ICA 14 BEIJING 1992

L3-7

BASIC STUDY ON SOUND FIELD SIMULATION BASED ON RUNNING INTERAURAL CROSS-CORRELATION

M. MORIMOTO* and K. IIDA**

* Environmental Acoustics Laboratory, Faculty of Engineering, Kobe University, Rokko, Nada, Kobe 657 Japan
**Environmental Acoustics Laboratory, Faculty of Engineering, and Matsushita Communication Industrial Co., Ltd. Saedo, Midori, Yokohama 226 Japan

1. INTRODUCTION

The purpose of sound field simulation is to make a listener perceive the same characteristics of sound image as those of a sound image which he perceives in an original (to be simulated) sound field. It is clear that this purpose can be achieved by simulating physically both of temporal and spatial structures of the original sound field. On the other hand, the purpose may be achieved by simulating the output of the binaural signal processing mechanism even if the simulated sound field is not physically equal to the original one. This paper investigates the possibility of the latter method of the sound field simulation by performing four psychological experiments. The simulation method is based on the running interaural cross-correlation model which is regarded as one of mechanisms in binaural signal processing and is called the RCC method.

2. PRINCIPLE OF SIMULATING A SOUND FIELD BASED ON THE RUNNING INTERAURAL CROSSCORRELATION MODEL

It is necessary to grasp the time delay, the level and the direction of all reflections, in order to simulate a sound field. The RCC method is applied to determine the direction of reflections. The procedure is as follows.

The running interaural crosscorrelation function $\Phi_{lr}(t)$ is calculated by Eqs.(1) and (2) by using impulse responses from a sound source to both entrances of ear canal in an original sound field.

$$\Phi_{lr}(t,\tau) = \int_{-\infty}^{t} p_l(\xi) p_r(\xi-\tau) G(t-\xi) d\xi \qquad (1)$$

where

$p_l(\zeta)$ and $p_r(\zeta-t)$ are the input signals to the left and the right entrance of ear canal, respectively.

τ is a time lag between left and right ear. $G(t-\zeta)$ is a temporal window.

Then the degree of interaural crosscorrelation for a temporal window at t, $RCC(t)$, is defined by

$$RCC(t) = |\phi_{lr}(t,\tau)|_{max}, \quad for\ |\tau| \le 680\mu s \qquad (2)$$

where

$\phi_{lr}(t,\tau)$ is the normalized interaural crosscorrelation function for a temporal window at t.

The interaural time difference for a temporal window at t, $RTD(t)$ is obtained by Eq.(3).

$$RTD(t) = \tau, \quad for\ |\phi_{lr}(t,\tau)| = RCC(t) \qquad (3)$$

The direction of the reflection is determined as the direction where the interaural time difference corresponds to the $RTD(t)$.

3. PSYCHOLOGICAL EXPERIMENTS

Four psychological experiments were performed to solve following technical problems in practical application. (1) What is the sound signal that should be provided from the direction of reflection determined as mentioned above? (2)How long is the duration of a temporal window that is proper for the simulation? (3)Whether the mechanism of the inner ear should be considered or not?

To first approximation, the shape of a temporal window $G(t-\xi)$ was rectangular and the adjacent temporal windows did not overlap each other but were in contact with each other. The source signal was a dry music motif (7sec). The original(to be simulated) sound field was created by 8(Exp.I) or 7(Exps.II,III,IV) loudspeakers located on the frontal half of the horizontal plane and consisted of early discrete 66 reflections within 200ms after the direct sound(Exp.I) or a lot of reflections of 2 sec. long (Exps.II,III,IV).

Kruscal's multidimensional scaling was used in Exps.I,III, and IV, and the method of complete triads was used in Exp.II.

3.1 Experiment I

The purpose of this experiment is to investigate the possibility of this simulation method, first of all.

In this Exp., one reflection of the simulated sound field is derived from each temporal window. The amplitude of the reflection is the equivalent sound pressure level for the window. The time delay of the reflection is the time when the amplitude is maximum in the window.

Four sound fields were created by the RCC method as a parameter of duration of the temporal window. The durations were 5, 10, 20, and 40ms. For comparison, other four sound fields were also created by choosing the greatest reflections among those of the original sound field(CGR method).

Figure 1 shows the two-dimensional configuration of the nine sound fields. The original sound field is plotted at the origin. The sound field RCC20 is the most similar sound field to the original one. The average of the judgments of the degree of dissimilarity between the sound field RCC20 and the original sound field was 1.63, which lies between "not different" and "a little different." This means that the sound image perceived by 67 reflections can almost be simulated by only 20 reflections.

The similarity between the sound field created by the RCC method and the original ones is not correlated to the number of reflections. From the result of this experiment, there seems to be a possibility of simulating almost the same characteristics of sound image as those perceived in an original(to be simulated) sound field by the RCC method and it seems to suggest that 10ms is a proper duration of the temporal window in calculating RCC.

3.2 Experiment II

The purpose of this experiment is to make clear the effect of the mechanism of inner ear in calculating RCC on the accuracy of simulation. The mechanism is practically approximated by a low pass filter(Fc=1600Hz)[1]. Two sound fields were simulated. One was simulated considering the filter(sound field:P) and another was not(sound field :N). But, in this simulation, all reflections included in each temporal window in calculating RCC, which were observed by one-point microphone at the center of a listener's head, were reproduced from the direction determined by the RCC method for each window. Both of them and the original sound field(sound field:O) were compared by the method of complete triads.

