Acoust. Sci. & Tech. **25**, 5 (2004)

TECHNICAL REPORT

A sound field simulation system for the study of ensemble performance on a concert hall stage

Motoko Rokutanda[1], Takako Kanamori[1], Kanako Ueno[2,*] and Hideki Tachibana[2]

[1]*Graduate School, University of Tokyo,*
4–6–1, Komaba, Meguro-ku, Tokyo, 153–8505 Japan
[2]*Institute of Industrial Science, University of Tokyo,*
4–6–1, Komaba, Meguro-ku, Tokyo, 153–8505 Japan

(*Received 3 October 2003, Accepted for publication 2 March 2004*)

Abstract: In order to investigate the effect of room acoustic conditions on music players on the stage, the authors have been conducting experimental studies of solo performances, using a newly developed sound field simulation method with a 6-channel recording/reproduction technique. By expanding this system, another simulation system to duplicate the situation where two musicians play in ensemble has been newly developed using two anechoic rooms and a 24 channel digital convolution system. Using this system, examinations were conducted to check the applicability of this sound field simulation technique to the experimental investigation of stage acoustics for ensemble performance between two players.

Keywords: Stage acoustics, Concert hall design, Sound field simulation, Ensemble performance

PACS number: 43.55.Fw, 43.55.Lb [DOI: 10.1250/ast.25.373]

1. INTRODUCTION

In concert hall design, the acoustic conditions for music players on the stage should be considered as well as those for the audience. The authors have been making experimental studies of "stage acoustics" focusing on solo performances by applying a newly developed sound field simulation technique using a 6-channel digital convolution system in an anechoic room [1,2].

In addition, the acoustic conditions for ensemble performances should be considered as an important factor of stage acoustics and several researchers have published studies on this topic [3–6]. The present authors have also started an experiment for ensemble performances by two players on a stage by expanding the sound field simulation technique mentioned above, using two anechoic rooms and a 24 channel digital convolution system. In this paper, the acoustic simulation system is first introduced and the results of the examination of its applicability to subjective experiments are presented.

*e-mail: ueno@iis.u-tokyo.ac.jp

2. SIMULATION SYSTEM FOR ENSEMBLE PERFORMANCES BY TWO PLAYERS

2.1. Modeling and Measurement Methods of the Sound Transmission between Two Players

The sound transmission between two players on a stage can be modeled as shown in Fig. 1. In this figure, P_{self} is the sound transmission from the instrument of a player to himself/herself and P_{cross} is that from the instrument of the co-player. To obtain these sound transmission characteristics, the impulse responses, $p_{\mathrm{AA}}(t)$, $p_{\mathrm{BB}}(t)$, $p_{\mathrm{AB}}(t)$ and $p_{\mathrm{BA}}(t)$ are measured in the configuration shown in Fig. 2. In the measurements, a dodecahedral loudspeaker system with an omni-directional characteristic up to about 2k Hz was used to model the musical instruments and an omni-directional or uni-directional microphone was located just behind the sound source as shown in Fig. 2. In the measurements using the uni-directional microphone (SONY C48), the impulse responses in six orthogonal directions ($p_{\mathrm{AA},i}(t)$, $p_{\mathrm{BB},i}(t)$, $p_{\mathrm{AB},i}(t)$ and $p_{\mathrm{BA},i}(t)$, $i = 1$ to 6) were measured by rotating the microphone in 90-degree increments [2]. The omni-directional impulse responses ($p_{\mathrm{AA,o}}(t)$, $p_{\mathrm{BB,o}}(t)$, $p_{\mathrm{AB,o}}(t)$ and $p_{\mathrm{BA,o}}(t)$) were also measured to obtain room acoustic indices. For this impulse response measurement, the time stretched pulse (TSP) technique [7,8] was applied by keeping the energy of the source signal constant.

2.2. Sound Field Simulation in the Laboratory

In order to simulate the sound field conditions for two players separately, two anechoic rooms (room-A: 7 meters cubed and room-B: 4.0 m × 6.8 m × 7.0 m) are coupled acoustically as shown in Fig. 3 and the four sets of 6-channel directional impulse responses installed in the digital convolution system (24 channels in total) are used to synthesize the sounds from six directions for each player.

