

# 少数のスピーカと波面合成法を用いた 三次元音場再生システムの提案及び性能評価

直江 宗紀<sup>†‡</sup> 木村 敏幸<sup>†‡</sup> 山肩 洋子<sup>†</sup> 勝本 道哲<sup>†</sup>

<sup>†</sup> 独立行政法人情報通信研究機構ユニバーサルメディア研究センター  
〒184-8795 東京都小金井市貫井北町 4-2-1

<sup>‡</sup> 東京農工大学工学府情報工学専攻ユビキタス&ユニバーサル情報環境専修  
〒184-8588 東京都小金井市中町 2-24-16

E-mail: <sup>†</sup> {nm.s512.ex, t-kimura, yamakata, katumoto}@nict.go.jp, <sup>‡</sup> 50007646206@st.tuat.ac.jp

**あらまし** 三次元音場を作る方法として波面合成法がある。従来の波面合成法は空間上に多数のスピーカを配置することで実現していたが、この場合だとシステムを構築するコストも高く、また、スピーカが聴取者の視線に入ってしまうので、スクリーン等と組み合わせた視聴覚システムを構築するのが困難になる。本研究では、立方体の頂点位置に配置した8個のスピーカと波面合成法を用いた三次元音場再生システムを提案し、数の少ないスピーカアレイでも性能が十分出ているかどうかを評価する。頭部を中心とした17方向のスピーカによる定位方向と、8個のスピーカアレイによる定位方向を比較したところ、立方体面の法線方向以外では定位方向は同じであることが分かった。

**キーワード** 三次元音場再生, スピーカアレイ, 波面合成法, 音像定位

## Proposal and Performance Evaluation of 3D Sound Field Reproduction System Using a Few Loudspeakers and Wave Field Synthesis

Munenori NAOE<sup>†‡</sup> Toshiyuki KIMURA<sup>†‡</sup> Yoko YAMAKATA<sup>†</sup> and Michiaki KATSUMOTO<sup>†</sup>

<sup>†</sup> Universal Media Research Center, National Institute of Information and Communication Technology  
4-2-1 Nukui-kitamachi, Koganei-shi, Tokyo, 184-8795 Japan

<sup>‡</sup> Department of Computer and Information Sciences, Graduate School of Engineering,  
Tokyo University of Agriculture and Technology 2-24-16 Naka-cho, Koganei-shi, Tokyo, 184-8588 Japan

E-mail: <sup>†</sup> {nm.s512.ex, t-kimura, yamakata, katumoto}@nict.go.jp, <sup>‡</sup> 50007646206@st.tuat.ac.jp

**Abstract** A conventional 3D sound field reproduction system using wave field synthesis places a lot of loudspeakers around the listener. However, since such a system is very expensive and loudspeakers come into the listener's field of vision, it is very difficult to construct an audio-visual system with it. We developed and evaluated a 3D sound field reproduction system using eight loudspeakers placed at the vertex of cube and wave field synthesis. We compared the sound localization of a loudspeaker array with that of seventeen loudspeakers placed around the listener and found that their localization capabilities were equivalent except the normal direction of cube's planes.

**Keyword** 3D sound field reproduction, Loudspeaker array, Wave field synthesis, Sound localization

### 1. はじめに

近年、聴覚バーチャルリアリティシステムのための立体音場再生技術が盛んに研究されている。この技術が実用化されると、遠隔通信会議システムや遠隔アンサンブルシステムのように、離れた場所にいる人々があたかも同じ会議室内で会議を行ったり、同じコンサートホール内で一緒に演奏したりしているように思わせることが可能になる。これらのシステムはテレビ電話のような従来のコミュニケーション手段とは比べ物にならないほどの

臨場感を与えることができるため、普及すれば社会全体において遠隔コミュニケーションがより一層活発になっていくことが期待される。

立体音場再生技術はマイクロホンで音を収録し、ヘッドホンもしくはスピーカで収録した音をそのまま再生するだけで聴取者に臨場感を感じさせることのできる技術である。中でも、波面合成法[1-8]は原音場に配置したマイクロホンアレイで音を収録し、再生音場に配置したスピーカアレイで収録した音をそのまま再生するもので、

