Trials for subjective evaluation of auditory virtual environment

Maori KOBAYASHI [1,2], Kanako Ueno[2,3], Yusuke Ikeda[2,4], and Shiro Ise[2,4]

[1] Organization for the strategic coordination of research and intellectual properties, Meiji University, Japan
[2] School of Science and Technology, Meiji University, Japan
[3] CREST, Japan Science & Technology Agency
[4] School of Information Environment, Tokyo Denki University, Japan

Abstract

It has been necessary to establish subjective measures for the performance of the virtual acoustic systems. In this paper, we report our trials to evaluate the performance of a three-dimensional sound field reproduction system based on the boundary surface control principle, the 'Sound Cask'. First, we introduce our investigations for the experts of audio engineering in order to clarify the difference of auditory impression between the Sound Cask and conventional audio systems. Second, we report psychological and physiological experiments focusing on the advantageous points of the Sound Cask, a clear sense of reality, that were pointed out in the investigations for the experts. Finally, we discuss the issues to be considered for subjective evaluation of virtual acoustic systems for future studies.

Keywords – 3-D sound reproduction, reality, subjective evaluation, the Boundary surface control system

1 Introduction

Many acoustic media technologies have been proposed for reproducing acoustic environments realistically. At present, media technology is expected to transmit not only voice and music but also "presence," which refers to the subjective experience of being there in the acoustic environments. It is necessary to establish objective and quantitative measures of presence and/or realism, in order to evaluate listeners' experience in virtual acoustic environments.

In this study, we report our attempt at evaluating the performance of a three-dimensional sound field reproduction system called the "Sound Cask," based on the boundary surface control principle. First, we introduced our investigations to experts in audio engineering in order to clarify the differences of auditory impressions between the Sound Cask and conventional audio systems. Second, we report psychological and physiological experiments focusing on the advantages of the Sound Cask, a clear sense of presence, which was one of factors pointed out in the investigations by the experts.

2 The Sound Cask and the BoSC system.

In this study, we used the "Sound Cask" (Fig. 1 bottom) to present acoustic virtual environments that are close to the real environments. The Sound Cask is a new version of the boundary surface control (BoSC) sound reproduction system that enables us to record and reproduce a sound field area [1].

Ise proposed the boundary surface control principle (BoSC principle) [2]. By integrating the Kirchhoff-Helmholtz integral and inverse system, the BoSC system can accurately reproduce a three-dimensional (3-D) sound field surrounded by a closed boundary surface. The BoSC system comprises the BoSC microphone system (Fig. 1 left) and the reproduction room, namely the Sound Cask (Fig. 1 right). First, the BoSC microphone system records the sound pressure on the surface of a volume defined by the BoSC microphone array. Next, the recorded signals are convoluted with a set of inverse filters. Finally, the loudspeaker array in the Sound Cask accurately recreates the sound field in other locations by reproducing the convoluted signals. In the BoSC system, the inverse filters are determined by an inverse system of a transfer function matrix measured between each loudspeaker and microphone pair.

The BoSC microphone array has the same configuration as a *C80* fullerene. Eighty omnidirectional microphones are installed at the nodes of the fullerene. The reproduction area in the BoSC system is the inner region enclosed by the microphone array. Therefore, the system can reproduce a 3-D sound field surrounding the listener's head. Consequently, listeners can perceive sound images created within sound fields without being adversely affected by their own head movements.

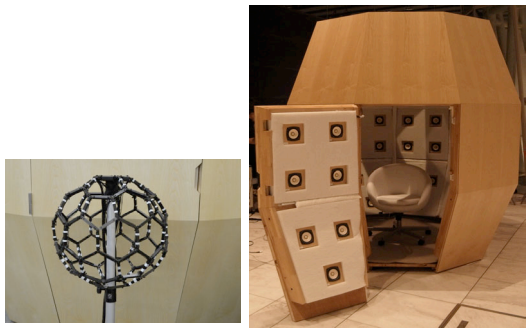

Figure 1. the BoSC microphone system (left panel) and the Sound Cask (right panel)

The Sound Cask has a nonagonal horizontal cross section, and the perimeter surfaces are set at three different angles in the vertical direction to avoid parallel surfaces except for the floor and ceiling planes. It has a 96-channel sound field reproduction system that uses a total of 96 full-range loudspeakers installed in all wall surfaces except for the floor plane.

3 Experiment 1:

We conducted research among audio and acoustics experts in order to examine the characteristics of the sound fields reproduced by the Sound Cask.

3.1 Method

A total of 54 adults participated in the experiment. They were expert listeners, either studying or working in the field of acoustics. We used seven different sound fields that were recorded by the BoSC recording system. After presentation of all stimuli, participants were asked to answer the characteristics of the Sound Cask compared to other conventional audio systems.

3.2 Results and Discussion

A total of 72 phrasings were obtained from 52 participants in the investigation. The all phrasings were classified in six semantic categories by social psychological methods. Table 1 shows the classification. From these results, it was clear that the most common positive comment was “a high presence”.

Table 1. The classification of the verbal reports in Experiment 1

Category group	Description	Positive	Negative	Neither
Sense of presence or realism	Being there, Natural	14	3	0
Localization	Perceive sounds position	10	7	1
Spatial reproducibility except for localization	Envelopment by sounds, Surround, Reverb	11	6	0
Sound quality	Frequency range, Volume	1	2	0
Other	Size of the system, Impression	1	11	0

4 Experiment 2: Psychological and physiological test

From the results of Experiment 1 with experts, it was suggested that the Sound Cask is able to present acoustical virtual environments with high reality. Then, we examine methods to measure “sense of presence” quantitatively and objectively.

