

音源情報を推定するための
拡散受音に関する研究

丹羽 健太

目次

第1章	序論	1
1.1	研究背景	1
1.2	従来のアレイ信号処理	6
1.2.1	観測信号のモデル化	6
1.2.2	線形フィルタリング	11
1.2.3	受音方式	17
1.3	本研究の考え方	21
1.3.1	着眼点と研究方針	21
1.3.2	本論文の構成	23
第2章	拡散受音の原理	25
2.1	相互情報量最大化のための空間相関行列の最適構造	25
2.2	拡散受音による相互情報量最大化	30
2.3	数値シミュレーションによる拡散受音の効果の検証	32
2.3.1	シミュレーション条件	33
2.3.2	シミュレーション結果	34
2.4	本章の結論	37
第3章	拡散受音に基づく信号強調	39
3.1	空間相関行列と信号出力パワーの関係	39

3.2	拡散受音におけるフィルタ設計	40
3.3	数値シミュレーションによる拡散受音に基づく信号強調効果の検証	46
3.3.1	シミュレーション 1: 拡散受音における雑音抑圧性能の調査	46
3.3.2	シミュレーション 2: フィルタ長と雑音抑圧性能の関係の調査	50
3.4	本章の結論	55
第 4 章	包囲型反射構造体を用いた拡散受音の実装	57
4.1	包囲型反射構造体アレイ	57
4.2	包囲型反射構造アレイの性能調査	59
4.2.1	実験条件	59
4.2.2	実験結果	61
4.3	本章の結論	62
第 5 章	拡散受音に基づく音源距離推定	73
5.1	室内における観測信号のモデル化	73
5.2	音源距離と空間相関行列の構造の関係	75
5.3	音源距離推定アルゴリズム	78
5.3.1	周波数領域 ICA に基づく空間相関行列の分離	78
5.3.2	空間相関行列の固有値分布を用いた音源距離推定	80
5.4	実験	81
5.4.1	実験条件	81
5.4.2	実験結果	82
5.5	本章の結論	87
第 6 章	結論	89
	謝辞	93

参考文献

95

研究業績

103

第1章 序論

1.1 研究背景

人がコミュニケーションしたり、周りの状況を判断するために、音は重要な役割を果たしている。人が状況を理解するために受聴した音信号から得る情報として、話し相手の発話音声（源信号）、位置関係（方向や距離）等がある。本論文では、源信号、音源までの位置（方向や距離）といった受信した信号から得られる情報を総じて「音源情報」と呼ぶ。本研究の目的は、欲しい音源情報を欲しい人に届けられるような音響信号処理を実現することである。例えば、スタジアムにいるサッカー選手の声やボールを蹴っている音を收音できるようになれば、あたかも選手と同じ位置にいるかのようなコンテンツ視聴が可能になるだろう。また、多人数が参加する遠隔会議では、発言者が收音装置から同じ方向にいたとしても、特定の発言者の声を收音することで、疎外感の少ない円滑なコミュニケーションを実現できるだろう。

音源情報を得るための手段として、音波を電気信号に変換するマイクロホンを用いる。電話、計算機、ビデオカメラ、カラオケ端末といった身の回りの多くの通信機器/家電製品に、マイクロホンが組み込まれている。音源情報を推定するための効果的な方法として、複数のマイクロホンを用いて受信した信号を解析するアレイ信号処理がある [1, 2, 3, 4]。アレイ信号処理では、マイクロホン間に生じる位相や振幅の差を利用することにより、雑音抑圧、残響除去、音源到来方向の推定といった技術を実現してきた。2000年代には、音声会議端末 [5, 6, 7]、携帯電話/スマートホン [8]、TV 電話やゲーム機の映像/音入力デバイス [9] を中心として、2~4本程度のマイクロホンを搭

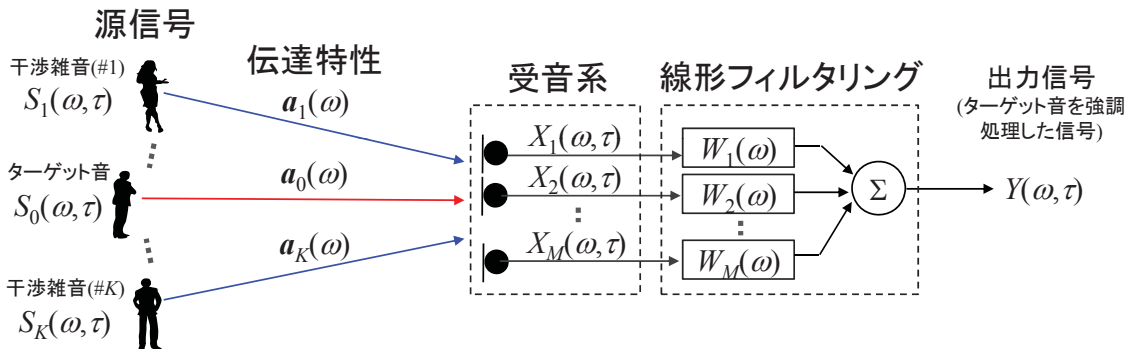


図 1.1: アレイ信号処理を用いた信号強調

載した製品が商用化されてきた。例えば，会議室のような静穏な環境下で，インターネットを通じて遠隔地と会話することを想定し，源信号やアレイに対する音声の到来方向を推定するための技術が確立し，実用化されてきた。また，2010年代には，Apple社の Siri や NTT docomo のしゃべってコンシェル [10] といった音声認識を用いた情報検索サービスが一般ユーザに普及し，雑音に対して頑健な音声認識を実現するために，アレイ信号処理の活用が期待されている。

図 1.1 に，アレイ信号処理を用いた信号強調の処理構成を示す。(i) 複数のマイクロホンを空間的に配置した受音系と，(ii) 複数マイクロホンの信号群を入力として音源情報を出力する信号処理の二部で構成される。これまでのアレイ信号処理に関する研究では，主に，2～4本程度のマイクロホンを中空に配置した受音系（中空配置型アレイ）を用いてきた。受音系の後段にある信号処理を高度化し，音源情報の推定性能を向上するアプローチが主流であった。信号処理に関する基本的な方法として，線形フィルタリングがある。線形フィルタリングでは，ターゲットとなる源信号を強調して收音するためのフィルタを設計し，観測信号と畳み込むことで出力信号を得る。線形フィルタリングによる信号強調処理については，1.2.1 節で説明する。フィルタを設計するための古典的な方式として，遅延和法，最小分散法，最尤法がある [1, 2, 3, 11, 12, 13, 14]。フィルタ設計に関する具体的な演算については，1.2.2 節で説明する。また，ターゲット音に関する情報（到来方向等）が事前に用意されていない場合に，音源情報を解析す

る方法として、ブラインド音源分離に関する研究例も多数報告されている。例えば、独立成分分析に基づくブラインド音源分離 [15, 16, 17, 18] を用いることで、各音源を分離して收音するためのフィルタを高精度に推定できることが知られている。また、線形フィルタリングの雑音抑圧性能をより高めるための信号処理として、非線形ポストフィルタに関して盛んに研究が行われている [19, 20, 21, 22, 23, 24]。

しかし、これまでの受信系の後段にある「信号処理を高度化する」というアプローチを適用するだけでは、音源情報の推定が困難な場合がある。推定困難と考えられる音源情報について2つの例を以下に示す。

1. 多雑音環境における源信号の推定：

多数の干渉雑音が存在する中で、目的とする特定の源信号（ターゲット音）を強調し、近接した位置にある干渉雑音を抑圧して收音することは困難である。それは、観測信号間の関係が音源の位置に応じて変化する割合が少ないため、ターゲット音とその他の干渉雑音を分離するための空間的な手掛かりを受音した信号からほとんど得られないことに起因する。

2. 音源とアレイ間の距離（音源距離）の推定：

音源距離に応じて、観測信号間の関係が変化する割合が少ないため、アレイに対して同方向にある複数の源信号を分離して收音したり、音源距離を推定することは困難であった。

以上の例で共通している課題は、音源情報を推定するための手掛かりが、受信信号から欠落してしまっているということである。この課題は、これまでの受信系の後段にある信号処理を高度化するというアプローチを単に継承したままでは解決されない。

受信系の後段にある信号処理を高度化することで音源情報を推定する従来研究とその課題について概説してきたが、2000年代で普及してきた受信系の進化に繋がる音響技術について説明する。(i) 多入出力デバイスの普及と(ii) 剛球バッフル型アレイの研

究開発に関して、技術概要と着目すべき考え方について概説する。

(i) 多入出力デバイスの普及:

多観測信号 (100 ~ 200ch 程度) をリアルタイムに信号処理するための音声入出力デバイスが市販されるようになってきた。その背景として、USB、IEEE 1394、Audio Engineering Society (AES) で標準化された MADI[25] といった高速伝送システム規格の普及や Steinberg 社の ASIO、Microsoft の Direct Sound といったオーディオデバイスのドライバインタフェースと開発キットの普及、Puredata[26, 27] 等のリアルタイム信号処理用途のプログラミングツールの普及による貢献が大きい。ここで着目したいのは、これまでのアレイ信号処理では、せいぜい4本程度のマイクロホンを用いた音源情報解析に関して研究されてきたのに対して、100 ~ 200本の多マイクロホンを用いて観測した信号をリアルタイムで解析し、音源情報を推定するという手段を扱えるようになったことである。2000年前後より、多観測信号を用いた収録/再生技術に関する検討に関して、多くの研究が行われている [28, 29, 30, 31, 32, 33, 34, 35, 36, 37]。

(ii) 剛球バッフル型アレイの研究開発:

2002年に G. W. Elko らが剛球バッフル型アレイに関する基本原理を発表し、多くの関連研究が見られるようになってきた [38, 39, 40, 41]。市販されている剛球バッフル型アレイを図 1.2 に示す。球面上に幾何対称性を持たせて 30本程度のマイクロホンを配置した中空配置型アレイ (球面对称アレイ) は、球面調和関数展開に基づく指向性フィルタの設計に適した性質を持っている。球面調和関数の直交性を利用すると、様々な次数の球面調和関数の組み合わせにより、所望の指向特性に基づく収録をするためのフィルタを解析的に導出可能である。しかし、球面对称アレイを用いた場合、特定の周波数において、数値計算上、不安定なフィルタが生成されることが知られている (禁止周波数の発生)。禁止周波数の発生を回避するための一つの方法として、剛球バッフル上にマイクロホンを配置した受信系 (剛球バッフル型アレイ) を用いる方法がある。剛球バッフルを設置することで発生した回折波により、禁止周波数における制御不安



図 1.2: 剛球バッフル型アレイ (mh acoustics 社)

定性を緩和できることが解析的に示されている。ここで着目したい考え方は、信号処理の高度化では回避できない問題（e.g. 禁止周波数の発生）を解決するために、受音系の構造を変えること（e.g. 剛球バッフルの設置）が有効であるという実験結果である。剛球バッフル型アレイに関する研究については、1.2.3 節で触れる。

多マイクロホンを用いることや受音系の構造を変えることで、信号処理の高度化では解決されない問題を解決するという考え方は、多くのアレイ信号処理の従来研究では、軽視されてきた考え方である。本論文では、信号処理を高度化させるのではなく、音源情報を解析するためにどのように信号を受音すべきなのかという問題に取り組む。受音系を進化させることで、これまで推定することが困難とされてきた音源情報を解析することが可能になり、欲しい音源情報を欲しい人に届けられるようになるだろう。

次節以降では、本論文全般で使用する記号の定義や、従来のアレイ信号処理について概説する。1.2 節では、線形フィルタリングに基づく信号強調について説明する。1.3 節では、これまで推定困難とされてきた音源情報を推定するための本研究の方針について述べる。

1.2 従来のアレイ信号処理

アレイを用いて観測した出力信号群を用いて、音源情報を推定するための基本的な信号処理として、線形フィルタリングがある。線形フィルタリングでは、 M 個のマイクロホンを用いて観測した信号それぞれに異なるフィルタを畳み込み、加算することで、出力信号を生成する。本節では、線形フィルタリングを用いた信号強調を取り上げ、受音系と信号処理の従来技術を概観する。1.2.1 節では観測信号をモデル化し、1.2.2 節では代表的なフィルタ設計方式について述べる。1.2.3 節では、受音系に関する従来研究について説明する。

1.2.1 観測信号のモデル化

本節では、線形フィルタリングを用いた信号強調における観測過程や線形フィルタリングによる信号出力についてモデル化する。なお、本節で定義した記号は、2章以降でも使用する。

(a) 受音系

マイクロホンアレイを用いた観測過程についてモデル化する。図 1.3 に示すように、強調して收音したい1つのターゲット音と抑圧して收音したい K 個の干渉雑音がある音場を想定する。これらの信号により構成される音場を M (≥ 2) 本のマイクロホンを用いて観測する。ターゲット音と m 番目のマイクロホンの間の時間応答(インパルス応答)を $a_{0,m}(l)$ 、 k ($k = 1, \dots, K$) 番目の干渉雑音と m 番目のマイクロホンの間のインパルス応答を $a_{k,m}(l)$ とする。ただし、インパルス応答長を L と表す。AD 変換後の時間インデックス l におけるターゲット音の源信号を $s_0(l)$ 、 k 番目の干渉雑音の源信号を $s_k(l)$ とする。また、 m 番目のマイクロホンに混在する背景雑音(センサの内部雑音を含む)を $n_m(l)$ とする。背景雑音は、チャンネル毎に独立で平均0のガウス分布に従うこととする。 m 番目のマイクロホンにおける観測信号 $x_m(l)$ は、次式でモデル化される。

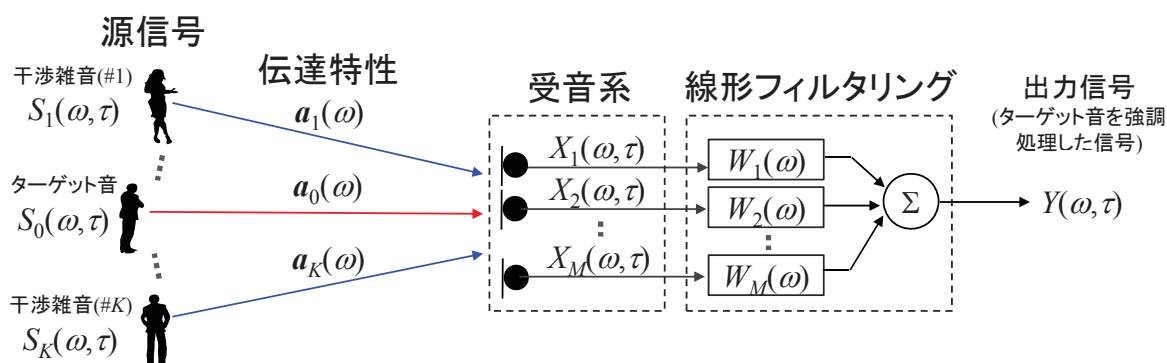


図 1.3: アレイ信号処理による信号強調 (再掲)

$$x_m(l) = \sum_{i=0}^{L-1} a_{0,m}(i) s_0(l-i) + \sum_{k=1}^K \sum_{i=0}^{L-1} a_{k,m}(i) s_k(l-i) + n_m(l) \quad (1.1)$$

短時間フーリエ変換により，周波数領域に変換すると，観測信号は次式でモデル化される。

$$X_m(\omega, \tau) = A_{0,m}(\omega) S_0(\omega, \tau) + \sum_{k=1}^K A_{k,m}(\omega) S_k(\omega, \tau) + N_m(\omega, \tau) \quad (1.2)$$

ここで， ω, τ は，それぞれ角周波数，フレーム時間のインデックスを表す。また， $X_m(\omega, \tau)$ ， $S_0(\omega, \tau)$ ， $S_k(\omega, \tau)$ ， $N_m(\omega, \tau)$ は，それぞれ $x_m(l)$ ， $s_0(l)$ ， $s_k(l)$ の時間周波数表現を表す。 $A_{0,m}(\omega)$ ， $A_{k,m}(\omega)$ は，それぞれ m 番目のマイクロホンとターゲット音， k 番目の干渉雑音の間の周波数応答 (伝達特性) を表す。式 (1.2) を行列形式で書き直すと，次式になる。

$$\begin{aligned} \mathbf{x}(\omega, \tau) &= \mathbf{a}_0(\omega) S_0(\omega, \tau) + \mathbf{A}_I(\omega) \mathbf{s}_I(\omega, \tau) + \mathbf{N}(\omega, \tau) \\ &= \mathbf{A}(\omega) \mathbf{s}(\omega, \tau) + \mathbf{n}(\omega, \tau) \end{aligned} \quad (1.3)$$

ここで，

$$\mathbf{x}(\omega, \tau) = [X_1(\omega, \tau), \dots, X_M(\omega, \tau)]^T \quad (1.4)$$

$$\mathbf{a}_k(\omega) = [A_{k,1}(\omega), \dots, A_{k,M}(\omega)]^T \quad (1.5)$$

$$\mathbf{A}_I(\omega) = [\mathbf{a}_1(\omega), \dots, \mathbf{a}_K(\omega)] \quad (1.6)$$

$$\mathbf{A}(\omega) = [\mathbf{a}_0(\omega), \dots, \mathbf{a}_K(\omega)] \quad (1.7)$$

$$\mathbf{s}_I(\omega, \tau) = [S_1(\omega, \tau), \dots, S_K(\omega, \tau)]^T \quad (1.8)$$

$$\mathbf{s}(\omega, \tau) = [S_0(\omega, \tau), \dots, S_K(\omega, \tau)]^T \quad (1.9)$$

$$\mathbf{n}(\omega, \tau) = [N_1(\omega, \tau), \dots, N_M(\omega, \tau)]^T \quad (1.10)$$

T は転置を表す。

(b) 線形フィルタリング

受音系から出力された M 個の観測信号 $x(\omega, \tau)$ を入力とし，ターゲット音を強調するための信号処理として，線形フィルタリングを説明する。観測信号 $x_m(l)$ と線形フィルタ $w_m(i)$ を畳み込み，加算することで，出力信号 $y(l)$ を得る。

$$y(l) = \sum_{m=1}^M \sum_{i=0}^{J-1} w_m(i) x_m(l-i) \quad (1.11)$$

ここで， J はフィルタ長を表す。

周波数領域で表現すると，出力信号 $Y(\omega, \tau)$ は次式で算出される。

$$Y(\omega, \tau) = \mathbf{w}^H(\omega) \mathbf{x}(\omega, \tau) \quad (1.12)$$

ここで， H は共役転置であり，

$$\mathbf{w}(\omega) = [W_1(\omega), \dots, W_M(\omega)]^T \quad (1.13)$$

である。

音源情報の解析精度を測るための尺度として，信号出力パワー $P(\omega)$ を定義する [1, 2]。

$$P(\omega) = \langle |Y(\omega, \tau)|^2 \rangle = \mathbf{w}^H(\omega) \mathbf{R}_X(\omega) \mathbf{w}(\omega) \quad (1.14)$$

ここで， $\langle \cdot \rangle$ は，期待値演算を表す。また，空間相関行列 $\mathbf{R}_X(\omega)$ は，観測信号を用いて計算される [1, 2]。

$$\begin{aligned} \mathbf{R}_X(\omega) &= \langle \mathbf{x}(\omega, \tau) \mathbf{x}^H(\omega, \tau) \rangle \\ &= \mathbf{A}(\omega) \langle \mathbf{s}(\omega, \tau) \mathbf{s}^H(\omega, \tau) \rangle \mathbf{A}^H(\omega) + \langle \mathbf{n}(\omega, \tau) \mathbf{n}^H(\omega, \tau) \rangle \\ &= \mathbf{A}(\omega) \mathbf{R}_S(\omega) \mathbf{A}^H(\omega) + \mathbf{R}_N(\omega) \end{aligned} \quad (1.15)$$

式(1.15)では， $S_k(\omega, \tau)$ と $N_m(\omega, \tau)$ が統計的に独立であることを仮定した。もし， $K+1$ 個の源信号が互いに無相関で，平均的に同等のパワーを持つことを仮定すると， $K+1$ 個の源信号の分散共分散行列 $\mathbf{R}_S(\omega)$ は，次式でモデル化される。

$$\mathbf{R}_S(\omega) = \langle \mathbf{s}(\omega, \tau) \mathbf{s}^H(\omega, \tau) \rangle = \sigma_S^2(\omega) \mathbf{I}_{K+1} \quad (1.16)$$

ここで， $\sigma_S^2(\omega)$ は，源信号の平均的なパワーを表す。同様に，背景雑音の分散共分散行列 $\mathbf{R}_N(\omega)$ は，次式でモデル化される。

$$\mathbf{R}_N(\omega) = \langle \mathbf{n}(\omega, \tau) \mathbf{n}^H(\omega, \tau) \rangle = \sigma_N^2(\omega) \mathbf{I}_M \quad (1.17)$$

ここで， $\sigma_N^2(\omega)$ は，背景雑音の平均的なパワーを表す。以上の仮定より， $\mathbf{R}_X(\omega)$ は，次式でモデル化される。

$$\mathbf{R}_X(\omega) = \sigma_S^2(\omega) \mathbf{R}_A(\omega) + \sigma_N^2(\omega) \mathbf{I}_M \quad (1.18)$$

各マイクロホンで受信した信号パワーが同等であることを仮定すると，伝達特性でモデル化された空間相関行列 $\mathbf{R}_A(\omega)$ は次式で表される。

$$\mathbf{R}_A(\omega) = \mathbf{A}(\omega)\mathbf{A}^H(\omega) = \begin{bmatrix} \sigma_A^2(\omega) & \Gamma_{1,2}(\omega) & \cdots & \Gamma_{1,M}(\omega) \\ \Gamma_{2,1}(\omega) & \sigma_A^2(\omega) & \cdots & \Gamma_{2,M}(\omega) \\ \vdots & \vdots & \ddots & \vdots \\ \Gamma_{M,1}(\omega) & \Gamma_{M,2}(\omega) & \cdots & \sigma_A^2(\omega) \end{bmatrix} \quad (1.19)$$

ここで， $\sigma_A^2(\omega)$ ， $\Gamma_{i,j}(\omega)$ は，受信パワーとチャンネル間相関を表す。

$$\sigma_A^2(\omega) = \sum_{k=1}^K |A_{m,k}(\omega)|^2 \quad (1.20)$$

$$\Gamma_{i,j}(\omega) = \sum_{k=1}^K A_{k,i}(\omega)A_{k,j}^*(\omega) \quad (1.21)$$

* は複素共役を表す。式 (1.18) を式 (1.14) に代入すると， $P(\omega)$ は次式で書き下すことができる。

$$\begin{aligned} P(\omega) &= \mathbf{w}^H(\omega) (\sigma_S^2(\omega)\mathbf{R}_A(\omega) + \sigma_N^2(\omega)\mathbf{I}_M) \mathbf{w}(\omega) \\ &= \sigma_S^2(\omega)\mathbf{w}^H(\omega)\mathbf{R}_A(\omega)\mathbf{w}(\omega) + \sigma_N^2(\omega)\|\mathbf{w}(\omega)\|^2 \end{aligned} \quad (1.22)$$

信号強調の性能を高めるためには，式 (1.23) のようにターゲット音に対する感度を拘束したまま， $P(\omega)$ を最小化するように， $\mathbf{w}(\omega)$ を設計することが求められる。

$$\mathbf{w}^H(\omega)\mathbf{a}_0(\omega) = 1 \quad (1.23)$$

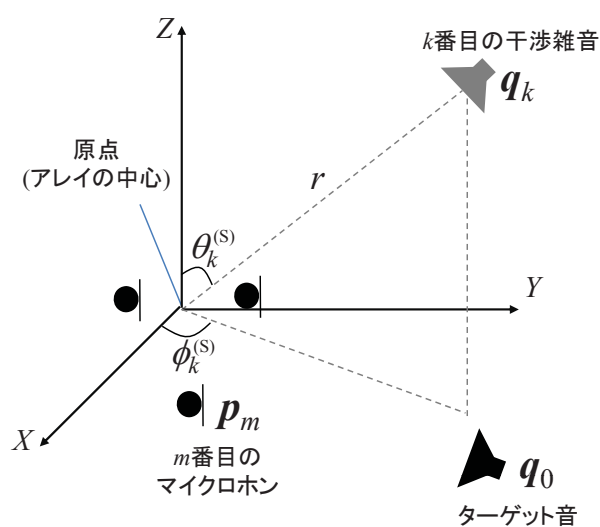


図 1.4: 座標系の定義

1.2.2 線形フィルタリング

本節では，1.2.1 節で説明した線形フィルタリングを用いて，ターゲット音を強調して收音するためのフィルタの設計方式について説明する。マイクロホン間に生じる位相/振幅の差を利用して，ターゲット音を強調したり（焦点形成），干渉雑音を抑圧する（死角形成）ことで，フィルタを設計する。代表的なフィルタ設計法として，(a) ターゲット音の到来方向に対して焦点を形成するための焦点形成法，(b) 干渉雑音の到来方向に対して死角を形成するための死角形成法，(c) 伝達特性そのものを用いて出力感度を制御する MINT (Multiple-input/output INverse Theorem) 法について概説する。

(a) 焦点形成法

ターゲット音の到来方向に対して焦点を形成するための焦点形成型のフィルタとして，遅延和法について説明する。源信号が受信位置に対して平面波伝搬することを仮定する。図 1.4 に示すように，アレイの中心を原点とした座標系を用いる。 m 番目のマイクロホンと k 番目の音源の位置ベクトルをそれぞれ p_m ， q_k とする。 k 番目の音源に

対する波数ベクトルを $\boldsymbol{\kappa}_k$ とすると，直接音の伝搬特性 (アレイ・マニフォールド・ベクトル) $\mathbf{h}_k(\omega)$ は，次式でモデル化される [1, 2]。

$$H_{k,m}(\omega) = \exp(-j\boldsymbol{\kappa}_k^T \mathbf{p}_m) \quad (1.24)$$

ここで，

$$\mathbf{h}_k(\omega) = [H_{k,1}(\omega), \dots, H_{k,M}(\omega)]^T \quad (1.25)$$

$$\mathbf{u}_k = [\sin \theta_k^{(S)} \cos \phi_k^{(S)}, \sin \theta_k^{(S)} \sin \phi_k^{(S)}, \cos \theta_k^{(S)}]^T \quad (1.26)$$

$$\kappa = \frac{\omega}{c} \quad (1.27)$$

$$\boldsymbol{\kappa}_k = -\kappa \mathbf{u}_k \quad (1.28)$$

$$\mathbf{q}_k = r_k^{(S)} \mathbf{u}_k \quad (1.29)$$

$$\mathbf{p}_m = r_m^{(R)} [\sin \theta_m^{(R)} \cos \phi_m^{(R)}, \sin \theta_m^{(R)} \sin \phi_m^{(R)}, \cos \theta_m^{(R)}]^T \quad (1.30)$$

c は音速を表す。 $r_k^{(S)}$, $\theta_k^{(S)}$, $\phi_k^{(S)}$ は，それぞれアレイの中心を原点とした k 番目の音源までの距離，水平角，仰角を表す。また， $r_m^{(R)}$, $\theta_m^{(R)}$, $\phi_m^{(R)}$ は，それぞれアレイの中心を原点とした m 番目のマイクロホンまでの距離，水平角，仰角を表す。

遅延和法 [1, 2] では，ターゲット音の到来方向に対する感度を強調する (焦点形成)。

$$\mathbf{w}^H(\omega) \mathbf{h}_0(\omega) = 1 \quad (1.31)$$

遅延和フィルタ $\mathbf{w}_{DS}(\omega)$ は，次式で計算される。

$$\mathbf{w}_{DS}(\omega) = \frac{\mathbf{h}_0(\omega)}{\mathbf{h}_0^H(\omega) \mathbf{h}_0(\omega)} \quad (1.32)$$

遅延和法では，アレイ・マニフォールド・ベクトルを用いて，各音源に対する感度を制御することを前提としている。つまり，マイクロホンの周りに反射・回折を誘発す

る物体を設置しない中空配置型アレイを用いて観測することを前提としている。また、干渉雑音の出力感度に対する制御は陽にしていなかったため、雑音抑圧性能が低いことが知られている。

(b) 死角形成法

雑音抑圧性能を高めるために、干渉雑音の到来方向に対して死角を形成する方法として最尤法 [11, 12, 13, 3] について説明する。ターゲット音とその他の雑音が互いに無相関であり、かつ観測信号が複素ガウス分布に従うことを仮定すると、ターゲット音 $S_0(\omega, \tau)$ を与えられたもとで、信号 $\mathbf{x}(\omega, \tau)$ を観測する確率密度関数 $p(\mathbf{x}(\omega, \tau)|S_0(\omega, \tau))$ は、次式で計算される。

$$p(\mathbf{x}(\omega, \tau)|S_0(\omega, \tau)) = \frac{1}{\det(\pi \mathbf{R}_{I,N}(\omega))} \cdot \exp\left[-(\mathbf{x}(\omega, \tau) - \mathbf{a}_0(\omega)S_0(\omega, \tau))^H \mathbf{R}_{I,N}^{-1}(\omega)(\mathbf{x}(\omega, \tau) - \mathbf{a}_0(\omega)S_0(\omega, \tau))\right] \quad (1.33)$$

雑音の空間相関行列 $\mathbf{R}_{I,N}(\omega)$ は、次式で定義される。

$$\begin{aligned} \mathbf{R}_{I,N}(\omega) &= \langle (\mathbf{x}(\omega, \tau) - \mathbf{a}_0(\omega)S_0(\omega, \tau))(\mathbf{x}(\omega, \tau) - \mathbf{a}_0(\omega)S_0(\omega, \tau))^H \rangle \\ &= \sigma_S^2(\omega) \mathbf{A}_I(\omega) \mathbf{A}_I^H(\omega) + \sigma_N^2(\omega) \mathbf{I}_M \end{aligned} \quad (1.34)$$

対数尤度関数は、次式で定義される。

$$\begin{aligned} LL(\mathbf{x}(\omega, \tau)|S_0(\omega, \tau)) &= -\log \det(\pi \mathbf{R}_{I,N}(\omega)) \\ &\quad - (\mathbf{x}(\omega, \tau) - \mathbf{a}_0(\omega)S_0(\omega, \tau))^H \mathbf{R}_{I,N}^{-1}(\omega)(\mathbf{x}(\omega, \tau) - \mathbf{a}_0(\omega)S_0(\omega, \tau)) \end{aligned} \quad (1.35)$$

対数尤度を最大とするためのフィルタを導出するために、尤度方程式を立てる。

$$\frac{\partial LL(\mathbf{x}(\omega, \tau)|S_0(\omega, \tau))}{\partial S_0^*(\omega, \tau)} = \mathbf{a}_0^H(\omega) \mathbf{R}_{I,N}^{-1}(\omega)(\mathbf{x}(\omega, \tau) - \mathbf{a}_0(\omega)S_0(\omega, \tau)) = 0 \quad (1.36)$$

最適解 $\mathbf{w}_{\text{ML}}(\omega)$ は、次式で表される。

$$\mathbf{w}_{\text{ML}}(\omega) = \frac{\mathbf{R}_{\text{I,N}}^{-1}(\omega)\mathbf{a}_0(\omega)}{\mathbf{a}_0^{\text{H}}(\omega)\mathbf{R}_{\text{I,N}}^{-1}(\omega)\mathbf{a}_0(\omega)} \quad (1.37)$$

なお、実用上の多くの場合には、中空配置型アレイで観測することを前提とし、式(1.37)に含まれる $\mathbf{a}_0(\omega)$ をアレイ・マニフォールド・ベクトル $\mathbf{h}_0(\omega)$ で置き換えて計算することが多い。

$$\mathbf{w}_{\text{NULL}}(\omega) = \frac{\mathbf{R}_{\text{I,N}}^{-1}(\omega)\mathbf{h}_0(\omega)}{\mathbf{h}_0^{\text{H}}(\omega)\mathbf{R}_{\text{I,N}}^{-1}(\omega)\mathbf{h}_0(\omega)} \quad (1.38)$$

これは、ターゲット音や干渉雑音の到来方向に対して、焦点や死角を形成することに対応する。死角形成法に基づくフィルタ設計は、ターゲット音の到来方向に関する感度を拘束したまま、雑音に対する感度を最小化するフィルタである。ただし、制御点の数 (M) よりも被制御点 (K) の数が多い場合 ($K > M$)、制御自由度が欠落することにより、信号出力パワーを小さくすることが困難となる。

