

小特集—頭部伝達関数とその応用—

# マイクロホンアレイと頭部伝達関数に基づくバイノーラル信号収音再生法の設計理論\*

サルバドル セザル, 坂本 修一, トレビーニョ ホルヘ,  
鈴木 陽一 (東北大学電気通信研究所)\*\*

43.60.Fg, Uv; 43.66.Pn

## 1. はじめに

何等かの方法で収音した信号を、その場にいたであろう聴取者の両耳に生成される音圧、すなわちバイノーラル信号として再現する方法をバイノーラル信号収音再生法と呼ぶ。本稿は、近年活発に研究が進む、マイクロホンアレイ収音信号と聴取者個々人の頭部伝達関数（以下 HRTF）の連成により両耳信号を再現する新しいバイノーラル信号収音再生法について述べる。なお、本稿は *Acoustical Science & Technology (AST)* 誌掲載の総説論文 [1] に基づいて、システム構築に関する部分を中心に要約したものである。詳細は原論文を参照されたい。

バイノーラル信号収音再生法の中で最も古典的な方法はダミーヘッド録音である。しかし、この方法は、HRTF の個人性 [2] の反映が困難であり、かつ、音の方向定位をより正確に行う上で重要な役割を果たしている聴取者の動き [3–5] に起因するバイノーラル信号の動的変化が実現できないという問題点が存在する。これに対し、聴取者個々人の形状に合ったダミーヘッドを聴取者の動きに応じて機械的に動かす TeleHead [6] も提案されているが、収音後に聴取する場合に聴取時の動きに対応できないという問題がある。

これに対し、マイクロホンアレイで収音した信号を処理することで、聴取者個々人に適した音空間情報を再現することができれば、聴取時の聴取者の動きは電子的に反映可能である。これにより、

聴取者の動きを反映したリアルな 3 次元音空間情報再生が複数の聴取者に対し、実時間のみならず収音後でも可能となる。MTB (motion tracked binaural) [7] はこのようなコンセプトに基づいて構築された最初の例と考えられる。

このようなマイクロホンアレイを用いたバイノーラル信号収音再生法において、リアルな音空間情報を実現するには聴取者個々人の HRTF を精密に合成することが重要となる。そのためには、各人の HRTF を全方位にわたり密度高く与えたデータセットが必須となる。近年の測定及び計算技術の向上により、そのような HRTF データセットの準備も困難ではなくなっており [8, 9]、実際に様々な研究機関において HRTF データセットが公開され、手軽に利用できるようになってきた [10, 11]。

このような背景から、マイクロホンアレイ収音信号と聴取者の HRTF データセットを連成させて両耳信号を再現する新しいバイノーラル信号収音再生法の研究が盛んになっている。この方法は、聴取者の動きを反映したリアルな音空間情報の再生がいつでも誰にでも実現できるという優れた特徴を持つ。以下、この新しいバイノーラル信号収音再生法について、従来研究を概観し、一般的定式化と分類、解説を行う。

## 2. 3次元聴覚ディスプレイとバイノーラル信号収音再生法

バイノーラル信号を聴取者に提示する技術は聴覚ディスプレイ (auditory display) の中でも音空間情報を提示する 3 次元聴覚ディスプレイの一類型と考えられる [12, 13]。3 次元聴覚ディスプレイで提示される音空間情報は、実際に収録した情報の場合もあれば、バーチャルなものの場合もある。バイノーラル信号合成法も例外ではないが、実音場再生でもバーチャル提示の場合でも、典型的に

\* Design theory for binaural synthesis: Combining microphone array recordings and head-related transfer function datasets.

\*\* César D. Salvador, Shuichi Sakamoto, Jorge Treviño and Yôichi Suzuki (Research Institute of Electrical Communication, Tohoku University, Sendai, 980–8577) e-mail: {salvador@ais., saka@ais., jorge@ais., yoh@}riec.tohoku.ac.jp

は、モノフォニックな音源信号にその音源の位置から耳までの伝達関数を反映させることにより出力信号を合成する。

一方、本稿が対象とする一連のバイノーラル信号收音再生法では、前段で述べた方法とは全く異なるアプローチを取る。両耳信号が、部屋やホールなどの特性も含めた音空間情報と、聴取者個人が持つ HRTF 特性との組み合わせにより形成されると考えて音空間情報の再現を行う。ここで、前者はマイクロホンアレイで収録される情報で、後者は個人々の HRTF データセットから得られるもの、言うなれば、聴取者の頭部や胴体によって生じる指向特性である。以下、このような考えに基づき論述を進める。

### 3. HRTF データセットとマイクロホンアレイを用いたバイノーラル信号收音再生

バイノーラル信号收音再生法は、(1) HRTF モデル化法、(2) マイクロホンアレイ收音信号モデル化法の 2 種類に大別できる。

(1) は、ある聴取者の HRTF データセットから頭部や胴体によって生じる指向特性を求め、アレイの個々のマイクロホン入力信号の重み付けフィルタの線形接合としてその指向特性を再現する手法である。従って、算出された重みフィルタは、HRTF データセットにおいて、個々のマイクロホンの位置から HRTF の測定位置への伝達関数に相当する。SENZI [14–16], VAH (virtual artificial head) [17, 18], バイノーラルビームフォーミング [19–23] 等がこの形の提案である。