In this system, the direct sound from each musical instrument is detected through a uni-directional microphone (SONY C48) set at a point close to the player in each room. The output signal is convolved with each impulse response of 2.7 s duration time measured in real concert halls. That is, the sound from the instrument of the player A is convolved with the sets of 6-channel directional impulse responses $p_{\mathrm{AA},i}(t)$ and $p_{\mathrm{AB},i}(t)$ and the sound from the instrument of the player B is convolved with the impulse responses $p_{\mathrm{BB},i}(t)$ and $p_{\mathrm{BA},i}(t)$. Here, for $p_{\mathrm{AA},i}(t)$ and $p_{\mathrm{BB},i}(t)$, the components of the direct sound from the sound source and the reflection from the stage floor are excluded from the directional impulse response signals, because the former is unnecessary for the simulation and the latter can not be simulated exactly in the simulation system used in this experiment.

The convolved signals are mixed as shown in Fig. 3 and reproduced through the six loudspeakers (TANNOY T12) set in each anechoic room. The loudspeakers are set on a spherical surface of 2 m radius in room-A and of 1.6 m radius in room-B around the performance positions. The frequency characteristics of the total recording and repro-

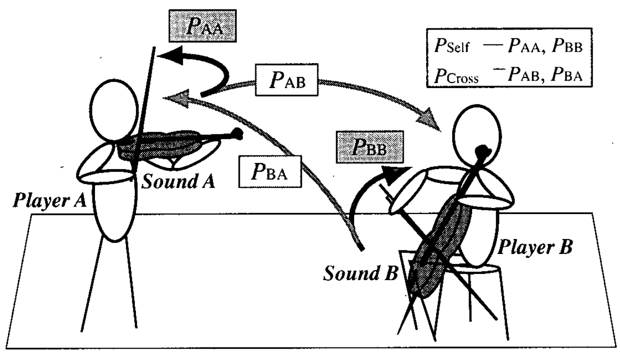


Fig. 1 Modeling of the sound transmission between two players.

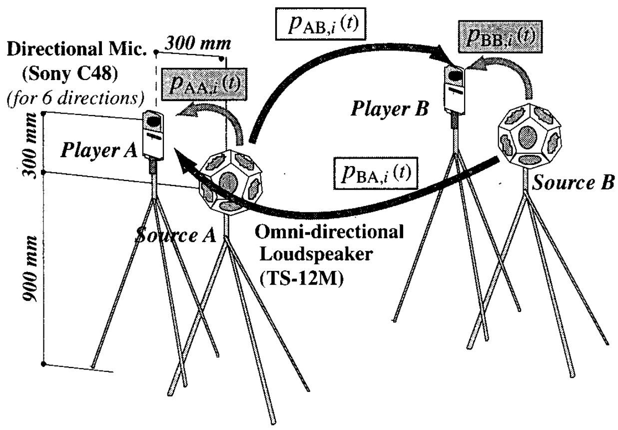


Fig. 2 Configuration of the measurement system.