Huygens の原理によって原音場の制御領域内の波面が再生音場の聴取領域内に忠実に再現される。この技術はバイノーラル[9]やトランスオーラル[10]とは異なり、聴取者が聴取領域内のどこに居ても構わないし、頭部を自由に動かしても自然な音像を得られるというところに特徴がある。

これまでに検討されていた波面合成システムは線形のスピーカアレイや聴取者の周りを水平面上に取り囲んだスピーカアレイによって構築されていた。しかし、これらのシステムでは二次元平面上の音場しか制御できないため、今までの研究では上下方向に関する音場が再生出来なかった。

そこで、上下方向の音場も制御する三次元音場を再生するために、スピーカアレイを三次元空間上に多数配置し、三次元の音場を制御するシステムも提案されている[11]。しかし、このシステムの場合だとシステムを構築するコストも高く、また、スピーカが聴取者の視線に入ってしまうので、スクリーン等と組み合わせた視聴覚システムを構築するのが困難になる。

一方、聴取者の聴覚性能に基づけば、非常に低い周波数でしか波面が忠実に合成されていなくてもシステムの構築に必要なマイクロホンやスピーカの数を少なくすることができることが報告されている[7]。ゆえに、聴取実験を実施すれば、スピーカの数が少なくても実用的なシステムを構築することができるようになると思われる。

そこで、本研究では、スピーカが聴取者の視線に入らないくらい数を少なくした場合でも三次元音場が再生できるようにするために、立方体の頂点位置に配置した 8 個のスピーカアレイと波面合成法を用いた三次元音場再生システムを提案する。また、提案したシステムでも聴覚性能が十分に出ているかどうかを定位実験によって評価する。

提案する三次元音場再生システムの構成を図 1 に示す。まず、図 1 の左側に示す立方体型マイクロホンアレイで音を収録する。次に、図 1 の右側に示す立方体型スピーカアレイで音を再生する。その際、マイクロホンとスピーカの位置は同じ位置に対応させる。その結果、マイクロホンアレイの中の三次元音場がスピーカアレイの中で再生されるので、マイクロホンアレイの上部で音を動かすと、スピーカアレイの中にいる聴取者は頭の上で音が動いているように感じるようになる。

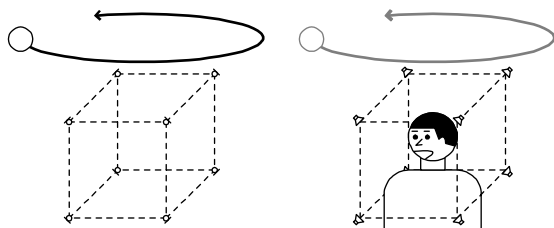


図 1 提案する三次元音場再生システム

Fig. 1 Proposed 3D sound field reproduction system

## 2. 定位実験

提案した三次元音場再生アレイシステムの聴覚性能を評価するために、定位実験を行う。

### 2.1. スピーカアレイの構築

まず、スピーカアレイに用いるスピーカを作製した。図 2 に示すようにスピーカボックスを設計し、スピーカユニット(AURASOUND: LCD034)をマウントすることによって図 3 に示すようなスピーカを作製した。