一方 (2) は、HRTF データセット中の個々の HRTF に掛け合わせるべき信号をアレイへの入力信号から計算する。言うなれば、HRTF の測定位置にバーチャル音源を配置して、その駆動信号を重ね合わせの理 [24] に基づいてアレイ入力信号から算出することに相当する。そのため、個々の HRTF の測定位置からアレイの各マイクロホンの位置への伝達関数を求める。バイノーラルアンビソニクス [25–29] はこの代表例である。また、BPLIC [30] などの收音手法と ADVISE [31] などの再生手法を組み合わせて実現することも可能となる。

以下、両手法について、マイクロホンと HRTF の位置が離散的な場合の定式化を行う。この場合、マイクロホンと HRTF の数の多少によって解く

表-1 バイノーラル信号收音再生設計法の分類

	手法	
	HRTF モデル化法	マイクロホンアレイ收音信号モデル化法
マイクロホン数 < HRTF 測定数 ( $M < L$ )	優決定系	劣決定系
マイクロホン数 > HRTF 測定数 ( $M > L$ )	劣決定系	優決定系

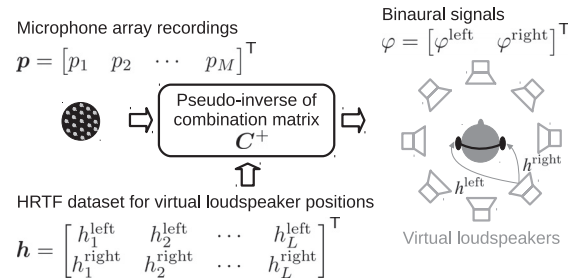


図-1 バイノーラル信号收音再生法の概念図

べき式の性質に違いが生ずる (表-1)。ちなみに、連続的な場合の定式化については AST 誌総説論文 [1] に記載した。なお、これ以降は周波数領域において定式化を進めるが、煩雑さを避け、周波数は陽に記載しないものとする。

#### 3.1 離散的な条件でのバイノーラル信号收音再生法の定式化

本節では、とある境界面  $\mathcal{A}$  上の任意の点  $\vec{a}$  で収録された音圧  $p$  と、別の領域面  $\mathcal{B}$  上の任意の点  $\vec{b}$  で測定された HRTF データセット  $h = \{h^{\text{left}}, h^{\text{right}}\}$  とを用いて、聴取者のバイノーラル信号  $\varphi = \{\varphi^{\text{left}}, \varphi^{\text{right}}\}$  を導出するための一般式の記述を行う。任意の連続した領域  $\mathcal{A}, \mathcal{B}$  上のそれぞれのすべての点で  $p, h$  が得られているものとする、バイノーラル信号  $\varphi$  は次式のように一般化して表現することができる。

$$\varphi = \int_{\vec{b} \in \mathcal{B}} \int_{\vec{a} \in \mathcal{A}} \overline{h(\vec{b})} C^\dagger(\vec{a}, \vec{b}) p(\vec{a}) d\vec{a} d\vec{b} \quad (1)$$

ここで、上線は複素共役を、 $\dagger$  は逆変換を表す。  $C$  は收音点  $\vec{a}$  と HRTF の測定点  $\vec{b}$  との音伝搬特性の変換関数を表しており、 $\mathcal{A}$  と  $\mathcal{B}$  の領域の形状、特性に依存したものとなる。この式は、バイノーラル信号  $\varphi$  の導出が、 $C^\dagger$  を求める逆問題として定式化されることを示している。

式 (1) は收音点も HRTF も領域上のすべての点

で連続に得られている場合を想定している。しかしシステムを構築する際には、HRTF の測定点もマイクロホンアレイ上の收音点（マイクロホンの位置）も離散化されることから、式 (1) を離散化して表現することが必要になる。すなわち、アレイ上に配置した  $M$  個のマイクロホンで収録された音信号と、空間上の  $L$  点で測定された HRTF を用いてバイノーラル信号を合成することとなる。概念図を図-1 に示す。このような条件において、バイノーラル信号、

$$\varphi = \begin{bmatrix} \varphi^{\text{left}} & \varphi^{\text{right}} \end{bmatrix}^{\text{T}} \quad (2)$$

は、

$$\varphi = \overline{\mathbf{h}}^{\text{T}} \mathbf{C}^+ \mathbf{p} \quad (3)$$

として離散化される。ここで、 $\text{T}$  は転置行列を表す。

式 (3) において、 $M$  個のマイクロホンで收音された音信号は次式で表すことができる。

$$\mathbf{p} = \begin{bmatrix} p_1 & p_2 & \cdots & p_M \end{bmatrix}^{\text{T}} \quad (4)$$

ここで、 $p_m$  ( $m = 1, 2, \dots, M$ ) は、境界面  $\mathcal{A}$  上の任意の位置  $\{\vec{a}_m\}_{m=1,2,\dots,M}$  上に離散的に配置されたマイクロホンで收音された音圧である。一方、HRTF に関しても、

$$\mathbf{h} = \begin{bmatrix} \mathbf{h}^{\text{left}} \\ \mathbf{h}^{\text{right}} \end{bmatrix}^{\text{T}} = \begin{bmatrix} h_1^{\text{left}} & h_2^{\text{left}} & \cdots & h_L^{\text{left}} \\ h_1^{\text{right}} & h_2^{\text{right}} & \cdots & h_L^{\text{right}} \end{bmatrix}^{\text{T}} \quad (5)$$