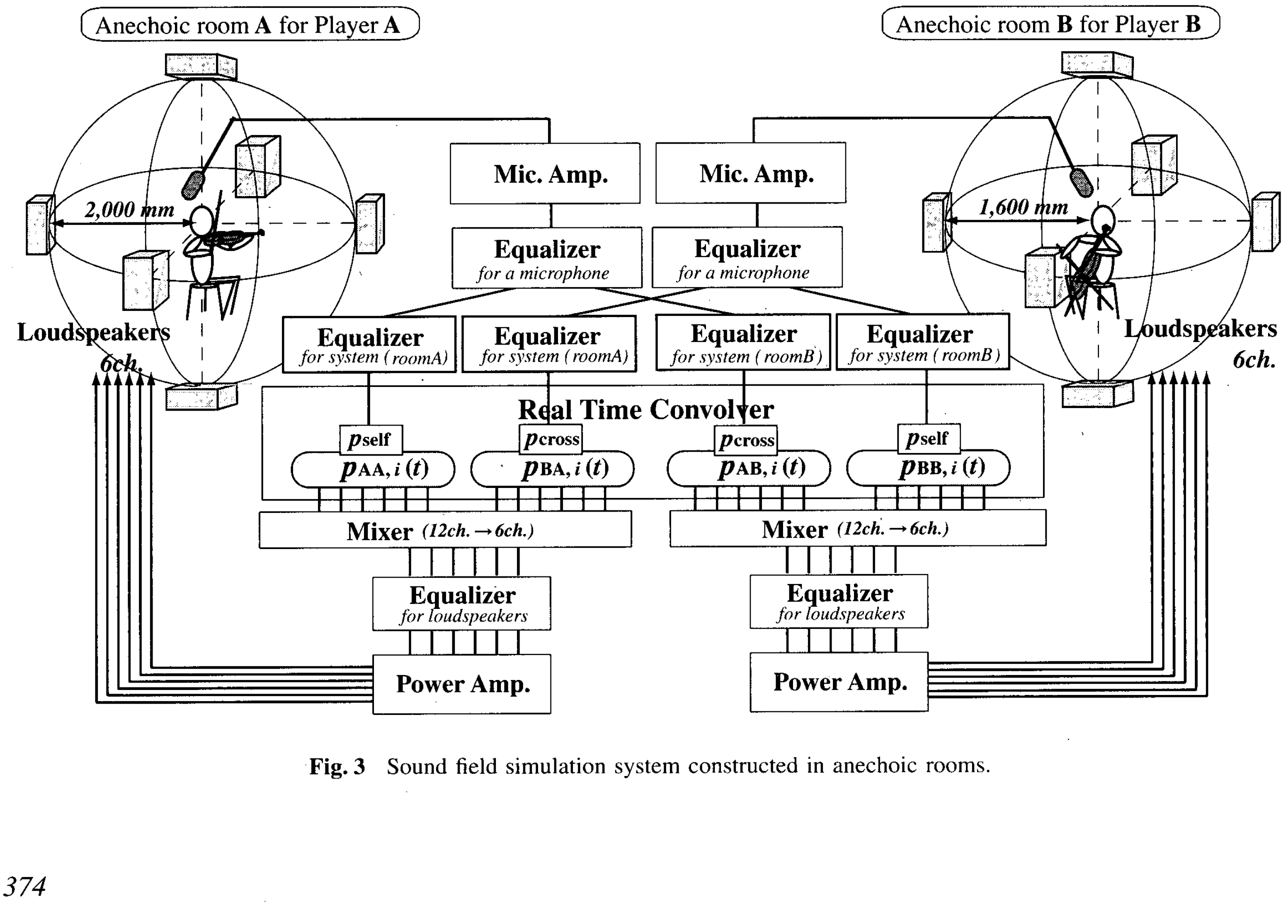


Fig. 3 Sound field simulation system constructed in anechoic rooms.

duction system, including the uni-directional microphone and the loudspeakers, are equalized over the frequency range from 125 Hz to 4k Hz in octave bands using spectrum equalizers.

3. EXAMINATION OF THE REPRODUCTION ACCURACY

3.1. Impulse Response Measurement in Real Concert Halls

The impulse response measurement mentioned above was performed in four concert halls in Japan as shown in Table 1. The measurement positions on a stage are shown in Fig. 4, where pattern A is for the players' positions for chamber music and pattern B are for those for an orchestra performance. The impulse responses between the two positions shown in Fig. 2 were measured for each pattern. Figure 5 shows the actual measurement on a stage.

Table 1 Data of the concert halls.

	Hall A	Hall B	Hall C	Hall D
Seats	440	761	1,502	2,020
Volume [m^3]	4,228	6,600	20,500	22,776
Stage area [m^2]	102	100	209	290
Style	shoebox	proscenium	proscenium	shoebox
RT (audience area) [s]	1.8	2.0	1.9	2.4
Measurement pattern	A	A	A, B	A, B

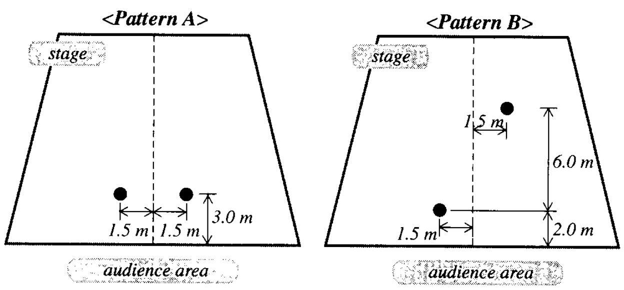


Fig. 4 Measurement positions on stages.

