

結晶型マイクロフォンアレイを用いた ポストフィルタ設計に基づく拡散性雑音抑圧

伊藤 信貴[†] 小野 順貴[†] 嵯峨山茂樹[†]

[†] 東京大学大学院 情報理工学系研究科 システム情報学専攻

〒 113-8656 東京都文京区本郷 7-3-1

E-mail: †{ito,onono,sagayama}@hil.t.u-tokyo.ac.jp

あらまし 我々は、拡散性雑音抑圧の新しい枠組として、結晶型のアレイ配置を用いた、拡散性雑音のブラインドな無相関化に基づく手法を研究している。拡散性雑音抑圧のための手法として近年注目されているのが多チャンネル Wiener フィルタであるが、これを実現するには、雑音を含む観測信号から目的信号のパワースペクトルを推定する必要がある。その際に、従来、素子間距離が小さい場合において、素子間の雑音相関が問題となっていた。これに対し、本発表では、結晶型アレイを用いた雑音無相関化により精確に推定された信号パワースペクトルを用いて、素子間距離が小さい場合でも Wiener ポストフィルタを適切に設計する方法を示す。

キーワード アレイ信号処理, 拡散性雑音, 多チャンネル Wiener フィルタ, ポストフィルタ無相関化, パワースペクトル推定

Diffuse noise suppression by crystal-array-based post-filter design

Nobutaka ITO[†], Nobutaka ONO[†], and Shigeki SAGAYAMA[†]

[†] Graduate School of Information Science and Technology, The University of Tokyo

7-3-1 Hongo Bunkyo-ku Tokyo 113-8656, Japan

E-mail: †{ito,onono,sagayama}@hil.t.u-tokyo.ac.jp

Abstract We are studying a new framework for diffuse noise suppression, which is based on blind decorrelation of diffuse noise using microphone arrays with crystal-like symmetry. Recently, many researchers have been studying the multichannel Wiener filter as a potential tool for diffuse noise suppression, where estimation of the power spectrum of the target signal from the observed signals is required. Formerly, noise correlation between sensors was main problem. To this end, this paper describes a method to design the multichannel Wiener filter properly even when the intersensor distance is small, utilizing signal power spectrum estimate obtained via noise decorrelation with crystal arrays.

Key words array signal processing, diffuse noise, multichannel Wiener filter, post-filter, decorrelation, power spectrum estimation

1. はじめに

マイクロフォンアレイを用いて、既知の方向から到来する目的音を強調し、他の音響信号（以下「雑音」）を抑圧する技術は、補聴器 [1]、テレビ会議 [2]、音声認識 [3]、携帯電話 [4] を始めとして多様な応用があり、数多くの研究がなされてきた [5], [6]。最も基本的な遅延和ビームフォーマは、アレイ開口を波長に対して十分大きくすれば、鋭いビームを形成して雑音を効果的に抑圧することができる。しかし、その際、空間的エリアシングを回避するためには多数のマイクロフォンを用いる必要がある。

また、多くの応用においては、比較的長い音波の波長に対して十分に大きな開口を取ることは困難である。これに対し、適応ビームフォーマは、少数の点音源から発生する雑音に対しては、指向性の零点を向けることにより、原理的には開口サイズによらず効果的な除去が可能である。一方、残響の多い室内、カクテルパーティー、空港、車内などのように、あらゆる方向から雑音が入り込む環境においては、適応ビームフォーマによる雑音抑圧は不十分であり、音響信号処理における課題の一つである。

このような拡散性雑音を抑圧するための枠組みとして近年注目されているのが、目的信号の線形最小平均二乗誤差 (Linear

Minimum Mean Square Error: LMMSE) 推定器 [7] である多チャンネル Wiener フィルタ [8] ~ [11] である。これを実現するには、雑音を含む観測信号から目的信号のパワースペクトルを推定する必要がある。この問題に対し、Zelinski は、拡散性の雑音場では素子間距離を波長に対して十分大きく取れば、素子間での雑音相関が小さくなることに着目し、観測信号のクロススペクトルから目的信号のパワースペクトルを推定する方法を提案した [12]。

