

少数のスピーカと波面合成法を用いた 三次元音場再生における合成音生成方法による主観的影響*

◎直江宗紀, 木村敏幸 (NICT/東京農工大), 山肩洋子, 勝本道哲 (NICT)

1 はじめに

波面合成法^[1]はKirchhoff-Helmholtz積分方程式に基づき制御領域の波面を聴取領域で忠実に合成する三次元音場再生技術である。この技術によって聴取者はヘッドホンのような音響デバイスを装着せずに三次元音場を体験することができる。

これまでに提案されている三次元音場再生システム^[2]は、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、視聴覚システムを構築するのが困難である。

そこでスピーカが視界に入らない視聴覚システムを構築するために、我々はこれまで Fig. 1 に示すようにマイクロホンとスピーカを 8 個用いた立方体型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している^[3]。前報告^[3]では立方体型アレイの大きさやマイクロホンの指向特性を固定していたが、本報告ではそれらを変化させたときに定位精度がどのように変化するかを検討する。

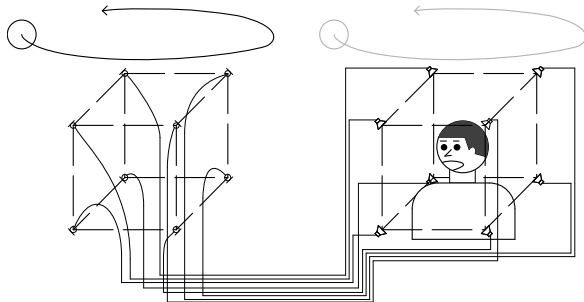


Fig. 1 Arrays of proposed 3D sound field reproduction system (Left: microphone array, Right: loudspeaker array)

2 定位実験

2.1 実験環境

実験は残響時間 115 ms、暗騒音レベル 20 dB(A)の部屋で行った。定位実験用に 8 個のスピーカアレイを Fig. 2 に示す位置に配置し、聴取位置をその中心に設置した。但し、Fig. 2 中に記載された a, b はそれぞれ立方体型アレイ一辺の幅 (40 cm or 50 cm) と聴取位置から立方体面までの長さ (20 cm or 25 cm) を表す。さらに、Fig. 3 に示すように、聴取位置を中心とした半径 1 m の球面上に 17 個の統制条件用スピーカを配置した。実際に配

置したシステムの状態を Fig. 4 に示す。聴取位置における音圧レベルは 60 dB(A)とした。

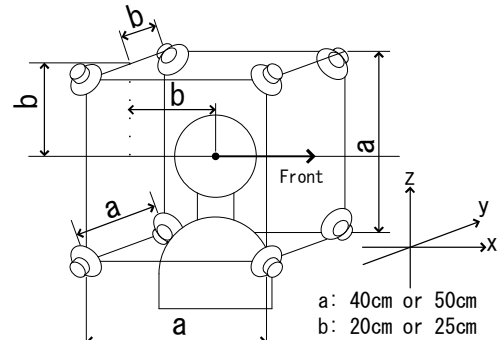


Fig. 2 Position of a listener and loudspeaker array in the localization test

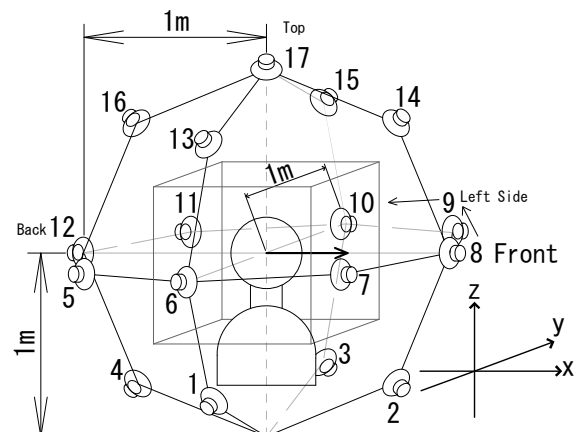


Fig. 3 Position of a listener and loudspeakers for control condition in the localization test



Fig. 4 Setup of the loudspeaker array and loudspeakers for control condition in the localization test

* The Subjective Effect of Synthesis Conditions in 3D Sound Field Reproduction System Using a Few Loudspeakers and Wave Field Synthesis, by Munenori Naoe, Toshiyuki Kimura (NICT/TUAT), Yoko Yamakata, and Michiaki Katsumoto (NICT)

2.2 マルチチャネル信号の作成

スピーカレイで再生するマルチチャ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($i=1..8$)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 kHz)はサンプリング周波数、 c (=340 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。

