

等方的雑音場を直交化するアレイ信号処理の理論と パワースペクトル推定への応用*

清水 光 (東大工), 松本 恭輔, 小野 順貴, 嵯峨山 茂樹 (東大院情報理工)

1 はじめに

種々の干渉音, 雑音が存在する環境下で目的方向の音源信号のみを取得するために, ビームフォーマや適応アレイなど, 様々な手法が提案されている [1] が, 従来は有限個の雑音源がつくる雑音場が仮定されることが多かった. これに対し本研究では, カクテルパーティや雑踏などにおける多数話者による背景雑音や室内の残響のような拡散性の雑音条件を対象に, 既知の目的方向の音源のパワースペクトルを取得するための新しい枠組を提案する. 具体的には, 拡散性雑音のモデルとして等方性を仮定し, 1) 対称性をもつアレイ配置, 2) 雑音場の直交化, 3) ノイズフリークロススペクトルからのパワースペクトルの回復, に基づく手法を論じ, 基礎的なシミュレーション実験の結果を報告する.

2 等方的雑音場の問題設定

いま, 目的音源方向を既知とし, 音場を N 個のマイクロフォンで観測するものとする. 観測信号の周波数表現 $O_n(\omega)$ (ただし n はマイクロフォン番号, ω は角周波数を表す) をベクトル表示すると,

$$O(\omega) = S(\omega)b(\omega) + N(\omega) \quad (1)$$

と表される. ただし, $S(\omega)$ は目的音源信号, $b(\omega)$ は目的音源方向で決まる既知のベクトルである.

従来のアレイ信号処理においては, 雑音 $N(\omega)$ は方向依存性の強い雑音場, すなわち有限個の音源により作られる雑音が想定されることが多かった. これに対し本稿ではこれと対照的な性格の音場として,

- パワースペクトルが観測位置に依らない
- クロススペクトルが 2 点の観測位置の方向に依らない

という性質を満たす雑音場を考え, これを等方的雑音場と定義する.

例えば正方形に配置されたマイクロフォンアレイにより等方的雑音場を観測した場合, その共分散行列は

$$V_N = \sigma^2 \begin{pmatrix} 1 & \alpha & \beta & \alpha \\ \alpha & 1 & \alpha & \beta \\ \beta & \alpha & 1 & \alpha \\ \alpha & \beta & \alpha & 1 \end{pmatrix} \quad (2)$$

と表すことができる. ここで, マイクロフォン 1-2, 2-3, 3-4, 4-1 間は正方形の一边を, 1-3, 2-4 間は正方形の対角線をなし, 距離が等しいため, 等方的雑音場の仮定の下では, クロススペクトルはそれぞれ α, β という等しい値となる. (ただし α, β の値自体は周波数に依存する.) このように行列のあるいくつかの成分同士が等しいという拘束条件を, 以下では構造と呼ぶ. この行列の構造は, センサの配置のみで決まることに注意する.

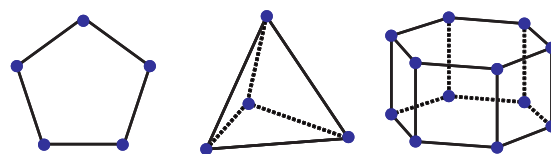


Fig. 1 対称アレイ配置: 左から正多角形, 正多面体, 正多角柱の例. この他に長方形, 直方体も含まれる

3 着眼点: 雑音場の直交化

式 (2) のような行列は一般にフルランクで零空間をもたないため, 波形領域での雑音抑圧性能には限界があるが, 音声認識への適用を想定すると, パワースペクトル領域での雑音抑圧 [2][3] も有用である. 式 (1) においてもスペクトルに着目し, 観測信号の共分散行列 V_O を考えると, 目的信号 S と雑音 N が無相関であれば,

$$V_O = \Phi_{SS}bb^h + V_N \quad (3)$$

と表される. ただし Φ_{SS} は目的信号のパワースペクトルである. またここでは簡単のため, 角周波数 ω は省略して表記した. 一般に V_N はエルミート行列なので, あるユニタリ行列 P により

$$P^h V_O P = \Phi_{SS}(P^h b)(P^h b)^h + \Lambda \quad (4)$$

と対角化される. ただし $\Lambda = P^h V_N P$ である. 理論的にはこの誤差の直交化により $P^h V_O P$ の非対角成分からは雑音が除かれることになる. また b も既知の量であるから, 式 (4) に基づき $P^h V_O P$ の非対角成分から目的音源信号のパワースペクトル Φ_{SS} を推定できないが, というのが我々の着眼点である.

