Optimal Design and Synthesis of Reverberators with a Fuzzy User Interface for Spatial Audio*

MINGSIAN R. BAI, *AES Member,* **AND GANYUAN BAI**
(msbai@mail.nctu.edu.tw) (Daniel.Bai@qsitw.com)

Department of Mechanical Engineering, National Chiao-Tung University, Hsin-Chu 300, Taiwan

Reverberators are key elements in three-dimensional spatial audio reproduction. The richness and spaciousness of reproduced sound can be enhanced with reverberators. Infinite impulse response (IIR) filters such as all-pass or comb filters are commonly used in reverberator design. An inadequate choice of the filter parameters in these reverberators often results in audible artifacts such as metallic and ringing sound. To minimize the effort of trial and error in parameter tuning when designing reverberators, an automatic search procedure based on genetic algorithms (GA) is presented. The architecture of the present reverberator consists of a finite impulse response (FIR) early reflection module and an IIR late reverberation module. To facilitate the choice of filter parameters according to the room modes specified by the user, an intelligent user interface is developed on the basis of fuzzy logic. Subjective listening tests were carried out to assess the performance of the proposed reverberators. The results indicate that, when compared to conventional reverberators, the presented reverberator is capable of delivering natural sounding reverberation.

0 INTRODUCTION

Reverberation results from the endless reflections of sound waves in a bounded space. Reverberation carries the cues of human perception of the acoustic environment, such as room size and absorptivity of the boundary. In audio reproduction and mixing, reverberation is necessary to enrich the dry recording, to make it sound more natural and spatial than the original. In some ill-conditioned listening environments, such as a car cabin, artificial reverberation is also useful in mitigating the coloration problems caused by a small space during surround presentation. Apart from the loudspeaker reproduction mentioned, reverberation can also be used in headphones to externalize the sound images so that listening discomfort can be minimized.

Instead of direct filtering using highly complex and diffuse room responses, which was considered computationally expensive and even intractable for early hardware, many artificial reverberators have been suggested in the past. Pioneering work on artificial reverberators can be attributed to Schroeder [1] and Moorer [2], who proposed several algorithms based on all-pass or comb filter networks for synthesizing room responses. Gardner [3] and Dahl and Jot [4] developed novel algorithms to produce reverberation that is natural sounding, free of tonal coloration, with high echo density and proper reverberation time. These infinite impulse response (IIR)–based reverberators can be implemented with moderate complexity. The feedback delay network (FDN) model suggested by Dahl and Jot [4] can be regarded as a generalized version of Schroeder's parallel comb filter. The network requires a unitary feedback matrix to maintain stability. This network is capable of producing reverberations with much higher echo densities than the parallel comb filters.

A common problem with artificial reverberators is that an inadequate choice of filter parameters could result in artifacts such as metallic and ringing sounds. It is then desirable to find a systematic and efficient way in the search of optimal parameters for reverberator filters. In this paper an automatic search procedure based on genetic algorithms (GA) is presented to minimize the effort of trial and error in designing reverberators. Early reflections are created using the image method [5], [6], whereas late reverberations are synthesized using the nested all-pass filter [7] and the comb filter. The nested all-pass filter produces reverberations with an echo density that increases with time, resembling realistic room responses. Comb filters are cascaded to the nested all-pass filters to further enhance the late reverberations. The optimum relative positions of source and receiver to produce the maximum spaciousness are also examined. To minimize the effort in adjusting the parameters of all-pass or comb filters, we

*Manuscript received 2004 August 30; revised 2005 July 12.

developed a procedure for searching the network parameters by using GA [8]. The GA technique is well suited for parameter optimization problems with many local minima. To facilitate the user's choice of system parameters for specified room modes, an intelligent user interface is also developed on the basis of fuzzy logic [8]. Five room modes can be selected with the aid of a user-friendly graphic interface.

Subjective listening tests are conducted to assess the performance of the proposed reverberators. The performance indices chosen for the test are spaciousness, reverberation, warmth, clarity, naturalness, ringing, and pleasantness. The results reveal that the proposed systems are capable of rendering natural sounding reverberations with minimal computational cost.

1 SYNTHESIS OF REVERBERATION

1.1 Attributes of Room Responses

A room response can in general be divided into two parts, as shown in Fig. 1. The early reflections, lasting approximately 80 ms, are composed of the direct sound accompanied by discrete reflections. Early reflections are dependent on the room geometry as well as the relative positions of the source and the receiver. As compared to the early reflections, the late reverberations are less structural and contain extremely dense echoes with exponentially decreasing amplitude. Both types of response are highly complex in typical diffuse sound fields and are difficult, if not impossible, to be modeled using modal analysis.

Instead of precise modal analysis, a statistical description is a more practical approach to tackle such a problem. Several statistics of the room response are relevant in the design of reverberators. First the echo density is defined in the time domain as the number of echoes per second in a room response. Kuttruff [6] derived the formula using a sphere model to estimate the number of echoes within time t,

$$N_t = \frac{4\pi(ct)^3}{3V} \tag{1}$$

where c is the speed of sound and V is the volume of the room. Differentiating Eq. (1) with respect to time t gives the echo density,

$$D_t \triangleq \frac{\mathrm{d}N_t}{\mathrm{d}t} = \frac{4\pi c^3}{V} t^2. \tag{2}$$

Note that the echo density is proportional to the square of time.

The second statistic is the modal density defined in the frequency domain as the number of normal modes per hertz. The number of normal modes N_f below frequency f, independent of the room geometry, can be estimated using the following formula [6]:

$$N_f = \frac{4\pi V}{c^3} f^3 + \frac{\pi S}{4c^2} f^2 + \frac{L}{8c} f \tag{3}$$

where S is the area of all walls and L is the sum of all edge lengths of the room. Differentiating Eq. (3) with respect to frequency leads to the expression of modal density,

$$D_f \triangleq \frac{\mathrm{d}N_f}{\mathrm{d}f} \approx \frac{4\pi V}{c^3} f^2. \tag{4}$$

Hence the modal density of a room grows in proportion to the square of the frequency.

Third the reverberation time T_{60} is the time required for the sound pressure level to decay by 60 dB after a steady-state source is switched off. The reverberation time can be estimated by the Sabine formula [9],

$$T_{60} = \frac{0.163V}{\sum_i a_i S_i} \tag{5}$$

where V is the volume of the room and S_i and a_i are the surface area and the associated absorption coefficient. The 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, the reverberation time decreases as the frequency increases. A method for estimating the reverberation time from the room impulse response was proposed by Schroeder [1]. His method requires finding the time for which the level of the energy decay curve (EDC) decays by 60 dB. Precisely, EDC is defined as

$$\mathrm{EDC}(t) = \frac{\int_t^\infty h^2(\tau)\,\mathrm{d}\tau}{\int_0^\infty h^2(\tau)\,\mathrm{d}\tau} \tag{6}$$

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

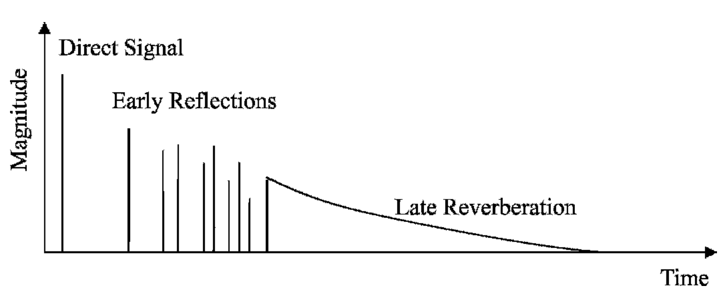

Fig. 1. Typical room response.

