

個人用三次元音場再生システム用収録装置の主観評価

羽生 史[†] 木村敏幸[‡]

[†] 東北学院大学大学院工学研究科電気工学専攻 〒985-8537 宮城県多賀城市中央 1-13-1

[‡] 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: [†] s1994208@g.tohoku-gakuin.ac.jp, [‡] t-kimura@m.ieice.org

あらまし 我々はこれまでに 8 個の指向性マイクロホンで収録した音を 8 個のスピーカで再生する個人用三次元音場再生システムを提案してきた。本報告では、収録装置を構築するために、收音位置が同定されたショットガンマイクロホンを用いたマイクロホンアレイを提案し、提案システムで収録した場合の音の定位方向をアンビソニックマイクと比較した。その結果、提案システムはアンビソニックマイクよりも音像の提示方向が優れていることが分かった。

キーワード 波面合成, 個人用三次元音場再生, ショットガンマイクロホン, アンビソニックマイク

Subjective Evaluation of Recording System for Personal 3D Sound Field Reproduction

Fumi HANYU[†] Toshiyuki KIMURA[‡]

[†] Graduate School of Engineering, Tohoku Gakuin University 1-13-1 Chuo, Tagajo, Miyagi, 985-8537 Japan

[‡] Faculty of Engineering, Tohoku Gakuin University 1-13-1 Chuo, Tagajo, Miyagi, 985-8537 Japan

E-mail: [†] s1994208@g.tohoku-gakuin.ac.jp, [‡] t-kimura@m.ieice.org

Abstract We have proposed the personal 3D sound field reproduction in which the sound recorded by eight directional microphones is replayed by eight loudspeakers. In this report, in order to construct the recording system for the personal 3D sound field reproduction, the microphone array using shotgun microphones in which the recording position was identified was proposed and the localized direction of the proposed microphone array was compared with that of an ambisonic microphone. As a result, it was shown that the localized direction of the proposed microphone array was better than that of an ambisonic microphone.

Keywords Wave field synthesis, Personal 3D sound field reproduction, Shotgun microphone, Ambisonic microphone

1. はじめに

近年、三次元音場再生技術が進歩しており、また遠隔操作システムによる技術も様々な現場で取り入れ始めている。建築現場での重機の操作などにおいて、三次元音場再生技術と遠隔操作システムを組み合わせることで、現場での危機察知能力の向上や作業の効率化が期待される。

これまで、遠隔操作への応用を目的とした個人用三次元音場再生システムが提案されてきた[1]。しかしながら、提案システムは収録装置の構築及び性能評価が行われてこなかった。本研究では、提案法として構築した収録装置、従来法としてアンビソニックマイクを用いて収録した音を使用し、音の定位性能の比較をおこなった。

2. 個人用三次元音場再生システム

個人用三次元音場再生システム[1]を Fig. 1 に示す。

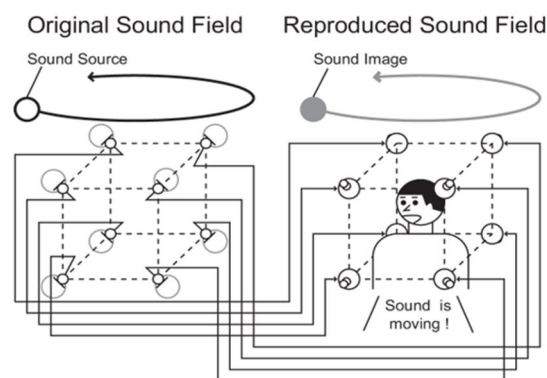


図 1 個人用三次元音場再生システムの構成図

Fig. 1 Configuration of
personal 3D sound field reproduction system

立方体の頂点の位置に指向特性を外向きで配置した 8 個の指向性マイクロホンで収録した音を、聴取者の頭



図 2 收音位置同定に用いた
ショットガンマイクロホン

Fig. 2 Shotgun microphone
used in the recording position identification

部の周りに指向性マイクロホンと同じ形で配置した 8 個のスピーカで再生することで、制御領域内の三次元音場を聴取領域内に再現するシステムである。信号の音響処理や作業者は頭部に何か装着する必要がなく、水平方向にスピーカを配置しないので視界を妨げないシステムとなっている。

