

結晶型アレイを用いた ポストフィルタリングに基づく拡散性雑音抑圧*

伊藤 信貴, 小野 順貴, 嵯峨山 茂樹 (東大院情報理工)

1 はじめに

我々は雑踏の喧騒, 室内の残響等の拡散性雑音の存在下で所望の信号を取得するための新しいアレイ信号処理として, 結晶のような形状をした対称性の高いアレイを用いた雑音成分の無相関化に基づく枠組の研究を進めている [1, 2, 3, 4, 5]。これまでに, この枠組による目的信号のパワースペクトル推定法を示し [1, 2], その推定値を用いて MVDR (Minimum Variance Distortionless Response) ビームフォーマ出力の振幅を置き換えるソフトマスキングによる雑音抑圧法を提案した [3]。今回, 更なる性能向上のため, 上記の信号パワースペクトル推定値を用いて, 目的信号の平均二乗誤差最小の推定器であるマルチチャンネル Wiener フィルタ [6] を設計した [5] ので報告する。

2 マルチチャンネル Wiener フィルタ

2.1 観測モデル

アレイを構成する M 個のマイクロフォンの各々が, 遅延・減衰した目的信号に拡散性雑音が重畳した信号を観測すると仮定する。これは時間周波数領域で,

$$\mathbf{X}(t, \omega) = d(\omega)S(t, \omega) + \mathbf{N}(t, \omega) \quad (1)$$
と書ける。ここで, $\mathbf{X}(t, \omega)$ は観測信号ベクトル, $S(t, \omega)$ は目的信号, $d(\omega)$ は既知のステアリングベクトル, $\mathbf{N}(t, \omega)$ は拡散性雑音成分である。 S と \mathbf{N} は零平均で互いに無相関であると仮定する。

2.2 マルチチャンネル Wiener フィルタ

目的信号 $S(t, \omega)$ を観測信号の荷重和として

$$\hat{S}(t, \omega) \triangleq \mathbf{w}^H(t, \omega) \mathbf{X}(t, \omega) \quad (2)$$

により推定することを考える。以下では簡単のため引数 t, ω を省略する。平均二乗誤差規範で最適な \mathbf{w} はマルチチャンネル Wiener フィルタ (MWF)

$$\mathbf{w}_{\text{opt}} \triangleq \Phi_{\mathbf{X}\mathbf{X}}^{-1} \phi_{SS} \mathbf{d} \quad (3)$$

である。ここで, $\Phi_{\mathbf{X}\mathbf{X}} \triangleq E[\mathbf{X}\mathbf{X}^H]$, $\phi_{SS} \triangleq E[|S|^2]$ であり, 以下でも同様の記法を用いる。

$$\Phi_{\mathbf{X}\mathbf{X}} = \mathbf{d}\mathbf{d}^H \phi_{SS} + \Phi_{\mathbf{N}\mathbf{N}} \quad (4)$$

に注意すると, \mathbf{w}_{opt} は次のように MVDR ビームフォーマ (\mathbf{w}_{MVDR}) と Wiener ポストフィルタと呼ばれる時間周波数マスク (H_{post}) の積に分解できる [6]。

$$\mathbf{w}_{\text{opt}} = \underbrace{\frac{\phi_{SS}}{\phi_{SS} + \phi_{N_0 N_0}}}_{H_{\text{post}}} \cdot \underbrace{\frac{\Phi_{\mathbf{X}\mathbf{X}}^{-1} \mathbf{d}}{\mathbf{d}^H \Phi_{\mathbf{X}\mathbf{X}}^{-1} \mathbf{d}}}_{\mathbf{w}_{\text{MVDR}}} \quad (5)$$

ここで, $N_0 \triangleq \mathbf{w}_{\text{MVDR}}^H \mathbf{N}$ である。 \mathbf{w}_{MVDR} は観測信号を用いて容易に得られるが, H_{post} を実装するためには, ϕ_{SS} を雑音が重畳した観測信号から推定する必要がある。