1.2 Image-Source Method

There are two commonly used methods for predicting the sound field in an enclosure: the ray-tracing method and the image-source method [5], [10]. The ray-tracing method is based on the assumption that sound behaves like rays in high frequencies. The distribution of the sound field is determined by keeping track of the frequency that passes each position in the space. The image-source method models the reflections from the room boundary as the sound waves emitted by virtual sources. Although image sources can be created indefinitely, only early reflections are computed in practice by utilizing low-order image sources. This method is particularly suited to the calculation of the responses of regularly shaped rooms. The image method is used here to calculate the early reflections of room responses. For a rectangular room the number of image sources required in the calculation of the nth-order reflection can be estimated as

$$N_n = 4n^2 + 2. \tag{7}$$

The number of image sources grows with the square of n. Fig. 2 illustrates the impulse response calculated using the image-source method up to the 30th-order reflection. Due to the limitations of computing power, the maximum order of reflections in calculating the early reflections is restricted to 6, corresponding to 376 image sources. Fig. 3 shows an example of the primary source and the first-order image sources distributed in a rectangular room. The impulse response of the room is constructed by recording all arrivals of the impulses from the primary source and the image sources as well. The impulse response due to the image sources is an attenuated and delayed version of the response due to the primary source. The resulting sum of impulses serves as the coefficients of the FIR filter for the early reflections of the room response.

1.3 Comb Filter and Nested All-Pass Filter Network

In addition to the early reflections the late reverberations are required to complete the room response. Referring to Fig. 4, the method we used for synthesizing the late reverberations is a comb filter and all-pass filter network [11]. The parallel comb filters serve to increase the modal density and the echo density, whereas series all-pass filters serve to further increase the echo density of reverberation. The filter parameters are dependent on the modal density and the echo density associated with a particular room. It can be shown that [12]

$$D_f = \frac{1}{F_s}\sum_{i=1}^{N} m_i = \frac{N\rho}{F_s} \tag{8}$$

$$D_t = \sum_{i=1}^{N} \frac{F_s}{m_i} \approx \frac{NF_s}{\rho} \tag{9}$$

where F_s is the sampling rate, N is the number of comb filters, m_i is the delay of the ith comb filter, and ρ is the average delay length (in samples). Combining Eqs. (8) and (9) leads to the number of comb filters [12],

$$N = \sqrt{D_f D_t}. \tag{10}$$

On the other hand, it can be shown that the reverberation time of a comb filter satisfies the following relation:

$$\frac{20\log_{10}(g_i)}{m_i F_s} = \frac{-60}{T_{60}}. \tag{11}$$

For a desired reverberation time we can choose an appropriate delay m_i and the feedback gain g_i to trade modal density for echo density. From Eq. (11) the feedback gains

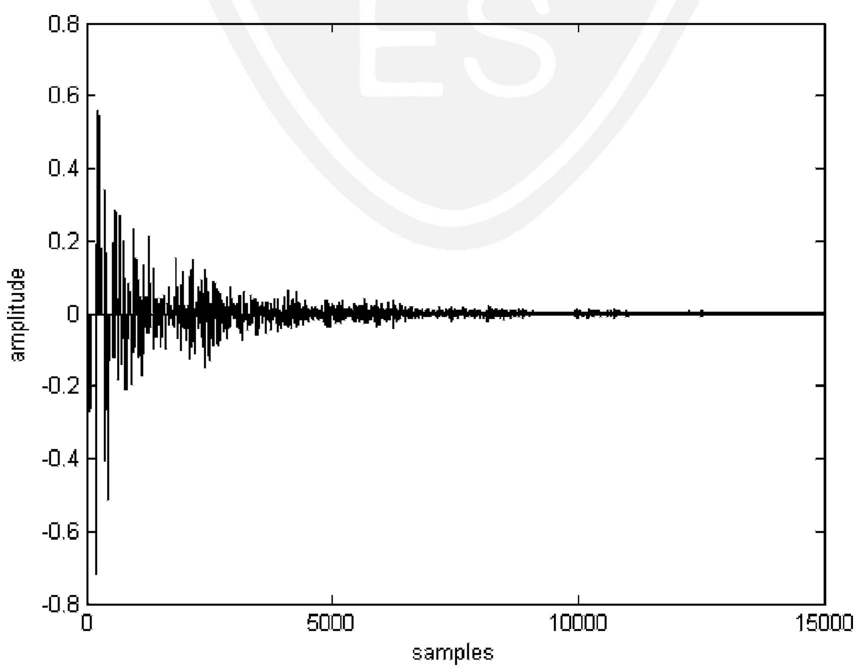

Fig. 2. Impulse response obtained by using image method with 30th-order reflections. Room dimension 10 m × 8 m × 3 m, absorption coefficient 0.8.

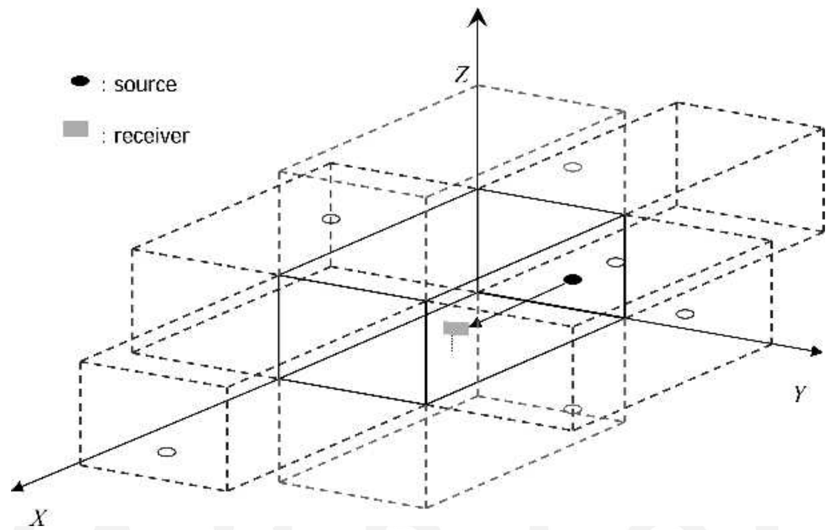

Fig. 3. Schematic of primary source and associated image sources (up to first-order reflection) corresponding to a rectangular room.

