全天球ステレオ録音再生方式の提案

電気・化学専攻　医用工学領域　2281214 大網　拓海

（主査：桐生　昭吾教授　副査：和多田　雅哉教授，桃沢　愛　准教授）

1. はじめに

近年仮想現実（VR）、拡張現実（AR）などの XR（クロスリアリティ）技術が注目されている。エンターテイメント産業や国土計画・都市計画、防災など幅広い分野で活用される中[1,2]、仮想空間の構築といった映像面における技術革新が目覚ましい一方、さらに臨場感や没入感を増すべくより簡便な空間音響技術が求められる。しかし原理的に 2 次元の空間情報を得るには 3 点計測が必須であり、同様に 3 次元の空間情報を得るには 4 点計測が必須である。[3,4,5]そのため既存の 4 チャンネル以上を用いる空間音響技術は多くのネットワークリソースやストレージを要するだけでなく、特殊なフォーマットを用いる必要があるといった課題が残る。[6]先行研究では空間音響を簡便に扱う手法として、一般的な録音機器のフォーマットであるステレオ 2 チャンネルで 2 次元の空間情報を扱うことが可能なパノラミックサウンド方式と、その収音システムであるパノラミックサウンド収音システムが提案された。[7,8,9,10]しかしマイクロフォンの種類や構造によってチャンネル間位相差の値が不安定になり左右方向の推定が困難となるだけでなく、推定できるのはあくまで 2 次元の空間情報であり、上下（仰角）の空間情報が失われてしまうという課題が残った。そこで 4 点計測に先行研究で提案されたシステムの概念を仮想的に取り入れる事で、4 チャンネルによって収録された 3 次元の空間情報を保持したまま、従来のステレオフォーマットと下位互換性を持つデータエンコード方式である、全天球ステレオを新たに提案した。本研究では提案手法を用いて実際に音源定位が可能であるか聴覚心理実験によって評価を行った。

2. 原理

2.1 正四面体マイクロフォンの構成

収音デバイスは Fig.1 のような 4 つの無指向性モノラルマイクロフォンが正四面体の各頂点に、3 次元的に配置されている。よってシステムの水平方向と仰角方向を含めた 3 次元の空間音響情報が推定可能である。本システムでは正四面体の各辺（隣り合うマイクロフォンの間隔）を 15 mm とした。

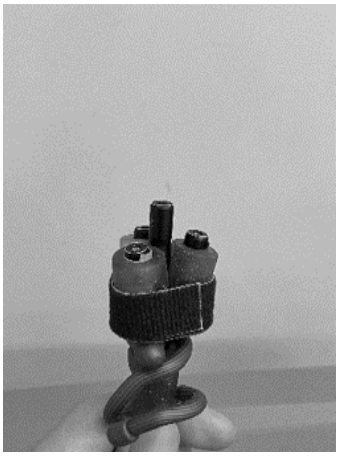

Fig.1　提案手法に用いた収音デバイス

本研究では 4 つの audio-technica 製 AT9903 で構成されている。

2.2 多点計測におけるチャンネル間位相差を用いた波の伝搬方向の推定

任意の観測点に置かれたセンサー素子に対する波の伝搬イメージ及び各種パラメータを Fig.2 に示す。説明を簡単にするため 3 点計測による算出で説明する。信号源はセンサー素子が設置された観測点から十分に離れていると仮定し、波の到来方向を表す法線ベクトルを Fig.中矢印で記す。隣り合うセンサー素子間の距離を r (本研究では 15mm)、Fig.の上側を正面方向（0°）とし、波の到来方向の角度をθとする。ここで Sensor A と Sensor C それぞれに波が到達するまでの伝搬距離差をL_{AC}とすると、θとL_{AC}の関係は以下の式で表すことができる。