3. 収録装置

3.1. ショットガンマイクロホンの收音位置同定

Fig. 1 より、マイクロホンアレイとスピーカアレイでは同じ形の立方体を組む必要がある。しかし本研究でマイクロホンアレイを構築する際に使用したショットガンマイクロホンの形状が棒状のため、どの位置で收音しているのかを測定する必要がある。そこで、先行研究[2]に倣い、スピーカからピンマイクロホンまでのインパルス応答をもとにスピーカの発音位置を特定した後に、スピーカから指向性マイクロホンまでのインパルス応答をもとにショットガンマイクロホンの收音位置を同定した。

收音位置同定に用いたショットガンマイクロホン (AZDEN SGM-1000)、ピンマイクロホン (オーディオテクニカ AT9903) 及びスピーカ (オーム電機 ASP-



図 3 收音位置同定に用いたピンマイクロホン

Fig. 3 Pin microphone
used in the recording position identification



図 4 收音位置同定に用いたスピーカ

Fig. 4 Loudspeaker
used in the recording position identification

204-N-K) の写真を Figs. 2~4 に、インパルス応答測定の際の配置を Fig. 5 に示す。

それぞれ測定したインパルス応答から算出した遅延時間を Table 1 に示す。Table 1 より、2 つのピンマイクロホン間の距離と遅延時間差は 1 m 及び 2.875 ms

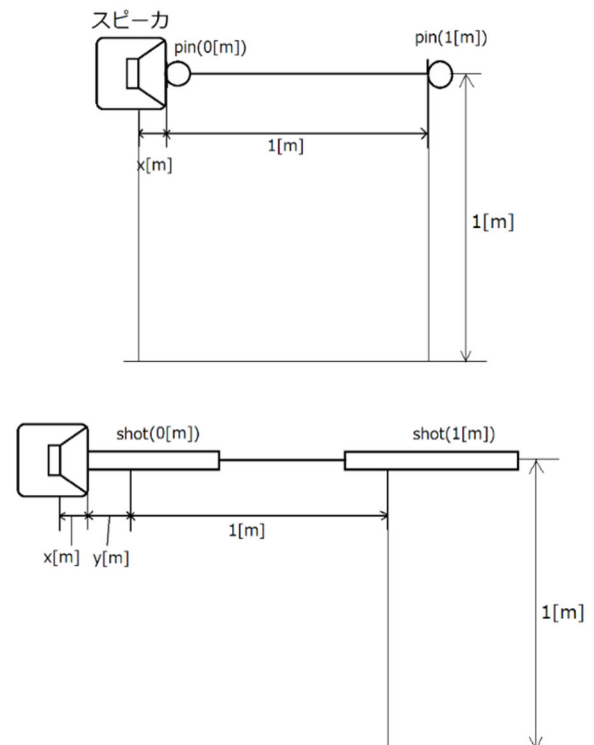


図 5 收音位置同定における
マイクロホン及びスピーカ配置

Fig. 5 Position of microphones and loudspeaker
in the recording position identification

表 1 收音位置同定における
インパルス応答の遅延時間

Table 1 Delay time of impulse response
in the recording position identification

	遅延時間[ms]
Pin [0 m]	0.104
Pin [1 m]	2.979
Shotgun [0 m]	0.521
Shotgun [1 m]	3.521

であるため、音速は

$$\frac{1[\text{m}]}{2.875[\text{ms}]} = 347.83[\text{m/s}]$$

となる。従って、スピーカの発音位置 x は、スピーカ表面より

$$x = 0.104[\text{ms}] \times 347.83[\text{m/s}] = 3.6174[\text{cm}]$$

奥行きとなる。この値は使用したスピーカの奥行き長 (6 cm) に収まっている。

一方、2つのショットガンマイクロホン間の距離と遅延時間差は 1 m 及び 3 ms であるため、音速は

$$\frac{1[\text{m}]}{3.000[\text{ms}]} = 333.33[\text{m/s}]$$

となり、收音位置 y は

$$\begin{aligned} y &= 0.521[\text{ms}] \times 333.33[\text{m/s}] - x \\ &= 17.3665[\text{cm}] - 3.6174[\text{cm}] \\ &= 13.7491 \approx 13.7[\text{cm}] \end{aligned}$$