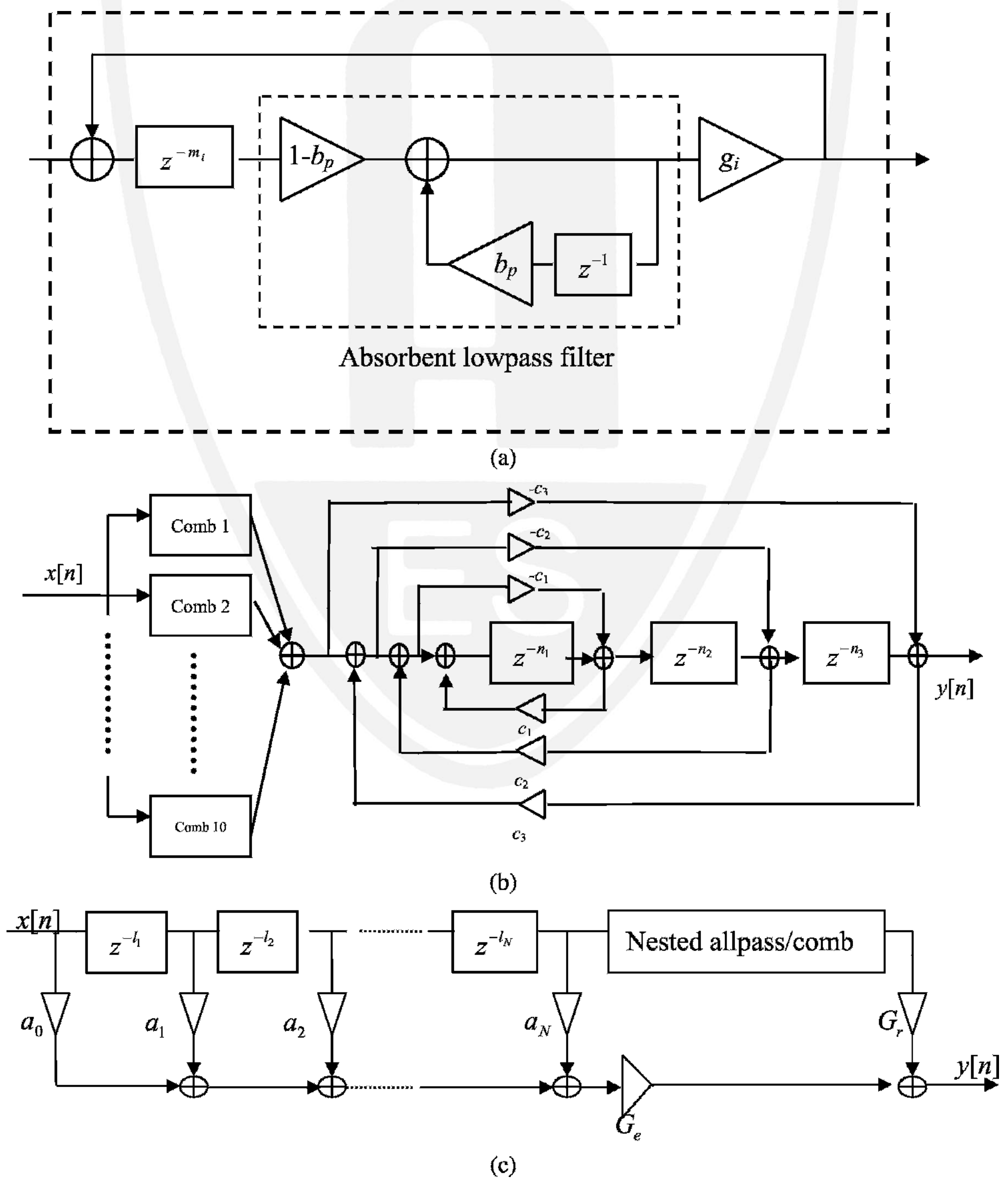

Fig. 4. IIR filter structures of reverberator. (a) Comb filter. b_{p}—gain of absorbent low-pass filter; g_i—gain of comb filter. (b) Ten parallel comb filters and three-layer nested all-pass filters. (c) Nested all-pass/comb reverberator in conjunction with early reflection module obtained using image method.

g_i of the comb filters can be chosen according to a desired reverberation time,

$$g_i = 10^{-3m_i F_s / T_{60}}. \quad (12)$$

The absorbent filter in Fig. 4(a) is a first-order low-pass filter whose parameters are determined by the ratio α of the reverberation times at the Nyquist frequency and direct current [4], [13]. The gain of the absorbent filter b_p is given as [12]

$$b_p = 1 - \frac{2}{1 + g^{1-1/\alpha}}. \quad (13)$$

The absorbent filter accounts for the high-frequency absorption in a room. The network structure comprising ten parallel comb filters and three-layer nested all-pass filters is shown in Fig. 4(b). There is an absorbent filter in the delay line of each comb filter. Fig. 4(c) shows the complete network structure of the nested all-pass/comb reverberator, combined with the early reflections designed using the image-source method. In the figure the parameters a_i and l_i ($i = 1, \ldots, n$) are the gains and delays of the early reflection module. The parameters G_e and G_r are the early reflection gain and the late reverberation gain, respectively.

Although recursive reverberation filters are advantageous in terms of computation cost, a word of caution is appropriate. Because the all-pass/comb networks are by nature recursive IIR filters, the numerical problems such as quantization and stability should be dealt with more carefully than the nonrecursive finite impulse response (FIR) filters during the implementation phase. Scaling is required to deal with these problems on some occasions.

2 OPTIMAL DESIGN OF REVERBERATORS

A further refinement is possible for the reverberator mentioned in the preceding section. In order to minimize the effort of trial and error, a systematic approach is presented to design the early reflection and late reverberation modules with the aid of optimization procedures.

2.1 Optimization of Early Reflections

In this section a measure relevant to the human perception of early reflections and spatial impression is introduced. The early lateral energy fraction (ELEF) proposed by Ouis [14] is defined as

$$\mathrm{ELEF} = \frac{\int_{5\mathrm{ms}}^{80\mathrm{ms}} P_L^2(t)\,\mathrm{d}t}{\int_{0}^{80\mathrm{ms}} P^2(t)\,\mathrm{d}t} \quad (14)$$

where P is the sound pressure, $P_L^2(t) = P^2(t)|\cos\theta|$, and θ is the angle subtended by the incident ray and the normal to the median plane of the listener's head. This index can be viewed as the ratio of the early lateral sound energy to the total sound energy during the first 80 ms of the room impulse response. This index has been found to be strongly correlated with the perception of spaciousness. It is used as the objective function for constructing early reflections using the image method. More precisely, the goal of optimization is to find the locations of the source and the receiver that maximize ELEF.

Consider a rectangular room of medium size with dimensions of 15 m (width) by 8 m (length) by 3 m (height). Fig. 5(a) shows the geometry of the room, where the sound source is located at (0.5, 0.5, 2). The ELEF is calculated for different receiver positions distributed uniformly in the room. Fig. 5(b) shows the contour of ELEF plotted versus the location of the receiver (in Cartesian coordinates). It is evident that the spaciousness increases as the receiver moves away from the source to the kitty-corner position. This source and receiver configuration is thus adopted in the following simulation.

2.2 Optimization of Late Reverberations Using the Genetic Algorithm

In the present work the parallel comb-nested all-pass filter network is used to model the late reverberations of the room response. The design parameters in the network consist of comb filter delays, comb filter attenuation gains, all-pass filter delays, all-pass filter attenuation gains, and the low-pass filter coefficient *a*. Obviously it would be rather time-consuming to adjust these eighteen parameters manually. To minimize the effort of trial and error, a systematic and efficient method based on the GA is presented.

In the GA method all parameters are encoded into binary strings called chromosomes [8]. The resolution of a parameter is dependent on the amount of bits per string and the search domain. In GA optimization the objective function is also termed the fitness function. A chromosome with high fitness has higher probability of surviving the natural selection and reproducing offspring in the next generation. Crossover enables exchanging genes in the chromosomes via probabilistic decisions in the mating pool. Mutation is required to prevent the problem of the genes becoming increasingly homogeneous, resulting in premature convergence. The mutation point is randomly chosen. Mutation is then carried out by alternating the gene from 0 to 1 and vice versa.

The flowchart of the optimization procedure is shown in Fig. 6. The GA procedure consists of four steps. In the first step three delays and three gains of the three-layered nested all-pass filters n_i and c_i, $i = 1, 2, 3$, are optimized. The upper and lower limits for the delays are chosen to be 1000 and 50, respectively. The upper and lower limits for the gains are chosen to be 1 and 0.1, respectively. The objective of optimization in this step is to attain high echo density D_t and high impulse energy E_n,

$$E_n = \int_{t=0}^{\infty} h^2(t)\,\mathrm{d}t \quad (15)$$

where $h(t)$ is the impulse response of reverberation. The fitness function in the GA is chosen as follows:

$$F_1(\theta_1) = D_t(\theta_1) + wE_n(\theta_1) \quad (16)$$

where $\theta_1 = [n_1 \quad n_2 \quad n_3 \quad c_1 \quad c_2 \quad c_3$ is the parameter vector, or the *chromosome* in the GA, and w is the weighting factor.

In the second step we wish to optimize those ten delays of the parallel comb filters m_i, $i = 1, 2, \ldots, 10$. The upper and lower limits of the delay are set to be 3528 samples (80-ms duration with a sampling rate of 44.1 kHz) and 441 samples (10-ms duration with a sampling rate of 44.1 kHz), respectively. The objective of optimization in this step is to find the best chromosome that attains the highest echo density and modal density. The modal density is estimated by counting the number of poles in the pole–zero map. The fitness function in the second step is defined as

$$F_2(\theta_2) = D_t(\theta_2) \times D_f(\theta_2) \tag{17}$$

where $\theta_2 = [m_1 \quad m_2 \quad \ldots \quad m_{10}]$ is the vector of the comb filter delays. Because modal density generally decreases as the echo density increases in synthesizing reverberation using comb filters, the product form is used in the fitness function to trade off these two parameters.