4 等方的雑音場を直交化するアレイ配置

一般に V_N を直交化する行列 P は V_N に依存し, V_N を得るためには雑音のみを観測する必要があるため, 通常は前節で述べたような手法を行なうことはできない. しかし我々は, 等方的な雑音場においては行列 V_N がセンサ配置によって式 (2) のような特別な構造をもつことに着目し, V_N の各成分の値に依らず, この構造だけで行列 P が定まるセンサ配置があることを発見した. 具体的にはマイクロフォンが, 正多角形, 長方形, 正多面体, 直方体, 正多角柱 (Fig. 1 参照) の頂点に配置されている場合には, そのような行列 P が存在することを証明し, また, その具体的な行列 P の形を導出した. 紙面の都合上詳細は略すが, 例えば正方形配置の場合には,

$$P = \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{pmatrix} \quad (5)$$

* A Theory of Array Signal Processing to Orthogonalize Isotropic Noise Field and its Application to Power Spectrum Estimation by SHIMIZU, Hikaru, MATSUMOTO, Kyosuke, ONO, Nobutaka and SAGAYAMA, Shigeaki (The University of Tokyo)

と表される．ただし，構造だけで P が定まるようなアレイ配置が，前述の配置で全てであるかどうかはまだ不明であり，今後の課題の1つである．

5 ノイズフリークロススペクトルを用いたパワースペクトル推定

P によって変換された観測信号の共分散行列 $\hat{V}_O = P^h V_O P$ の (i, j) 成分を Φ_{ij} と表す．理論的にはこの非対角成分はノイズフリーであるが，実際の観測値には観測時間が有限であることや，雑音場の等方性からのずれなどのモデル化誤差が含まれる．これを ε_{ij} と表すと Φ_{ij} は

$$\Phi_{ij} = b_{ij}\Phi_{SS} + \varepsilon_{ij} \quad (6)$$

と表される．ここで， b_{ij} は式 (4) 中の bb^h の (i, j) 成分であり，例えば，

$$b_{12} = -(j\sin\omega\tau_1 + \sin\omega\tau_2)(\cos\omega\tau_1 + \cos\omega\tau_2) \quad (7)$$

$$\tau_1 = \frac{D\cos\theta}{c}, \quad \tau_2 = \frac{D\sin\theta}{c} \quad (8)$$

などとなる．ただし， D はマイクロフォンと原点の距離， c は音速， θ は音源方向である．

モデル化誤差 ε_{ij} の大きさも未知であり，また成分 ij によって異なると考えられるので，ここではこれら分散の異なるガウス分布に従うと仮定し，これらの分散も推定すべきパラメータとして最尤法を適用する．具体的には定数項を除いた対数尤度関数

$$L(\Phi_{SS}, \sigma) = \sum_{i \neq j}^K -\log\sigma_{ij} - \frac{|\Phi_{ij} - b_{ij}\Phi_{SS}|^2}{2\sigma_{ij}^2} \quad (9)$$

をパワースペクトル Φ_{SS} ，分散 σ_{ij} に関して最大化する．ここで， K は非対角成分の総数を表す（等価成分，つまり対称成分は除く）．これは非線形であるが，対数尤度関数の各変数での偏微分を 0 とおくことにより求まる以下の更新式

$$\sigma_{ij}^{2(t+1)} = |\Phi_{ij} - b_{ij}\Phi_{SS}^{(t)}|^2, \quad \Phi_{SS}^{(t+1)} = \frac{\sum_{i \neq j}^K \frac{b_{ij}\Phi_{ij}}{\sigma_{ij}^{2(t+1)}}}{\sum_{i \neq j}^K \frac{|b_{ij}|^2}{\sigma_{ij}^{2(t+1)}}} \quad (10)$$