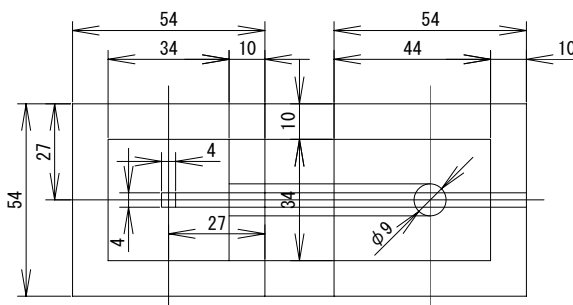


図 2 スピーカボックスの設計図

Fig. 2 Design of loudspeaker box

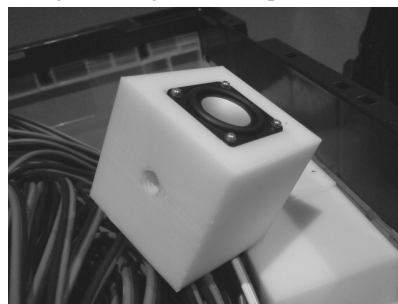


図 3 作製したスピーカユニットの概観図

Fig. 3 Image of the manufactured loudspeaker

スピーカを計 25 個作製し、図 4 に示す位置に配置した。その際、8 個のスピーカアレイを一辺 0.4 m の立方体の頂点位置に配置し、17 個の統制条件用スピーカを半径 1 m の球面上に配置した。実際に配置したスピーカアレイ及び統制条件用スピーカの状態を図 5 に示す。但し、図 5 中のスタンドの先に取り付けられた白いボックスは図に示したスピーカである。

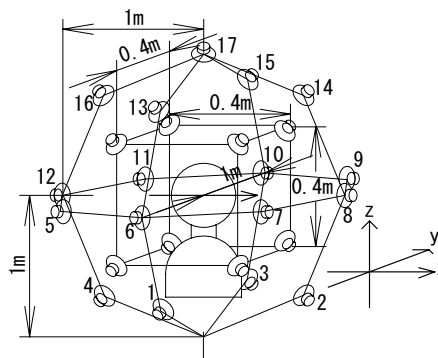


図 4 聴取者とスピーカの位置

Fig. 4 Position of a listener and loudspeakers



図 5 スピーカアレイと統制条件用スピーカ  
Fig. 5 Setup of the loudspeaker array and loudspeakers for control condition

## 2.2. マルチチャンネル信号の合成

本実験における、スピーカアレイで再生するマルチチャンネル信号は計算機上で合成した。音源から  $i$  番目のマイクロホン ( $i=1\dots 8$ ) までの室内インパルス応答  $g_i(n)$  は以下のように表される。

$$g_i(n) = \frac{1}{d_i} \delta \left\{ n - \text{round} \left( \frac{d_i F_s}{c} \right) \right\} \quad (1)$$

但し、 $F_s (=48 \text{ kHz})$  はサンプリング周波数、 $c (=340 \text{ m/s})$  は音速、 $\delta(n)$  は Dirac のデルタ関数、 $d_i (=|\mathbf{r}_0 - \mathbf{r}_i|)$  は音源から  $i$  番目のマイクロホンまでの距離である。また、 $\mathbf{r}_0$ 、 $\mathbf{r}_i$  は音源及び  $i$  番目のマイクロホンの位置ベクトルであり、実験では以下のように設定した。

$$\mathbf{r}_0 = (d \cos \theta \cos \phi \quad d \sin \theta \cos \phi \quad d \sin \theta \sin \phi)^T$$

$$\mathbf{r}_i = \begin{pmatrix} -0.2 & 0.2 & 0.2 & -0.2 & -0.2 & 0.2 & 0.2 & -0.2 & -0.2 \\ -0.2 & -0.2 & 0.2 & 0.2 & -0.2 & -0.2 & 0.2 & 0.2 & 0.2 \\ -0.2 & -0.2 & -0.2 & -0.2 & 0.2 & 0.2 & 0.2 & 0.2 & 0.2 \end{pmatrix} \quad (2)$$

但し、 $d (=1, 3 \text{ m})$  は音源から聴取位置までの距離、 $\theta$ 、 $\phi$  は聴取位置における音源の方位角及び仰角であり、本実験では表 1 に示すような 17 種類を設定した。

表 1 音源の方位角と仰角

Table 1 Azimuth and elevation angles of sound sources

Index	$\theta [^\circ]$	$\phi [^\circ]$	Index	$\theta [^\circ]$	$\phi [^\circ]$
1	-90	-45	10	90	0
2	0	-45	11	135	0
3	90	-45	12	180	0
4	180	-45	13	-90	45
5	-135	0	14	0	45
6	-90	0	15	90	45
7	-45	0	16	180	45
8	0	0	17	---	90
9	45	0			

音源信号を  $s(n)$  とすると、各マイクロホンによって収録されるチャンネル信号  $x_i(n)$  は以下のように合成される。

$$x_i(n) = D_i \{ g_i(n) * s(n) \} = \frac{D_i}{d_i} s \left\{ n - \text{round} \left( \frac{d_i F_s}{c} \right) \right\} \quad (3)$$