In the third step the feedback gain g of the comb filters is determined according to the reverberation time T_{60} of the room. The reverberation time must be specified prior to the GA search for the optimal gain of the comb filters. The procedure will continue until the T_{60} constraint is met.

In the fourth step the parameter α is adjusted using GA according to the prespecified room response templates, such as church, living room, or auditorium. For each room mode the best chromosome of the parameter α is calculated to match the frequency responses of the template and that synthesized by the network. The fitness function at this stage is given as

$$F_4 = |P(f) - \hat{P}(f)| \tag{18}$$

where $P(f)$ and $\hat{P}(f)$ are the desired frequency response template and the synthesized response, respectively.

Overall there are 40 populations formed randomly for each parameter. Each population consists of eight chromosomes. The upper and lower limits of the parameters are summarized in Table 1. In the GA procedure the crossover rate and the mutation rate are selected to be 0.85 and 0.008, respectively. After 100 generations the GA procedure converged to the optimal parameters shown in Table 2. The learning curve in Fig. 7 shows the convergence behavior of the first four comb filter delays during the GA iteration.

3 FUZZY USER INTERFACE FOR REVERBERATORS

3.1 Fuzzy Interface System

To facilitate synthesizing the room responses, a user-friendly interface based on fuzzy logic [8] is developed. The architecture of the fuzzy user interface is shown in

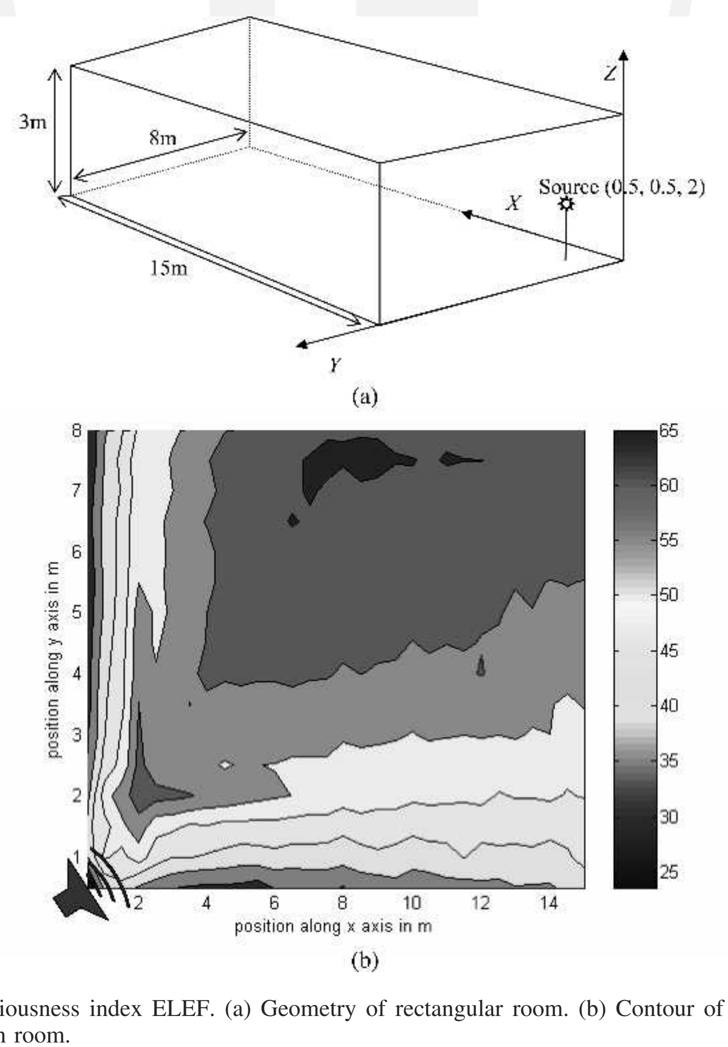

Fig. 5. Computation of spaciousness index ELEF. (a) Geometry of rectangular room. (b) Contour of ELEF in percent plotted for different receiver positions in room.

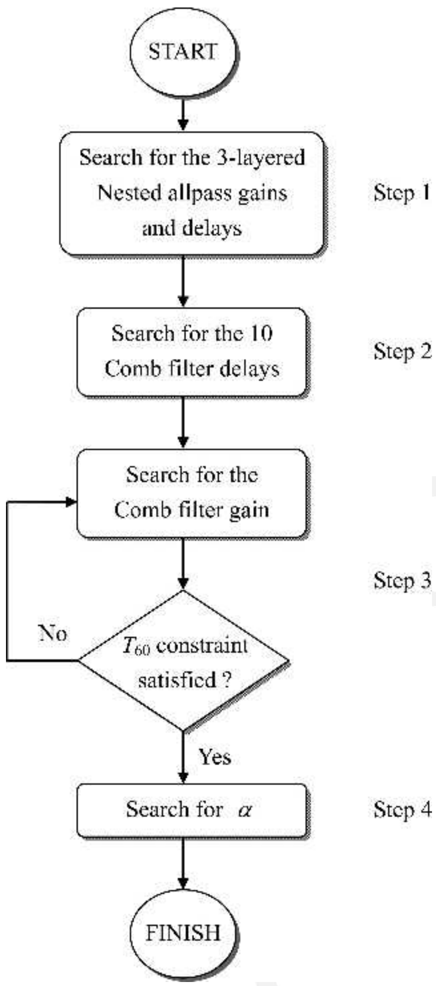

Fig. 6. Flowchart of optimization procedure of GA.

Fig. 8. The subjective scales of the fuzzy logic are chosen based on objective acoustic parameters related to the reverberation characteristics. The fuzzy user interface operates in two stages. In the first stage the room mode R is the only fuzzy variable that takes on five different values: 0.2, 0.4, 0.6, 0.8, and 1.0 for LR (living room), SC (small club), CH (church), LA (large auditorium), and GY (gymnasium), respectively. Fuzzy rules of the form (A_i, B_i) represent the linguistic rule "if X is A_i, then Y is B_i." Four fuzzy membership functions, S (small), M (medium), L (large), and V (very large), are assigned to different input levels. The following Gaussian function is adopted as the membership function for the subjective indices shown in Fig. 9,

$$\mathrm{gauss}(x,a,b) = \exp\left[\frac{-(x-b)^2}{2a^2}\right] \tag{19}$$

where x, a, and b represent the input parameter, the standard deviation, and the mean, respectively. The subjective indices include room size, diffuseness, warmth, clarity, and reverberation.

In the second stage these five subjective indices are used as the fuzzy inputs and eight system parameters are used as the fuzzy outputs. These fuzzy outputs are denoted by Dim (room dimension), Comb_d (comb filter delay), Comb_g (comb filter gain), Apd (all-pass filter delay), Alpha (high-frequency ratio), Fc (cutoff frequency of early reflections), Ge (early reflection gain), and Gr (late reverberation gain). The range space in which the membership function of a fuzzy set is defined is termed the universe of discourse [8]. Dim ranges from 0 to 50, corresponding to

Table 1. Upper and lower limits of each optimized parameter.

Parameter	Upper Limit	Lower Limit
d_i (i = 1, 2, 3)	1000	50
g_i (i = 1, 2, 3)	1	0.1
M_i (i = 1, 2, . . . , 10)	3528	441
k_p (comb gain)	NA	NA
α	1	0.1
G_e	1	0
G_r	1	0

Table 2. Optimal results of system parameters obtained by using GA.