本発表では、結晶構造のような対称的なアレイ [13] を用いた、雑音無相関化に基づく目的音の正確なパワースペクトル推定 [14], [15] に基づく、素子間隔が小さい場合にも適用可能な多チャンネル Wiener フィルタの設計法 [15] について述べる。

2. 背景

2.1 雑音抑圧の問題設定

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 (音源位置既知) に拡散性雑音加わった信号を観測すると仮定する。観測モデルは、時間周波数領域で

$$X_m(\tau, \omega) = S(\tau, \omega)D_m(\omega) + N_m(\tau, \omega) \quad (1)$$

と表される。ここで、 $m = 1, 2, \dots, M$ はマイクロフォンの番号、 $X_m(\tau, \omega)$ は観測信号、 $S(\tau, \omega)$ は目的信号、 $D_m(\omega)$ は S から X_m への伝達関数 (既知)、 $N_m(\tau, \omega)$ は雑音である。

$$\mathbf{X}(\tau, \omega) \triangleq [X_1(\tau, \omega) \quad X_2(\tau, \omega) \quad \dots \quad X_M(\tau, \omega)]^T \quad (2)$$

とし、 $D(\omega)$ 、 $N(\tau, \omega)$ も同様に定義すると、(1) はまとめて

$$\mathbf{X}(\tau, \omega) = S(\tau, \omega)\mathbf{D}(\omega) + \mathbf{N}(\tau, \omega) \quad (3)$$

と書ける。 $S(\tau, \omega)$ 、 $N(\tau, \omega)$ は零平均で互いに無相関と仮定する。本稿で扱う問題は、 $\mathbf{X}(\tau, \omega)$ から $S(\tau, \omega)$ を推定することである。推定量のクラスとしては、線形推定量

$$\hat{S}(\tau, \omega) \triangleq \mathbf{W}^H(\tau, \omega)\mathbf{X}(\tau, \omega) \quad (4)$$

を考える。以下では次の表記を用いる ($E[\cdot]$ は期待値) :

$$\Phi_{\mathbf{X}\mathbf{X}}(\tau, \omega) \triangleq E[\mathbf{X}(\tau, \omega)\mathbf{X}^H(\tau, \omega)] \quad (5)$$

$$\Phi_{\mathbf{N}\mathbf{N}}(\tau, \omega) \triangleq E[\mathbf{N}(\tau, \omega)\mathbf{N}^H(\tau, \omega)] \quad (6)$$

$$\Phi_{SS}(\tau, \omega) \triangleq E[|S(\tau, \omega)|^2] \quad (7)$$

$$\Phi_{\mathbf{X}S}(\tau, \omega) \triangleq E[\mathbf{X}(\tau, \omega)S^*(\tau, \omega)] \quad (8)$$

2.2 MVDR ビームフォーマ

適応ビームフォーマの一種である MVDR (Minimum Variance Distortionless Response) ビームフォーマは、目的信号を歪ませないことを意味する線形拘束条件

$$\mathbf{W}^H(\tau, \omega)\mathbf{D}(\omega) = 1 \quad (9)$$

の下で、出力パワー

$$\mathbf{W}^H(\tau, \omega)\Phi_{\mathbf{X}\mathbf{X}}(\tau, \omega)\mathbf{W}(\tau, \omega) \quad (10)$$

を最小化する $\mathbf{W}(\tau, \omega)$ であり、

$$\mathbf{W}_{\text{MVDR}}(\tau, \omega) \triangleq \frac{\Phi_{\mathbf{X}\mathbf{X}}^{-1}(\tau, \omega)\mathbf{D}(\omega)}{\mathbf{D}^H(\omega)\Phi_{\mathbf{X}\mathbf{X}}^{-1}(\tau, \omega)\mathbf{D}(\omega)} \quad (11)$$

で与えられる。

MVDR ビームフォーマは、雑音が少数 ($M - 1$ 個以下) の点音源からのみ到来する場合には、雑音源方向に指向性の零点を向けて効果的に抑圧を行う。一方、あらゆる方向から雑音が到来する拡散性の雑音環境においては、その性能は著しく低下する。