(c) MINT(Multiple-input INverse Theorem) 法

焦点形成法や死角形成法では、ターゲット音や干渉雑音の到来方向に対して焦点や死角を形成することで、各音源に対する出力感度を制御してきた。本節で説明する MINT 法 [42] は、低残響場で観測することを想定し、伝達特性に対する感度を制御するという点において、焦点形成法や死角形成法とは異なる。伝達特性に対する感度を制御することは、アレイ・マニフォールド・ベクトルに含まれる直接音の伝搬特性のみならず、初期反射や後部残響を含めて制御することに対応する。故に、MINT 法は、残響抑圧技術として認知されている。本論文では、 $a_{0,m}(l)$ や $a_{k,m}(l)$ に対する出力感度を制御することを伝達特性を強調/抑圧するという表現を用いることとする。

インパルス応答長 L とフィルタ長 J の関係について議論するため、時間領域でフィルタを算出する。フィルタ \mathbf{w} は、 M チャネル分のフィルタが格納されているベクトル

を表す。

$$\mathbf{w} = [\mathbf{w}_1^T, \dots, \mathbf{w}_M^T]^T \quad (1.39)$$

$$\mathbf{w}_m = [w_m(0), \dots, w_m(J-1)]^T \quad (1.40)$$

ターゲット音の伝達特性を強調し，干渉雑音の伝達特性を抑圧するためのフィルタの条件を時間領域で表す。

$$\mathbf{a}\mathbf{w} = \mathbf{g} \quad (1.41)$$

ここで，

$$\mathbf{a} = \begin{bmatrix} \mathbf{a}_{0,1} & \cdots & \mathbf{a}_{0,M} \\ \vdots & \ddots & \vdots \\ \mathbf{a}_{K,1} & \cdots & \mathbf{a}_{K,M} \end{bmatrix} \quad (1.42)$$

$$\mathbf{a}_{k,m} = \left[\begin{array}{cccc} & \overbrace{\hspace{10em}}^J & & \\ a_{k,m}(0) & & & \mathbf{O} \\ a_{k,m}(1) & a_{k,m}(0) & & \\ \vdots & \vdots & \ddots & \\ a_{k,m}(L-1) & a_{k,m}(L-2) & \ddots & a_{k,m}(0) \\ & a_{k,m}(L-1) & \ddots & \vdots \\ & & \ddots & a_{k,m}(L-2) \\ \mathbf{O} & & & a_{k,m}(L-1) \end{array} \right] \Bigg\}^{J+L-1} \quad (1.43)$$

$$\mathbf{g} = [\mathbf{g}_0^T, \mathbf{g}_1^T, \dots, \mathbf{g}_K^T]^T \quad (1.44)$$

$$\mathbf{g}_0 = [\underbrace{0, \dots, 0}_{J-1}, 1, \underbrace{0, \dots, 0}_{L-1}]^T \quad (1.45)$$

$$\mathbf{g}_k = \underbrace{[0, \dots, 0]_{J+L-1}}^T \quad (1 \leq k \leq K) \quad (1.46)$$

である。

式(1.41)の逆問題を解くために、干渉雑音よりもマイクロホンが少ないこと ($K < M$) を仮定する。 J の条件で場合分けして、フィルタを導出する。

(i) $J = (L - 1)(K + 1)/(M - (K + 1))$ の場合

$J = (L - 1)(K + 1)/(M - (K + 1))$ を満たす場合、 \mathbf{a} は正方行列となる。伝達特性に共通な零点が存在しない場合、 \mathbf{w} は式(1.47)で算出される。

$$\mathbf{w}_{\text{MINT}} = \mathbf{a}^{-1} \mathbf{g} \quad (1.47)$$

(ii) $J > (L - 1)(K + 1)/(M - (K + 1))$ の場合

$J > (L - 1)(K + 1)/(M - (K + 1))$ を満たす場合、式(1.41)から \mathbf{w} を導出する問題は、過決定型の逆問題となる。伝達特性に共通な零点が存在しない時、 \mathbf{w} は式(1.48)で算出される。

$$\mathbf{w}_{\text{MINT}} = \mathbf{a}^+ \mathbf{g} = \mathbf{a}^T (\mathbf{a} \mathbf{a}^T)^{-1} \mathbf{g} \quad (1.48)$$

MINT法では、低残響場で観測することを想定し、音源とマイクロホン間の伝達特性が事前に既知であることを前提としている。干渉雑音数がマイクロホン数よりも少ない条件 ($K < M$) において、雑音抑圧性能が高くなることが示されている。

表 1.1: フィルタ設計に関する前提条件の比較

フィルタ設計方式	伝搬特性に関する既知条件				干渉雑音の数		観測場の仮定
	ターゲット音		干渉雑音		少ない	多い	
	直接音	伝達特性	直接音	伝達特性	$K \leq M$	$K > M$	
焦点形成法 (遅延和法)	✓	-	-	-	✓	-	自由場
死角形成法 (最尤法)	✓	-	-	✓	✓	-	自由場
MINT 法	-	✓	-	✓	✓	-	低残響場

本節で説明してきた焦点形成法 (遅延和法), 死角形成法 (最尤法), MINT 法の関係について, 表 1.1 に示す。3 手法において, 事前に既知とするパラメータの条件や観測場の仮定が異なるが, 多雑音下 ($K > M$) において源信号を推定することが困難であることは共通している。その理由として, 観測信号間の関係が, 音源の位置に応じて変化する割合が少なく, ターゲット音とその他の干渉雑音を分離するための空間的な手掛かりをほとんど得られないことがある。この課題は, 受信後の信号処理の高度化するというアプローチだけでは解決することが難しい。

1.2.3 受信方式

本節では, 受信方式を工夫し, 観測する信号そのものに含まれる空間的性質を変えるための知見や研究例について説明する。まず, マイクロホンを中空に設置する中空配置型アレイに関する知見について述べる。次に, 中空配置型アレイの一種である球対称アレイで発生する禁止周波数を取り除くことを目的とした剛球バッフル型アレイについて概説する。

(a) 中空配置型アレイ

反射，回折，拡散といった解析的にモデル化しづらい物理現象を生じさせないために，マイクロホンを中空に設置したアレイ（中空配置型アレイ）に関する知見について説明する。中空配置型アレイでは，主にマイクロホンの配置方法について検討されてきた。マイクロホンを一直線上に設置するラインアレイ，3次元上の任意の方向から到来する波源を特定するための四面体アレイ，球対称アレイ等がある [1, 2]。また，フィルタ設計方式が同一であったとしても，マイクロホン配置に応じて，雑音抑圧性能は変化する。

マイクロホン配置を決める際に，空間サンプリング定理 [1, 2] を考慮することが多い。空間サンプリング定理に従えば，空間周波数領域でエリアジングを生じさせないための最大マイクロホン間隔が与えられる。解析する信号に含まれる最大の周波数に対応する波長を λ_{Min} とする。空間サンプリング定理では， λ_{Min} の半波長以下でマイクロホンを設置することにより，空間エリアジングが生じないことを規定している。

$$\|p_i - p_j\| \leq \frac{\lambda_{\text{Min}}}{2} \quad (1.49)$$

ただし， p_i は i 番目のマイクロホンの位置ベクトルである。音速 c は約 340 m/s (摂氏 15 度) であるので，可聴帯域 (20 Hz ~ 20 kHz) に含まれる波長幅は，0.02 ~ 1 m 以上と広い。可聴帯域において空間エリアジングが生じないようにするためには，0.01 m 間隔でマイクロホンを配置することが求められる。しかし，マイクロホン間隔を狭めることは，音源情報の解析精度の向上に直接寄与しない。それは，狭間隔でマイクロホンを配置することにより，周波数帯域によっては，マイクロホン間に生じる位相/振幅差が小さくなり，音源情報を推定するための空間的な手掛かりが少なくなるためである。広帯域に渡って，音源情報の解析精度を高めるための最適な中空配置型アレイの形状は知られていない。

(b) 剛球バッフル型アレイ

半径 ζ の球面上に M 本のマイクロホンを幾何対称性を持たせて配置した球面对称アレイ (中空配置型アレイの一種) は, 球面調和関数に基づいて, 解析的にフィルタを設計するために用いられている。本節では, 球面对称アレイに対して生じる問題 (禁止周波数の発生) とそれを解決するための剛球バッフル型アレイに関して説明する [38, 39, 40, 41]。

式 (1.24) で定義した平面波の伝搬特性を球面調和関数展開すると, 次式で表される。

$$H_{k,m}(\omega) = \sum_{n=0}^{\infty} B_n(\zeta\kappa) \sum_{l=-n}^n S_n^{l*}(\theta_m^{(R)}, \phi_m^{(R)}) S_n^l(\theta_k^{(S)}, \phi_k^{(S)}) \quad (1.50)$$

ここで,

$$B_n(\zeta\kappa) = 4\pi j^n J_n(\zeta\kappa) \quad (1.51)$$

$$S_n^l(\theta, \phi) = \sqrt{\frac{(2n+1)(n-l)!}{4\pi(n+l)!}} P_n^l(\cos\theta) \exp(jl\phi) \quad (1.52)$$

$J_n(x)$, $P_n^l(x)$ は, それぞれ第一種球ベッセル関数, ルジャンドル陪関数を表す。球面調和関数群 $S_n^l(\theta, \phi)$ は互いに直交することから, k 番目の音源を強調するためのフィルタリングは次式で表される。

$$Y(\omega) = \sum_{n=0}^N B_n(\zeta\kappa) \sum_{l=-n}^n S_n^{l*}(\theta_m^{(R)}, \phi_m^{(R)}) W_{n,l}^*(\omega) \quad (1.53)$$

ここで, 球面調和展開されたフィルタ $W_{n,l}^*(\omega)$ は,

$$W_{n,l}^*(\omega) = \frac{S_n^l(\theta_k^{(S)}, \phi_k^{(S)})}{B_n(\zeta\kappa)} \quad (1.54)$$

である。また, N は球面調和関数の展開次数である。式 (1.54) の分母を構成する $B_n(\zeta\kappa)$ には, 第一種球ベッセル関数が含まれている。そのため, 分母が 0 になる周波数 (禁止

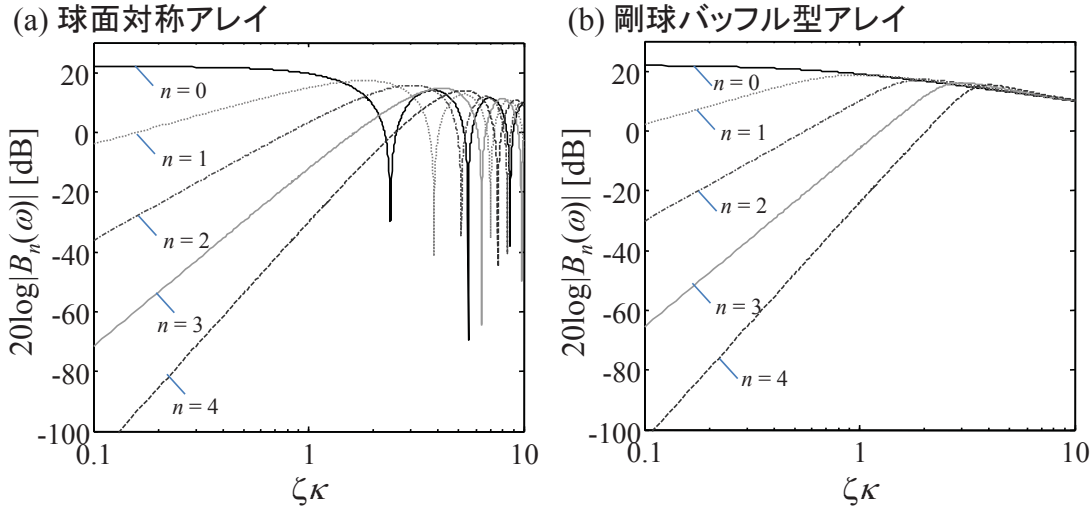


図 1.5: (a) 球対称アレイにおける $B_n(\zeta\kappa)$, (b) 剛球バッフル型アレイにおける $B_n(\zeta\kappa)$

周波数)が発生する(図 1.5(a)を参照)。禁止周波数では、フィルタ係数が無限大に発散してしまうため、安定的にフィルタを計算することが困難となる。

禁止周波数を生じさせないための一つの方法として、図 1.2 に示したように、剛球バッフル上にマイクロホンを設置する受音系(剛球バッフル型アレイ)を用いる方法がある。剛球バッフルを設置することにより生じる回折波の影響を考慮すると、 $B_n(\zeta\kappa)$ は、次式で表される。

$$B_n(\zeta\kappa) = 4\pi j^n \left(J_n(\zeta\kappa) - \frac{J_n'(\zeta\kappa)}{H_n'(\zeta\kappa)} H_n(\zeta\kappa) \right) \quad (1.55)$$

ここで、 $H_n(x)$ は、第二種球ハンケル関数である。剛球バッフルを設置することにより、 $B_n(\omega)$ が 0 まで低減することを回避することができる(図 1.5(b)を参照)。故に、球面調和展開されたフィルタ係数を安定的に計算することが可能になる。

1.3 本研究の考え方

1.3.1 着眼点と研究方針

1.2.2 節では、代表的な信号処理として、遅延和法、死角形成法、MINT 法に基づく線形フィルタリングについて説明してきた。一方、1.2.3 節では、代表的な受音系として、中空配置型アレイ、剛球バツフル型アレイについて説明してきた。従来のアレイ信号処理では、受音系と信号処理を組み合わせることで、音源情報を解析することを試みてきた。しかし、1.1 節で述べたように、多雑音下において源信号を推定することや音源距離を推定することは困難とされている。その理由は、音源の位置に応じて、観測信号間の関係が変化する割合が少ないためだと考えられる。

本研究では、剛球バツフル型アレイに関する研究から得た「アレイ構造を工夫することで、観測信号間の関係を理的に変える」という着想に基づいて、これまで推定することが困難とされてきた音源情報を得るために受音すべき信号の性質を導出し、その性質を満たすための受音方式を明らかにするための研究を行う。なお、受音方式に関する研究と信号処理に関する研究の独立性は高いため、双方の技術進展を組み合わせることで、音源情報の推定精度のさらなる向上が見込めるだろう。

音源情報を推定するために受音すべき信号の性質に関して、一つの着想を人間の聴覚特性から得た。聴覚の分野では、人間は、受聴信号に含まれる反射・拡散音の混在度合を手掛かりとして、音源までの距離を知覚していると考えられている [43]。四方を壁に囲まれた室内では、受聴位置から音源までの距離に応じて、伝達特性に含まれる直接音と反射・拡散音のパワー比（直間比）が変化する [44]。故に、アレイを用いて観測した受聴した信号から直間比を検知することができれば、人間と同様に、音源距離を推定できると考えられる。ここで着目したい考え方は、残響が付加された音場では、観測信号間の関係が変化する割合が大きくなり、これまで推定困難とされてきた音源距離の推定に繋がるということである。

表 1.2: 受音方式の比較

受音方式	音源の位置に対する 観測信号間の関係の変化	論理的解析の容易さ
中空配置型アレイ	変化が小さい	容易
剛球バッフル型アレイ	変化が小さい	容易
拡散受音	変化が大きい	困難

従来のアレイ信号処理では、信号処理で操作対象とするアレイ・マニフォールド・ベクトルに反射・拡散といった、モデル化することが困難な物理現象を含めてこなかった。言い換えると、従来技術では、反射・拡散して到来した音波は、音源情報を推定する上で、性能を低下させる要因として扱われてきた。一方、本研究では、反射・拡散して到来した音波に音源情報を解析するために有用な情報が含まれていると考えており、反射・拡散音を積極的に活用して音源情報を推定する。この点において、従来のアレイ信号処理とは大きくアプローチが異なる。本論文で提案する方法は、あらゆる方向から反射・拡散音が到来する拡散場の性質に着目した受音方式であるので、提案する受音方式を拡散受音と命名した。表 1.2 に、他の受音方式と拡散受音の違いを示す。中空配置型アレイや剛球バッフル型アレイといった従来の受音方式では、アレイ・マニフォールド・ベクトルのモデル化について論理的解析がしやすいことが特徴である。一方、提案する拡散受音では、多重に反射・拡散を繰り返した波源を観測するため、アレイ・マニフォールド・ベクトルを解析的に導出することは難しい。しかし、音源の位置に対する観測信号間の関係の変化が大きくなるような信号を受音するので、これまで推定することができなかった音源情報の解析（多雑音環境下における源信号推定、音源距離の推定）が可能になるだろう。

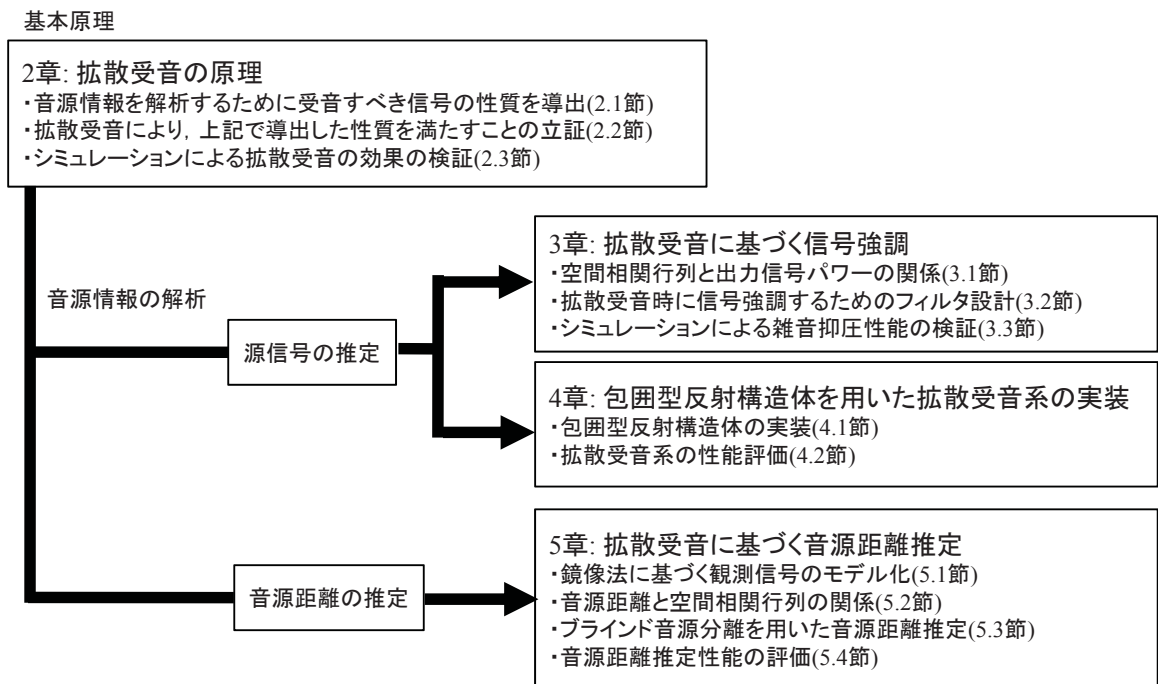


図 1.6: 本論文の構成

1.3.2 本論文の構成

本論文の構成を図 1.6 に示す。2 章では, 音源情報を解析するために受音すべき信号の性質について導出し, 提案方式(拡散受音)がその性質を満たす受音方式であることを示す。また, シミュレーションにより, 理想条件における拡散受音の効果を検証する。3 章では, 拡散受音を信号強調に適用した際の基本的な性質やフィルタ設計法について論ずる。4 章では, 拡散受音の一つの実装方式として, 包囲型反射構造体を用いたアレイについて説明する。5 章では, 残響室の中に設置したバツフル型アレイを用いて観測することを拡散受音として見立て, 音源距離の推定に適用する。6 章では, 本論文の結論を述べる。

第2章 拡散受音の原理

本章では、(i) 音源情報を解析するためにアレイで受音すべき信号の性質を明らかにし(2.1節)、(ii) (i)の性質を理論的に満たす受音方式の一つが拡散受音であることを示す(2.2節)。2.3節にて、数値シミュレーションによる拡散受音の効果を検証し、2.4節で本章の結論を述べる。

2.1 相互情報量最大化のための空間相関行列の最適構造

本節では、音源情報を解析するために、マイクロホンアレイで受音すべき信号の性質を導出する。移動通信の分野では、多入力多出力系(MIMO: Multiple Input Multiple Output)における情報伝送論が確立している[45, 46, 47]。MIMOの情報伝送論の一部を音響アレイ信号処理に取り入れて、音源情報を解析するためにアレイで受音すべき信号の性質を導く。

多音源 $s(\omega, \tau)$ から発生した波源が空気伝播されて、アレイを用いて多信号 $\mathbf{x}(\omega, \tau)$ を受音する MIMO 系を考える。音源とアレイ間の伝達特性は、MIMO では通信路特性に対応する。通信路を介して伝送される情報量(相互情報量) $I(s(\omega, \tau), \mathbf{x}(\omega, \tau))$ は、次式で与えられる。

$$I(s(\omega, \tau), \mathbf{x}(\omega, \tau)) = \iint p_{XS}(\mathbf{x}(\omega, \tau), s(\omega, \tau)) \log_2 \frac{p_{XS}(\mathbf{x}(\omega, \tau), s(\omega, \tau))}{p_X(\mathbf{x}(\omega, \tau))p_S(s(\omega, \tau))} dX dS \quad (2.1)$$

ここで, $p_{XS}(\mathbf{x}(\omega, \tau), \mathbf{s}(\omega, \tau))$ は, $\mathbf{x}(\omega, \tau)$ と $\mathbf{s}(\omega, \tau)$ の結合確率密度関数, $p_X(\mathbf{x}(\omega, \tau))$ と $p_S(\mathbf{s}(\omega, \tau))$ は, それぞれ $\mathbf{x}(\omega, \tau)$ と $\mathbf{s}(\omega, \tau)$ の確率密度関数である。相互情報量 $I(\mathbf{s}(\omega, \tau), \mathbf{x}(\omega, \tau))$ が大きいほど, 音源から伝送される情報量が増大し, 結果として音源情報を解析するために適した信号受音に繋がるという仮説を立てる。なお, 3.1 節では, 本節で導き出した結論を満たす信号受音により, 信号出力パワーが最小化し, 音源情報の一つである信号源の推定(信号強調)の性能向上に寄与することを示す。以後, $I(\mathbf{s}(\omega, \tau), \mathbf{x}(\omega, \tau))$ を最大化するための受信信号の性質を導出する。

$I(\mathbf{s}(\omega, \tau), \mathbf{x}(\omega, \tau))$ の最大値をチャンネル容量 $C(\omega)$ として定義する。

$$C(\omega) = \max\{I(\mathbf{s}(\omega, \tau), \mathbf{x}(\omega, \tau))\} \quad (2.2)$$

$C(\omega)$ を算出するために, 式 (2.1) を変形する。

$$\begin{aligned} I(\mathbf{s}(\omega, \tau), \mathbf{x}(\omega, \tau)) &= - \int p_X(\mathbf{x}(\omega, \tau)) \log_2 p_X(\mathbf{x}(\omega, \tau)) dX \\ &+ \iint p_S(\mathbf{s}(\omega, \tau)) p_{XS}(\mathbf{x}(\omega, \tau) | \mathbf{s}(\omega, \tau)) \log_2 p_{XS}(\mathbf{x}(\omega, \tau) | \mathbf{s}(\omega, \tau)) dS dX \end{aligned} \quad (2.3)$$

ここで, $p_{XS}(\mathbf{x}(\omega, \tau) | \mathbf{s}(\omega, \tau))$ は, $\mathbf{s}(\omega, \tau)$ が与えられた時の $\mathbf{x}(\omega, \tau)$ の条件付き確率密度関数を表す。背景雑音 $\mathbf{n}(\omega, \tau)$ が平均 0 の複素ガウス分布に従うという仮定から, $p_{XS}(\mathbf{x}(\omega, \tau) | \mathbf{s}(\omega, \tau))$ は次式で表される。

$$\begin{aligned} p_{XS}(\mathbf{x}(\omega, \tau) | \mathbf{s}(\omega, \tau)) &= \frac{1}{\pi^M \det(\mathbf{R}_N(\omega))} \\ &\cdot \exp \left[- (\mathbf{x}(\omega, \tau) - \mathbf{A}(\omega) \mathbf{s}(\omega, \tau))^H \mathbf{R}_N^{-1}(\omega) (\mathbf{x}(\omega, \tau) - \mathbf{A}(\omega) \mathbf{s}(\omega, \tau)) \right] \end{aligned} \quad (2.4)$$

したがって, 式 (2.3) の第二項目は, 次式で展開される。

$$\begin{aligned}
& \int p_S(\mathbf{s}(\omega, \tau)) dS \int p_{XS}(\mathbf{x}(\omega, \tau) | \mathbf{s}(\omega, \tau)) \log_2 p_{XS}(\mathbf{x}(\omega, \tau) | \mathbf{s}(\omega, \tau)) dX \\
&= \int p_S(\mathbf{s}(\omega, \tau)) dS [-\log_2 (\pi^M \det(\mathbf{R}_N(\omega))) \\
&\quad - \frac{1}{\log 2} \text{tr} \langle (\mathbf{x}(\omega, \tau) - \mathbf{A}(\omega) \mathbf{s}(\omega, \tau)) (\mathbf{x}(\omega, \tau) - \mathbf{A}(\omega) \mathbf{s}(\omega, \tau))^H \mathbf{R}_N^{-1}(\omega) \rangle] \\
&= -\log_2 (\pi^M \det(\mathbf{R}_N(\omega))) - \frac{M}{\log 2} \tag{2.5}
\end{aligned}$$

つまり、式 (2.3) の第二項目は、 $p_S(\mathbf{s}(\omega, \tau))$ に関係なく一定になる。一方、式 (2.3) の第一項目は、 $\mathbf{x}(\omega, \tau)$ のエントロピーである。多数の音源から発生した波源が混在して到来し、 M 個のマイクロホンで観測する時、 $\mathbf{x}(\omega, \tau)$ は平均 0 の複素ガウス分布に従うと想定できる。

$$p_X(\mathbf{x}(\omega, \tau)) = \frac{1}{\pi^M \det(\mathbf{R}_X(\omega))} \exp(-\mathbf{x}^H(\omega, \tau) \mathbf{R}_X^{-1}(\omega) \mathbf{x}(\omega, \tau)) \tag{2.6}$$

この時、式 (2.3) の第一項目は、次式で計算される。

$$-\int p_X(\mathbf{x}(\omega, \tau)) \log_2 p_X(\mathbf{x}(\omega, \tau)) = \log_2 (\pi^M \det(\mathbf{R}_X(\omega))) + \frac{M}{\log 2} \tag{2.7}$$

なお、観測信号が複素ガウス分布に従うという仮定は、 $\mathbf{x}(\omega, \tau)$ のエントロピーが最大であることに対応する（最大エントロピーの定理 [47]）。式 (2.5), (2.7) より、 $C(\omega)$ は次式で展開される。

$$\begin{aligned}
C(\omega) &= \log_2 \det(\mathbf{R}_N^{-1}(\omega) \mathbf{R}_X(\omega)) \\
&= \log_2 \det(\mathbf{R}_N^{-1}(\omega) \mathbf{A}(\omega) \mathbf{R}_S(\omega) \mathbf{A}^H(\omega) + \mathbf{I}_M) \\
&= \log_2 \det(\sigma_{SN}^2(\omega) \mathbf{R}_A(\omega) + \mathbf{I}_M) \tag{2.8}
\end{aligned}$$

ここで,

$$\sigma_{\text{SN}}^2(\omega) = \frac{\sigma_{\text{S}}^2(\omega)}{\sigma_{\text{N}}^2(\omega)} \quad (2.9)$$

である。なお, ここまでの MIMO における情報伝送論については, 文献 [48] を参照した。

式 (2.9) を構成する $\mathbf{R}_A(\omega)$ の構造を調査するために, 固有値分解する。

$$\mathbf{R}_A(\omega) = \mathbf{V}(\omega)\mathbf{\Lambda}(\omega)\mathbf{V}^H(\omega) \quad (2.10)$$

ここで, $\mathbf{V}(\omega)$ は固有フィルタ, $\mathbf{\Lambda}(\omega)$ は固有値で構成される。

$$\mathbf{V}(\omega) = [\mathbf{v}_1(\omega), \dots, \mathbf{v}_M(\omega)] \quad (2.11)$$

$$\mathbf{v}_m(\omega) = [V_{m,1}(\omega), \dots, V_{m,M}(\omega)]^T \quad (2.12)$$

$$\mathbf{\Lambda}(\omega) = \text{diag}\{\Lambda_1(\omega), \dots, \Lambda_M(\omega)\} \quad (2.13)$$

M 個の固有値は, 次式のように整列している。

$$\Lambda_1(\omega) \geq \dots \geq \Lambda_M(\omega) \geq 0 \quad (2.14)$$

また, $\mathbf{R}_A(\omega)$ のトレースは, 次式のように保存される。

$$\text{tr}(\mathbf{R}_A(\omega)) = \sum_{m=1}^M \Lambda_m(\omega) = \sum_{m=1}^M \sum_{k=1}^K |A_{m,k}(\omega)|^2 = M\sigma_A^2(\omega) \quad (2.15)$$

式 (2.14), (2.15) より, 最大固有値 $\Lambda_1(\omega)$ は, 以下のように制限される。

$$\sigma_A^2(\omega) \leq \Lambda_1(\omega) \leq M\sigma_A^2(\omega) \quad (2.16)$$

式 (2.10) を式 (2.8) に代入すると, $C(\omega)$ は次式で書き直せる。

$$\begin{aligned}
C(\omega) &= \log_2 \det (\mathbf{V}(\omega) (\sigma_{\text{SN}}^2(\omega) \boldsymbol{\Lambda}(\omega) + \mathbf{I}_M) \mathbf{V}^{\text{H}}(\omega)) \\
&= \log_2 \det (\sigma_{\text{SN}}^2(\omega) \boldsymbol{\Lambda}(\omega) + \mathbf{I}_M) \\
&= \log_2 \prod_{m=1}^M (\sigma_{\text{SN}}^2(\omega) \Lambda_m(\omega) + 1)
\end{aligned} \tag{2.17}$$

マイクロホン数 M や受音時の SN 比 $\sigma_{\text{SN}}^2(\omega) \sigma_{\text{A}}^2(\omega)$ を一定とした場合に, $C(\omega)$ を最大化するための $\mathbf{R}_{\text{A}}(\omega)$ の最適な構造条件を導く。相加相乗平均の関係 [49] より, 式 (2.17) を構成する項の最大値は, 次式で算出される。

$$\sqrt[M]{\prod_{m=1}^M (\sigma_{\text{SN}}^2(\omega) \Lambda_m(\omega) + 1)} \leq \frac{1}{M} \sum_{m=1}^M (\sigma_{\text{SN}}^2(\omega) \Lambda_m(\omega) + 1) = \sigma_{\text{SN}}^2(\omega) \sigma_{\text{A}}^2(\omega) + 1 \tag{2.18}$$

式 (2.18) の不等式において, $\mathbf{R}_{\text{A}}(\omega)$ の固有値が次式の条件を満たす時のみ等号条件が成立する。

$$\Lambda_1(\omega) = \dots = \Lambda_M(\omega) \tag{2.19}$$

式 (1.19) より, $\mathbf{R}_{\text{A}}(\omega)$ の対角項は定数であることを仮定している。式 (2.19) のように固有値分布が平滑化するためには, $\mathbf{R}_{\text{A}}(\omega)$ の非対角項を構成するチャネル間相関 $\Gamma_{i,j}(\omega)$ が 0 に低減するように受音することが求められる。

$$\lim_{\Gamma_{i,j} \rightarrow 0} \mathbf{R}_{\text{A}}(\omega) \rightarrow \begin{bmatrix} \sigma_{\text{A}}^2(\omega) & 0 & \cdots & 0 \\ 0 & \sigma_{\text{A}}^2(\omega) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{\text{A}}^2(\omega) \end{bmatrix} = \sigma_{\text{A}}^2(\omega) \mathbf{I}_M \tag{2.20}$$