	Comb Filter		All-Pass Filter
1	$m_i = 3007$	1	$n_i = 430$ $c_i = 0.77765$
2	$m_i = 2825$	2	$n_i = 248$ $c_i = 0.88$
3	$m_i = 2753$	3	$n_i = 329$ $c_i = 0.59059$
4	$m_i = 2656$		
5	$m_i = 2499$		
6	$m_i = 2341$		
7	$m_i = 1784$		
8	$m_i = 1712$		
9	$m_i = 1482$		
10	$m_i = 441$		

four fuzzy sets: SL (short), ML (moderate), LL (long), and VL (very long). Each fuzzy set has its own membership function with the characteristic values shown in Fig. 10(a). The value of Dim determined by the fuzzy interface system corresponds to the length dimension of a rectangular room in the early reflection model. The width and height of the room are set to be 0.6 and 0.15 times the value of the length, respectively. The universe of discourse of Comb_d ranges from 0 to 3600, corresponding to four fuzzy sets with the membership functions shown in Fig. 10(b). The value of Comb_d is set to be the average of the ten comb filter delays. The universe of discourse of Comb_g ranges from 0 to 10, corresponding to four fuzzy sets with the membership functions shown in Fig. 10(c). The value of Comb_g is set to be the comb filter gain. Large (small) comb filter gain results in a long (short) filter impulse response. The universe of discourse of Apd ranges from −150 to 300, corresponding to four fuzzy sets with the membership functions shown in Fig. 10(d). In general a high degree of diffuseness implies high echo density, and hence shorter all-pass filter delay. The choice of Apd is based on the optimal setting of the all-pass filter delays determined in the GA procedure.

The subjective index Warmth is related to the proportion of low-frequency content in the reverberation. A high degree of Warmth entails a small high-frequency ratio (Alpha) and a low cutoff frequency (*Fc*). The universe of discourse of Alpha ranges from 0 to 0.9, corresponding to four fuzzy sets with the membership functions shown in Fig. 10(f). The universe of discourse Fc ranges from 6 to 16 kHz, corresponding to four fuzzy sets with the membership functions shown in Fig. 10(e).

The subjective indices Clarity and Reverb are related to the relative proportion of early reflections and late reverberations. However, these two indices are complementary to one another. A high degree of Clarity or a low degree of Reverb generally requires a high early reflection gain (Ge) and a low late reverberation gain (Gr), and vice versa. The universe of discourse of Ge and Gr ranges from 0 to 0.9, corresponding to four fuzzy sets with the membership functions shown in Fig. 10(f). The fuzzy decision rules for the inference system are summarized in Tables 3 and 4. It is noted that a weighting factor is used to trade off Clarity and Reverb in rules 13–16 of Table 4. The defuzzification is carried out by calculating the center of the area (COA),

$$Z_{\mathrm{COA}} = \frac{\sum_{j=1}^{n} \mu_C(Z_j) Z_j}{\sum_{j=1}^{n} \mu_C(Z_j)} \tag{20}$$

where n is the number of quantization levels of the output, Z_j is the value of the control output at the quantization level j, and $\mu_c\ (Z_j)$ represents its membership value in the output fuzzy set C.

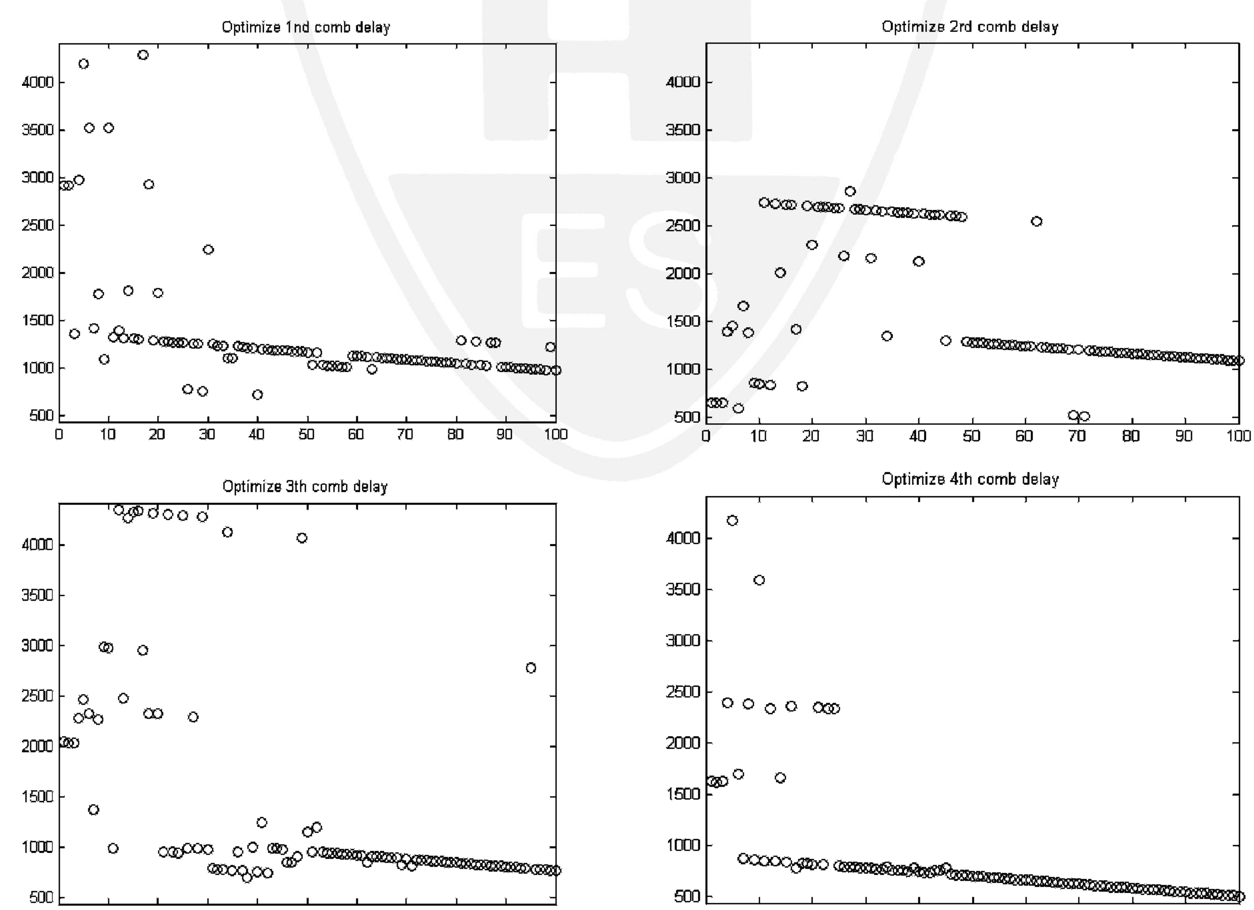

Fig. 7. Learning curve of first four comb filter delays after 100 generations in GA procedure.

3.2 Graphic User Interface

In order to facilitate the operation of the aforementioned reverberator, a graphic user interface (GUI) was developed, as shown in Fig. 11. This GUI serves as a user-friendly front end of the fuzzy system and the GA-based room response synthesizer. With the GUI it is relatively convenient for users to select the reverberator that best

