

高次相関を用いた非線形 MUSIC による高分解能方位推定*

杉本侑哉, 宮部滋樹, 山田武志, 牧野昭二 (筑波大)

1 はじめに

音源信号の存在する空間を分析するマイクロホンアレー信号処理において, 音源の方位推定は, 事前情報に音源の方位情報を要する音源分離や雑音抑圧など, 幅広い分野で求められ, 様々な研究が行われている. 空間共分散行列の高位主成分ベクトルとステアリングベクトルの相似性を利用した MUSIC (Multiple Signal Classification) [1] は, 高分解能に方位推定を行う手法として知られている. しかし, MUSIC は共分散行列の次元数に拘束される. N 個の音源の方位を推定するためには, $M (> N)$ 個の素子を用いる必要があり, また N が M に近付くと推定精度が低下する. 従って, 多数の音源の方位を推定する場合, アレーの大規模化が避けられないという問題がある. このような背景から, 統計量として高次キュムラントを用い, 次元数を増やした分析へと MUSIC を拡張するいくつかの手法が提案されてきた. 例えば, 統計量に 4 次キュムラントを用いる 4-MUSIC [2] や, 4-MUSIC をさらに高次元に拡張する $2q$ -MUSIC [3] が知られている. これらの手法は, 音源数が素子数よりも多い劣決定条件の下での方位推定が可能であり, 同時に方位推定の分解能を向上させる.

本稿では, 4-MUSIC などとは異なったアプローチで劣決定条件に MUSIC を拡張する手法を提案する. パターン認識の分野では, カーネル関数を用いた高次元ヒルベルト空間への非線形写像により, 識別器の線形識別性を高めるカーネル法が知られている. また, カーネル法を用いて観測信号を高次元化することで, ビームフォーミングの性能を高める手法 [4] が提案されている. カーネル法により処理の次元数を増やすアプローチは, 共分散行列の次元数に拘束を受ける MUSIC にとって極めて有効であると考えられる. そこで我々は, カーネル法を用いて高次元ヒルベルト空間で MUSIC を行うカーネル MUSIC を提案し, 定式化を行う. どのような条件のカーネル関数が MUSIC の分析に適しているかを定め, その中で最も単純な二次カーネルに注目する. 統計的・幾何学的な視点からの議論を通じて, 二次カーネルを用いた二次カーネル MUSIC が 4 次モーメントの分析であることと, $(M^2 - 1)$ までの音源を高分解能に方位推定可能であることを示す. さらに, 観測信号の時間長が長い場合に計算が効率的な, 二次カーネルの与える非線形写像を解析的に表現して用いたエルミート MUSIC を提案し, 従来手法に比べてより少ない計算量で方位推定が可能であることを示す. 実験により, 音源の方位推定性能が従来手法とほぼ同等で, 残響に対してやや頑健であることを確認する.

2 問題設定

本稿を通して信号は周波数領域で表されることとする. 観測信号は次のようにモデル化できる.

$$\begin{aligned} \mathbf{x}(\omega, t) &= [x_1(\omega, t), \dots, x_M(\omega, t)]^T \\ &= \sum_{i=1}^N \mathbf{a}_i(\omega) s_i(\omega, t) + \mathbf{n}(\omega, t), \end{aligned} \quad (1)$$

$$\mathbf{n}(\omega, t) = [n_1(\omega, t), \dots, n_M(\omega, t)]^T, \quad (2)$$

$$\mathbf{a}_i(\omega) = [a_{1,i}(\omega), \dots, a_{M,i}(\omega)]^T. \quad (3)$$

ここで, $t = 1, \dots, L$ はある処理区間における時間フレームインデックス, ω は周波数, M は素子数, N は音源数, $s_i(\omega, t)$ は i 番目の音源信号の複素振幅, $x_j(\omega, t)$ は j 番目の素子における観測信号の複素振幅, $n_j(\omega, t)$ は j 番目の素子における観測雑音の複素振幅, $a_{j,i}(\omega)$ は i 番目の音源から j 番目のマイクロホンへの伝達関数を表す. $\mathbf{a}_i(\omega)$ は, 各音源の到来方位を表現するベクトルであることから, 本稿ではこれを i 番目の音源の方向ベクトルと呼ぶ.