と行列を用いて表現できる。ここで、 $h_\ell^{\text{left}}$ 、及び、 $h_\ell^{\text{right}}$  ( $\ell = 1, 2, \dots, L$ ) は、領域面  $\mathcal{B}$  上に離散的に配置された任意の点  $\{\vec{b}_\ell\}_{\ell=1,2,\dots,L}$  から両耳までの HRTF を表している。先にも述べたように、 $\mathbf{C}$  はすべての  $\vec{a}_m$  と  $\vec{b}_\ell$  との間の音伝搬特性の変換行列であることから、本論文で取り扱うバイノーラル信号收音再生法の本質は、この行列の擬似逆行列  $\mathbf{C}^+$  を求めることと言い換えることができる。 $\mathbf{C}^+$  を求める際に、HRTF に着目して算出した  $L \times M$  行列の  $\mathbf{C}_{\text{HRTF}}$  を用いる手法が HRTF モデル化法であり、マイクロホンアレイでの収録音に着目して算出した  $M \times L$  行列である  $\mathbf{C}_{\text{mic}}$  を用いる手法がマイクロホンアレイ收音信号モデル化法となる。これらの二つの手法の信号処理の概略を図-2 に示す。以下、それぞれの手法について詳述する。

### 3.1.1 HRTF モデル化法

HRTF モデル化法では、マイクロホンアレイ收音信号に HRTF から算出された重みフィルタをかけることでバイノーラル信号  $\varphi$  を算出する。これを式に表すと以下ようになる。

$$\varphi = \overline{\mathbf{w}}^{\text{T}} \mathbf{p} \quad (6)$$

ここで重みフィルタ  $\mathbf{w}$  は、

$$\mathbf{w} = \begin{bmatrix} \mathbf{w}^{\text{left}} \\ \mathbf{w}^{\text{right}} \end{bmatrix}^{\text{T}} = \begin{bmatrix} w_1^{\text{left}} & w_2^{\text{left}} & \cdots & w_M^{\text{left}} \\ w_1^{\text{right}} & w_2^{\text{right}} & \cdots & w_M^{\text{right}} \end{bmatrix}^{\text{T}} \quad (7)$$

で表現され、

$$\mathbf{C}_{\text{HRTF}} \mathbf{w} = \mathbf{h} + \epsilon_{\text{HRTF}} \quad (8)$$

を解くことにより次のように求められる。

$$\mathbf{w} = \mathbf{C}_{\text{HRTF}}^+ \mathbf{h} \quad (9)$$

なお、 $\mathbf{C}_{\text{HRTF}}^+$  は、 $\vec{b}_\ell$  から  $\vec{a}_m$  までの音響伝達関数を表す行列であることに注意する必要がある。式 (9)、(6) と、式 (3) を見比べると、式 (3) で求めるべき  $\mathbf{C}^+$  は次式となる。

$$\mathbf{C}^+ = \overline{\mathbf{C}_{\text{HRTF}}^+}^{\text{T}} \quad (10)$$

### 3.1.2 マイクロホンアレイ收音信号モデル化法

一方、マイクロホンアレイ收音信号モデル化法では、HRTF データセットに含まれる個々の HRTF に対する駆動信号を、アレイ收音信号から求めることでバイノーラル信号  $\varphi$  を算出する。これを式に表すと以下ようになる。

$$\varphi = \overline{\mathbf{h}}^{\text{T}} \mathbf{u} \quad (11)$$

ここで

$$\mathbf{u} = \begin{bmatrix} u_1 & u_2 & \cdots & u_L \end{bmatrix}^{\text{T}} \quad (12)$$

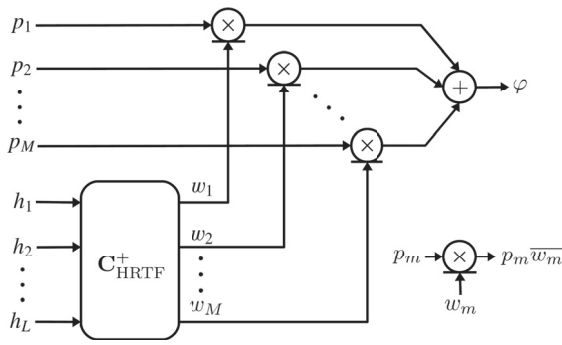
で表される式 (11) 中の駆動信号  $\mathbf{u}$  は、

$$\mathbf{C}_{\text{mic}} \mathbf{u} = \mathbf{p} + \epsilon_{\text{mic}} \quad (13)$$

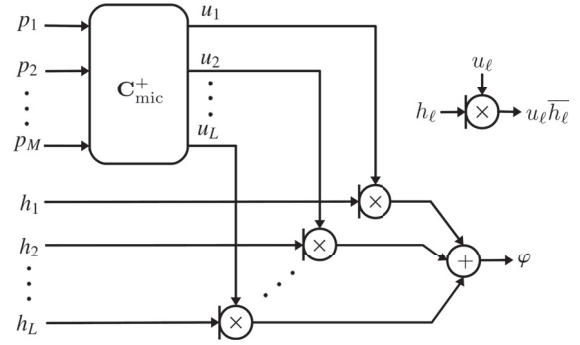
を解くことで次式のように求められる。

$$\mathbf{u} = \mathbf{C}_{\text{mic}}^+ \mathbf{p} \quad (14)$$

なお、 $\mathbf{C}_{\text{mic}}^+$  は、 $\mathbf{C}_{\text{HRTF}}^+$  とは逆に、 $\vec{a}_m$  から  $\vec{b}_\ell$  ま