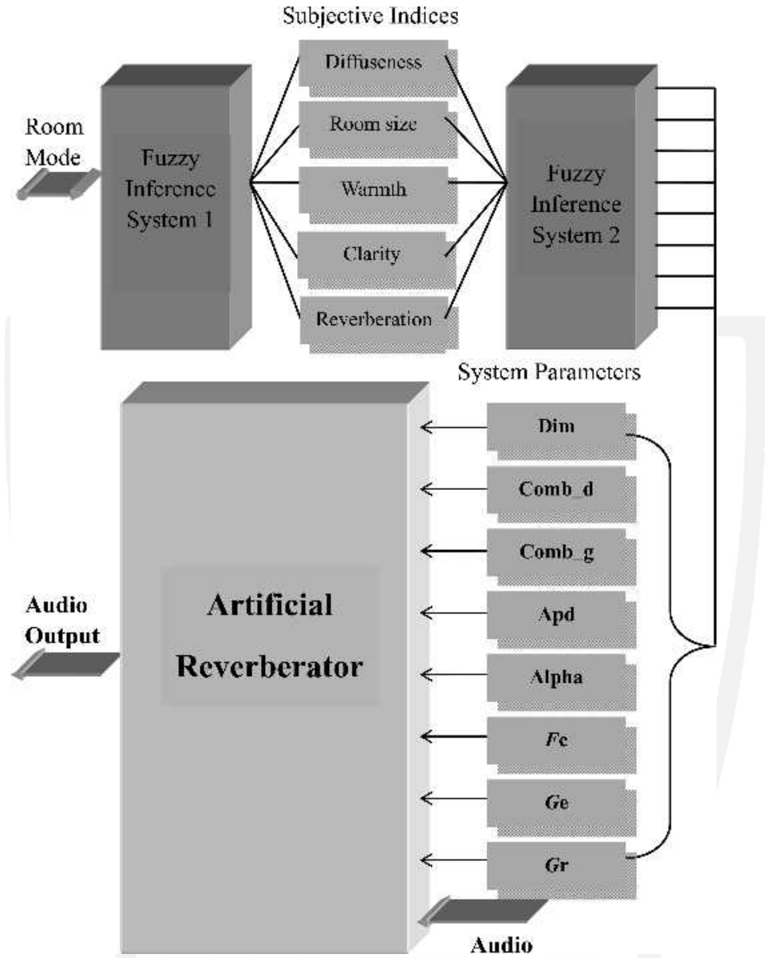

Fig. 8. Two-stage scheme of fuzzy user interface for artificial reverberator.

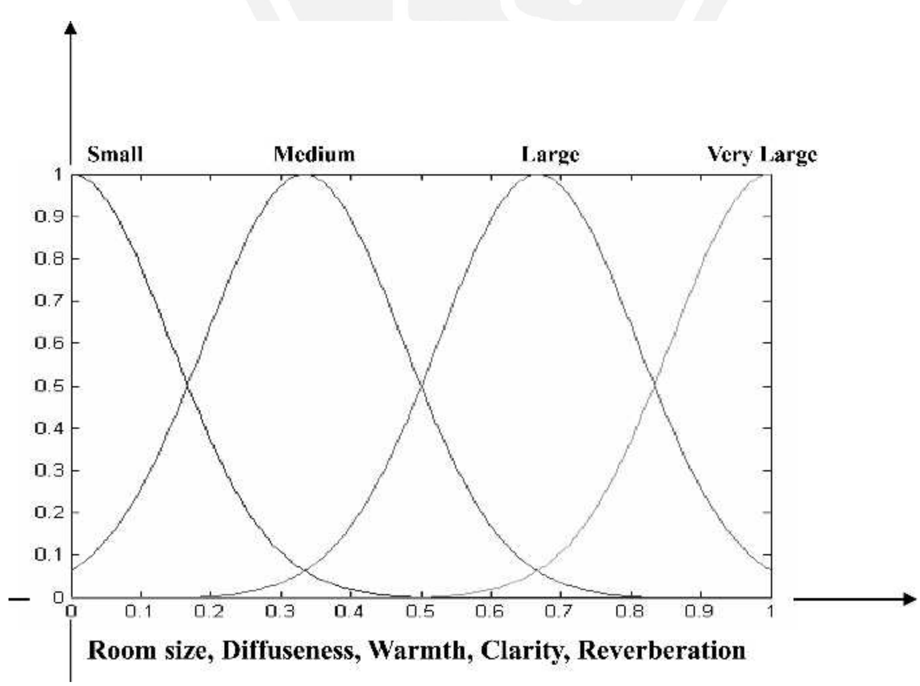

Fig. 9. Gaussian membership functions for five subjective indices: room size, diffuseness, warmth, clarity, and reverberation.

suits a particular sound venue. Five room modes, including living room, small club, church, large auditorium, and gymnasium, can be selected via the GUI. Upon pushing the Run button, the five subjective indices associated with the sound venue and the eight parameters of the artificial reverberator are set automatically. In addition, the impulse response, the frequency response, and the energy decay curve of the chosen room will be shown in a display window. The corresponding reverberation time will be shown in the bottom left-hand corner. When entering the file name of the input audio program, the output signal processed by the intended reverberator can be played back by using a headset or stereo loudspeakers in real-time fashion.

4 SUBJECTIVE LISTENING EXPERIMENTS

In order to assess the performance of the proposed reverberators, subjective listening experiments were conducted. Ten human subjects participated in the experiments. Seven subjective indices, including spaciousness, reverberation, warmth, clarity, naturalness, ringing, and pleasantness, were selected in a questionnaire for the listening tests. Seven room modes, including the optimal GA reverberator (1), five fuzzy-based reverberators (2–6), and the reverberation produced by direct convolution (7) were compared in the listening tests. It should be noted that except for the seventh reverberator, which is an FIR filter, the rest of the reverberators were all IIR all-pass or comb filters. A 1-minute passage of female speech was used as the audio input. The original signal was played prior to the signals being processed by the reverberators. Ten subjects participated in the subjective listening test. The listeners' perception in terms of the subjective indices was answered on a scale ranging from 0 to 10. The score of 5 indicates "no difference" between the processed signal and the unprocessed signal. A score greater than 5 indicates that the processed signal perceptually outperforms the unprocessed signal in terms of a particular subjective index. Conversely, a score less than 5 indicates that the unprocessed signal perceptually outperforms the processed signal in terms of a particular subjective index. Fig. 12 illustrates the average score of the subjective indices. It can be seen that all indices, except ringing, are greater than 5. To compare the overall performance of these seven reverberators, the total score is summarized in Fig. 13. As expected, the result of the direct convolution had the best performance. However, the remaining five reverberators, especially the fourth reverberator (church), performed almost equally well, as compared to the direct convolution. The second reverberator (living room), however, failed to generate significant reverberation because of the small room size, and it received the lowest score.

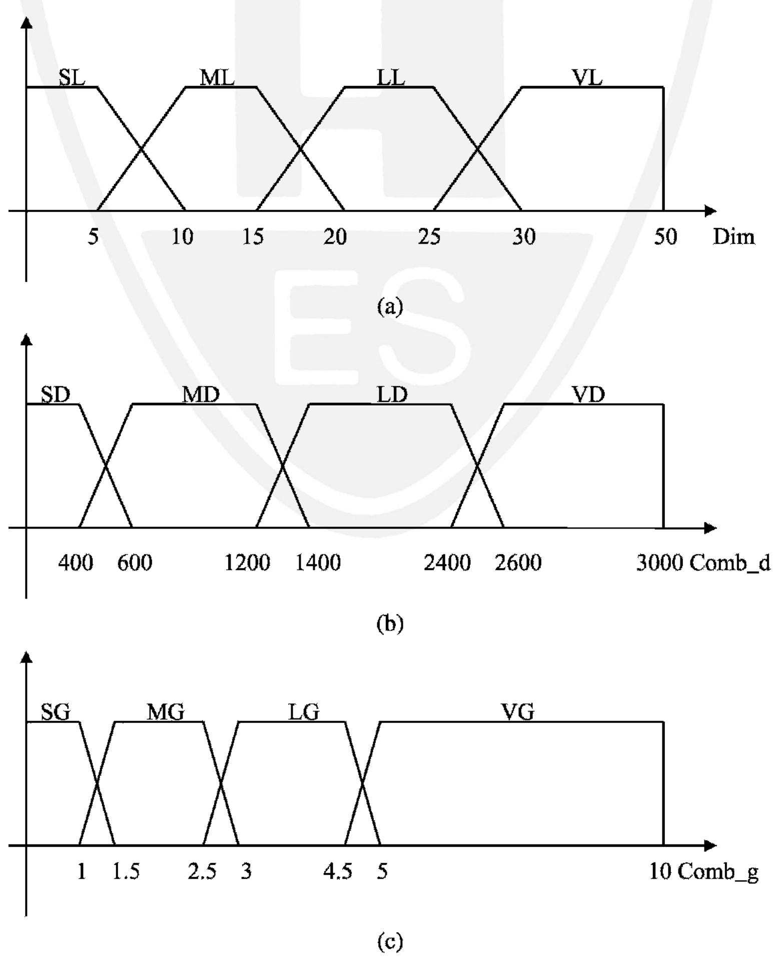

Fig. 10. Membership functions of filter parameters. (a) Output parameter Dim. (b) Output parameter Comb_d. (c) Output parameter Comb_g. (d) Output parameter Apd. (e) Output parameter Fc. (f) Output parameters Alpha, Ge, and Gr.