となる。よって、ショットガンマイクロホンの收音位置は、先端より 13.7 cm である。この値はショットガンマイクロホンの全長 (31 cm) に収まっている。

3.2. マイクロホンアレイの構築

前節の結果から実際に構築したマイクロホンアレイを Fig. 6 に示す。收音位置の点を頂点とし、一辺が 0.4 m の立方体となるように設置した。その際にシヨ

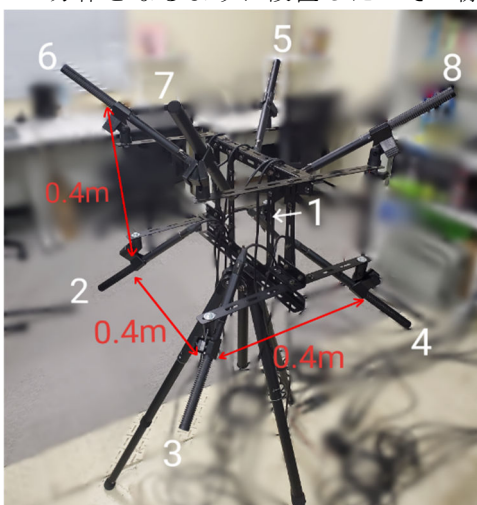


図 6 構築したマイクロホンアレイ
Fig. 6 Constructed microphone array

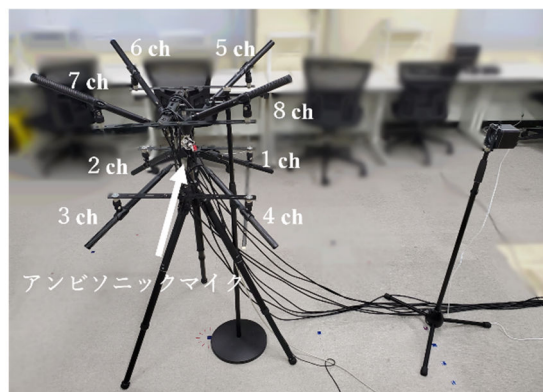


図 7 主観評価におけるインパルス応答の測定
(音源位置 180 度)

Fig. 7 Impulse response measurement
in subjective evaluation
(Sound source direction : 180°)

ットガンマイクロホンの指向特性は外側を向くよう設置した。ショットガンマイクロホンの角度は分度器で 45 度 ($= \tan^{-1} \frac{1}{\sqrt{2}}$) を測定し、だいたいの角度を定めた後に目測で調整している。

4. 主観評価

構築したマイクロホンアレイにアンビソニックマイク (Zoom H3-VR) を加えた計 2 種類の収録条件において、音の定位性能を比較する。

4.1. インパルス応答の測定

主観評価実験で実験参加者に聞かせる音は、マイクロホンアレイとアンビソニックマイクを用いて測定したインパルス応答に白色雑音と音声を畳み込んで作成した。本節ではインパルス応答の測定に関して述べる。

4.1.1. 測定環境

測定は実験室においておこなった。残響時間は 500 ms、測定の際の暗騒音レベルは 31.6 dBA であった。音源から 1 m 地点での音圧レベルは 74.2 dB とした。音を収録している写真を Fig. 7 に示す。収録ではマイクロホンアレイの中心にアンビソニックマイクを置いた。音源の位置を Fig. 8 のように設定し、マイクロホンアレイの中心から水平方向に 1 m 離れた地点へ 30 度間隔でスピーカを設置した。スピーカからはサンプリング周波数 48 kHz の 65536 点 TSP 信号 (Time Stretched Pulse, 時間引き延ばしパルス) を再生した。その際、9 回の同期加算をおこなった。