(a) HRTF のモデル化法



(b) マイクロホンアレイ收音信号モデル化法

図-2 マイクロホンアレイ收音信号 ( $\mathcal{A}$  上の  $p_m$ ) と HRTF データセット ( $\mathcal{B}$  上の  $h_\ell$ ) により構成される 2 種類のバイノーラル信号收音再生法の構成図

での音響伝達関数を表す行列であることに注意したい。式 (14), (11) と、式 (3) を見比べると、式 (3) で求めるべき  $C^+$  は次式となる。

$$C^+ = C_{\text{mic}}^+ \quad (15)$$

もし、 $C_{\text{HRTF}}^+$  と  $C_{\text{mic}}^+$  が同一空間上で定義される、すなわち、マイクロホンが存在する領域  $\mathcal{A}$  と HRTF を測定した領域  $\mathcal{B}$  が同じで、かつ、マイクロホン数  $M$  と HRTF の測定点数  $L$  が同じ場合、式 (10) と式 (15) から、次の関係が導出される。

$$C_{\text{HRTF}}^+ = \overline{C_{\text{mic}}^+}^\top \quad (16)$$

しかし、実際にシステムを構築する際には、マイクロホン数と HRTF の測定数は必ずしも等しくなるわけではない。このような場合、数の多少と用いる手法の違いによって、いずれか一方が劣決定系となり、もう一方が優決定系となる (表-1)。

#### 4. 球状マイクロホンアレイを用いたバイノーラル信号收音再生法

前節で定式化したバイノーラル信号收音再生システムを実現する上で、マイクロホンアレイの形 [32, 33] と、HRTF データセット [10, 11] の測定位置のいずれも球状にすることが一般的である。特に、收音用球状アレイでは、バイノーラル信号收音再生法で重要な役割を果たす擬似逆行列を安定して求められることから、剛球上にマイクロホンを等密度に配置したものが多い [32, 33]。そのような球状アレイを用いることで、全方向にわたり等しい空間解像度で操作が可能であることに加え、波動方程式を解く際に空間解像度をスケール

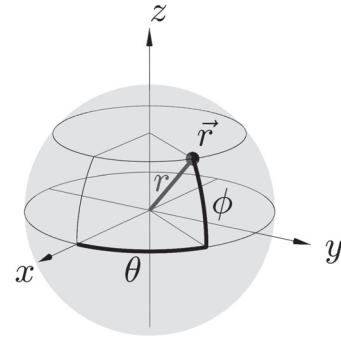


図-3 本論文で使用する球座標系  
原点はマイクロホンアレイ中心、及び、聴取者の頭部中心である。

ブルに変更することが可能となるため、目的に応じた音空間再現を実現することができる [34]。特に後者は重要な利点であり、マイクロホンアレイと HRTF データセットに応じて、收音側と再生側で適切な空間解像度を独立に選択することができる [19–23, 25–29]。これ以降、この空間解像度を次数  $\nu$  とする。

これ以降の定式化に用いる球座標系を図-3 に示す。空間上の点  $\vec{r}$  は、半径  $r$ 、方位角  $\theta \in [-\pi, \pi]$ 、仰角  $\phi \in [-\frac{\pi}{2}, \frac{\pi}{2}]$  を用い、 $\vec{r} = (r, \theta, \phi)$  と表される。方位角と仰角を一つの変数  $\Omega = (\theta, \phi)$  で表せば  $\vec{r} = (r, \Omega)$  となる。

半径  $a$  の剛球上にマイクロホンを配置した球状アレイにおいて、位置  $\vec{a}_m = (a, \Omega_m)$  に配置されたマイクロホンで收音した音圧は式 (4) の  $\mathbf{p}$  における  $p_m$  として表される。一方、無響室内に設置された半径  $b$  のスピーカアレイにおいて、位置  $\vec{b}_\ell = (b, \Omega_\ell)$  に設置されたスピーカから測定された左右耳の HRTF は式 (5) の  $h_\ell^{\text{left}}$  と  $h_\ell^{\text{right}}$  で表される。これらから  $C_{\text{HRTF}}^+$  と  $C_{\text{mic}}^+$  を解析的に

求める際は、剛球による音の散乱モデル [34] から  $\mathbf{C}_{\text{HRTF}}$  と  $\mathbf{C}_{\text{mic}}$  を求めた上で、擬似逆行列を計算して求めることになる。ただし、擬似逆行列が安定して求まらないことがあるため、Tikhonov の正則化法 [35] を用いて求めることが多い。この部分の詳細に関しては、AST 誌総説論文 [1] の第 4 節を参照されたい。