式 (2.20) は，相互情報量を最大化するために，最適な空間相関行列の構造を表す。式 (2.20) における固有値は， $\Lambda_1(\omega) = \dots = \Lambda_M(\omega) = \sigma_A^2(\omega)$ となり，確かに式 (2.19) を満たす。また，式 (2.20) を満たす受音が達成されたとき， $C(\omega)$ は最大となる。

$$C_{\text{MAX}}(\omega) = M \log_2 (\sigma_{\text{SN}}^2(\omega) \sigma_A^2(\omega) + 1) \quad (2.21)$$

なお，MIMO の情報伝送論を音響アレイ信号処理に適用し，式 (2.10) 以降の式展開により，式 (2.20) で示した相互情報量を最大化するための受信信号の性質を明らかにしたことは，本研究成果に寄るものである。

式 (2.20) を満たす信号受信について，物理的な解釈を与える。式 (2.20) を満たす場合， $K = M - 1$ であると仮定すると， $R_A(\omega)$ を構成する $A(\omega)$ は正規直交行列になる。その時， $R_A(\omega)$ は，次式で展開される [50]。

$$R_A(\omega) = A(\omega) A^H(\omega) = A^H(\omega) A(\omega) = \sigma_A^2(\omega) I_M \quad (2.22)$$

式 (2.22) より，伝達特性 $\mathbf{a}_k(\omega)$ は互いに直交することが分かる。

$$\mathbf{a}_0(\omega) \perp, \dots, \perp \mathbf{a}_K(\omega) \quad (2.23)$$

各源信号と受信点群の間の伝達特性が互いに直交するほど，伝達特性が音源に対して固有なため，音源に関する多くの情報を取り出せると解釈できる。

2.2 拡散受音による相互情報量最大化

式 (2.20) を満たし，相互情報量を最大化するための受信方式として，拡散受音を提案する。拡散受音は，方向性のある音波を拡散させて受信する方法である。拡散受音

における $R_A(\omega)$ をモデル化するために，チャンネル間相関 $\Gamma_{i,j}(\omega)$ をモデル化する。各波源が，多重に反射・拡散して到来することを想定すると，受信位置では，等方位から均一なエネルギー密度で音波が伝搬することで表される [51, 52]。拡散受音における $\Gamma_{i,j}(\omega)$ の期待値は，次式で算出される。

$$\begin{aligned} \langle \Gamma_{i,j}(\omega) \rangle &= \frac{\sigma_A^2(\omega)}{4\pi} \int_{-\pi}^{\pi} \int_{-\pi/2}^{\pi/2} \exp(-j\boldsymbol{\kappa}^T(\mathbf{p}_i - \mathbf{p}_j)) d\theta d\phi \\ &= \sigma_A^2(\omega) \cdot \text{sinc}(\kappa \|\mathbf{p}_i - \mathbf{p}_j\|) \end{aligned} \quad (2.24)$$

ここで，

$$\kappa = \frac{\omega}{c} \quad (2.25)$$

$$\boldsymbol{\kappa} = -\kappa[\sin\theta \cos\phi, \cos\theta \cos\phi, \sin\phi]^T \quad (2.26)$$

である。音源やマイクロホンの数が十分に多いと仮定すると，式 (1.19) の非対角項を $\langle \Gamma_{i,j}(\omega) \rangle$ で置き換えて計算しても問題ないと考えられる。拡散受音における $R_A(\omega)$ は，次式でモデル化される。

$$\mathbf{R}_A(\omega) = \sigma_A^2(\omega) \begin{bmatrix} 1 & \text{sinc}(\kappa \|\mathbf{p}_1 - \mathbf{p}_2\|) & \cdots & \text{sinc}(\kappa \|\mathbf{p}_1 - \mathbf{p}_M\|) \\ \text{sinc}(\kappa \|\mathbf{p}_2 - \mathbf{p}_1\|) & 1 & \cdots & \text{sinc}(\kappa \|\mathbf{p}_2 - \mathbf{p}_M\|) \\ \vdots & \vdots & \ddots & \vdots \\ \text{sinc}(\kappa \|\mathbf{p}_M - \mathbf{p}_1\|) & \text{sinc}(\kappa \|\mathbf{p}_M - \mathbf{p}_2\|) & \cdots & 1 \end{bmatrix} \quad (2.27)$$

ここで，拡散受音の広義的な解釈として，マイクロホン間隔が十分に広いことを含める。マイクロホン間隔が十分広い時， $\langle \Gamma_{i,j}(\omega) \rangle$ は 0 に収束する。

$$\lim_{\|\mathbf{p}_i - \mathbf{p}_j\| \rightarrow \infty} \langle \Gamma_{i,j}(\omega) \rangle = \lim_{\|\mathbf{p}_i - \mathbf{p}_j\| \rightarrow \infty} \sigma_A^2(\omega) \cdot \text{sinc}(\kappa \|\mathbf{p}_i - \mathbf{p}_j\|) \rightarrow 0 \quad (2.28)$$

拡散受音における空間相関行列は，定数倍された単位行列に漸近する。

$$\lim_{\|\mathbf{p}_i - \mathbf{p}_j\| \rightarrow \infty} \mathbf{R}_A(\omega) = \sigma_A^2(\omega) \mathbf{I}_M \quad (2.29)$$

従って，拡散受音は，式(2.20)を満たし，相互情報量を最大化するための受音方式である。

方向性のある音波を拡散させて收音するためには，様々な実装方式があると考えられる。2.3節では，理想的な拡散受音を実現するために，残響室の中にアレイを設置することを拡散受音とし，シミュレーションにより効果を検証する。しかし，拡散受音を実用上の音源情報解析の問題に適用するためには，任意の環境下で，方向性のある音波を拡散させることが求められる。4.1節では，任意の環境で拡散受音を実現するための受音系として，包囲型反射構造体を用いたアレイの実装について触れる。また，拡散受音は，式(2.20)を満たすための十分条件であることに注意されたい。相互情報量を最大化し，音源情報を解析するために適した受音系の性質が，式(2.20)で表されることが認知されるようになれば，拡散受音以外のアプローチで，受音系構造を最適化する研究が，今後現れるのではないかと考えられる。

2.3 数値シミュレーションによる拡散受音の効果の検証

本節では，拡散受音によって，チャンネル間相関 $\Gamma_{i,j}(\omega)$ が低減し，空間相関行列の構造が式(2.20)に漸近するかどうかを検証する。

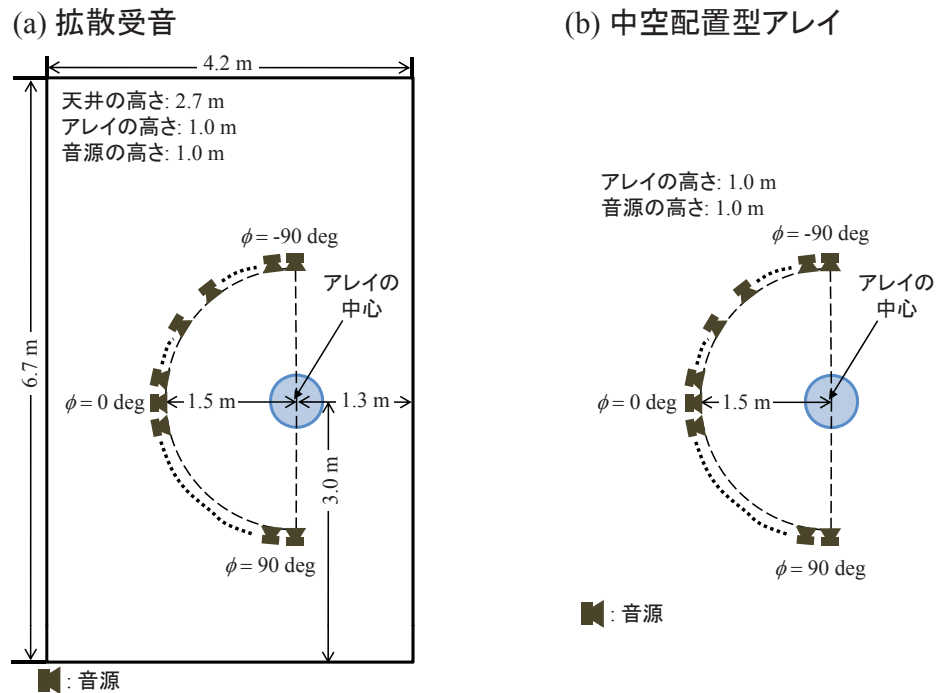


図 2.1: 拡散受音, 中空配置型アレイの設置条件

2.3.1 シミュレーション条件

本節では, 残響室の中にマイクロホンアレイを設置することを拡散受音とする。鏡像法 [53] に基づいて, 各音源からアレイまでのインパルス応答を算出した。図 2.1 (a) に, 拡散受音で用いた残響室の大きさ, 音源やアレイの配置条件を示す。壁, 床, 天井の反射係数は $\xi(\omega) = 0.85$ とし, サンプル周波数 48.0 kHz の場合にインパルス応答長は $L = 16384$ taps (341 ms) であった。直径 0.6 m の球上に M 個のマイクロホンを配置したアレイを用いることとし, マイクロホン数の異なる (A1) ステレオマイク ($M = 2$) ~ (A7) 切頂二十面体 ($M = 60$) の計 7 パターンを用意した。各アレイにおけるマイクロホンの位置や最小間隔に関して, 表 2.1 に示す。(A1) ステレオマイクの場合を除いて, 多面体の頂点にマイクロホンが配置されている。アレイの直径を一定としたため, マイクロホン本数 M に応じて, マイクロホンの最小間隔が変化することに注意されたい。アレイを囲むように, 181 個の音源を 1 度間隔で設置した。その

他のシミュレーション条件を表 2.2 に示す。

比較対象として、マイクロホンを中空に設置する中空配置型アレイを用いた。なお、1.2.3 節で説明した剛球バッフル型アレイは、比較対象としない。剛球バッフル型アレイは、球面調和関数展開に基づく解析的なフィルタ設計において、数値不安定性を緩和するための方式であり、チャンネル間相関を低減するために提案された方式ではない。本論文で扱う信号処理では、球面調和関数展開によるフィルタ設計を扱わないため、剛球バッフル型アレイを比較対象から外した。中空配置型アレイにおける音源やアレイの設置条件を図 2.1 (b) に示す。拡散受音と異なる条件は、残響室の有無のみである。鏡像法に基づいて算出したインパルス応答を周波数領域に変換し、受音系毎に空間相関行列 $R_A(\omega)$ を計算した。

2.3.2 シミュレーション結果

空間相関行列 $R_A(\omega)$ の構造を評価するための尺度として、条件数 $\Psi(\omega)$ を導入する。

$$\Psi(\omega) = 10 \log_{10} \frac{\Lambda_1(\omega)}{\Lambda_M(\omega)} \quad [\text{dB}] \quad (2.30)$$

チャンネル間相関が低減し、式 (2.20) に漸近するほど、最小固有値 $\Lambda_M(\omega)$ と最大固有値 $\Lambda_1(\omega)$ は同じような値をとるようになる。従って、 $\Psi(\omega)$ が小さくなるほど相互情報量が増大する。もし、式 (2.20) を満たす最適な信号受音が達成された時、 $\Psi(\omega)$ は最小化する ($\Psi(\omega) = 0$ [dB])。

図 2.2 に実験結果を示す。(A1) ステレオマイクや(A2) 正四面体を用いた場合、 $\Psi(\omega)$ に対する受音方式の違いは、ほとんど見られなかった。これは、受音系の構造を工夫しなくても、少ない数のマイクロホンで観測し得る最大の情報を受信しているためだと考えられる。一方(A3) 正六面体～(A7) 切頂二十面体を用いた場合、中空配置型アレイと比較して、拡散受音を用いることで、 $\Psi(\omega)$ が低減することを確認した。中空配置型アレイを用いた場合、マイクロホン数の増加に伴って、チャンネル間相関が高いマ

表 2.1: マイクロホンの設置条件

マイクロホンの設置条件	数	最小間隔
(A1) ステレオマイク (球の両端)	$M = 2$	0.60 m
(A2) 正四面体 (頂点)	$M = 4$	0.49 m
(A3) 正八面体 (頂点)	$M = 6$	0.42 m
(A4) 正六面体 (頂点)	$M = 8$	0.35 m
(A5) 正二十面体 (頂点)	$M = 12$	0.32 m
(A6) 正十二面体 (頂点)	$M = 20$	0.21 m
(A7) 切頂二十面体 (頂点)	$M = 60$	0.12 m

表 2.2: 鏡像法によるインパルス応答の計算条件

サンプリング周波数	48.0 kHz
解析周波数帯域	0.3–16.0 kHz
マイクロホン数, M	2, 4, 6, 8, 12, 20, 60 (7パターン)
アレイの直径	0.6 m
音源数, $K + 1$	181 (1度間隔)
インパルス応答長, L	16384 taps (341 ms)
フィルタ長, J	16384 taps (341 ms)
残響室の大きさ	4.2 m (W) × 6.7 m (D) × 2.7 m (H)
音源とアレイの中心の距離, $r_k^{(S)}$	1.5 m
壁, 床, 天井の反射係数, $\xi(\omega)$	0.85

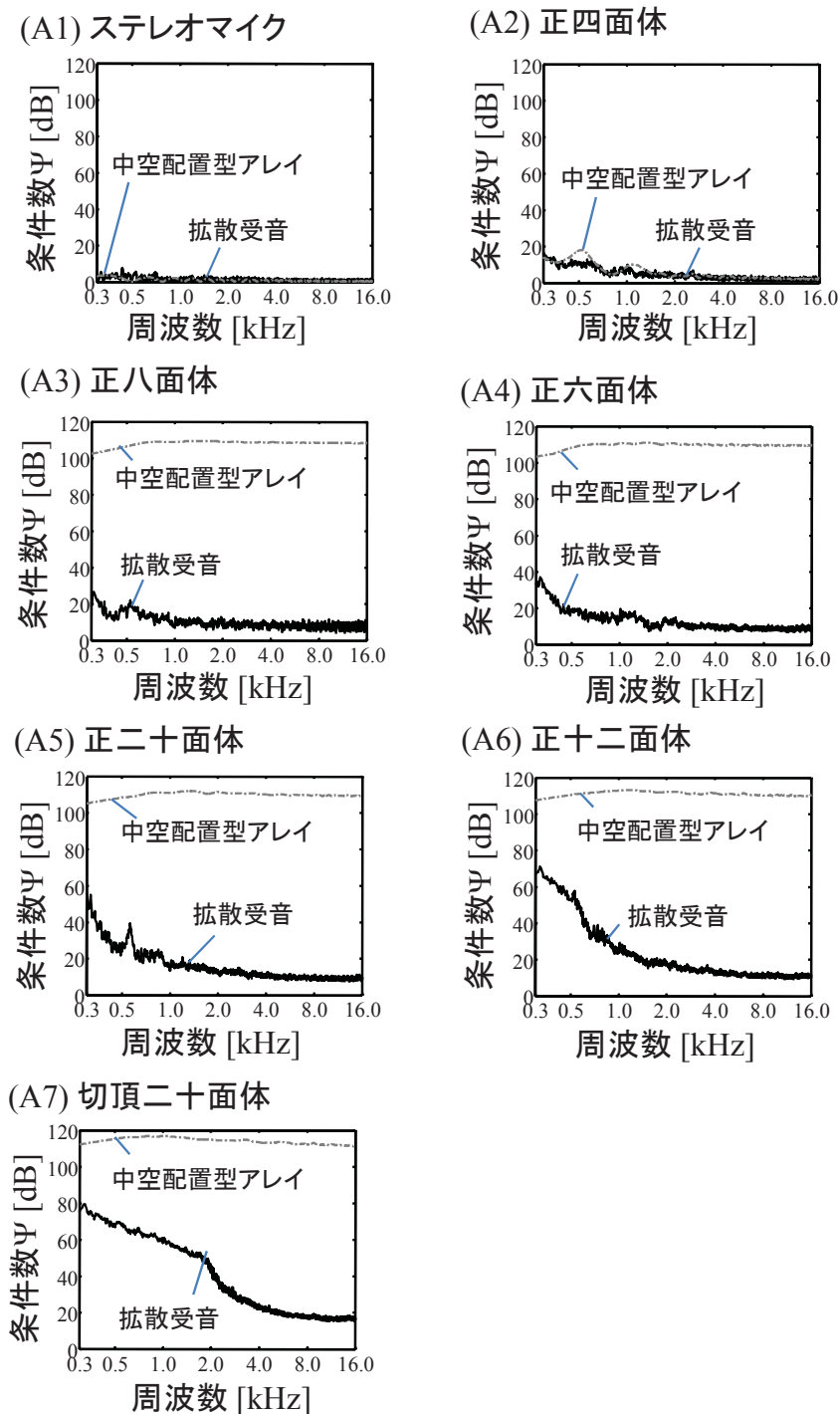


図 2.2: 受音方式と条件数の関係

マイクロホンペアが存在する可能性が高まる。その結果、空間相関行列 $R_A(\omega)$ の次元数が縮退し、 $\Psi(\omega)$ が増加してしまうと考えられる。拡散受音を用いた場合、マイクロホン数が増加しても、広帯域に渡って $\Psi(\omega)$ が低減した状態を保つことができた。2.1 節で説明と照らし合わせると、拡散受音により、各マイクロホンが異なる性質の信号を受音し、相互情報量が増加するような情報伝送を実現できていることを示唆している。(A6) 正二十面体や (A7) 切頂二十面体において、0.3~2.0 kHz の低周波帯域で $\Psi(\omega)$ が増大する傾向が見られた。これは、表 2.1 に示したように、 M が増加するに従ってマイクロホンの最小間隔が狭まるため、式 (2.27) を構成するチャンネル間相関が増加することに起因する。本節のシミュレーション条件では、 $\Psi(\omega)$ を最小化するまでには至らなかったが、拡散受音を用いることで、空間相関行列の構造が式 (2.20) に漸近し、相互情報量を増大することを確認した。

2.4 本章の結論

音源情報を解析するために、アレイで受音すべき信号の性質を導出した。マイクロホン数と受信時の SN 比が一定である時、チャンネル間相関が低減させるように受音することで、音源群と観測信号間の相互情報量が最大化されることを明らかにした。また、チャンネル間相関を低減させるように受信することは、伝達特性が音源に対して固有となり、音源に関する多くの情報を取り出せることに対応することを説明した。チャンネル間相関を低減させるための受音方法の 1 つとして、方向性のある音波を多重に反射・拡散させて収音する拡散受音を提案した。拡散受音では、等方位から均一なエネルギー密度で音波が到来するため、広帯域に渡ってチャンネル間相関が低減し、空間相関行列の固有値分布が平滑化することを示した。残響室にアレイを設置することで拡散受音を実装し、シミュレーションにより拡散受音の効果を検証した。実験結果から、拡散受音により、広帯域に渡って空間相関行列の条件数が低減し、チャンネル間相関を低減させる効果があることを確認した。以後の 3~5 章では、拡散受音により出力され

た多観測信号群に対して信号処理を適用することにより，信号強調や音源距離推定の問題に取り組む。

第3章 拡散受音に基づく信号強調

音源情報の一つである源信号を推定するために、2章で説明した拡散受音の多出力信号に対して信号強調法を適用する。なお、信号強調法は、線形フィルタリングに限定するものとし、拡散受音と組み合わせた場合の基本的な性質の調査やフィルタ設計に関して考察する。3.1節では、空間相関行列 $\mathbf{R}_A(\omega)$ と信号出力パワー $P(\omega)$ の関係について説明し、拡散受音の信号強調に対する効果について調査する。3.2節では、拡散受音に基づく信号強調におけるフィルタ設計に関して考察する。3.3節では、数値シミュレーションにより性能調査を行い、3.4節で本章の結論を述べる。

3.1 空間相関行列と信号出力パワーの関係

式(2.20)を満たすような拡散受音が実現されたと仮定した時の、信号出力パワー $P(\omega)$ の基本的な性質について述べる。なお、1.2.1節で説明したように、ターゲット音に対するゲインが拘束された条件下で $P(\omega)$ が小さくなるほど、雑音抑圧性能が高まり、源信号の推定精度が高まる。式(2.10)を式(1.22)に代入すると、 $P(\omega)$ は次式で書き表される。

$$\begin{aligned} P(\omega) &= \sigma_S^2(\omega) \mathbf{w}^H(\omega) \mathbf{R}_A(\omega) \mathbf{w}(\omega) + \sigma_N^2(\omega) \|\mathbf{w}(\omega)\|^2 \\ &= \sigma_S^2(\omega) \mathbf{w}^H(\omega) \mathbf{V}(\omega) \mathbf{\Lambda}(\omega) \mathbf{V}^H(\omega) \mathbf{w}(\omega) + \sigma_N^2(\omega) \|\mathbf{w}(\omega)\|^2 \end{aligned} \quad (3.1)$$

固有空間の性質より、 $\mathbf{w}(\omega)$ と $\mathbf{v}_1(\omega)$ が平行関係にあるとき、 $P(\omega)$ は最大となる。

$$\mathbf{w}(\omega) = \|\mathbf{w}(\omega)\| \mathbf{v}_1(\omega) \quad (3.2)$$

その時、 $P(\omega)$ の最大値は、次式で得られる。

$$\begin{aligned} P_{\text{MAX}}(\omega) &= \sigma_S^2(\omega) \|\mathbf{w}(\omega)\|^2 \mathbf{v}_1^H(\omega) \mathbf{V}(\omega) \mathbf{\Lambda}(\omega) \mathbf{V}^H(\omega) \mathbf{v}_1(\omega) + \sigma_N^2(\omega) \|\mathbf{w}(\omega)\|^2 \\ &= \|\mathbf{w}(\omega)\|^2 (\sigma_S^2(\omega) \Lambda_1(\omega) + \sigma_N^2(\omega)) \end{aligned} \quad (3.3)$$

$P_{\text{MAX}}(\omega)$ は、最大固有値 $\Lambda_1(\omega)$ に比例して増大することが分かる。

$\Lambda_1(\omega)$ の一般的な性質として、式 (2.16) のように制約される。拡散受音により、チャネル間相関が低減されて、固有値分布が平滑化することは、 $\Lambda_1(\omega)$ が最小化することを意味する。つまり、拡散受音には、信号出力パワーを最小化する効果がある。 $P_{\text{MAX}}(\omega)$ の最小値は次式で計算される。

$$\lim_{\Lambda_1 \rightarrow \sigma_A^2} P_{\text{MAX}}(\omega) \rightarrow \|\mathbf{w}(\omega)\|^2 (\sigma_S^2(\omega) \sigma_A^2(\omega) + \sigma_N^2(\omega)) \quad (3.4)$$

信号出力パワーが最小化することで、出力信号に含まれる雑音レベルが低減し、源信号の推定精度が向上する。つまり、拡散受音が音源情報を解析するために適した信号受音であることの裏付けとなる。

3.2 拡散受音におけるフィルタ設計

ターゲット音を強調し、信号出力パワーを最小化するためのフィルタ設計条件について考察する。拡散受音では、各音源から放射された音波が多重に反射・拡散してマイクロホンに伝搬する。故に、式 (3.5) のように、ターゲット音の伝達特性を強調し、 $P_N(\omega)$ を最小化するフィルタが、信号強調に適している。

$$\begin{cases} \min_{\mathbf{w}(\omega)} P(\omega) \\ \text{subject to } \mathbf{w}^H(\omega) \mathbf{a}_0(\omega) = 1 \end{cases} \quad (3.5)$$

式 (3.5) の条件を満たすフィルタは、次式で算出される。

$$\mathbf{w}(\omega) = \frac{\mathbf{R}_X^{-1}(\omega) \mathbf{a}_0(\omega)}{\mathbf{a}_0^H(\omega) \mathbf{R}_X^{-1}(\omega) \mathbf{a}_0(\omega)} \quad (3.6)$$

なお、1.2.2 節で説明した死角形成型のフィルタと構成が同じであるため、以後、式 (3.6) で計算されたフィルタを死角形成法として扱うこととする。

2.2 節において、拡散受信にはチャンネル間相関の抑圧効果があることを説明した。式 (2.29) を式 (1.18) に代入すると、空間相関行列は次式でモデル化される。

$$\begin{aligned} \mathbf{R}_X(\omega) &= \sigma_S^2(\omega) \sigma_A^2(\omega) \mathbf{I}_M + \sigma_N^2(\omega) \mathbf{I}_M \\ &= (\sigma_S^2(\omega) \sigma_A^2(\omega) + \sigma_N^2(\omega)) \mathbf{I}_M \end{aligned} \quad (3.7)$$

その時、式 (3.6) に含まれる $\mathbf{R}_X^{-1}(\omega)$ は、次式で計算される。

$$\mathbf{R}_X^{-1}(\omega) = \frac{1}{\sigma_S^2(\omega) \sigma_A^2(\omega) + \sigma_N^2(\omega)} \mathbf{I}_M \quad (3.8)$$

故に、拡散受信により、チャンネル間相関が十分に抑圧されていれば、式 (3.6) は、次式で近似して計算される。

$$\mathbf{w}(\omega) = \frac{\mathbf{a}_0(\omega)}{\mathbf{a}_0^H(\omega) \mathbf{a}_0(\omega)} \quad (3.9)$$

なお，ターゲット音の直接音の伝搬特性（アレイ・マニフォールド・ベクトル）ではなく，伝達特性に対して出力ゲインが拘束されているものの，1.2.2節で説明した焦点形成型のフィルタと構成が同じであるため，以後，式(3.9)で計算されるフィルタを焦点形成法として扱うこととする。焦点形成法では， K 個の干渉雑音に関する伝達特性が未知であったとしても，ターゲット音の伝達特性さえ既知であれば，自然と雑音が抑圧されて，信号強調可能であることを示唆している。これは，式(2.23)で表したように，拡散受信により，伝達特性が音源に対して固有になるためと考えられる。

焦点形成法に基づいてフィルタを設計する場合，フィルタ長 J をどの程度の長さに設定すべきかについて考察する。 J とインパルス応答長 L の関係を調査するために，時間領域でフィルタを設計する。式(3.9)と同様に，ターゲット音のインパルス応答を強調するための条件式は，式(3.10)で表される。

$$\mathbf{a}_0 \mathbf{w} = \mathbf{g}_0 \quad (3.10)$$

ここで，

$$\mathbf{a}_0 = [\mathbf{a}_{0,1}, \dots, \mathbf{a}_{0,M}] \quad (3.11)$$

$$\mathbf{a}_{0,m} = \left[\begin{array}{cccc} \overbrace{\hspace{10em}}^J & & & \\ a_{0,m}(0) & & & \mathbf{O} \\ a_{0,m}(1) & a_{0,m}(0) & & \\ \vdots & \vdots & \ddots & \\ a_{0,m}(L-1) & a_{0,m}(L-2) & \ddots & a_{0,m}(0) \\ & a_{0,m}(L-1) & \ddots & \vdots \\ & & \ddots & a_{0,m}(L-2) \\ \mathbf{O} & & & a_{0,m}(L-1) \end{array} \right] \quad (3.12)$$

$$\mathbf{w} = [\mathbf{w}_1^T, \dots, \mathbf{w}_M^T]^T \quad (3.13)$$

$$\mathbf{w}_m = [w_m(0), \dots, w_m(J-1)]^T \quad (3.14)$$

$$\mathbf{g}_0 = [\underbrace{0, \dots, 0}_{J-1}, 1, \underbrace{0, \dots, 0}_{L-1}]^T \quad (3.15)$$

である。1.2.2節で説明したMINT法を参考にすると、ターゲット音の応答が無歪で出力されるための条件は、伝達特性に共通零点が存在しない、かつ $J \geq (L-1)/(M-1)$ を満たすことで達成される。 \mathbf{a}_0 が正方行列となる場合 ($J = (L-1)/(M-1)$) とその他の場合 ($J > (L-1)/(M-1)$) に分けて、フィルタを算出する。

(i) $J = (L-1)/(M-1)$ の場合、フィルタは式 (3.16) で算出される。

$$\mathbf{w} = \mathbf{a}_0^{-1} \mathbf{g}_0 \quad (3.16)$$

(ii) $J > (L-1)/(M-1)$ の場合、フィルタは式 (3.17) で算出される。

$$\mathbf{w} = \mathbf{a}_0^+ \mathbf{g}_0 = \mathbf{a}_0^T (\mathbf{a}_0 \mathbf{a}_0^T)^{-1} \mathbf{g}_0 \quad (3.17)$$

k 番目の干渉雑音に対する応答 ε_k を評価することで、雑音信号の出力パワーが小さくなるためのフィルタ長の条件を導く。

$$\varepsilon_k = \mathbf{a}_k \mathbf{w} = \mathbf{a}_k \mathbf{a}_0^T (\mathbf{a}_0 \mathbf{a}_0^T)^{-1} \mathbf{g}_0 \quad (3.18)$$

ここで、

$$\varepsilon_k = [\varepsilon_k(0), \dots, \varepsilon_k(J+L-2)]^T \quad (3.19)$$

$$\mathbf{a}_k = [\mathbf{a}_{k,1}, \dots, \mathbf{a}_{k,M}] \quad (3.20)$$

$$\mathbf{a}_{k,m} = \left[\begin{array}{cccc} \overbrace{a_{k,m}(0)}^J & & & \mathbf{O} \\ a_{k,m}(1) & a_{k,m}(0) & & \\ \vdots & \vdots & \ddots & \\ a_{k,m}(L-1) & a_{k,m}(L-2) & \ddots & a_{k,m}(0) \\ & a_{k,m}(L-1) & \ddots & \vdots \\ & & \ddots & a_{k,m}(L-2) \\ \mathbf{O} & & & a_{k,m}(L-1) \end{array} \right] \quad (3.21)$$

である。式 (3.18) を構成する \mathbf{a}_0 の分散共分散行列を展開する。

$$\mathbf{a}_0 \mathbf{a}_0^T = \begin{bmatrix} \sigma_0 & \gamma_{0,1} & \cdots & \gamma_{0,J+L-2} \\ \gamma_{1,0} & \sigma_1 & \cdots & \gamma_{1,J+L-2} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{J+L-2,0} & \gamma_{J+L-2,1} & \cdots & \sigma_{J+L-2} \end{bmatrix} \quad (3.22)$$

ここで,

$$\sigma_i = \sum_{m=1}^M \sum_{l=0}^i a_{0,m}^2(i-l) \quad (3.23)$$

$$\gamma_{i,j} = \sum_{m=1}^M \sum_{l=0}^{\min(i,j)} a_{0,m}(i-l)a_{0,m}(j-l) \quad (3.24)$$

であり, $i > L-1$ の時, $a_{0,m}(i) = 0$ とする。拡散受信では, 多重に反射・拡散した音波を受信するので, インパルス応答 $a_{0,m}(l)$ はランダムで長い時系列となることが想定される [54]。その時, $a_{0,m}(l)$ の自己相関関数は鋭いピークを持つと考えられる。また,

チャンネル間相関が広帯域に渡って小さくなるので， $\gamma_{i,j}$ もまた十分小さくなると仮定する。その時，式 (3.22) は，式 (3.25) によって近似計算される。

$$\mathbf{a}_0 \mathbf{a}_0^T = \begin{bmatrix} \sigma_0 & & & \mathbf{O} \\ & \sigma_1 & & \\ & & \ddots & \\ \mathbf{O} & & & \sigma_{J+L-2} \end{bmatrix} \quad (3.25)$$

式 (3.25) を式 (3.18) に代入することで， ε_k は次式で計算される。

$$\varepsilon_k = \mathbf{a}_k \mathbf{a}_0^T \left[\overbrace{0, \dots, 0}^{J-1}, 1/\sigma_{J-1}, \overbrace{0, \dots, 0}^{L-1} \right]^T \quad (3.26)$$

ε_k の i 番目の振幅値は，次式で計算される。

$$\varepsilon_k(i) = \frac{\sum_{m=1}^M \sum_{l=0}^{\min(i, J-1)} a_{k,m}(i-l) a_{0,m}(J-l-1)}{\sum_{m=1}^M \sum_{l=0}^{J-1} (a_{0,m}(J-l-1))^2} \quad (3.27)$$

ただし， $i > L-1$ の時， $a_{k,m}(i) = 0$ である。拡散受信の効果として，各源信号に対する伝達特性の直交性が高まることを 2.1 節で説明した。伝達特性間の相関が低減することで， $a_{0,m}(l)$ と $a_{k,m}(l)$ の内積が小さくなることが期待されるので，式 (3.27) の分子は自ずと小さくなると考えられる。一方，式 (3.27) の分母は， $J \geq L$ とすることで最大となる。つまり，雑音の信号出力パワーを低減するためのフィルタ長 J の条件は，インパルス応答長 L 以上の長さに設定することが求められる。

