

# 個人用三次元音場再生用収録システムの三次元定位の評価

木村 敏幸<sup>†</sup>

<sup>†</sup> 東北学院大学工学部情報基盤工学科 〒985-8537 宮城県多賀城市中央 1-13-1

E-mail: †t-kimura@m.ieice.org

**あらまし** これまでに提案してきた個人用三次元音場再生では、ショットガンマイクロホンを用いた収録システムを構築し、二次元定位性能をアンビソニックマイクと比較してきた。本報告では、三次元方向の定位性能を評価するために、インパルス応答を測定し、再生音を合成したうえで定位実験を実施した。その結果、提案システムの三次元定位性能がアンビソニックマイクよりも良いことが分かった。

**キーワード** 波面合成, 個人用三次元音場再生, インパルス応答測定, 定位実験, アンビソニックマイク

## 3D Localization Evaluation of Recording System for Personal 3D Sound Field Reproduction

Toshiyuki KIMURA<sup>†</sup>

<sup>†</sup> Department of Information Technology, Faculty of Engineering, Tohoku Gakuin University

1-13-1 Chuo, Tagajo-shi, Miyagi, 985-8537 Japan

E-mail: †t-kimura@m.ieice.org

**Abstract** In the proposed personal 3D sound field reproduction, the recording system using shotgun microphones have been constructed and the two-dimensional localized performance of the proposed system have been compared with that of an ambisonic microphone. In this report, in order to evaluate the three-dimensional localized performance of the proposed system, impulse responses were measured, the sound for replay was synthesized, and the localization test was performed. As a result, it was shown that the three-dimensional localized performance of the proposed system was better than that of an ambisonic microphone.

**Key words** wave field synthesis, personal 3D sound field reproduction, impulse response measurement, localization test, ambisonic microphone

### 1. はじめに

近年、三次元音場再生技術が盛んに研究開発されている。この技術はあたかも目の前に対象物があるかのような感覚（本報告ではこの感覚のことを「臨場感」と呼ぶ）を体験させることができるので、これらの技術を遠隔操作システムに適用すれば、現場での危機察知能力が向上したり、より効率的に作業できるようになることが期待される。

これまでに遠隔操作への応用を目的とした個人用三次元音場再生技術を提案し [1], 収録システムを構築して定位性能を評価してきた [2]。しかしながら、これまでに評価してきた定位性能は二次元方向に限定されていた。本報告では、提案する収録システムの三次元方向の定位性能を評価する実験を実施し、定位性能を従来のアンビソニックマイクと比較する。

### 2. 個人用三次元音場再生 [1]

個人用三次元音場再生の構成を Fig. 1 に示す。まず、Fig. 1 の左側に示すように、原音場において聴取者の頭部周辺に設定した立方体型の制御領域の頂点位置に 8 個の超指向性マイクロホンを設置し、音を収録する。その際、マイクロホンの指向特性（図中の灰色線の円）は制御領域の外側に向ける。次に、Fig. 1 の右側に示すように、再生音場において制御領域と同じ形に設定した聴取領域の頂点位置に配置した 8 個のスピーカから収録した音を直接再生する。その際、スピーカの位置はそれぞれの指向性マイクロホンの位置と同じになるように対応させる。その結果、制御領域内の三次元音場が聴取領域内に再現されるので、例えば制御領域の上側で音が動くと、聴取領域内にいる聴取者は Fig. 1 の右側に示すように頭の上で音が動いているように感じるようになる。

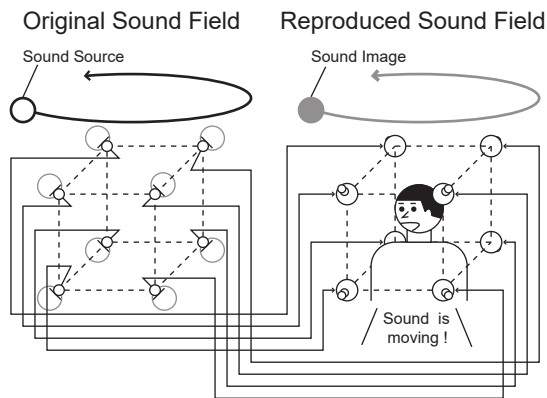


図1 個人用三次元音場再生の基本構成 [1]

Fig. 1 Basic configuration of the personal 3D sound field reproduction [1].