但し、 $*$  は畳み込み演算を表わす。また、 $D_i$  は  $i$  番目のマイクロホンの指向特性係数で、制御領域の外側からの音のみ収録するように設定する [8]。本実験では過去の研究成果 [8] から以下のような超指向性とした。

$$D_i = \begin{cases} \cos \theta_i & (|\theta_i| \leq 90^\circ) \\ 0 & (|\theta_i| > 90^\circ) \end{cases} \quad (4)$$

但し、 $\theta_i$  は  $i$  番目のマイクロホンにおける音源の入射角であり、以下のように定義される。

$$\theta_i = \cos^{-1} \left\{ \frac{\mathbf{r}_i \cdot (\mathbf{r}_0 - \mathbf{r}_i)}{|\mathbf{r}_i| |\mathbf{r}_0 - \mathbf{r}_i|} \right\} \quad (5)$$

また、半径 1 m の球面上に配置した 17 個のスピーカの中から選ばれた 1 個だけを再生する信号も統制条件として合成した。

## 2.3. 実験環境及び手順

実験は残響時間 180 ms、暗騒音レベル 23 dB(A) の部屋で行った。また、聴取位置における音圧レベルは 60 dB(A) とした。

被験者は聴力会話域正常な 6 名の男性及び 1 名の女性である。実験の流れを図 6 に示す。本実験では、音源として白色雑音と音声の 2 種類を用いた。2 種類の音源ごとに 3 条件 (統制条件、音像距離 1 m、音像距離 3 m) 及び 17 方向ごとの刺激を計 51 種類用意した。被験者には音源ごとに練習試行を 34 回行った後、本試行を 153 回行った。その際、51 回ごとに休憩を設けた。また、音源の順序は聴取者ごとにランダム化した。練習試行においては、17 (方向)  $\times$  2 (統制条件、音像距離 1 m)  $\times$  1 (繰り返し) = 34 回の試行を、本試行においては 17 (方向)  $\times$  3 (統制条件、音像距離 1 m、音像距離 3 m)  $\times$  3 (繰り返し) = 153 回の試行をランダム化し、順に被験者に提示した。

### Listening Test

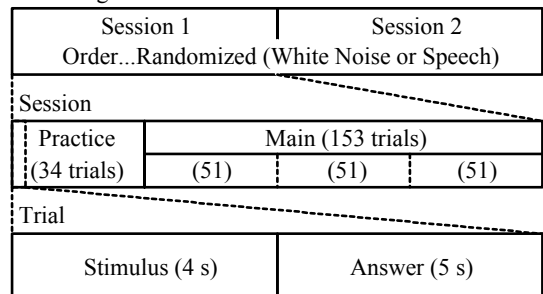


図 6 聴取実験の流れ図

Fig. 6 Flow chart of listening test

被験者には、図 7 に示すような知覚方向と番号の割り当て図を見せ、知覚した方向を割り当てた番号で解答用

紙に記入してもらった。また、音を聞いている際には頭部は動かしても良い状態とした。

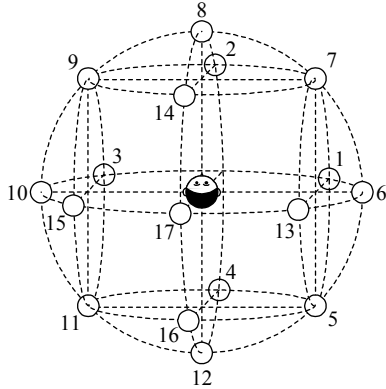


図 7 知覚方向と回答番号の対応図  
Fig.7 Relation between perceived directions and direction numbers