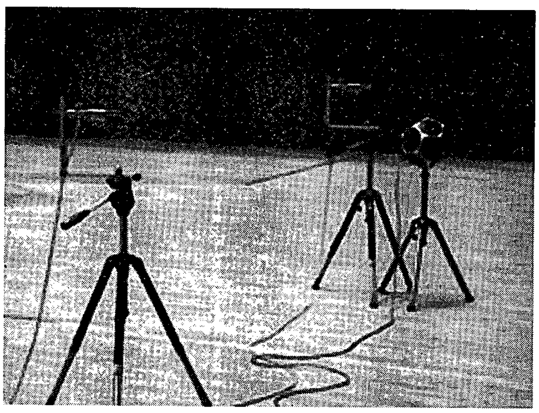


Fig. 5 Measurement on a stage.

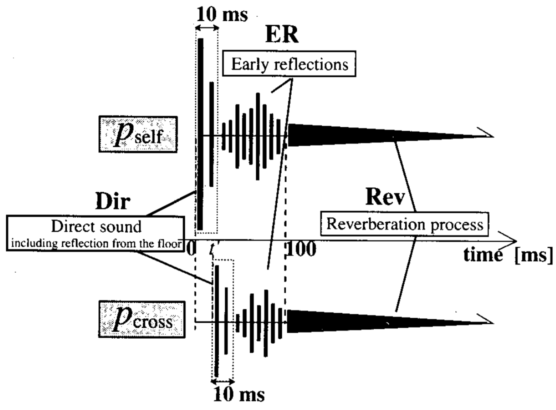


Fig. 6 Division of impulse responses into three components.

3.2. Methods for Analysis of the Impulse Responses

In order to examine the transient processes, the impulse responses measured with an omni-directional microphone were divided into three components as shown in Fig. 6; the direct sound including the reflection from the floor (Dir), the early reflections (ER) and the reverberation process (Rev). These components of an impulse response were separated in the time domain as follows.

$$p_T(t) = w_T(t) \cdot p(t) \tag{1}$$

where $p_T(t)$ is the respective part of the impulse response, $w_T(t)$ is the gating function and $p(t)$ is the over-all impulse response. This processing was performed for $p_{\text{self}}(t)$ and $p_{\text{cross}}(t)$ as shown in Table 2.

Next, the respective $p_T(t)$ was filtered through band-pass filters by digital signal processing. Here, it is expressed as $p_{T,N}(t)$: N indicates the N-th band. The energy (squared and time-integrated sound pressure), $E_{T,N}$, was calculated for each frequency band for the respective components as follows.

$$E_{T,N} = \int_0^\infty p_{T,N}^2(t)dt \tag{2}$$

From these results, the following indices normalized to $E_{\text{Dir},N,\text{self}}$, the energy of the direct sound component of $p_{\text{self}}(t)$, were calculated.

$$L_{\text{ER},N,\text{self}} = 10\lg\left(E_{\text{ER},N,\text{self}}/E_{\text{Dir},N,\text{self}}\right) \tag{3}$$

$$L_{\text{Rev},N,\text{self}} = 10\lg\left(E_{\text{Rev},N,\text{self}}/E_{\text{Dir},N,\text{self}}\right) \tag{4}$$

$$L_{\text{Dir},N,\text{cross}} = 10\lg\left(E_{\text{Dir},N,\text{cross}}/E_{\text{Dir},N,\text{self}}\right) \tag{5}$$

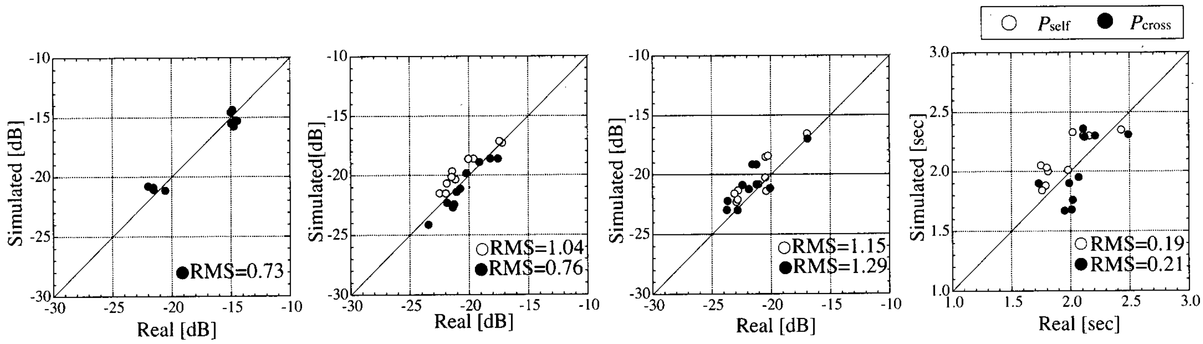
$$L_{\text{ER},N,\text{cross}} = 10\lg\left(E_{\text{ER},N,\text{cross}}/E_{\text{Dir},N,\text{self}}\right) \tag{6}$$

