SUBJECTIVE ASSESSMENT FOR THE NUMBER OF CHANNEL SIGNALS TO REALIZE SOUND FIELD BASED ON WAVEFIELD SYNTHESIS

Toshiyuki Kimura, Kazuhiko Kakehi, Kazuya Takeda and Fumitada Itakura*

Center for Integrated Acoustic Information Research (CIAIR),
Nagoya University, Japan

ABSTRACT

It is very important to study the number of required channel signals in the sound field reproduction technique based on wavefield synthesis. Although there are many studies of objective affect such as wavefront accuracy, there are not so much studies of the subjective affect such as sound field perception. Therefore, the subjective assessment was designed in order to evaluate the number of required channel signals, that were synthesized by convolving a sound source to room transfer functions of free field, on the directional perception.

In the low-reverberation room, a circle loudspeaker array of which radius is 2m was set on the horizontal plane. The subject evaluated the directional perception of the sound image that was reproduced at the distance of 3 and 4m by playing simultaneously channel signals from the loudspeakers of which intervals were set at 10, 15, 20, 30 and 45 degrees, respectively. As a result, it was subjectively confirmed that 24 channel signals were enough to reproduce the directional perception of 5 degrees accuracy if the control area was the circle of radius 2m.

1. INTRODUCTION

Wavefield synthesis method [1, 2, 3] is the technique to reproduce the sound field as that in other place. A conceptual diagram is shown in Fig.1. When a microphone array placed at the control area records channel signals and a loudspeaker array placed at the reproduction area plays channel signals, wavefronts of the control area are synthesized in the reproduction area based on Huygens principle. Since it needs to transmit a great number of channel signals from the control area to the reproduction area, it is very important to evaluate the number of required channel signals to design the reproduction system.

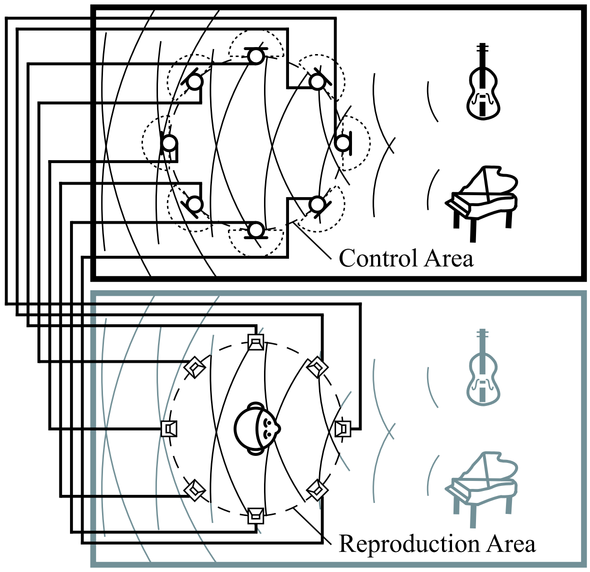

Fig. 1. Conceptual diagram of wavefield synthesis method

There are 2 approaches in the evaluation of the

*E-mail: kimura@cog.human.nagoya-u.ac.jp

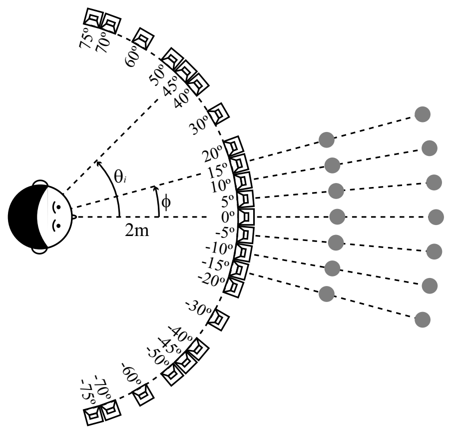

Fig. 2. Position of loudspeakers and sound images

number of required channel signals, e.g. objective or subjective one. Although there are many studies [4, 5] due to the objective approach, in which the physical accuracy of synthesized wavefronts is evaluated, there are not so much studies due to the subjective approach, in which the accuracy of the sound field perception is evaluated. In this paper, the number of channel signals is evaluated according to the subjective approach.

Sound field perception has dimensions as follows [6]: directional perception, distant perception and spatial impression. Although it is necessary to conduct the subjective assessment for each perception to evaluate the number of required channel signals, only the directional perception is treated in this paper.

