

電子情報通信学会 応用音響研究会に参加して

大野 瑞季
Mizuki OHNO

情報メディア学専攻修士課程 2015 年度修了

1. はじめに

私は、2015 年 11 月 12 日から 13 日までの 2 日間、熊本大学黒髪南キャンパスで開催された電子情報通信学会応用音響研究会に参加し、「異なる空間特性の複数雑音環境における空間ポストフィルタ推定」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。そこで、複数のマイクロホンを任意に配置した装置であるマイクロホンアレイを用いることで音波を物理的に解析し、所望の音だけを集音する技術の研究が行われている。本発表では、音場に存在する音を有指向性、無指向性のものに分解し、得られた有指向性/無指向性音響モデルから Wiener ポストフィルタを推定する手法の提案を行った。

3. 空間クロススペクトルを用いた空間ポストフィルタ推定

実環境の音場には図 1 のように、目的音以外に方向性雑音、拡散性雑音と呼ばれる空間特性の異なる雑音が存在する場合が多い。これらの雑音に頑健な目的音強調を実現するため、式 (1) のような Wiener ポストフィルタを設計する。

$$H = \frac{\phi_{ss}}{\phi_{ss} + \phi_{NN} + \phi_{mm}} \quad (1)$$

ここで、 ϕ_{ss} は目的音、 ϕ_{NN} は方向性雑音、 ϕ_{mm} は拡散性雑音の推定モデルである。提案する Wiener ポストフィルタは、目的音と方向性雑音を有指向性、

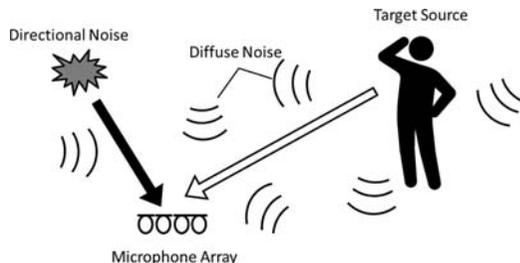


図 1 空間特性の異なる複数雑音環境

拡散性雑音を無指向性の音として音響モデル推定を行う。

3.1 無指向性モデル推定

無指向性の音が音場に一樣分布すると仮定すると、マイクロホンアレイの各チャンネルに様々な方向から到来すると考えられる。つまり、異なる方向の音を強調したビームフォーマ出力に含まれる共通成分を無指向性の音として推定できる。したがって、本研究で扱う無指向性の音である拡散性雑音は、式 (2) のように推定される。

$$\hat{\phi}_{mm} = \frac{2}{L(L-1)} \sum_{i=1}^{L-1} \sum_{j=i+1}^L \Re\{Y_{DS}(\theta_i)Y_{DS}^*(\theta_j)\} \quad (2)$$

ここで、 $Y_{DS}(\theta)$ はマイクロホンアレイの入力信号を遅延和フィルタによって任意の方向 θ にビームフォーミングした信号である。

3.2 有指向性モデル推定

無指向性モデル $\hat{\phi}_{mm}$ を用いて、有指向性である目的音および方向性雑音の音響モデル推定を行う。有指向性の音はマイクロホンアレイのチャンネル間で平面波の位相差が生じると仮定すると、ビームフォーマ出力 $Y_{DS}(\theta)$ から推定できる。また、 $Y_{DS}(\theta)$ から拡散性雑音成分を低減するため、式 (2) のようにビームフォーマ出力をフィルタリングする。

$$Y'_{DS}(\theta) = H'Y_{DS}(\theta) \quad (3)$$

ここで、フィルタ H' は以下の式で求められる拡散性雑音抑圧フィルタである。

$$H' = (\phi_{zz} - \phi_{mm})/\phi_{zz} \quad (4)$$

$$\hat{\phi}_{zz} = \frac{1}{L} \sum_{i=1}^L Y_{DS}(\theta_i) Y_{DS}^*(\theta_i) \quad (5)$$