2.3 Zelinski の方法

Zelinski [7] による Wiener ポストフィルタの設計法では, 雑音がチャンネル間で無相関であるという仮定の下, $\Phi_{\mathbf{X}\mathbf{X}}$ の非対角成分 (上記仮定によるとノイズフリー) から ϕ_{SS} を取得していた。しかし実際には雑音はチャンネル間で相関を持ち, 特に, マイクロフォン間隔が小さい場合や低周波数帯域で大きな誤差要因となっていた。

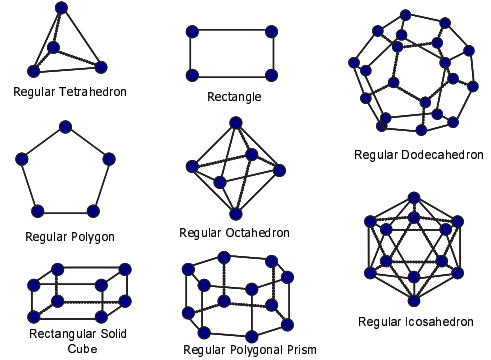


Fig. 1 結晶型アレイの例

3 提案手法

3.1 ブラインド雑音無相関化による信号パワースペクトル推定

Zelinski の方法の問題点を踏まえ, 線形変換により N を無相関化することを考える。 e_1, e_2, \dots, e_M を $\Phi_{\mathbf{N}\mathbf{N}}$ の正規直交な固有ベクトルとし, $\mathbf{E} \triangleq [e_1, e_2, \dots, e_M]$ とすると,

$$\tilde{\mathbf{N}} \triangleq \mathbf{E}^H \mathbf{N} \quad (6)$$

の各成分は無相関である。 $\tilde{\mathbf{X}}, \tilde{\mathbf{d}}$ を $\tilde{\mathbf{N}}$ と同様に定義すると, (4) はこの変換により

$$\Phi_{\tilde{\mathbf{X}}\tilde{\mathbf{X}}} = \tilde{\mathbf{d}}\tilde{\mathbf{d}}^H \phi_{SS} + \Phi_{\tilde{\mathbf{N}}\tilde{\mathbf{N}}} \quad (7)$$

となり, $\Phi_{\tilde{\mathbf{X}}\tilde{\mathbf{X}}}$ の非対角成分はノイズフリーである。従って, 行列 \mathbf{E} を求めることができれば, $\Phi_{\tilde{\mathbf{X}}\tilde{\mathbf{X}}}$ の非対角成分から ϕ_{SS} を取得することができる。

通常 \mathbf{E} を求めるためには $\Phi_{\mathbf{N}\mathbf{N}}$ の値を知る必要がある。これは, 雑音が非音声である場合には VAD (Voice Activity Detection) による雑音区間の切り出しにより可能だが, 雑音も音声である場合には困難である。ところが我々は, 拡散性雑音場のモデルとして

- パワースペクトルが観測点に依らない。
- クロススペクトルが 2 点間の方向に依らない。

なる等方性の仮定を置くと, 結晶型アレイと呼ぶ対称性の高いアレイ (Fig. 1 を参照) を用いることにより, $\Phi_{\mathbf{N}\mathbf{N}}$ の値に依存しない E を取ることができることを示し, 結晶型アレイの各々に対して E を求めた [1, 2, 4]。例えば正多角形アレイの場合, E として DFT 行列を取ることができる。従って, 結晶型アレイを用いると, $\Phi_{\mathbf{N}\mathbf{N}}$ の値が未知でも等方的雑音は無相関化することができ, これをブラインド雑音無相関化 (Blind Noise Decorrelation: BND) と呼ぶ。

提案法では, BND を適用して, (7) の非対角成分

$$\phi_{\tilde{\mathbf{X}}_m \tilde{\mathbf{X}}_n} = \tilde{d}_m \tilde{d}_n^* \phi_{SS} \quad (m \neq n) \quad (8)$$