上記の手続きを HRTF モデル化法で実現すると、式 (9) の  $\mathbf{C}_{\text{HRTF}}^+ = [c_{m\ell}^{\text{HRTF}+}]$  の個々の要素は、

$$c_{m\ell}^{\text{HRTF}+} = \frac{\exp(-jkb)}{b} \alpha_m \beta_\ell \times \sum_{\nu=0}^{[\sqrt{L}]-1} (2\nu+1) \overline{R_\nu^{\text{reg}}(a, b, k)} P_\nu(\cos \Theta_{m\ell}) \quad (17)$$

と表されることになる。ここで、 $\Theta_{m\ell}$  は、マイクロホン位置  $\vec{a}_m$  とスピーカ位置  $\vec{b}_\ell$  とのなす角である。式 (17) 中、総和項の角度成分は、 $\nu$  次のルジャンドル多項式  $P_\nu$  によって定まることが分かる。一方、距離成分は次に示す正則化フィルタによって定まる。

$$R_\nu^{\text{reg}} = \frac{|R_\nu|^2}{|R_\nu|^2 + \lambda^2} \times \frac{1}{R_\nu} \quad (18)$$

$$R_\nu = \frac{h_\nu(kb)}{ka^2 h'_\nu(ka)}$$

ここで、 $\lambda$  は正則化パラメータであり、 $h_\nu$  は次数  $\nu$  の第 2 種球ハンケル関数である。また  $'$  は微分を表す。式 (17) 中の  $\alpha_m$  と  $\beta_\ell$  は、連続場での表現である式 (1) における  $d\vec{a}$  と  $d\vec{b}$  を離散化したものと考えることができる。更に、この  $\alpha_m$ 、 $\beta_\ell$  は、いずれも正で、 $\sum_m \alpha_m = \sum_\ell \beta_\ell = 1$  と正規化された値となる。また、球面上の空間サンプリングに依存するため、球状マイクロホンアレイのマイクロホンの配置や、HRTF 測定点の空間的な配置に大きく関係する。

一方、マイクロホンアレイ收音信号モデル化法では、式 (14) の  $\mathbf{C}_{\text{mic}}^+ = [c_{\ell m}^{\text{mic}+}]$  の個々の要素は、

$$c_{\ell m}^{\text{mic}+} = \frac{\exp(jkb)}{b} \alpha_m \beta_\ell \times \sum_{\nu=0}^{[\sqrt{M}]-1} (2\nu+1) R_\nu^{\text{reg}}(a, b, k) P_\nu(\cos \Theta_{m\ell}) \quad (19)$$

と表されることになる。式 (17) と同様に、式 (19) の角度成分は  $\nu$  次のルジャンドル多項式  $P_\nu$  で表され、距離成分は式 (18) で表される。更に、HRTF モデル化法と同様に、 $\alpha_m$  と  $\beta_\ell$  は球面上の空間サンプリングに依存した値となる。

以上に示した 2 種類のバイノーラル信号收音再生法は、球座標系で実現されている様々なバイノーラル信号收音再生法の理論的基盤を与えるものである。前者の HRTF モデル化法は  $\mathbf{C}_{\text{HRTF}}$  を対角化する手法であり、バイノーラルビームフォーミング [19, 22] として知られているのに対し、後者は  $\mathbf{C}_{\text{mic}}$  を対角化する手法でありバイノーラルアンビソニックス [25, 27, 28] として知られているものである。

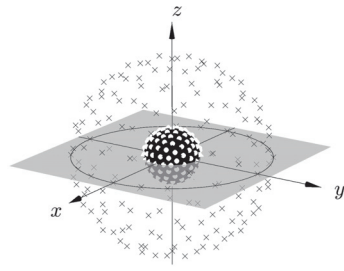
## 5. バイノーラル信号收音再生法実装例

本節では、マイクロホンの個数  $M$  と HRTF の測定点数  $L$  を変えたときの音空間再現精度を検証する。使用するマイクロホンアレイ ( $a$ ) と HRTF 測定位置 ( $b$ ) の球の半径は、それぞれ 8.5 cm, 1.5 m とした。マイクロホンの位置と HRTF の測定位置は、正 20 面体の各面を等分割する手法に基づいて決定した [14–16, 36]。決定した配置の例を図-4(a) に示す。図中の白い点がマイクロホンの位置を表し、×印が HRTF 測定点を示す。

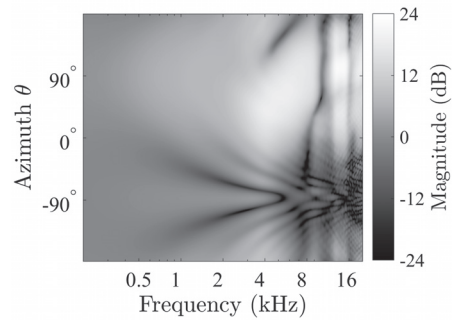
このような配置で、マイクロホンアレイを剛球と仮定して各マイクロホンの收音信号を解析的に求め [37]、境界要素法を用いて HRTF を計算した [9]。その上で、各マイクロホン数、HRTF 測定数で図-4(a) 中の水平面上に記載した円上にインパルス音源を配置した際のバイノーラル信号を合成し、同じくその点から測定した HRTF と比較した。誤差がなければこれらは一致することになる。

ターゲットとなるバイノーラル信号を図-4(b) に示す。また、各手法で合成されたバイノーラル信号を、図-4(c)~4(h) に示す。なお、擬似逆行列を求める際の正則化パラメータ  $\lambda$  は  $1 \times 10^{-3}$  に設定した。

