

Internal Use Only (非公開)

TR-SLT-0003

小規模マイクロホンアレーを用いた雑音除去システムの
構築

A noise reduction system using a small-scale
microphone array

野本 克
Katsumi Nomoto

水町 光徳
Mitsunori Mizumachi

2002.2.22

本稿では、実環境で有効な雑音除去技術の確立を目指し、小規模マイクロホンアレーを用いた雑音除去アルゴリズム最適化し、処理速度を向上させることを試みた。アルゴリズム最適化の結果、目的信号と雑音信号がともに音声である場合に、ほぼ実時間で雑音除去処理を実現し、あらゆる雑音到来方向に対して約 10dB の SNR 改善量が得られることがわかった。

(株) 国際電気通信基礎技術研究所

音声言語コミュニケーション研究所

〒 619-0288 京都府相楽郡精華町光台二丁目 2 番地 2

TEL: 0774-95-1301 FAX: 0774-95-1308

Advanced Telecommunication Research Institute International

Spoken Language Translation Research Laboratories

2-2-2 Hikaridai Seika-cho Souraku-gun Kyoto 619-0288 Japan

TEL: +81-774-95-1301 FAX: +81-774-95-1308

©2002 (株) 国際電気通信基礎技術研究所

©2002 Advanced Telecommunication Research Institute International

目次

1	背景及び目的	1
2	雑音除去アルゴリズムの概要	2
2.1	受信信号の定式化	2
2.2	目的信号及び雑音の到来方向推定	2
2.3	雑音スペクトル推定	3
2.4	雑音除去	4
3	雑音除去アルゴリズムの最適化	5
3.1	サンプリング周波数と雑音除去性能との関係	5
3.2	方向推定精度と雑音除去性能との関係	9
	(3.2.1) 目的信号の到来方向のみ固定した場合	9
	(3.2.2) 目的信号及び雑音の到来方向を固定した場合	10
3.3	ビームフォーマ実装法の最適化	13
3.4	使用プログラミング言語の最適化	14
4	最適化アルゴリズムの雑音除去性能	15
5	まとめ	17
6	今後の課題	18
	参考文献	19

1 背景及び目的

様々な音響機器にとって、環境雑音は脅威的存在であり、雑音の影響を低減することは必要不可欠な課題である。例えば、音声認識システムや補聴器などは、雑音により性能が大幅に低下することが知られている。音声認識システムの場合、認識対象の音声以外はすべて雑音であり、それらを除去することが可能であれば、実環境における音声認識率の飛躍的な向上が期待できる。また、補聴器の場合、目的信号と雑音を区別し、雑音を除去して目的信号のみを強調することができれば、明瞭度が向上することにより補聴器の普及にも貢献できる。また、実環境で有効な雑音除去システムを構築する場合、信号に対する制約条件が少なく、演算量が少なく、ハードウェアも小規模であることが望ましい。

本稿では、先に提案されている小規模マイクロホンアレーを用いた雑音除去アルゴリズム [1] の適用範囲を限定してアルゴリズムの最適化を行い、雑音除去性能の低下を最小限に抑え、演算量を削減することを目的とする。つまり、雑音除去性能と演算量との相互関係から、雑音除去アルゴリズムの最適化を行う。従来のアルゴリズムは、移動音源にも正確に対応することを目指しており、MATLABにより実装した結果、演算時間は実時間の40倍～50倍であった。本稿では、本アルゴリズムの音声認識システムの雑音除去フロントエンドとしての利用を想定し、ラップトップPCやPDAなどの携帯機器への搭載を前提として最適な雑音除去アルゴリズムについて検討する。

2 雑音除去アルゴリズムの概要

本雑音除去アルゴリズムは、直線上に等間隔で配置した3つのマイクロホンを用い、信号の各マイクロホン対への到来時間差を手掛かりに雑音除去を行う [1]。まず受信信号の相互相関より目的信号及び雑音の到来方向を推定し、次に減算型ビームフォーミング技術により雑音スペクトルを推定し、最後にスペクトルサブトラクション [2] により雑音成分を除去する。以下に、その概要を説明する。