8 個のスピーカは聴取者頭部の水平面上には配置されていないので、聴取者の水平面方向の視界はスピーカに妨げられていない。遠隔操作システムにおける作業者の目線は水平面方向を向いていることが多いので、将来的には映像と組み合わせた遠隔操作システムも容易に構築することができるようになる。

### 3. 収録システム [2]

先行研究 [2] にて構築した収録用マイクロホンアレイを Fig. 2 に示す。映像と組み合わせた遠隔操作システムを構築するために、ショットガンマイクロホン配置するアレイ用フレームは水平面の前方及び後方に柱を設置せず、側方に配置した 4 枚のプレートで支える構造となっている。さらに、アレイ中心に映像用カメラを配置するために、4 枚のプレートはアレイ中心に設置しない構造となっている。

一辺が 0.4 m の立方体の頂点位置にショットガンマイクロホン (AZDEN : SGM-1000) の收音位置 (先行研究 [3] において提案した手法により、先端より 13.7 cm であると同等) を配置した。その後、ショットガンマイクロホンのペア (図中では 1 と 7, 2 と 8, 3 と 5, 4 と 6) を目視で一直線上に並べることで、ショットガンマイクロホンの指向特性の方向を制御領域の外側方向に調整した。

### 4. 評価実験

#### 4.1 インパルス応答の測定

まず評価実験において実験参加者に聞かせる再生音を作成するために、構築した収録用マイクロホンアレイとアンビソニックマイク (ZOOM : H3-VR) を用いてインパルス応答を測定した。

測定は実験室内において実施した。実験室の残響時間は 500 ms, 測定の際の暗騒音レベルは 32.6 dBA であった。音圧レベルは音源から 1 m 地点において 75.1dBA とした。Fig. 3 に示すように、収録用マイクロホンアレイの中心にアンビソニックマイクを配置し、計 25 カ所の音源位置を設定した。設定した音源位置の方位角及び仰角を Table 1 に示す。

測定の様子を Fig. 4 に示す。スピーカ (オーム電機 : ASP-204N-K) からはサンプリング周波数 48kHz の 65536 点 TSP 信

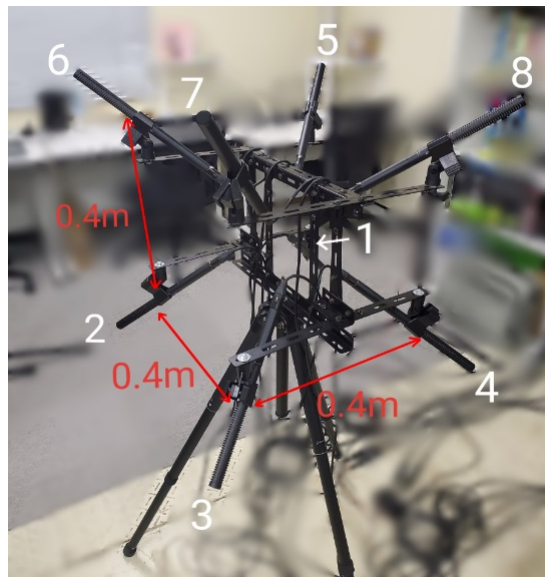


図2 個人用三次元音場再生の収録用マイクロホンアレイ [2]

Fig. 2 Recording microphone array of the personal 3D sound field reproduction [2].

表1 インパルス応答測定における音源位置の方位角及び仰角

Table 1 Azimuth and elevation angles of sound source position in the measurement of impulse responses.

Index	$\theta$	$\phi$	Index	$\theta$	$\phi$
1	-180°	-45°	14	45°	0°
2	-135°	-45°	15	90°	0°
3	-90°	-45°	16	135°	0°
4	-45°	-45°	17	-180°	45°
5	0°	-45°	18	-135°	45°
6	45°	-45°	19	-90°	45°
7	90°	-45°	20	-45°	45°
8	135°	-45°	21	0°	45°
9	-180°	0°	22	45°	45°
10	-135°	0°	23	90°	45°
11	-90°	0°	24	135°	45°
12	-45°	0°	25	—	90°
13	0°	0°			