$$L_{AC} = r\cos\left(\theta + \frac{\pi}{6}\right)$$

同様に、Sensor B と Sensor C それぞれに波が到達するまでの伝搬距離差をL_{BC}とすると、以下の式で表すことができる。

$$L_{BC} = r\cos\left(\theta - \frac{\pi}{6}\right)$$

各素子で計測された波の位相（α_A, α_B, α_C）を用いて, 3 つの素子の内 1 つを基準（ここでは Sensor C とする）としたチャンネル間位相差（$\alpha_A - \alpha_C$, $\alpha_B - \alpha_C$）を求める。チャンネル間位相差は各素子間の伝搬距離差に置き換えることができるので、チャンネル間位相差の比は以下の式で表すことができる。

$$\frac{\alpha_A - \alpha_C}{\alpha_B - \alpha_C} = \frac{L_{AC}}{L_{BC}} = \frac{rcos\left(\theta + \frac{\pi}{6}\right)}{rcos\left(\theta - \frac{\pi}{6}\right)} = \frac{cos\left(\theta + \frac{\pi}{6}\right)}{cos\left(\theta - \frac{\pi}{6}\right)}$$

右辺の変数は θ のみであることがわかる。つまりチャンネル間位相差の比は波の伝搬方向 θのみに依存する。よってチャンネル間位相差から波の伝搬方向を推定することが出来る。

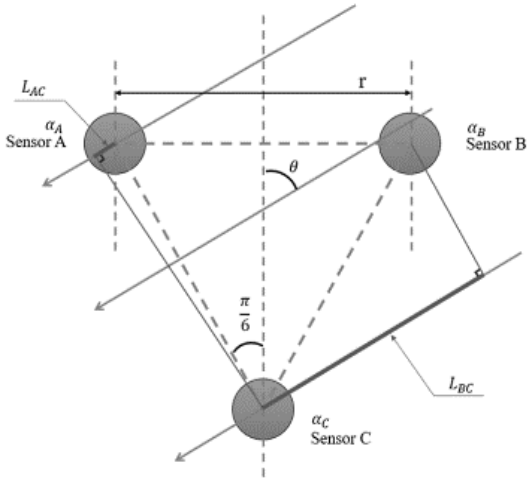

Fig.2 任意の観測上のセンサー素子に対する波の伝搬イメージ
矢印は波の伝搬方向、θ はセンサーを垂直に通る点線と伝搬方向とのなす角、
r はセンサー間の距離を示す。

2.3 ステレオフォーマットへの符号化

収音デバイスで計測した信号は 4 チャンネル信号となっており、全天球型の空間情報を含んでいる。この空間情報を損なうことなくステレオフォーマットへ符号化するため、仮想ステレオマイクロフォンシステムを用い符号化を行う。Fig.3 の様な仮想のステレオマイクロフォンシステムを想定する。各マイクロフォンの感度特性は単一指向性であり、位相特性は方向によらず一定とする。また各マイクロフォンは片方が前方に、他方が後方に向けられており、距離 r の間隔を挟んで左右に配置されている。

4 点計測によって求めた周波数 bin ごとの伝搬方向と伝搬速度から、仮想ステレオマイクロフォンシステムの出力トラックである Track A、Track B の信号を推定することが出来る。前後方向に関する 1 次元の空間情報はチャンネル間レベル差、左右方向に関する 1 次元の空間情報はチャンネル間位相差として記録される。よって 2 次元の空間情報、つまり水平角の伝搬方向をチャンネル間レベル差とチャンネル間位相差から一意に求めることが出来る。また音波が仰角方向から伝搬する際、伝搬方向の仰角の絶対値が大きくなればなるほどチャンネル間レベル差とチャンネル間位相差は共に小さくなる。しかし仰角の符号、つまり波の伝搬方向が上半球からなのか、下半球からなのかを判定することができない。よって仰角の符号情報を別の手段でエンコードする必要がある。まず仰角の符号は＋か－かの二択となるので、周波数 bin ごとに 1 ビットの情報である。また FFT のポイント数と周波数 bin の数は同じである。よって仰角の符号情報は、オーディオデータ 1 サンプルあたり 1 ビットの情報として埋め込めばよい。そこで仮想ステレオマイクロフォンシステムの Track A のオーディオ信号（16 ビット）の最小ビット値を、伝搬方向が下半球（仰角の符号が - ）なら 0、上半球（仰角の符号が ＋ ）なら 1 と信号処理プログラム上で変換することで、 上下方向の判定が可能になるようにした。こ