2.3 多チャンネル Wiener フィルタ

多チャンネル Wiener フィルタ (multichannel Wiener filter) は、平均二乗誤差

$$E[|\hat{S}(\tau, \omega) - S(\tau, \omega)|^2] \quad (12)$$

を最小化する $\mathbf{W}(\tau, \omega)$ であり、

$$\mathbf{W}_{\text{MWF}}(\tau, \omega) \triangleq \Phi_{\mathbf{X}\mathbf{X}}^{-1}(\tau, \omega)\Phi_{\mathbf{X}S}(\tau, \omega) \quad (13)$$

で与えられる。

多チャンネル Wiener フィルタは、次のように MVDR ビームフォーマとその出力における Wiener フィルタ $M(\tau, \omega)$ に分解できることが知られている [9]。

$$\mathbf{W}_{\text{MWF}}(\tau, \omega) = M(\tau, \omega)\mathbf{W}_{\text{MVDR}}(\tau, \omega) \quad (14)$$

ここで、

$$M(\tau, \omega) \triangleq \frac{\Phi_{SS}(\tau, \omega)}{\Phi_{YY}(\tau, \omega)} \quad (15)$$

である。但し、

$$Y(\tau, \omega) \triangleq \mathbf{W}_{\text{MVDR}}^H(\tau, \omega)\mathbf{X}(\tau, \omega) \quad (16)$$

は MVDR ビームフォーマの出力であり、

$$\Phi_{YY}(\tau, \omega) \triangleq E[|Y(\tau, \omega)|^2] \quad (17)$$

である。従って、

Φ_{YY} は

$$\hat{\Phi}_{YY}(\tau, \omega) \triangleq \frac{1}{M} \sum_{m=1}^M \hat{\Phi}_{X_m X_m}(\tau, \omega) \quad (18)$$

により推定できる [12] から、分子の $\Phi_{SS}(\tau, \omega)$ の推定が重要である ($\hat{\cdot}$ は推定値)。

3. 目的信号のパワースペクトルの推定

信号のパワースペクトル Φ_{SS} を推定する必要がある。ところが観測できるのは \mathbf{X} であるから、観測パワースペクトル $\Phi_{\mathbf{X}\mathbf{X}}$ の推定値から Φ_{SS} を推定する必要がある。

$$\Phi_{\mathbf{X}\mathbf{X}} = \Phi_{SS}\mathbf{D}\mathbf{D}^H + \Phi_{\mathbf{N}\mathbf{N}} \quad (19)$$

であるが、雑音パワースペクトル $\Phi_{\mathbf{N}\mathbf{N}}$ の存在下で Φ_{SS} をいかにして推定するかが問題である。

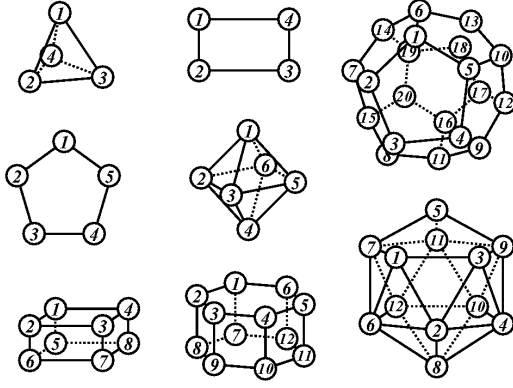


図1 結晶型アレイの例

3.1 Zelinski の方法 [12]

Zelinski は、拡散性の雑音環境においては、波長に対して素子間距離を十分大きくすれば、素子間での雑音の相関が小さくなることに着目した。素子間で雑音が無相関であると近似すると、 Φ_{NN} は対角行列ゆえ、(19) の非対角成分より、

$$\Phi_{X_m X_n} = \Phi_{SS} D_m D_n^* \quad (m \neq n) \quad (20)$$

を得る。従って、各非対角成分から、 Φ_{SS} の推定値

$$\hat{\Phi}_{SS}^{\text{Zel}(m,n)} \triangleq \Re \left[\frac{\hat{\Phi}_{X_m X_n}}{D_m D_n^*} \right] \quad (m \neq n) \quad (21)$$

