

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

木村 敏幸^{†a)}

Personal Compact 3D Sound Field Reproduction System Based on Concept of Wave Front Synthesis Technique Using Eight Directional Microphones

Toshiyuki KIMURA^{†a)}

あらまし 従来の波面合成技術を用いた個人向け三次元音場再生システムは数多くのスピーカを聴取者の頭部の周りに配置することを前提としている。しかし、遠隔操作システムを実現するには伝送チャンネル数が多すぎるし、スピーカが聴取者の視界に入るので、視聴覚システムを構築するのが困難という問題点がある。本論文では、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向け三次元音場再生システムを提案する。8個の指向性マイクロホンを聴取者の頭部の周りに設定した立方体の頂点位置に配置する。再生用スピーカは聴取者の頭部の水平面上に配置されていないので、聴取者の水平方向の視界はスピーカによって妨げられていない。また、必要な伝送チャンネル数は8チャンネルで良い。提案システムの定位性能を定位実験によって検討したところ、想定する立方体再生空間の一边を40 cmにし、超指向性マイクロホンを用いればシステムを構築するには十分な定位性能が得られることが分かった。

キーワード 遠隔操作, 音場再生, 波面合成法, 指向性マイクロホン, 定位実験

1. ま え が き

2011年3月11日に発生した東日本大震災では地震や津波の被害のみならず、福島第一原子力発電所の事故も引き起こされた。それにより、今もなお原子力発電所の周囲が立入禁止区域となっており、区域内で瓦礫除去や除染などの作業を実施することが非常に制限されている。このような問題を解決するためには、作業用の重機を遠隔で操作し、作業者が区域内に立ち入らなくても作業を実施できるようにすることが非常に重要となる。

一方で、独立行政法人情報通信研究機構(NICT)では立体映像、立体音響、多感覚インタラクションなどを用いた超臨場感コミュニケーション技術を研究開発

している[1],[2]。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(本論文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技術に適用すれば、より効率的に作業を遂行できるようになることが期待される。本論文ではその中でも立体音響技術に着目し、あたかも遠隔地の対象物から音が出ているかのように聞こえる遠隔操作システムについて検討する。

遠隔操作システムに適用できる立体音響技術としては以下の技術要件を満たす必要がある。

- (1) 三次元音響空間を忠実に表現できること。
- (2) 伝送チャンネル数ができる限り少量であること。

(3) 操作によるインタラクション性を求められるので、過度な遅延を生じるような音響信号処理はしないこと。例えば、レイグジスタンス(ロボットの遠隔操作)システムの場合、遅延の許容限は74 ms以下であるので[3]、音響システムにおいても遅延時間は、その程度に抑えるのが望ましいと考える。

(1)と(2)の条件から想定される立体音響技術はバイノーラル[4],[5]やトランスオーラル[6],[7]といった

[†] 独立行政法人情報通信研究機構ユニバーサルコミュニケーション研究所, 京都府

Universal Communication Research Institute, National Institute of Information and Communications Technology, Hikaridai 3-5, Seika-cho, Soraku-gun, Kyoto, 619-0289 Japan

a) E-mail: t-kimura@nict.go.jp

技術である。これらの技術は遠隔地において作業者を想定したダミーヘッドの左右耳位置にマイクロホンを設置し、収録した音響信号を伝送すればよいので、伝送チャンネル数が2チャンネルで済むという利点がある。しかしながら、バイノーラル技術では作業者がヘッドホンを装着する必要があるし、ダミーヘッドの頭部形状が作業者と一致しなければ頭内定位や前後誤りを引き起こすという問題点がある。トランスオーラル技術は頭内定位が起これないという利点があるものの、スピーカから作業者の頭部の左右耳位置への音響経路をキャンセルする逆フィルタを収録した音響信号にあらかじめ畳み込む必要がある。音響経路を完全にキャンセルするためには作業者がいる室内の残響時間の1/4の時間遅延を必要とするので[8]、室内の残響時間が約300ms(通常の会議室の残響時間に相当)以上の場合、逆フィルタによって過度な遅延を引き起こす可能性がある。何よりも、どちらの技術においても作業者の頭部が動いた時にはダミーヘッドも同じように動かす必要があるため、作業者の頭部の位置や向きを検出する機構やダミーヘッドを動かすための駆動機構[9]を導入しなければならない。

サラウンド録音に応用されている收音技術[10]を適用することも考えられるが、現在提案されているマイクロホン配置は5.1chシステム[11]に適用するために2次元平面を想定しているため、遠隔操作に必要な音響情報(三次元空間上における特定の方向から危険を察知する音が鳴っているように聞こえること)を表現するには適していない。また、一方では3次元空間を想定した收音技術(22.2chシステム[12]、高次アンビソニクス[13]、境界音場制御[14])も提案されているが、忠実な音響情報を表現しようとすればするほどマイクロホンの数が多くなり、その結果伝送チャンネル数が多くなってしまふ。

ところで、これまでに指向性マイクロホンを用いた波面合成技術が提案されている[15]~[18]。この技術は制御領域内の波面を別の領域(聴取領域)内に忠実に再現する三次元音場再生技術である。まず、原音場において制御領域の境界面上に指向性マイクロホンを配置する。その際、マイクロホンの指向特性は制御領域の外側に向けるようにする。次に、制御領域と同じ形状をした聴取領域の境界面上にスピーカを配置する。その際、スピーカの位置と指向性マイクロホンの位置は同じにする。そして指向性マイクロホンで収録した音をスピーカで直接再生することにより、