2. SUBJECTIVE ASSESSMENT FOR DIRECTIONAL PERCEPTION

2.1. Environment

Subejctive assessment was performed in the low-reverberation room of which reverberation time is about 80ms. Twenty-three loudspeakers were set on the arc of radius 2m as shown in Fig.2. Grey circles indicate sound images reproduced by the loudspeaker array. A background room noise level was 25.0dB(A) and a sound pressure level was set to about 60dB(A) at the position of the subject. The subject's head was fixed by a head-rest and the subject was not able to see the loudspeakers by an acoustical transparent curtain in front of the loudspeakers.

2.2. Synthesis of Multi-channel Signals

White noise and speech of which the duration was 1sec were used as a dry source. Since the directional perception mainly depends on the direct sound from a sound source, a free space was assumed to caluculate ith channel signal $x_i(n)$ from the source signal $s(n)$.Then $x_i(n)$ is calculated as shown in Equation (1),

$$x_i(n) = \frac{d-r}{d_i} s(n - \text{round}\left(\frac{d_i F_s}{c}\right)), \quad (1)$$

where r(= 2m) is the distance between the subject and each loudspeaker , d(= 3 & 4m) is the distance between the subject and the sound source, F_s(= 48kHz) is a sampling frequency and c(= 340m/s) is a sound velocity. d_i, the distance between the sound source and the loudspeaker i, is calculated as shown in Equation (2),

$$d_i = \sqrt{d^2 + r^2 - 2dr\cos(\phi - \theta_i)}, \quad (2)$$

where ϕ and θ_i are the azimuth of the sound source and the ith loudspeaker.

Six experimental conditions about the number of channel signals are shown in Fig.3. In control condition (a), that is the ordinary condition of sound source localization, the subject listens the sound source from each loudspeaker. In other 5 conditions (b)-(f), the subject listens the sound image synthesized with 15, 11, 7, 5 and 3 loudspeakers, respectively. The direction was set from -15° to 15° at 5° intevals and the distance of the sound image was set on 3 and 4m.

2.3. Experimental Design

Subjects were 8 graduate students (4 males and 4 females). The design of the subjective assessment