2.1 受信信号の定式化

目的信号を $s(t)$ 、雑音を $n(t)$ 、マイクロホンアレーの3つのマイクロホンのうち左端、中央、右端のマイクロホンで受信した信号をそれぞれ $l(t)$ 、 $c(t)$ 、 $r(t)$ とする。ここで、目的信号及び雑音が、左端と中央のマイクロホンから構成されるマイクロホン対へ到来するときの時間差を ζ 及び δ とすると、以下の関係が成り立つ。

$$l(t) = s(t - \zeta) + n(t - \delta) \quad (1)$$

$$c(t) = s(t) + n(t) \quad (2)$$

$$r(t) = s(t + \zeta) + n(t + \delta) \quad (3)$$

このとき、 $l(t)$ 、 $c(t)$ 、 $r(t)$ のフーリエ変換 $L(\omega)$ 、 $C(\omega)$ 、 $R(\omega)$ は

$$L(\omega) = S(\omega)e^{-j\omega\zeta} + N(\omega)e^{-j\omega\delta} \quad (4)$$

$$C(\omega) = S(\omega) + N(\omega) \quad (5)$$

$$R(\omega) = S(\omega)e^{j\omega\zeta} + N(\omega)e^{j\omega\delta} \quad (6)$$

となる。ここで、 $S(\omega)$ 及び $N(\omega)$ は、目的信号及び雑音のスペクトルである。

2.2 目的信号及び雑音の到来方向推定

信号の到来方向は、信号のマイクロホン対への到来時間差と考えて、受信信号の相互相関より推定することが可能である。そこで、両端のマイクロホンで受信した信号より白色化相互相関関数 $\phi_{lr}(t)$ を計算し、まずパワーが最大の信号が、マイクロホンアレーの両端マイクロホンにより構成されるマイクロホン対へ到来するときの時間差 d_1 を推定する。

$$\phi_{lr}(t) = \text{IFFT} \left[\frac{L(\omega)R^*(\omega)}{|L(\omega)||R(\omega)|} \right] \quad (7)$$

$$d_1 = \underset{t}{\text{argmax}} [\phi_{lr}(t)] \quad (8)$$

次に、2番目にパワーの大きな信号の到来時間差 d_2 を推定する。これは、白色化相互相関関数 $\phi_{lr}(t)$ から推定可能ではあるが、最もパワーの大きな信号の影響を強く受ける。そこで、

信号の到来方向推定に減算型ビームフォーマを導入する。具体的には、マイクロホンアレーの左端と中央、中央と右端のマイクロホンで構成される2つのマイクロホン対に対して、先に推定した d_1 方向に死点を持つような減算型ビームフォーマをそれぞれ設計し、それら出力信号を用いて白色化相互相関関数を計算し、2番目にパワーの大きな信号の到来時間差 d_2 を推定する。例えば、左端と中央のマイクロホンで構成されるマイクロホン対に対しては、減算型ビームフォーマ $g_{lc}(t)$ は以下のように設計すればよい。

$$g(t) = \frac{\{l(t + d_1 + \tau_0) - l(t + d_1 - \tau_0)\} - \{c(t + \tau_0) - c(t - \tau_0)\}}{4} \quad (9)$$

τ_0 は0以外の任意定数とする。

ここで、目的信号と雑音は同一方向から到来せず、目的信号よりも大きな雑音は2つ以上は存在せず、目的信号及び雑音の初期到来方向のうち正中面方向により近いものを目的信号の到来方向と仮定する。これより、短時間フレームごとに、目的信号及び雑音の到来方向を推定することができる。