号 (Time Stretched Pulse : 時間引き延ばしパルス) [4] を再生した。その際、9 回の同期加算をおこなった。1 か所の音源位置において測定が終わるごとに音源のスピーカを移動し、インパルス応答の測定を計 25 回繰り返した。

#### 4.2 再生音の合成

次に、測定したインパルス応答に先行研究 [5] で用いた 4 秒間の白色雑音と音声を畳み込んで 8 チャンネルの再生音を合成した。その際、測定環境の残響による定位性能への影響を除去するため、インパルス応答の直接波成分以降の振幅を 0 にすることで、測定環境の残響音を除去した。

構築した収録用マイクロホンアレイの再生音はアレイを構成するそれぞれのマイクロホンで測定したインパルス応答に音源を畳み込んだ。一方、アンビソニックの再生音は先行研究 [6] に基づき、測定したインパルス応答に音源を畳み込んだ後で下

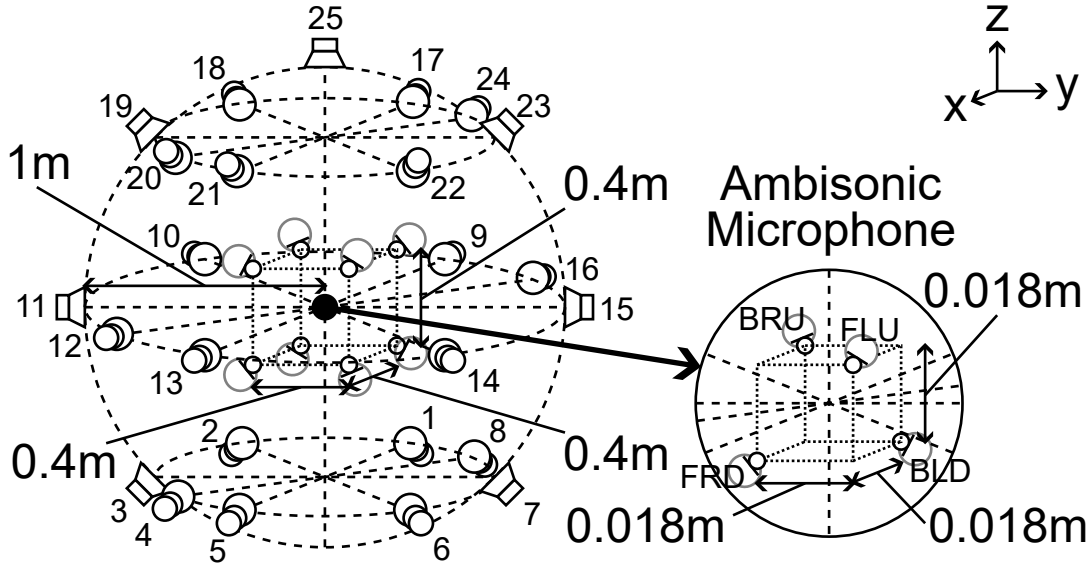


図3 インパルス応答測定におけるスピーカ及びマイクロホンの配置

Fig. 3 Placement of loudspeakers and microphones in the measurement of impulse responses.

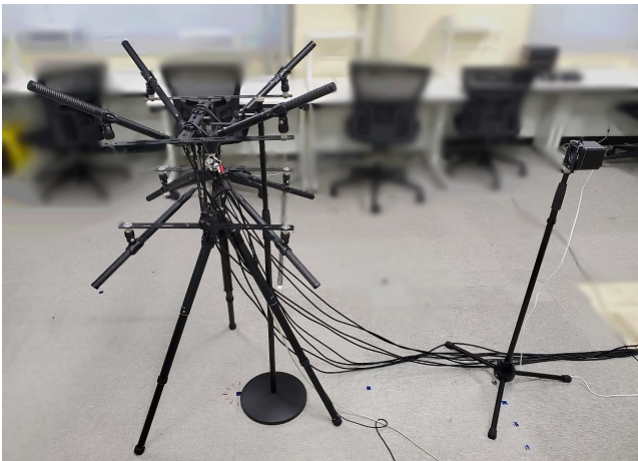


図4 インパルス応答測定の様子（音源位置番号9のとき）

Fig. 4 Image of the measurement of impulse responses  
(Sound position index: 9).