式 (1) に示される信号モデルにおいて, 各素子では複数の音源から到来する信号と, 雑音が線形に混ざり合って観測される. 方位推定技術の目標は, このような混合信号から各音源の方向ベクトル $\mathbf{a}_1(\omega), \dots, \mathbf{a}_N(\omega)$ を個別に同定することである. 従って, 本稿での問題設定を観測信号ベクトル $\mathbf{x}(\omega, t)$ を用いて, 真の音源方位に対応する方向ベクトルと類似のステアリングベクトルを推定することと定める.

3 従来手法

3.1 MUSIC

ここでは, MUSIC について述べる. MUSIC は, 空間共分散行列の上位主成分ベクトルとステアリングベクトルの相似性を利用し, マイクロホンアレーで観測された混合信号から各音源の方位推定を行う手法である. マイクロホンアレーの素子数が M の場合, $(M - 1)$ までの音源方位を高い分解能で推定できる. 前節でモデル化された観測系において, $\mathbf{n}(\omega, t)$ の各要素が互いに無相関であると仮定すると, 観測信号の共分散行列は以下のように表される.

$$\begin{aligned} \mathbf{R}(\omega) &= E[\mathbf{x}(\omega, t)\mathbf{x}^H(\omega, t)] \\ &= \sum_{i=1}^N P_i(\omega) \mathbf{a}_i(\omega) \mathbf{a}_i^H(\omega) + \sigma^2(\omega) \mathbf{I}_M. \end{aligned} \quad (4)$$

ここで, $P_i(\omega) = E[|s_i(\omega, t)|^2]$ は, i 番目の音源の信号パワー, $\sigma^2(\omega)$ は雑音パワー, \mathbf{I}_j は j 次元の単位行列を表す. 共分散行列 $\mathbf{R}(\omega)$ の固有値分解により, 次式を得る.

$$\mathbf{R}(\omega) = \mathbf{V}(\omega) \mathbf{E}(\omega) \mathbf{V}^H(\omega), \quad (5)$$

$$\mathbf{V}(\omega) = [\mathbf{v}_1(\omega), \dots, \mathbf{v}_M(\omega)], \quad \mathbf{V}^H(\omega) \mathbf{V}(\omega) = \mathbf{I}_M, \quad (6)$$

$$\begin{aligned} \mathbf{E}(\omega) &= \text{diag}[e_1(\omega), \dots, e_M(\omega)] + \sigma^2(\omega) \mathbf{I}_M, \\ e_1(\omega) &\geq \dots \geq e_N(\omega), \\ e_{N+1}(\omega), \dots, e_M(\omega) &= 0. \end{aligned} \quad (7)$$

ここで, $e_1(\omega), \dots, e_M(\omega)$ は $\mathbf{R}(\omega)$ の固有値, $\mathbf{v}_1(\omega), \dots, \mathbf{v}_M(\omega)$ は各固有値に対応する固有ベクトルとし, $\text{diag}[\dots]$ は引数の要素からなる対角行列を表す. 音源信号に対応する固有ベクトル

*High-resolution DOA estimation with non-linear MUSIC based on higher-order covariance. by Yuya SUGIMOTO, Shigeki MIYABE, Takeshi YAMADA, and Shoji MAKINO. (University of Tsukuba)

$\mathbf{v}_1(\omega), \dots, \mathbf{v}_N(\omega)$ によって張られる部分空間は、各音源の方向ベクトル $\mathbf{a}_1(\omega), \dots, \mathbf{a}_N(\omega)$ が張る部分空間と一致する。この部分空間を信号部分空間と呼び、 $S(\omega)$ と表す。

$$S(\omega) \triangleq \text{span}[\mathbf{a}_i(\omega)]_{i=1}^N = \text{span}[\mathbf{v}_i(\omega)]_{i=1}^N. \quad (8)$$