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

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

但し、 d (=1, 3 m)は音源から聴取位置までの距離、 θ, ϕ 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 Table 1 に示すような 17 種類を設定した。これは、Fig. 3 に示した 17 個の統制条件用スピーカの方角角及び仰角と同じである。

Table 1 Azimuth and elevation angles of sound sources in the localization test

Direction index	θ [°]	ϕ [°]	Direction index	θ [°]	ϕ [°]
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			

音源には長さ 4 秒の白色雑音を用いた。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[4]。本実験では以下の 2 種類の指向特性を設定した。

$$D_i = \frac{(1 + \cos \theta_i)}{2} \quad (4)$$

(単一指向性)

$$D_i = \begin{cases} \cos \theta_i & (|\theta_i| \leq 90^\circ) \\ 0 & (|\theta_i| > 90^\circ) \end{cases} \quad (5)$$

(超指向性)

但し、 θ_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。

$$\theta_i = \cos^{-1} \left\{ \frac{\mathbf{r}_i \cdot (\mathbf{r}_0 - \mathbf{r}_i)}{|\mathbf{r}_i| |\mathbf{r}_0 - \mathbf{r}_i|} \right\} \quad (6)$$

また、Fig. 3 に示す半径 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 実験手順

被験者は聴力会話域正常な 7 名の男性及び 3 名の女性である。実験の流れ図を Fig. 5 に示す。まずアレイ幅 40 cm 及び 50 cm の 2 つのセッションに分けた。セッションの提示順序は被験者ごとにランダム化した。それぞれのセッションでは、練習試行として 17(方向)×2(統制条件、単一指向性マイクロホンによる合成音像で距離 1 m)×1(繰り返し)=34 回の試行、本試行においては 17(方向)×5(統制条件、単一指向性マイクロホンによる合成音像で距離 1 m、単一指向性マイクロホンによる合成音像で距離 3 m、超指向性マイクロホンによる合成音像で距離 1 m、超指向性マイクロホンによる合成音像で距離 3 m)×2(繰り返し)=170 回の試行を順に被験者に提示した。その際、42 回及び 43 回ごとに休憩を設けた。なお、刺激の提示手順について被験者毎にランダム化した。

被験者には、Fig. 6 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答えてもらった。また、音を聞いている際には頭を自由に動かしてもらった。

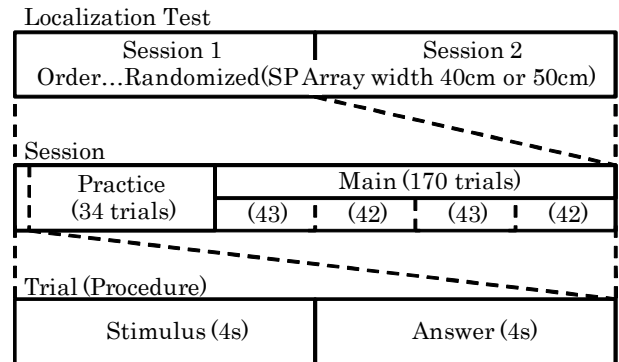


Fig. 5 Flow chart of the localization test

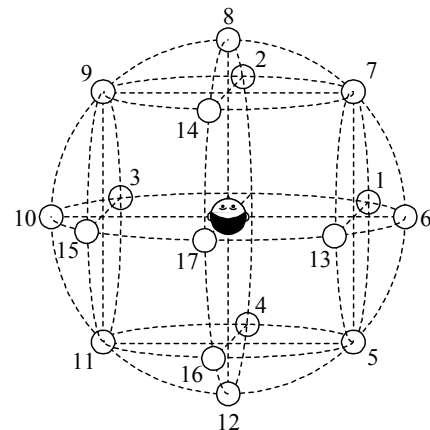


Fig. 6 Relation between perceived directions and direction indices

2.4 実験結果及び考察

実験結果の分析を行うにあたり、統制条件及び各合成条件における各方向の正答率を以下のように算出した。

$$\text{正答率 [\%]} = \frac{\text{正答数}}{\text{提示回数}} \quad (7)$$

統制条件及び各合成条件の正答率を Fig. 7 に示す。尚、Fig. 7 中のエラーバーは 95%信頼区間を示す。尚、Fig. 7 の 40 cm、50 cm は、それぞれ立方体型アレイの一辺の幅が 40 cm と 50 cm であることを、U, S は、それぞれ、Unidirectional microphone (単一指向性マイクロホン) と Shotgun microphone (超指向性マイクロホン) であることを示す。