次に、拡散性雑音成分を低減した各ビームフォーマ出力 $Y'_{DS}(\theta)$ から空間コヒーレンス関数を算出し、目的音および方向性雑音のモデル化を行う。空間コヒーレンス関数 $|\gamma(\theta)|^2$ は目的音源方向 θ_s に関するビームフォーマ出力と他方向に関するビームフォーマ出力の相関である。この性質から目的音と方向性雑音の SNR を算出することができ、各ビームフォーマ出力 $Y'_{DS}(\theta)$ の目的音に関する SNR を $\eta_s(\theta)$ 、方向性雑音に関する SNR を $\eta_n(\theta)$ とすると、それぞれ式 (6) と式 (7) のように表すことができる。

$$\eta_s(\theta) = \frac{|\gamma(\theta)|^2}{1 - |\gamma(\theta)|^2} \quad (6)$$

$$\eta_n(\theta) = \frac{1 - |\gamma(\theta)|^2}{|\gamma(\theta)|^2} \quad (7)$$

以上で求めた $\eta_s(\theta)$ および $\eta_n(\theta)$ を用いて各ビームフォーマ出力 $Y'_{DS}(\theta)$ の重み付け平均を計算すると、目的音モデル ϕ_{ss} および方向性雑音モデル ϕ_{NN} が式 (8) と式 (9) で推定できる。

$$\hat{\phi}_{ss} = \frac{1}{L} \sum_{i=1}^L \eta_s(\theta_i) Y'_{DS}(\theta_i) Y'_{DS}{}^*(\theta_i) \quad (8)$$

$$\hat{\phi}_{NN} = \frac{1}{L} \sum_{i=1}^L \eta_n(\theta_i) Y'_{DS}(\theta_i) Y'_{DS}{}^*(\theta_i) \quad (9)$$

4. 性能評価

4.1 実験条件

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、方向性雑音用のスピーカはマイクロホンアレイ（チャンネル数 $M=8$ 、チャンネル間隔 5 [cm]）の方向に向け、拡散性雑音用のスピーカを壁面に向けて再生することで方向性／拡散性雑音環境を再現した。

4.2 実験結果

入力 SNR の変化に対する各手法の SNR 改善量を図 2 に示す。処理性能は MVDR、McCowan ポス

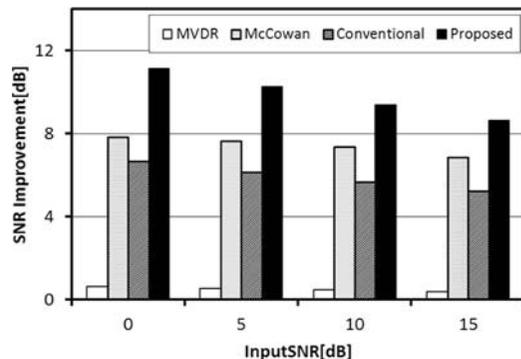


図 2 入力 SNR の変化に対する SNR 改善量

トフィルタ^[1]、従来手法^[2]、提案手法を比較して評価している。

図から、MVDR の SNR 改善量が非常に低く、方向性／拡散性雑音環境下では拡散性雑音の影響から方向性雑音を抑圧できていないことがわかる。提案手法は方向性雑音と拡散性雑音をそれぞれ抑圧対象として Wiener フィルタを設計しているため、従来手法より高い SNR 改善量を示している。また、McCowan ポストフィルタと比較しても優れた数値をとっており、VAD を使用せずにより高い抑圧性能が得られていることがわかる。

以上より、提案手法は方向性雑音と拡散性雑音を十分に抑圧でき、従来手法と比較しても優れた抑圧性能が得られることが確認できた。

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。

参考文献

- [1] I. McCowan, H. Bourland, TSA, pp.709-716, 2003
- [2] 大野瑞季, 片岡章俊, 信学技法, vol.115, no.126, EA 2015-3, pp.11-16, 2015. 7