さらに、 $\text{span}[\mathbf{v}_i(\omega)]_{i=1}^M$ において信号部分空間の直交補空間を雑音部分空間と呼ぶと、雑音部分空間の基底ベクトル $\mathbf{v}_{N+1}(\omega), \dots, \mathbf{v}_M(\omega)$ と各音源の方向ベクトルの間には以下の関係が成り立つ。

$$\begin{aligned} \mathbf{a}_i^H(\omega) \mathbf{v}_j(\omega) &= 0, \\ (i, j) &\in \{i = 1, \dots, N, j = N + 1, \dots, M\}. \end{aligned} \quad (9)$$

式 (9) の直交性を利用して、次式に示す MUSIC スコア $f(\omega; \theta)$ を考える。

$$f(\omega; \theta) \triangleq \frac{1}{\sum_{i=N+1}^M |\mathbf{b}^H(\omega; \theta) \mathbf{v}_i(\omega)|}. \quad (10)$$

ここで、 $\mathbf{b}(\omega; \theta)$ ($\|\mathbf{b}(\omega; \theta)\| = 1$) は方向 θ のステアリングベクトルを表す。この時、 $i = 1, \dots, N$ いずれかについて $\mathbf{b}^H(\omega; \theta) \mathbf{a}_i(\omega) = \|\mathbf{a}_i(\omega)\|$ に近づく θ に対して、 $f(\omega; \theta)$ は大きな値を取る。MUSIC は、この性質を利用して、 $f(\omega; \theta)$ が大きな値を取る方向を音源方向として推定する手法である。なお、式 (10) によって求まるのはある狭帯域での方位推定結果であるが、所望の周波数帯域における単一の推定結果は、各周波数ビンでの推定結果を周波数平均することで求められる。

MUSIC の注意すべき性質として、式 (4) の第二項が単位行列の定数倍でない時に固有ベクトルの値が変わることから、各素子の観測雑音に相関がある場合推定性能が低下する。また、雑音部分空間を表現するために少なくとも一次元の部分空間が必要となるため、推定する音源数 N は素子数 M 未満に拘束される。高分解能な方位推定を行うためには雑音部分空間の高い表現能力が求められ、 M は N より十分大きくなくてはならない。

3.2 4-MUSIC

4-MUSIC は、高次統計量として 4 次キュムラントを用いた音源の方位推定手法である。マイクロホンアレーの素子数が M の場合、音源数 ($M^2 - 1$) までの方向推定が可能であり、方位推定の分解能も MUSIC より高い。4-MUSIC では、各要素が 4 次キュムラントよりなる $M^2 \times M^2$ 行列 $\mathbf{C}(\omega)$ を固有値分解する。4 次キュムラントは式 (11) のように計算される。

$$\begin{aligned} \text{Cum}[X_i, X_j, X_k, X_l] &= \\ E[X_i X_j \overline{X_k} \overline{X_l}] - E[X_i X_j] E[\overline{X_k} \overline{X_l}] \\ &\quad - E[X_i \overline{X_k}] E[X_j \overline{X_l}] - E[X_i \overline{X_l}] E[X_j \overline{X_k}]. \end{aligned} \quad (11)$$

方位推定は固有ベクトルの分析によって MUSIC と同様の手順で行われる。

4 提案手法

パターン認識の分野では、カーネル関数を用いた高次元のヒルベルト空間への写像により、識別器の線形識別性を高めるカーネル法が知られている。本節ではカーネル法を MUSIC に応用したカーネル MUSIC を提案する。雑音部分空間の次元数を増やすことによ

り方位推定の分解能が向上し、また推定する音源数が素子数よりも多い劣決定条件でも方位推定が可能である。写像される空間の次元はカーネルの選択に依存し、無限次元への写像も可能であるが、本稿ではパワー推定型のビームフォーマ [4] で用いられるもっとも単純な二次カーネルによる M^2 次元空間への写像に焦点を当てて、幾何学的・統計的な視点から議論する。さらに、二次カーネルを用いたカーネル MUSIC と等価で、 $M^2 < L$ (L はある処理区間における時間フレームインデックスの総数) の場合に計算が効率的なエルミート MUSIC アルゴリズムを提案する。