$$L_{\text{Rev},N,\text{cross}} = 10\lg\left(E_{\text{Rev},N,\text{cross}}/E_{\text{Dir},N,\text{self}}\right) \tag{7}$$

Table 2 Separation of impulse response in the time domain.

Components	$p_T(t)$	Gate-on time: $w_T(t) = 1$
The direct sound of $p_{\mathrm{self}}(t)$	$p_{\mathrm{Dir,\,self}}(t)$	$0 \le t < 10\,\mathrm{ms}$
The early reflections of $p_{\mathrm{self}}(t)$	$p_{\mathrm{ER,\,self}}(t)$	$10\,\mathrm{ms} \le t < 100\,\mathrm{ms}$
The reverberation process of $p_{\mathrm{self}}(t)$	$p_{\mathrm{Rev,\,self}}(t)$	$100\,\mathrm{ms} \le t < \infty$
The direct sound of $p_{\mathrm{cross}}(t)$	$p_{\mathrm{Dir,\,cross}}(t)$	$t' \le t < t' + 10\,\mathrm{ms}$
The early reflections of $p_{\mathrm{cross}}(t)$	$p_{\mathrm{ER,\,cross}}(t)$	$t' + 10\,\mathrm{ms} \le t < 100\,\mathrm{ms}$
The reverberation process of $p_{\mathrm{cross}}(t)$	$p_{\mathrm{Rev,\,cross}}(t)$	$100\,\mathrm{ms} \le t < \infty$

$*$ $t = 0$ is the starting point of $p_{\mathrm{self}}(t)$; $t = t'$ is the starting point of $p_{\mathrm{cross}}(t)$.



(a) L_{Dir} (b) L_{ER} (c) L_{Rev} (d) RT

Fig. 7 Reproduction accuracy.

As mentioned above, the over-all impulse responses were divided in the time domain and filtering processing was performed for each part, and the total energy of each part in each frequency band was calculated by squaring and time-integration processing. This is because after band-pass filtering the impulse response becomes long in time especially in the low frequency bands and it becomes impossible to divide it into the components: the direct sound, early reflections and the reverberation process. In Gade's definition, $L_{\mathrm{ER,\,self}}$ corresponds to ST_1 and $L_{\mathrm{Rev,\,self}}$ corresponds to ST_{late} [5].

3.3. Reproduction Accuracy

In order to determine the reproduction level of the simulation system, impulse responses were measured at the center point (performance position) of the simulated sound field in each room using the dodecahedral loudspeaker source and the omni-directional microphone in the same way as in the measurements in real concert halls (see Fig. 2). The reproduction level of the simulation system in each room was adjusted so that $L_{\mathrm{ER,\,self}}$ and $L_{\mathrm{Rev,\,self}}$ measured in the simulated sound field were equal to those measured in the real concert hall, respectively. Next, the reproduction levels for the cross components between the two rooms were adjusted so that $L_{\mathrm{Dir,\,cross}}$ measured in the simulated sound field were equal to that measured in the real hall. By making such level adjustments, the sound field conditions at the two performance positions can be simulated in the artificial sound field in each of the anechoic rooms.

To see the similarity between the real sound field and the simulated one, the correspondences of the indices defined by Eq. (3) to (7) and *RT* between the real sound field and the simulated one were examined. Here, when obtaining the value of E_{ER} for the simulated sound field, the starting time in the calculation was changed from 10 ms to 20 ms for $E_{\mathrm{ER,\,self}}$ and from $t' + 10$ ms to $t' + 20$ ms for $E_{\mathrm{ER,\,cross}}$, respectively, because reflections caused by the loudspeakers were observed just after the direct sound. The results for room-A (for the subject) in the middle frequency range (2-octave band including 500 Hz and 1 kHz octave bands) are shown in Fig. 7(a), (b) and (c), respectively. In each figure, the RMS value indicates the extent of the scattering of the plots. In each of these results, almost all plots lie on the 45 degrees line and a fairly good correspondence can be seen. Figure 7(d) shows the correspondence of *RT* between the real sound field and the simulated one for P_{self} and P_{cross} in room-A. In this result, the correspondence is not as good as the former cases and more deviations from the 45 degrees line are seen. The RMS value for P_{self} and P_{cross} are 0.19 s and 0.21 s, respectively. The reason of this discrepancy might be attributed to error in the evaluation of the decay curve and that occurred in the retouching processing of the