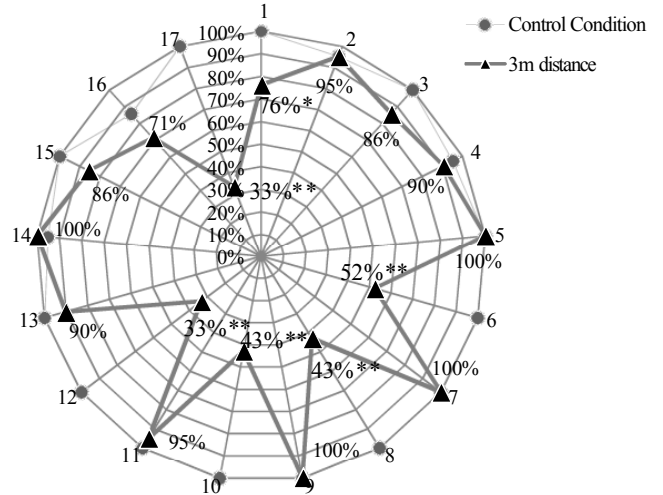


図 9 各方向の正答率(白色雑音, 音像距離 3m)  
Fig. 9 Accuracy rates of each direction in white noise (case of 3m distance)

## 2.4. 実験結果

統制条件、音像距離 1 m 及び音像距離 3 m の各条件における全方向の平均正答率を表 2 に示す。

表 2 から、統制条件における正答率はほぼ 100%である。これは音が再生されている間に頭を自由に動かすことが可能だったために後方や上下方向の音源を正しく定位できたためと考えられる。一方、提案システムの定位正答率は、全体的に 75%程度であり、絶対的な正答率として判断するならば、このシステムの定位性能は十分であると考えられる。

表 2 条件ごとの正答率

Table 2 Accuracy rate of each condition

	White Noise	Speech	Average
Control Condition	98%	96%	97%
1m distance	76%	77%	76%
3m distance	76%	74%	75%

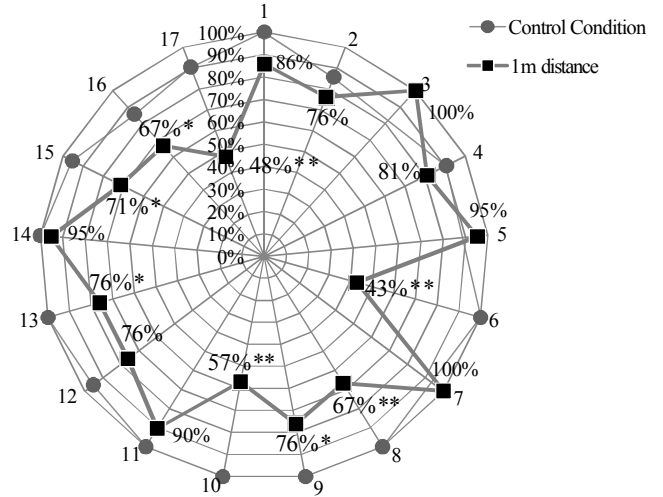


図 10 各方向の正答率 (音声, 音像距離 1m)  
Fig. 10 Accuracy rates of each direction in speech (case of 1m distance)

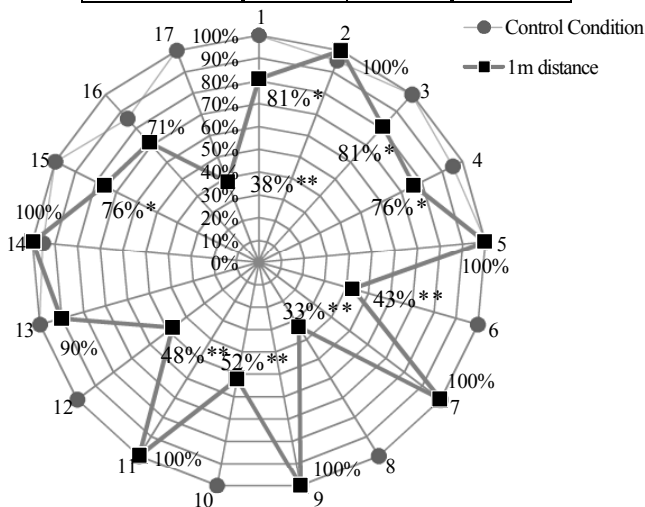


図 8 各方向の正答率(白色雑音, 音像距離 1m)  
Fig. 8 Accuracy rates of each direction in white noise (case of 1m distance)