以上より，拡散受信におけるフィルタ設計では，(i) ターゲット音の伝達特性 $a_0(\omega)$ を強調するようにフィルタを設計することと，(ii) フィルタ長 J をインパルス応答長 L 以上の長さに設定することが求められる。焦点形成法，死角形成法とも，ターゲット

音の伝達特性に強調するようにフィルタを設計するため，上記の条件を満たす。もし，拡散受音によって，チャンネル間相関が十分に抑圧されていない場合には，信号処理によってチャンネル間相関を抑圧する効果が含まれている死角形成法のほうが，焦点形成法と比較して雑音抑圧性能が高まると考えられる。

3.3 数値シミュレーションによる拡散受音に基づく信号強調効果の検証

本節では，拡散受音に基づく信号強調に関して，2種類の実験を行う。1つ目の実験は，拡散受音により出力された多観測信号に対して，信号強調法（焦点形成法，死角形成法）を適用した場合の雑音抑圧性能，及び指向特性に関して調査を行う。なお，本シミュレーションでは，2.3節と同様に，残響室の中にマイクロホンアレイを設置することを拡散受音とする。2つ目の実験では，焦点形成法を用いて，フィルタ長と雑音抑圧性能の関係について調査を行う。

3.3.1 シミュレーション 1: 拡散受音における雑音抑圧性能の調査

[シミュレーション条件]

中空配置型アレイと信号強調法（焦点形成法，死角形成法）を組み合わせたアレイ信号処理を比較対象とし，拡散受音と信号強調法（焦点形成法，死角形成法）を組み合わせたアレイ信号処理の雑音抑圧性能を調査する。なお，性能調査にあたって，2.3.1節のシミュレーションで算出したインパルス応答群を用いることとする。2.3.2節のシミュレーション結果では，マイクロホン数が多くなるほど，条件数に対する中空配置型アレイと拡散受音の差異が大きかった。本節では，(A5) 正二十面体 ($M = 12$)，(A6) 正十二面体 ($M = 20$)，(A7) 切長二十面体 ($M = 60$) の計3種類のマイクロホン配

表 3.1: アレイ信号処理の条件比較

アレイ信号処理方式	伝搬特性の既知条件		観測場の性質	
	ターゲット音	干渉雑音	自由音場	拡散音場
(S1) 中空配置型アレイ+焦点形成法	✓	-	✓	-
(S2) 中空配置型アレイ+死角形成法	✓	✓	✓	-
(S3) 拡散受信+焦点形成法	✓	-	-	✓
(S4) 拡散受信+死角形成法	✓	✓	-	✓

表 3.2: シミュレーション条件 (その1)

サンプリング周波数	48.0 kHz
解析周波数帯域	0.3 ~ 16 kHz
マイクロホン数, M	12, 20, 60 (3種類)
アレイの直径	0.6 m
ターゲット音の到来方向, $\phi_0^{(S)}$	-45, 0度 (2種類)
音源数, $K + 1$	181 (1度間隔)
インパルス応答長, L	16384 taps (341 ms)
フィルタ長, J	16384 taps (341 ms)
残響室の大きさ	4.2 m(W) × 6.7 m(D) × 2.7 m(H)
音源とアレイ間の距離, $r_k^{(S)}$	1.5 m
壁, 床, 天井の反射係数, $\xi(\omega)$	0.85

置に限定して実験を進める。2種類の受信系（中空配置型アレイ，拡散受信）に対して，それぞれ2種類の信号処理を適用し，(S1)～(S4)の4種類のアレイ信号処理を用意する。4種類のアレイ信号処理における条件比較を表3.1に示す。(S1)中空配置型アレイ＋焦点形成法，(S2)中空配置型アレイ＋死角形成法の2種類については，従来のアレイ信号処理方式である。(S2)の信号処理には出力信号パワーを最小化するためのコスト関数がフィルタ設計に含まれているので，(S1)よりも信号強調性能は高くなることが想定される。一方，(S3)拡散受信＋焦点形成法，(S4)拡散受信＋死角形成法の2種類については，拡散受信に基づく信号強調法（提案方式）である。(S3)では，理論上，チャンネル間相関を低減させるための信号受信が実現されている。また，(S4)では，受信系の工夫に加えて，出力信号パワーを最小化するための信号処理が含まれている。つまり，(S4)を用いることで，雑音抑圧性能が最も高くなると予想される。4種類のアレイ信号処理において，ターゲット音を $\phi_0^{(S)} = -45, 0$ 度に配置した時の雑音抑圧性能や指向特性を算出した。他のシミュレーション条件について，表3.2に示す。

[シミュレーション結果]

雑音抑圧性能を評価するための尺度として，信号出力パワーの期待値 $\bar{P}(\omega)$ を用いる。

$$\bar{P}(\omega) = 10 \log_{10} (\mathbf{w}^H(\omega) \mathbf{R}_A(\omega) \mathbf{w}(\omega)) \quad [\text{dB}] \quad (3.28)$$

式(3.28)では，背景雑音に対して源信号のパワーが十分大きい（受信時のSN比が高い）ことと，アレイを取り囲むように配置した干渉雑音源の出力パワーが均一であることを仮定している。(S1)～(S4)までのいずれのアレイ信号とも，ターゲット音に対する出力ゲインが拘束されているので， $\bar{P}(\omega)$ が小さいほど，雑音抑圧性能が高いと言える。 $\phi_0^{(S)} = -45, 0$ 度にあるターゲット音を強調した時の信号出力パワーをそれぞれ図3.1, 3.2に示す。(A5)正二十面体($M = 12$)，(A6)正十二面体($M = 20$)では， $\bar{P}(\omega)$ を低減するために，(S2)と(S4)の2種類のアレイ信号処理が有効であった。一方，(A7)切頂二十面体($M = 60$)では， $\bar{P}(\omega)$ を低減するために，(S2)よりも(S4)のほうが有効

であった。なお，(S4) を用いた場合，信号出力パワーは， $\bar{P}(\omega) = 10 \log_{10}(K/M)$ dB 程度に低減した。以上より，拡散受信と死角形成法を組み合わせることで，高い雑音抑圧性能を得ることができ，特にマイクロホン数が多くなるほど，その傾向が顕著であることを確認した。

各音源の位置に対するフィルタの感度 $\bar{G}_k(\omega)$ を算出し，指向特性を示す。

$$\bar{G}_k(\omega) = 20 \log_{10} |\mathbf{w}^H(\omega) \mathbf{a}_k(\omega)| \quad [\text{dB}] \quad (3.29)$$

(S1) ~ (S4) までのいずれのアレイ信号とも，ターゲット音に対する出力感度は， $\bar{G}_0(\omega) = 0$ dB で拘束されている。(S1) ~ (S4) に対する指向特性を図 3.3 ~ 3.8 に示す。マイクロホン数やターゲット音の到来方向に対して，図番号を分けた。(S1) を用いた場合，2.0 kHz 以下の低周波帯域ではメインローブの幅が広がり，4.0 kHz 以上の高周波帯域では空間エリアジングが生じる傾向が見られた。そのため，周波数帯域によっては，干渉雑音に対する感度が高まり，雑音抑圧性能が低下してしまう。(S2) を用いた場合，(S1) と比較して，メインローブの幅が狭くなる傾向が見られた。しかし，低周波帯域でメインローブの幅が広がり，高周波帯域で空間エリアジングが生じる傾向は変わらない。これは，ターゲット音とその他の干渉雑音を区別するための情報が，受信した信号から欠落しているためだと考えられる。(S3) では，チャンネル間相関を低減する拡散受信の効果により，(S1) と比較してメインローブの幅が狭くなり，高周波帯域で目立った空間エリアジングが生じなかった。しかし，2.3.1 節のシミュレーションで示したように，チャンネル間相関が完全に 0 まで低減していないため，干渉雑音に対する出力感度が高い帯域も存在した。そのため，(S2) と (S3) の雑音抑圧性能に対する優劣はつけ難い。一方，(S4) では，拡散受信の多出力信号群に対して，出力信号パワーを低減するような信号強調処理がされている。そのため，(S3) と比較して，干渉雑音に対する感度を広帯域に渡って低減することができた。以上より，(i) チャンネル間相関を低減するように受信系を設計すること (e.g. 拡散受信) と，(ii) ターゲット音の伝達特性を強調

表 3.3: シミュレーション条件 (その2)

サンプリング周波数	12.0 kHz
解析周波数帯域	0.3 ~ 5.5 kHz
マイクロホン数, M	12, 20, 60 (3種類)
アレイの直径	0.6 m
ターゲット音の到来方向, $\phi_0^{(S)}$	-45, 0度 (2種類)
音源数, $K + 1$	181 (1度間隔)
インパルス応答長, L	4096 taps (341 ms)
フィルタ長, J	512, 1024, 2048, 4096, 8192 taps (5種類)
残響室の大きさ	4.2 m(W) × 6.7 m(D) × 2.7 m(H)
音源とアレイ間の距離, $r_k^{(S)}$	1.5 m
壁, 床, 天井の反射係数, $\xi(\omega)$	0.85

し, 死角形成法を組み合わせることで, 雑音抑圧性能を大幅に向上できることを確認した。

3.3.2 シミュレーション2: フィルタ長と雑音抑圧性能の関係の調査

[シミュレーション条件]

(S3) のアレイ信号処理を用いて, 拡散受信時におけるフィルタ長と雑音抑圧性能の関係について調査を行う。なお, 性能調査にあたって, 2.3.1 節のシミュレーションで算出したインパルス応答群を 12 kHz までダウンサンプリングして実験を行う。これは, 長さの異なる焦点形成フィルタを生成するために, 式 (3.17) を用いて, 時間領域でフィルタを設計するためである。実験では, 5 種類のフィルタ長 $J = 512, 1024, 2048, 4096 (= L), 8192$

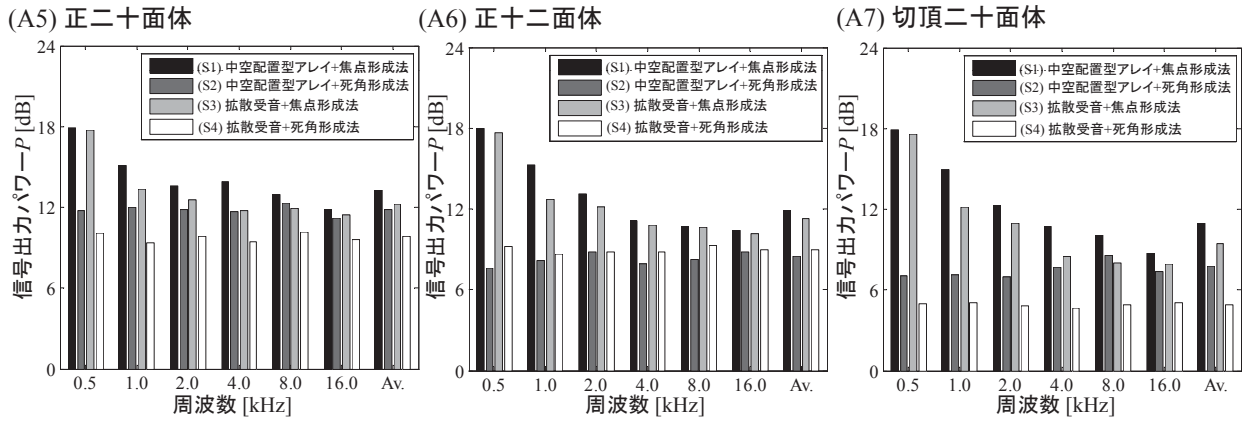


図 3.1: アレイ信号処理と信号出力パワーの関係 ($\phi_0^{(S)} = -45$ [deg])

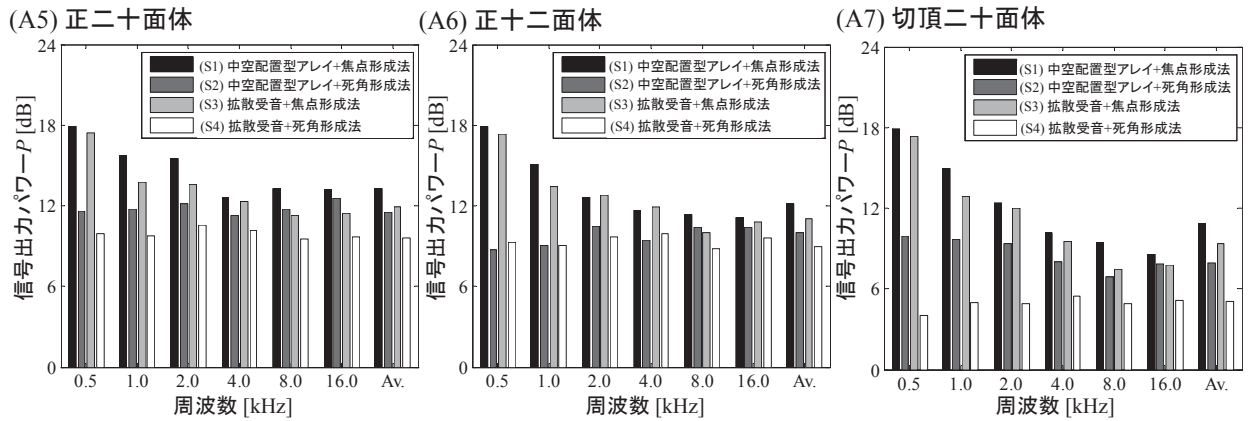


図 3.2: アレイ信号処理と信号出力パワーの関係 ($\phi_0^{(S)} = 0$ [deg])

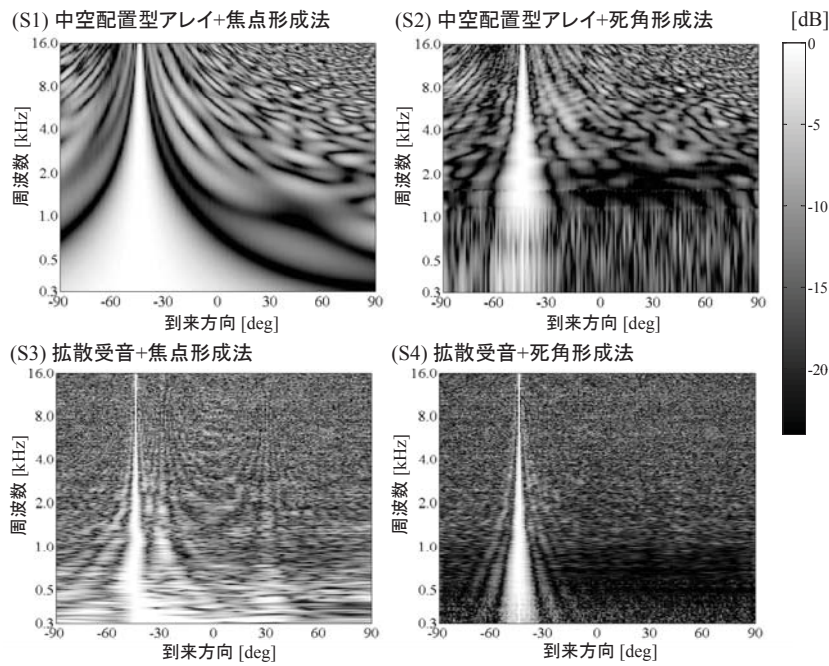


図 3.3: アレイ信号処理と指向特性の関係 ((A5) 正二十面体 ($M = 12$), $\phi_0^{(S)} = -45$ 度)

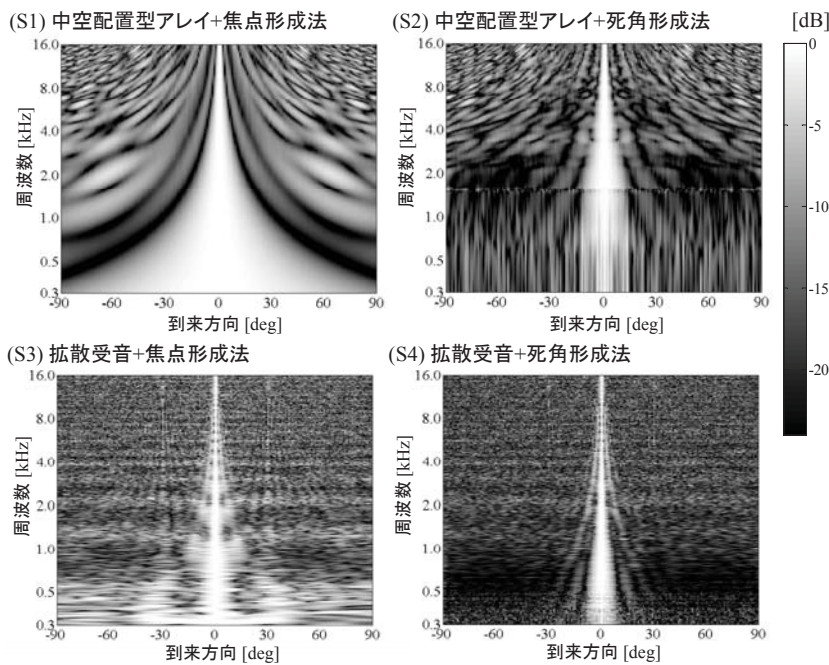


図 3.4: アレイ信号処理と指向特性の関係 ((A5) 正二十面体 ($M = 12$), $\phi_0^{(S)} = 0$ 度)

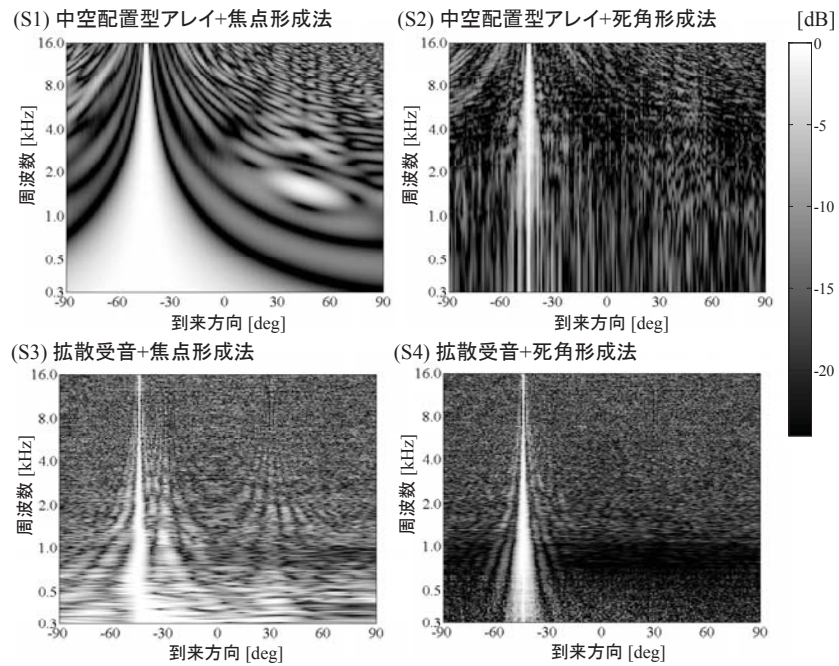


図 3.5: アレイ信号処理と指向特性の関係 ((A6) 正十二面体 ($M = 20$), $\phi_0^{(S)} = -45$ 度)

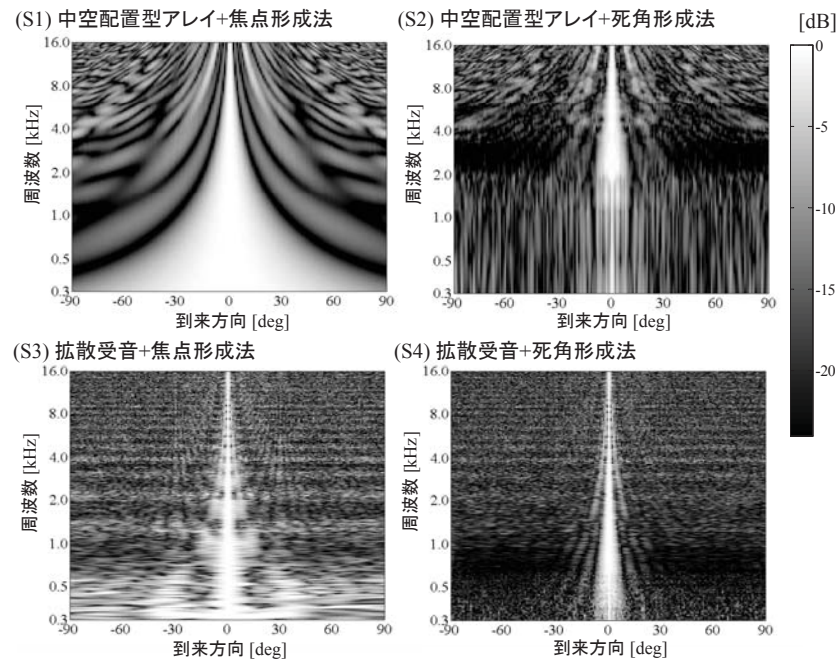


図 3.6: アレイ信号処理と指向特性の関係 ((A6) 正十二面体型配置 ($M = 20$), $\phi_0^{(S)} = 0$ 度)

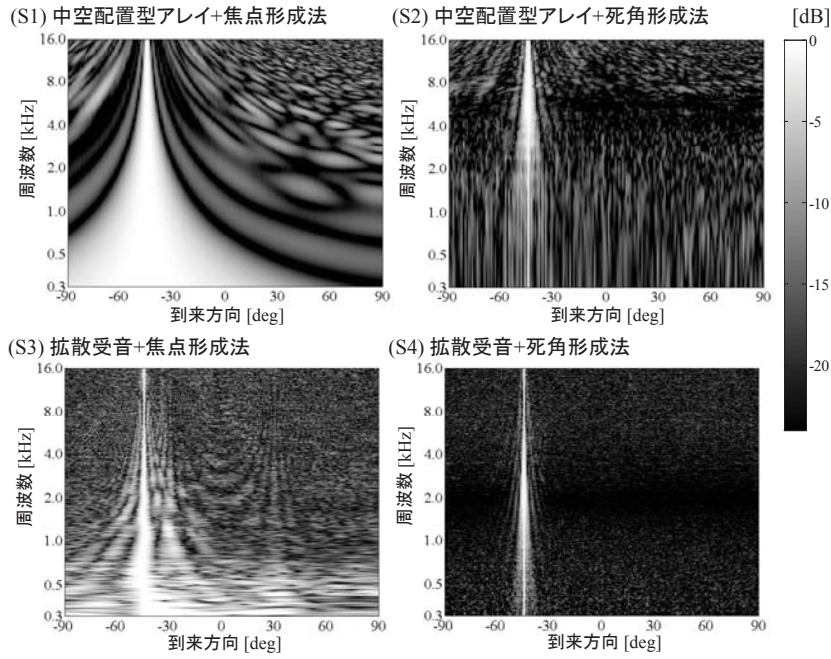


図 3.7: アレイ信号処理と指向特性の関係 ((A7) 切頂二十面体 ($M = 60$), $\phi_0^{(S)} = -45$ 度)

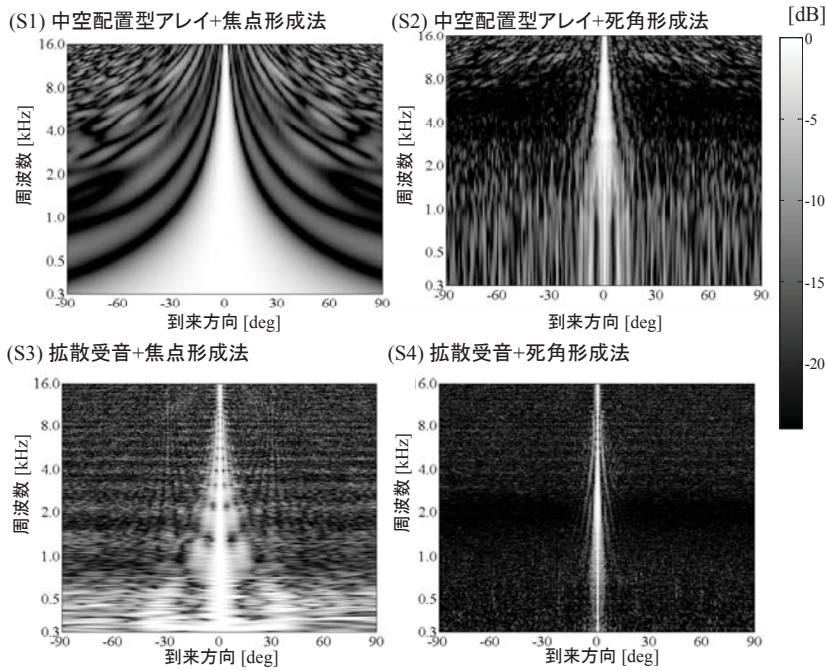


図 3.8: アレイ信号処理と指向特性の関係 ((A7) 切頂二十面体 ($M = 60$), $\phi_0^{(S)} = 0$ 度)

を用いた。なお、いずれの受信系、フィルタ長ともに、 $J > (L-1)/(M-1)$ の条件を満たす。その他のシミュレーション条件を表 3.3 に示す。

[シミュレーション結果]

雑音抑圧性能を評価するための尺度として、式 (3.28) で定義した信号出力パワーの期待値 $\bar{P}(\omega)$ を用いた。(S4) を用いて、 $\phi_0^{(S)} = -45, 0$ 度にあるターゲット音を強調した時の信号出力パワーをそれぞれ図 3.9, 3.10 に示す。マイクロホン数、ターゲット音の到来方向、周波数帯域に関わらず、フィルタ長を長くするほど $\bar{P}(\omega)$ が低減する傾向を確認した。これにより、拡散受信時には、 $J \geq L$ となるように、フィルタ長を設定すべきであることが明らかになった。

3.4 本章の結論

拡散受信の出力信号群に対して、信号強調するための方式に関して調査した。拡散受信の出力信号群には、信号出力パワーを低減するための性質が自然に含まれることを示した。この事実は、2.1 節で説明した相互情報量を最大化するための受信により、音源情報を解析しやすい信号受信が実現されるという仮説と一致する。数値シミュレーションにより、(i) チャンネル間相関を低減するための拡散受信と、(ii) ターゲット音とアレイ間の伝達特性（反射・拡散を含む）を強調しつつ、信号出力パワーを最小化する死角形成法を組み合わせることで、広帯域に渡って雑音抑圧性能の高いアレイ信号処理を実現できることを確認した。

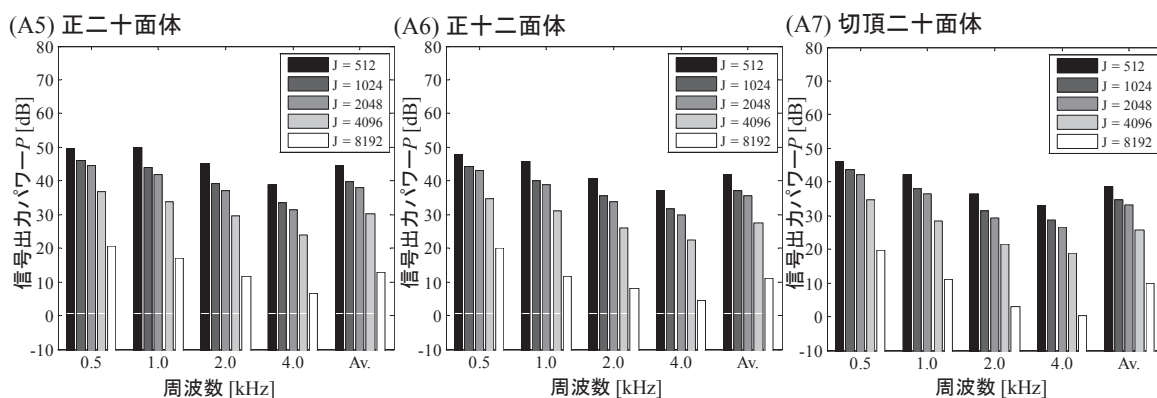


図 3.9: 信号出力パワーとフィルタ長の関係 ($\phi_0^{(S)} = -45$ 度)

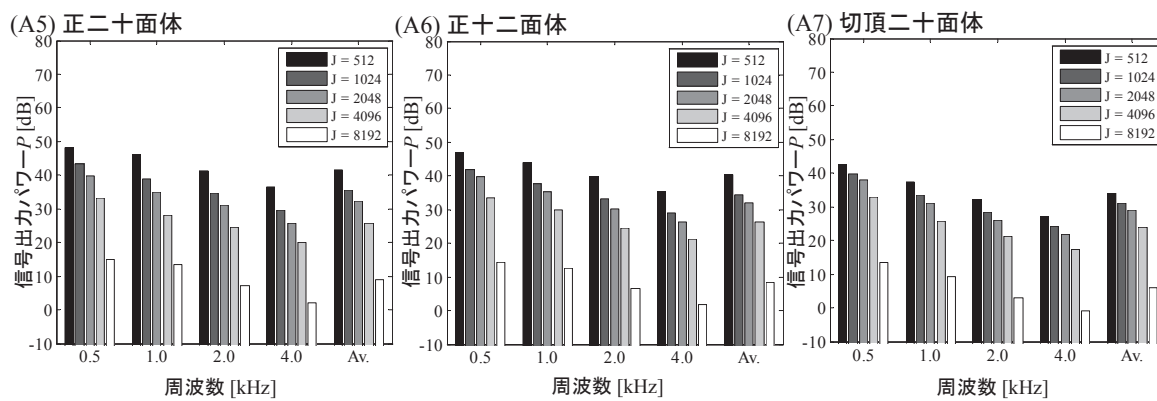


図 3.10: 信号出力パワーとフィルタ長の関係 ($\phi_0^{(S)} = 0$ 度)

第4章 包囲型反射構造体を用いた拡散 受音の実装

3章では、残響室の中にアレイを設置することを拡散受音とし、焦点形成法や死角形成法に従ってフィルタ設計することで信号強調できることを示した。しかし、焦点形成法や死角形成法では、音源とアレイ間の伝達特性を用いてフィルタを設計するため、環境に応じて伝達特性を測定したり、音場シミュレーションによって伝達特性を事前に用意する必要がある。本章では、任意の環境で、拡散受音の効果を利用した音源情報解析を実現するために、包囲型反射構造体を用いた拡散受音の実装に関して説明する。4.1節では、包囲型反射構造体を用いたアレイの構造について説明する。4.2節では、包囲型反射構造体を用いたアレイと死角形成法を用いた場合の雑音抑圧性能について調査する。4.3節で本章の結論を述べる。

4.1 包囲型反射構造体アレイ

残響室の中にアレイを設置して観測することと同様に、任意の環境において、反射・拡散して到来した音を受音するための手段として、包囲型反射構造体アレイを提案する。包囲型反射構造体アレイは、多マイクロホンとアレイを囲うような形状の反射構造体で構成される。反射構造体は、音を多重に反射させるために、硬度の高い素材で構成することが望ましい。実験に用いた受音系では、8 mm 厚の ABS 樹脂 (Acrylonitrile Butadiene Styrene) の板を組み合わせて、包囲型反射構造体を組み立てた。