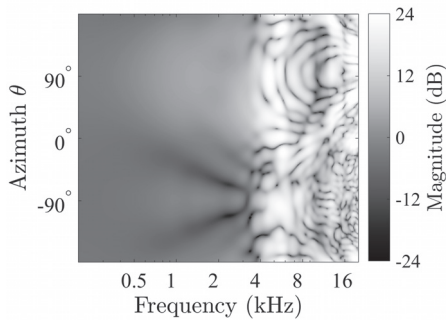
マイクロホン数と HRTF 測定数が同数 ( $M = L$ ) の場合、 $\mathbf{C}_{\text{HRTF}}$  と  $\mathbf{C}_{\text{mic}}$  は同じ次元数の正方行列となるため、図-4(e) と図-4(f) を見ると分かる通り、両手法での合成結果は同一となる。それに対し、マイクロホン数が HRTF 測定数より少ない ( $M < L$ ) 場合、 $\mathbf{C}_{\text{HRTF}}$  は優決定系となる一方



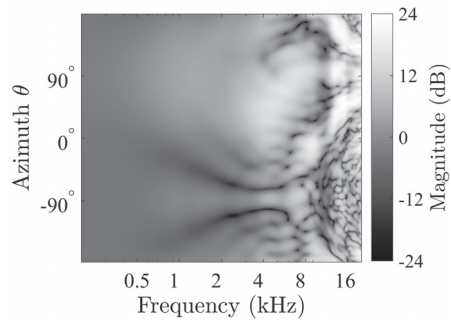
(a) 音源 (×) とマイクロホンアレイ (図中中央) の配置図



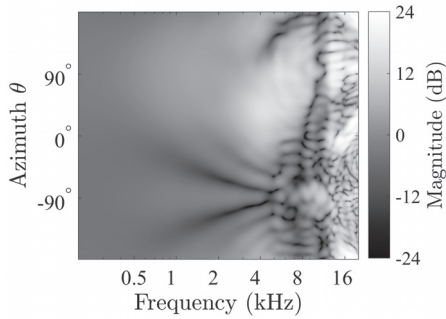
(b) バイノーラル信号 (ターゲット)



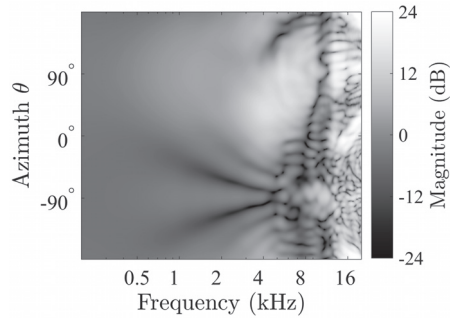
(c) HRTF モデル化法 ( $M = 92, L = 252$ )



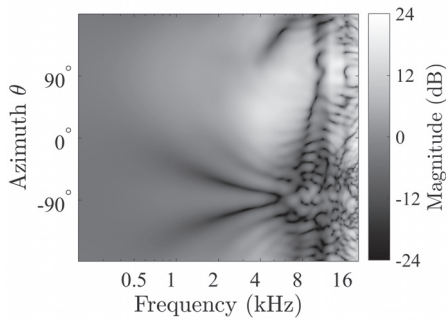
(d) マイクロホンアレイ收音信号モデル化法 ( $M = 92, L = 252$ )



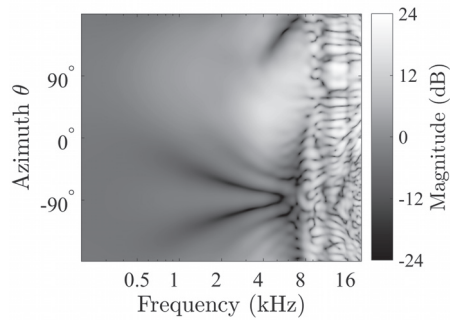
(e) HRTF モデル化法 ( $M = L = 252$ )



(f) マイクロホンアレイ收音信号モデル化法 ( $M = L = 252$ )



(g) HRTF モデル化法 ( $M = 492, L = 252$ )



(h) マイクロホンアレイ信号モデル化法 ( $M = 492, L = 252$ )

図-4 2種類の手法で実現したバイノーラル信号收音再生法の空間再現精度

で  $C_{mic}$  は劣決定系となる。図-4(c) と図-4(d) との比較から分かるとおり、高周波数領域において HRTF モデル化法では大きな合成誤差が観測されることになる。反対に、マイクロホン数が HRTF 測定数より多い ( $M > L$ ) 場合、 $C_{HRTF}$  が劣決定系になる一方で  $C_{mic}$  が優決定系となる。その結果、図-4(g) と図-4(h) に示されるとおり、逆にマイクロホン收音信号モデル化法において高い周波数領域で合成精度が低下する。

### 6. ま と め

本稿では、マイクロホンアレイと HRTF データセットを用いて聴取者の両耳の信号を合成するバイノーラル信号收音再生法について概観した。この方法は、マイクロホンの位置と HRTF の測定点との音伝搬特性の変換関数を対象に、その逆を求める逆問題として定式化される。定式化においては、HRTF から聴取者の指向特性を求めてマイクロホン入力信号の重み付けフィルタとして表現する HRTF モデル化法と、HRTF の測定位置に配置したバーチャル音源の駆動信号として表現するマイクロホンアレイ收音信号モデル化法の二つのアプローチがあることを述べ、マイクロホンの個数と HRTF の測定数の関係で合成される音空間の精度が異なることを示した。これらの結果をまとめると表-2 のようになる。