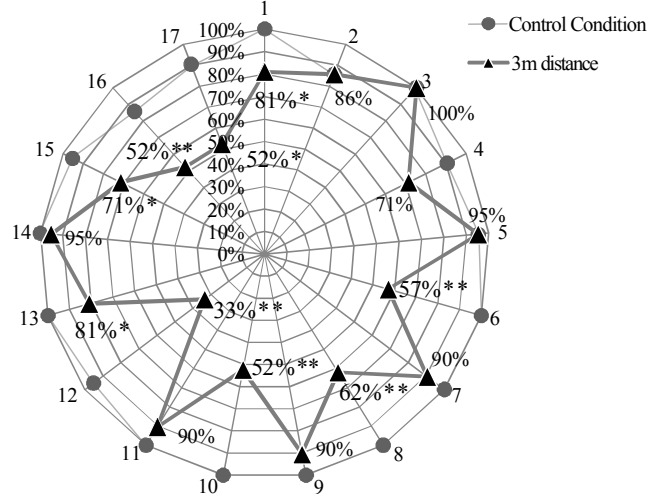


図 11 各方向の正答率 (音声, 音像距離 3m)  
Fig. 11 Accuracy rates of each direction in speech (case of 3m distance)

表 3 各方向のカイ二乗検定結果(白色雑音)

Table 3 Results of  $\chi^2$  test in white noise

Index	Control condition	1m distance	3m distance
1	100%	81%*	76%*
2	95%	100%	95%
3	100%	81%*	86%
4	95%	76%*	90%
5	100%	100%	100%
6	100%	43%**	52%**
7	100%	100%	100%
8	100%	33%**	43%**
9	100%	100%	100%
10	100%	52%**	43%**
11	100%	100%	95%
12	100%	48%**	33%**
13	100%	90%	90%
14	95%	100%	100%
15	100%	76%*	86%
16	86%	71%	71%
17	100%	38%**	33%**

表 4 各方向のカイ二乗検定結果(音声)

Table 4 Results of  $\chi^2$  test in speech

Index	Control condition	1m distance	3m distance
1	100%	86%	81%*
2	86%	76%	86%
3	100%	100%	100%
4	90%	81%	71%
5	95%	95%	95%
6	100%	43%**	57%**
7	100%	100%	90%
8	100%	67%**	62%**
9	100%	76%*	90%
10	100%	57%**	52%**
11	100%	90%	90%
12	95%	76%	33%**
13	100%	76%*	81%*
14	100%	95%	95%
15	95%	71%*	71%*
16	86%	67%*	52%**
17	90%	48%**	52%*

各方向における正答率を図 8-11 に、カイ二乗検定による結果を表 3-4 に示す。但し、表中の\*、\*\*は統制条件の正答率に比べて有意水準 5%及び 1%で有意差が見られることを示している。検定結果を見た場合、6、8、10、12、17 の方向では有意水準 1%で有意差が見られるので、正答率は統制条件よりも低くなっている。しかし、それ以外の方向では全体的に有意差は見られない。有意差が見られたとしても有意水準 1%までには達していない。従って、正答率は統制条件とほぼ同じであると考えられる。

ゆえに、これらの方向に関して提案システムは良い定位性能を示していることがわかる。

正答率が非常に低く、有意差も見られる方向に対してさらに細かく分析するために、正答率の低い 5 方向(6, 8, 10, 12, 17)における音源ごとの回答率を表 5-6 及び図 12-13 に示す。

提示方向が 6(被験者の右側)の時、誤答の多くは 7 や 13 であった。その中でも 13 に誤答する率が高かった。他の番号を誤答する率が 7, 13 以外にほとんど無いため、このシステムでは 6 から出ている音がやや上方向や前方方向に定位が偏っていると考えられる。

提示方向が 8(被験者正面)の時、14 と誤答する率が極めて高い。このことから、8 の方向を提示した時、やや上方向に定位が偏っているのではないかと考えられる。

提示方向が 10(被験者の左側)の時、3 と誤答する率が高い。このことから、10 の方向を提示した時、やや下方向に定位が偏っているのではないかと考えられる。

表 5 5 方向の回答率(白色雑音)