Kirchhoff-Helmholtz 積分方程式[19]に従って波面が聴取領域内に忠実に再現される。この技術では、聴取者が聴取領域内で自由に頭部を動かすことができるし、逆フィルタのような音響信号処理を導入する必要がないのが特徴であるが、高周波数域にまで忠実に波面を再現するためには指向性マイクロホンやスピーカを緊密に配置する必要がある。また、指向性マイクロホンやスピーカの数は領域の表面積に比例する。その結果、複数人が同時に聴取することを想定した広い領域で再現しようとするとき指向性マイクロホンやスピーカの数が非常に多くなり、伝送チャンネル数が非常に多くなってしまふという問題点があった。

しかしながら、約300~400Hz以下の範囲でしか波面が正確に再現されていないような場合においても、それ以上指向性マイクロホンやスピーカの数を増やして再現される周波数域を上げて聴取者は知覚上の違いを感じることはできないので、結果的に波面を合成するのに必要な指向性マイクロホンやスピーカの数を減らすことができるということが示されている[20],[21]。従って、聴取領域を作業者の頭部近傍に限定することにより領域の表面積を減少させ、聴取実験を実施することにより知覚上の違いを感じることができないかを検討すれば、必要な指向性マイクロホンやスピーカの数を必要最小限に抑えることができ、伝送チャンネル数が十分に実用的な遠隔操作システムを構築することができるようになることが期待される。その際、音場知覚の構成要素(方向感、距離感、空間印象[22])ごとに聴取実験を実施した場合、方向感を再現するのに必要なチャンネル数の方が空間印象よりも多かったため[20],[21]、方向感を検討する定位実験を実施することが特に重要であると言える。

聴取領域を作業者の頭部近傍に限定する場合、領域形状を作業者の頭部を囲む球形にすればどの三次元方向に対しても空間解像度を均一にすることができる。その際、境界面上の制御点は以下の条件を満たす必要がある。

- (1) 制御点を球面上に均等に配置すること。
- (2) スピーカの位置が作業者の胴体の位置と重なるので、球の真下に制御点を配置しないこと。
- (3) スピーカの位置が作業者の視界に入るので、球の真正面に制御点を配置しないこと。
- (4) 約300~400Hzまでの波面を再現するために、制御点間隔を少なくとも $0.567(=\frac{340}{300} \times \frac{1}{2})$ m以下にすること。

(1) の条件から正多面体 (正四面体 (頂点数 4), 正六面体 (頂点数 8), 正八面体 (頂点数 6), 正十二面体 (頂点数 20), 正二十面体 (頂点数 12)) の頂点位置が想定されるが, (2) や (3) の条件から正八面体と正二十面体は除外される。また, (4) の条件に基づいて正四面体の頂点位置でスピーカアレイ用のフレームを構築した場合, 内接球の半径が小さすぎるので ($0.567 \times \frac{\sqrt{6}}{12} = 0.116\text{m}$), スピーカアレイの内部に作業者の頭部が入らない可能性がある。さらには, 遠隔操作システムのためには制御点数が少ない方が望ましい。従って, 正六面体の 8 個の頂点の位置に制御点を配置するのが遠隔操作システムを構築するには最も適した制御点配置であると考えられる。

そこで, 本論文では 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遠隔操作のための個人向け三次元音場再生システムを提案する。作業者は聴取領域内で頭を自由に動かすことができるので, 作業者の頭部の位置や向きを検出する機構や遠隔先でダミーヘッドを動かすための駆動機構を導入する必要がないことが提案システムの特徴であると言える。伝送チャネル数はバイノーラル技術やトランスオーラル技術といった従来の三次元音場再生システムを適用した場合 (2 チャネル) に比べて 8 チャネルと多くなるが, 最近では 8 チャネルの音響伝送サービスも実用化されているので, 十分に実用的なシステムであると考えられる。第 2 章では提案システムの原理について説明する。第 3 章では提案システムの定位性能を検討するために実施した定位実験について説明する。第 4 章では定位実験によって得られた結果とその考察について記述する。

2. 提案システムの原理

提案するシステムの構成を図 1 に示す。まず, 図 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 個の指向性マイクロホンを配置し, 音を取録する。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, 図 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 個のスピーカから取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, 低周波数域において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

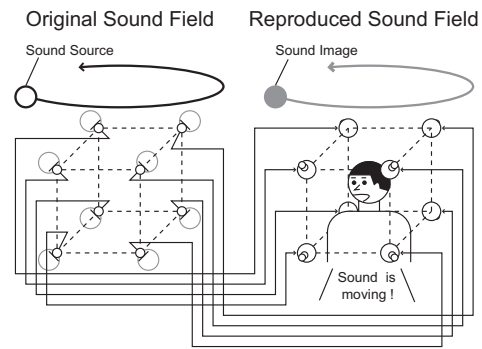


図 1 提案する個人向け三次元音場再生システム
Fig.1 Proposed personal 3D sound field reproduction system.

制御領域の上側で音が動くとき, 聴取領域内にいる聴取者は図 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ことが期待される。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と考えられるので, 将来的には立体映像技術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と考えられる。

3. 定位実験

三次元空間の定位実験を実施する場合, 従来は 2 種類の実験 (水平面方向に限定した方位角の定位実験, 正中面に限定した仰角の定位実験) による定位性能の検討が行われている。しかし, それぞれの実験では回答基準を方位角か仰角のどちらかに限定して知覚方向を回答させ, 実験結果から得られた定位性能を検討しているため, 全ての方向に対して方位角と仰角に関する定位性能を同時に検討しているとは言えない。従って, 上記の 2 種類の実験を実施するだけでは三次元空間の定位性能を検討するには不十分であると考えられる。本論文では, 聴取者に知覚方向を球面上の点の位置で回答させることで, 全ての定位方向において方位角と仰角に関する定位性能を同時に検討することができる定位実験を実施する。

3.1 実験環境

定位実験は暗騒音レベルが A 特性による音圧レベルで 20 dB で, 残響時間が 115 ms の室内にて実施した。まず, 聴取する頭部中心位置を設定し, 聴取する

表 1 定位実験における音源の方位角及び仰角
Table 1 Azimuth and elevation angles of sound sources in the localization test.