4.1.2. 測定手順

音源位置を 12 か所としたので、1 か所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繰り返した。

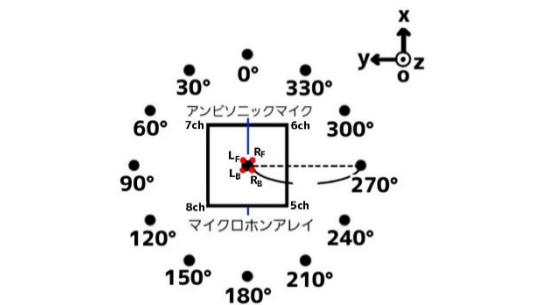
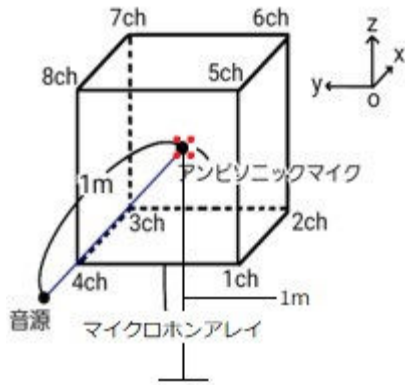


図 8 インパルス応答測定時のスピーカ配置
(上：音源位置 180 度の時の斜めから見た図，
下：すべての音源位置を上から見た図)

Fig. 8 Loudspeaker placement
in impulse response measurement
(Up : Sound source direction 180°,
Down : All sound source directions)

4.2. 再生音の合成

実験参加者に聞かせるための再生音は測定したインパルス応答に 4 s の長さの白色雑音と音声を畳み込んだものを使用した。アンビソニックスは合計 4 ch で音を収録していたが、今回使用するスピーカアレイは 8 ch なので先行研究[3]に従い、8 ch に展開している。

4.3. 定位性能実験

4.3.1. 実験環境

実験は実験室にて実施し、Fig. 9 に示すスピーカアレイ（1 辺が 40 cm 四方の頂点にスピーカを配置）を用いた。実験参加者は構築した視聴覚提示装置の椅子に座り HMD を装着した。配置図を Fig. 10 に示す。スピーカアレイで出力する音は、前述したものを利用している。実験参加者の視界である HMD に表示する映像は先行研究[4]で作成された三次元仮想空間を用いた。収録の際と同様に実験参加者の位置に相当するカメラ位置を中心とする円周上に番号が振られたオブジェクトを 30 度間隔で計 12 個設置した。実験参加者の位置からオブジェクトまでの距離は 10 m とした。Fig. 11 にオブジェクト及び作成した三次元仮想空間を示す。

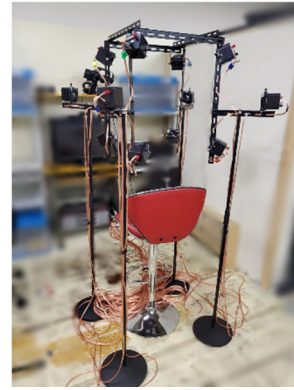


図 9 主観評価に用いたスピーカアレイ

Fig. 9 Loudspeaker array
used in the subjective evaluation

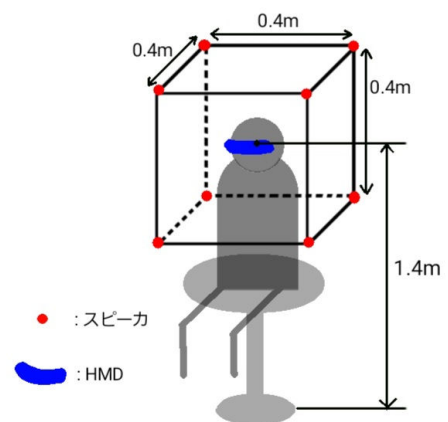


図 10 主観評価における実験参加者及び
スピーカアレイの配置

Fig. 10 Position of participant and loudspeaker array
in the subjective evaluation