Fig. 7 から、各合成条件における正答率の平均を比べると、アレイ幅が 40 cm で超指向性マイクロホンの時が一番高い正答率を出していることが分かる。

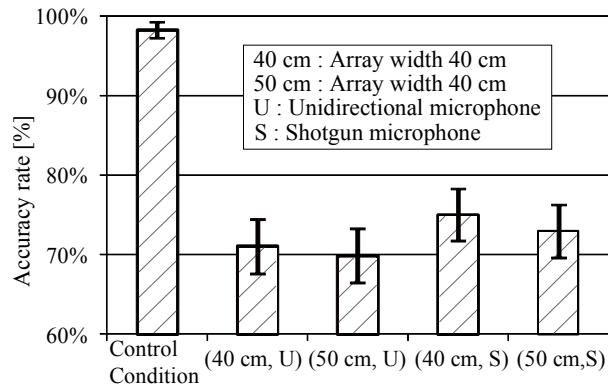


Fig. 7 Accuracy rates of each synthesis condition in the localization test

次に方向ごとの正答率と χ^2 検定の結果を Table 2 に示す。但し、*, **は χ^2 検定の結果、統制条件に比べて正答率に 5%有意差、1%有意差があることを示す。

Table 2 より、聴取者から見て水平方向(5-12)では、検定結果に違いは見られない。一方、聴取者から見て下部方向(1-4)の時、アレイ幅 40 cm かつ超指向性マイクロホンの場合以外の条件において 1%有意差があることが見られる。たとえば、方向 1 に関して、アレイ幅 40 cm、超指向性マイクロホンの条件では、唯一の 90% 以上という非常に高い正答率を出し、さらに有意差も見られない。しかし、それ以外の条件では、有意差が見られる。また、上部方向(13-17)では、下部方向(1-4)のような明確な違いは出ていないが、マイクロホン指向性の条件において差が出ていることが分かる。たとえば、方向 16 では、単一指向性マイクロホンの合成音の場合、両方向ともに 1%有意差が見られるのに比べ、超指向性マイクロホンの合成音の場合では、1%有意差が見られない。

つまり、Fig. 7 において合成条件間における正答率の違いが表れた原因は、下部方向及び上部方向における正答率の違いによるものであると分かる。

Table 2 Accuracy rates and results of the χ^2 -test of each direction in the localization test

Direction index	Control	U, 40 cm	U, 50 cm	S, 40 cm	S, 50 cm
1	100%	70%**	65%**	95%	88%*
2	98%	73%**	88%	85%*	88%
3	100%	75%**	55%**	88%*	65%**
4	98%	75%**	63%**	70%**	70%**
5	100%	98%	100%	100%	98%
6	100%	50%**	33%**	40%**	30%**
7	100%	100%	100%	100%	100%
8	100%	40%**	43%**	33%**	28%**
9	100%	100%	98%	98%	100%
10	100%	35%**	28%**	38%**	18%**
11	100%	95%	98%	95%	100%
12	100%	45%**	35%**	53%**	40%**
13	100%	80%**	88%*	90%*	83%**
14	98%	93%	90%	95%	95%
15	95%	78%*	88%	85%	90%
16	88%	50%**	55%**	65%*	78%
17	95%	53%**	65%**	48%**	73%**

ここで、アレイ幅 40 cm、超指向性マイクロホンの条件において、正答率が悪い方向についてさらに分析を行う。

Fig. 8 に各方向における回答率をグラフにしたものを示す。また、聴取者が回答した方向について、以下の式から回答率を算出した。

$$\text{回答率 [\%]} = \frac{\text{回答数}}{\text{提示回数}} \quad (8)$$

Fig. 8 の示すグラフは、刺激として提示した方向を中心として、聴取者がどの方向を答えたのかについて、各方向の回答率を円の大きさにより表現している。参考のために、右端に 25%, 50%, 75% の時の円を提示した。また、聴取者の顔の向きは、提示した方向を中心として見た場合の聴取者の顔の向きを示している。Fig. 8 の左端を例で言うと、方向 6 を中心として見たとき、聴取者は方向 8 が正面に来るように向いているため、聴取者の顔が左を向いているように見える。よって、Fig. 8 上の方向 6 の図では、聴取者の顔の向きも左向きとなっている。

Fig. 8 から、水平方向 (6, 8, 10, 12) における回答率を見ると、上下方向に回答がばらついていることが分かる。これらの方向から、音像を提示した場合、音像の水平方向は正しく知覚しているが、音像が上下方向にぼけているために、音像の上下方向を正しく知覚できず、上下方向に誤って回答していると考えられる。こ

これは、前報告^[3]でも述べたように、これらの方向では、4個のスピーカから同じ音を再生しているの、それによって合成されたファンタム音源による影響が大きいと考えられる。