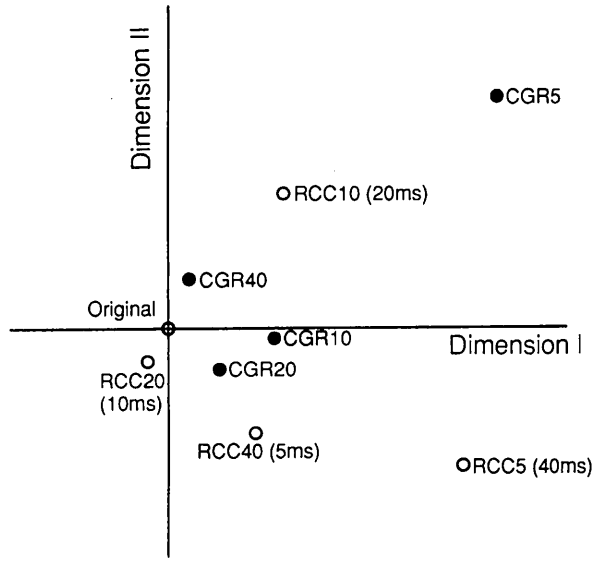

Fig.1 Result of Exp.I. The number represents the number of reflections, and the time in the bracket is the duration of the temporal window.

Table 1 indicates the psychological distance between them. According to the subjects' reports after the Exp. the subjects were divided to two groups. Namely subjects of group I took notice of the spatial characteristics and those of group II took notice of the tone color in comparing. It can be considered that the subjects of group I can discriminate between O and N and between P and N and that the subjects of group II cannot. These results conclude that the low pass filter should be considered in calculating RCC.

Table 1 Psychological distance between three sound fields in Exp.I

	Between sound fields		
Subject	O - P	O - N	P - N
Group I	0.05	0.87	0.82
Group II	0.51	0.42	0.09

O: Original sound field, P: simulated sound field with low pass filter, N: simulated sound field without low pass filter

3.3 Experiment III

The purpose of this experiment is to investigate the difference between the simulated sound fields where one reflection is derived from each temporal window like Exp.I(method of impulse response simplified) and where all reflections included in each temporal window were reproduced like Exp.II(method of impulse response not simplified).

Four sound fields were created by using each method as a parameter of duration of the temporal window. The durations were 2, 5, 10, and 20ms. The low pass filter was considered in both simulation methods.

The result shows that simulated sound fields fall into two clusters as shown in **Fig.2**. This means a devoluted solution in Kruscal's analysis. According to the subjects' reports after the Exp. the difference was caused by tone color. The original sound field belongs to the identical cluster that sound fields simulated by method of impulse response not simplified. This means that the impulse response should not be simplified to simulate tone color of sound image.

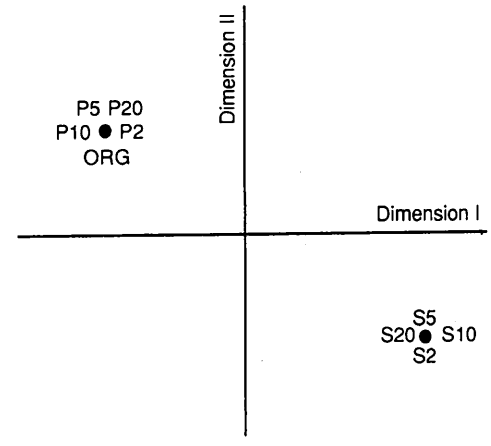

Fig.2 Result of Exp.II. ORG:Original sound field, P: sound field simulated by method of impulse response not simplified, S: sound field simulated by method of impulse response simplified. The number is the duration of the temporal window.

3.4 Experiment IV

The purpose of this experiment is to investigate how long the duration of temporal window is proper for the simulation by method of impulse response not simplified. Four sound fields were created as a parameter of duration of the temporal window. The durations were 2, 5, 10, and 20ms. The low pass filter was considered in the simulation.

Figure 3 shows the one-dimensional configuration of the original and four simulated sound fields, according to the responses of the subjects who took notice of the spatial characteristics of sound image. The result shows that 2ms or 10ms is more close to the original sound field than 5ms and 20ms. The rectangular window of 8ms long is proposed regarding to temporal masking[2]. Therefore, it can be considered that 10ms is proper as the duration of temporal window, from a standpoint of the simulation based on the auditory mechanism.

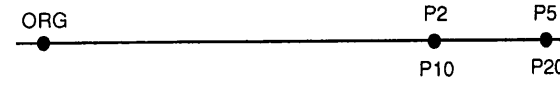

Fig.3 Result of Exp.III. ORG:Original sound field. The number is the duration of the temporal window.

4. CONCLUSIONS

Results of four psychological experiments show; (1)there is a possibility of the simulation based on RCC, (2) the low pass filter should be considered in calculating RCC, (3) the impulse response should not be simplified, and (4) 10ms is proper as the duration of temporal window in calculating RCC.

REFERENCES

[1] Yamanaka: Master thesis(Kobe Univ. 1991)
[2] Moor et.al: JASA, 83 (1988) 1102