Table 5 Answer rates of five directions in white noise

		6		8		10		12		17	
		1m	3m	1m	3m	1m	3m	1m	3m	1m	3m
Perceived Index	1	0	10	0	0	0	5	0	0	0	0
	2	0	0	5	5	5	0	0	0	0	0
	3	0	0	0	0	29	48	0	0	0	0
	4	0	0	0	0	0	0	10	24	0	0
	5	5	0	0	0	0	0	0	5	0	0
	6	43	52	0	0	0	0	0	0	0	0
	7	14	14	0	0	0	0	0	0	0	0
	8	0	0	33	43	0	0	0	0	5	5
	9	0	0	0	0	14	0	0	0	0	0
	10	0	0	0	0	52	43	0	0	0	0
	11	0	0	0	0	0	0	5	0	0	0
	12	0	0	0	0	0	0	48	33	0	5
	13	38	24	0	0	0	0	0	0	0	0
	14	0	0	62	52	0	0	0	0	52	52
	15	0	0	0	0	0	5	0	0	0	0
	16	0	0	0	0	0	0	14	24	5	5
	17	0	0	0	0	0	0	24	14	38	33

Unit: [%]

表 6 5 方向の回答率(白色雑音)

Table 6 Answer rates of five directions in speech

		6		8		10		12		17	
		1m	3m	1m	3m	1m	3m	1m	3m	1m	3m
Perceived Index	1	10	5	0	0	0	0	0	0	0	5
	2	0	0	0	0	0	0	0	0	0	0
	3	5	10	0	0	33	38	0	0	0	0
	4	0	0	0	0	0	0	5	19	0	0
	5	10	5	0	0	0	0	5	0	0	0
	6	43	57	0	0	0	0	0	0	0	0
	7	19	10	0	0	0	0	0	0	5	0
	8	0	0	67	62	0	0	0	0	10	19
	9	0	0	0	0	5	5	0	0	0	0
	10	0	0	0	0	57	52	0	0	0	0
	11	0	0	0	0	0	0	0	5	0	0
	12	0	0	0	0	0	0	76	33	0	0
	13	10	14	0	0	0	0	0	5	0	0
	14	5	0	24	29	0	0	0	0	29	19
	15	0	0	0	0	5	5	0	0	0	0
	16	0	0	0	0	0	0	14	14	10	5
	17	0	0	10	10	0	0	0	24	48	52

Unit: [%]

提示方向が 12(被験者後方)の時、音声音源の音像距離 1 m の場合を除き、16 や 17 と誤答する率が高い。また、

音像距離が3mでは、4と誤答する率も高くなっている。このことから、音像の定位がやや上方向に偏っていたり、音像が上下方向にぼやけていたりするのではないかと考えられる。

提示方向が17(被験者真上)の時、14と誤答する率が極めて高い。このことから、定位が前方に非常に偏っていると考えられる。

以上のことから、これら5方向の回答率が低い原因について、いずれの方向も4個のスピーカから同じ音を同時に鳴らしている条件であることから、ファントム音源による音像のぼけと定位の偏りが発生しているのではないかと考えられる。また、本実験では聴取者に対し実験中首を動かすことを許可していたが、聴取者とスピーカアレイとの距離が非常に近かったために、少し首を動かすだけでスピーカアレイによる音像位置が変わってしまった可能性も考えられる。従って、頭を常に固定した状態で定位実験を行えば、5方向の正答率が改善される可能性がある。

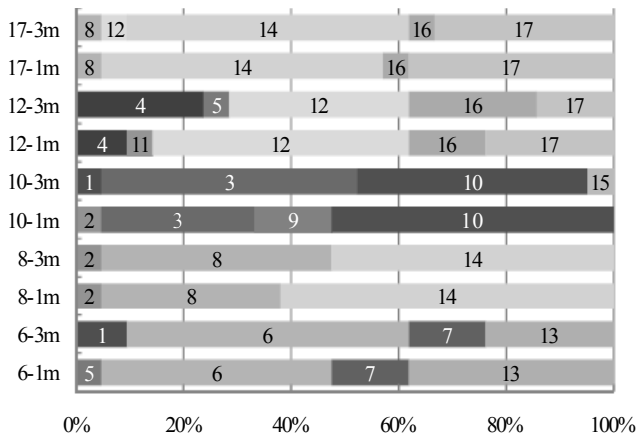


図 12 5 方向の回答率(白色雑音)