本稿で述べたバイノーラル信号收音再生法を更に進めて、人工的に残響を付加したり、ある方向の音を強調、抑制するといった、様々な空間特性を付加、編集するような手法へと展開することも可能である。これは、本稿で扱った変換行列  $C_{HRTF}$ 、

表-2 バイノーラル信号收音再生における 2 種類の手法の定式化

		手法	
		HRTF モデル化法	マイクロホンアレイ收音信号モデル化法
定式化		$C_{HRTF}$ は HRTF データセット $\mathbf{h}$ と線形結合： $C_{HRTF} \mathbf{w} = \mathbf{h} + \epsilon_{HRTF}$	$C_{mic}$ はマイクロホンアレイ収録信号 $\mathbf{p}$ と線形結合： $C_{mic} \mathbf{u} = \mathbf{p} + \epsilon_{mic}$
	合成アルゴリズム	1) 重みフィルタ： $\mathbf{w} = C_{HRTF}^+ \mathbf{h}$ 2) バイノーラル信号： $\varphi = \overline{\mathbf{w}^T} \mathbf{p}$	1) 駆動信号： $\mathbf{u} = C_{mic}^+ \mathbf{p}$ 2) バイノーラル信号： $\varphi = \overline{\mathbf{h}^T} \mathbf{u}$

$C_{mic}$  にパラメータを加えて操作することに相当する。このような観点に基づき、近年の関連研究では、 $C_{HRTF}$ 、 $C_{mic}$  の最適な表現法を模索する試みが続いている。マイクロホンアレイ信号のスパース性を減らす圧縮センシング [38] などはその一例であり、HRTF データセットの空間特性を損なうことなく、アレイの指向特性を最適化するなどといった手法が提案されているのが現状である。

### 謝 辞

本研究の一部は、科研費 (No. 24240016, 16H01736)、及び、日中韓フォーサイト事業による HRTF データセットの計算にあたりご協力いただいた京都大学大谷真准教授に感謝する。

### 文 献

- [1] C. D. Salvador, S. Sakamoto, J. Treviño and Y. Suzuki, "Design theory for binaural synthesis: Combining microphone array recordings and head-related transfer function datasets," *Acoust. Sci. & Tech.*, **38**, 51–62 (2017).
- [2] M. Morimoto and Y. Ando, "On the simulation of sound localization," *J. Acoust. Soc. Jpn. (E)*, **1**, 167–174 (1980).
- [3] H. Wallach, "On sound localization," *J. Acoust. Soc. Am.*, **10**, 270–274 (1939).
- [4] 川浦淳一, 鈴木陽一, 浅野 太, 曾根敏夫, "頭部伝達関数の模擬によるヘッドホン再生音像の定位," *音響学会誌*, **45**, 756–766 (1989).
- [5] Y. Iwaya, Y. Suzuki and D. Kimura, "Effects of head movement on front-back error in sound localization," *Acoust. Sci. & Tech.*, **24**, 322–324 (2003).
- [6] I. Toshima, H. Uematsu and T. Hirahara, "A steerable dummy head that tracks three-dimensional head movement: TeleHead," *Acoust. Sci. & Tech.*, **24**, 327–329 (2003).
- [7] V. R. Algazi, R. O. Duda and D. M. Thompson, "Motion-tracked binaural sound," *J. Audio Eng. Soc.*, **52**, 1142–1156 (2004).
- [8] D. N. Zotkin, R. Duraiswami, E. Grassi and N. A. Gumerov, "Fast head-related transfer function measurement via reciprocity," *J. Acoust. Soc. Am.*, **120**, 2202–2215 (2006).
- [9] M. Otani and S. Ise, "Fast calculation system specialized for head-related transfer function based on boundary element method," *J. Acoust. Soc. Am.*, **119**, 2589–2598 (2006).
- [10] P. Majdak, Y. Iwaya, T. Carpentier, R. Nicol, M. Parmentier, A. Roginska, Y. Suzuki, K. Watanabe, H. Wierstorf, H. Ziegelwanger and M. Noisternig, "Spatially oriented format for acoustics: A data exchange format representing head-related transfer functions," *AES 134th Convention*, May (2013).
- [11] K. Watanabe, Y. Iwaya, Y. Suzuki, S. Takane and S. Sato, "Dataset of head-related transfer functions measured with a circular loudspeaker array," *Acoust. Sci. & Tech.*, **35**, 159–165 (2014).
- [12] R. D. Shilling and B. Shinn-Cunningham, "Virtual auditory displays," in *Handbook of Virtual Environment Technology* (Lawrence Erlbaum Associates, New York, 2002), pp. 65–92.