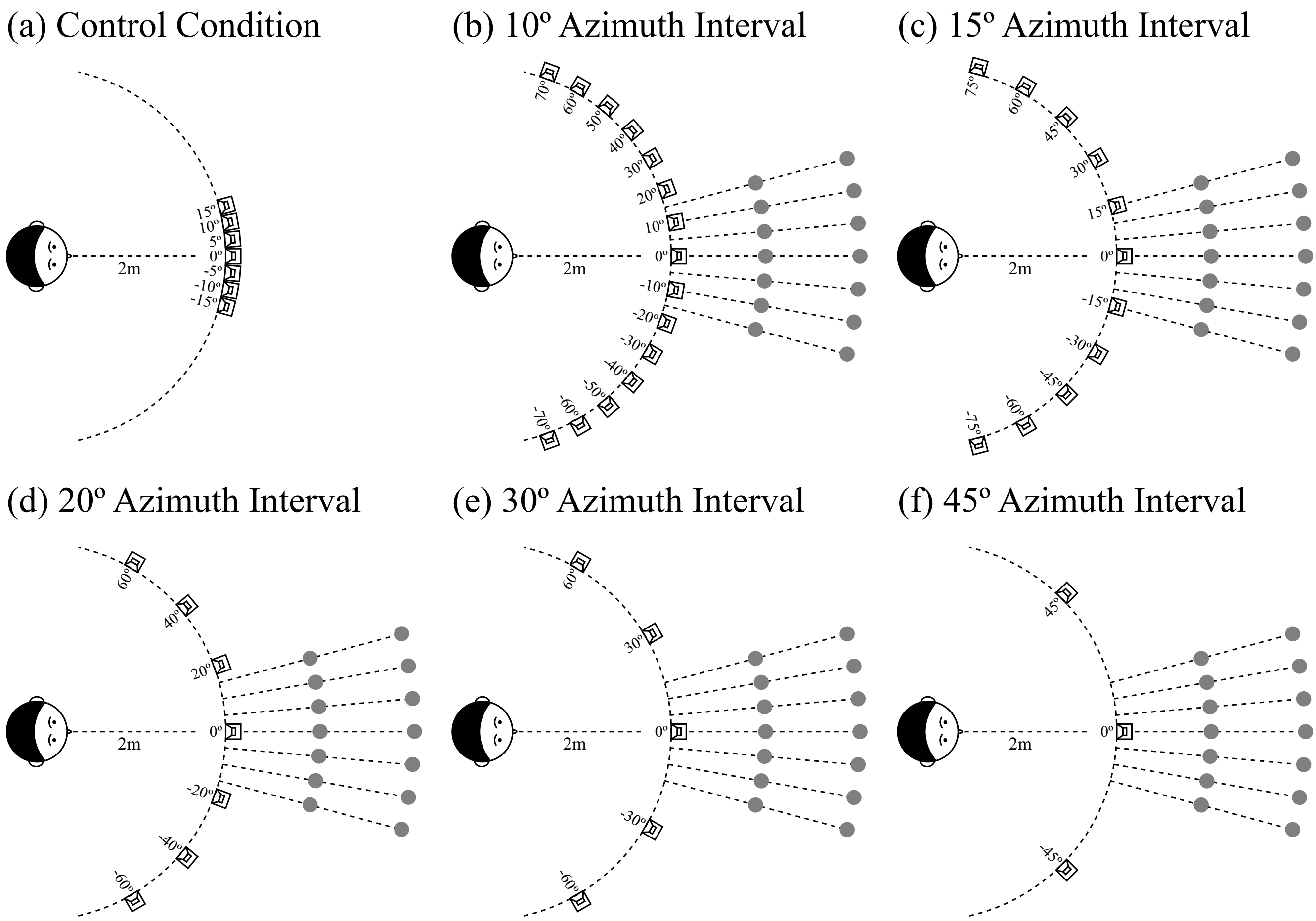

Fig. 3. Conditions of the number of channel signals

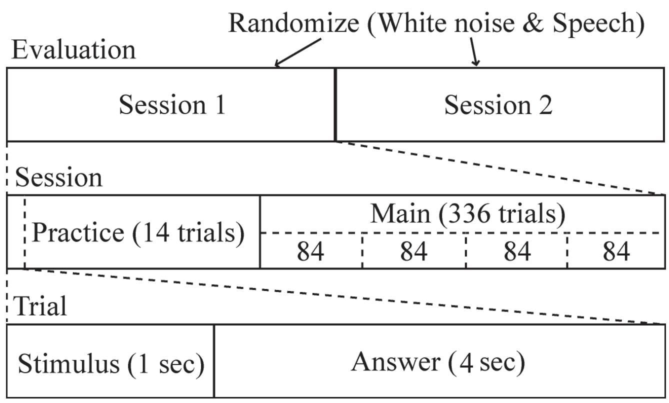

Fig. 4. Design of the subjective assessment

is shown in Fig.4. The subjective assessment was divided into 2 sessions for each dry source. The dry source's order was randomized in each subject. In each session, after 14 practice trials, 336 main trials were performed. Rest times were introduced in every 84 main trials. The breakdown of practice trials and main trials are shown in Table 1.

2.4. Experimental Procedure

The subject was instructed to report the direction of sound during 4sec after listening the stimulus 1-second long. Subjects reported tthe direction due to a scale which is placed in front of the subject and marked from -25° to 25° at 2.5° intervals.