4.1 一般化カーネル MUSIC の定式化

正定値カーネル関数: $\mathbb{C}^M \times \mathbb{C}^M \rightarrow \mathbb{C}$ は、カーネル関数が決定する写像 $\phi: \mathbb{C}^M \rightarrow \mathcal{H}$ で与えられるヒルベルト空間 \mathcal{H} における内積となることが知られている。

$$k(\mathbf{x}, \mathbf{y}) = \langle \phi(\mathbf{x}), \phi(\mathbf{y}) \rangle_{\mathcal{H}}. \quad (12)$$

ここで、 $\mathbf{x}, \mathbf{y} \in \mathbb{C}^M$ 、 $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ はヒルベルト空間 \mathcal{H} における内積を表す。非線形なカーネル関数により与えられるヒルベルト空間 \mathcal{H} は一般に \mathbb{C}^M よりも高次元であるため、カーネル関数は識別問題において特徴ベクトルの線形識別性を高めるために頻繁に用いられる。MUSIC における共分散行列 $\mathbf{R}(\omega) = E[\mathbf{x}(\omega, t) \mathbf{x}^H(\omega, t)]$ の固有値問題をヒルベルト空間 \mathcal{H} における相関線形作用素 $\Gamma_{\omega}: \mathcal{H} \rightarrow \mathcal{H}$ 、 $\Gamma_{\omega} = E[\langle \cdot, \phi(\mathbf{x}(\omega, t)) \rangle_{\mathcal{H}} \langle \phi(\mathbf{x}(\omega, t)), \cdot \rangle_{\mathcal{H}}]$ の固有値問題に、ステアリングベクトル $\mathbf{b}(\omega; \theta)$ を写像 $\phi(\mathbf{b}(\omega; \theta))$ に置き換えることにより、高次元写像を用いて分解能を高めたカーネル MUSIC が実現される。

一般に写像関数 ϕ は解析的に記述できないため、相関線形作用素 Γ_{ω} もまた解析的に記述できない。しかし、 Γ_{ω} の固有値問題

$$\begin{aligned} \Gamma_{\omega}(\mathbf{v}_i(\omega)) &= \lambda_i(\omega) \mathbf{v}_i(\omega), \\ \langle \mathbf{v}_i(\omega), \mathbf{v}_j(\omega) \rangle_{\mathcal{H}} &= \delta_{ij}, \quad \lambda_1(\omega) \geq \lambda_2(\omega) \geq \dots \geq 0. \end{aligned} \quad (13)$$

(δ_{ij} は、クロネッカ のデルタ) における固有値 $\lambda_i(\omega)$ と固有ベクトル $\mathbf{v}_i(\omega) \in \mathcal{H}$ を、観測信号 $\mathbf{x}(\omega, t)$ 、 $t = 1, \dots, L$ の $L \times L$ グラム行列 $\mathbf{K}(\omega)$ の固有値分解を用いて以下のように記述する方法がカーネル PCA として提案されている [5]。

$$\lambda_i(\omega) = d_i(\omega)/L, \quad (14)$$

$$\mathbf{v}_i(\omega) = \frac{1}{\sqrt{d_i(\omega)}} \sum_{j=1}^L u_{ji}(\omega) \phi(\mathbf{x}(\omega, j)), \quad (15)$$

$$\begin{aligned} \mathbf{K}(\omega) &= [k(\mathbf{x}(\omega, i), \mathbf{x}(\omega, j))]_{ij} \\ &= \mathbf{U}(\omega) \mathbf{D}(\omega) \mathbf{U}^H(\omega), \end{aligned} \quad (16)$$

$$\mathbf{U}(\omega) = [u_{ji}]_{ji}, \quad \mathbf{U}^H(\omega) \mathbf{U}(\omega) = \mathbf{I}_L, \quad (17)$$

$$\begin{aligned} \mathbf{D}(\omega) &= \text{diag}[d_1(\omega), \dots, d_L(\omega)], \\ d_1(\omega) &\geq \dots \geq d_L(\omega) \geq 0. \end{aligned} \quad (18)$$