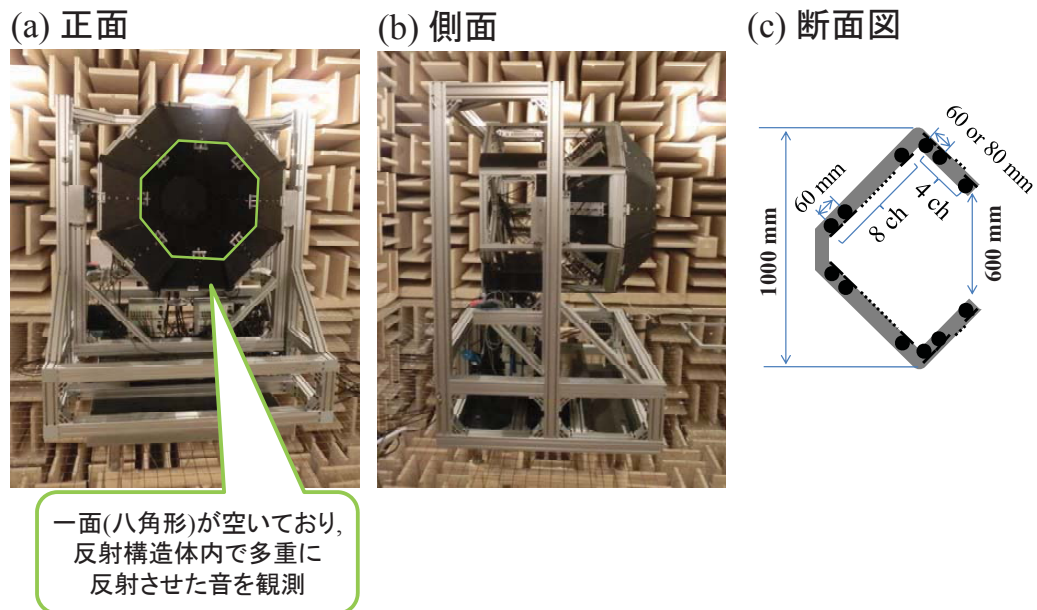


図 4.1: 包囲型反射構造体アレイ

構築した包囲型反射構造体アレイを図 4.1 に示す。構築したアレイの大きさは、直径 1 m の球程度の大きさである。正面の一面(八角形)は取り外されていて、そこから入射した音波が、包囲型反射構造体内で多重に反射し、マイクロホンに到達する。包囲型反射構造体アレイの効果について、図 4.2 を用いて説明する。鏡像法 [53] に基づいて、反射音をモデル化する。各反射音は、音源やマイクロホンの位置、包囲型反射構造体の形状によって決まる鏡像の位置に、仮想的にマイクロホン(鏡像マイクロホン)を配置して收音することでモデル化される。多重に反射し拡散することで、図 4.2(b) のように鏡像マイクロホンが様々な位置に展開されることになる。最終的に受信系から出力される信号数は固定値 (M) であるが、仮想的に M 以上のマイクロホンを様々な間隔で並べ、巨大なアレイを用いて観測したかのような効果を得ることができる。この効果により、チャンネル間相関が広帯域に渡って低減し、音源情報を推定するために適した信号受音に繋がると考えられる。

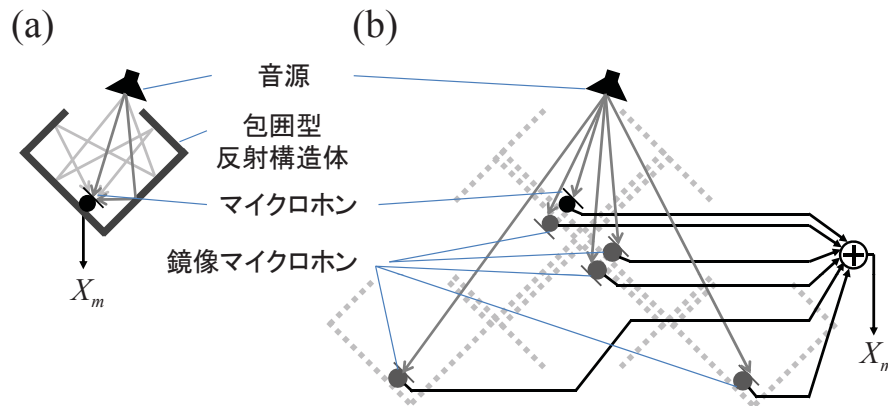


図 4.2: 鏡像法に基づく反射構造体アレイの等価システム

4.2 包囲型反射構造アレイの性能調査

本節では、包囲型反射構造体アレイを用いて実装した拡散受音と中空配置型アレイを比較し (i) 条件数を用いた受音系そのものの性能評価、及び (ii) 包囲型反射構造体アレイと死角形成法を組み合わせた時の雑音抑圧性能や方向/距離感度特性について調査する。

4.2.1 実験条件

4.1 節で説明した包囲型反射構造体アレイを無響室に設置した (図 4.3)。包囲型反射構造体アレイの正面の領域 (5.75 m(W)×4.5 m(D)×1.25 m(H)) に 0.25 m 間隔でスピーカを配置し、計 2,448 点の伝達特性を測定した。計測した全ての伝達特性を用いて $R_A(\omega)$ を計算し、ターゲット音を設置した 1 点を強調するフィルタを死角形成法で算出した。これは、3.3 節のシミュレーションにおける (S4) 拡散受音 (包囲型反射構造体アレイ) + 死角形成法に対応する。図 4.4 に示す 16 箇所 (P1 ~ P16) にターゲット音を設置し、雑音抑圧性能や指向性を算出した。

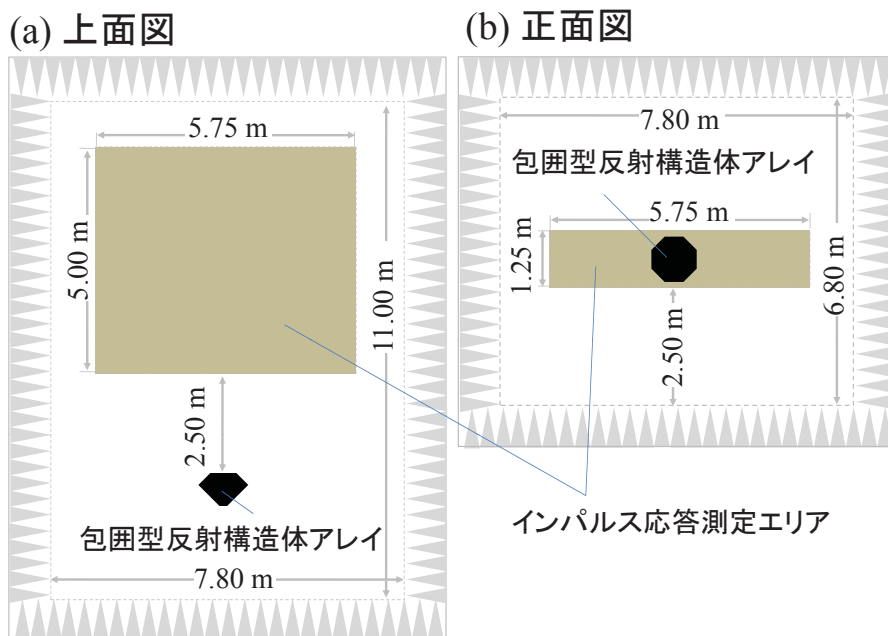


図 4.3: 無響室内の包囲型反射構造体アレイの設置位置とインパルス応答の測定領域

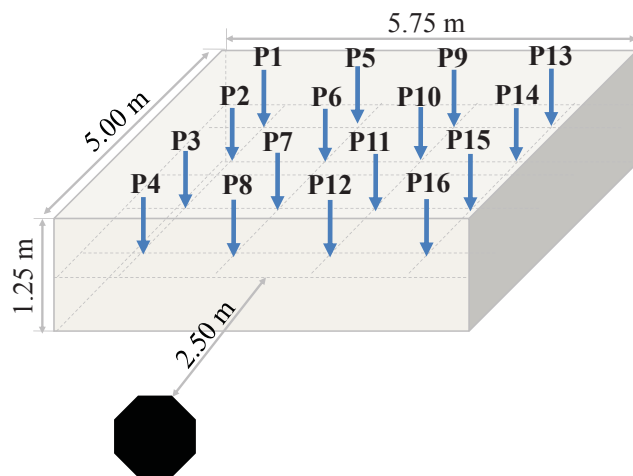


図 4.4: ターゲット音の設置位置 (16 箇所)

表 4.1: 実験条件

サンプリング周波数	48.0 kHz
解析周波数帯域	0.3 ~ 16.0 kHz
マイクロホン数, M	96
音源数 (伝達特性の測定点数), $K + 1$	2,448
インパルス応答長, L	16384 点 (341 ms)
フィルタ長, J	16384 点 (341 ms)
ターゲット音の設置位置	16 箇所

また, 比較対象として, 包囲型反射構造体アレイと同数, 同位置にマイクロホンを配置した中空配置型アレイの伝播特性をシミュレーションで算出し, 死角形成法でフィルタを生成した。これは, 3.3 節のシミュレーションにおける (S2) 中空配置型アレイ + 死角形成法に対応する。その他の実験条件を表 4.1 に示す。

4.2.2 実験結果

2 種類の受信系 (包囲型反射構造体アレイ, 中空配置型アレイ) と条件数の関係を図 4.5 に示す。包囲型反射構造体をアレイの近傍に設置することで, 広帯域に渡って, 条件数が低減することを確認した。この傾向は, 2.3 節のシミュレーション結果と同様であった。従って, 包囲型反射構造体アレイは, 拡散受信の実装方式の一つと言える。

雑音抑圧性能を調査するために, (P1) ~ (P16) までの 16 パターンで信号出力パワー $\bar{P}(\omega)$ を計算し, その平均値を図 4.6 に示した。(S4) 包囲型反射構造体アレイ + 死角形成法を用いることで, (S2) 中空配置型アレイ + 死角形成法と比較して, 信号出力パワーが低減することを確認した。また, 式 (3.29) のように, フィルタと伝達特性を掛け合わせることで, 方向/距離感度特性 $\bar{G}_k(\omega)$ を計算した。図 4.7 ~ 4.38 に, (S2) 中空配置

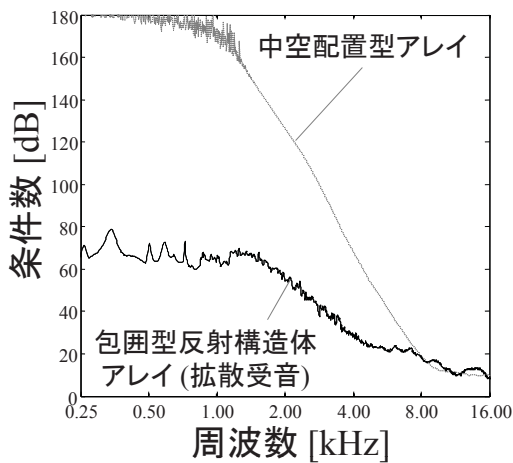


図 4.5: 受音方式と条件数の関係

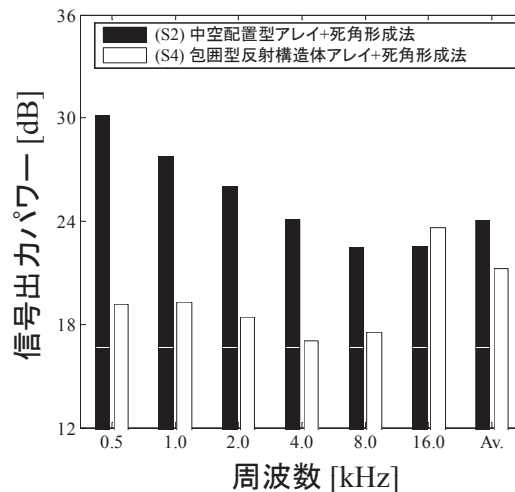


図 4.6: アレイ信号処理と信号出力パワーの関係

型アレイ + 死角形成フィルタと (S4) 包囲型反射構造体アレイ + 死角形成フィルタを用いた場合の方向/距離感度特性を示す。なお、図中の“ ”で囲われた位置にターゲット音を設置した。(S2)では低周波帯域ではメインローブの幅が広がり、高域では空間エリアジングが生じてしまう傾向が見られた。よって、任意の位置にあるターゲット音を強調して收音することは難しい。一方、(S4)では、包囲型反射構造体アレイを用いることで、広帯域に渡って、ターゲット音に対して高い感度形成(指向制御)が可能であった。ターゲット音の到来方向に対して指向性を形成し、距離に対してはターゲット音の前後である程度低減するような感度を形成することを確認した。

4.3 本章の結論

包囲型反射構造体と多マイクロホンを用いた拡散受音の実装について説明した。反射構造体内で多重に反射した音を收音することを鏡像法でモデル化すると、仮想的に巨大なアレイを用いて観測することと同等の効果を得ることを示した。包囲型反射構造体アレイを無響室に設置して、信号強調に対する性能を調査するための実験を行なっ

た。包囲型反射構造体で観測した信号に対して、死角形成法で設計したフィルタを適用することで、ターゲット音を強調し、その他の方向から到来する干渉雑音の出力パワーを抑圧するように感度を形成可能であることを確認した。

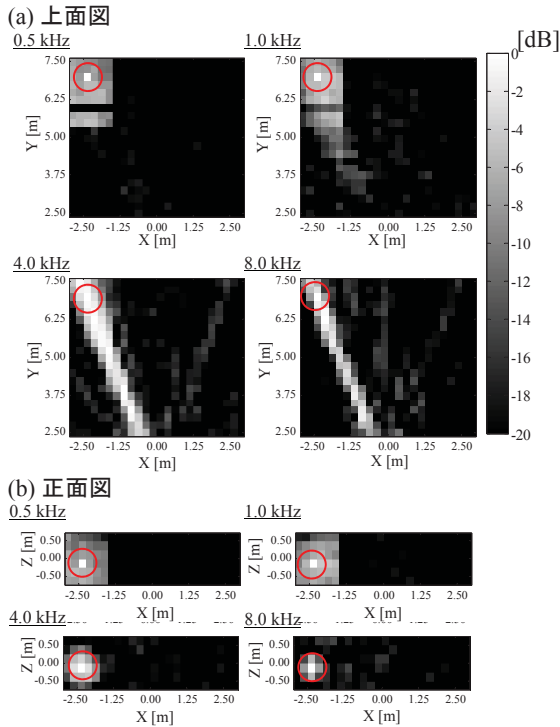


図 4.7: P1 を強調した場合の方向/距離感度特性 ((S4) 包囲型反射構造体アレイ + 死角形成法)

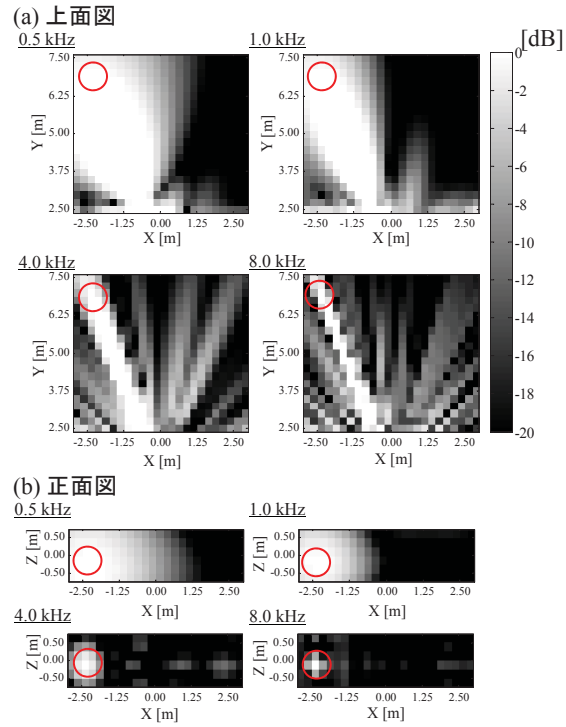


図 4.8: P1 を強調した場合の方向/距離感度特性 ((S2) 中空配置型アレイ + 死角形成法)

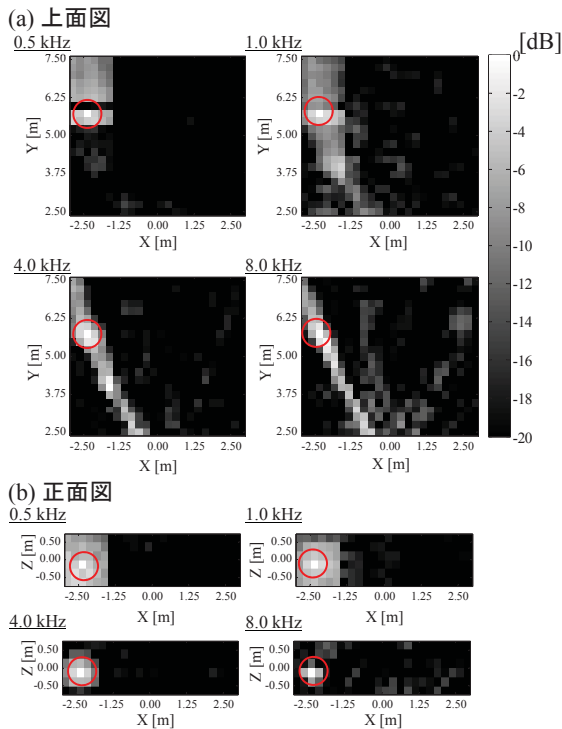


図 4.9: P2 を強調した場合の方向/距離感度特性 ((S4) 包囲型反射構造体アレイ + 死角形成法)

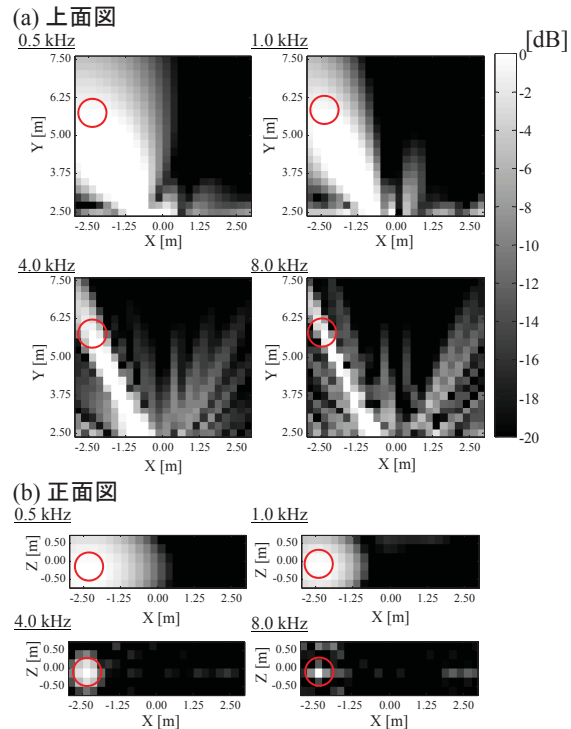


図 4.10: P2 を強調した場合の方向/距離感度特性 ((S2) 中空配置型アレイ + 死角形成法)

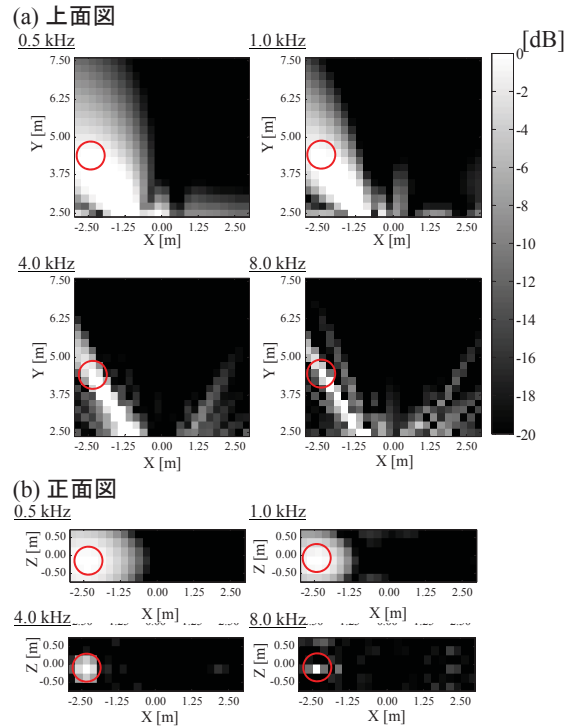
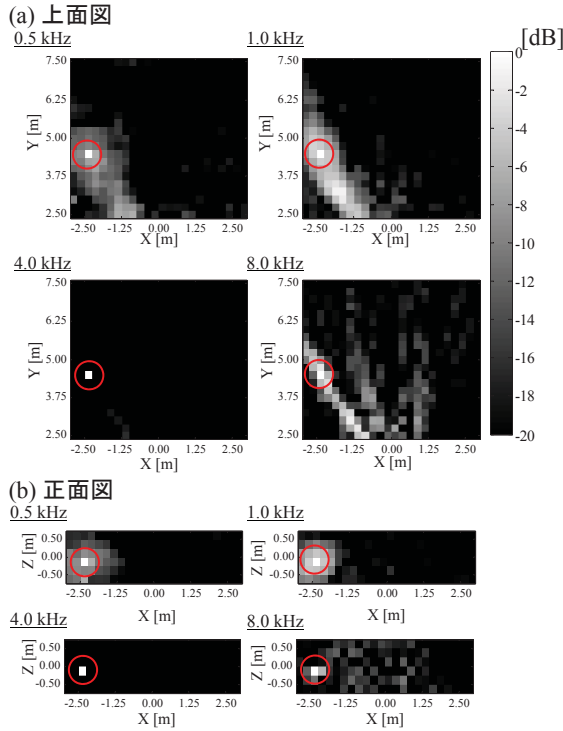


図 4.11: P3 を強調した場合の方向/距離感度特性 ((S4) 包囲型反射構造体アレイ + 死角形成法) 図 4.12: P3 を強調した場合の方向/距離感度特性 ((S2) 中空配置型アレイ + 死角形成法)

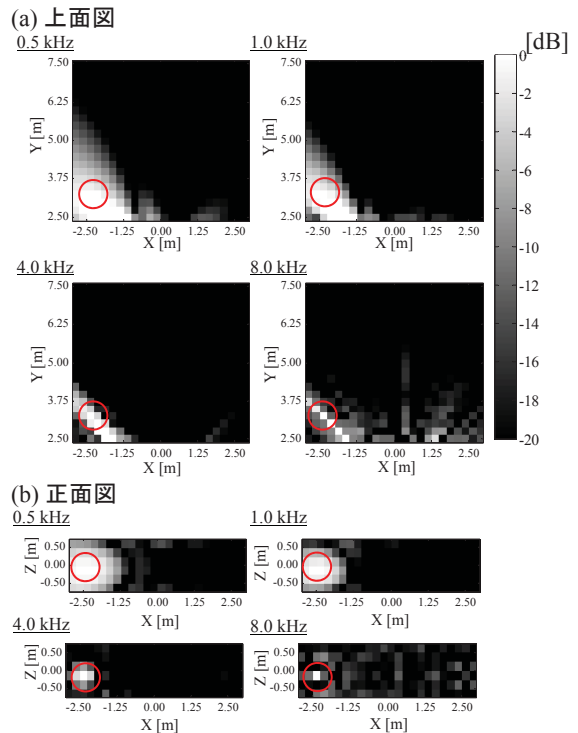
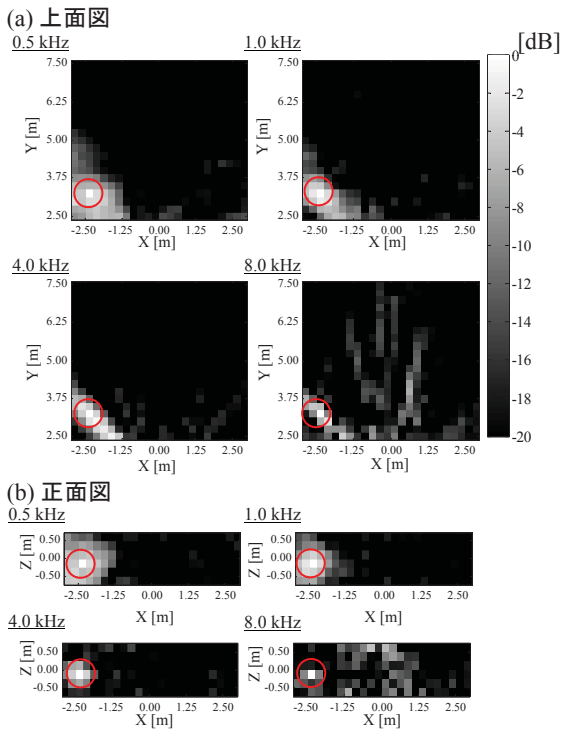


図 4.13: P4 を強調した場合の方向/距離感度特性 ((S4) 包囲型反射構造体アレイ + 死角形成法) 図 4.14: P4 を強調した場合の方向/距離感度特性 ((S2) 中空配置型アレイ + 死角形成法)

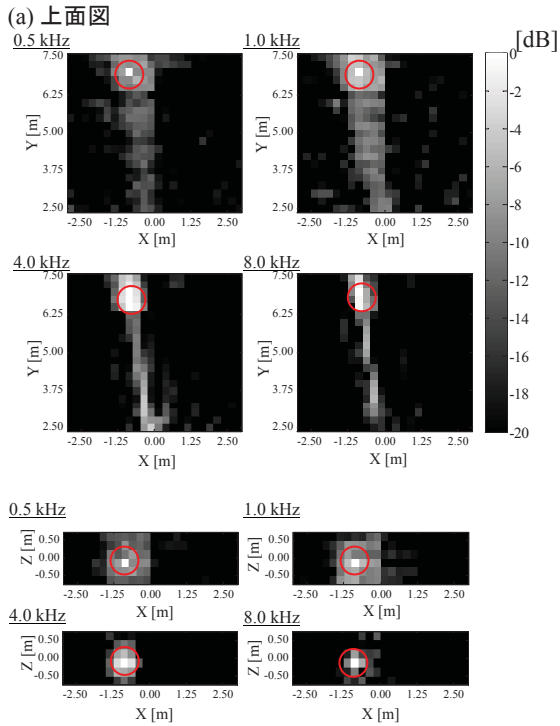


図 4.15: P5 を強調した場合の方向/距離感度特性 ((S4) 包囲型反射構造体アレイ + 死角形成法)

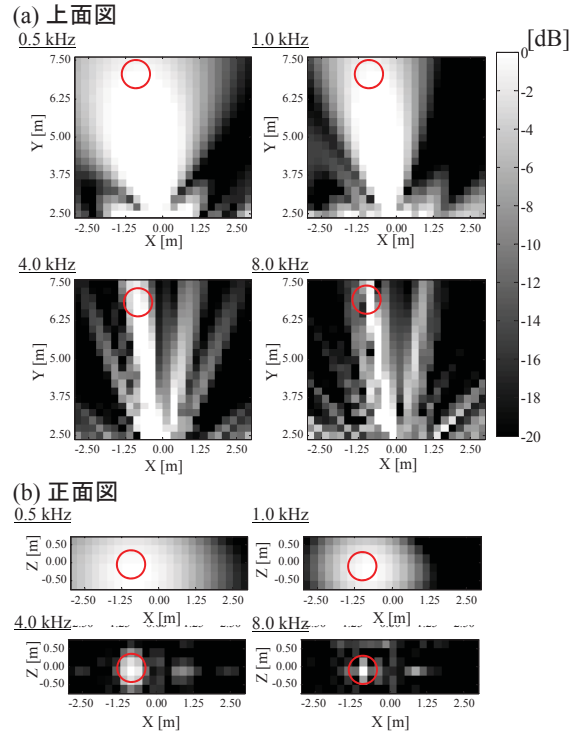


図 4.16: P5 を強調した場合の方向/距離感度特性 ((S2) 中空配置型アレイ + 死角形成法)

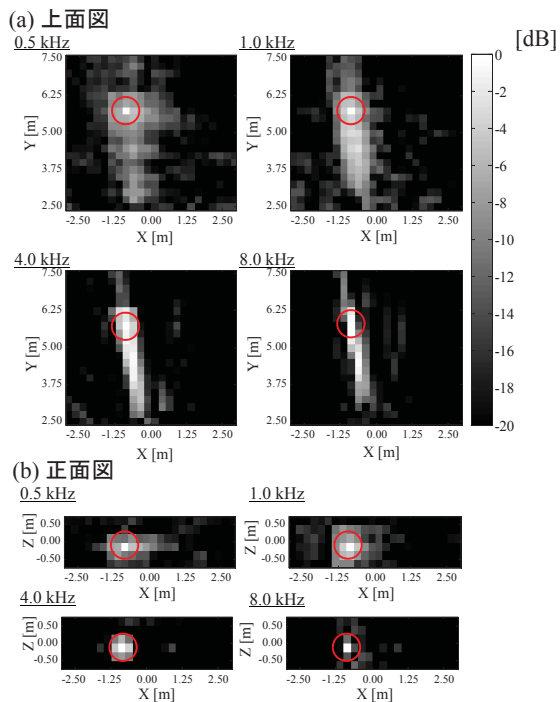


図 4.17: P6 を強調した場合の方向/距離感度特性 ((S4) 包囲型反射構造体アレイ + 死角形成法)

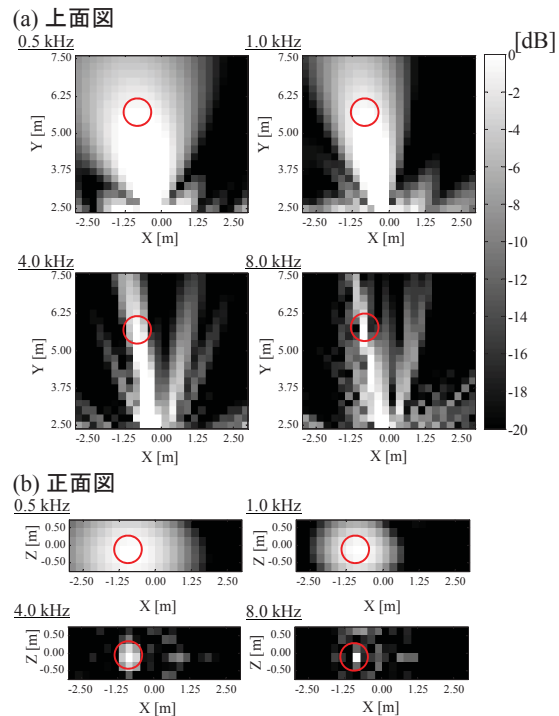


図 4.18: P6 を強調した場合の方向/距離感度特性 ((S2) 中空配置型アレイ + 死角形成法)

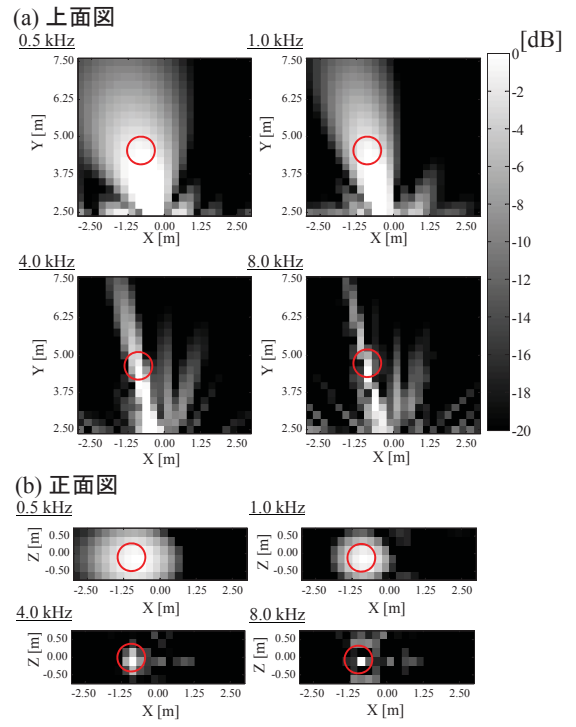
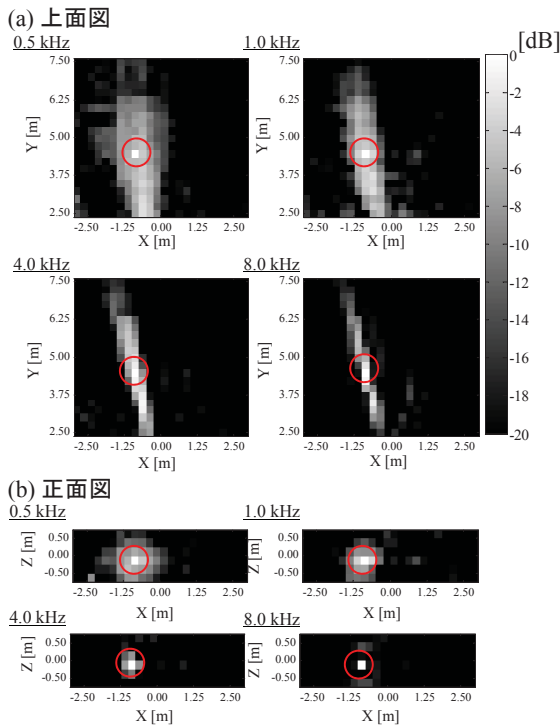


図 4.19: P7 を強調した場合の方向/距離感度特性 ((S4) 包囲型反射構造体アレイ + 死角形成法) 性 ((S2) 中空配置型アレイ + 死角形成法)