のようにして 3 次元の伝搬方向（水平角と仰角）をステレオ信号に埋め込むことができる。

しかしここで求められる伝搬方向は伝搬速度を音速とみなして求めたものである。伝搬速度が速くなるほど、仰角が大きくなるのと同様に、チャンネル間レベル差、チャンネル間位相差はともに小さくなる。よって伝搬速度が決まらなければ伝搬方向の仰角を決めることが出来ない。

そこで本研究では、伝搬速度の情報量を切り詰め、C を音速としたとき、比較的遅い（$\frac{2c}{\sqrt{3}}$ 未満）か、比較的速い（$\frac{2c}{\sqrt{3}}$ 以上）かの 2 値とし、これを仮想ステレオマイクロフォンシステムの Track B のオーディオ信号（16 ビット）の最小ビット値に符号化した。Track B の最小ビット値を伝搬速度が比較的遅ければ 0、比較的早ければ 1 と信号処理プログラム上で変換することで、伝搬速度を 2 値化した。再生する際には Track B の最小ビット値を参照して仰角を補正することで、伝搬方向と伝搬速度を再現した。

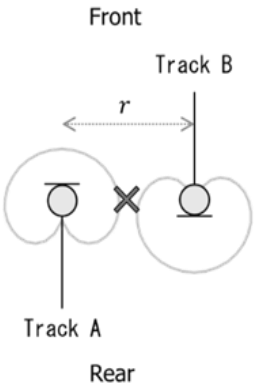

Fig.3 仮想マイクロフォンシステムの模式図

3. 研究方法

3.1 提案手法を用いた検査用音源の録音

録音は一般的な残響のある空間とされる東京都市大学世田谷キャンパス 3 号館医用電子工学研究室内の会議室で行った。正四面体マイクロフォンシステムを会議室の中央に配置した後、スピーカーを Fig.4 の様に正四面体マイクロフォンシステムから 100cm 離した位置に配置し、信号を再生した。検査用音源として量子化ビット数 24、サンプリング周波数が 44.1 kHz の信号を録音した。使用した音声[11]とスピーカーの配置は Table 1 に示す。