2.3 雑音スペクトル推定

一般に、マイクロホン数が少ない場合、加算型ビームフォーマは極めてブロードな指向特性しか形成できないが、減算型ビームフォーマは鋭い死角を形成できる。そこで、本雑音除去アルゴリズムは、マイクロホン対での受音信号より単一の死角を持つビームフォーマを設計し、これを用いて雑音スペクトルを推定する。雑音スペクトルを推定するためには、目的信号の方向に死角を形成すればよい。まず、マイクロホンアレーの両端のマイクロホンで構成されるマイクロホン対の場合、以下のように減算型ビームフォーマ $g_{lr}(t)$ を設計する。

$$g_{lr}(t) = \frac{\{l(t + \zeta + \tau_1) - l(t + \zeta - \tau_1)\} - \{r(t - \zeta + \tau_1) - r(t - \zeta - \tau_1)\}}{4} \quad (10)$$

τ_1 は0以外の任意定数である。減算型ビームフォーマ $g_{lr}(t)$ のフーリエ変換 $G_{lr}(\omega)$ は、

$$G_{lr}(\omega) = N(\omega)\sin\omega(\delta - \zeta)\sin\omega\tau_1 \quad (11)$$

となり、式(11)は目的信号を含んでいない、すなわち目的信号が完全に抑圧されていることが確認できる。ここで、 τ_1 は0以外の任意定数であるため

$$\tau_1 = \delta - \zeta \quad (12)$$

とすると、

$$G(\omega) = N(\omega)\sin^2\omega(\delta - \zeta) \quad (13)$$

となり、雑音スペクトル $N(\omega)$ は、

$$N(\omega) = G(\omega)/\sin^2\omega(\delta - \zeta), \quad (14)$$

$$\omega(\delta - \zeta) \neq n\pi \quad (n: \text{整数}) \quad (15)$$

と推定できる。但し、式(15)は、雑音スペクトルのうち

$$\omega = \frac{n\pi}{\delta - \zeta} \quad (n: \text{整数}) \quad (16)$$

となる周波数成分は推定できない。そこで、 $l(t)$ と $c(t)$ を用いて目的信号の到来方向に死角を持つ新たなビームフォーマ $g_{cr}(t)$ を設計する。

$$g_{cr}(t) = \frac{\{c(t + \tau_2) - c(t - \tau_2)\} - \{r(t - \zeta + \tau_2) - r(t - \zeta - \tau_2)\}}{4} \quad (17)$$

ここで、 τ_2 は0以外の任意定数であるため、

$$\tau_2 = \frac{\delta - \zeta}{2} \quad (18)$$

と定めると、ビームフォーマ $g_{cr}(t)$ の周波数表現は

$$G_{cr}(\omega) = N(\omega)e^{j\omega\frac{\delta-\zeta}{2}} \sin^2\omega\frac{\delta-\zeta}{2} \quad (19)$$

となる。ビームフォーマ $g_{cr}(t)$ を用いて雑音スペクトルが推定できない周波数帯域は

$$\omega = \frac{2m\pi}{\delta - \zeta} \quad (m: \text{整数}) \quad (20)$$

となる。

本雑音除去アルゴリズムは、各周波数において、ビームフォーマ $g_{lr}(t)$ とビームフォーマ $g_{cr}(t)$ のより適切な方を自動選択することにより、高精度の雑音スペクトル推定を実現している。具体的には、 ϵ_1 と ϵ_2 の2つの閾値を設け、式(21)により雑音スペクトル $\hat{N}(\omega)$ を推定する。

$$\hat{N}(\omega) = \begin{cases} G_{lr}(\omega)/\sin^2\omega(\delta - \zeta) & \sin^2\omega(\delta - \zeta) > \epsilon_1 \\ G_{cr}(\omega)e^{-j\omega\frac{\delta-\zeta}{2}}/\sin^2\omega\frac{\delta-\zeta}{2} & \sin^2\omega(\delta - \zeta) \leq \epsilon_1 \quad \text{and} \quad \sin^2\omega\frac{\delta-\zeta}{2} > \epsilon_2 \\ G_{lr}(\omega)/\epsilon_2^2 & \sin^2\omega\frac{\delta-\zeta}{2} \leq \epsilon_2 \end{cases} \quad (21)$$