を反復的に適用することで Φ_{SS} が求まる．

6 シミュレーション実験による評価

Fig. 2 に示すような条件で等方的雑音場をシミュレートし，音源方向が既知である単一目的信号のパワースペクトルを推定する実験を行なった．フレーム長は 48ms とし，観測データからの相関関数、クロススペクトルの算出及び平滑化は FFT 法を用いた [5]．求めたパワースペクトルの推定精度は，以下の SD (Spectral Distortion Measure) [6] を用いて評価した．

$$SD^2 = \frac{10^2}{N} \sum_k^N \left(\log_{10} |\Phi_{SS,k}| - \log_{10} |\hat{\Phi}_{SS,k}| \right)^2 \quad (11)$$

ここで， N はサンプル点数， $\hat{\Phi}_{SS}$ は推定されたパワースペクトルを表す．

SN 比を変化させた時の本手法，DS 法，及び雑音を

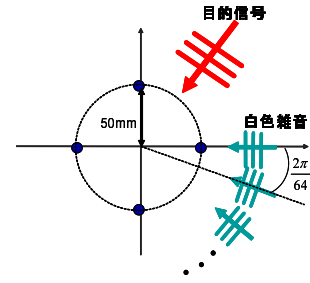


Fig. 2 雑音抑圧シミュレーションの実験条件．半径 50mm の円周上に 4 個のマイクロフォンを等間隔に配置し，一周 2π を 64 等分した各方向から統計的にパワーが等しい白色雑音を加え，等方的雑音場をシミュレートした．

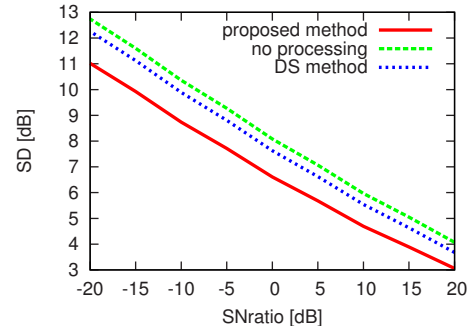


Fig. 3 SN 比の変化に対する SD の変化

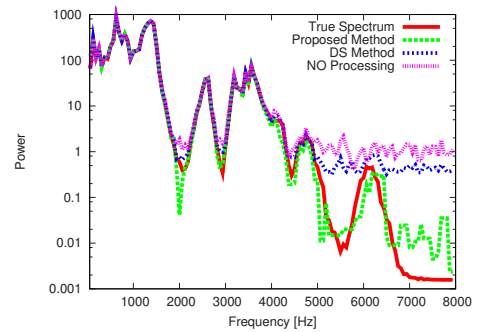


Fig. 4 推定されたパワースペクトルの例
基本的な DS 法では，ノイズのパワーに埋もれてしまうスペクトルの微細な構造が，本手法により復元されている

除去しない場合の推定されたパワースペクトルと，ノイズを加えないパワースペクトルとの SD の変化を Fig. 3 に示す．また，実際に推定されたパワースペクトルを Fig. 4 に示す．これらの実験結果から，同一の個数のマイクロフォンを用いた場合の基本的な DS 法と比較して，スペクトルの推定精度の向上が確認された．今後は実環境での検証も行なう予定である．

参考文献

- [1] M. Brandstein et al., *Microphone Arrays*, Springer, 2001.
- [2] S. F. Boll, *IEEE Trans. ASSP.*, Vol. 27, No. 2, pp. 113–120, Apr. 1979.
- [3] H. Saruwatari et al., *IEICE Trans. Fundamentals*, Vol. E82-A, No. 8, Aug. 1999.
- [4] 伊理正夫，一般線形代数，岩波書店，2003.
- [5] 日野幹雄，スペクトル解析，朝倉書店，2005.
- [6] 古井貞熙，デジタル音声処理，東海大学出版会，1985.