- [13] 館 暲, 佐藤 誠, 廣瀬通孝, バーチャルリアリティ学 (コロナ社, 東京, 2010).
- [14] S. Sakamoto, S. Hongo, R. Kadoi and Y. Suzuki, "SENZI and ASURA: New high-precision sound-space sensing systems based on symmetrically arranged numerous microphones," *Proc. 2nd Int. Symp. Universal Communication*, pp. 429–434 (2008).
- [15] S. Sakamoto, S. Hongo and Y. Suzuki, "3D sound-space sensing method based on numerous symmetrically arranged microphones," *IEICE Trans. Fundam.*, **E97-A**, 1893–1901 (2014).
- [16] S. Sakamoto, S. Hongo, T. Okamoto, Y. Iwaya and Y. Suzuki, "Sound-space recording and binaural presentation system based on a 252-channel microphone array," *Acoust. Sci. & Tech.*, **36**, 516–526 (2015).
- [17] E. Rasumow, M. Blau, M. Hansen, S. van de Par, S. Doclo, V. Mellert and D. Püschel, "Smoothing individual head-related transfer functions in the frequency and spatial domains," *J. Acoust. Soc. Am.*, **135**, 2012–2025 (2014).
- [18] E. Rasumow, M. Hansen, S. v. d. Par, D. Püschel, V. Mellert, S. Doclo and M. Blau, "Regularization approaches for synthesizing HRTF directivity patterns," *IEEE/ACM Trans. Audio Speech Lang. Process.*, **24**, 215–225 (2016).
- [19] W. Song, W. Ellermeier and J. Hald, "Using beamforming and binaural synthesis for the psychoacoustical evaluation of target sources in noise," *J. Acoust. Soc. Am.*, **123**, 910–924 (2008).
- [20] C. D. Salvador, S. Sakamoto, J. Treviño, J. Li, Y. Yan and Y. Suzuki, "Accuracy of head-related transfer functions synthesized with spherical microphone arrays," *Proc. Meet. Acoust.*, Vol. 19, No. 1, Apr. (2013).
- [21] N. R. Shabtai and B. Rafaely, "Generalized spherical array beamforming for binaural speech reproduction," *IEEE/ACM Trans. Audio Speech Lang. Process.*, **22**, 238–247 (2014).
- [22] N. R. Shabtai, "Optimization of the directivity in binaural sound reproduction beamforming," *J. Acoust. Soc. Am.*, **138**, 3118–3128 (2015).
- [23] C. D. Salvador, S. Sakamoto, J. Treviño and Y. Suzuki, "Evaluation of white noise gain in a binaural system for microphone arrays," 音講論集, 3-7-12, pp. 401–404 (2016.9).
- [24] G. H. Koopmann, L. Song and J. B. Fahnlne, "A method for computing acoustic fields based on the principle of wave superposition," *J. Acoust. Soc. Am.*, **86**, 2433–2438 (1989).
- [25] R. Duraiswami, D. N. Zotkin, Z. Li, E. Grassi, N. A. Gumerov and L. S. Davis, "High order spatial audio capture and its binaural head-tracked playback over headphones with HRTF cues," *AES 119th Convention*, New York, USA, Oct. (2005).
- [26] A. Avni, J. Ahrens, M. Geier, S. Spors, H. Wierstorf and B. Rafaely, "Spatial perception of sound fields recorded by spherical microphone arrays with varying spatial resolution," *J. Acoust. Soc. Am.*, **133**, 2711–2721 (2013).
- [27] C. D. Salvador, S. Sakamoto, J. Treviño and Y. Suzuki, "Numerical evaluation of binaural synthesis from rigid spherical microphone array recordings," *AES Int. Conf. Headphone Technology*, Aalborg, Denmark, Aug. (2016).
- [28] C. D. Salvador, S. Sakamoto, J. Treviño and Y. Suzuki, "Spatial accuracy of binaural synthesis from rigid spherical microphone array recordings," *Acoust. Sci. & Tech.*, **38**, 23–30 (2017).
- [29] B. Bernschutz, A. V. Giner, C. Porschmann and J. Arend, "Binaural reproduction of plane waves with reduced modal order," *Acta Acust. united Ac.*, **100**, 972–983 (2014).
- [30] S. Takane, Y. Suzuki and T. Sone, "A new method for global sound field reproduction based on Kirchhoff's integral equation," *Acta Acust. united Ac.*, **85**, 250–257 (1999).
- [31] S. Takane, Y. Suzuki, T. Miyajima and T. Sone, "A new theory for high definition virtual acoustic display named ADVISE," *Acoust. Sci. & Tech.*, **24**, 276–283 (2003).
- [32] J. Meyer and G. Elko, "A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield," *Proc. IEEE ICASSP*, II, Orlando, FL, USA, pp. 1781–1784 (2002).
- [33] B. Rafaely, "Analysis and design of spherical microphone arrays," *IEEE Trans. Speech Audio Process.*, **13**, 135–143 (2005).
- [34] E. G. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography* (Academic Press, London, 1999).
- [35] A. Neumaier, "Solving ill-conditioned and singular linear systems: A tutorial on regularization," *SIAM Rev.*, **40**, 636–666 (1998).
- [36] R. Sadourny, A. Arakawa and Y. Mintz, "Integration of the nondivergent barotropic vorticity equation with an icosahedral-hexagonal grid for the sphere 1," *Mon. Weather Rev.*, **96**, 351–356 (1968).
- [37] R. O. Duda and W. L. Martens, "Range dependence of the response of a spherical head model," *J. Acoust. Soc. Am.*, **104**, 3048–3058 (1998).
- [38] E. J. Candes and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, **25**, 21–30 (2008).