inter.noise 2000
The 29th International Congress and Exhibition on Noise Control Engineering
27-30 August 2000, Nice, FRANCE

I-INCE Classification: 7.9

A REAL-TIME AURALIZATION SYSTEM FOR PC

J.-H. Wang, J.C. Ferng

Department of Power Mechanical Enginneering, National Tsing-Hua University, 300, Hsinchu, Taiwan

Tel.: 886-3-5715131-3759 / Fax: 886-3-5727873 / Email: jhwang@mx.nthu.edu.tw

Keywords:
AURALIZATION

ABSTRACT
A real-time auralization system for PC was developed in this work. In order to realize the convolution in real time, a method was proposed in this work to reduce the amount of data points of the BRIR. According to the methods proposed in this work, the computational loading was reduced to 1/360. The auralization system can get the original signal from the HD or CD-ROM of PC and then convolve the signal with a selected BRIR in real time. The results are output via normal sound card and one can listen 3-D sound by headphones.

1 - INTRODUCTION

The development of binaural sound field simulation is got more and more attention recently because audio is very important in virtual reality. If the auralization system is presented via headphones, i.e., the sound field is listened by headphones, the auralization can be regarded as two steps of process: (1) to generate the binaural room impulse response (BRIR) (2) to convolve the original signal with the BRIR. Theoretically, the BRIR can be obtained by experimental measurement [1] or by theoretical simulation [2]. Depending on the acoustic characteristics of the space (room), the duration of the BRIR can be longer than 2s (for instance, a typical concert hall) for the response energy to decay 60dB. If a audio signal should be convolved with the BRIR by PC in real time, the PC should finish 176400 multiplications (88200×2 for left and right ears) and 176400 summations within 0.000023s (1/44100). It is evident that a general purpose PC available nowadays can not meet this basic requirement for auralization.
In order to realize the convolution in real time by PC, some criteria were proposed in this work to reduce the data points of the BRIR. The auralization system developed in this work can get the signal from the HD or CD-ROM and then convolve it in real time with the selected BRIR which is stored in PC in advance.

2 - CRITERIA FOR BRIR REDUCTION

For simplicity, the BRIR for the left and right ears are denoted by $h_\ell(\tau)$ and $h_r(\tau)$, respectively. If the input anechoic signal is $x(t)$, then the output signals for left and right channels should be

$$y_\ell(t) = \int_0^\infty h_\ell(\tau)\, x(t-\tau)\, d\tau, y_r(t) = \int_0^\infty h_r(\tau)\, x(t-\tau)\, d\tau, \tag{1}$$

or in digital summation form

$$y_\ell(k) = \sum_{j=0}^{N-1} x(k-j)\, h_\ell(j)\,, y_r(k) = \sum_{j=0}^{N-1} x(k-j)\, h_r(j) \tag{2}$$

where N is the number of data points of h_r or h_ℓ, for instance, N=88200 for 2s (at 44.1KHz sampling rate). As discussed before, it is impossible to calculate the convolution in real time by the PC nowadays. Therefore, the data points of the BRIR should be properly reduced, however, without degrading the aural impression of the 3-D sound field significantly.
The basic concept of data reduction can be stated as: "No matter how to reduce the data of a BRIR, the simplified BRIR should still provide the following information: (1) the location (or position) of sound

source, (2) the dimensions of the space, (3) the sound absorption characteristic of the space." According to this basic concept, some criteria were proposed in this work to reduce the data of a BRIR.

2.1 - Criteria for keeping the information of source location

The hearing system of humans can identify the location of the source from the infinite reflections by the so called first wavefront theory. In other words, the hearing system of humans uses the information of the direct sound to determine the location of the sound source. Therefore, the direct sound in the BRIR should be kept.

2.2 - Criteria for keeping the spatial impression

According to the results of Kuttruff [3], Barron [4], et. al., the reflections in a space can be roughly classified into three categories, i.e., (1) the direct sound, (2) the early reflections, (3) the later reverberation. The spatial impression is mainly contributed by the early reflections, especially due to the reflections from the lateral walls. However, so far there is no definite criterion to define quantitatively what the early reflection is. In this work, a factor was proposed to define the duration of the early reflections quantitatively. It is known that the path length of a reflection is proportional to the delayed time of the reflection. For a typical rectangular space, the delayed time generally is proportional to the order of reflection. The order of reflection means that the number of time the reflection hits the boundary before it reaches the receiver. The number of reflections of each order is generally proportional to the order of reflection. Therefore, one can define a factor K_n for each order of reflection as:

$$K_n = \sum_{i=1}^{N_n} \left[\frac{1}{(d_i)_n^2} \right] \tag{3}$$

where N_n represents the number of reflections with order of reflection n, and d_i indicates the path length of the reflection i. A typical example of the factor K_n vs. the order of reflection is shown in Fig. 1. One can find that K_n becomes constant as the order of reflection is larger than 5. Therefore, it is proposed in this work that the early reflections are defined as: all the reflections which arrive the receiver earlier than the time as the K_n becomes constant.