2.4 雑音除去

雑音スペクトルは、式(21)によって短時間フレームごとに正確に推定できる。そこで、中央のマイクロホンでの受信信号 $c(t)$ の振幅スペクトルから、先に推定した雑音の振幅スペクトル $|\hat{N}(\omega)|$ を引きさることにより、目的信号の振幅スペクトル $|\hat{S}(\omega)|$ が得られる。なお、雑音スペクトルの推定において、推定誤差が生じる可能性もあるため、サブトラクション係数 α とフロアリング係数 β を導入した非線形スペクトルサブトラクションを採用する。

$$|\hat{S}(\omega)| = \begin{cases} |C(\omega)| - \alpha \cdot |\hat{N}(\omega)|, & |C(\omega)| \geq \alpha \cdot |\hat{N}(\omega)| \\ \beta \cdot |C(\omega)|, & \text{otherwise} \end{cases} \quad (22)$$

3 雑音除去アルゴリズムの最適化

本章では、第2章で概説した雑音除去アルゴリズムを携帯型音声認識システムの雑音除去フロントエンドとして利用することを前提に、演算量を削減するための最適化アルゴリズムについて検討する。ここでは主に、サンプリング周波数と雑音除去性能との関係、方向推定精度と雑音除去性能との関係に着目する。

3.1 サンプリング周波数と雑音除去性能との関係

従来のアルゴリズムでは、受信信号のサンプリング周波数を48kHzと設定して性能評価を行っていた。しかし、現在の音声認識システムを想定した場合、有効周波数帯域は8kHzまでで十分であるため、サンプリング周波数を16kHzへと変更する。本雑音除去アルゴリズムは、信号のマイクロホン対への到来時間差を手掛かりとするため、サンプリング周波数を低くすると空間分解能も粗くなり、雑音除去性能が変化することが予想される。

そこで、まずサンプリング周波数と空間分解能との関係を調査する。表1に、信号到来方向 [degree] と、各サンプリング周波数における信号の隣接マイクロホン間への到来時間差 [pt]¹ との関係を示す。

表 1: サンプリング周波数と空間分解能との関係

arrival direction of signal [degree]	4	8	12	16	21	25	30	35	40	45	51	58	67	83
time difference in 48kHz [pt]	1	2	3	4	5	6	7	8	9	10	11	12	13	14
time difference in 16kHz [pt]		1			2			3			4			

次に、サンプリング周波数の変更による雑音除去性能の変化について検討する。本稿では、雑音除去前後の信号から算出する SNR 改善量を用いて雑音除去性能の評価を行う。

評価条件:

- 目的信号: 日本語母音 /a/、マイクロホンアレーの正中面方向 (0°) から到来
- 雑音信号: 125Hz ~ 1kHz、1kHz ~ 3kHz、3kHz ~ 5kHz のバンドパスノイズ、15° から 5° 刻みで 90° のいずれかから到来
- 信号継続時間長: 0.2s

¹1pt は 1 サンプル間隔を意味する。

- 入力信号の SNR: 0dB

本稿では、隣接マイクロホン間隔は 10cm、方向推定を行う際のフレーム長はサンプリング周波数に関わらず 1024pt、雑音スペクトル推定及び雑音除去を行う際のフレーム長は、サンプリング周波数が 48kHz の場合に 256pt、サンプリング周波数が 16kHz の場合に 64pt と設定する。一般に、フレーム長は大きいほど方向推定の精度は向上する。一方、スペクトルサブトラクションにより雑音除去を行う場合、フレーム長を大きくすると音声の子音部など時間変化の速い信号への対応が困難になる。そこで、本稿では、各処理において異なるフレーム長を設定した。また、式 (21) の閾値 ϵ_1 、 ϵ_2 はそれぞれ 0.5、0.3 とし、式 (22) のサブトラクション係数は 1.0、フロアリング係数は 0.001 とした。