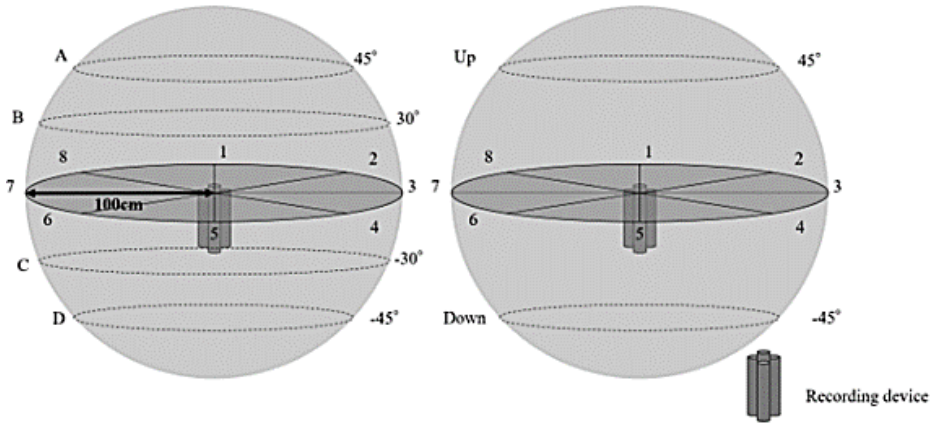

Fig.4 録音デバイスと音源の位置関係

音源は収音デバイスから半径 100cm に設置される。

Table 1 使用音源と配置

	使用音源、音源の位置	
問1	ドラム（水平面315°）	
問2	ドラム（水平面90°）	
問3	アコースティックギター（水平面0°、仰角30°）	
問4	セミの鳴き声（水平面 225°、仰角45°）	
問5	ドラム（水平面135°、俯角45°）	
問6	男性の朗読（水平面270°、俯角30°）	
問7	男性の朗読（水平面180°）	セミの鳴き声（水平面0°、仰角45°）
問8	男性の朗読（水平面180°）	エレキギター（水平面225°、仰角45°）
問9	グランドピアノ（水平面0°）	コウロギ（水平面90°、俯角45°）
問10	カラスの鳴き声（水平面315°）	エレキギター（水平面225°、俯角45°）

3.2 音源定位実験

20 歳代の健聴者 7 人（男性：5 人、女性：2 人）を被験者とし、音源定位実験を実施した。実験では検査用音源を、再生ソフトを通じてイヤホンで聴取し、どの音源がどの方向に存在するかを推定した。録音した検査用音源は開発した信号処理プログラムによりステレオフォーマットに符号化した。再生時には開発した再生ソフトを用いてステレオフォーマットへ符号化された音源を再生した。聴取者は再生ソフト上で水平面内に対しパンニング操作をすることが可能となっている。また仰角の推定を容易にするため、使用した再生ソフトでは、向いている方向に対して±60°のロール（前後軸まわりの回転）が可能となっている。被験者は刺激音を 10 種聞き、どの音源がどの方向から鳴っているか推測した。

3.3 実験結果と考察

水平角に対する問 1~10 の結果を Fig.5 左図に示す。各グラフの横軸は音源が水平角のどこに配置されたかを示し、縦軸は被験者の水平角に対する回答を示す。また音源の正しい位置は△で、被験者の回答は○で示される。Fig.5 左図より聴取実験全体として、実際に配置された音源と被験者の回答には正の相関が示された。このことから音源の水平面方向を推定することは可能であり、 2 次元の空間情報を扱え

ると考えられる。 次に仰角方向に対する問 1~10 の結果を Fig.5 右図に示す。横軸は音源が仰角方向のどこに配置されたかを示し、縦軸は被験者の仰角方向に対する回答を示す。Fig.5 右図から、仰角方向に対して実際に配置された音源と被験者の回答には相関関係が示されなかった。原因として考えられるのは、配置された 2 つの音源が仰角方向に対して最大でも 45° の角度差であり、聴取者にとって上下の判別が困難であったと考えられる。

そこで 2 つの音源を同じ水平面方向で上下別々に（仰角差を大きくして）配置することで、聴取者が上下判別を行うことが出来るかを調査するため、追加で音源定位実験を行った。