ここで、 $[x]_{ij}$ は第 i 行 j 列要素が x である行列を表す。固有値 $\lambda_i(\omega)$ の比零の値の数を $r(\omega)$ とすると、 $\text{span}[\mathbf{v}_i(\omega)]_{i=N+1}^{r(\omega)}$ が雑音部分空間に相当するため、ヒルベルト空間 \mathcal{H} における MUSIC スコア $g_{\mathcal{H}}(\omega)$ は以下のように求められる。

$$g_{\mathcal{H}}(\omega) \triangleq \frac{1}{\sum_{i=N+1}^{r(\omega)} |\langle \phi(\mathbf{b}(\omega; \theta)), \mathbf{v}_i(\omega) \rangle_{\mathcal{H}}|}$$

$$= \frac{1}{\sum_{i=N+1}^r(\omega) \left| \sum_{j=1}^L \frac{u_{ji}(\omega)}{\sqrt{d_i(\omega)}} k(\mathbf{b}(\omega; \theta), \mathbf{x}(\omega, j)) \right|} \quad (19)$$

4.2 カーネル関数の適切な選択

カーネル MUSIC はヒルベルト空間 \mathcal{H} における相関の分析であり、写像 ϕ の選択によってその性質が大幅に変化する。そのためカーネル関数を適切に選ぶ必要がある。まず信号の写像は、1) ノルムの大小関係が維持され、 $\|\mathbf{x}\| > \|\mathbf{y}\|$ のときに $\|\phi(\mathbf{x})\|_{\mathcal{H}} > \|\phi(\mathbf{y})\|_{\mathcal{H}}$ ($\|\cdot\|_{\mathcal{H}}$ はヒルベルト空間におけるノルム) となり、2) 写像が原点を移動させず、 $\phi(0)$ が \mathcal{H} における原点に一致し、かつ 3) 信号間の直交性が維持されて $\mathbf{x}^H \mathbf{y} = 0$ のときに $\langle \phi(\mathbf{x}), \phi(\mathbf{y}) \rangle_{\mathcal{H}} = 0$ となるという 3 つの条件を満たす必要がある。また、ヒルベルト空間への写像のためにはカーネル関数が正定値カーネルである必要がある。これらの条件を満たすカーネル関数の例として、次式が考えられる。

$$k(\mathbf{x}, \mathbf{y}) = \sum_{i=0}^{\infty} p_i |\mathbf{x}^H \mathbf{y}|^{2i} \mathbf{x}^H \mathbf{y} + \sum_{i=1}^{\infty} q_i |\mathbf{x}^H \mathbf{y}|^{2i}, \quad p_i, q_i \geq 0 \text{ for } \forall i. \quad (20)$$

本稿では、式 (20) を満たすカーネル関数の中で最も単純で、またパワー推定型の非線形ビームフォーマに用いられる以下の二次カーネル $k_2(\cdot, \cdot)$ に焦点を当てる。

$$k_2(\mathbf{x}, \mathbf{y}) \triangleq |\mathbf{x}^H \mathbf{y}|^2. \quad (21)$$

二次カーネル $k_2(\cdot, \cdot)$ によって再生されるヒルベルト空間 \mathcal{H}_2 は \mathbb{R}^{M^2} であり、与えられる写像 $\phi_2: \mathbb{C}^M \rightarrow \mathbb{R}^{M^2}$ は以下のようにユークリッドベクトル関数として解析的に表すことができる。

$$\phi_2(\mathbf{x}) \triangleq [\phi_{\text{abs}}^T(\mathbf{x}), \phi_{\text{re}}^T(\mathbf{x}), \phi_{\text{im}}^T(\mathbf{x})]^T, \quad (22)$$

$$\phi_{\text{abs}}(\mathbf{x}) \triangleq [\sqrt{|x_i|^2} |1 \leq i \leq M]^T, \quad (23)$$

$$\phi_{\text{re}}(\mathbf{x}) \triangleq [\sqrt{2} [\text{Re}[x_i x_j^*] | 2 \leq i \leq M, 1 \leq j \leq i-1]^T, \quad (24)$$

$$\phi_{\text{im}}(\mathbf{x}) \triangleq [\sqrt{2} [\text{Im}[x_i x_j^*] | 2 \leq i \leq M, 1 \leq j \leq i-1]^T. \quad (25)$$