次に、真上方向である、方向17では、主に前方方向である、方向14と誤答する率が非常に高いことがわかる。これは、4個のスピーカによるファンタム音源の影響から、音像が前方に偏っているためであると考えられるが、他に考えられる事象として、スピーカの配置による影響も考えられる。聴取者の耳に近い位置にあるスピーカアレイでは、聴取者の上の面で4つのスピーカが鳴った時、聴取者の後ろに位置しているスピーカからの音が聴取者の外耳などによって音が回折し、音が十分に耳に届いていないのではないかと考えられる。

以上のことから、音像定位の性能としては、アレイ幅やマイクロホン指向性の条件を用意して正答率を調べたが、5方向(6, 8, 10, 12, 17)ではどの条件でも、共通して正答率が悪くなる傾向があり、ファンタム音源による影響は前報告^[3]から改善することができなかった。しかし、一番性能が良い条件として、アレイ幅 40 cm かつ超指向性マイクロホンの場合であることが分かった。

3 まとめ

本報告では、先に提案した 8 個のマイクロホンとスピーカを用いたシステムにおいて、アレイの大きさやマイクロホンの指向性を変化させて合成音を作った場

合の主観的な変化について、定位実験による評価を行った。その結果、音像定位の性能としては、アレイ幅の幅が 40 cm で、かつ超指向性マイクロホンを使用した場合が、一番正答率が高いことが分かった。また、正答率の低い方向については、誤答は主に4個のスピーカによるファンタム音源による影響であることが分かった。今後は、正答率が低かった方向の定位性能を改善するための手法を検証していく予定である。

参考文献

- [1] A. J. Berkhout, D. de Vries, and P. Vogel, "Acoustic control by wave field synthesis," *Journal of Acoustical Society of America*, Vol.93, No. 5, pp. 2764-2778, May 1993.
- [2] 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.
- [3] 直江宗紀, 木村敏幸, 山肩洋子, 勝本道哲, "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", *電子情報通信学会技術研究報告*, No. EA-2008-1, pp. 1-6, April 2008.
- [4] 木村敏幸, 笈一彦, "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.

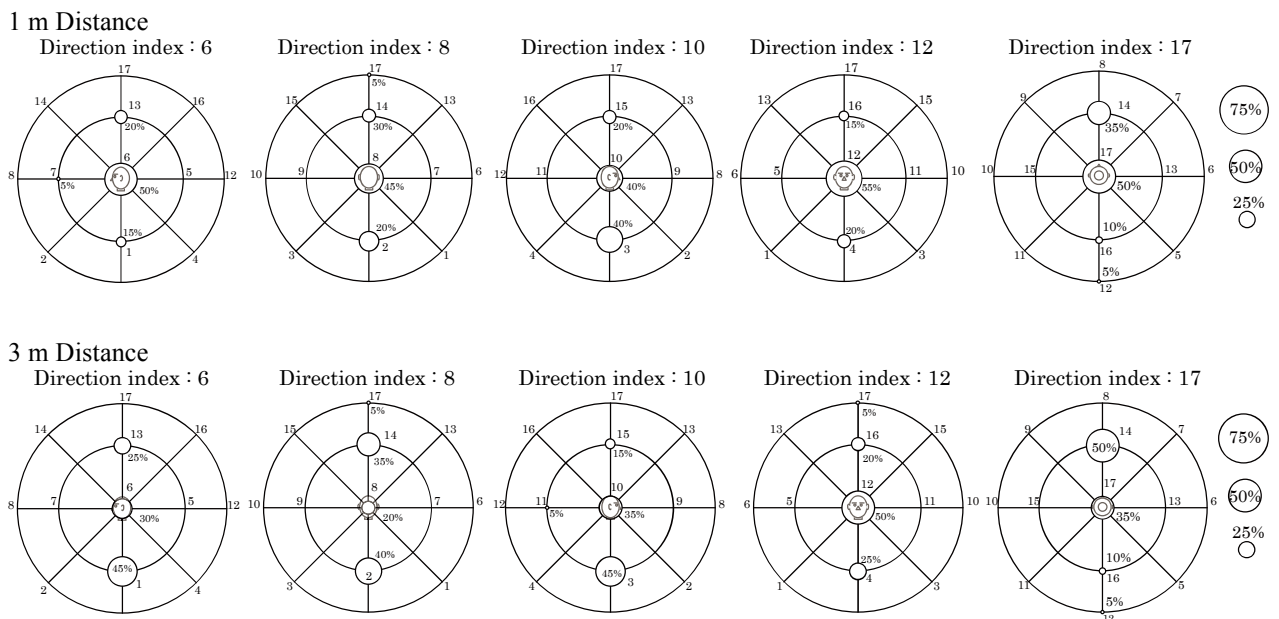


Fig.8 Answer rates of five presented directions in the localization test (Array width 40 cm, shotgun microphone)