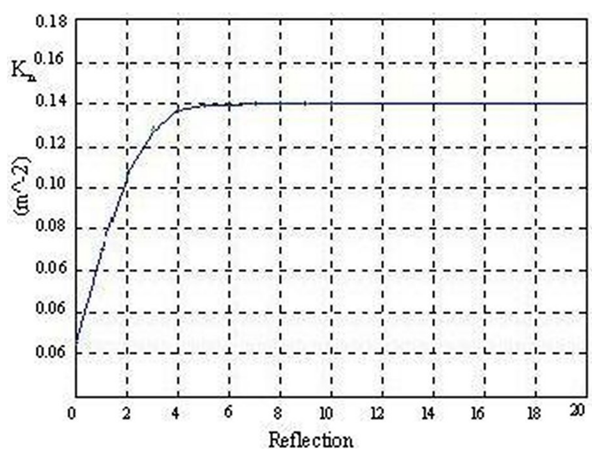

Figure 1: The definition of early reflection.

Although the spatial impression is mainly contributed by the early reflections, not all the early reflections have the same contribution to the spatial impression. According to the works of Kuttruff [3] and Barron [4], the early reflections coming from the lateral walls are the main contributors of spatial impression. Therefore, only the lateral reflections in the BRIR are kept in this work. However, according to our listening tests if all the reflections coming from the back wall, front wall, the floor and the ceiling are all eliminated, the perception of the dimensions of the space may be distorted, i.e., the perception of the dimensions of the space becomes larger than the actual size. Therefore, it is suggested that the first reflections coming from the back wall, front wall, the floor and ceiling are kept.

In summary, the criteria to keep the spatial impression are:

- All the early lateral reflections are kept except that the energy difference between two reflections within 20ms is larger than 15dB, then the smaller reflection is omitted.

- All the early reflections coming from the back wall, front wall, the floor and ceiling are omitted except the first reflections.

2.3 - Criteria to keep the later reverberation

All the reflections which arrive the receiver later than the early reflections are defined as the later reverberation. In this work, we use four parallel comb filters for each channel to synthesize the later reverberation. The parameters of the comb filters were so designed that the decayed rate of the filter was the same as the original decayed rate of the later reverberation of the BRIR.

According to the above proposed criteria, the proposed procedure to generate the simplified BRIR is shown in Fig. 2. Besides the reduction of the data points, the MMX instructions developed by Intel Corporation were used in this work to speed up the convolution.

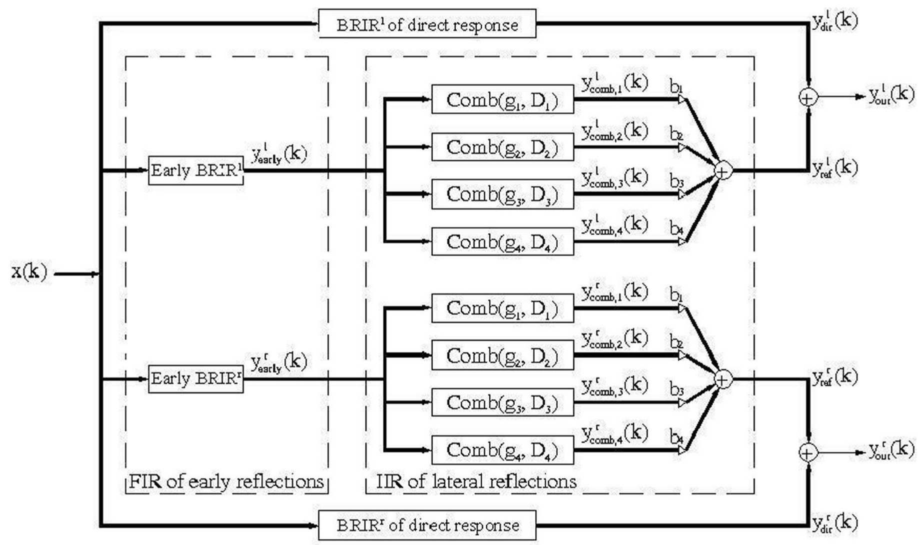

Figure 2: Proposed procedure to generate the simplified BRIR.

3 - HEADPHONE LISTENING TESTS

In order to evaluate the quality of the simplified BRIR, two music samples were chosen to convolve with the original BBIR and the reduced BRIR. One of the music samples, Minuet De L'arlesienne by Bizet, was recorded in the anechoic chamber of Tsing Hua University with flute as the instrument, and the other sample was a piece of piano sonata by Beethoven, and was generated by general MIDI device. Because the results convolved with the original BRIR and with the reduced BRIR can be compared via headphones, it is very easy for a normal person to feel whether there is a noticeable difference between the two results. The test people are the students of the Tsing Hua University. The testing results indicate that more than 85% of the tested students can't point out the difference between the two results, and only a few students said they could feel the difference but could not point out which one was better.