ここで $\mathbf{x} = [x_1, \dots, x_M]^T$, $x_i \in \mathbb{C}$ である。この等価性は以下の内積の一致により証明される。

$$\begin{aligned} k_2(\mathbf{x}, \mathbf{y}) &= |\mathbf{x}^H \mathbf{y}|^2 \\ &= \sum_{i=1}^M |x_i|^2 |y_i|^2 + 2 \sum_{i=2}^M \sum_{j=1}^{i-1} \text{Re}[x_i x_j^*] \text{Re}[y_i y_j^*] \\ &\quad + 2 \sum_{i=2}^M \sum_{j=1}^{i-1} \text{Im}[x_i x_j^*] \text{Im}[y_i y_j^*] \\ &= \phi_2^T(\mathbf{x}) \phi_2(\mathbf{y}). \end{aligned} \quad (26)$$

ここで式 (22)–(25) に注目する。 $\phi_{\text{abs}}(\mathbf{x})$ は、エルミート行列 $\mathbf{x} \mathbf{x}^H$ の対角成分よりなるベクトルとなっている。同様に、 $\phi_{\text{re}}(\mathbf{x})$, $\phi_{\text{im}}(\mathbf{x})$ は $\mathbf{x} \mathbf{x}^H$ の対角成分を除いた下三角要素の実部と虚部よりなるベクトルとなっている。入力ベクトル \mathbf{x} より得られるエルミート行列 $\mathbf{x} \mathbf{x}^H$ の成分に対応した写像となっていることから、

本稿では写像 ϕ_2 をエルミート展開写像と呼ぶこととする。

以下では、二次カーネルを用いたカーネル MUSIC の幾何学的・統計的性質について議論する。エルミート展開写像 $\phi_2: \mathbb{C}^M \rightarrow \mathbb{R}^{M^2}$ の定義より、特徴ベクトルから高次ユークリッド空間への写像によって次元数が M から M^2 へと増加していることが分かる。このとき、エルミート展開写像 ϕ_2 の与える相関線形作用素 Γ_ω を用いた MUSIC は、次元数の増加による雑音部分空間の表現能力の向上によって $(M^2 - 1)$ までの音源の方位推定が可能である。また、式 (22)–(25) から分かるように、エルミート展開写像は観測信号の $\mathbf{x}(\omega, t)$ の二次関数による写像であり、エルミート展開写像の与える相関線形作用素を用いた二次カーネル MUSIC は 4 次のモーメントの分析となっていることが分かる。同様に、式 (20) においてより高次の係数 p_i, q_i が非零となるカーネル関数を用いたカーネル MUSIC は、より高い偶数次モーメントを用いた分析となることが分かる。

4.3 エルミート MUSIC

4.1 節で議論したカーネル MUSIC に二次カーネル $k_2(\cdot, \cdot)$ を用いる場合は L 次元の固有値問題を解く必要があるが、 $L > M^2$ となる条件では、エルミート展開写像の与える $M^2 \times M^2$ 共分散行列 $\mathbf{R}_2(\omega)$:

$$\mathbf{R}_2(\omega) = E[\phi_2(\mathbf{x}(\omega, t)) \phi_2^T(\mathbf{x}(\omega, t))] \quad (27)$$

