living room. (**b**) Medium room. (**c**) Theater. (**d**) Church. (**e**) Gymnasium. (**f**) Great hall. (**g**) Overall data.

Table 4.　The MANOVA Results of Experiment-2

Test Room Mode	Significance Value		
	Total Preference	**Existence**	**Timbral Quality**
Living Room	0.5058	0.8139	0.005234
Theater	0.5645	0.09555	0.240863
Great Hall	0.58778	0.000463	0.593482
Total Data	0.2615	0.049977	0.62343

Cases with significance value below 0.05 indicate that statistically significant difference exists among methods.

zero. Positive scores indicate that the method under investigation provides perceptual improvement compared to the known reference, whereas negative scores indicate that the method results in perceptual degradation. In order to justify the statistical significance, the scores were further processed by using the MANOVA. In this test, the hidden reference case is also taken into account when calculating the statistical significance. Cases with significance levels below a low limit, say 0.05, indicate that statistically significant difference exists among methods. The test results are plotted in Figs. (**11a-d**) and the associated significance level in the MANOVA output is summarized in Table **4**. The significance levels of the small and the medium rooms are greater than 0.05, indicating that there is no significant difference among cases, albeit early reflections have been removed. However, the significance level of the large room is below 0.05, meaning there is the statistically significant difference among methods. The listening test results revealed that the filter network approximation method outperform the other methods in modeling early reflections. In addition, the timbral quality has not been compromised after adding the early reflections. The filter network approximation is found useful in modeling early reflections, especially for large rooms. However, for small and medium rooms, modeling early reflections does not seem to be as critical.

- *Experiment-3*. Experiment-1 and experiment-2 revealed that modeling early reflections using filter network approximation seemed to be a preferred approach. However, about 23 KB of memory is required for storing the delayed samples in the full reverberator comprising the early reflection and the late reverberation modules. In order to reduce storage requirement, multirate signal processing technique is exploited, as illustrated in the block diagram of Fig. (**12**). The down-sampler and the up-sampler are connected before and after the artificial reverberator. Therefore, the memory requirement is decreased to 1/4*23 KB = 5.75 KB for a down-sampling ratio 4. Another benefit of using up/down sampling is that computational efficiency is enhanced because major processing is carried out at a lower sample rate. However, the price to pay using this approach is that the

effective bandwidth will be reduced to 6 kHz. To address the problem, a SBR technique is employed to recover the lost high-frequency signals. The SBR processor involves the use of the following nonlinear transformation to generate harmonics in high frequencies [8]:

$$y = c_1 \tanh(c_2 x), \tag{13}$$

where x and y are input and output signals, respectively, and c_1 and c_2 are set to be 0.828 3, respectively. The third listening test is carried out according to the MUSHRA procedure. There are three cases to compare in this test. In the first case, the preceding artificial reverberator was used as the reference. The artificial reverberator with up/down sampling is examined in the second case, denoted as "bandlimited". The third case is the bandlimited reverberator followed by SBR processing. The anchor is a mono signal which is the average of the two-channel stereo signal, filtered by a 128-tap FIR lowpass filter of 3 kHz cut-off. The order of cases is randomly arranged. The grading scale ranges from -3 to +3. Six types of rooms are investigated in this test. Five subjective indices including *total preference*, *spaciousness*, *artifacts*, *fullness* and *brightness* were employed in this listening test. In this test, the hidden reference is also taken into account when calculating the statistical significance. The listening test results are plotted in Fig. (**13a-g**) and the significance level in the MANOVA output is summarized in Table **5**. Although the SBR reverberator attained higher grades in *brightness* than the bandlimited method, the timbral quality seems to have been compromised using the SBR approach. There is still room for improvement with regard to how to eliminate the artifacts introduced by SBR.

5. CONCLUSIONS

Methods have been presented for reverberation synthesis that makes use of filter network approximation based on comb/nested allpass filters. To facilitate the tuning of artificial reverberators, a systematic and efficient method on the basis of the GA procedure has been developed to find the optimal parameters of reverberators. The optimization procedure hinges on several characteristics pertaining to natural room reverberations. Up/down sampling technique is exploited to reduce memory storage requirement and to enhance the processing efficiency. In addition, SBR technique is employed to recover the lost high frequency signals after up/down sampling.