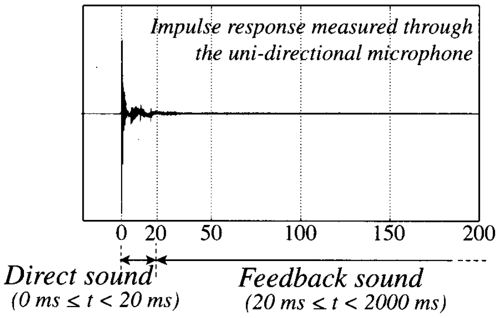


Fig. 8 Definition of the feedback sound in the simulated sound field.

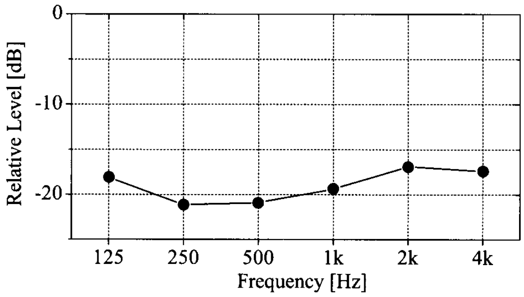


Fig. 9 Relative level of the feedback sound to the direct sound.

impulse responses by using the multiplying exponential function to depress the effect of background noise in the later part of the reverberation process [9].

Here, to examine the feedback effect from the loudspeakers to the uni-directional microphone used to detect the musical instrument's sound, the relative level between the energy of the direct sound (within 20 ms from the start of the direct sound) and that of the feedback sound (after 20 ms) in the impulse response measured through the uni-directional microphone in room-A was analyzed (see Fig. 8). Among the results, the relative level in each octave band for the most serious condition in which the reproduced sound from the loudspeakers was the strongest among the data shown in Fig. 7 is shown in Fig. 9. In this result, the energy of the false signal caused by the feedback effect is lower than that of the direct sound by 16 to 22 dB. For the other conditions, the relative level is lower, i.e. the feedback effect is weaker than this case. These results indicate that the feedback effect can be neglected in this sound field simulation system.

4. PRELIMINARY STUDY FOR SUBJECTIVE ASSESSMENT EXPERIMENT

Using these simulated conditions, a subjective assessment experiment was conducted to check the validity of the simulation system. The subject, a professional musician, performed one or two phrases in a piece at the center point of the simulated field in room-A with a co-player (an

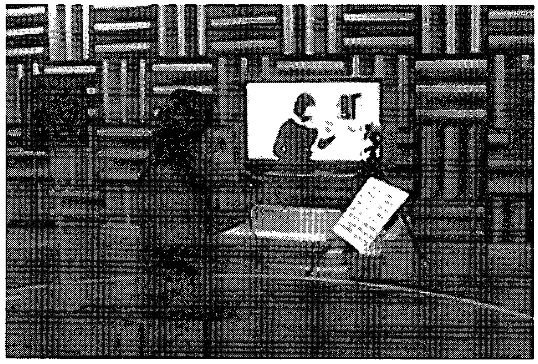


Fig. 10 Ensemble performance in the simulated sound field.

amateur violinist with 16 years' experience) in room-B by imaging that they were playing together on a real stage. The two players could see each other through a video-monitoring display set in each room (see Fig. 10).

During the performance, the experimenter stayed in room-A and questioned the subject's auditory impressions about the naturalness and presence (reality) of the sound field and the effectiveness of the visual presentation. In this preliminary experiment, variation of the hall conditions (the set of the directional impulse responses, see Table 1) for the same combination of performance positions and the difference of the performance positions for the same concert hall were examined (see Fig. 4). Eleven professional players (4 violinists, 2 violist, 1 cellist, 2 flutists, 1 clarinetist and 1 oboist) participated in this experiment. For each subject, four to eight variations among the test conditions described above were presented.