Index	θ_0	ϕ_0	Index	θ_0	ϕ_0
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°			

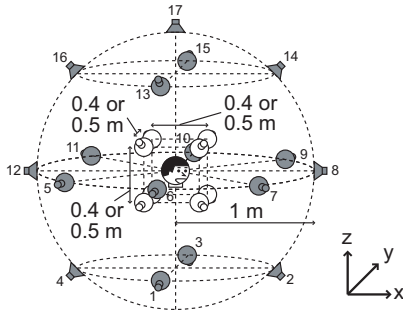


図 2 定位実験における聴取者及びスピーカの配置
Fig. 2 Positions of a listener and loudspeakers in the localization test.

頭部中心位置を重心とする立方体の頂点位置に 8 個のスピーカを配置した。それと同時に、定位性能の比較のために半径 1 m の球上に統制条件用スピーカを配置した。実験に用いたマルチチャンネル音響再生装置の都合上、配置できる統制条件用スピーカの数は最大 24 個であり、方位角と仰角に関する定位性能を同時に検討するためには統制条件用スピーカはできる限り球面上に均一に配置した方が望ましいことから、本論文では統制条件用スピーカは表 1 に示すような 17 箇所に配置した。図 2 に配置した計 25 個のスピーカを示す。白色のスピーカは音場再生用スピーカを示し、灰色のスピーカは半径 1 m の球上に配置した統制条件用スピーカを示している。スピーカは図 3 に示すようにスピーカユニット (AURASOUND 社製 NSW1-205-8A 相当品) をエンクロージャにマウントして製造した。音圧レベルは聴取する頭部中心位置において A 特性による音圧レベルで 60 dB に設定した。

3.2 マルチチャンネル信号の作成

8 個の音場再生用スピーカから再生したマルチチャンネル信号は計算機上で作成した。方向感はず主に音源からの直接音に依存するので、原音場は自由空間とした。

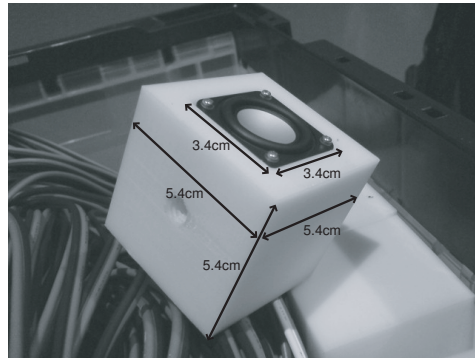


図 3 製造したスピーカ
Fig. 3 Manufactured loudspeaker.

音源から i 番目の指向性マイクロホンまでの室内インパルス応答 $g_i(n)$ は以下のように表わされる。

$$g_i(n) = \frac{1}{d_i} \delta \left\{ n - \text{round} \left(\frac{d_i F_s}{c} \right) \right\} (i = 1 \dots 8) \quad (1)$$

但し、 F_s ($=48$ kHz) は標準化周波数、 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 番目の指向性マイクロホンまでの距離を表す。音源と i 番目の指向性マイクロホンの位置ベクトル (\mathbf{r}_0 , \mathbf{r}_i) はそれぞれ以下のように設定される。

$$\mathbf{r}_0 = \begin{pmatrix} d \cos \theta_0 \cos \phi_0 \\ d \sin \theta_0 \cos \phi_0 \\ d \sin \phi_0 \end{pmatrix} \quad (2)$$

$$\mathbf{r}_i = \begin{cases} \left(-\frac{A}{2}, -\frac{A}{2}, -\frac{A}{2}\right)^T & (i = 1) \\ \left(\frac{A}{2}, -\frac{A}{2}, -\frac{A}{2}\right)^T & (i = 2) \\ \left(\frac{A}{2}, \frac{A}{2}, -\frac{A}{2}\right)^T & (i = 3) \\ \left(-\frac{A}{2}, \frac{A}{2}, -\frac{A}{2}\right)^T & (i = 4) \\ \left(-\frac{A}{2}, -\frac{A}{2}, \frac{A}{2}\right)^T & (i = 5) \\ \left(\frac{A}{2}, -\frac{A}{2}, \frac{A}{2}\right)^T & (i = 6) \\ \left(\frac{A}{2}, \frac{A}{2}, \frac{A}{2}\right)^T & (i = 7) \\ \left(-\frac{A}{2}, \frac{A}{2}, \frac{A}{2}\right)^T & (i = 8) \end{cases} \quad (3)$$

ただし、 d ($=1, 3$ m) は音源から聴取する頭部中心位置までの距離、 θ_0 と ϕ_0 はそれぞれ聴取する頭部中心位置における方位角と仰角を表す。 A は想定する立方体再生空間の一辺の長さを表す。少なくとも 300 Hz までの波面を忠実に再現するために想定する立方体再生空間の一辺の長さは 0.567 m 以下にする方が望ましいが、本論文では、少なくとも 300 Hz までの波面

を忠実に再現し、なおかつスピーカアレイを実際に構築した際に聴取者が半径約 10 cm の頭部を聴取領域内に安全に入れることができる条件として 2 種類の $A(=0.4, 0.5 \text{ m})$ を設定した。

音源は MATLAB 上で合成された白色雑音を用いた。音源信号を $s(n)$ とした場合、 i 番目の指向性マイクロホンで収録される信号 $x_i(n)$ は以下のように表わされる。

$$\begin{aligned} x_i(n) &= D_i \{g_i(n) * s(n)\} \\ &= \frac{D_i}{d_i} s \left\{ n - \text{round} \left(\frac{d_i F_s}{c} \right) \right\} \end{aligned} \quad (4)$$