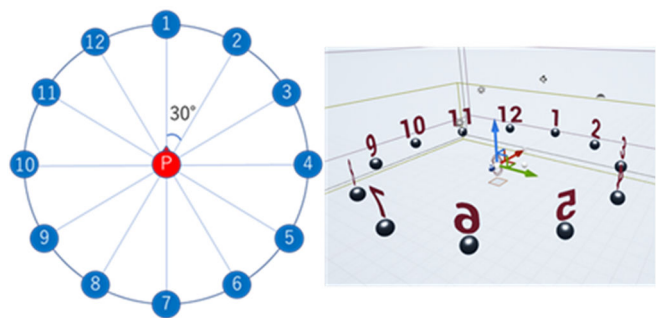


図 11 主観評価における
オブジェクト（左）及び三次元仮想空間（右）

Fig. 11 Position of objects (Left) and
3D virtual space (Right) in the subjective evaluation

4.3.2. 実験手順

Fig. 12 に実験の流れ図を示す。実験は音声を流す条件をセッション A とし、白色雑音を流す条件をセッション B とした。実験参加者は 10 名で、実験参加者に

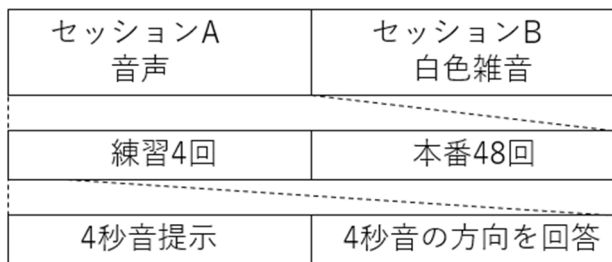


図 12 主観評価実験の流れ図

Fig. 12 Flowchart of the subjective evaluation

は各セッションにおいて、4回の練習試行の後に計48回の本番試行(=収録システム2条件×12音像箇所×2回)を実施した。セッションおよび音像の提示角度の順序、収録装置の種類は実験参加者ごとにランダムイズした。

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、実験者は実験参加者の回答した番号を記録した。

4.3.3. 実験結果

まず、白色雑音を提示した時の実験結果を Fig. 13 に示す。また、実験結果から2つの収録方法の結果を比較するため、完全正答率(提示角度と同じ回答角度を正答とした割合)と方向正答率(提示角度±30度の誤差までの回答を正答とした割合)を集計し、それぞれ χ^2 検定を実施した。検定結果を Table 2 に示す。どちらも p 値は0.05を下回り、有意な差が認められた。

Fig. 13 及び Table 2 より、白色雑音を提示した場合、提案法の個人用三次元音場再生システムは完全正答率が50%をきったものの、方向正答率は92.1%という結果となった。そのため提案法で音を収録した場合、音像のおおまかな方向を聴取者に伝えることができると推察される。従来法のアンビソニックマイクは方向正答率が53.8%と提案法に比べはるかに性能が劣っていた。実験者が従来法の音を実際に聞いた時、音像がぼやけていて定位性能が悪いと感じたが、実験結果にもそれが表れていた。

次に、音声を提示した時の実験結果を Fig. 14 に示す。実験結果から2つの収録方法の結果を比較するため、完全正答率と方向正答率についてそれぞれ χ^2 検定を実施した。検定結果を Table 3 に示す。どちらも p 値は0.05を下回り、有意な差が認められた。Fig. 13 及び Table 3 より、音声の場合においても白色雑音の時と同じ傾向となった。

さらに、回答から完全正答率と方向正答率を音源位置ごとに集計し、それぞれ χ^2 検定を実施した。結果をそれぞれ Tables 4-5 に示す。Tables 4-5 より、音源位置が210~330度の場合において有意な差が出にくい結果となっている。特に音源位置が270度の場合、すべ