から最小二乗法に基づき, ϕ_{SS} を

$$\hat{\phi}_{SS} = \Re \left[\frac{\sum_{m < n} \tilde{d}_m^* \tilde{d}_n \phi_{\tilde{\mathbf{X}}_m \tilde{\mathbf{X}}_n}}{\sum_{m < n} |\tilde{d}_m|^2 |\tilde{d}_n|^2} \right] \quad (9)$$

により推定する [3]。ここで * は複素共役を表す。

* Diffuse Noise Suppression Based on Post-filtering Using Crystal Arrays by ITO, Nobutaka, ONO, Nobutaka, and SAGAYAMA, Shigeki (The University of Tokyo)

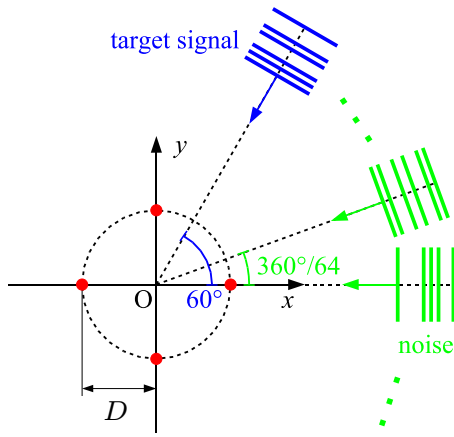


Fig. 2 検証実験におけるマイクロフォンと音源配置

3.2 Wiener ポストフィルタの設計

(9) の $\hat{\phi}_{SS}$ と Zelinski による分母の推定法 [7] を組み合わせて、Wiener ポストフィルタを

$$H_{\text{post}} = \frac{\hat{\phi}_{SS}}{\frac{1}{M} \sum_{m=1}^M \phi_{X_m X_m} / |d_m|^2} \quad (10)$$

と設計する。また、定義より $0 \leq H_{\text{post}} \leq 1$ であるから、(10) の後処理として

$$H_{\text{post}} \leftarrow \begin{cases} 0 & (H_{\text{post}} < 0) \\ 1 & (H_{\text{post}} > 1) \end{cases} \quad (11)$$

とすることにより、上記範囲に含まれるよう修正する。

4 検証実験と結果

Fig. 2 に示すように、自由空間において無数の方向から相異なる干渉信号（音声または白色雑音）が到来する拡散性雑音環境下で、特定方向から到来する目的信号（音声）を提案法及び Zelinski の方法により設計した Wiener ポストフィルタを用いた MVDR により抽出するシミュレーションを行った。目的信号、干渉信号とも平面波伝播を仮定した。アレイ配置は正方形とした。音声は ATR 音声データベース B セットの連続音声を 16kHz にダウンサンプリングしたものをを用いた。分析条件は、フレーム長: 2^8 点、フレームシフト: 2^4 点とし、Hamming 窓を用いた。 w_{MVDR} における ϕ_{XX} は全データから算出した。行列 E は DFT 行列を用いた。(9)、(10) における観測信号のパワースペクトル及びクロススペクトルは FFT 法 [8] により 2^4 フレーム毎に推定した。 H_{post} の値域制限処理 (11) は Zelinski の方法に対しても適用した。

Fig. 3 に目的信号と両手法による分離結果のスペクトログラムを示す。条件は、干渉信号: 音声、アレイ半径: 0.015m, 入力 SN 比: 0.1dB とした。Zelinski の方法では低周波数帯域において雑音が余り抑圧されずに残っている。これは雑音がチャンネル間で無相関であるという仮定が低周波数帯域では成立していないからであると考えられる。一方、提案法では低周波数帯域においても雑音を効果的に抑圧できている。