但し、 $*$ は畳み込みを表す。過去の研究において単一指向性や超指向性マイクロホンを用いた方が波面がより忠実に合成されることが示されているので [16], [17], 本論文では i 番目の指向性マイクロホンの指向特性 D_i は図 4 に示すような 2 種類の条件 (単一指向性, 超指向性) とした。指向特性は以下のように定義される。

$$\text{(単一指向性)} \quad D_i = \frac{1 + \cos \theta_i}{2} \quad (5)$$

$$\text{(超指向性)} \quad D_i = \begin{cases} \cos \theta_i & (|\theta_i| \leq 90^\circ) \\ 0 & (|\theta_i| > 90^\circ) \end{cases} \quad (6)$$

但し、 i 番目の指向性マイクロホンにおける音源の入射角 θ_i は以下のように定義される。

$$\theta_i = \cos^{-1} \left\{ \frac{\mathbf{r}_i \cdot (\mathbf{r}_0 - \mathbf{r}_i)}{|\mathbf{r}_i| |\mathbf{r}_0 - \mathbf{r}_i|} \right\} \quad (7)$$

単一指向性マイクロホンを用いる場合は、再生できない音源の方向が存在する可能性は生じない。一方、超指向性マイクロホンを用いる場合は、下記の式で示される領域内に音源 \mathbf{r}_0 が存在するときには 8 個のマイクロホンの指向特性が全て 0 になるため、スピーカアレイから音源が再生されなくなる。

$$|r_{0x}| + |r_{0y}| + |r_{0z}| \leq \frac{3}{2} A \quad (8)$$

ただし、 r_{0x} , r_{0y} , r_{0z} は \mathbf{r}_0 の三次元座標を表す。2 種類の立方体の大きさと 2 種類の指向特性条件から、今回の定位実験では表 2 に示すような計 4 種類の合成条件が設定されることになり、実験結果からこれらの合成条件ごとに定位性能の検討を行う。

3.3 実験条件

定位実験で用いた 5 種類の実験条件を図 5 に示す。

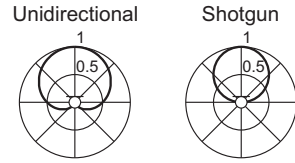


図 4 定位実験におけるマイクロホンの指向特性
Fig. 4 Directivity patterns of microphones in the localization test.

表 2 定位実験で用いた合成条件

Table 2 Synthesis conditions used in the localization test.

Index	Cubic size A	Microphone directivity D_i
(i)	0.4 m	Unidirectional
(ii)	0.5 m	Unidirectional
(iii)	0.4 m	Shotgun
(iv)	0.5 m	Shotgun

統制条件 (a) では、17 個の統制条件用スピーカの中から選ばれた 1 個のスピーカから音源信号が直接再生される。その結果、聴取者は図 5(a) に示すスピーカの位置から音が鳴っているように感じる事が期待される。その他の 4 条件 (b)-(e) では、(4) 式によって作成された 8 チャンネル信号 $x_i(n)$ が 8 個の音場再生用スピーカから再生される。図 5(b)-(e) の左側におけるマイクロホンの灰色の線はマイクロホンの指向特性を表す。その結果、聴取者は図 5(b)-(e) の右側に示す灰色の円の位置から音が鳴っているように感じる事が期待される。なお、表 2 で示した 2 つの合成条件 (i), (ii) は 2 つの実験条件 (b), (c) に含まれ、2 つの合成条件 (iii), (iv) は 2 つの実験条件 (d), (e) に含まれている。

3.4 実験手順

本実験には 7 名の男性と 3 名の女性が聴取者として参加した。実験の流れを図 6 に示す。実験は表 2 に示す立方体の大きさと 2 つのセッションに分けられた。立方体の大きさの提示順序は聴取者ごとにランダム化した。セッションごとに 34 回の練習試行と 170 回の本試行を実施した。その際、本試行においては 42 回もしくは 43 回ごとに休憩時間を設けた。試行の提示順序は聴取者ごとにランダム化した。練習試行と本試行の内訳を表 3 に示す。

聴取者には知覚した音の方向を回答紙に番号で記入するように教示した。その際、図 7 に示したような知覚した方向と方向番号の対応関係図を回答紙に常に見せたうえで知覚方向を回答させた。また、聴取者は音を聞いている間頭部や椅子を自由に回転させることができた。

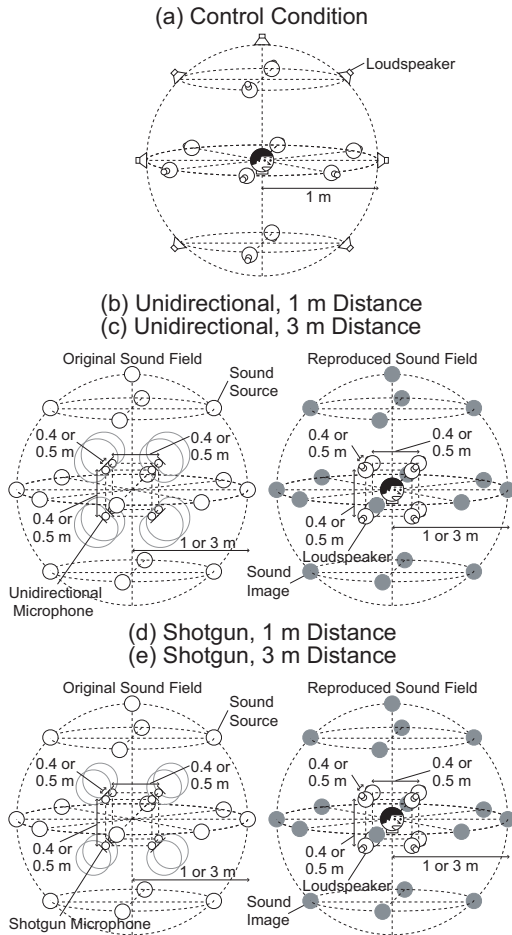


図 5 定位実験で用いた実験条件. 条件 (a) では音源信号を 17 個のスピーカの中から選ばれた 1 個のスピーカから再生. 条件 (b)-(e) では算出した 8 チャネル信号を 8 個のスピーカで再生.