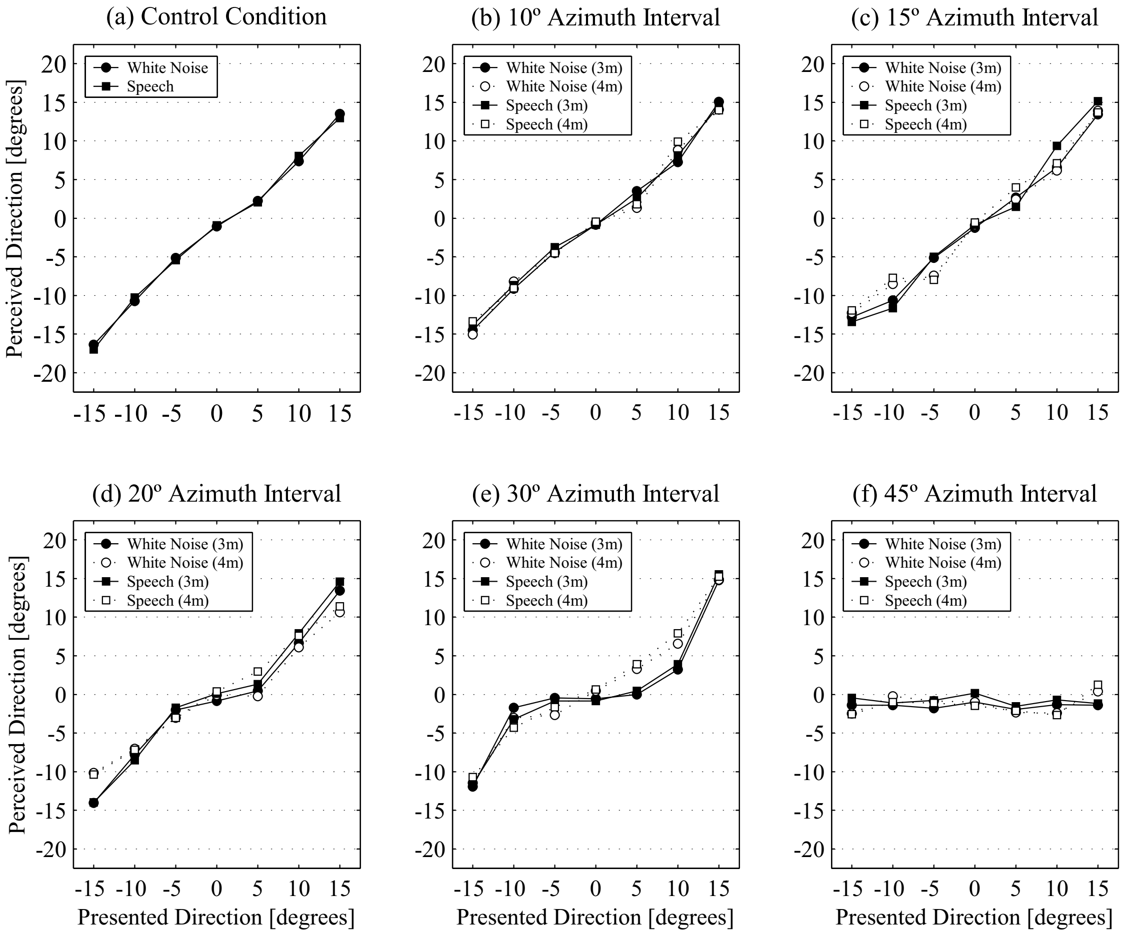

Fig. 5. Results of the localization experiment

2.5. Results and Discussions

Experimental results of localization are shown in Fig.5. In the control condition, perceived directions are almost same as the presented directions. In the 10° and 15° azimuth interval, perceived directions are near to those of the control condition. On the other hand, in 20°, 30° and 45° azimuth interval condiitons perceived directions tend to be biased towards 0°.

The reason of bias is explained by using the example of 45° azimuth interval as shown in Fig.6. The subject localizes the sound image which is synthesized based on channel singals from 3 (0°

Table 1. Breakdown of trials

	Number	Value
Practice (14)	$=$ 1 distance	3m
	$\times$ 7 directions	0°, ±5°, ±10 ° & ±15 °
	$\times$ 2 conditions	(a) and (b) shown in Fig.3
Main (336)	$=$ 2 distances	3 & 4m
	$\times$ 7 directions	0°, ±5°, ±10 ° & ±15 °
	$\times$ 6 conditions	From (a) to (f) shown in Fig.3
	$\times$ 4 repetitions	