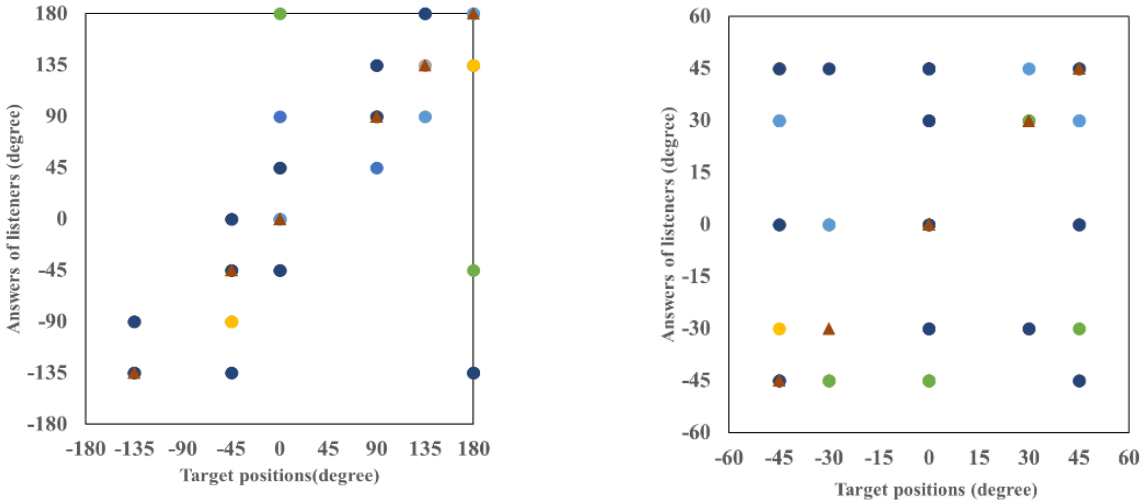

Fig.5 水平角に対する被験者の回答（左）と
仰角に対する被験者の回答（右）

3.4 追加実験用音源の録音

録音環境や機材は変えず Table 2 のようにスピーカーを配置し音源を録音した。Table 2 より 2 つの音源は同一の水平方向に配置されているが、仰角方向の符号を変えて配置した。本実験では 2 種類の音源を用意した。1 つはグランドピアノ単体で構成される、曲中の強弱が少ない、ドイツ民謡「かえるの合唱」とした。もう 1 つは「かえるの合唱」とは対照的に、曲調の激しい BGM をフリー音源サイト[12]でダウンロードし、利用規約に則り使用した。

Table 2 追加実験における音源の種類と配置

	音源の位置	
	BGM	ピアノ
問1	水平面0°、仰角60°	水平面0°、俯角60°
問2	水平面135°、俯角60°	水平面135°、仰角60°
問3	水平面315°、仰角60°	水平面315°、俯角60°
問4	水平面45°、俯角60°	水平面45°、仰角60°
問5	水平面180°、俯角60°	水平面180°、仰角60°
問6	水平面225°、仰角60°	水平面225°、俯角60°
問7	水平面270°、俯角60°	水平面270°、仰角60°
問8	水平面90°、仰角60°	水平面0°、俯角60°

3.5 実験結果と考察

20 歳代の健聴者 11 人（男性：9 人、女性：2 人）を被験者とし、音源定位実験を実施した。被験者は刺激音を指定された水平方向の音源を 8 種類聞き、同一の水平方向に配置された 2 つの音源が、それぞれ上下どちらの方向から鳴っているか推測した。聴取実験の結果を Fig.に示す。横軸は 2 つの音源が配置された水平面方向を示し，縦軸は被験者の回答率を示す。Fig.より収音システムの 0°、135°、315°方向以外の水平角に音源が配置された場合、被験者 11 人全員が 2 つの音源の上下判別が可能であったことが示された。また 0°、135°、315°方向の水平角に音源が配置された場合にも、8 割以上の被験者が正しい伝搬方向を回答した。

以上の結果から、2 つの音源の仰角差を 120°にした際、聴取者は上下の推定が可能であると考えられ、提案手法によって全天球型空間情報を扱えることが出来ると考えられる。