Fig. 5 Experimental conditions used in the localization test. The sound source signal is played from one loudspeaker selected from seventeen loudspeakers in condition (a). Calculated eight channel signals are played from eight loudspeakers in conditions (b)-(e).

4. 実験結果及び考察

4.1 前後誤り

バイノーラルやトランスオーラル技術といった従来の三次元音場再生システムでは、聴取者の頭部の位置や向きを検出する機構を導入しない場合、聴取者の頭部伝達関数が収録時のものと一致していなければ前後誤りが引き起こされる [23]. そこで、聴取者の個人特性が提案システムに及ぼす影響を評価するために、定

Localization Test

Session 1		Session 2	
Order...Randomized (Array size 0.4 m or 0.5 m)			
Session			
Practice (34 trials)	Main (170 trials)		
	(43)	(42)	(43)
Trial			
Stimulus (4 s)		Answer (4 s)	

図 6 定位実験の流れ図
Fig. 6 Flowchart of the localization test.

表 3 定位実験における練習試行及び本試行

Table 3 Practice and main trials in the localization test.

	Element	Note
Practice (34)	= 17 directions × 2 conditions	(a) and (b) in Fig. 5
Main (170)	= 17 directions × 5 conditions × 2 repetitions	(a)-(e) in Fig. 5

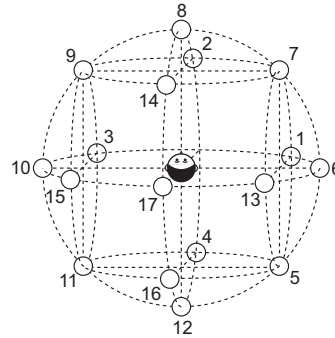


図 7 知覚方向と方向番号の対応関係
Fig. 7 Correspondence relation between perceived directions and direction numbers.

位実験の結果から前後誤りをした回答数を集計した. その際、本実験における前後誤りの判定基準は下記の 5 種類とした.

- 2 ($\theta = 0^\circ, \phi = -45^\circ$) の方向を 4 ($\theta = 180^\circ, \phi = -45^\circ$) と回答. もしくはその逆.
- 5 ($\theta = -135^\circ, \phi = 0^\circ$) の方向を 7 ($\theta = -45^\circ, \phi = 0^\circ$) と回答. もしくはその逆.
- 8 ($\theta = 0^\circ, \phi = 0^\circ$) の方向を 12 ($\theta = 180^\circ, \phi = 0^\circ$) と回答. もしくはその逆.
- 9 ($\theta = 45^\circ, \phi = 0^\circ$) の方向を 11 ($\theta = 135^\circ, \phi = 0^\circ$) と回答. もしくはその逆.
- 14 ($\theta = 0^\circ, \phi = 45^\circ$) の方向を 16 ($\theta = 180^\circ,$

表 4 定位実験における前後誤りの結果

Table 4 Results of front-back confusions in the localization test.

Condition	Distance	Number (Rate)
Control	-	1/400 (0.25%)
(i)	1 m	3/200 (1.50%)
	3 m	2/200 (1.00%)
(ii)	1 m	0/200 (0.00%)
	3 m	1/200 (0.50%)
(iii)	1 m	0/200 (0.00%)
	3 m	0/200 (0.00%)
(iv)	1 m	0/200 (0.00%)
	3 m	1/200 (0.50%)

$\phi=45^\circ$) と回答. もしくはその逆.

前後誤りの結果を表 4 に示す. 統制条件における前後誤り率は 0.25% と非常に低い. 従って, 聴取者はほとんど前後誤りを起こしていないことが分かる. これは白色雑音という定位が非常にしやすい音を音源として用いたことと聴取者が頭部や椅子を自由に回転させることができたことによるものと考えられる. 一方, 全ての合成条件において前後誤り率は 1.5% 以下である. 統制条件と 8 条件 (=4 合成条件 \times 2 距離) とでカイ二乗検定を実施したところ, 8 条件すべてにおいて統制条件と間に有意差は見られなかった. 従って, 提案システムの前後誤り性能は全ての合成条件において統制条件と同等であるということが言える. これは聴取者が頭部や椅子を自由に回転させることができたことによるものが大きいと考えられる.

4.2 知覚方向誤差

定位実験の結果を知覚した音像位置に基づいて分析する場合, 音像位置が球面上にあるときには知覚方向を方位角と仰角で評価することがまず考えられる. しかしながら, 方位角と仰角によって分析するには問題がある. 例えば, 提示方向が真上 (仰角 $\pm 90^\circ$) の場合は方位角は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像が真上からどの方向に変位しているかを仰角だけで分析することはできない. 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. その際, 提示方向が正面方向になるように極座標系を回転させれば, 実験結果は常に正面の提示方向を中心とした分布になるので, 全ての提示方向における知覚方向を同じ条件で分析できるようになることが期待される.

そこで, 本論文では下記の式に従って極座標系を回

転させることにより, 回答方向の方位角及び仰角 (θ, ϕ) を水平角及び垂直角 (θ', ϕ') に変換した. 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像の左右方向の変位に相当し, 変換した垂直角は音像の上下方向の変位に相当する.

$$\theta' = \tan^{-1} \frac{y'}{x'} \quad (9)$$

$$\phi' = \sin^{-1} \frac{z'}{\sqrt{x'^2 + y'^2 + z'^2}} \quad (10)$$

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.

