

Title	小規模マイクロホンアレーを用いた複数雑音にロバストな雑音除去法に関する研究
Author(s)	籠, 貴司
Citation	
Issue Date	2001-09
Type	Thesis or Dissertation
Text version	author
URL	<a href="http://hdl.handle.net/10119/1549">http://hdl.handle.net/10119/1549</a>
Rights	
Description	Supervisor:赤木 正人, 情報科学研究科, 修士

# 小規模マイクロホンアレーを用いた 複数雑音にロバストな雑音除去法に関する研究

籠 貴司

北陸先端科学技術大学院大学 情報科学研究科

2001年8月15日

キーワード: 小規模マイクロホンアレー, 減算型ビームフォーマ, 複数雑音, 周波数帯域分割型アルゴリズム.

## 1 はじめに

雑音除去技術の需要はデジタル音響処理技術の発展に伴い益々高まっており、近年では指向特性の設計(ビームフォーミング)が可能なマイクロホンアレーを用いた雑音除去法も数多く提案されている。しかしながら、シミュレーションなどでは有効な手法であっても、実環境下においてはその雑音除去精度が著しく低下してしまう。これは以下の影響が原因であると考えられる。

- (1) 突発的な非定常雑音
- (2) 異なる方向から到来する複数雑音
- (3) 反射音による残響

(1)の問題に関して水町らは、3ch. 解析型アレーを用いた手法を提案した [1]。この手法は時々刻々推定した雑音成分を受音信号から減算するため、非定常雑音に対してロバストな雑音除去法である。しかしながら、雑音源を1つと仮定しているため、複数雑音には対応していない。

そこで本研究では、周波数帯域を分割し、各帯域で水町らの雑音除去アルゴリズムを繰り返し適用することにより、特に複数雑音に対応可能な雑音除去法の構築を試みる。これによって非定常雑音および複数雑音の両方にロバストで、より実環境に対応した雑音除去法が確立できる。またハンズフリー型 ASR (自動音声認識器) のフロントエンドへの応用などを想定し、雑音除去アルゴリズムを小規模マイクロホンアレーで実現する。

## 2 雑音除去アルゴリズム

以下に本雑音除去アルゴリズムの大まかな手順を示す。基本的なアルゴリズムは水町らの提案法とほぼ同等であるが、周波数帯域を分割し、各帯域で最も優勢な雑音スペクトルを推定することが本手法の特徴である。

## 2.1 雑音の到来時間差推定

目的信号の到来方向は正中面であるとする。ある周波数帯域  $\omega_{k-1} < |\tilde{\omega}| \leq \omega_{k+1}$  ( $k = 1 \sim K$ ) において、隣接マイクロホン間の到来時間差が  $\delta_k$  となる雑音  $N_k(\tilde{\omega})$  を想定する。3ch. マイクロホンアレーにおけるマイクロホンの受信信号  $l(t)$  (左側) と  $c(t)$  (中央)、および  $c(t)$  と  $r(t)$  (右側) より目的音を完全に抑圧した信号 (減算型ビームフォーマ)  $G_{lc}(\tilde{\omega})$ 、 $G_{cr}(\tilde{\omega})$  を計算する。

$$G_{xy}(\tilde{\omega}) = \text{FFT} \left[ \frac{\{x(t+\tau) - x(t-\tau)\} - \{y(t+\tau) - y(t-\tau)\}}{4} \right] \quad \omega_{k-1} < |\tilde{\omega}| \leq \omega_{k+1} \quad (1)$$

ここで  $xy$  の組み合わせは  $lc$ 、 $cr$ 、 $lr$  のいずれかである。また  $\tau$  は 0 以外の任意の定数であり、焦点 (信号を抽出する方向) を制御することができる。 $G_{lc}(\tilde{\omega})$ 、 $G_{cr}(\tilde{\omega})$  より、白色化相互相関関数を用いて到来時間差  $\delta_k$  を推定する。