を得ることができる。これを m, n に関して平均して、最終的な推定値

$$\hat{\Phi}_{SS}^{\text{Zel}} \triangleq \frac{1}{M(M-1)} \sum_{m < n} \hat{\Phi}_{SS}^{\text{Zel}(m,n)} \quad (22)$$

を得る。

3.2 提案手法

素子間隔が小さい場合には、雑音は素子間で強い相関を持ち、これが Zelinski の方法の大きな誤差要因となっていた。仮に線形変換により Φ_{NN} を対角化できれば、変換によりノイズフリーとなった (19) の非対角成分から Φ_{SS} をより精度良く推定できると期待される。問題は未知の Φ_{NN} をいかにして対角化するかということである。

我々は、雑音分散は素子に依らず、雑音共分散は2素子間の距離のみに依存するという仮定の下、図1に示す結晶型アレイを用いれば、 Φ_{NN} を定行列により対角化できることを示した [13], [14]。これにより、 Φ_{NN} を未知のまま、いわばブラインドに対角化することができる。これをブラインド雑音無相関化と呼んでいる。例えば正方形アレイを用いた場合、 Φ_{NN} は巡回行列

$$\Phi_{NN} = \begin{bmatrix} \alpha & \beta & \gamma & \beta \\ \beta & \alpha & \beta & \gamma \\ \gamma & \beta & \alpha & \beta \\ \beta & \gamma & \beta & \alpha \end{bmatrix} \quad (23)$$

であるから、対角化行列 P は、DFT (Discrete Fourier Transform) 行列

$$P = \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & \zeta & \zeta^2 & \zeta^3 \\ 1 & \zeta^2 & \zeta^4 & \zeta^6 \\ 1 & \zeta^3 & \zeta^6 & \zeta^9 \end{bmatrix} \quad (24)$$

である (Φ_{NN} , α, β, γ の引数 τ, ω は省略)。ここで、 $\zeta \triangleq \exp(j\frac{2\pi}{4}) = j$ である (j は虚数単位)。

(19) にブラインド雑音無相関化の基底変換を施すと

$$\Phi_{\tilde{X}\tilde{X}} = \Phi_{SS} \tilde{D} \tilde{D}^H + \Phi_{\tilde{N}\tilde{N}} \quad (25)$$

となる ($\tilde{\cdot}$ は基底変換行列 P の共役転置を左から掛ける変換)。 $\Phi_{\tilde{N}\tilde{N}}$ は対角行列ゆえ、(25) の非対角成分より

$$\Phi_{\tilde{X}_m \tilde{X}_n} = \Phi_{SS} \tilde{D}_m \tilde{D}_n^* \quad (m \neq n) \quad (26)$$

を得る。従って、最小二乗法を用いて、

$$\hat{\Phi}_{SS}^{\text{prop}} = \Re \left[\frac{\sum_{m < n} \tilde{D}_m^* \tilde{D}_n \hat{\Phi}_{\tilde{X}_m \tilde{X}_n}}{\sum_{m < n} |\tilde{D}_m|^2 |\tilde{D}_n|^2} \right] \quad (27)$$

により Φ_{SS} を推定できる [15]。

4. 検証実験と結果

提案法の有効性を実証するためにシミュレーション実験を行った。平面内の64方向から異なる音声が入射する雑音環境において、提案法および Zelinski の方法による信号パワースペクトル推定に基づく多チャンネル Wiener フィルタにより目的音声 (到来方向既知) を取得した。音声は ATR 音声データベース B セット [16] の連続音声を用いた。アレイは半径 1.5cm の正方形アレイを用いた。

信号パワースペクトルの推定精度を確認するため、推定値 (縦軸) とクリーンな信号から計算した「真値」 (横軸) の散布図を図2に示す。各点はある時間周波数スロット (τ, ω) における Φ_{SS} に対応する。赤い点 (提案法) が緑の点 (Zelinski の方法) よりも直線 $y = x$ の近くに集中していることから、提案法による推定がより精確であると分かる。