記の式に従って8チャンネルに展開した.

$$h_{BRD}(n) = w(n) + 0.707\{-x(n) - y(n) - z(n)\}, \quad (1)$$

$$h_{FRD}(n) = w(n) + 0.707\{x(n) - y(n) - z(n)\}, \quad (2)$$

$$h_{FLD}(n) = w(n) + 0.707\{x(n) + y(n) - z(n)\}, \quad (3)$$

$$h_{BLD}(n) = w(n) + 0.707\{-x(n) + y(n) - z(n)\}, \quad (4)$$

$$h_{BRU}(n) = w(n) + 0.707\{-x(n) - y(n) + z(n)\}, \quad (5)$$

$$h_{FRU}(n) = w(n) + 0.707\{x(n) - y(n) + z(n)\}, \quad (6)$$

$$h_{FLU}(n) = w(n) + 0.707\{x(n) + y(n) + z(n)\}, \quad (7)$$

$$h_{BLU}(n) = w(n) + 0.707\{-x(n) + y(n) + z(n)\} \quad (8)$$

ただし,  $w(n)$ ,  $x(n)$ ,  $y(n)$ ,  $z(n)$  は以下の式のように合成される [7].

$$w(n) = \{g_{FLU}(n) + g_{FRD}(n) + g_{BLD}(n) + g_{BRU}(n)\} * s(n), \quad (9)$$

$$x(n) = \{g_{FLU}(n) + g_{FRD}(n) - g_{BLD}(n) - g_{BRU}(n)\} * s(n), \quad (10)$$

$$y(n) = \{g_{FLU}(n) - g_{FRD}(n) + g_{BLD}(n) - g_{BRU}(n)\} * s(n), \quad (11)$$

$$z(n) = \{g_{FLU}(n) - g_{FRD}(n) - g_{BLD}(n) + g_{BRU}(n)\} * s(n) \quad (12)$$

ここで,  $s(n)$  は音源信号,  $*$  は畳み込み演算,  $g_{FRD}(n)$ ,  $g_{BLD}(n)$ ,  $g_{BRU}(n)$ ,  $g_{FLU}(n)$  は Fig. 3 中の4本のアンビソニックマイクで測定したインパルス応答を表す.

### 4.3 定位実験

合成した再生音を用いて定位実験を実施した. 実験は4.1節におけるインパルス応答の測定の際に使用した実験室内の一角にて実施した. スピーカユニット (ELECOM: MS-P06A より流用) を密閉型エンクロージャ (オーム電機: ASP-204N-K より流用) に取り付けられたスピーカを音場再生用のスピーカアレイに使用した. 実験参加者は Fig. 5 に示されるスピーカアレイ (1辺が40cm四方の頂点にスピーカを配置) 内の椅子に座り, ヘッドマウントディスプレイ (Microsoft: HoloLens2) を装着した. 実験参加者の頭部中心の高さはアレイ中心の高さと同じ (地面から1.4m) にし, 音圧レベルはアレイ中心において約70 dBA に設定した.

実験の間, ヘッドマウントディスプレイには Fig. 6 に示すような三次元仮想空間の映像を Unreal Engine 4.25 で作成し, 提示した. 三次元仮想空間では実験参加者の頭部位置にカメラを固定し, 番号が振られた球形のオブジェクトを Table 1 に示す方向に向かって計25個配置した. その際, カメラ位置からオブジェクトまでの距離は10mとした.

実験参加者は10名である. 実験参加者は実験開始前に実験



図5 定位実験におけるスピーカアレイ  
Fig. 5 Loudspeaker array in the localization test.

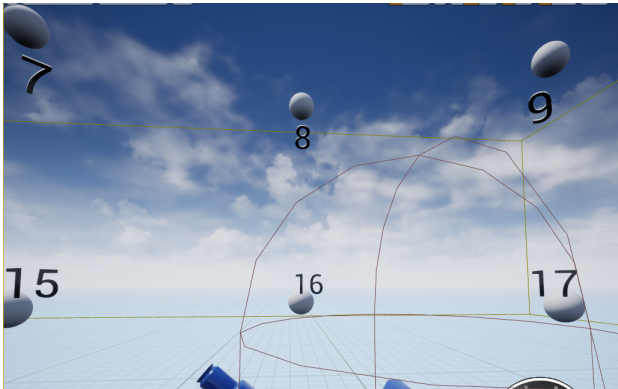


図6 定位実験における三次元仮想空間  
Fig. 6 3-D virtual space in the localization test.