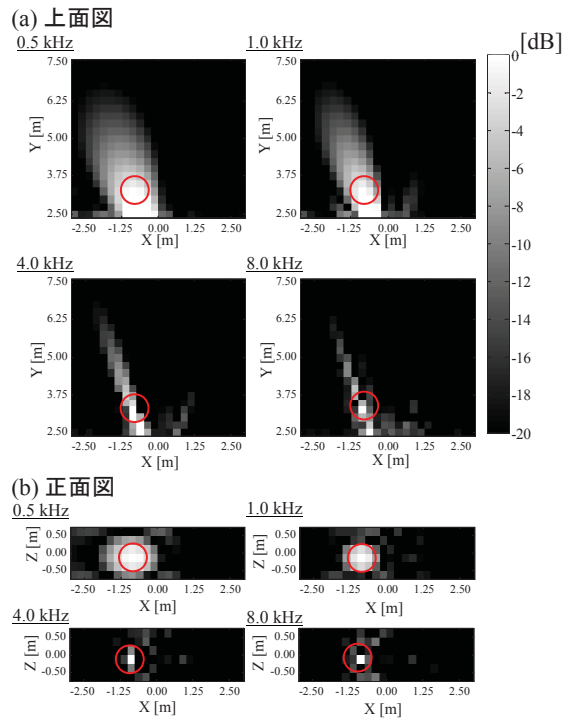
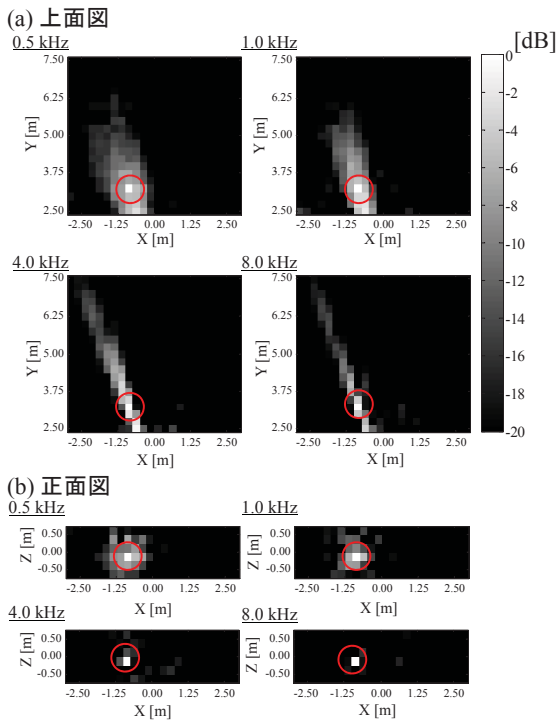


図 4.21: P8 を強調した場合の方向/距離感度特性 ((S4) 包囲型反射構造体アレイ + 死角形成法) 性 ((S2) 中空配置型アレイ + 死角形成法)

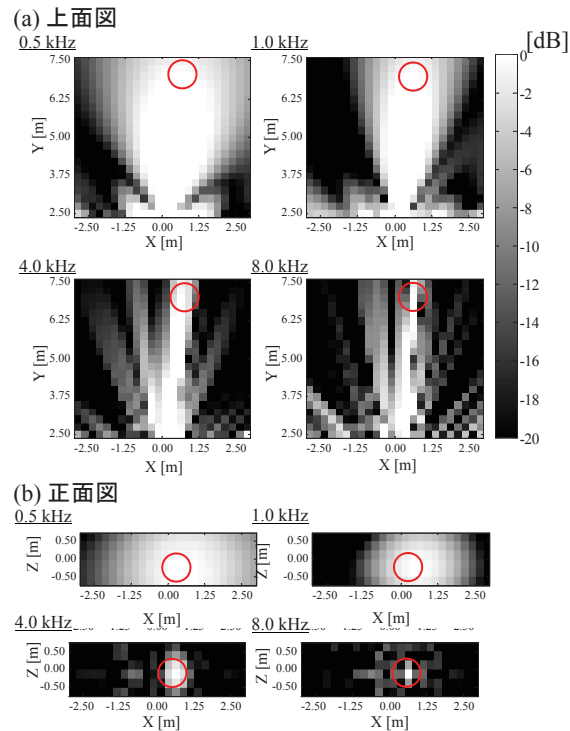
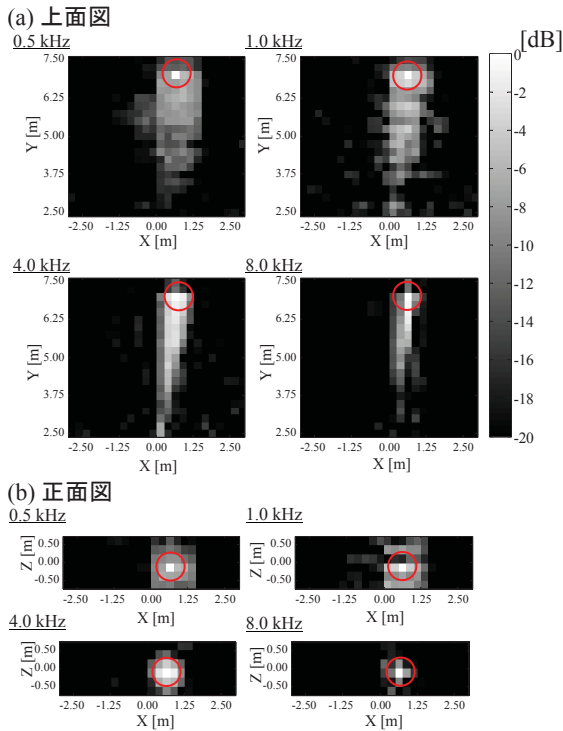


図 4.23: P9 を強調した場合の方向/距離感度特性 ((S4) 包囲型反射構造体アレイ + 死角形成法)

図 4.24: P9 を強調した場合の方向/距離感度特性 ((S2) 中空配置型アレイ + 死角形成法)

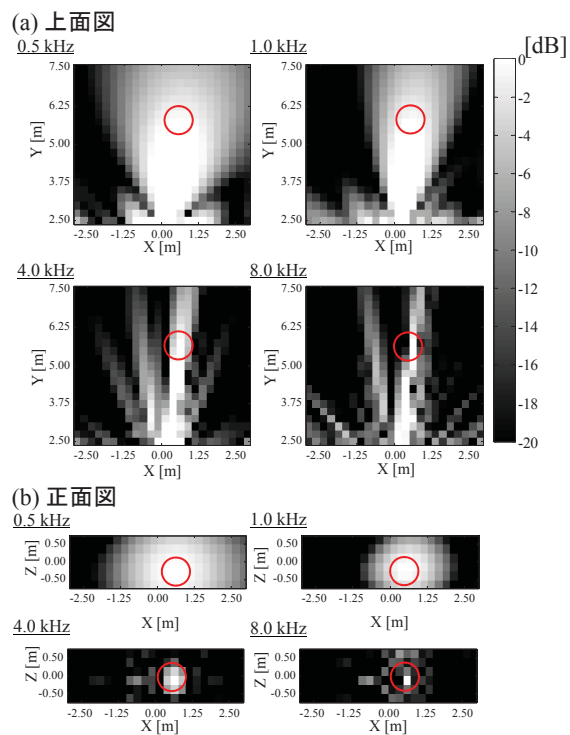
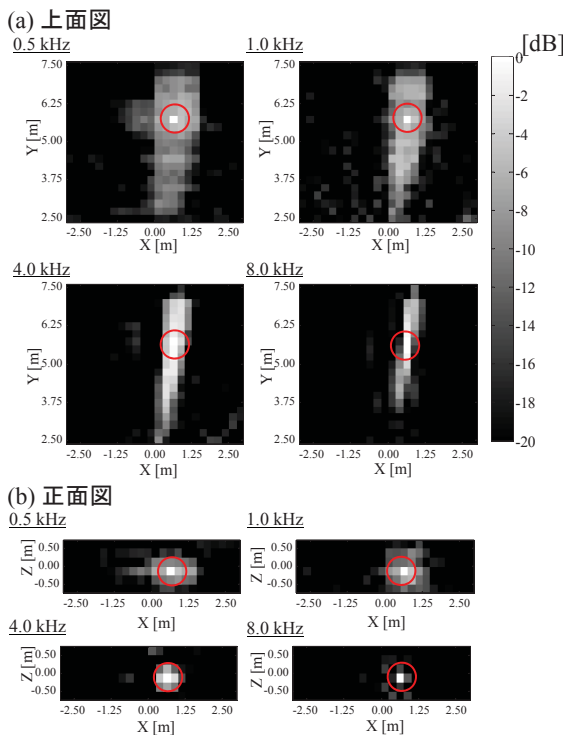


図 4.25: P10 を強調した場合の方向/距離感度特性 ((S4) 包囲型反射構造体アレイ + 死角形成法)

図 4.26: P10 を強調した場合の方向/距離感度特性 ((S2) 中空配置型アレイ + 死角形成法)

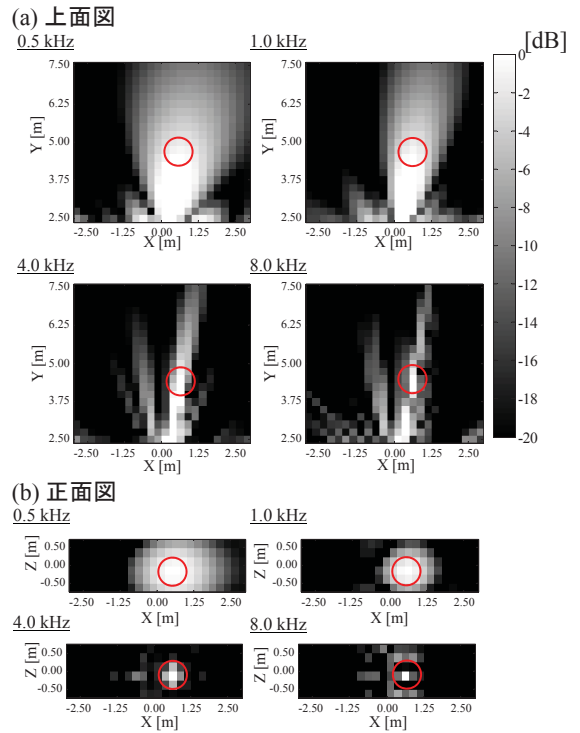
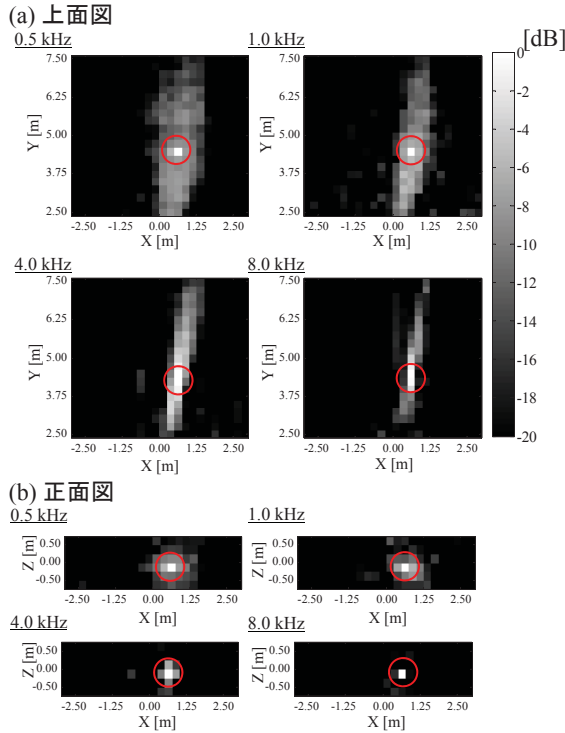


図 4.27: P11 を強調した場合の方向/距離感度特性 ((S4) 包囲型反射構造体アレイ + 死角形成法) 図 4.28: P11 を強調した場合の方向/距離感度特性 ((S2) 中空配置型アレイ + 死角形成法)

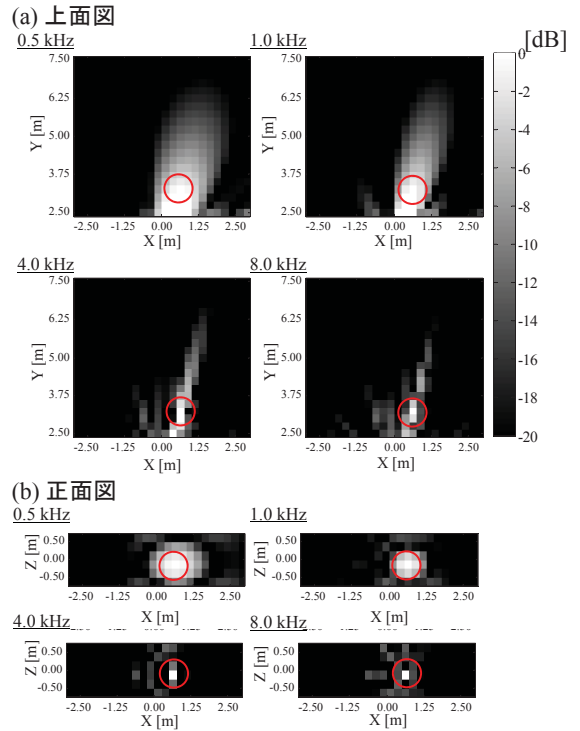
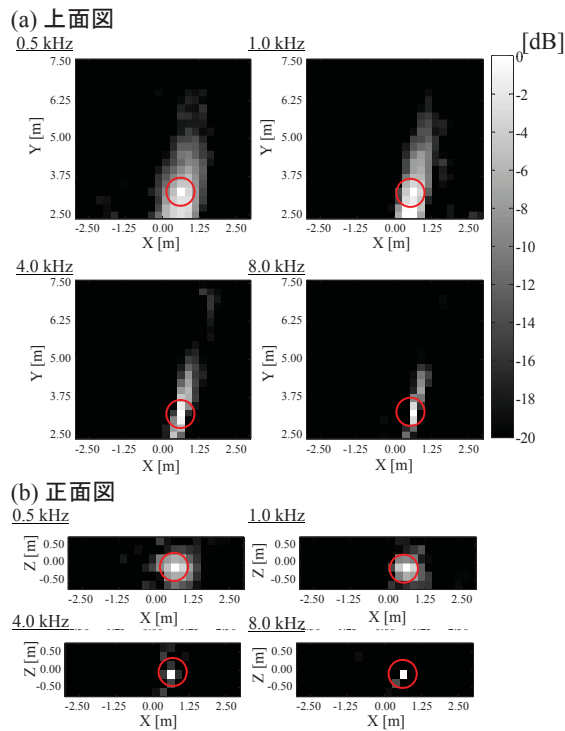


図 4.29: P12 を強調した場合の方向/距離感度特性 ((S4) 包囲型反射構造体アレイ + 死角形成法) 図 4.30: P12 を強調した場合の方向/距離感度特性 ((S2) 中空配置型アレイ + 死角形成法)

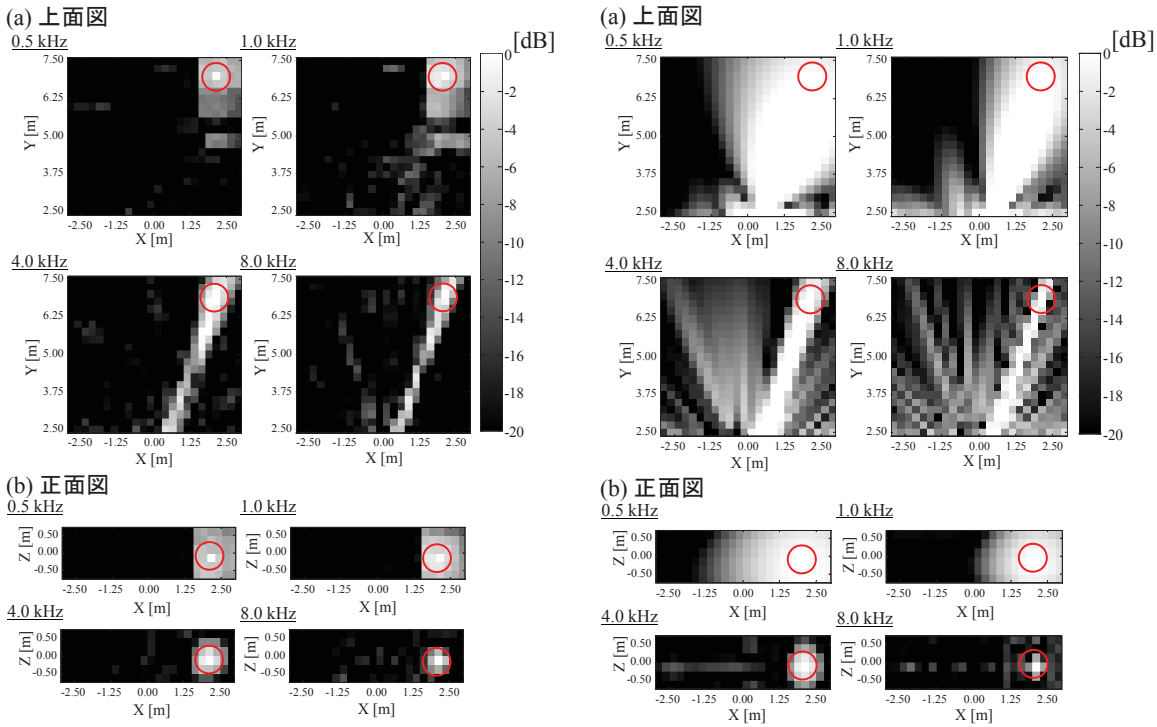


図 4.31: P13 を強調した場合の方向/距離感度特性 ((S4) 包囲型反射構造体アレイ + 死角形成法) 図 4.32: P13 を強調した場合の方向/距離感度特性 ((S2) 中空配置型アレイ + 死角形成法)

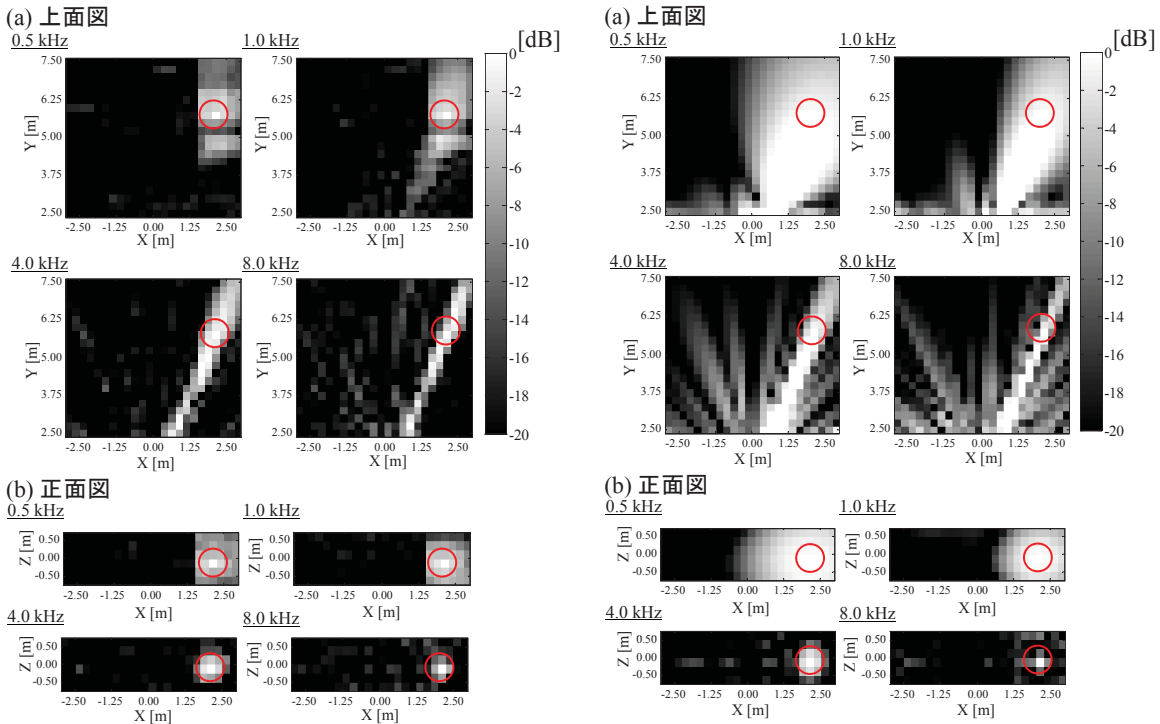


図 4.33: P14 を強調した場合の方向/距離感度特性 ((S4) 包囲型反射構造体アレイ + 死角形成法) 図 4.34: P14 を強調した場合の方向/距離感度特性 ((S2) 中空配置型アレイ + 死角形成法)

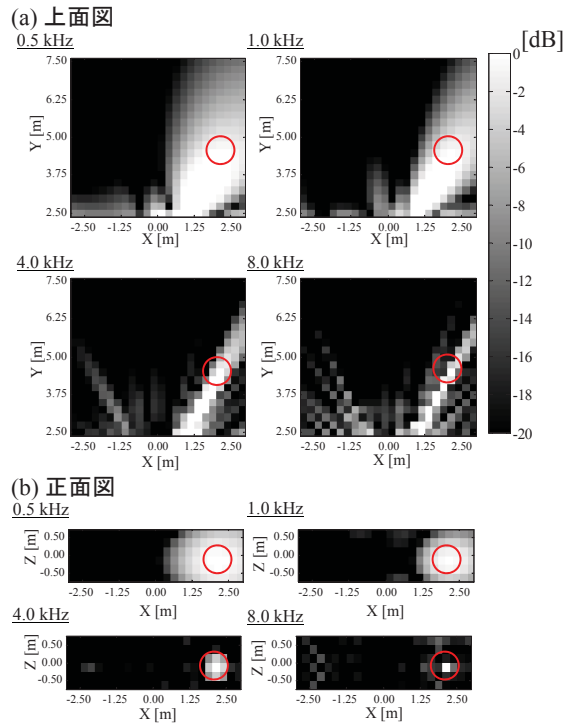
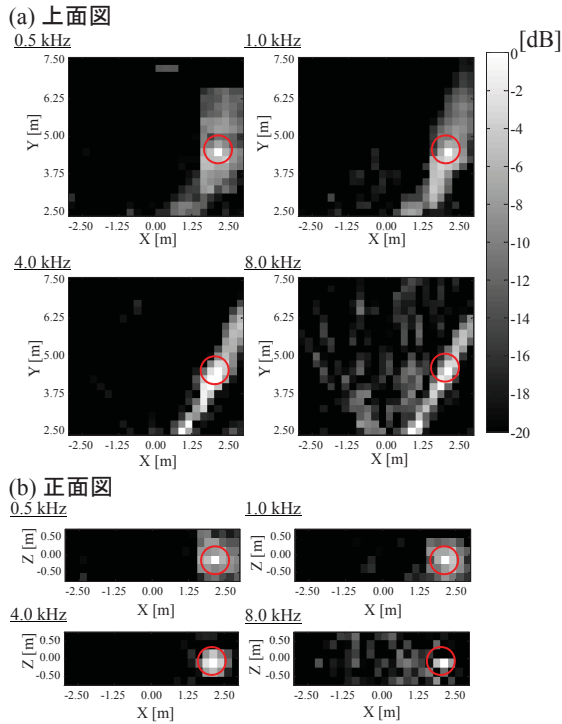


図 4.35: P15 を強調した場合の方向/距離感度特性 ((S4) 包囲型反射構造体アレイ + 死角形成法) 図 4.36: P15 を強調した場合の方向/距離感度特性 ((S2) 中空配置型アレイ + 死角形成法)

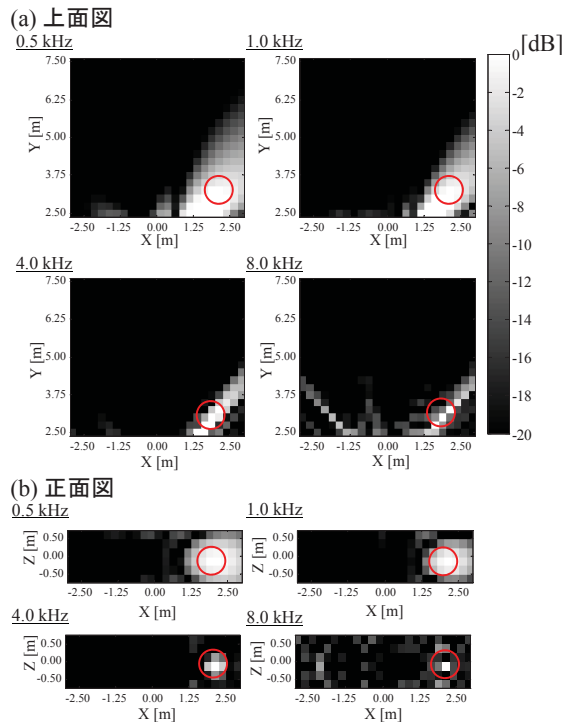
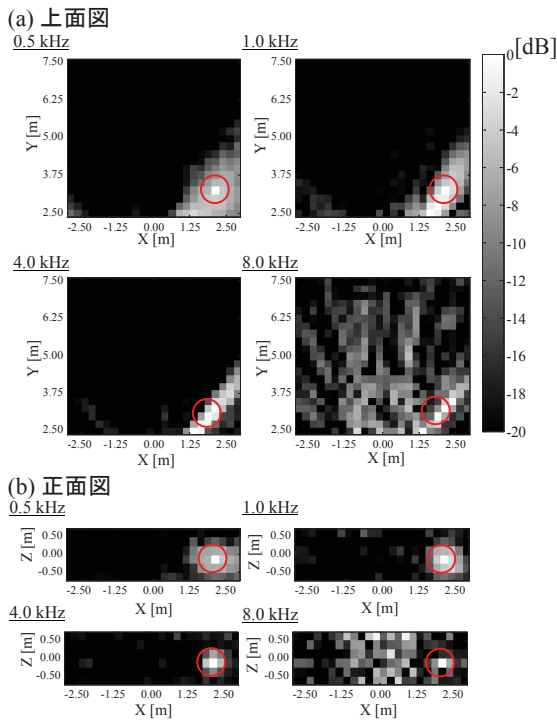


図 4.37: P16 を強調した場合の方向/距離感度特性 ((S4) 包囲型反射構造体アレイ + 死角形成法) 図 4.38: P16 を強調した場合の方向/距離感度特性 ((S2) 中空配置型アレイ + 死角形成法)

第5章 拡散受音に基づく音源距離推定

本章では、拡散受音に基づく音源距離推定に関して説明する。室内にアレイを設置して観測することを拡散受音とする。アレイから音源までの距離に応じて変化する伝達特性を事前に既知とし、伝達特性を利用した信号処理により音源距離を推定する。5.1節で室内における観測信号をモデル化し、5.2節で音源距離と空間相関行列の構造の関係について説明する。5.3節では、観測信号からブラインドでターゲット音の空間相関行列を分離し、音源距離を推定するための一信号処理について述べる。5.4節では、実験により音源距離の推定性能を検証し、5.5節で本章の結論を述べる。

5.1 室内における観測信号のモデル化

室内に設置したアレイを用いて音波を観測する。音場には、複数の干渉雑音源があるとするとする。本章での目標は、観測信号を解析し、ターゲット音までの距離 $r_0^{(s)}$ を推定することである。鏡像法 [53] を用いて、室内における伝達特性をモデル化する。図 5.1 は、室内に音源を設置した時の鏡像音源（反射音）の分布を示す。図 5.1 の丸の大きさは、各鏡像音源のパワーを表す。伝達特性 $A_{k,m}(\omega)$ は、直接音と D 個の鏡像音源の重畳でモデル化される。

$$A_{k,m}(\omega) = \sum_{d=0}^D H_{k,m,d}(\omega) \quad (5.1)$$

ここで、

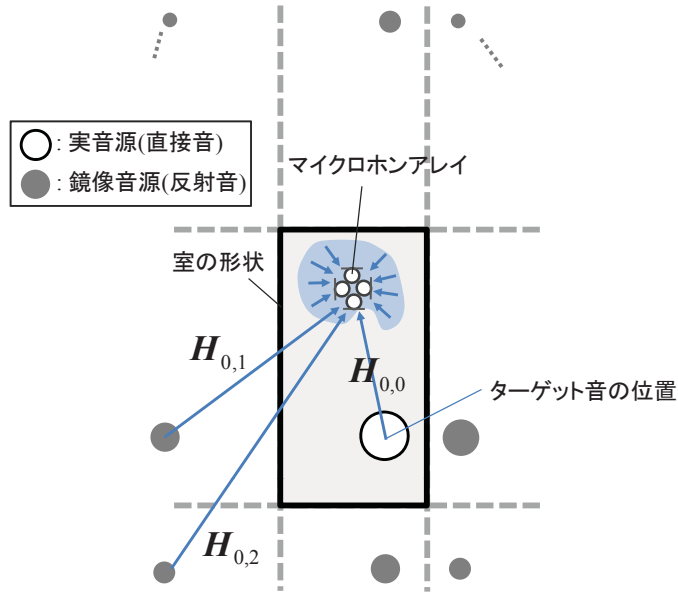


図 5.1: 鏡像法に基づく伝達特性のモデル化

$$H_{k,m,d}(\omega) = \frac{\kappa^{n(d)}(\omega)}{\|\mathbf{q}_{k,d} - \mathbf{p}_m\|} \exp\left(j\omega \frac{\|\mathbf{q}_{k,d} - \mathbf{p}_m\|}{c}\right) \quad (5.2)$$

である。なお， $H_{k,m,0}(\omega)$ は， k 番目の音源の直接音の伝搬特性を表す。また， $\kappa(\omega)$ ， $n(d)$ は，それぞれ壁の反射係数，反射次数を表す。 k 番目の音源とアレイ間の伝達特性 $\mathbf{a}_k(\omega)$ は以下で計算される。

$$\mathbf{a}_k(\omega) = \mathbf{H}_k(\omega) \overbrace{[1, \dots, 1]}^{D+1}{}^T \quad (5.3)$$

ただし，

$$\mathbf{H}_k(\omega) = [\mathbf{h}_{k,0}(\omega), \mathbf{h}_{k,1}(\omega), \dots, \mathbf{h}_{k,D}(\omega)] \quad (5.4)$$

$$\mathbf{h}_{k,d}(\omega) = [H_{k,1,d}(\omega), \dots, H_{k,M,d}(\omega)]^T \quad (5.5)$$

である。音源がアレイから十分離れていて、直接音が平面波として到来する場合、 $\mathbf{h}_{k,0}(\omega)$ は音源の到来方向 $\theta_k^{(S)}, \phi_k^{(S)}$ で決まる関数である。一方で、空間に三次元的に分布している D 個の鏡像音源の位置は、部屋の形状や音源の位置に応じて変化するので、 $\mathbf{h}_{k,d}(\omega)$ は、音源距離 $r_k^{(S)}$ に依存して変化する関数と言える。つまり、室内では、 $\mathbf{a}_k(\omega)$ は $r_k^{(S)}$ に依存して変わる。ターゲット音と K 個の干渉雑音を観測する時、観測信号 $\mathbf{x}(\omega, \tau)$ は、式 (5.6) でモデル化される。

$$\mathbf{x}(\omega, \tau) = \mathbf{b}_0(\omega, \tau) + \sum_{k=1}^K \mathbf{b}_k(\omega, \tau) \quad (5.6)$$

ここで、 $\mathbf{b}_k(\omega, \tau) = [B_{k,1}(\omega, \tau), \dots, B_{k,M}(\omega, \tau)]^T$ は、伝達特性が畳み込まれた k 番目の音源信号を表す。

$$\mathbf{b}_k(\omega, \tau) = \mathbf{a}_k(\omega) Z_k(\omega, \tau) \quad (5.7)$$