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \mathbf{R}_y(-\phi_0) \mathbf{R}_z(-\theta_0) \begin{pmatrix} \cos \theta \cos \phi \\ \sin \theta \cos \phi \\ \sin \phi \end{pmatrix} \quad (11)$$

$$\mathbf{R}_y(-\phi_0) = \begin{pmatrix} \cos \phi_0 & 0 & \sin \phi_0 \\ 0 & 1 & 0 \\ -\sin \phi_0 & 0 & \cos \phi_0 \end{pmatrix} \quad (12)$$

$$\mathbf{R}_z(-\theta_0) = \begin{pmatrix} \cos \theta_0 & \sin \theta_0 & 0 \\ -\sin \theta_0 & \cos \theta_0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (13)$$

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, θ_0 と ϕ_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 ϕ' が 90° になった場合には, 変換した水平角 θ' のデータはこの後の集計には用いなかった. また, 4.1 節で記述したような前後誤りが起こった場合には, 変換した水平角 θ' と変換した垂直角 ϕ' 両方のデータともこの後の集計には用いなかった.

統制条件の変換した水平角及び垂直角の結果を図 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. 図中の灰色の丸は提示方向の方位角と仰角を表しているため, 平均値のプロットが灰色の丸に近ければ, 聴取者は提示方向と同じ方向に音像を知覚していると言える.

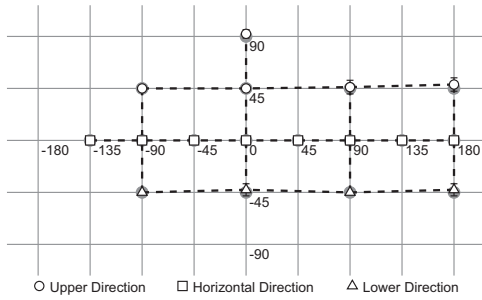


図 8 統制条件における変換した水平角及び垂直角の結果
Fig. 8 Results of transformed horizontal and vertical angles in control condition.

全ての方向において知覚方向は提示方向に近い。さらに、灰色の丸は信頼区間の中に入っているため、知覚方向と提示方向との間に有意差は見られない。従って、統制条件では聴取者は提示方向に正しく音像を定位していると言える。これは白色雑音という定位が非常にしやすい音を音源として用いたことと聴取者が頭部や椅子を自由に回転させることができたことによるものと考えられる。

一方、合成条件ごとの変換した水平角及び垂直角の結果を図 9~12 に示す。縦及び横方向のエラーバーは水平角及び垂直角の 95%信頼区間を表す。その際、縦及び横方向のエラーバーの中に灰色の丸が入っていない(つまり、知覚方向と提示方向との間に 5%有意差が見られる)提示方向は黒色の破線で囲んでいる。合成条件が (i), (ii), (iv) の場合は有意差が見られる提示方向が全 34(=17 方向×2 距離) 方向中 5 方向である。特に、水平面より下側の方向において有意差が見られる。これは水平面より下側の方向から音波が聴取する頭部中心位置に到来する際には聴取者の胴体による回折の影響が大きくなるのに対し、今回の実験では聴取者の胴体の回折の影響を考慮せずにマルチチャンネル信号を合成したためであると考えられる。それに対し、合成条件が (iii) の場合は有意差が見られる提示方向が全 34 方向中 1 方向である。従って、提案システムによって三次元音場再生システムを構築しようとする場合、合成条件は (iii)(すなわち、想定する立方体再生空間の一边を 40 cm にし、超指向性マイクロホンを用いること) にすると、他の合成条件に比べて最も良い定位性能が得られるということが言える。

さらに定量的な考察を行うため、以下の式で定義される変換した水平角及び垂直角の平均二乗誤差を算出する。

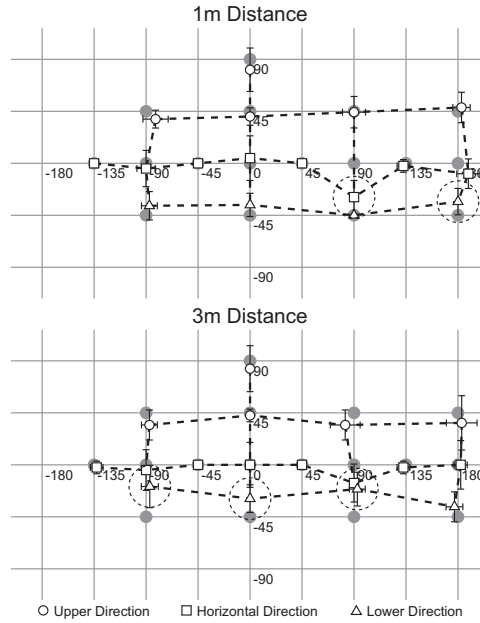


図 9 合成条件 (i) における変換した水平角及び垂直角の結果
Fig. 9 Results of transformed horizontal and vertical angles in synthesis condition (i).

(Horizontal angle)

$$MSE = \sqrt{\frac{1}{N} \sum_{\theta_0} (\theta' - \theta'_{CTRL})^2} \quad (14)$$

(Vertical angle)

$$MSE = \sqrt{\frac{1}{N} \sum_{\phi_0} (\phi' - \phi'_{CTRL})^2} \quad (15)$$

但し、 θ' , θ'_{CTRL} は提示方向を θ_0 及び ϕ_0 とした場合の合成条件及び統制条件の変換した水平角、 ϕ' , ϕ'_{CTRL} は提示方向を θ_0 及び ϕ_0 とした場合の合成条件及び統制条件の変換した垂直角、 $N(=17)$ は提示方向の総数である。合成条件ごとに得られた平均二乗誤差を変換した水平角及び垂直角ごとに図 13 に示す。合成条件が (iii) の場合は水平角の平均二乗誤差が 3 度未満、垂直角の平均二乗誤差が 9 度未満である。これらの値は白色雑音のような広帯域音における正面方向の左右方向の最小弁別角度(約 3 度 [24])や上下方向の最小弁別角度(約 9 度 [24])よりも小さい。一方、他の合成条件では、水平角の平均二乗誤差が 3 度以上であるか垂直角の平均二乗誤差が 9 度以上の場合が存在する。従って、合成条件を (iii) にしてシステムを構築すれば、聴