参加者に対して実験内容に関する説明を受けた。実験内容に関する説明においては、実験参加者は自分の意思で実験に参加できることや実験参加者の個人情報保護されることを説明した。説明終了後、実験参加者が同意した場合は、実験参加者は用意した同意書に署名した。

Fig. 7に実験の流れ図を示す。実験は音源（白色雑音、音声）ごとにセッションを設定した。実験参加者には各セッションにおいて、4回の練習試行の後に100回（=25方向×2収録装置×2繰り返し）の本試行を実施した。セッション、音源の方向、収録装置の提示順序は実験参加者ごとにランダム化した。実験参加者には「試行ごとに音を4秒間聞き、音の鳴っている方向の番号を回答する」ように教示し、実験実施者は実験参加者の回答した番号を記録した。

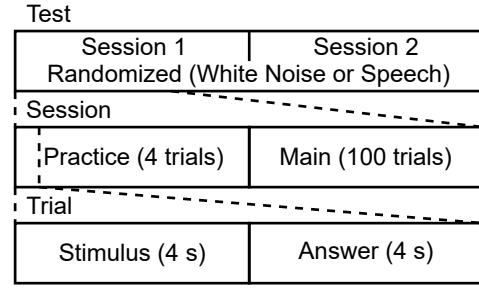


図7 定位実験の流れ図  
Fig. 7 Flowchart of the localization test.

#### 4.4 定位実験結果

まず、知覚した音源の方向に基づいて定位実験の結果を分析する。この場合、知覚方向を方位角と仰角で評価することが考えられるが、提示方向が真上（仰角 $\pm 90^\circ$ ）の場合は方位角が定義されていないので、知覚方向が真上以外だった場合に音源が真上からどの方向に変位しているのかを仰角だけで分析することができない。従って、提示方向が真上の場合も含めた定位実験の知覚方向を分析しようとする場合は、全ての提示方向を同じ条件で分析しているとみなせる座標系に実験結果を変換する必要がある。

そこで、先行研究[1]と同様に、下記の式に従って極座標系を回転させ、回答方向の方位角及び仰角 $(\theta, \phi)$ を水平角及び垂直角 $(\theta', \phi')$ に変換する。変換した水平角は聴取者が提示方向に頭部を向けた時の音源の左右方向の変位に相当し、変換した垂直角は音源の上下方向の変位に相当する。

$$\theta' = \tan^{-1} \frac{y'}{x'} \quad (13)$$

$$\phi' = \sin^{-1} \frac{z'}{\sqrt{x'^2 + y'^2 + z'^2}} \quad (14)$$

但し、変換した回答方向の三次元座標 $(x', y', z')$ は下記のように定義される。

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \mathbf{R}_y(-\phi_0) \mathbf{R}_z(-\theta_0) \begin{pmatrix} \cos \theta \cos \phi \\ \sin \theta \cos \phi \\ \sin \phi \end{pmatrix} \quad (15)$$

$$\mathbf{R}_y(-\phi_0) = \begin{pmatrix} \cos \phi_0 & 0 & \sin \phi_0 \\ 0 & 1 & 0 \\ -\sin \phi_0 & 0 & \cos \phi_0 \end{pmatrix} \quad (16)$$

$$\mathbf{R}_z(-\theta_0) = \begin{pmatrix} \cos \theta_0 & \sin \theta_0 & 0 \\ -\sin \theta_0 & \cos \theta_0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (17)$$

但し、 $\mathbf{R}_z(-\theta_0)$ は頭部及び椅子の水平方向の回転操作、 $\mathbf{R}_y(-\phi_0)$ は頭部の上下方向の回転操作に相当し、 $\theta_0$ と $\phi_0$ は提示方向の方位角及び仰角を示す。なお、変換した垂直角 $\phi'$ が $\pm 90^\circ$ になった場合には、変換した水平角 $\theta'$ のデータはこの後の集計には用いなかった。