The reverberators were compared *via* subjective listening experiments. The results revealed that the reverberator using network approximation is capable of delivering the most natural-sounding reverberation. A complete artificial reverberator should include both early reflection and late reverberation modules, in particular for simulating large rooms. On the other hand, the third subjective listening test also
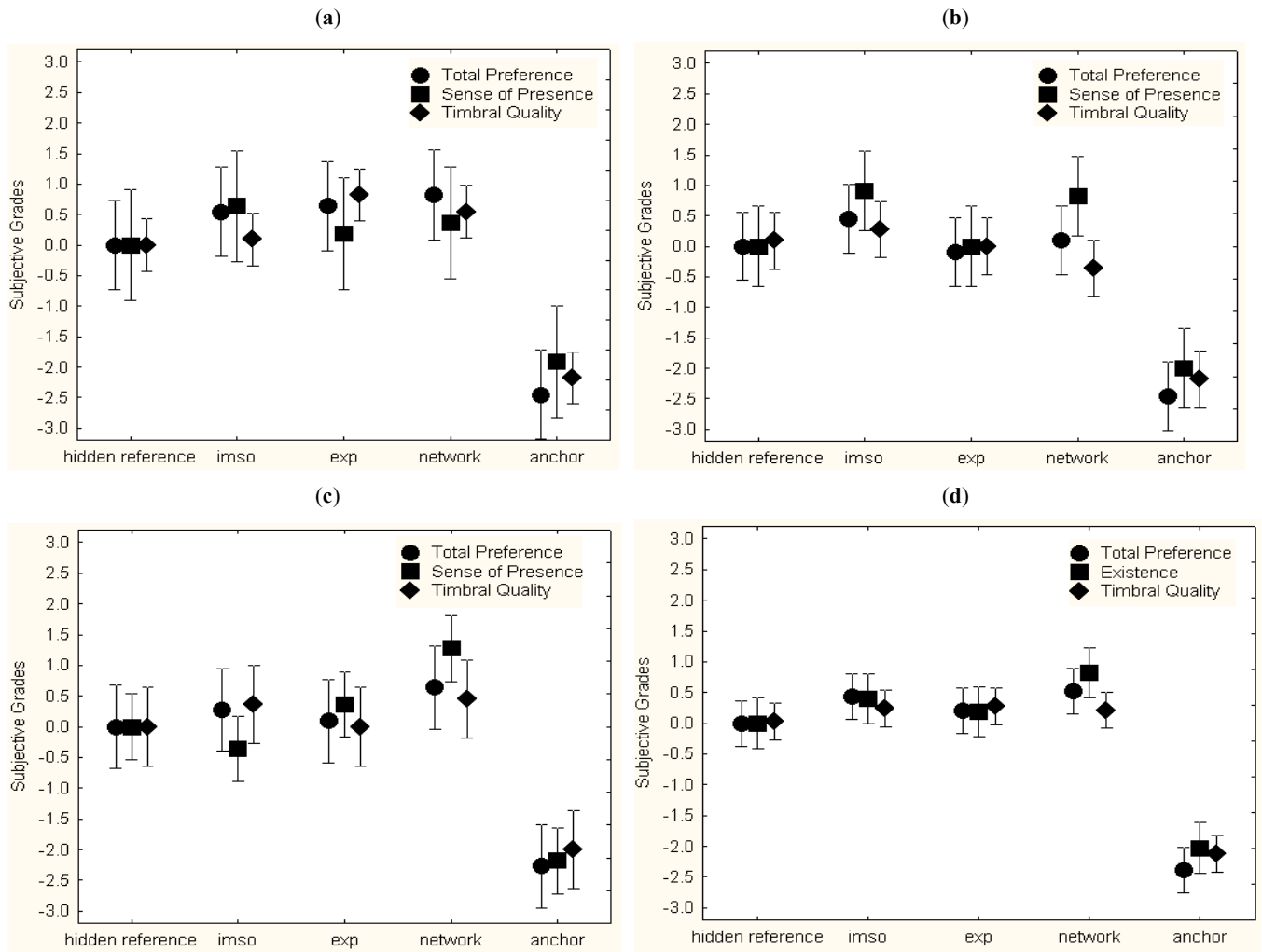
(**a**)

(**b**)

(**c**)

(**d**)

