

日本音響学会 2015 年春季研究 発表会に参加して

大野 瑞季

Mizuki OHNO

情報メディア学専攻修士課程 2年

1. はじめに

私は、2015年3月16日から18日までの3日間、中央大学後楽園キャンパスで開催された日本音響学会2015年春季研究発表会に参加し、「空間クロススペクトルを用いた拡散性雑音抑圧」という題目で発表を行った。

2. 研究背景

近年、音声認識技術や音声会議記録の発展に伴い、雑音環境下で所望の話者の音声のみを抽出する技術が求められている。マイクロホンアレイを用いた雑音抑圧技術としてビームフォーマが知られており、マイクロホンアレイに対して単一の方向から到来する方向性雑音の抑圧に効果的であるが、複数の方向から到来する拡散性雑音が存在する環境では十分な効果を得ることができない。そこで、ビームフォーマの出力に対して拡散性雑音抑圧を目的としたWienerフィルタを適用する手法が提案されている。従来手法は素子間隔の大きい大規模アレイを用いた場合に拡散性雑音を抑圧可能であるが、小規模アレイの場合に抑圧性能が大幅に低下するという問題がある。そこで、小規模アレイを用いた場合に高い抑圧性能が得られる空間クロススペクトルに基づく拡散性雑音抑圧手法を提案する。

3. 空間クロススペクトルを用いた Wiener フィルタ推定

小規模アレイによる拡散性雑音抑圧を目的としたWienerフィルタ $H(\omega)$ の設計手法を提案する。提案手法は、拡散性雑音が音場全体に一様に分布しているという仮定に基づき、マイクロホンアレイ入力を

遅延処理によって空間ごとに強調した信号から空間クロススペクトルを算出し、拡散性雑音の推定を行う。雑音の推定には素子間で拡散性雑音が無相関であることを要さないため、小規模アレイによる実装が可能である。提案手法のWienerフィルタ $H(\omega)$ は、入力信号のパワースペクトルを $P_{s+n}(\omega)$ 、雑音のパワースペクトルを $P_n(\omega)$ と表すと、式(1)のように計算される。

$$H(\omega) = \frac{\hat{P}_{s+n}(\omega) - \hat{P}_n(\omega)}{\hat{P}_{s+n}(\omega)} \quad (1)$$

ここで、 $P_{s+n}(\omega)$ および $P_n(\omega)$ は、任意の角度 θ にビームを向けた遅延和フィルタ $\mathbf{W}_{DS}(\theta, \omega)$ によって強調された信号 $Y_{DS}(\theta, \omega)$ を用いて推定される。角度 θ ごとに設計した遅延和フィルタ $\mathbf{W}_{DS}(\theta, \omega)$ を用いて受信信号 $\mathbf{Z}(\omega)$ を式(2)のようにフィルタリングすると、角度ごとに強調された信号 $Y_{DS}(\theta, \omega)$ が得られ、 $\mathbf{W}_{DS}(\theta, \omega)$ は式(3)で計算される。拡散性雑音が音場に一樣に分布する場合、拡散性雑音のパワースペクトル $P_n(\omega)$ は式(4)のように $Y_{DS}(\theta, \omega)$ のクロススペクトルで計算され、入力信号のパワースペクトル $P_{s+n}(\omega)$ は式(5)のように $Y_{DS}(\theta, \omega)$ のパワースペクトルで計算される、

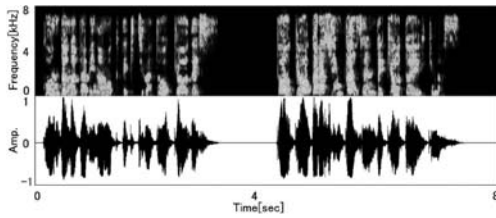
$$Y_{DS}(\theta, \omega) = \mathbf{W}_{DS}^H(\theta, \omega) \mathbf{Z}(\omega) \quad (2)$$

$$\mathbf{W}_{DS}(\theta, \omega) = \frac{\mathbf{a}(\theta, \omega)}{\mathbf{a}^H(\theta, \omega) \mathbf{a}(\theta, \omega)} \quad (3)$$

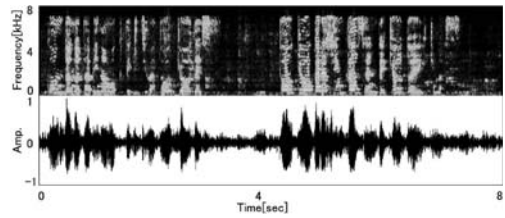
$$\hat{P}_n(\omega) = \frac{2}{(L-l)(L-l+1)} \cdot \sum_{i=1}^{L-1} \sum_{j=i+1}^L \Re\{Y_{DS}(\theta_i, \omega) Y_{DS}^*(\theta_j, \omega)\} \quad (4)$$

$$\hat{P}_{s+n}(\omega) = \frac{1}{L-l+1} \sum_{i=1}^L Y_{DS}(\theta_i, \omega) Y_{DS}^*(\theta_i, \omega) \quad (5)$$

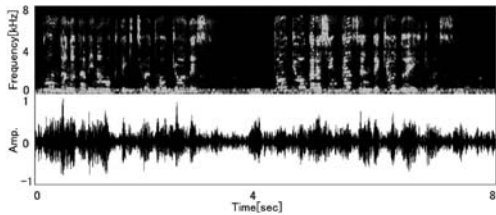
ここで、 θ_L および θ_1 はそれぞれ、制御範囲の最大角、最小角を表している。



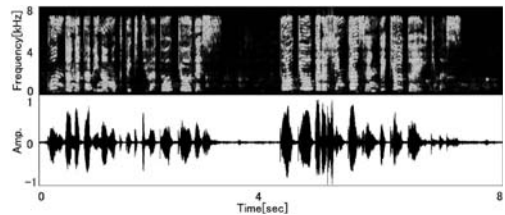
(a) クリーン音声



(b) 雑音付加音声 (SNR = 5 dB)



(c) 従来手法



(d) 提案手法

Fig. 1 各手法の波形とスペクトログラムによる評価

4. 性能評価

提案手法の性能を評価するため、実環境における実験を行った。目的音および雑音の収録は残響時間が約 700 ms の会議室で行い、雑音用のスピーカーを壁面に向けて再生することで拡散性雑音を再現した。また、雑音には室内環境で発生し得る電話や書類、話声などが合わさった音である office 雑音を用いた。評価には音声波形とスペクトログラムおよび評価指標である SD (Spectrum Distortion) と SNR 改善量を使用し、SD は音声歪み量、SNR 改善量は雑音抑圧量を表している。Fig. 1 に波形とスペクトログラムの例を示す。Zelinski の手法で処理した音声は雑音の低周波数成分が抑圧されておらず、高周波数成分は抑圧されているものの目的音まで抑圧していることがわかる。一方、提案手法においては、

周波数に関わらず雑音が抑圧されていることが確認できる。また、SNR 改善量は従来手法に比べ 15 dB、SD は約 4 dB 改善しており、小規模のマイクロホンアレイを用いて拡散性雑音の高精度な抑圧を実現している。今後は、高周波数帯域における抑圧性能と低周波数帯域における音声歪みのトレードオフに対するアプローチを行い、提案手法に適したマイクロホン配置の検討を行っていきたい。

5. おわりに

今回の発表を終えて、音響に関する知識をさらに広げ、様々なアドバイスを頂くことで自らの研究における課題の発見や意欲向上につなげることができました。最後に、研究や発表に関して多大なご指導を頂いた片岡章俊教授ならびに片岡研究室の皆様へ深く感謝いたします。