5 CONCLUSIONS

Reverberation is known to be an important element of human perception in acoustical environments. Appropriate reverberation should be incorporated into dry input signals to create proper audio impressions of distance as well as spaciousness in three-dimensional audio reproduction. Although direct convolution using measured room responses can produce the best quality of reverberation, high computation costs prevent practical real-time implementation. Traditionally all-pass or comb networks have been a useful alternative of reverberation synthesis. The tuning process of this type of artificial reverberators, however, is often time-consuming and heuristic. To alleviate this problem, this paper presents a reverberator composed of an FIR-based early reflection module and an IIR-based late reverberation module. In the system a systematic and efficient method based on the GA procedure was also developed to design the optimal reverberators. The optimization procedure hinges on several characteristics pertaining to natural room reverberations. A fuzzy logic and graphic user interface was also developed to facilitate

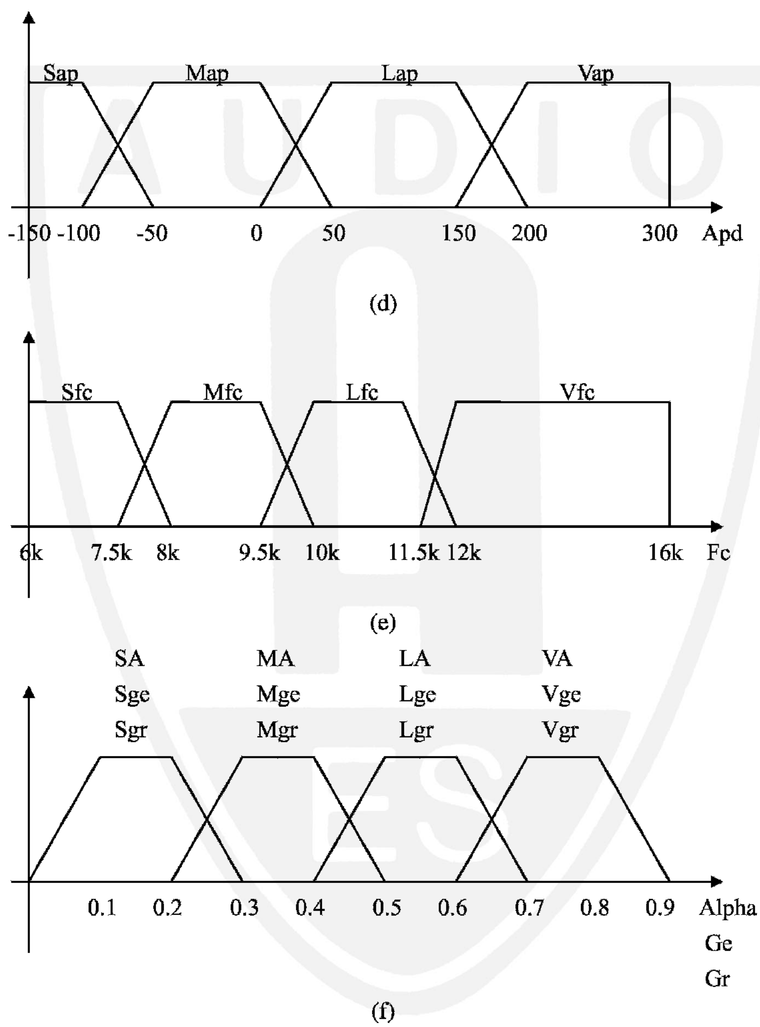

Fig. 10. *Continued.*

Table 3. Fuzzy rules for fuzzy inference stage 1.*

IF	Then				
R	Room Size	Diffusion	Warmth	Clarity	Reverb
LR	S	M	S	L	L
SC	M	L	M	S	M
CH	L	VL	L	VL	VL
LA	VL	L	VL	M	S
GY	L	S	M	VL	M

* R—room mode; LR—living room; SC—small club; CH—church; LA—large auditorium; GY—gymnasium; S—small; M—medium; L—large; VL—very large; reverb—reverberation.

the selection of a user's favorite room modes. Subjective listening experiments revealed that the present fuzzy system is capable of delivering natural sounding reverberations for various sound venues, as compared to those generated by direct convolution using the measured room response.

Table 4. Fuzzy rules for fuzzy inference stage 2.*

If					Then							
Rs	Diff	Wa	Cla	Rev	Dim	C d	C g	Apd	α	Fc	Ge	Gr
S					S	S	S					
M					M	M	M					
L					L	L	L					
V					V	V	V					
	S							V				
	M							L				
	L							M				
	V							S				
		S							V	V		
		M							L	L		
		L							M	M		
		V							S	S		
			S								0.3V	S
			M								0.3L	M
			L								0.3M	L
			V								0.3S	V
				S								S
				M								M
				L								L
				V								V

* Rs—room size; Diff—diffuseness; Wa—warmth; Cla—clarity; Rev—reverberation; Dim—dimension; Comb_d—comb filter delay; Comb_g—comb filter gain; Apd—all-pass filter delay; α—high-frequency ratio; Fc—cutoff frequency; Ge—early reflection gain; Gr—late reverberation gain.

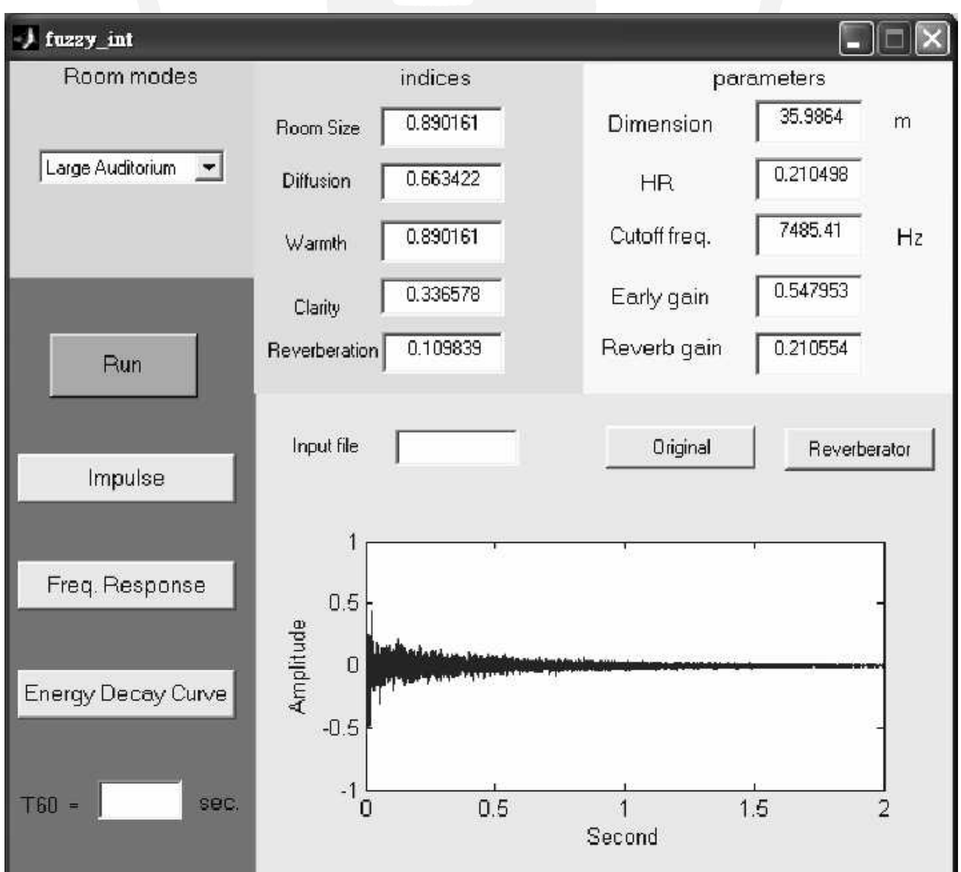

Fig. 11. Fuzzy logic and graphic user interface of artificial reverberator. Room impulse response is shown in display window.