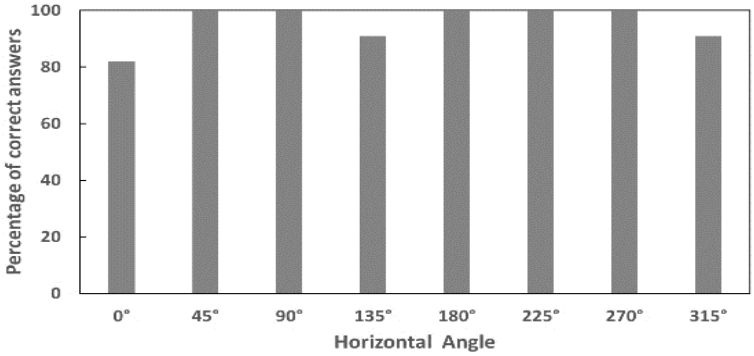

Fig.6 追加実験における各水平面方向に対する上下判断課題の正答率

4. まとめ

本研究では、4 チャンネル録音により得られた空間情報を一般的なステレオフォーマットに符号化する全天球ステレオ方式を提案し、音源定位が可能か聴覚心理実験により評価を行った。聴取実験の結果から、提案手法により，ステレオフォーマットでも全天球型空間情報のうち，少なくとも 2 次元の空間情報は安定して扱えることが示唆された。一方仰角方向を含めた 3 次元の空間情報は 2 つの信号源の仰角差が 120°の際、扱うことが出来ると示唆された。しかし本稿の結果はあくまで特定の音源および実験環境下におけるものであり、信号源の仰角を変える等様々な条件での動作検証が必要である。

参考文献

[1]総務省,「(1) メタバース」, 「第 1 部　特集　新時代に求められる強靭・健全なデータ流通社会の実現に向けて 」, 2023,
<https://www.soumu.go.jp/johotsusintokei//whitepaper/ja/r05/html/nd131210.html ＞（参照 2024-1-17）

[2]総務省，「(2) デジタルツイン」，「第 1 部　特集　新時代に求められる強靭・健全なデータ流通社会の実現に向けて 」, 2023, <https://www.soumu.go.jp/johotsusintokei//whitepaper/ja/r05/html/nd131220.html> （参照 2024-1-17）

[3] 宮浦尊士，内海伶香，蘆原 郁，桐生 昭吾，“ミニチュアヘッドシミュレータアレイによる騒音制御，”信学技報, 117 (170), EA2017-31, pp. 27-32, Aug. 2017

[4] Takashi Miyaura, Reika Uchimi, Kaoru Ashihara, Shogo Kiryu，“Spatial sound morphing by using a miniature head simulator array,” IEICE Tech. Rep., EA2017-68, pp. 55-60, Nov. 2017

[5] Reika Uchimi, Hiroaki Sato, Kaoru Ashihara, Shogo Kiryu，“Sound sensing by using panoramic sound morphing and panoramic beamformer，” Proc. Auditory Res. Meeting, The Acoustical Society of japan, 49, 8, H-2019-107, Dec. 2019.

[6] 西村竜一，アンビソニックス， 映像情報メディア学会誌 68, 8, 616-620, 2014

[7] 堀地悠矢，亀井文太，蘆原 郁，桐生昭吾，“2 チャンネルシステムによる 2 次元空間音響の収音・再生，”第 23 回 日本音響学会 関西支部 若手研究者交流研究発表会，Dec. 2020.

[8] 藤原汐里，佐藤 渉，大網拓海，蘆原 郁，桐生昭吾，“パノラミックステレオ信号のマルチスピーカー再生”第 24 回 日本音響学会 関西支部 若手研究者交流研究発表会，Dec. 2021.

[9] 大網拓海，佐藤　渉，桐生昭吾，蘆原郁，“自由音場における 2 チャンネル 2 次元パノラミックサウンド収音システムのチャンネル間特性”，信学技報 Vol. 122No. 144，2022 年 7 月 28 日，pp.1-6

[10] 佐藤　渉，大網拓海，桐生昭吾，蘆原郁，“実環境における 2 チャンネル 2 次元パノラミックサウンド収音システムの特性評価”，信学技報 Vol. 122No. 144，2022 年 7 月 28 日，pp.7-12

[11] On-Jin ~音人~ <https://on-jin.com/sound/ani.php> (参照 2024-1-11)

[12]BGMer<https://bgmer.net/songlist?genre=&keyword=edm&type=music>
(参照 2024-1-11)