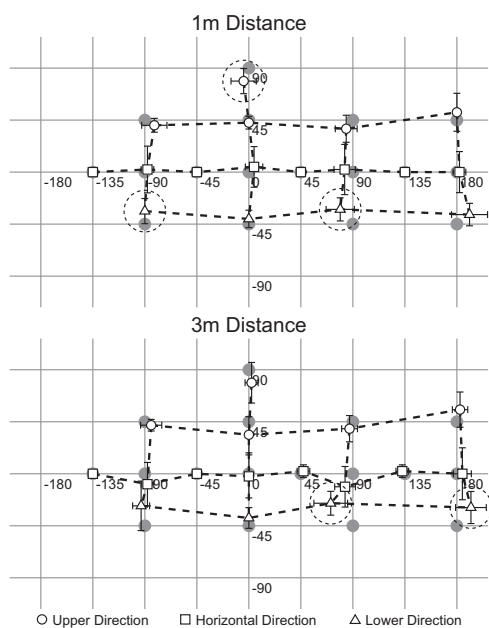


図 10 合成条件 (ii) における変換した水平角及び垂直角の結果

Fig. 10 Results of transformed horizontal and vertical angles in synthesis condition (ii).

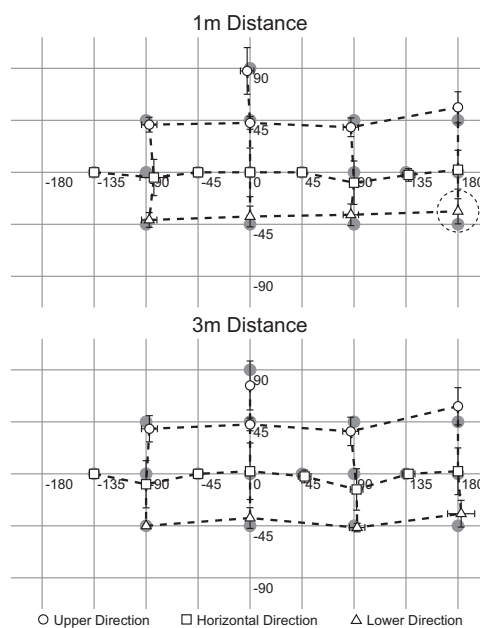


図 11 合成条件 (iii) における変換した水平角及び垂直角の結果

Fig. 11 Results of transformed horizontal and vertical angles in synthesis condition (iii).

取者は合成された音像の方向の違いを識別することができないので、聴覚的に十分な定位性能が得られるということが言える。

5. まとめ

本論文では、8 個の指向性マイクロホンと波面合成技術を用いた遠隔操作のための個人向け三次元音場再生システムを提案した。提案システムでは 8 個の指向性マイクロホンとスピーカを聴取者の頭部の周りに設定した立方体の頂点位置に配置した。提案システムの定位性能を評価するため、白色雑音を音源として用いた定位実験を実施した。変換した水平角及び垂直角に基づいて実験結果を分析したところ、想定する立方体再生空間の一辺を 40 cm にし、超指向性マイクロホンを用いれば、システムを構築するには十分に問題のない定位性能が得られることが分かった。

今後の課題としては音源信号を他の種類に変えた場合や実際のマイクロホンを用いてシステムを構築した場合の定位性能を検討することが挙げられる。また、今回提案した三次元音場再生システムを眼鏡なし立体映像システムと組み合わせることによって、作業者がメガネやヘッドホンを付けずに作業できる視聴覚の遠

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。

謝辞

本研究の遂行にあたり、定位実験を実施して下さった直江宗紀氏に感謝の意を申し上げる。また、ご支援下さった勝本道哲博士、山肩洋子博士に感謝の意を申し上げる。本論文における定位実験は独立行政法人情報通信研究機構倫理規定に従い実施した。

文 献

- [1] 井ノ上直己, “超臨場感コミュニケーション技術の研究開発概要,” 情報通信研究機構季報, vol.56, no.1/2, pp.3-7, March/June 2010.
- [2] N. Inoue, “Outline of ultra-realistic communication research,” J. NICT, vol.57, no.1/2, pp.3-8, March/June 2010.
- [3] 竹下, 渡邊, 佐藤, 南澤, 館, “テレイグジスタンスの研究 (第 63 報)-TELESAR3 において許容される通信遅延の検討-,” 日本バーチャルリアリティ学会大会論文集, no.1C1-4, pp.146-149, Sept. 2010.
- [4] 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, Aug. 1990.
- [5] J. Blauert, “Spatial hearing,” pp.372-392, MIT Press, Cambridge, Mass, revised edition, 1997.
- [6] 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

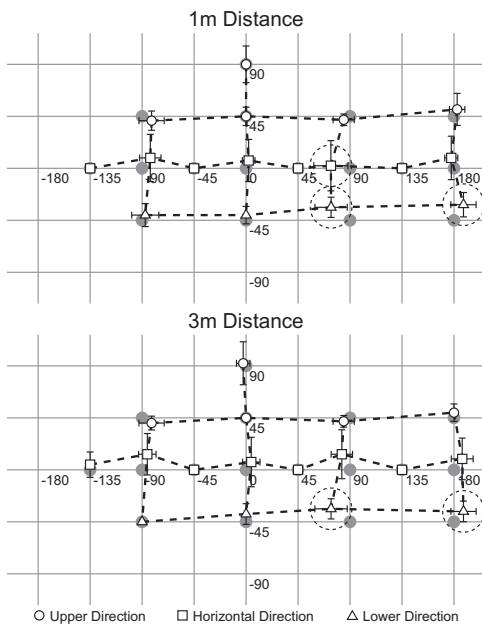


図 12 合成条件 (iv) における変換した水平角及び垂直角の結果

Fig. 12 Results of transformed horizontal and vertical angles in synthesis condition (iv).