The convolution results also indicate that the ratio of the CPU times for the convolution with the original BRIR and with the reduced BRIR is about 360:1. Note that the MMX technology was not used during the convolution with the original BRIR.

4 - REAL-TIME AURALIZATION SYSTEM FOR PC

The goal of this work is to develop a real-time auralization system for PC, it was decided to implement it on Microsoft Windows 98. According to our testing results, a real time convolution is possible for the original BRIR longer than 2s with Pentium II 300MHz. The convolution algorithm applied the MMX instructions to speed up the calculation and the API of the Direct Sound was used to get and output the waveform data via normal sound card. The main selection menu of the system is shown in Fig. 3.

The main functions in the menu are described briefly below:

- Input Signal: One can choose the signal to be convolved with BRIR from a data file stored as wave format in PC or from the mixer of a sound card.
- Angle: On can choose the source location in horizontal plane with 45° as a step. The 0° indicates the source locates right in front of the receiver.

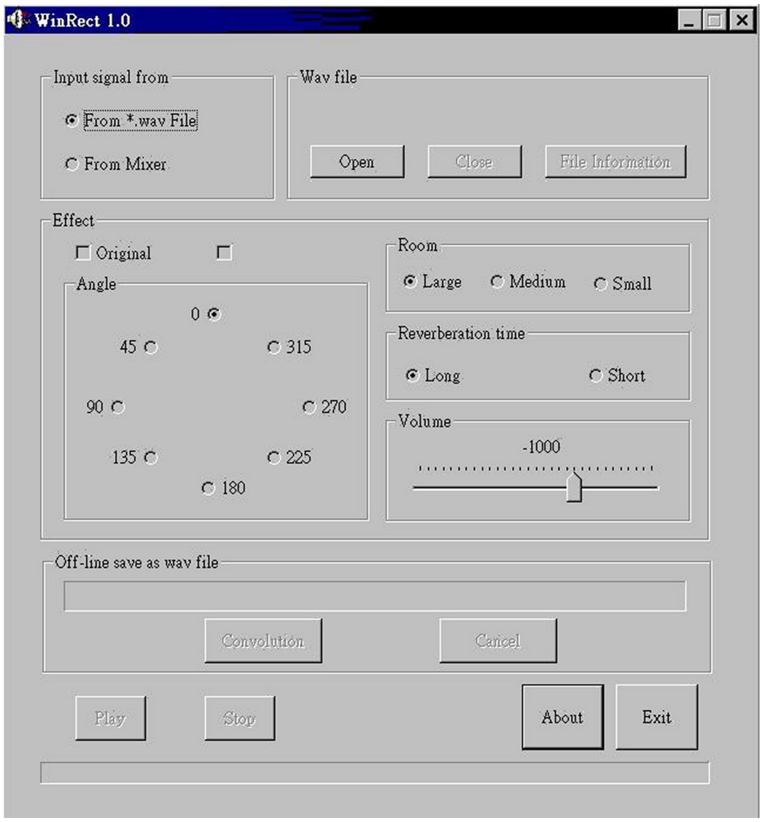

Figure 3: The main selection menu of the auralization system.

- Room type: There are three typical room types for choice. The volumes of the large, medium and small rooms are 7400m^3, 2630 m^3 and 75m^3, respectively.
- Reverberation time: For each room type there are two reverberation times for choice.
- Volume: One can adjust the volume of sound by mouse.

The system developed in this work is very easy to use and does not need any extra instrument for PC. The system can get the signal from CD-ROM via mixer of a sound card or direct from the memory of the PC and output the convolved signal via normal sound card.

5 - CONCLUSION

A real-time auralization system for PC was developed in this work. In this work, some criteria based on the psychroacoustics and room acoustics were proposed to reduce the data amount of the BRIR so that the computational loading was reduced to about 1/360. The system developed in this work is very easy to use and does not need any extra instrument for a normal PC. The system is software-based and can change the parameters easily in contrast to the hardware-based sound effect provided by some sound cards.

ACKNOWLEDGEMENTS

Part of this work is supported by the National Science Council, R.O.C. (Contract No.NSC 89-2212-E007-010). This support is hereby gratefully acknowledged.

REFERENCES

1. **J.H. Wang and C.L. Chen**, Measurement and Verification of the Binaural Room Impulse Response Function, In *The 7th National Conference on the Chinese Society of Sound and Vibration, Hsinchu, Taiwan*, pp. 145-149, 1999
2. **R. Heinz**, Binaural Room Simulation Based on an Image Source Model with Addition of Statistical Methods to Included the Diffuse Sound Scattering of Walls and to Predict the Reverberant Tail, *Applied Acoustics*, Vol. 38, pp. 145-149, 1993

3. **H. Kuttruff**, *Room Acoustics*, Third Edition, Elsevier Applied Science, London, 1991

4. **M. Barron**, *Auditorium Acoustics and Architectural Design*, E & FN SPON, London, 1993