5.2 音源距離と空間相関行列の構造の関係

本節では、何らかの信号処理によって、 $\mathbf{x}(\omega, \tau)$ から $\mathbf{b}_k(\omega, \tau)$ が分離されたこと仮定する。 $\mathbf{b}_k(\omega, \tau)$ を解析することで、 $r_k^{(S)}$ の推定に有効な手がかりを得られることを示す。なお、 $\mathbf{x}(\omega, \tau)$ から $\mathbf{b}_k(\omega, \tau)$ を分離するための一方式については、5.3 節で説明する。 $\mathbf{b}_k(\omega, \tau)$ に含まれる空間的な特性を解析するために、空間相関行列 $\mathbf{R}_k(\omega)$ を計算する。

$$\begin{aligned}
\mathbf{R}_k(\omega) &= \langle \mathbf{b}_k(\omega, \tau) \mathbf{b}_k^H(\omega, \tau) \rangle \\
&= \langle |Z_k(\omega, \tau)|^2 \rangle \cdot \mathbf{a}_k(\omega) \mathbf{a}_k^H(\omega) \\
&= \langle |Z_k(\omega, \tau)|^2 \rangle \cdot \mathbf{H}_k(\omega) \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix} \mathbf{H}_k^H(\omega) \quad (5.8)
\end{aligned}$$

アレイから音源までの距離が離れるほど、直接音に対する受音パワーが低減し、鏡像音源のパワーが相対的に大きくなる。つまり、 $\mathbf{H}_k(\omega)$ の構成要素は、 $r_k^{(S)}$ に応じて変化する。故に、 $\mathbf{R}_k(\omega)$ の構造もまた $r_k^{(S)}$ に応じて変化する。

空間相関行列から $r_k^{(S)}$ を推定するための有効な手がかりを得るために、 $\mathbf{R}_k(\omega)$ を固有値分解する。

$$\mathbf{R}_k(\omega) = \mathbf{V}_k(\omega) \mathbf{\Lambda}_k(\omega) \mathbf{V}_k^H(\omega) \quad (5.9)$$

ここで、

$$\mathbf{V}_k(\omega) = [\mathbf{v}_{k,1}(\omega), \dots, \mathbf{v}_{k,M}(\omega)] \quad (5.10)$$

$$\mathbf{v}_{k,m}(\omega) = [V_{k,m,1}(\omega), \dots, V_{k,m,M}(\omega)]^T \quad (5.11)$$

$$\mathbf{\Lambda}_k(\omega) = \text{diag}\{[\Lambda_{k,1}(\omega), \dots, \Lambda_{k,M}(\omega)]\} \quad (5.12)$$

である。固有値は、 $\Lambda_{k,1}(\omega) \geq \dots \geq \Lambda_{k,M}(\omega) \geq 0$ の順で整列しているとする。

図5.2に、音源距離に対する $\mathbf{R}_k(\omega)$ の固有値の変化のモデルを示す。図5.2(a)は、音源までの距離が近い場合の直接音や鏡像音源の伝搬特性と $\mathbf{R}_k(\omega)$ の固有空間の関係を表す。なお、 $\mathbf{R}_k(\omega)$ は、本来 M 次元空間であるが、可視化するために三次元に縮退

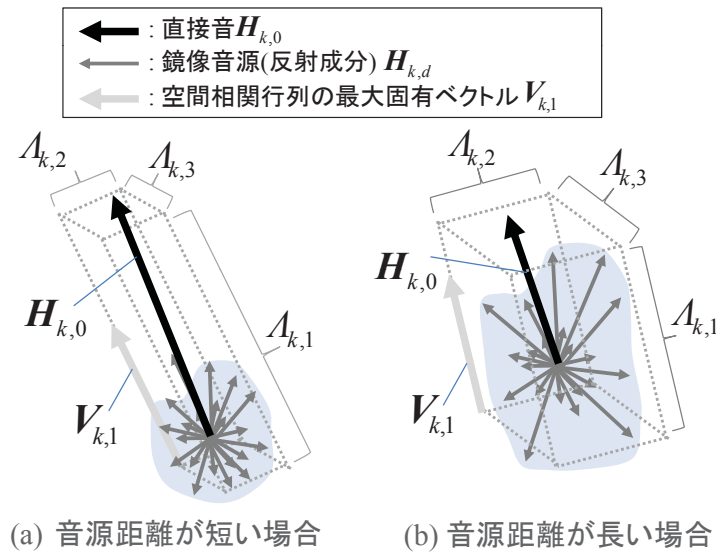


図 5.2: 音源距離に対する空間相関行列の構造変化のモデル, (a) アレイから音源までの距離 $r_k^{(S)}$ が短い場合, (b) $r_k^{(S)}$ が長い場合

して表す。図中の矢印の長さとは向きは、それぞれ伝搬特性のパワーと到来方向を示す。観測信号には大きなパワーを持つ直接音が含まれる。そのため、 $R_k(\omega)$ の固有空間は、 $h_{k,0}(\omega)$ と平行方向に長辺を持つ形状となるだろう。この時、 $\Lambda_{k,1}(\omega)$ が他の固有値と比較して顕著に大きくなる。一方、音源までの距離が遠くなるにつれて、直接音のパワーが減衰し、鏡像音源のパワーが相対的に増大する。鏡像音源は様々な方向から到来するので、図 5.2(b) のように、第二以降の固有値の値が大きくなる。つまり、残室内では $r_k^{(S)}$ に応じて、 $R_k(\omega)$ の固有値分布が変化する。そのため、 $\Lambda_{k,1}(\omega), \dots, \Lambda_{k,M}(\omega)$ は、音源距離を推定するための有用な手掛かりであると考えられる。

5.3 音源距離推定アルゴリズム

$x(\omega, \tau)$ から $R_k(\omega)$ を分離し, $r_k^{(S)}$ を推定するための信号処理の実装方式は様々あると考えられる。5.3.1節では, 干渉雑音がある中で $R_k(\omega)$ を観測信号から分離・抽出するための方式として, 周波数領域ICAを用いた方法について説明する。また, 5.3.2節では, $R_k(\omega)$ から $r_k^{(S)}$ を推定するためのアルゴリズムについて述べる。

5.3.1 周波数領域ICAに基づく空間相関行列の分離

複数の信号源が混在した観測信号から空間特性を含む信号源を得るための方式として, 周波数領域ICAに基づくブラインド音源分離を適用する。分離フィルタ $W(\omega)$ と $x(\omega, \tau)$ を周波数領域で掛け合わせることで, \hat{K} 個の分離信号 $y(\omega, \tau)$ を出力する。

$$\mathbf{y}(\omega, \tau) = \mathbf{W}(\omega)\mathbf{x}(\omega, \tau) \quad (5.13)$$

ここで,

$$\mathbf{y}(\omega, \tau) = [Y_0(\omega, \tau), Y_1(\omega, \tau), \dots, Y_K(\omega, \tau)]^T \quad (5.14)$$

$$\mathbf{W}(\omega) = [\mathbf{w}_0(\omega), \dots, \mathbf{w}_K(\omega)]^H \quad (5.15)$$

$$\mathbf{w}_k(\omega) = [W_{k,1}(\omega), \dots, W_{k,M}(\omega)]^T \quad (5.16)$$

である。周波数領域ICAに関する従来研究では, $x(\omega, \tau)$ から音源数を推定する問題と $W(\omega)$ を推定する問題を併せて議論されることがあった[55]。本節では, 音源数が事前に既知であることを仮定する。自然勾配法[56, 57]を用いてフィルタを学習する場合, 式(5.17)に従って $W(\omega)$ を逐次的に更新する。

$$\mathbf{W}(\omega) \leftarrow \mathbf{W}(\omega) + \eta \cdot \text{off-diag} [\langle \Phi(\mathbf{y}(\omega, \tau)) \mathbf{y}^H(\omega, \tau) \rangle] \mathbf{W}(\omega) \quad (5.17)$$

ここで, η , $\text{off-diag}[\cdot]$ は, それぞれ更新係数, 行列の非対角項零化演算を表す。また,

$$\Phi(\mathbf{y}(\omega, \tau)) = [\phi(Y_0(\omega, \tau)), \dots, \phi(Y_K(\omega, \tau))]^T \quad (5.18)$$

である。非線形関数 $\phi(\cdot)$ の一例として以下を用いる。

$$\phi(Y_k(\omega, \tau)) = \tanh\{\Re(Y_k(\omega, \tau))\} + j \cdot \tanh\{\Im(Y_k(\omega, \tau))\} \quad (5.19)$$

$W(\omega)$ を構成する K 個のフィルタには, 出力順序に不定性が生じることが知られている。推定したフィルタに含まれる空間特性の類似性を基準として, 全帯域に渡ってグルーピングする方法を用いる [18, 58]。

空間特性を含む分離信号 $b_k(\omega, \tau)$ を生成するために, $W(\omega)$ の振幅特性を補正する。 $W(\omega)$ の振幅特性を補正するための方法として, 観測信号と出力信号の二乗誤差を最小化するフィルタ $\mathbf{c}_k(\omega) = [C_{k,1}(\omega), \dots, C_{k,M}(\omega)]^T$ を生成する [60, 61]。 $\mathbf{c}_k(\omega)$ は, 式 (5.20) で計算される。

$$\mathbf{c}_k(\omega) = \frac{\mathbf{R}(\omega)\mathbf{w}_k(\omega)}{\mathbf{w}_k^H(\omega)\mathbf{R}(\omega)\mathbf{w}_k(\omega)} \quad (5.20)$$

ここで, $\mathbf{R}(\omega)$ は, 観測信号から計算した空間相関行列である。

$$\mathbf{R}(\omega) = \langle \mathbf{x}(\omega, \tau)\mathbf{x}^H(\omega, \tau) \rangle \quad (5.21)$$

ターゲット音から m 番目のマイクロホンまでの伝達特性を含む分離信号を生成するために, フィルタの振幅特性を式 (5.22) で補正する [59]。

$$\mathbf{w}_k^{(m)}(\omega) \leftarrow C_{k,m}(\omega)\mathbf{w}_k(\omega) \quad (5.22)$$

$\mathbf{w}_k^{(m)}(\omega)$ を離散フーリエ逆変換により，時間領域の分離フィルタ $w_{k,1}^{(m)}(j), \dots, w_{k,M}^{(m)}(j)$ に変換し，観測信号と畳み込むことで，空間特性を含む分離信号 $b_{k,m}(t)$ を得る。

$$b_{k,m}(t) = \sum_{n=1}^M \sum_{j=0}^{J-1} w_{k,n}^{(m)}(j) x_n(t-j) \quad (5.23)$$

$b_{k,m}(t)$ を短時間離散フーリエ変換して $B_{k,m}(\omega, \tau)$ を算出し，その共分散行列を計算することでターゲット音の空間相関行列 $\mathbf{R}_k(\omega)$ を推定した。

$$\mathbf{R}_k(\omega) = \frac{1}{|\xi|} \sum_{\tau=1}^{|\xi|} \mathbf{b}_k(\omega, \tau) \mathbf{b}_k^H(\omega, \tau) \quad (5.24)$$

ここで， $|\xi|$ は総フレーム数を表す。

5.3.2 空間相関行列の固有値分布を用いた音源距離推定

5.2節にて， $\mathbf{R}_k(\omega)$ の固有値分布に， $r_k^{(S)}$ を推定するための手掛かりが含まれることを説明した。本節では，音源距離に対する固有値分布の振舞いを事前に観測した信号から学習し，観測した信号から得られる特徴との類似性を計算することにより， $r_k^{(S)}$ を推定する。式(5.24)で算出した $\mathbf{R}_k(\omega)$ を固有値分解し， $\Lambda_{k,1}(\omega), \dots, \Lambda_{k,M}(\omega)$ を得る。源信号の性質が与える固有値分布への影響を抑えるために，各周波数毎に固有値を正規化する。

$$\bar{\Lambda}_{k,m}(\omega) = \frac{\Lambda_{k,m}(\omega)}{\sum_{m=1}^M \Lambda_{k,m}(\omega)} \quad (5.25)$$

固有値分布を評価するための一つの尺度として，最大固有値とその他の固有値の和の比(固有値比) Θ_k を用いる。

$$\Theta_k = \frac{1}{|\Omega|} \sum_{\omega \in \Omega} \frac{\bar{\Lambda}_{k,1}(\omega)}{\sum_{m=2}^M \bar{\Lambda}_{k,m}(\omega)} \quad (5.26)$$

ここで、 Ω は解析帯域のインデックス集合を表し、 $|\Omega|$ はインデックスの総数を表す。

残響時間やアレイ構造といった条件に応じて、反射音の到来方向や受信パワーは変化し、空間相関行列の固有値分布の振舞いは変化する。事前に測定して得られたデータから音源距離と固有値比の対応関係を学習し、観測信号から計算した Θ_k と比較することで、 $r_k^{(S)}$ を推定することができるだろう。アレイから距離 d だけ離れた位置にある音源の固有値比モデルを $\tilde{\Theta}_d$ とすると、観測信号から得られた固有値比との差が最小となる距離 d を $r_k^{(S)}$ の推定結果として出力する。

$$\hat{r}_k^{(S)} = \arg \min_{d \in \Psi} \{|\Theta_k - \tilde{\Theta}_d|\} \quad (5.27)$$

ここで、 Ψ は、事前に固有値比のモデルを用意する音源距離のリストを表す。

5.4 実験

室内で収録した音声信号を用いて、本章で説明してきた音源距離推定方式の有効性を検証する。

5.4.1 実験条件

図 5.3 に収録環境を示す。残響時間 $T_{60} = 300$ msec の室にバフル型アレイを設置して収録した。十二面体型バフルの表面には、 $M = 60$ 本の無指向性マイクロホンが設置されている [63]。音源距離や音源間の角度差といった条件により、音源分離性能が変わることが考えられる。音源分離性能と音源距離の推定精度の関係を調査するために、音源間の角度差を 90 、 180 度の 2 パターン用意した。

$r_k^{(S)}$ の組み合わせを変えて、 $K = 2$ 個の音源を配置した。 $r_k^{(S)}$ のリストとして、 0.5 m から 2.5 m の間で 0.5 m 間隔で計 5 種類を用意した。また、評価用の音源信号とし

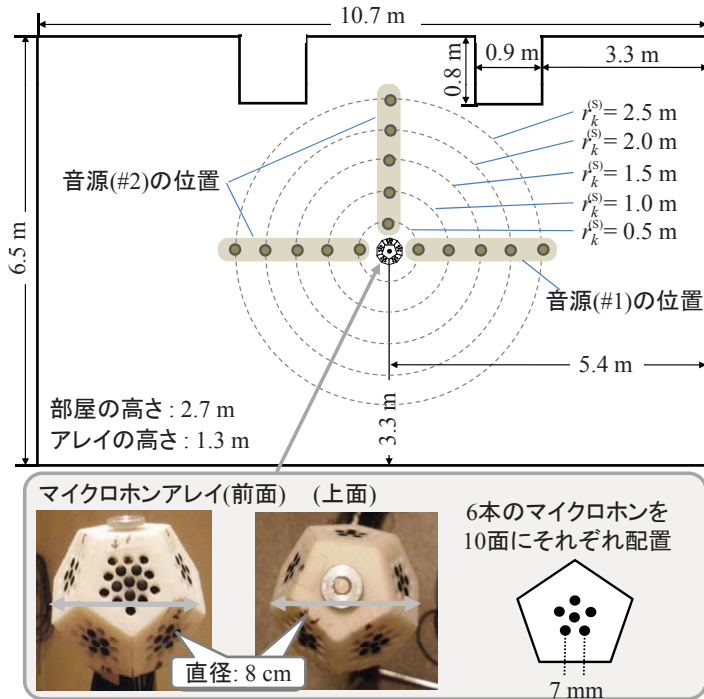


図 5.3: 収録環境とマイクロホンアレイの構成

て、全6種類の男女の発話音声を用いた。その他の実験条件を表 5.1 に示す。

5.4.2 実験結果

周波数領域 ICA により分離した信号源 $\hat{b}_{k,m}(t)$ における分離性能を評価するための尺度として、SDR (Signal to Distortion Ratio) を用いた。SDR は、式 (5.28) で計算される。

$$\text{SDR} = 10 \log_{10} \left(\frac{\sum_t b_{k,\tilde{M}}^2(t)}{\sum_t \{b_{k,\tilde{M}}(t) - \hat{b}_{k,\tilde{M}}(t)\}^2} \right) [\text{dB}] \quad (5.28)$$

ここで、 \tilde{M} は参照するマイクロホンのインデックスであり、ターゲット音源から最も近い位置にあるマイクロホンを選択した。SDR が大きいほど、空間特性を含めた音源

表 5.1: 実験条件

サンプリング周波数	16 kHz
音源数, K	2
マイクロホン数, M	60
音源距離, r_k	0.5, 1.0, 1.5, 2.0, 2.5 m
室の残響時間, T_{60}	300 ms
評価用の音源信号	男女発話 (全 6 種)
音源信号長	12 s
分析窓長	2048 taps (128 ms)
分析シフト長	1024 taps (64 ms)
分離フィルタ長, J	2048 taps (128 ms)
空間相関行列の解析帯域, Ω	1.0 - 4.0 kHz
更新係数, μ	0.001

信号が忠実に再現されているので、高い音源分離性能が得られていることを示す。 $r_k^{(S)}$ を推定するための手がかりである固有値比 Θ_k は、 M 個の分離信号 $\hat{b}_{k,1}(t), \dots, \hat{b}_{k,M}(t)$ から $\hat{R}_k(\omega)$ を生成することで算出した。全ての音源配置パターンで計算した Θ_k を5種類の音源距離毎に整理し、その平均値を各音源距離 $r_k^{(S)}$ における固有値比のモデル $\tilde{\Theta}_d$ として用いることとした。各評価用データ毎に、式(5.27)で音源距離を推定した。また、固有値比のモデルの学習用データによる結果の差異を確認するために、クローズドセットでの評価だけでなく、評価用のセットとは異なる音源信号で学習した $\tilde{\Theta}_d$ を用いるオープンセットでの評価を行った。

ターゲット音や干渉雑音の音源距離による音源分離性能への影響について、図5.4に示す。音源数よりも多くのマイクロホンを用いて音源分離処理を行っているため、全体的に高い分離性能であった。音源間の角度差によるSDRへの影響は十分に小さかつ

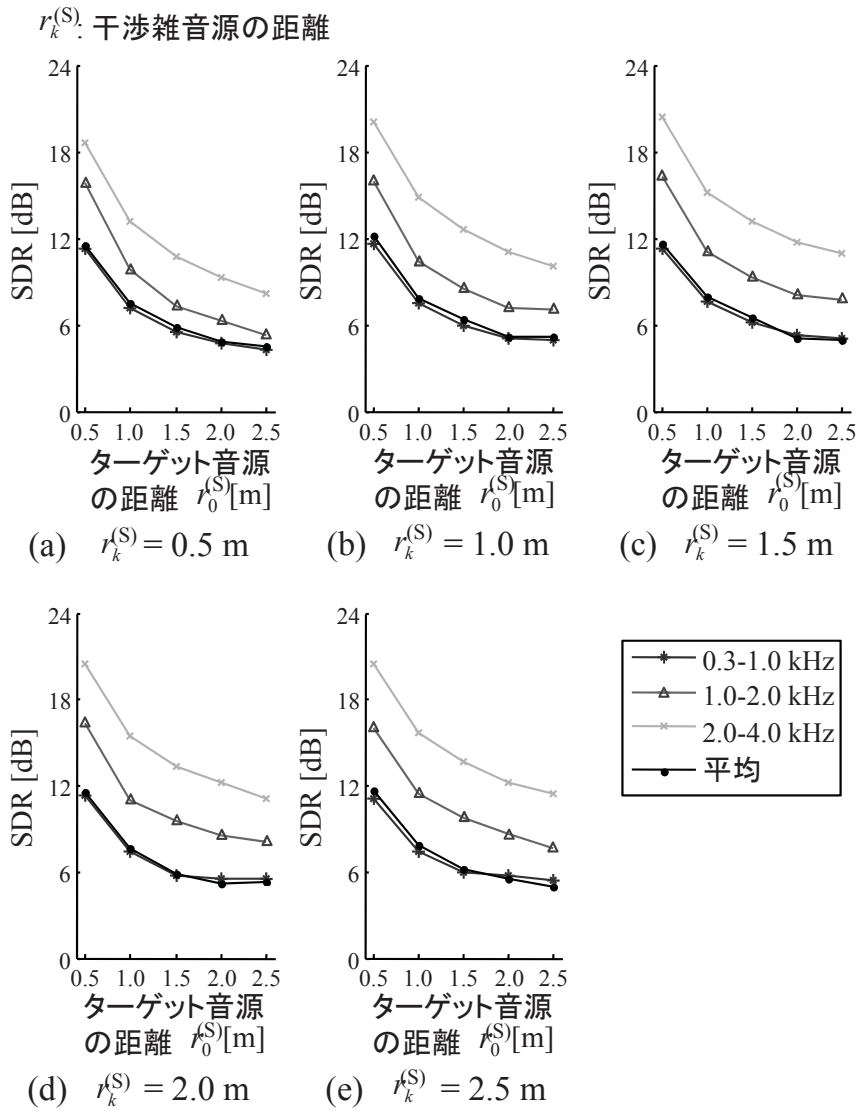


図 5.4: SDR を用いた分離性能の評価

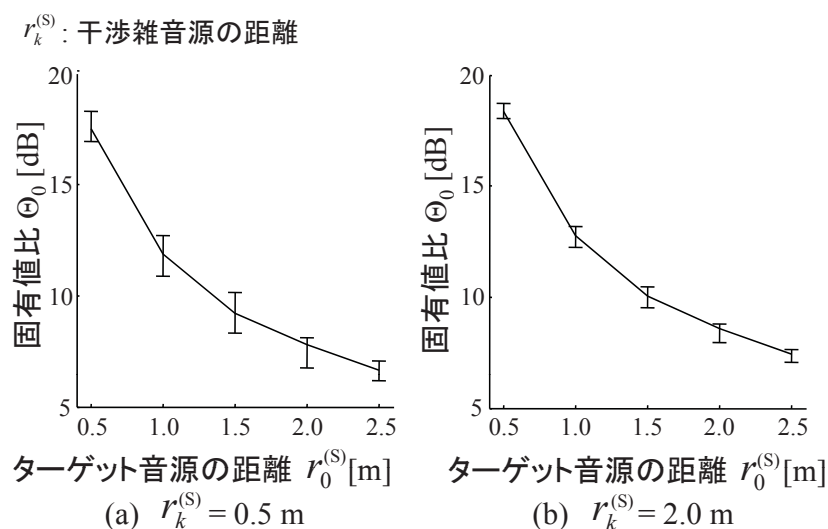


図 5.5: 音源距離と固有値比の関係

たので, $r_k^{(S)}$ と SDR の関係を示した。干渉雑音の音源距離に関わらず, ターゲット音までの距離 $r_0^{(S)}$ が近いほど, SDR が高くなる傾向が見られた。これは, 観測時における干渉雑音に対するターゲット音の SN 比が高いためだと考えられる。周波数帯域ごとに SDR を算出したところ, 干渉雑音が 1.0 m 以上離れている場合に, 1.0–4.0 kHz の帯域において SDR が高かったので, Θ_k を求めるための解析帯域 Ω として用いることとした。

図 5.5 に, Θ_k と $r_k^{(S)}$ の関係を示す。図中のバーは, 源信号や音源方向の違いにより, ある程度ばらつきがあることを示している。実験結果から, 干渉雑音の距離に関わらず, ターゲット音源がアレイから離れるほど, Θ_k が小さくなる傾向が見られた。これは, 5.2 節で説明したように, $r_k^{(S)}$ に応じて観測信号に含まれる反射音のパワーが変化し, $\mathbf{R}_k(\omega)$ の固有値分布が変化するためだと考えられる。一方, 源信号や音源方向の違いによるばらつきは小さかった。つまり, 音源距離に対する固有値比の振る舞いをあらかじめ学習することで, $r_k^{(S)}$ を推定できると考えられる。

図 5.6, 5.7 は, それぞれ, クローズドセットを用いた場合, オープンセットを用い

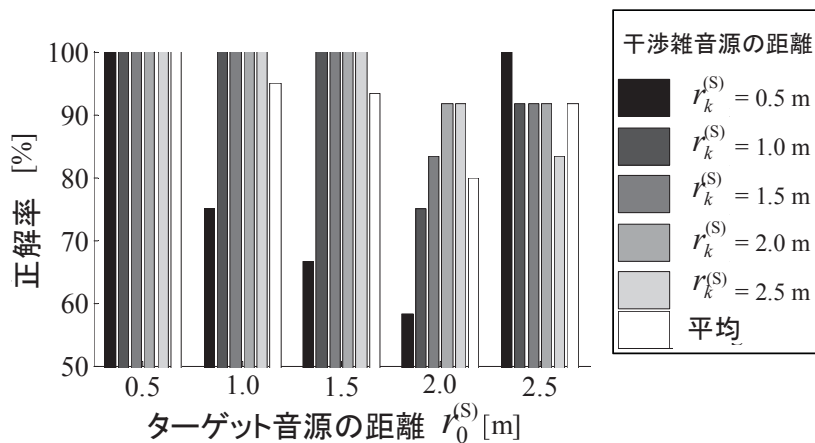


図 5.6: 音源距離の推定精度 (クローズドセットでの評価)

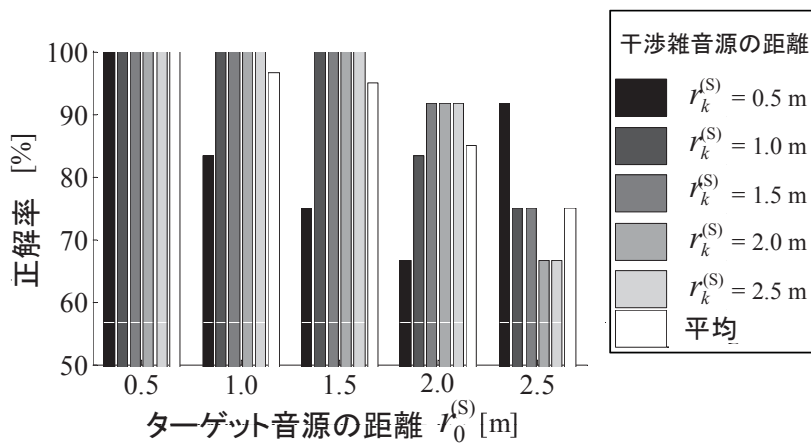


図 5.7: 音源距離推定精度 (オープンセットでの評価)

た場合で学習した固有値比を用いて算出した音源距離の推定精度を表す。5 パターンの距離 ($r_k^{(S)} = 0.5, 1.0, 1.5, 2.0, 2.5$ m) の中から正確に評価用データにおける音源距離を選択できる精度を距離推定精度として定義した。ターゲット音までの距離が近い場合、分離性能が高いため、 $r_k^{(S)}$ に対して Θ_k の変化が大きく、距離推定精度が高まる傾向が見られた。一方、アレイと干渉雑音間の距離が 1.0 m 以下である場合、分離性能が劣化し、距離推定精度もまた低下する傾向が見られた。音源距離毎に平均化した推定精度は、クローズドセットを用いた場合で 92 %、オープンセットを用いた場合で 90 % であった。1 音源のみを観測した信号から固有値を計算し、 $r_k^{(S)}$ を推定したところ、ほぼ誤差なく推定できたことから、音源分離性能の影響が音源距離推定に影響を与えていることが明らかになった。また、クローズドセット、オープンセットを用いた場合の音源距離推定における誤差分布をそれぞれ図 5.8, 5.9 に示す。各距離での頻度を丸の大きさで表した。ターゲット音源がアレイから遠ざかるほど、音源距離に対して Θ_k の変化幅が小さくなるので、誤差が大きくなる傾向が見られた。しかし、隣接した距離に誤差が分布しており、空間相関行列の固有空間の性質を用いた音源距離推定精度が高いことを確認した。

5.5 本章の結論

室内に設置したアレイを用いて観測することを拡散受信とし、音源距離を推定するための方式について説明した。空間相関行列の構造が音源距離に応じて特徴的に変わること示した。空間相関行列の構造と音源距離の対応関係についてあらかじめ学習すれば、音源距離を推定できることを示した。

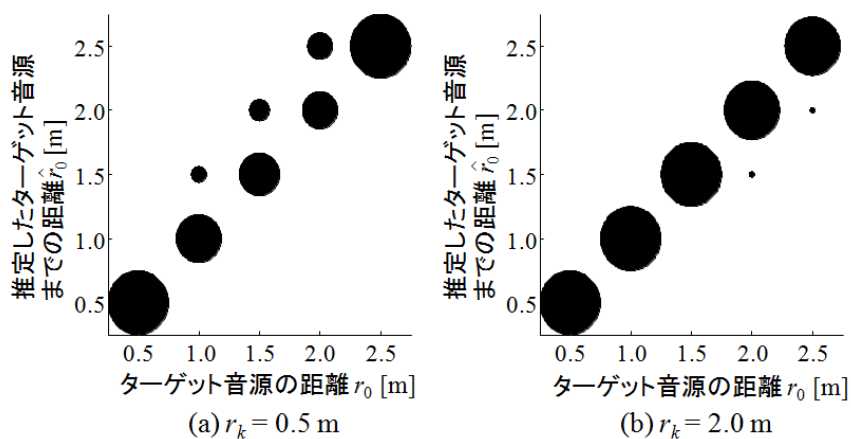


図 5.8: 音源距離推定における誤差分布 (クローズドセットでの評価)

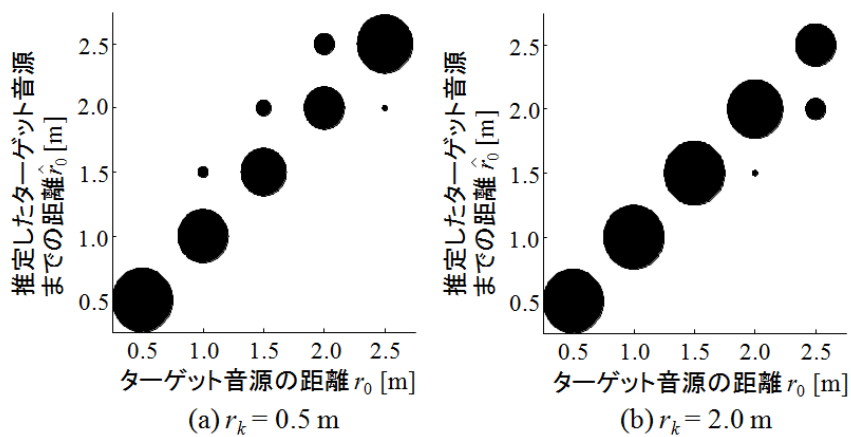


図 5.9: 音源距離推定における誤差分布 (オープンセットでの評価)

第6章 結論

本論文は、音源情報を解析するために、どのように受音すべきかについて研究した成果をまとめたものである。従来のアレイ信号処理では、中空に配置した複数のマイクロホンで構成された中空配置型アレイ受音することを前提とし、直接音の伝搬特性（アレイ・マニフォールドベクトル）を操作することで、音源情報を推定してきた。しかし、特定の周波数を除いて音源情報を高精度に解析することが困難であり、多雑音環境下における信号源の推定や音源距離を推定することが困難であった。本研究では、従来のアレイ信号処理では性能を低下させる要因として扱われてきた反射・拡散音を有用な情報として活用することで、これまで推定できなかった音源情報の推定に繋がるのではないかとこの着想から研究を始めるに至った。

本研究の主張点は、反射・拡散させた音を観測すること（拡散受音）により、広帯域に渡ってチャンネル間相関が低減し、音源情報を分離・抽出しやすい性質を持つ信号を受音できるという点にある。また、ターゲット音とアレイ間の伝達特性が事前に既知であれば、多雑音環境下で信号源を強調したり、音源までの距離や向きを推定できることを実証した。アレイ信号処理に関する従来研究では、主に信号処理の改良に関するアプローチがほとんどであった。しかし、本研究成果により、受音すべき信号の性質や拡散現象を利用した系実装の効果が示されたことにより、従来のアレイ信号処理を性能を超えることや新しい応用技術を創出するための新しいアプローチを与えた。

1章では、研究背景、従来のアレイ信号処理、本研究の方針について述べた。

2章では、多入力多出力（MIMO）系における相互情報量を定義し、相互情報量が最

大となるような受信信号の性質について導出した。その結果、受信時のSN比が大きくなり、チャンネル間相関が低減するような受信が望ましいことが明らかになった。相互情報量を最大化するための具体的な受信方法として、方向性のある音波を反射・拡散させて受信する拡散受信を提案した。拡散受信における空間相関行列をモデル化すると、チャンネル間相関が sinc 関数で表される。マイクロホンが十分離されて配置されている場合、広帯域に渡ってチャンネル間相関が低減し、相互情報量が最大化するような受信を実現できることを示した。数値シミュレーションにより、拡散受信の効果を検証し、拡散受信によりチャンネル間相関が低減し、空間相関行列の固有値分布が平滑化することを確認した。