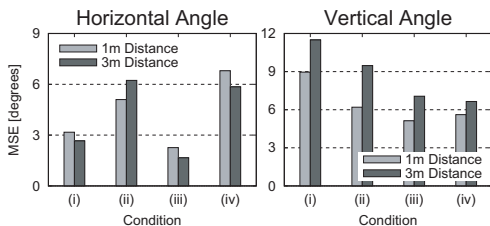


図 13 定位実験における平均二乗誤差の結果

Fig. 13 Results of mean square errors in the localization test.

音響学会誌, vol.29, no.3, pp.151–159, March 1973.

[7] M.R. Schroeder, D. Gottlob, and K.F. Siebrasse, “Comparative study of european concert halls: Correlation of subjective preference with geometric and acoustic parameters,” *J. Acoust. Soc. Am.*, vol.56, no.4, pp.1195–1201, Oct. 1974.

[8] 木村敏幸, 笈一彦, 武田一哉, 板倉文忠, “音場再現におけるマルチチャンネル信号の空間符号化,” *日本バーチャルリアリティ学会論文誌*, vol.8, no.4, pp.433–442, Dec. 2003.

[9] I. Toshima, S. Aoki, and T.Hirahara, “Sound localization using an auditory telepresence robot: Tele-Head II,” *Presence*, MIT Press, vol.17, no.4, pp.392–404, Aug. 2008.

[10] 深田晃, “音楽録音におけるマイクロフォンテクニック,”

映像情報メディア学会誌, vol.64, no.9, pp.1344–1348, Sept. 2010.

[11] ITU-R Recommendation BS.775-1, “Multichannel stereophonic sound system with and without accompanying picture,” 1992-1994.

[12] 小野一穂, 西口敏行, 松井健太郎, 濱崎公男, “音響遮蔽板を利用した 22.2ch 音響小型球形マイクロホン,” *日本音響学会春季研究発表会講演論文集*, no.1-P-34, pp.899–900, March 2013.

[13] 岡本拓磨, 岩谷幸雄, 坂本修一, 鈴木陽一, “高次アンビソニクス型音場収録のための 121ch 球状マイクロホンアレイの構築,” *電子情報通信学会技術研究報告*, no.EA2010-43, pp.1–6, Aug. 2010.

[14] S. Enomoto, Y. Ikeda, S. Ise, and S. Nakamura, “Three-dimensional sound field reproduction and recording systems based on boundary surface control principle,” *Proc. Int. Conf. Auditory Display*, pp.1–8, June 2008. <http://www.icad.org/Proceedings/2008/EnomotoIkeda2008.pdf>.

[15] M. Camras, “Approach to recreating a sound field,” *J. Acoust. Soc. Am.*, vol.43, no.6, pp.1425–1431, Nov. 1968.

[16] T. Kimura and K. Takehi, “Effects of directivity of microphones and loudspeakers in sound field reproduction based on wave field synthesis,” *Proc. Int. Cong. Acoust.*, no.RBA-15-011, pp.1–6, Madrid, Spain, Sept. 2007.

[17] 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.

[18] 木村敏幸, “波面合成技術の研究動向,” *日本音響学会誌*, vol.67, no.11, pp.538–543, Nov. 2011.

[19] B.B. Baker and E.T. Copson, “The mathematical theory of Huygens’ principle,” pp.23–26, Oxford University Press, London, UK, second edition, 1950.

[20] T. Kimura, K. Takehi, K. Takeda, and F. Itakura, “Subjective assessments for the effect of the number of channel signals on the sound field reproduction used in wavefield synthesis,” *Proc. Int. Cong. Acoust.*, no.Th.P1.18, pp.3159–3162, Kyoto, Japan, April 2004.

[21] 木村敏幸, 笈一彦, 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.

[22] M. Morimoto, “The relation between spatial impression and the precedence effect,” *Proc. Int. Conf. Auditory Display*, no.SS#2-1, pp.297–306, Kyoto, Japan, July 2002.

[23] M. Morimoto and Y. Ando, “On the simulation of sound localization,” *J. Acoust. Soc. Jpn. (E)*, vol.1, no.3, pp.167–174, July 1980.

[24] J. Blauert, “Spatial hearing,” pp.39–50, MIT Press, Cambridge, Mass., revised edition, 1997.

(平成 xx 年 xx 月 xx 日受付)



木村 敏幸 (正員)

1998 年名大・工・物化卒. 2000 年同大学院・人間情報・修士課程了. 2005 年同大学院・人間情報・博士課程了. 日本学術振興会特別研究員 (PD), 名大研究員, 東京農工大特任助手を経て, 2007 年より NICT・有期研究員, 現在に至る. 3D システム, 空間知覚, アレイ信号処理に関する研究に従事. 博士 (学術). 日本音響学会, 日本バーチャルリアリティ学会, ヒューマンインタフェース学会, Audio Engineering Society 各会員.

Abstract In a conventional personal 3D sound field reproduction system using wave front synthesis technique, numerous loudspeakers are placed around the listener. However, it is difficult to construct an audio-visual system since the number of transmitted channels for remote control is enormous and loudspeakers are visible in the listener's field of vision. In this paper, a personal 3D sound field reproduction system based on the concept of wave front synthesis technique using eight directional microphones is proposed for remote control. The localized performance of the proposed system was evaluated by the localization test. As a result, it was indicated that the localized performance of the proposed technique was enough to construct the system.

Key words remote control, sound field reproduction, wave front synthesis, directional microphone, localization test