図3にスペクトログラムの例を示す。1) は目的信号、2) は観測信号 (SN 比: 5.0dB) であり、3) および 4) は Zelinski の方法および提案法による多チャンネル Wiener フィルタの出力である。Zelinski の方法では十分に雑音を抑圧できていない (特に素子間の雑音相関が顕著な低周波帯域) が、提案法では効果的に抑圧できていることが分かる。

5. おわりに

結晶型アレイによるブラインド雑音無相関を用いて、素子間距離が小さい場合にも有効な多チャンネル Wiener フィルタを設計する方法を述べた。

展望としては、結晶型アレイを実際に製作して実環境で実験を行う準備を進めており、更に、ブラインド雑音無相関化の音源定位への応用も検討中である。

文 献

- [1] J. M. Kates and M. R. Weiss, "A comparison of hearing-aid array-processing techniques," *JASA*, vol. 99, no. 5, pp.

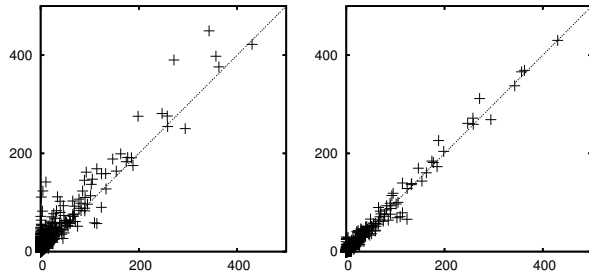


図2 信号パワースペクトルの真値 (横軸) と推定値 (縦軸) の散布図。
左: Zelinski の推定法、右: 提案法

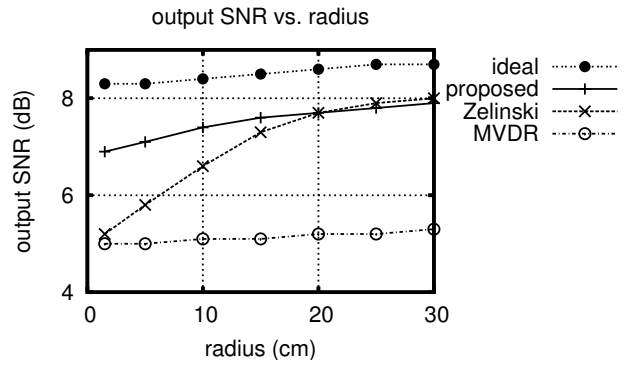


図4 SNR vs. radius

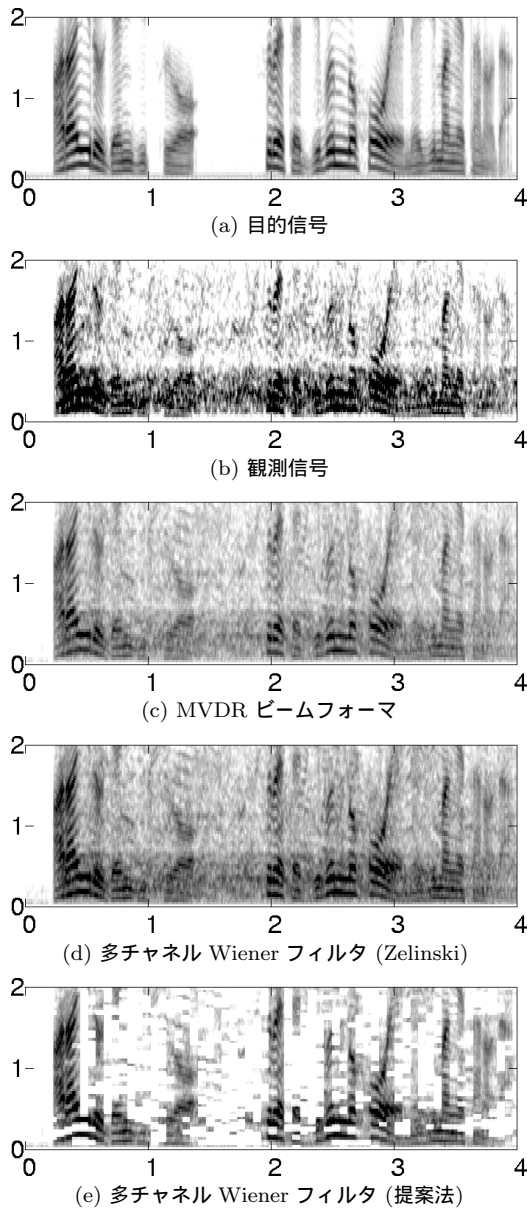


図3 スペクトログラムの例

- 3138–3148, May 1996.
- [2] F. Khalil, J. P. Jullien, and A. Gilloire, “Microphone array for sound pickup in teleconference systems,” *Journal of the Audio Engineering Society*, vol. 42, no. 9, pp. 691–700, Sep. 1994.
 - [3] M. Omologo, M. Matassoni, and P. Svaizer, “Speech recognition with microphone arrays,” in *Microphone Arrays*, M.

Brandstein and D. Ward, Eds., Springer-Verlag, ch. 15, pp. 331–353, 2001.

- [4] J. G. Ryan and R. A. Goubran, “Application of near-field optimum microphone arrays to hands-free mobile telephony,” *IEEE Trans. Vehicular Technology*, vol. 52, no. 2, pp. 390–400, Mar. 2003.
- [5] D. H. Johnson and D. E. Dudgeon, *Array Signal Processing: Concepts and Techniques*, Prentice-Hall, 1993.
- [6] M. Brandstein and D. Ward, Eds., *Microphone Arrays*, Springer-Verlag, 2001.
- [7] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*, Prentice-Hall, 1993.
- [8] C. Marro, Y. Mahieux, and K. U. Simmer, “Analysis of noise reduction and dereverberation techniques based on microphone arrays with postfiltering,” *IEEE Trans. SAP*, vol. 6, no. 3, pp. 240–259, May 1998.