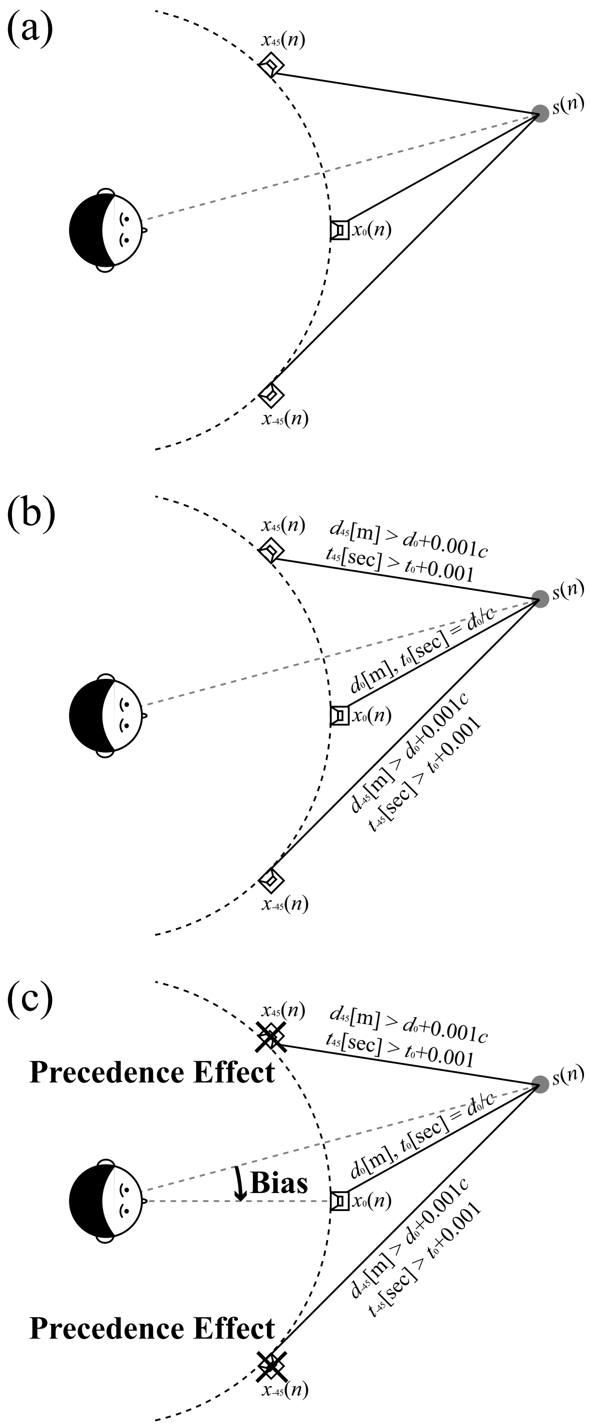

Fig. 6. Cause of the bias

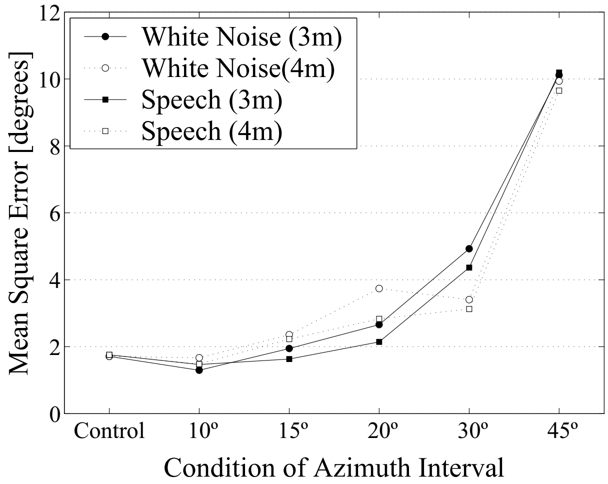

Fig. 7. Results of the mean square error

and ±45°) loudspeakers (Fig.6(a)). The channel signal from the 0° loudspeaker $x_0(n)$ is the fastest sound to be heard because the distance between the sound source and the 0° loudspeaker d_0 is the shortest. On the other hand, the other sounds to be heard from ±45° loudspeakers are delayed more than 1ms from the sound radiated by the 0° loudspeaker because the distance between the sound source and the ±45° loudspeakers $d_{\pm 45}$ is the 0.001c[m] longer than d_0 (Fig.6(b)). Because of the precedence effect of $x_0(n)$, only $x_0(n)$ contributes the perceptual localization (Fig.6(c)).

The accuracy threshold is discussed by calculating the mean square error between presented directions and perceived directions as shown in Equation (3),

$$MSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(D_i - D_i')^2}, \tag{3}$$

where D, D' and N are the presented direction, the perceived direction and the number of directions. Results of the mean square error is shown in Fig.7. When the azimuth interval is less than 15°, the MSE is same as the control condition at 2°. Thus, it is considered that the accuracy of 2 ° in terms of MSE is practical threshold. Fifteen degrees azimuth interval correspons to 24 loudspeakers on the circle of radius 2m.

3. CONCLUSION

The number of required channel signals in the directional perception was evaluated. As a result, it was subjectively confirmed that the mean square error of the practical accuracy threshold was about 2 degrees. It means that 24 channels is enough to realize the sound field inside a circle of 2m radius. The future works involve the evaluation of the number of required channel signals for the spatial impression and the development of the localization model that is able to predict the results of this subjective assessment.

4. REFERENCES

[1] M. Camras, “Approach to recreating a sounf field,” *J. Acoust. Soc. Am.*, **43**, pp.1425-1431 (1968).

[2] A. J. Berkhout, D. de Vries and P. Vogel, “Acoustic control by wave field synthesis,” *J. Acoust. Soc. Am.*, **93**, pp.2764-2778 (1993).

[3] M. A. Poletti, “A Unified Theory of Horizontal Holographic Sound Systems,” *J. Audio Eng. Soc.*, **48**, pp.1155-1182 (2000).

[4] O. Kirkeby and P. A. Nelson, “Reproduction of plane wave sound fields,” *J. Acoust. Soc. Am.*, **94**, pp.2992-3000 (1993).

[5] D. B. Ward and T. D. Abhayapala, “Reproduction of a Plane-Wave Sound Field Using an Array of Loudspeakers,” *IEEE Trans. SAP*, **9**, pp.697-707 (2001).

[6] M. Morimoto, “The relation between spatial impression and the precedence effect,” *Proc. of International Conf. on Auditory Display*, pp.297-306 (2002).