Fig. 12 Answer rates of five directions in white noise

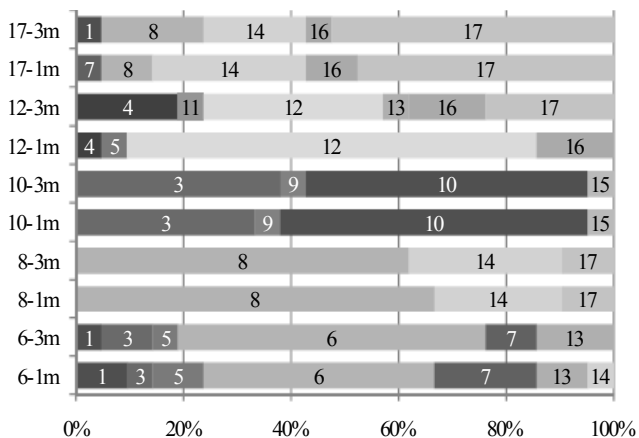


図 13 5 方向の回答率(音声)

Fig. 13 Answer rates of five directions in speech

### 3. まとめ

本研究では、スピーカ聴取者の視線に入らないよう少

なくした場合でも三次元音場が再生できるようにするため、8個のスピーカと波面合成法を用いたシステムを提案した。さらに、定位実験による聴覚性能の評価を行った。その結果、音像定位全体の性能としては、75%という結果が得られた。また、今回検討した17方向のうち、12方向については非常に良い定位性能が得られた。今後は、今回の実験では良い性能が得られなかった5方向の定位性能を改善する手法を提案し、定位実験によってその性能を評価していくことを予定している。

### 文 献

- [1] H. Fletcher, "Symposium on wire transmission of symphonic music and its reproduction on auditory perspective: Basic requirement," Bell System Technical Journal, vol.13, no.2, pp.239-244, April 1934.
- [2] 伊藤毅, 音響工学原論, コロナ社, 東京, 1973.
- [3] M. Camras, "Approach to recreating a sound field," J. Acoust. Soc. Am., vol.43, no.6, pp.1425-1431, June 1968.
- [4] A.J. Berkhout, D. de Vries, and P. Vogel, "Acoustic control by wave field synthesis," J. Acoust. Soc. Am., vol.93, no.5, pp.2764-2778, May 1993.
- [5] 及川靖広, 馬場貞如, 山崎芳男, "人間の聴覚を考慮した波面合成法による音場制御と音場再現," 日本バーチャルリアリティ学会論文誌, vol.5, no.3, pp.919-926, September 2000.
- [6] 平井徹, 本地由和, 池田雅弘, 清水寧, 川上福司, 東山三樹夫, "遠隔地音楽レッスン用 AVW システムとその原理," 電子情報通信学会技術研究報告, no.EA2000-17, pp.25-32, June 2000.
- [7] 木村敏幸, 箕一彦, 武田一哉, 板倉文忠, "波面合成におけるチャンネル数の主観的影響—音源が正面付近にある場合—," 日本バーチャルリアリティ学会論文誌, vol.10, no.2, pp.257-266, June 2005.
- [8] 木村敏幸, 箕一彦, "波面合成法による立体音場再生におけるマイクロホン及びスピーカの指向特性による波面の合成精度への影響," 日本バーチャルリアリティ学会論文誌, vol.12, no.2, pp.191-198, June 2007.
- [9] 岡部馨, "ダミーヘッドを用いた音場再生," 日本音響学会誌, vol.46, no.8, pp.650-656, August 1990.
- [10] 安藤四一, 設楽貞樹, 前川純一, 城戸健一, "電子計算機による室内音響設計法に関する二, 三の基礎的考察," 日本音響学会誌, vol.29, no.3, pp.151-159, March 1973.
- [11] 伊勢史郎, 豊田政弘, 榎本成悟, 中村哲, "深いコミュニケーションを可能とする空間創造の試み—プロジェクトの基本方針—," 日本音響学会研究発表会講演論文集, No. 2-1-20, pp. 585-586, March 2007.