## 2.2 雑音スペクトルの推定

次に推定した雑音  $N_k(\tilde{\omega})$  の到来時間差  $\delta_k$  と減算型ビームフォーマ  $G_{lr}(\tilde{\omega})$ 、 $G_{cr}(\tilde{\omega})$  を用いて、中央のマイクロホン位置における雑音のスペクトルを推定する。 $G_{lr}(\tilde{\omega})$  において  $\tau = \delta_k$ 、 $G_{cr}(\tilde{\omega})$  において  $\tau = \frac{\delta_k}{2}$  を代入し、焦点を雑音方向に合わせることにより

$$G_{lr}(\tilde{\omega}) = N_k(\tilde{\omega}) \sin^2 \tilde{\omega} \delta_k \quad \omega_{k-1} < |\tilde{\omega}| \leq \omega_k \quad (2)$$

$$G_{cr}(\tilde{\omega}) = N_k(\tilde{\omega}) e^{j\tilde{\omega} \frac{\delta_k}{2}} \sin^2 \tilde{\omega} \frac{\delta_k}{2} \quad \omega_{k-1} < |\tilde{\omega}| \leq \omega_k \quad (3)$$

となるが、 $G_{lr}(\tilde{\omega})$  においては  $\tilde{\omega} = \frac{n\pi}{\delta_k}$  ( $n$ : 整数)、 $G_{cr}(\tilde{\omega})$  においても  $n$  が偶数となる周波数成分付近は正確に雑音のスペクトルが推定できない。よって 2 つの閾値  $\varepsilon_1$ 、 $\varepsilon_2$  を設けて、次式により雑音のスペクトル  $\hat{N}_k(\tilde{\omega})$  を推定する。

$$\hat{N}_k(\tilde{\omega}) = \begin{cases} G_{lr}(\tilde{\omega}) / \sin^2 \tilde{\omega} \delta_k & \sin^2 \tilde{\omega} \delta_k > \varepsilon_1 \\ G_{cr}(\tilde{\omega}) e^{-j\tilde{\omega} \frac{\delta_k}{2}} / \sin^2 \tilde{\omega} \frac{\delta_k}{2} & (\sin^2 \tilde{\omega} \delta_k \leq \varepsilon_1) \wedge (\sin^2 \tilde{\omega} \frac{\delta_k}{2} > \varepsilon_2) \\ G_{lr}(\tilde{\omega}) & \sin^2 \tilde{\omega} \frac{\delta_k}{2} \leq \varepsilon_2 \end{cases} \quad (4)$$

以上の雑音の到来時間差推定およびスペクトル推定を、各周波数帯域で繰り返し行い、各帯域で最も優勢な雑音スペクトル  $\hat{N}_1(\tilde{\omega}) \sim \hat{N}_K(\tilde{\omega})$  を求め、それらを統合することにより全帯域の雑音スペクトル  $\hat{N}(\omega)$  とする [2]。

## 2.3 雑音スペクトルの除去

最後に Boll のスペクトルサブトラクション (SS) [3] を改良した非線形 SS を用いて、中央のマイクロホンにおける受信信号の振幅スペクトル  $|C(\omega)|$  から、推定した雑音の振幅スペクトル  $|\hat{N}(\omega)|$  を減算することにより雑音を除去する。

$$|\hat{S}(\omega)| = \begin{cases} |C(\omega)| - \alpha |\hat{N}(\omega)| & |C(\omega)| \geq \alpha |\hat{N}(\omega)| \\ \beta |C(\omega)| & \text{otherwise} \end{cases} \quad (5)$$

ここで  $\alpha$ 、 $\beta$  は定数であり、それぞれサブトラクション係数、フロアリング係数と呼ばれる。

### 3 雑音除去実験

本雑音除去アルゴリズムの精度検証のため、複数雑音に対する雑音除去実験を行った。受信系は隣接マイクロホン間隔が 10 cm の 3ch. 等間隔直線配列マイクロホンアレーである。