の固有値問題を解く方が計算が効率的である。本節では、 $\mathbf{R}_2(\omega)$ を用いるエルミート MUSIC アルゴリズムを提案し、4-MUSIC より計算が効率的であることを示す。共分散行列 $\mathbf{R}_2(\omega)$ の計算のためには、式 (22)–(25) の写像 $\phi_2(\mathbf{x}(\omega, t))$, $t = 1, \dots, L$ を求めるための $2M^2 L$ 回の実数乗算と、 $\phi_2(\mathbf{x}(\omega, t))$ を $\mathbf{R}_2(\omega)$ に変換するための $M^2(M^2 + 1)L/2$ 回の実数乗算のみが必要であり、4-MUSIC におけるキュムラント行列 $\mathbf{C}(\omega)$ の計算よりも乗算の回数が少ない。また、4-MUSIC における $M^2 \times M^2$ 次元キュムラント行列 $\mathbf{C}(\omega)$ が複素行列であるのに対し、 $\mathbf{R}_2(\omega)$ は同じ $M^2 \times M^2$ 行列であるが実数行列となるため、固有値問題を解くための計算が大きく単純化される。

エルミート展開写像を直接用いたエルミート MUSIC の MUSIC スコア $g_2(\omega; \theta)$ は以下のように求められる。

$$g_2(\omega; \theta) \triangleq \frac{1}{\sum_{i=N+1}^{M^2} |\phi_2^T(\mathbf{b}(\omega; \theta)) \mathbf{w}_i(\omega)|}, \quad (28)$$

$$\mathbf{R}_2(\omega) = \mathbf{W}(\omega) \mathbf{\Lambda}(\omega) \mathbf{W}^T(\omega), \quad (29)$$

$$\mathbf{W}(\omega) = [\mathbf{w}_1(\omega), \dots, \mathbf{w}_{M^2}(\omega)],$$

$$\mathbf{W}^T(\omega) \mathbf{W}(\omega) = \mathbf{I}_{M^2}, \quad (30)$$

$$\mathbf{\Lambda}(\omega) = \text{diag}[\lambda_1(\omega), \dots, \lambda_{M^2}(\omega)],$$

$$\lambda_1(\omega) \geq \dots \geq \lambda_{M^2}(\omega) \geq 0. \quad (31)$$

5 実験

本章では、提案するエルミート MUSIC の有効性を検証するために、計算機によって行ったシミュレーション実験について述べる。比較のための従来手法として、3 章に示した MUSIC, 4-MUSIC に加えて、遅延和アレー, GCC-Eckart[6] を用いた。

5.1 実験条件

提案するエルミート MUSIC の方位推定性能の評価を、音声混合のシミュレーションを用いて行った。

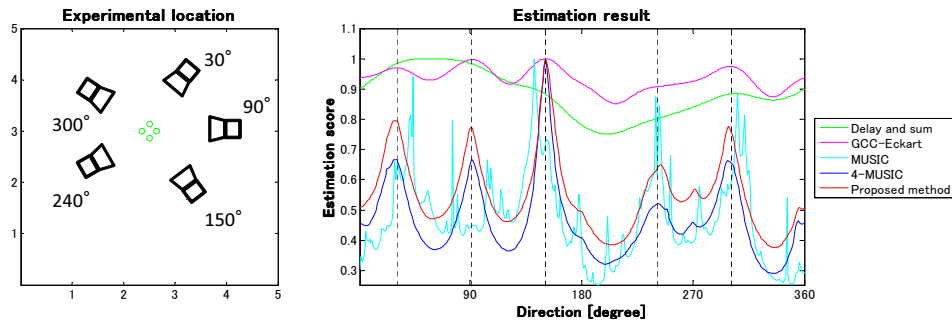


Fig. 1 Experimental results of estimation performance.

Table 1 Experimental conditions

アレー形状	円形アレー
音源	アレーから距離 1.5 m の点音源
素子数	4
音源数	5
アレー半径	0.2 m
部屋サイズ	5 m × 5 m
残響時間 (T_{60})	0, 0.12, 0.3 [s]
SNR	0, 10, 20 [dB]
サンプリング周波数	16 kHz
フレーム長	512 for T_{60} of 0 s
	1024 for T_{60} of 0.12 s
	2048 for T_{60} of 0.3 s
シフト長	1/4 フレーム
窓関数	ハミング窓