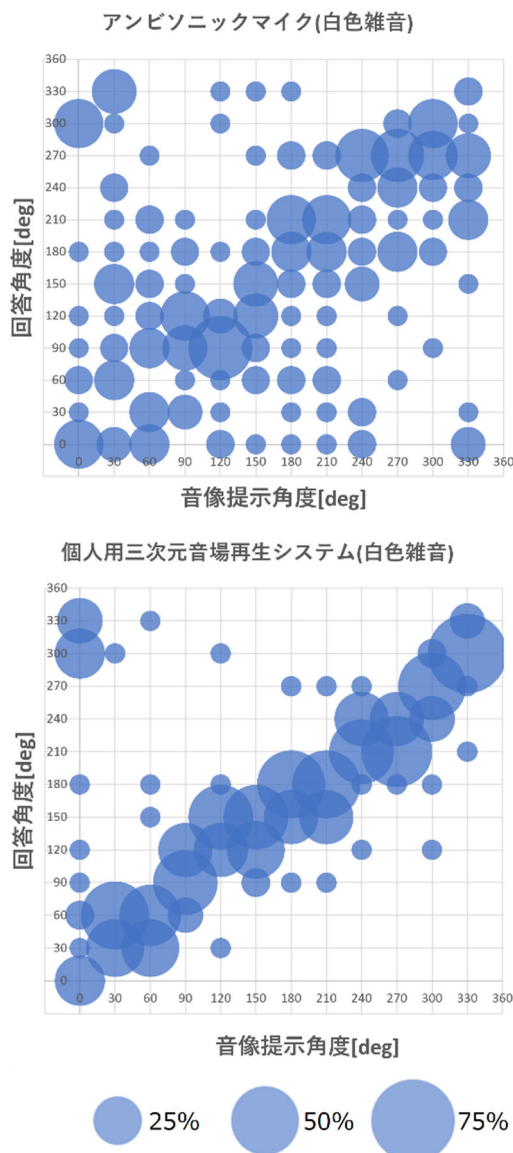


図 13 主観評価の実験結果(白色雑音)
Fig. 13 Experimental result
of the subjective evaluation (White noise)

ての場合において有意差がでなかった。これは210~330度の方向に壁があり、障害物による反射音が結果に影響を及ぼしたのではないかと考えられる。

5. まとめ

本研究ではこれまでに提案されている個人用三次元音場再生システムの収録装置を構築するため、收音

表 2 主観評価における χ^2 検定結果(白色雑音)

Table 2 Result of χ^2 test
in the subjective evaluation (White noise)

	従来法	提案法	p 値
完全正答率[%]	19.2	46.7	1.45×10^{-10}
方向正答率[%]	53.8	92.1	3.41×10^{-21}

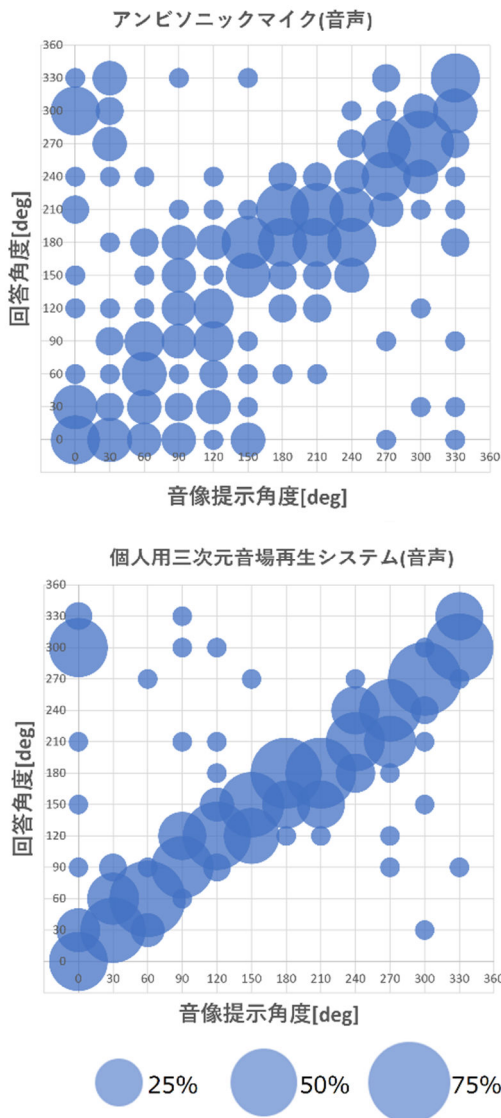


図 14 主観評価の実験結果 (音声)
Fig. 14 Experimental result
of the subjective evaluation (Speech)