#### 3.1 実験条件

実験は防音室（残響時間 高周波数帯域：約 50 msec、500 Hz 以下の低周波数帯域：100 ~ 300 msec）で行い、目的信号を放射するスピーカー（直径 10 cm）を正中面方向（距離 0.5 m）、雑音 1、2 を放射するスピーカーをそれぞれ左 45 °、右 30 °方向（距離は共に 2.0 m）に設置した。目的信号として ATR 音声データベース男性話者 mau の/aoi/を用い、雑音 1 として RWCP 実環境音声・音響データベース非音声ドライソースの「ガスプレーの噴射音」（4~8 kHz にわたる帯域雑音）雑音 2 として「携帯電話の着信音」（約 2.8、5.6 kHz 付近に大きなパワーを持つ断続音）を用いた。これらを同時に音場に放射したものをサンプリング周波数 48 kHz、16 bit 量子化でデータ収録した。またパラメーター設定値は表 1 の通りであり、周波数帯域は最も細かく分割し、48000/1024 Hz ごとに雑音の到来時間差およびスペクトルの推定を行った。

表 1: パラメーター設定値

パラメーター	設定値
フレーム長	1024 pt (21.3 msec)
フレーム周期	512 pt (10.6 msec)
窓関数	Hanning window
サブトラクション係数 $\alpha$	1.0
フロアリング係数 $\beta$	0.001
閾値 $\varepsilon_1$	0.5
閾値 $\varepsilon_2$	0.2

また雑音除去精度の評価尺度としては、SNR、LSD を以下の式を用いて求めた。LSD は音声区間のフレームのみで計算し、平均をとった値とした。 $s(t_n)$ 、 $S(\omega)$  はクリーンな音声、 $\hat{s}(t_n)$ 、 $\hat{S}(\omega)$  は評価対象の音声を表す。単位は共に dB であり、 $W = 6$  kHz とした。

$$\text{SNR} = 10 \log_{10} \frac{\sum_n s^2(t_n)}{\sum_n \{s(t_n) - \hat{s}(t_n)\}^2} \quad (6)$$

$$\text{LSD} = \sqrt{\frac{1}{W} \sum_{\omega}^W \left( 20 \log_{10} \frac{|S(\omega)|}{|\hat{S}(\omega)|} \right)^2} \quad (7)$$

#### 3.2 実験結果

図 1 に雑音除去結果を示す。上図から順に (a) クリーンな音声、(b) 雑音を付加した音声 (c) 本手法で推定した雑音、(d) 本手法による雑音除去後の音声、(e) 水町らの提案法による雑音除去後