アンビソニックマイクの変換した水平角及び垂直角の結果をFig. 8に示す。縦及び横方向のエラーバーは水平角及び垂直角



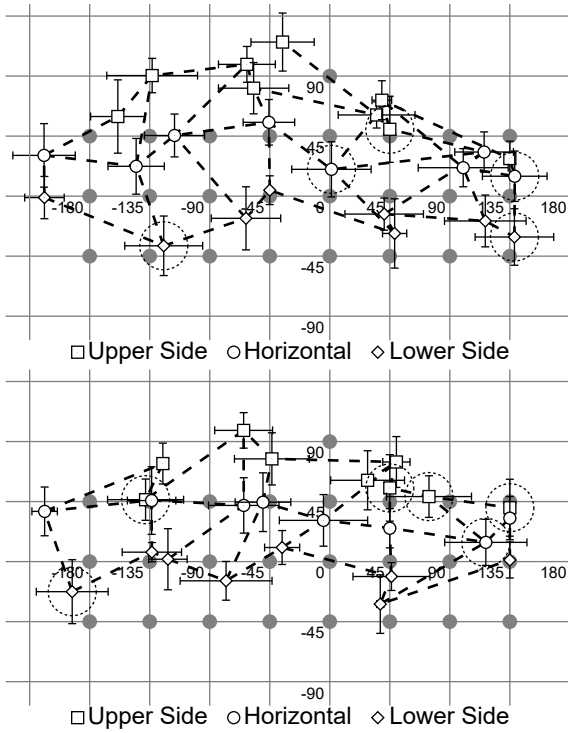


図8 定位実験におけるアンビソニックマイクの知覚方向の結果  
(上：白色雑音，下：音声)  
Fig. 8 Results of perceived directions of the ambisonic microphone in the localization test (Upper: White Noise, Lower: Speech).

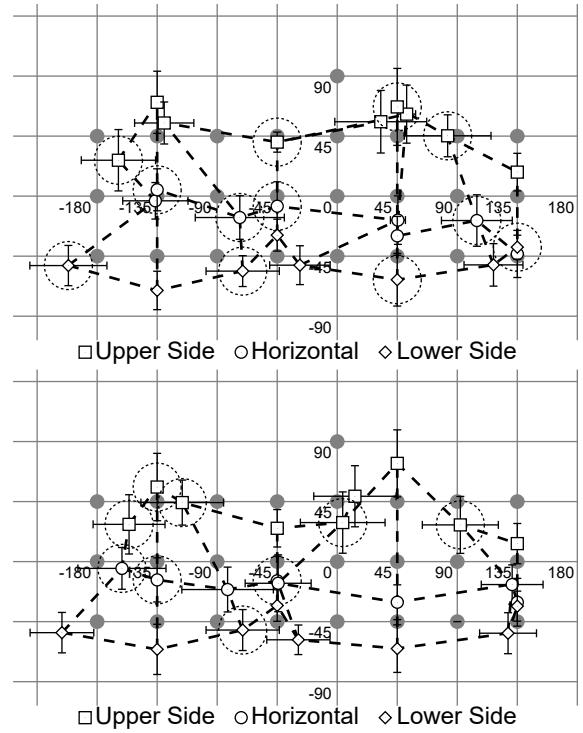


図9 定位実験における提案する収録システムの知覚方向の結果  
(上：白色雑音，下：音声)  
Fig. 9 Results of perceived directions of the proposed recording system in the localization test (Upper: White Noise, Lower: Speech).

の95%信頼区間を表す。但し、変換した水平角及び垂直角の平均値は常に0°に近い値となるが、図では提示方向の上下及び左右の関係を分かり易くするために、平均は提示方向の方位角及び仰角の分だけ横及び縦方向にシフトし、方位角もしくは仰角が同じ値で互いに隣り合っている提示方向同士を黒色の太い点線で結んでいる。さらに、縦及び横方向のエラーバーの中に提示方向を示す灰色の丸が入っている（つまり、変換した水平角及び垂直角ごとに母平均のt検定を実施した結果、知覚方向と提示方向との間に有意差が見られない）提示方向は黒色の破線で囲んでいる。

どちらの音源の場合においても、黒色の破線で囲まれている提示方向の数が少ないため（白色雑音：5，音声：6），アンビソニックマイクにおいては良い定位性能を示しているとは言えない。これはアンビソニックは音場の球面調和解析における0次と1次の項のみを用いた音場再生手法であることと、理想的にはアレイの中心に配置すべきアンビソニックマイクの各マイクホロンが中心に配置できないためであると考えられる。