- [9] K. U. Simmer, J. Bitzer, and C. Marro, “Post-filtering techniques,” in *Microphone Arrays*, M. Brandstein and D. Ward, Eds., Springer-Verlag, ch. 3, pp. 39–60, 2001.
- [10] A. Spriet, M. Moonen, and J. Wouters, “Robustness analysis of multichannel Wiener filtering and generalized sidelobe cancellation for multimicrophone noise reduction in hearing aid applications,” *IEEE Trans. SAP*, vol. 13, no. 4, pp. 487–503, Jul. 2005.
- [11] A. Spriet, M. Moonen, and J. Wouters, “The impact of speech detection errors on the noise reduction performance of multi-channel Wiener filtering and generalized sidelobe cancellation,” *Sig. Pro.*, vol. 85, no. 6, pp. 1073–1088, Jun. 2005.
- [12] R. Zelinski, “A microphone array with adaptive post-filtering for noise reduction in reverberant rooms,” *Proc. ICASSP*, pp. 2578–2581, Apr. 1988.
- [13] N. Ono, N. Ito, and S. Sagayama, “Five classes of crystal arrays for blind decorrelation of diffuse noise,” *Proc. SAM*, Jul. 2008. (to appear)
- [14] H. Shimizu, N. Ono, K. Matsumoto, and S. Sagayama, “Isotropic noise suppression in the power spectrum domain by symmetric microphone arrays,” *Proc. WASPAA*, pp. 54–57, Oct. 2007.
- [15] N. Ito, N. Ono, and S. Sagayama, “A blind noise decorrelation approach with crystal arrays on designing post-filters for diffuse noise suppression,” *Proc. ICASSP*, pp. 317–320, Apr. 2008.
- [16] A. Kurematsu, K. Takeda, Y. Sagisaka, S. Katagiri, H. Kuwabara, and K. Shikano, “ATR Japanese speech database as a tool of speech recognition and synthesis,” *Speech Commun.*, vol. 9, no. 4, pp. 357–363, Aug. 1990.