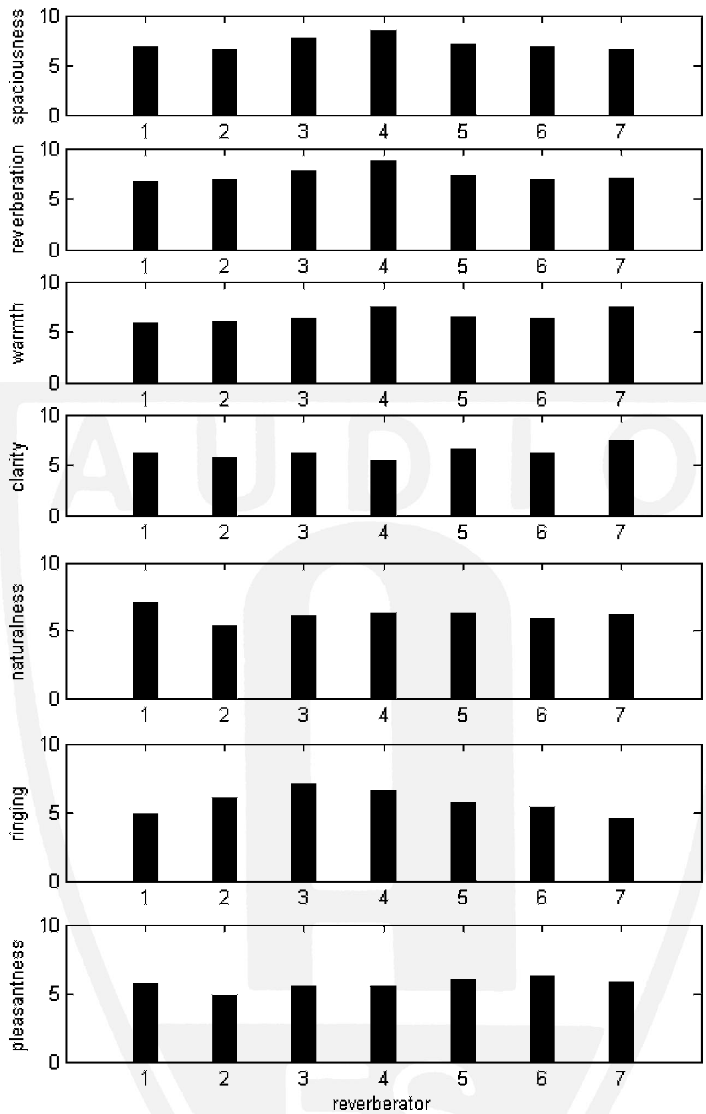

Fig. 12. Results of subjective listening experiments shown in terms of each subjective index for seven reverberators.

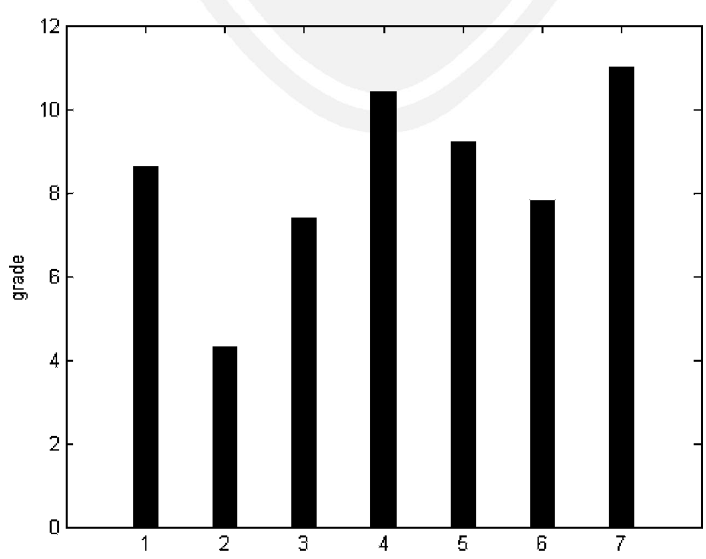

Fig. 13. Total score of subjective assessment for seven reverberators.

6 ACKNOWLEDGMENT

The work was supported by the National Science Council of the Republic of China under project NSC 92-2212-E009-030.

7 REFERENCES

[1] M. R. Schroeder, "New Method of Measuring Reverberation Time," *J. Acoust. Soc. Am.,* vol. 37, pp. 409–412 (1965).

[2] J. A. Moorer, "About This Reverberation Business," *Computer Music J.,* vol. 3, no. 2, pp. 13–28 (1979).

[3] W. G. Gardner, "Reverberation Algorithms," in *Applications of DSP to Audio and Acoustics,* M. Kahrs and K. Brandenburg, Eds. (Kluwer Academic, Boston, MA, 1998), pp. 85–131.

[4] L. Dahl and J. M. Jot, "A Reverberator Based on Absorbent All-Pass Filters," in *Pro. COST G-6 Conf. on Digital Audio Effects (DAFX-00),* (Verona, Italy, 2000 Dec. 7–9).

[5] L. Savioja, "Modeling Techniques for Virtual Acoustics," Ph.D. thesis, Espoo, Finland (1999).

[6] H. Kuttruff, *Room Acoustics* (Elsevier, New York, 1991).

[7] W. G. Gardner, "A Real-Time Multi-Channel Room Simulator," *J. Acoust. Soc. Am.,* vol. 92, pp. 2395–2404 (1992).

[8] C. T. Lin, and G. C. S. Lee, *Neural Fuzzy Systems* (Prentice-Hall, Englewood Cliffs, NJ, 1999).

[9] W. C. Sabine, 1972, "Reverberation," in *Acoustics: Historical and Philosophical Development,* R. B. Lindsay, Ed. (Dowden, Hutchinson, and Ross, Stroudsburg, PA, 1972).

[10] H. Lehnert and J. Blauert, "Principles of Binaural Room Simulation," *Appl. Acoust.,* vol. 36, pp. 259–291 (1992).

[11] W. G. Gardner, "*The Virtual Acoustic Room,*" Master's thesis, MIT Media Lab, Cambridge, MA (1992).

[12] J. Frenette, "Reducing Artificial Reverberation Requirements Using Time-Variant Feedback Delay Networks," Master Research Project, University of Miami, Coral Gables, FL (2000).

[13] F. A. Beltràn and J. R. Beltràn, "Implementing Reverberation Algorithms in Matlab," *J. New Music Res.,* vol. 31, pp. 153–161 (2002).

[14] D. Ouis, "Study on the Relationship between Some Room Acoustical Descriptors," *J. Audio Eng. Soc.* (*Engineering Reports*), vol. 51, pp. 518–533 (2003 June).

THE AUTHORS

Ganyuan Bai was born in 1976 in Tainan, Taiwan, ROC. He received a bachelor's degree in aeronautics and astronautics engineering from the National Cheng-Kung University, Taiwan, in 1999 and a master's degree in mechanical engineering from the National Chiao-Tung University, Hsin-Chu, Taiwan, in 2004. His master's thesis is on optimal design of reverberation with fuzzy user interface for spatial audio. He is currently an engineer at Quanta Storage Inc.

The biography of Mingsian R. Bai was published in the 2005 April issue of the *Journal.*