Fig. (11). The listening test results of Experiment-2. (imso = image method, exp = exponential truncation method, network = filter network approximation method.) (**a**) Living room. (**b**) Theater. (**c**) Great hall. (**d**) Overall data.
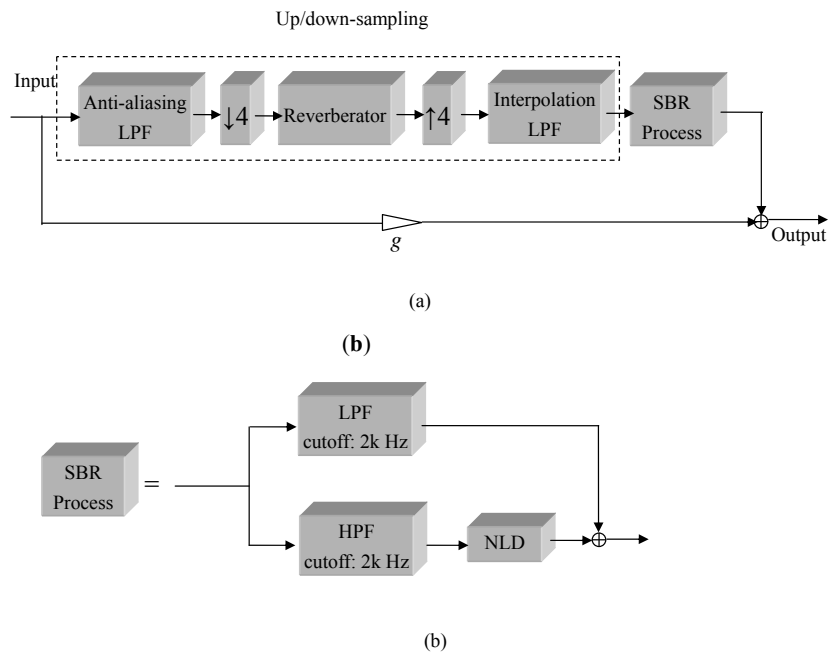
(a)

(**b**)

(b)

Fig. (12). The block diagram of the reverberator using up/down sampling and SBR. (**a**) The overall procedure. (**b**) The block diagram of SBR.
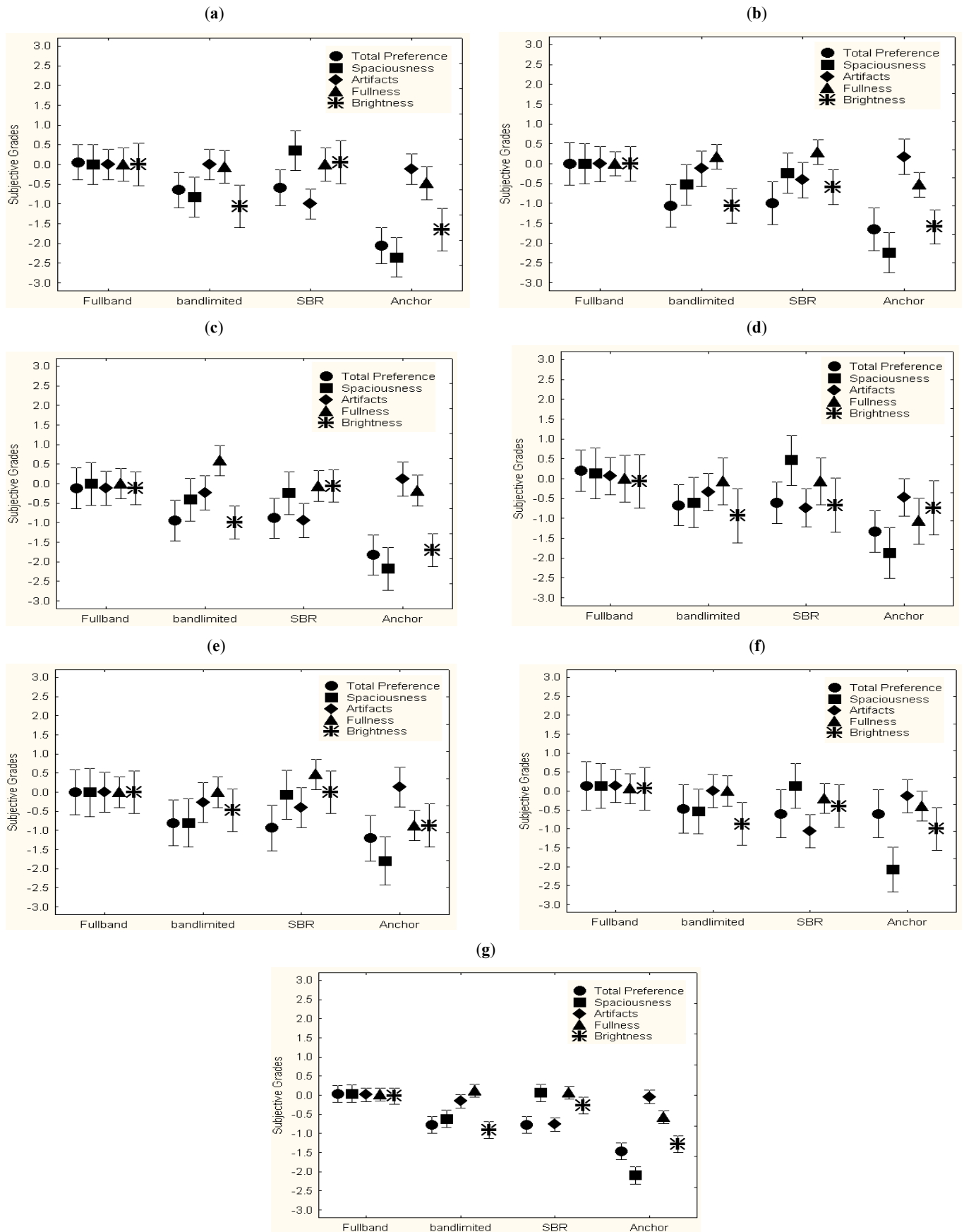
Fig. (13). The listening test results of Experiment-3. (**a**) Living room. (**b**) Medium room. (**c**) Theater. (**d**) Church. (**e**) Gymnasium. (**f**) Great hall. (**g**) Overall data.