Fig. 4 にアレイ半径を変化させたときの両手法による分離結果の SN 比のグラフを示す。ポストフィルタリングによる SN 比増分が分かるように、MVDR ビームフォーマ単独（ポストフィルタなし）での結果の SN 比も共に示す。条件は、干渉信号: 音声または白色雑音、入力 SN 比: 0.0dB とした。Zelinski の方法では、アレイ半径を小さくしてゆくと性能が低下して MVDR ビームフォーマ単独での結果に近づいてゆく。これは、アレイ半径が小さい場合には上述の仮定が崩れ、従って ϕ_{SS} が過大評価されて H_{post} が 1 に近い値に見積もられるためであると考えられる。それ

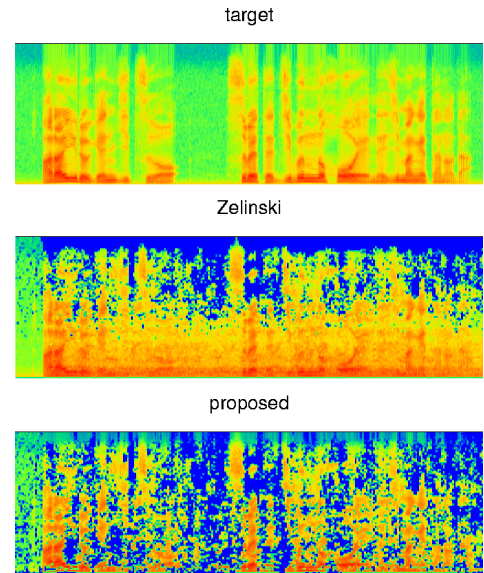


Fig. 3 目的信号及び分離信号のスペクトログラム

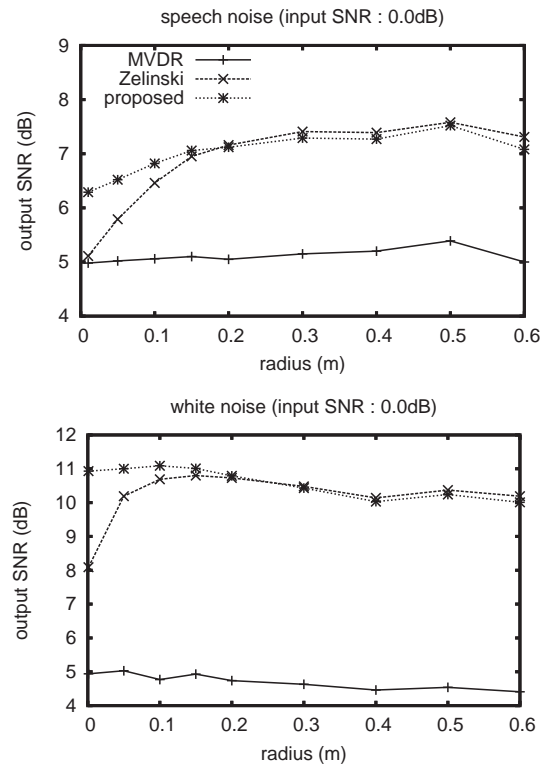


Fig. 4 アレイ半径に対する出力 SN 比の変化

に対し提案法ではアレイ半径が小さい場合にも MVDR ビームフォーマ単独での結果と比べて大幅に SN 比が改善されている。

参考文献

- [1] 清水他, 音講論 (春), pp. 569–570, 2007.
- [2] H. Shimizu *et al.*, Proc. WASPAA, pp. 54–57, 2007.
- [3] 伊藤他, 音講論 (秋), pp. 619–620, 2007.
- [4] 小野他, 音講論 (秋), pp. 621–622, 2007.
- [5] N. Ito *et al.*, Proc. ICASSP, 2008. (accepted)
- [6] K. Simmer *et al.*, in *Microphone Arrays*, M. Brandstein *et al.* (Eds.), Springer, ch. 2, pp. 19–38, 2001.
- [7] R. Zelinski, Proc. ICASSP, pp. 2578–2581, 1988.
- [8] 日野幹雄, スペクトル解析, 朝倉書店, 2005.