一方、提案する収録システムの変換した水平角及び垂直角の結果をFig. 9に示す。どちらの音源の場合においても、黒色の破線で囲まれている提示方向の数がアンビソニックよりも多いため（白色雑音：12，音声：9），提案する収録システムの定位性能はアンビソニックマイクよりも良いと言える。これは先行研究[1]に従ってショットガンマイクロホンを1辺が0.4mの立方体の頂点位置に配置したことにより、スピーカアレイ内に音場が再現されたためであると考えられる。

さらに定量的な考察を行うため、実験条件ごとに正答率を

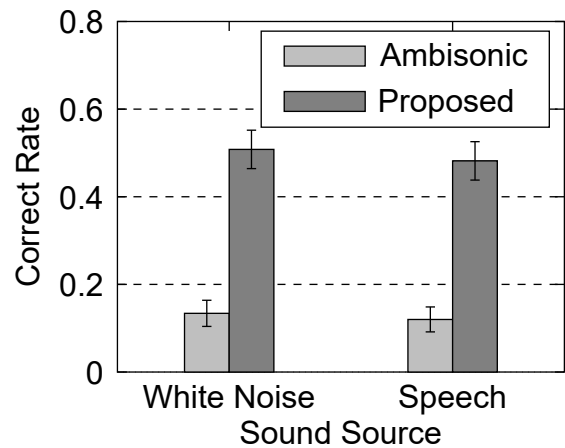


図10 定位実験における正答率の結果  
Fig. 10 Results of the correct rates in the localization test.

算出する。その際、回答数は500（=25方向×2繰り返し×10名）、正答数は提示方向に正しく回答した回数とした。実験条件ごとの正答率をFig. 10に示す。js-STAR[8]を用いて音源ごとにFisherの正確確率検定を実施したところ、ともに0.1%水準において有意差が見られた。ゆえに、正答率の観点においても、提案する収録システムの定位性能はアンビソニックマイクよりも良いと言える。

## 5. まとめ

本報告では、これまでに提案してきた個人用三次元音場再生のための収録システムの三次元方向の定位性能を評価するため

に、インパルス応答を測定し、再生音を合成したうえで定位実験を実施し、定位性能を従来のアンビソニックマイクと比較した。その結果、提案する収録システムの定位性能はアンビソニックマイクよりも良いことが分かった。

今後の課題としては、今回提案した収録システムを映像システムと組み合わせた遠隔操作システムを構築し、その性能を評価することが挙げられる。一方で、視聴覚の定位性能に基づけば三次元音場再生に必要な条件が緩和される可能性があるため、ヘッドマウントディスプレイとスピーカアレイを組み合わせた視聴覚提示システムを用いて検討することが挙げられる。

#### 謝辞

本研究の遂行にあたり、インパルス応答の測定、再生音の合成及び定位実験を実施して下さった平田憲汰氏に感謝の意を申し上げます。

#### 文 献

- [1] 木村敏幸, “8 個の指向性マイクロホンを用いた波面合成技術のコンセプトに基づいた個人用コンパクト三次元音場再生システム,” 電子情報通信学会論文誌, vol.J97-A, no.4, pp.284–294, April 2014.
- [2] 羽生史, 木村敏幸, “個人用三次元音場再生システム用収録装置の主観評価,” 電子情報通信学会技術研究報告, no.EA2020-81, pp.128–133, March 2021.
- [3] 木村敏幸, 萩野広佑, “超指向性マイクロホンの收音位置同定手法,” 電子情報通信学会技術研究報告, no.EA2017-27, pp.1–4, Aug. 2017.
- [4] 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, vol.97, no.2, pp.1119–1123, Feb. 1995.
- [5] T. Kimura and H.Ando, “3d audio system using multiple vertical panning for large-screen multiview 3d video display,” ITE Transactions on Media Technology and Applications, vol.2, no.1, pp.33–45, Jan. 2014.
- [6] D.G. Malham and A. Myatt, “3-d sound spatialization using ambisonic techniques,” Computer Music Journal, vol.19, no.4, pp.58–70, Winter 1995.
- [7] K. Farrar, “Soundfield microphone,” Wireless World, vol.85, pp.48–51, Oct. 1979.
- [8] js-STARWebsite. <https://www.kisnet.or.jp/nappa/software/star/>.