位置が同定されたショットガンマイクロホンを用いてマイクロホンアレイを構築した。構築した収録装置の主観性能を評価するため、音の定位方向に関する主観評価を実施し、従来法のアンビソニックマイクと比較した。その結果、構築したマイクロホンアレイの方がアンビソニックマイクよりも定位性能が優れていることが分かった。また、構築したマイクロホンアレイで

表 3 主観評価における x^2 検定結果 (音声)
Table 3 Result of x^2 test
in the subjective evaluation (Speech)

	従来法	提案法	p 値
完全正答率[%]	26.7	57.9	4.23×10^{-12}
方向正答率[%]	61.3	90.0	2.21×10^{-13}

表 4 主観評価における音源位置ごとの完全正答率と x^2 検定結果 (左: 白色雑音, 右: 音声)

Table 4 Result of fully correct rate and x^2 test of each position in the subjective evaluation (Left : White noise, Right : Speech)

提示角度	従来法[%]	提案法[%]	p値	提示角度	従来法[%]	提案法[%]	p値
0	30	55	0.1098	0	30	45	0.3272
30	0	40	0.0016	30	10	55	0.0024
60	0	45	0.0007	60	25	75	0.0016
90	25	50	0.1025	90	15	50	0.0181
120	15	35	0.1441	120	15	60	0.0033
150	25	50	0.1025	150	25	55	0.0528
180	20	55	0.0222	180	30	65	0.0267
210	10	55	0.0024	210	25	45	0.1848
240	20	35	0.3355	240	30	50	0.1967
270	30	55	0.1098	270	55	70	0.3272
300	40	70	0.0565	300	30	95	2.18×10^{-5}
330	15	15	1.0000	330	30	30	1.0000
合計	19.2	46.7		合計	26.7	57.9	

表 5 主観評価における音源位置ごとの方向正答率と x^2 検定結果 (左: 白色雑音, 右: 音声)

Table 5 Result of directional correct rate and x^2 test of each position in the subjective evaluation (Left : White noise, Right : Speech)

提示角度	従来法[%]	提案法[%]	p値	提示角度	従来法[%]	提案法[%]	p値
0	60	95	0.0080	0	60	80	0.1675
30	35	95	4.32×10^{-8}	30	40	90	0.0009
60	40	85	0.0033	60	60	95	0.0080
90	60	100	0.0016	90	35	85	0.0012
120	65	85	0.1441	120	40	85	0.0033
150	60	90	0.0285	150	60	95	0.0080
180	60	90	0.0285	180	75	95	0.0765
210	30	95	6.31×10^{-5}	210	70	95	0.0375
240	60	95	0.0080	240	75	85	0.4292
270	70	90	0.1138	270	87	85	1.0000
300	70	95	0.0375	300	75	100	0.0168
330	35	90	0.0003	330	60	90	0.0285
合計	53.8	92.1		合計	61.3	90.0	

は音像の方向をおおまかに定位できることも分かった。

本研究の主観評価実験では実験参加者の水平方向に音源を用意したが、今後は上下方向にも音源を置き、定位性能を評価する必要がある。

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," 信学論(A), Vol. J97-A, No. 4, pp. 284-294 (2014).
- [2] 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," 信学技報, No. EA2017-27, pp. 1-4 (2017).
- [3] K. Farrar, "Soundfield microphone," Wireless World, 85, pp. 48-51 (Oct.1979).
- [4] 繁泉宥斗, 木村敏幸, "個人用音場再生とバイノーラルの音像定位比較実験," 平成 31 年東北地区若手研究者研究発表会講演資料, No. YS-17-A8, pp. 15-16 (2019).