3章では、拡散受信に基づく信号強調について述べた。チャンネル間相関を低減するような受信と信号強調法を組み合わせることで、広帯域に渡って信号出力パワーを低減できることを示した。この知見は、アレイ信号処理に関する多くの従来研究にみられる信号処理の高度化により雑音抑圧性能を改善するアプローチでは得難いものである。また、拡散受信時のフィルタ設計に関して検討した。数値シミュレーションより、ターゲット音とアレイ間の伝達特性を強調し、信号出力パワーを最小化するように死角形成フィルタを設計することで、高い雑音抑圧性能を得られることを明らかにした。

4章では、拡散受信の具体的な実装方式として、包囲型反射構造体と多マイクロホンを用いたアレイ（包囲型反射構造体アレイ）について説明し、信号強調に適用した。ABS板で構成した直径1.0 mの球程度の包囲型反射構造体の一面を取り外し、反射構造体内で多重に反射させた音を96本のマイクロホンで受信する系を構築した。無響室に包囲型反射構造体アレイを設置してインパルス応答を計測し、信号出力パワーや方向/距離感度特性に関して調査を行なった。包囲型反射構造体アレイを用いることで、広帯域に渡って信号出力パワーが低減し、ターゲット音の到来方向に対して高い感度を形成することを確認した。

5章では、室内に設置したアレイで観測することを拡散受音とし、音源距離を推定するための方式に関して検討した。音源とアレイ間の距離に応じて変化する空間相関行列の構造をモデル化した。事前測定やシミュレーション等の手段により、音源距離に対する空間相関行列の構造の変化が既知であれば、音源距離を推定することができる。残響時間 300 ms の会議室で行った実験では、1つの干渉雑音がある中で、音源距離を90%以上の精度で推定できた。

本研究では、多入出力を備える受音系の構造を決めるための一つの指針を示すことができたと考えている。多入出力デバイスが一般的に利用できるようになった現代において、受音系に関する研究成果の重要性が高まると考えられる。本論文では、反射・拡散に着目した受音系とその基本的な性質を明らかにすることを主題としたため、できるだけ単純な信号処理を適用し、実験結果を示してきた。今後、発展した信号処理技術と組み合わせることで、より高性能で音源情報解析を実現できるようになるだろう。

4本程度のマイクロホンを用いた收音技術が、TV電話やゲーム機を中心に普及しつつある。世の中で、音響アレイ信号処理が広く使われるようになれば、将来的にはアレイの多素子化が進む可能性は十分にあると思われる。本研究は、扱う波長幅が広い音響アレイ信号処理を用いて、どのように受音すべきかの指針を明らかにした先駆的研究であり、アレイの多素子化が求められる時代にこそ役立つ知見である。アレイの多素子化に対して訴求するために、遠くの音を收音するズームアップマイクを始めとして、様々な音響技術を提案していきたいと考えている。

謝辞

本研究を遂行するにあたって、数多くの方々に御指導と御援助を賜りました。特に御世話になった方々をここに記し、深い感謝の意を表します。

名古屋大学 武田一哉教授には、本論文の構成や内容について、丁寧な御教示や御指導を賜り、本研究の進むべき方向性を決定づけるアイデアを多数頂きました。名古屋大学 中岩浩巳客員教授には、異分野の立場から、本論文の構成に関して有益な意見を頂きました。東京大学 猿渡洋教授には、研究の本質に関わる御意見をいただき、博士論文を纏める上で参考とさせて頂きました。名古屋大学 北岡教英准教授には、本論文をまとめるにあたり、数々の有益な御意見を頂くとともに、数多くの御配慮を頂きました。三重大学 西野隆典准教授、名古屋大学 宮島千代美助教には、学部生時代より、研究の進め方について熱心にご指導頂きました。名古屋大学 江崎知様 (2012年博士課程前期卒) には、著者の研究の一部に関して、検討して頂きました。

NTT-AT アプリケーションソリューション事業本部 児島治彦ユニット長には、上長として本研究にご理解頂き、本研究遂行の機会を与えて頂きました。NTTメディアインテリジェンス研究所 稲垣博人所長には、本研究ご理解頂くと共に、温かい激励と熱心に議論させて頂きました。

電気通信大学 羽田陽一 教授には、本研究の全般に渡って御指導と御助言を頂きました。拡散受音の核となる考え方は、空間相関行列の固有ベクトルの物理的意味について深く考察したことから生み出されたと考えています。羽田教授との熱心な議論なくし

では、本研究は成りたち得なかったと考えています。大分大学 古家賢一教授には、拡散受音の原理を論理的に説明する上で、多大な御助言を賜りました。NTT メディアインテリジェンス研究所 大室伸グループリーダーには、数々の有益な御意見を頂くとともに、本研究遂行の機会を与えて頂きました。NTT サービスエボリューション研究所 岡本学博士、NTT メディアインテリジェンス研究所 阪内澄宇グループリーダーには、本研究を進める際の上司として、様々な御助言を頂きました。NTT メディアインテリジェンス研究所 小林和則博士には、遠くの音を收音するズームアップマイクの基本的なソフトウェア構築で数々の御援助を頂きました。また、アレイ信号処理の開発に関して数々の技術を指導して頂きました。カンタベリー大学 日岡裕輔博士には、アレイ信号処理に関して熱心に議論させて頂くと共に、論文の書き方について懇切丁寧にご指導頂きました。著者が困難に直面した際に相談する相手であり、多大な御指導、御助言を頂きました。NTT アドバンステクノロジー 関根弘美様には、本理論から派生したズームアップマイクのシステム構築において、数々の御支援を頂きました。NTT コミュニケーション科学基礎研究所 中谷智広博士、荒木章子博士には、拡散受音の研究アプローチに関して熱心に議論をさせて頂き、收音技術の最新動向に関する情報を提供して頂きました。

遠くの音をクリアに收音するズームアップマイクに関して、数々のメディア関係者の皆様にお世話になりました。また、個々の御名前は省略させて頂きませんが、NTT メディアインテリジェンス研究所 音声言語メディアプロジェクトの方々には数多くの御討論をさせて頂きました。国際会議・学会・研究会を通じて、数多くの諸先輩方から温かい激励やコメントを頂きました。本研究の遂行は、そのような方々の熱心な御支援により成し得たものです。未筆ながら、関係者の皆様方に感謝の意を表します。

参考文献

- [1] D. H. Johnson, D. E. Dudgeon, *Array processing: concepts and techniques*, Englewood Cliffs, NJ: Prentice-Hall, 1993.
- [2] H. L. Van Trees, *Optimum array processing*, Wiley (Part IV ed.), 2002.
- [3] 浅野 太, 音のアレイ信号処理 –音源の定位・追跡と分離–(第1版), コロナ社, 2011.
- [4] J. L. Flanagan, D. A. Berkley, G. W. Elko, J. E. West, and M. M. Sondhi, “Autodirective microphone systems,” *Acoustica*, vol. 73, no. 2, pp. 58–71, 1991.
- [5] 小林 和則, 羽田 陽一, 日和崎 祐介, 大室 伸, 入島 勉, “方向別 AGC 機能の IP 電話会議装置への実装,” 電子情報通信学会大会講演論文集, pp. 136, 2006.
- [6] 日和崎 祐介, 大室 伸, 小林 和則, 羽田 陽一, 入島 勉, 中山 圭一, 阿部 匡伸, “広帯域音声コーデック (UEMCLIP) の IP 電話会議装置への実装,” 電子情報通信学会総合大会講演論文集, , pp. 137, 2006.
- [7] Y. Hioka, K. Furuya, K. Kobayashi, S. Sakauchi, and Y. Haneda, “Angular region-wise speech enhancement for hands-free speakerphone,” *IEEE Transactions on Consumer Electronics*, vol. 58, no. 4, pp. 1403–1410, 2012.
- [8] 天野 文雄, “らくらくホンにおける聞こえ支援機能について,” 日本音響学会講演論文集 (春), pp. 1563–1564, 2012.

- [9] 鎌土 記良, 大沼 侑司, 猿渡 洋, 鹿野 清宏, “Kinect のマイクロホン・アレーによる音声信号処理”, *INTERFACE* 1月号, CQ 出版社, pp. 112–116, 2012.
- [10] 辻野 孝輔, 栄藤 稔, 磯田 佳徳, 飯塚 真也, “実サービスにおける音声認識と自然言語インタフェース技術,” *人工知能学会誌*, vol. 28, no. 1, pp. 75–81, 2013.
- [11] J. Capon, “High-resolution frequency-wavenumber spectrum analysis,” *Proceeding of IEEE*, vol. 57, pp. 1408–1418, 1969.
- [12] O. L. Frost III, “An algorithm for linearly constrained adaptive array processing,” *Proceeding of IEEE*, vol. 60, no. 8, pp. 926–935, 1972.
- [13] K. Takao, M. Fujita, and T. Nishi, “An adaptive antenna array under directional constraint,” *IEEE Transaction on Antennas & Propagation*, vol. AP-24, no. 5, pp. 662–669, 1976.
- [14] Y. Kaneda and J. Ohga, “Adaptive microphone-array system for noise reduction,” *IEEE Transaction on Acoustics, Speech, and Signal Processing*, vol. ASSP-34, pp. 1391–1400, 1986.
- [15] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent component analysis*, Wiley, 2001.
- [16] P. Smaragdis, “Blind separation of convolved mixtures in the frequency domain,” *Neurocomputing*, vol. 22, no. 1–3, pp. 21–34, 1998.
- [17] A. Belouchrani, K. Abed-Meraim, J. F. Cardoso, and E. Moulines, “A blind source separation technique using second-order statistics,” *IEEE Transaction on Signal Processing*, vol. 45, no. 2, pp. 434–443, 1997.

- [18] H. Sawada, S. Araki, R. Mukai and S. Makino, “Grouping separated frequency components with estimating propagation model parameters in frequency-domain blind source separation,” *IEEE Transaction on Audio, Speech, and Language Processing*, vol. 15, no. 5, pp. 1592–1604, 2007.
- [19] M. Aoki, M. Okamoto, S. Aoki, H. Matsui, T. Sakurai, and Y. Kaneda, “Sound source segregation based on estimating incident angle of each frequency component of input signals acquired by multiple microphones,” *Acoustical science and technology* vol. 22, no. 2, pp. 149–157, 2001.
- [20] O. Yilmaz and S. Rickard, “Blind separation of speech mixtures via time-frequency masking,” *IEEE Transaction on Signal Processing*, vol. 52, no. 7, pp. 1830–1847, 2004.
- [21] C. Marro, Y. Mahieux, K. U. Simmer, “Analysis of noise reduction and dereverberation techniques based on microphone arrays with postfiltering,” *IEEE Transaction on Speech and Audio Processing*, 6, 240–259, 1998.
- [22] R. Zelinski, “A microphone array with adaptive post-filtering for noise reduction in reverberant rooms,” *ICASSP 1988*, 5, 2578–2581, 1988.
- [23] I. A. McCowan, H. Bourslard, “Microphone array post-filter based on noise field coherence,” *IEEE Transaction on Audio, Speech, Language Processing*, 11, 709–716, 2003.
- [24] Y. Hioka, K. Furuya, K. Kobayashi, K. Niwa, Y. Haneda, “Underdetermined sound source separation using power spectrum density estimated by combination of directivity gain,” *IEEE Transaction on Audio, Speech, and Language Processing*, vol. 21, no. 6, pp. 1240–1250, 2013.

- [25] “AES10-2008: AES recommended practice for digital audio engineering – serial multichannel audio digital interface (MADI),” (revision of AES10-1991)
- [26] M. Puckette, “Pure data,” Proceedings, International Computer Music Conference, pp. 224–227, 1996.
- [27] 岡本 拓磨, “Pure Data によるリアルタイムマルチメディアプログラミング,” 日本音響学会誌, vol. 69, no. 6, pp. 285–290, 2013.
- [28] 伊勢 史郎, “キルヒホッフ-ヘルムホルツ積分方程式と逆システム理論に基づく音場制御の原理,” 日本音響学会誌, vol. 53, no. 9, pp. 706–713, 1997.
- [29] H. F. Silverman, W. R. Patterson, and J. L. Flanagan, “The huge microphone array,” *IEEE Concurrency*, vol. 6, Issue 4, pp. 36–46, 1998.
- [30] K. Hamasaki, T. Nishiguchi, R. Okumura, Y. Nakayama, and A. Ando, “A 22.2 multichannel sound system for ultrahigh-definition TV,” *SMPTE Motion Imaging Journal*, vol. 117, no. 3, pp. 40–49, 2008.
- [31] D. H. Cooper, T. Shiga, “Discrete-matrix multichannel stereo,” *Journal of Audio Engineering Society (AES)*, vol. 20, no. 5, pp. 346–360, 1972.
- [32] D. B. Ward, T. D. Abhayapala, “Reproduction of a plane-wave sound field using an array of loudspeakers,” *IEEE Transaction on Speech and Audio Processing*, vol. 9, no. 6, pp. 697–707, 2001.
- [33] M. Poletti, “Three-dimensional surround sound systems based on spherical harmonics,” *Journal of Audio Engineering Society*, vol. 53, no. 11, pp. 1004–1025, 2005.

- [34] J. Ahrens and S. Spors, “An analytical approach to sound field reproduction using circular and spherical loudspeaker distribution,” *ACTA Acoustica United with Acustica*, vol. 94, pp. 988–999, 2008.
- [35] M. Camras, “Approach to recreating a sound field,” *Journal of Acoustic Society of America*, vol. 43, no. 6, pp. 1425–1431, 1968.
- [36] S. Koyama, K. Furuya, Y. Hiwasaki, and Y. Haneda, “Analytical approach to wave field reconstruction filtering in spatio-temporal frequency domain,” *IEEE Transaction on Audio, Speech, and Language Processing*, vol. 21, Issue 4, pp. 685–696, 2012.
- [37] 安藤 彰男, “高臨場感音響技術とその理論,” 電子情報通信学会 Fundamental Review, vol. 3, no. 4, pp. 33–46, 2010.
- [38] J. Meyer, and G. W. Elko, “A spherical microphone array for spatial sound recordings,” *Journal of Acoustical Society of America*, vol. 111, Issue 5, 2346, 2002.
- [39] T. D. Abhayapala, and D. B. Ward, “Theory and design of high order sound field microphones using spherical microphone array,” *ICASSP 2002*, vol. II, pp. 1949–1952, 2002.
- [40] B. Rafaely, “Plane-wave decomposition of the sound field on a sphere by spherical convolution,” *Journal of Acoustical Society of America*, vol. 116, Issue 4, pp. 2149–2157, 2004.
- [41] B. Rafaely, “Open–sphere designs for spherical microphone arrays,” *IEEE Trans. Audio, Speech, and Language Process.*, vol. 15, no. 2, pp. 727–732, 2007.
- [42] M. Miyoshi and Y. Kaneda, “Inverse filtering of room acoustics,” *IEEE Transaction on Speech and Audio Processing*, vol. 36, no. 2, pp. 145–152, 1988.

- [43] P. Zahorik, D. S. Brungrat, and A. W. Bronkhorst, “Auditory distance perception in humans: A summary of past and present research,” *Acta Acustica united with Acustica*, vol. 91, no. 3, pp. 409–420, 2005.
- [44] J. Jo and M. Koyasu, “Measurement of reverberation time based on the direct–reverberant sound energy ratio in steady state,” *Inter-noise 1975*, pp. 579–582, 1975.
- [45] G. J. Foschini, “Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas,” *Bell Labs Technical Journal*, vol. 1, no. 2, pp. 41–59, 1996.
- [46] G. J. Foschini and M. J. Gans, “On limits of wireless communications in a fading environment when using multi-element antennas,” *Wireless Personal Communications*, vol. 6, no. 3, pp. 311–335, 1998.
- [47] 今井 秀樹, 情報理論, 昭晃堂, 2000.
- [48] 府川 和彦, 電子情報通信学会 知識ベース (4群1編 無線通信基礎 7章 MIMO 伝送 7.1節 通信路容量), pp. 1-5, 2010.
- [49] G. H. Hardy, J. E. Littlewood, G. Polya, 細川 尋史 (翻訳), 不等式, 丸善出版, 2012.
- [50] K. Petersen and M. Pedersen, “The matrix cookbook,” 2007.
- [51] R. K. Cook, R. V. Waterhouse, R. D. Berendt, S. Edelman, and M. C. Thompson, Jr., “Measurement of correlation coefficients in reverberant sound fields,” *J. Acoust. Soc. Am.*, vol. 27, Issue 6, pp. 1072–1077, 1955.
- [52] M. Tohyama, H. Suzuki, Y. Ando, *The nature and technology of acoustic space*, Academic Press, 1995.

- [53] J. B. Allen and D. A. Berkley, “Image method for efficiently simulating small-room acoustics,” *Journal of Acoustical Society of America*, vol. 65, no. 4, pp. 943–950, 1979.
- [54] H. Kuttruff, *Room Acoustics (5th edition)*, CRC Press, 2009
- [55] H. Sawada, R. Mukai, S. Araki, S. Makino, “Estimating the number of sources using independent component analysis,” *Acoustic Science and Technology*, vol. 26, no. 5, pp. 450–452, 2005.
- [56] A. Bell and T. Sejnowski, “An information-maximization approach to blind separation and blind deconvolution,” *Neural Computation*, vol. 7, pp. 1129–1159, 1995.
- [57] S. Choi, S. Amari, A. Cichocki, and R. Liu, “Natural gradient learning with a non-holonomic constraint for blind deconvolution of multiple channels,” *International Workshop on ICA and BSS*, pp. 371–376, 1999.
- [58] S. Kurita, H. Saruwatari, S. Kajita, K. Takeda, and F. Itakura, “Evaluation of blind signal separation method using directivity pattern under reverberant conditions,” *ICASSP 2000*, pp. 3140–3143, 2000.
- [59] A. Araki, H. Sawada, and S. Makino, “Blind speech separation in a meeting situation with maximum SNR beamformers,” *ICASSP*, pp. 41–44, 2007.
- [60] K. Matsuoka and S. Nakashima, “Minimal distortion principle for blind source separation,” *ICA 2001*, pp. 722–727, 2001.
- [61] N. Murata, S. Ikeda, and A. Ziehe, “An approach to blind source separation based on temporal structure of speech signals,” *Neurocomputing*, vol. 41, no. 1–4, pp. 1–24, 2001.

- [62] T. Takatani, T. Nishikawa, H. Saruwatari, and K. Shikano, “High-fidelity blind separation of acoustic signals using SIMO-model-based independent component analysis,” *IEICE Transaction on Fundamentals*, vol. E87-A, no. 8, pp. 2063–2072, 2004.
- [63] M. Ogasawara, T. Nishino, K. Takeda, “A small dodecahedral microphone array for blind source separation,” in *Proceeding of ICASSP 2010*, pp. 229–232, 2010.

研究業績

主な研究業績

受賞

1. 日本音響学会 栗屋潔学術奨励賞 (2010)
2. 日本音響学会 ポスター賞 (2008)
3. IEEE International Workshop on Multimedia Signal Processing 2008 Best Paper Award (2008)
4. 情報学ワークショップ (WiNF2007) 優秀論文賞 (2007)

論文

1. 丹羽 健太, 江崎 知, 日岡 裕輔, 西野 隆典, 武田 一哉, “空間相関行列の固有値分布に着目した音源別距離推定,” 電子情報通信学会論文誌 A, Vol. J97-A, No. 2, pp. 68–76, 2014.
2. K. Niwa, Y. Hioka, K. Furuya, and Y. Haneda, “Diffused sensing for sharp directive beamforming,” *IEEE Transactions on Audio, Speech and Language Processing*, Vol. 21, Issue 11, pp. 2346–2355, 2013.
3. K. Niwa, Y. Hioka, S. Sakauchi, K. Furuya, and Y. Haneda, “An estimation

- method of sound source orientation using eigenspace variation of spatial correlation matrix,” *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, Vol. E96–A, No. 9, pp. 1831–1839, 2013.
4. K. Niwa, Y. Hioka, S. Sakauchi, K. Furuya, and Y. Haneda, “Sharp directive beamforming using microphone array and planar reflector,” *Acoustical Science and Technology*, Vol. 34, No. 4, pp. 253–262, 2013.
 5. Y. Hioka, K. Furuya, K. Kobayashi, K. Niwa, and Y. Haneda, “Underdetermined sound source separation using power spectrum density estimated by combination of directivity gain,” *IEEE Transactions on Audio, Speech and Language Processing*, Vol. 21, No. 6, pp. 1240–1250, 2013.
 6. Y. Hioka, K. Niwa, S. Sakauchi, K. Furuya, and Y. Haneda, “Estimating direct-to-reverberant energy ratio using D/R spatial correlation matrix model,” *IEEE Transactions on Audio, Speech and Language Processing*, Vol. 19, Issue. 8, pp. 2374–2384, 2011.
 7. K. Niwa, T. Nishino, and K. Takeda, “Blind source separation of musical signals applied to selectable–listening–point audio reconstruction,” *IEICE Transactions on Information and Systems*, Vol. E92–D, No. 3, pp.469–476, 2009.

本

1. T. Nishino, M. Ogasawara, K. Niwa, and K. Takeda, “Selective listening point audio based on blind signal separation and 3D audio effect,” Principles and applications of spatial hearing, (IWPASH 2009), Y. Suzuki, D. Brungart, Y. Iwaya, K. Iida, D. Cabrera, H. Kato edtions, pp. 277–291, World Scientific, ISBN: 978-981-4313-87-2, 2011.

レター

1. 福嶋 慶繁, 丹羽 健太, 圓道 知博, 藤井 俊彰, 谷本正幸, 西野 隆典, 武田 一哉, “多視点・多聴点データ取得システムを用いた自由視聴点映像生成,” 電子情報通信学会論文誌 D, Vol. J91-D, No. 8, pp. 2039–2041, 2008.

査読付国際会議

1. K. Niwa, Y. Hioka, K. Kobayashi, K. Furuya, and Y. Haneda, “Evaluation of microphone array based on diffused sensing with various filter design methods”, Proceedings of European Signal Processing Conference (EUSIPCO 2013)
2. Y. Hioka, K. Furuya, K. Niwa, and Y. Haneda, “Estimation of direct-to-reverberation energy ratio based on isotropic and homogeneous propagation model,” Proceedings of International Workshop on Acoustic Echo and Noise Control (IWAENC 2012), pp. 4–6, 2012.
3. K. Niwa, S. Sakauchi, K. Furuya, M. Okamoto, and Y. Haneda, “Diffused sensing for sharp directivity microphone array”, Proceedings of International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2012), pp. 225–228, 2012.
4. K. Niwa, Y. Hioka, K. Furuya, and Y. Haneda, “Telescopic microphone array using reflector for segregating target source from noises in same direction,” Proceedings of International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2012), pp. 5457–5460, 2012.
5. S. Esaki, K. Niwa, T. Nishino, and K. Takeda, “Estimating sound source depth using a small-size array,” Proceedings of International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2012), pp. 401–404, 2012.
6. K. Niwa, S. Sakauchi, K. Furuya, M. Okamoto, and Y. Haneda, “Superdirective

- beamforming using microphone array and single reflector,” Proceedings of International Conference on Consumer Electronics (ICCE 2012), pp. 17–18, 2012.
7. Y. Hioka, K. Niwa, K. Furuya, and Y. Haneda, “Estimating direct-to-reverberant energy ratio based on spatial correlation model segregating direct sound and reverberation,” Proceedings of International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2010), pp. 149–152, 2010.
 8. K. Niwa, Y. Hioka, S. Sakauchi, K. Furuya, and Y. Haneda, “Estimation of sound source orientation using eigenspace of spatial correlation matrix,” Proceedings of International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2010), pp. 129–132, 2010.
 9. M. P. Tehrani, K. Niwa, N. Fukushima, Y. Hirano, T. Fujii, M. Tanimoto, K. Takeda, K. Mase, A. Ishikawa, S. Sakazawa and A. Koike, “3DAV integrated system featuring arbitrary listening-point and viewpoint generation,” International Workshop on Multimedia Signal Processing (MMSP 2008), Vol. PID-213, 2008.
 10. K. Niwa, T. Nishino, and K. Takeda, “Encoding large array signals into a 3D sound field representation for selective listening point audio based on blind source separation,” Proceedings of International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2010), pp. 181–184, 2008.
 11. K. Niwa, T. Nishino, and K. Takeda, “Development of selectable viewpoint and listening point system for musical performance,” Proceedings of the 19-th International Congress on Acoustics (ICA 2007), No. PPA-06-011, 2007.

招待講演

1. 丹羽 健太, “ランダム行列で音を聞き分ける,” 日本物理学会秋季大会シンポジウム講演 (21pAF-7), 2012.

技術報告

1. 江崎 知, 丹羽 健太, 西野 隆典, 北岡 教英, 武田 一哉, “反射音の到来方向分布を用いた音源距離推定,” 電子情報通信学会技術研究報告 (EA), Vol. 111, No. 306, pp. 19-24, 2011.
2. 丹羽 健太, 阪内 澄宇, 古家 賢一, 岡本 学, 羽田 陽一, “反射音を利用したマイクロホンアレーによる距離別收音技術,” 電子情報通信学会技術研究報告 (EA), Vol. 111, No. 89, pp. 31-36, 2011.
3. 丹羽 健太, 阪内 澄宇, 古家 賢一, 岡本 学, 羽田 陽一, “多重反射を利用した超指向性收音技術の検討,” 電子情報通信学会技術研究報告 (EA), Vol. 111, No. 89, pp. 25-30, 2011.
4. 丹羽 健太, 西野 隆典, 武田 一哉, “受聴位置選択型音場再現のためのブラインド音源分離を用いた多マイクロホンアレー信号の符号化,” 電子情報通信学会技術研究報告 (EA), Vol. 107, No. 532, pp. 49-54, 2008.
5. 丹羽 健太, 西野 隆典, 武田 一哉, “ブラインド音源分離と頭部伝達関数を用いた音楽演奏を題材とした自由視聴点コンテンツの制作,” 電子情報通信学会技術研究報告 (EA), Vol. 107, No. 120, pp. 31-36, 2007.

学会発表

1. 丹羽 健太, 日岡 裕輔, 小林 和則, 鎌土記良, “雑音下での音声認識率向上を目的としたマイクロホンアレーの実装,” 日本音響学会春季講演論文集, pp. 717-718, 2014.
2. 丹羽 健太, 日岡 裕輔, 小林 和則, 古家 賢一, 羽田 陽一, “伝達特性間の独立性を用いた拡散センシング型アレーの構造評価,” 日本音響学会春季講演論文集, pp. 745-746, 2013.
3. 丹羽 健太, 日岡 裕輔, 古家 賢一, 羽田 陽一, “拡散センシングに基づく指向制御 ~

- 雑音出力パワーに対するフィルタ長の影響～,” 日本音響学会秋季講演論文集, pp. 645–646, 2012.
4. 丹羽 健太, 日岡 裕輔, 荒木 章子, 古家 賢一, 羽田 陽一, “最大 SN 比法への拡散センシングの適用,” 日本音響学会秋季講演論文集, pp. 761–762, 2012.
 5. 江崎 知, 丹羽 健太, 西野 隆典, 武田 一哉, “空間相関行列を用いた音源距離推定,” 日本音響学会春季講演論文集, pp. 759–760, 2012.
 6. 日岡 裕輔, 古家 賢一, 丹羽 健太, 羽田 陽一, “残響音の等方到来を仮定した直間比推定,” 日本音響学会春季講演論文集, pp. 827–828, 2011.
 7. 丹羽 健太, 阪内 澄宇, 古家 賢一, 岡本 学, 羽田 陽一, “大型多重反射板付きマイクロホンアレーを用いた超指向性收音,” 日本音響学会秋季講演論文集, pp. 611–612, 2011.
 8. 岡本 拓磨, 岩谷 幸雄, 鈴木 陽一, 丹羽 健太, 阪内 澄宇, 羽田 陽一, “多様な室内残響下における音源放射指向特性と残響成分の分析,” 日本音響学会春季講演論文集, pp. 911–912, 2011.
 9. 丹羽 健太, 阪内 澄宇, 古家 賢一, 羽田 陽一, “狭指向性を実現する反射板利用型マイクロホンアレー,” 日本音響学会春季講演論文集, pp. 659–660, 2011.
 10. 丹羽 健太, 阪内 澄宇, 古家 賢一, 羽田 陽一, “最小分散ビームフォーマを用いた空間周波数解析による反射音情報推定,” 日本音響学会秋季講演論文集, pp. 691–692, 2010.
 11. 丹羽 健太, 日岡 裕輔, 阪内 澄宇, 古家 賢一, 羽田 陽一, “二次元格子状マイクロホンアレーを用いた反射音情報の推定,” 日本音響学会春季講演論文集, pp. 735–736, 2010.
 12. 丹羽 健太, 日岡 裕輔, 阪内 澄宇, 古家 賢一, 羽田 陽一, “音源向きと空間相関行列の固有空間の関連,” 日本音響学会秋季講演論文集, pp. 789–790, 2009.
 13. 日岡 裕輔, 丹羽 健太, 阪内 澄宇, 古家 賢一, 羽田 陽一, “受信信号の直間比に基づ

- く距離別收音の検討,” 日本音響学会秋季講演論文集, pp. 633–634, 2009.
14. 丹羽 健太, 阪内 澄宇, 古家 賢一, 岡本 学, 羽田 陽一, “空間相関行列の固有値の比率に着目した発話者向きの推定,” 日本音響学会春季講演論文集, pp. 775–776, 2009.
 15. 丹羽 健太, 西野 隆典, 武田 一哉, “自由視点 TV における受聴位置選択型オーディオのための音響空間の符号化法,” 日本音響学会春季講演論文集, pp. 805–806, 2008.
 16. 福嶋 慶繁, 丹羽 健太, 圓道 知博, 藤井 俊彰, 谷本 正幸, 西野 隆典, 武田 一哉, “多視点・多聴点データからの自由視聴点映像生成,” 第 12 回映像メディア処理シンポジウム講演論文集 (I-4.14), pp. 95–96, 2007.
 17. 丹羽 健太, 西野 隆典, 武田 一哉, “自由視聴点 TV のための自由聴点音場の構成方法,” 情報学ワークショップ (WiNF2007), pp. 143–146, 2007.
 18. 丹羽 健太, 西野 隆典, 武田 一哉, “自由視点・聴点システムのためのコンテンツ制作,” 日本音響学会春季講演論文集, pp. 675–676, 2007.
 19. K. Niwa, T. Nishino, C. Miyajima, and K. Takeda, “Blind signal separation of musical signals applied to selectable-listening-point audio reconstruction,” 4-th Joint Meeting of ASA and ASJ (3pSP20), 2006.
 20. 丹羽 健太, 西野 隆典, 宮島 千代美, 武田 一哉, “楽器音の音源信号分離と HRTF を用いた音像の再配置の検討,” 日本音響学会秋季講演論文集, pp. 525–526, 2006.

メディア報道

(遠くの音をクリアに收音するズームアップマイクに関して)

1. NHK 教育テレビ, “大!天才テレビくん,” (2012.9.26 放送)
2. テレビ東京, “所さんの学校では教えてくれないそこんトコロ,” (2012.2.17 放送)
3. TBS ラジオ, “森本毅郎のスタンバイ!!”, (2011.11.8 放送)

4. 日経ビジネス, “技術&トレンド,” (2011.10.21 発売号)
5. 電経新聞, (2011.9.5 発売号)
6. 日本テレビ, “不可思議探偵団!!,” (2011.8.29 放送)
7. 中国中央テレビ, “第一時間,” (2011.7.16 放送)
8. 毎日コミュニケーションズ, “PCFan,” (2011.5.24 発売号)
9. テレビ朝日, “さきっちょ,” (2011.5.3 放送)
10. 放送ジャーナル社, “放送ジャーナル,” (2011.4.15 発売号)
11. テレビ東京, “ワールドビジネスサテライト (トレンドたまご),” (2011.2.21 放送)
12. フジテレビ, “とくダネ (オープニングトーク),” (2011.2.17 放送)
13. フジサンケイビジネスアイ, (2011.2.17 発売号)