Table 3 shows the representative comments made by the subjects through the experiment. In these comments, different opinions are seen concerning the visual monitoring system. This might be attributed to the way of setting up the video-monitoring system in this experiment and the differences between the playing attitudes of the subjects related to their usual activity and musical instrument. As was expected, unnaturalness because the co-player was spatially separated was pointed out. Nevertheless, as seen in the comments on the subject's feelings about acoustic presence, it can be considered that the subjects can get the feeling of playing on the real stage of a concert hall and can distinguish differences between stage conditions. It was also commented that it is possible to concentrate on creating music with a co-player after becoming accustomed to the simulated sound field.

5. CONCLUSIONS

In order to simulate the 3-dimensional sound field on the stage of a concert hall, a technique using a 24 channel digital convolution system and two anechoic rooms was developed and its applicability to psycho-acoustic experiments on ensemble performance was examined by con-

Table 3 Comment of subjects on applicability of the simulation system.

Visual presentation	Unnaturalness	Acoustical presence
Not necessary ・I can play without the visual presentation because I usually play using ear. ・When concentrating on sound, the visual presentation is rather disturbing. Necessary to make sign ・I see the video-monitor only to see the starting sign from the co-player and when the ensemble is out of step. Indispensable ・Eye contact is essential in ensemble performance.	Presence of the co-player ・Physical contact is lacking because the co-player is not in the same space. ・Body movement is also necessary to take contact with the co-player; it can not be realized in this system. ・Breathing is also an important sign in ensemble and this system seems inorganic on this point.	Reality ・The difference of the distance between the players (3 m and the distance between the concert master and the wind instrument player) can be sensed as being accustomed in real halls. ・The easiness and difficulties in playing experienced in real halls can be sensed in this simulation field. ・The influence of the reverberation can be judged as in a real hall. Creation of music ・It is possible to concentrate on creating music with the co-player. ・When accustomed to this system, it may be possible to create music as usual in real halls.

ducting a preliminary subjective assessment experiment. As a result, it was found that musicians can get a feeling of the acoustic reality of performing on a stage and can distinguish the differences between acoustic conditions, though the absence of a co-player from the same room was felt to be unnatural.

From these results, it has been decided to pursue a more rigorous subjective assessment experiment to investigate the relationship between acoustic conditions on a stage and the musicians' impressions of ensemble performances by two players. This experiment is now being performed with the cooperation of professional musicians.

ACKNOWLEDGEMENT

The authors wish to generously thank the musicians who participated in this study for their willing cooperation during the subjective assessment experiments.

REFERENCES

[1] S. Yokoyama, K. Ueno, S. Sakamoto and H. Tachibana, "6-channel recording/reproduction system for 3-dimensional auralization of sound fields," *Acoust. Sci. & Tech.*, **23**, 97–103 (2002).

[2] K. Ueno and H. Tachibana, "Experimental study on the evaluation of stage acoustics by musicians using a 6-channel sound simulation system," *Acoust. Sci. & Tech.*, **24**, 130–138 (2003).

[3] A. H. Marshall, D. Gottlob and H. Alrutz, "Acoustical conditions preferred for ensemble," *J. Acoust. Soc. Am.*, **64**, 1437–1442 (1978).

[4] J. Meyer, "Problems of mutual hearing of musicians," *Proc. ICA '86*, pp. 33–38 (1986).

[5] A. C. Gade "Investigations of musicians' room acoustics conditions in concert halls, Part 1: Methods and laboratory experiments," *Acustica*, **69**, 193–203 (1989).

[6] J. Meyer, "Influence of communication on stage on the Musical quality," *Proc. ICA '95*, pp. 573–576 (1995).

[7] N. Aoshima, "Computer-generated pulse signal applied for sound measurement," *J. Acoust. Soc. Am.*, **69**, 1484–1488 (1981).

[8] Y. Suzuki, F. Asano, H. Y. Kim and T. Sone, "An optimum computer-generated pulse signal suitable for the measurement of very long impulse responses," *J. Acoust. Soc. Am.*, **97**, 1119–1123 (1995).

[9] M. Kaite, Y. Hidaka and H. Tachibana, "Measurement of impulse responses used for synthesizing test signals for the subjective evaluation of room acoustics," *Proc. 1990's Meet. Archit. Instit. Jpn.*, pp. 251–252 (1990).