鏡像法 [7] によりインパルス応答を与え、異なる 3 つの残響時間で音声混合した。観測信号には、拡散性雑音を模擬した 3 つの異なるピンクノイズを重ねた [8]。Table 1 に各実験条件について示す。性能評価には、 $\{0^\circ, 30^\circ, 60^\circ, \dots, 330^\circ\}$ の 30° 間隔 12 方向における全組み合わせで求めた、真の音源方位と各手法の推定方位の間の最小二乗誤差平方根平均 (RMSE) を用いた。また、音源数分のピーク探索結果が得られず、正しく方位推定が行えなかった場合は、全ての推定方位のエラーを平均したペナルティを与えた。

なお、 $M \leq N$ の場合における MUSIC の方位推定は、各帯域で最小固有値に対応する一次元部分空間を雑音部分空間とみなし、広帯域のピークを検出することにより行った。

5.2 結果と考察

Fig. 1 にある方位の組み合わせのもとで、残響時間が 0.12 s の場合の方位推定の結果を示す。左側が、音源とマイクロホンアレーの配置図、右側が推定結果のグラフとなっている。Fig. 2 に RMSE の全試行における平均を示す。結果から、提案手法と 4-MUSIC がほぼ同等の優れた方位推定性能を示し、また高残響化において提案手法が優れていることが分かる。計算が効率的であることから、提案手法は十分有効な手法であることが確認できる。

6 おわりに

本稿では、劣決定条件で高分解能に複数音源の方位を推定する手法として、カーネル法と MUSIC を組み合わせたカーネル MUSIC を提案した。また、二次カーネル MUSIC と等価で、 $M^2 < L$ の場合に計算

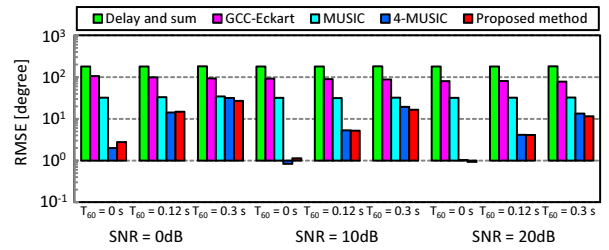


Fig. 2 Experimental results of RMSE.

が効率的なエルミート MUSIC を提案し、二次カーネルの幾何学的・統計的な性質について述べた。従来手法との性能比較から、ほぼ同等の推定性能であり、高残響下においては提案手法の推定性能が優れていることを示した。計算が効率的であることから、提案手法は十分有効な手法であることを確認した。

今後、より高次に拡張したカーネル関数での性能評価に取り組む予定である。

謝辞 この研究は、科研費 23760325 の支援を受けたものである。

参考文献

- [1] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," IEEE Trans. Antennas and Propagation, vol.AP-34, no.3, 1986.
- [2] B. Porat and B. Friedlander, "Direction finding algorithms based on higher order statistics," IEEE Trans. Signal Processing, vol.39, no.9, pp.2016-2024, 1991.
- [3] P. Chevalier, A. Ferreol and L. Albera, "High resolution direction finding from higher order statistics: The 2q-MUSIC algorithm," IEEE Trans. Signal Processing, vol.54, no.8, pp.2986-2997, 2006.
- [4] S. Miyabe, et al., "Kernel-based nonlinear independent component analysis for underdetermined blind source separation," Proc. ICASSP, pp.1641-1644, 2008.
- [5] B. Scholkopf, A. Smola and K. R. Muller, "Support vector methods in learning and feature extraction," Australian J. of Intelligent Information Processing Systems, vol.10, no.5, pp.1299-1319, 1998.
- [6] C. H. Knapp and G. C. Carter, "The generalized correlation method for estimation of time delay," IEEE Trans. Acoustics, Speech and Signal Processing, vol.24, pp.320-327, 1976.
- [7] J. A. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," J. of the Acoustical Society of America, vol.65, no.4, pp.943-950, 1979.
- [8] I. A. McCowan and H. Bourslard, "Microphone array post-filter based on noise field coherence," IEEE Trans. Speech and Audio Processing, Vol.11, No.6, pp.709-716, 2003.