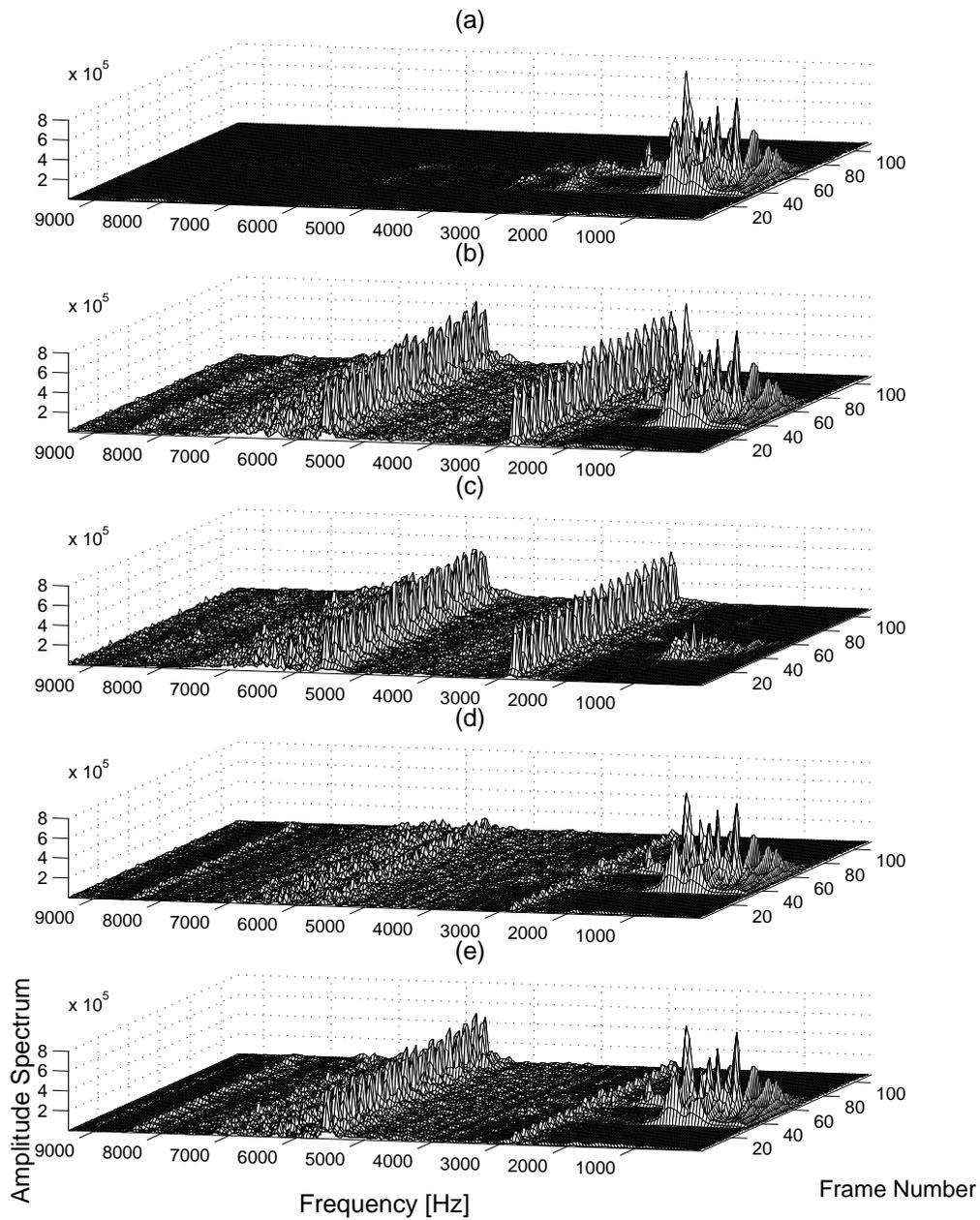


図 1: 雑音除去結果 —スペクトル— (a) クリーンな音声 (b) 雑音を付加した音声 (c) 推定した雑音 (本手法) (d) 雑音除去後の音声 (本手法) (e) 雑音除去後の音声 (水町らの提案法)

の音声のスペクトルを表す。ここでクリーンな音声は、音声のみをスピーカーから音場に放射し、中央のマイクロホンで収録したものとした。

本手法においては、2つの雑音スペクトルがはっきりと推定されており(図1(c))、ほとんど雑音が除去されたといえる(図1(d))。しかしながら、水町らの提案法においては、本手法では除去されている携帯電話の着信音の5.6 kHz成分が除去されていないことがわかる(図1(e))。

また表2に雑音除去前および雑音除去後のSNR、LSDの値を示す。SNRを約10 dB増加させ、LSDを約5 dB減少させたことから、本手法が複数雑音に対して有効に働いたといえる。また水町らの提案法との値の比較により、本手法の優位性も確認できた。

表 2: SNR、LSD による雑音除去評価 (LSD は音声区間のみで計算した値)

	雑音除去前	雑音除去後 (本手法)	雑音除去後 (水町らの提案法)
SNR [dB]	-4.0	6.5	0.9
LSD [dB]	17.9	12.8	14.4

## 4 まとめ

周波数帯域分割型アルゴリズムを水町らの提案法に導入した、複数雑音にロバストな雑音除去法を構築した。また雑音除去実験により、複数雑音に対する本手法の有効性、および水町らの提案法との比較による優位性を示した。

## 参考文献

- [1] 水町 光徳, “マイクロホン対を用いた雑音除去法に関する研究”, JAIST 博士論文 (2000)
- [2] Masato Akagi and Mitsunori Mizumachi, “NOISE REDUCTION BY PAIRED MICROPHONES”, Proc.EUROSPEECH '97, Rhodes, Greece, Vol.1, pp.335-338 (1997)
- [3] S.F.Boll, “Suppression of Acoustic Noise in Speech Using Spectral Subtraction”, IEEE Trans.ASSP, Vol.27, No.2, pp.113-120 (1979)