Table 5. **The MANOVA Results of Experiment-3**

Test Room Mode	Significance Value				
	Total Preference	**Spaciousness**	**Artifacts**	**Fullness**	**Brightness**
Living Room	0.06385	0.0051	0.000047	0.96593	0.005649
Medium Room	0.000237	0.224906	0.346596	0.37776	0.002241
Theater	0.027584	0.511201	0.009839	0.05136	0.00226
Church	0.01466	0.056863	0.0329	0.98146	0.1008
Gym	0.003935	0.1099	0.2554	0.1346	0.2688
Great Hall	0.0976	0.18545	0.00024	0.6228	0.0165
Total Data	0.00006	0.003573	0.00001	0.0252	0.00001

Cases with significance value below 0.05 indicate that statistically significant difference exists among methods.

Indicated that the SBR processing introduced artifacts. There is still room for improvement with regard to how to eliminate the artifacts introduced by SBR.

Although the proposed technique is effective in rendering room effects with remarkable realism and moderate computation complexity, some limitations should be mentioned. First, it was found during the listening tests that ringing problems may occur for some music involving percussion instruments, which deserves further study. Second, the procedure of GA optimization is still somewhat time-consuming. More efficient methods should be sought in the future for finding optimal parameters of filter networks.

ACKNOWLEDGEMENT

REFERENCES

[1] Savioja L. Modeling techniques for virtual acoustics. Doctorate thesis, Espoo: Finland 1999.
[2] Schroeder MR. Natural sounding artificial reverberation. J Acoust Soc Am 1961; 10(3): 219-33.
[3] Gerzon MA. Unitary (energy preserving) multichannel networks with feedback. Electron Lett 1976; 12(11): 278-9.
[4] Vaidyanathan PP. Multirate systems and filter banks. Prentice-Hall, Englewood Cliffs: NJ 1993.
[5] Stautner J, Puckette M. Designing multichannel reverberators. Comput Music J 1982; 6(1): 52-65.
[6] Jot JM, Dahl L. A reverberator based on absorbent all-pass filters. Creative Advanced Technology Center, Scotts Valley: CA 2000.
[7] Bai MR, Bai G. Optimal design and synthesis of reverberators with a fuzzy user interface for spatial audio. J Audio Eng Soc 2005; 53: 812-25.
[8] Aarts RM, Larsen E. Audio bandwidth extension: application of psychoacoustic, signal processing and loudspeaker design. Wiley, Chichester: UK 2004.
[9] Oppenheim AV, Schafer RW. Digital signal processing. Prentice-Hall, Englewood Cliffs: NJ 1975.
[10] Kuttruff H. Room acoustics. Elsevier Science Publishing, New York 1991.
[11] Sabine WC. Reverberation in acoustics: historical and philosophical development, Lindsay RB, Ed. Dowden, Hutchinson, Ross, Stroudsburg: PA 1972.
[12] Schroeder MR. New method of measuring reverberation time. J Acoust Soc Am 1965; 37: 409-12.
[13] Moorer JA, Chauveau A, Abbott C, Eastty P, Lawson J. The 4C machine. Comput Music J 1979; 3: 244-50.
[14] Gardner WG. The virtual acoustic room. Master's thesis, MIT Media Lab, Cambridge: MA 1992.
[15] Lin CT, Lee GCS. Neural fuzzy systems. Prentice-Hall, Eaglewood Cliffs: NJ 1999.
[16] ITU-R Recommendation BS.1116-1. Methods for the subjective assessment of small impairments in audio system including multichannel sound systems. International Telecommunications Union, Geneva: Switzerland 1994.
[17] ITU-R Recommendation BS.1534-1. Method for the subjective assessment of intermediate sound quality (MUSHRA). International Telecommunications Union, Geneva: Switzerland 2001.
[18] ITU-R Recommendation BS.1284-1. General methods for the subjective assessment of sound quality. International Telecommunications Union, Geneva: Switzerland 1997.