4.1 Method

Twelve healthy adults participated in the experiment. Auditory stimulus recorded by the BoSC recording system was the sound of 7 men preceded by a guitar player approaching and moving away from the microphone while clapping. When we presented this auditory stimulus in the Sound Cask, it seemed that the listener is surprised to the approaching sounds in our other investigation. Then, we measured the skin conductance level (SCLs) and the blood volume pulse (BVP) that indicated physiological and psychological arousal response [3,4], as the physiological measure. Because the SCLs and the BVP are influenced by the sound intensity levels or the sound frequency characteristics, we have to set conditions so that the same characteristics between conditions are kept. Then, we set two conditions, the valid and the invalid condition.

Testing was conducted in the Sound Cask. Participants sat in the chair and listened to the auditory stimulus. While the stimulus was being presented, participants were measured his BVP and SCL. After each auditory stimulus ended, participants were asked to answer four questions: Question 1, Did you feel the sound stimulus was unnatural? Question 2, Did you feel as if you were there? Question 3, Did you feel as if someone passed by there? Question 4, What was the distance between you and the clapping man when he was closest to you? A 5-point rating scale was used: 1 was labeled “not at all”; 5 was labeled “very much”; and 3 was labeled “neither.” For Question 4, 1 was labeled “within 50 cm”; 2 was labeled “50–100 cm”; 3 was labeled “100–150 cm”; 4 was labeled “150–200 cm”; 5 was labeled “more than 200 cm.”

4.2 Results and Discussion

Figure 2 shows mean score and standard error of each questionnaire for all the participants. A repeated-measure ANOVA (2 stimulus conditions × 4 questions) was performed on the mean score of all participants. A main effect of the question and the interaction between the stimuli condition and the question was found (main effect: $F(3,33) = 11.71$, $p < 0.01$; interaction: $F(3,33) = 12.49$, $p < 0.01$). The post-hoc test revealed the significant differences between the valid and the invalid condition in Question 2: ($p < 0.01$), Question 3 ($p < 0.05$), and Question 4 ($p < 0.01$). These results show that participants perceived higher presence of spaces and objects for the valid condition than the invalid condition. This is consistent with previous studies stating that the accuracy of the spatialized sounds is important for presence in acoustic virtual environments [5].

The BVP across all participants decreased in the valid condition more than in the invalid condition (BVPs: $t(11) = 6.02$, $p < 0.05$). The SCLs increased in the valid condition more than in the invalid condition (SCLs: $t(11) = 6.02$, $p < 0.05$). These results suggest that the BVP and SCL were influenced by the spatial reproduced condition in the auditory environments.

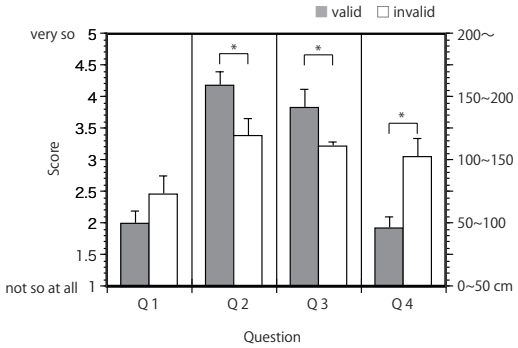

Figure 2. Results of psychological test in Experiment 2

5 Summary

In this study, we report two trials to evaluate listeners’ experience in an acoustic virtual environment in a quantitative and objective way by using the Sound Cask. In Experiment 1, we investigated characteristics of the Sound Cask using verbal reports from experts in audio engineering in order to clarify the differences of auditory impression between the Sound Cask and conventional audio systems. As a result, the experts pointed out localization performance and a clear sense of presence as advantageous features of the Sound Cask in the investigations. In Experiment 2, we examined the methods to quantify the sense of presence by psychological and physiological measures. As a result, both indexes showed different responses between conditions of spatial reproduction accuracy, suggesting the possibility to be the objective measures to evaluate presence or subjective experience in virtual acoustic environments.

Acknowledgements

This work was supported by CREST from Japan Science Technology Agency.

References

[1] Ikeda, Y., Ise, S. “Sound Cask: Music and voice communications system with three-dimensional sound reproduction based on boundary surface control principle” Proc. ICA 2013, 2013

[2] Ise, S. “A principle of sound field control based on the kirchhoff-helmholtz integral equation and the theory of inverse systems”. *Acoustica*, **85**, 78–87, 1999.

[3] Baumgartner, T., Volko, L., Esslen, M., Jancke, L. “Neural correlate of spatial presence in an arousing and noninteractive virtual reality: An EEG and Psychophysiology study” Cyberpsychol. Behav., **9**, 30-45, (2006)

[4] Barry, R.J., Clark A.R., McCarthy, R., Seikowitz, M., Rushby, J. A. “Arousal and activation in continuous performance task: An exploration of state effects in normal children” J. Psychohysiol., **19**, 91-99, (2005)

[5] Lombard, M., Ditton, T. “At the hear of it all: the concept of presence” J Computer-Mediated Com 3 (2), 1-43, (1997)