評価結果:

受音信号のサンプリング周波数と SNR 改善量との関係を図 1～図 3 に示す。ここで、各図の横軸は雑音の到来方向、縦軸は SNR 改善量を表す。○印 (点線) はサンプリング周波数を 48kHz に設定した場合、□印 (実線) はサンプリング周波数を 16kHz に設定した場合の SNR 改善量を表す。

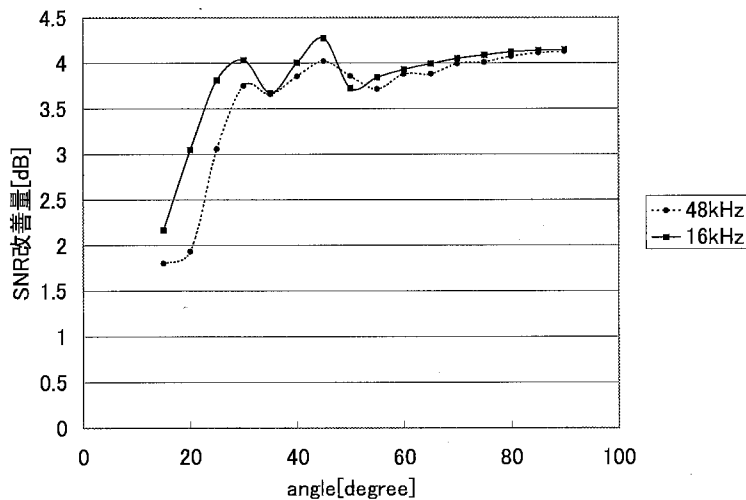


図 1: サンプリング周波数と SNR 改善量との関係 (バンドパスノイズ 125Hz ~ 1kHz の場合)

考察:

(a) 125Hz ~ 1kHz バンドパスノイズ (図 1) について

本周波数帯域 (低周波数帯域) に関しては、他の雑音周波数帯域と比較し、あらゆる雑音到来方向において SNR 改善量が小さい。この要因は、例えば式 (15) により雑音スペクトルを推定する際、分母の \sin 項が 0 に近い値となるためである。参考のため、 τ を信号の到来時間差 [pt] とした場合の周波数 f [Hz] と $\sin^2 2\pi f\tau$ との関係を図 4 に示す。サンプリング周波数は 48kHz

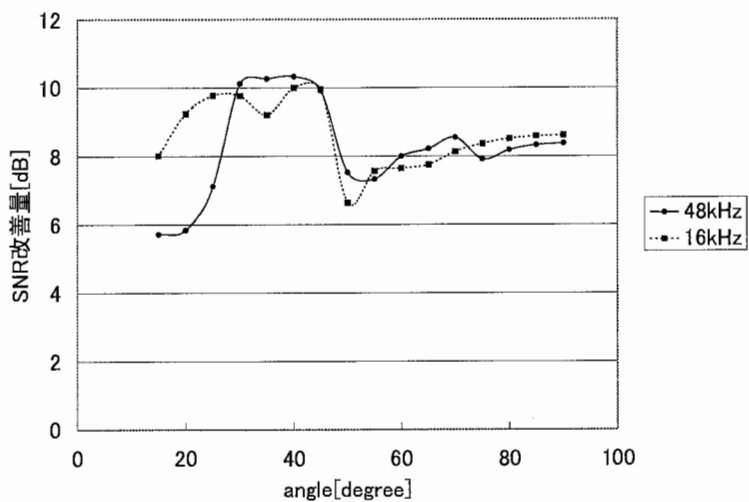


図 2: サンプリング周波数と SNR 改善量との関係 (バンドパスノイズ 1kHz ~ 3kHz の場合)

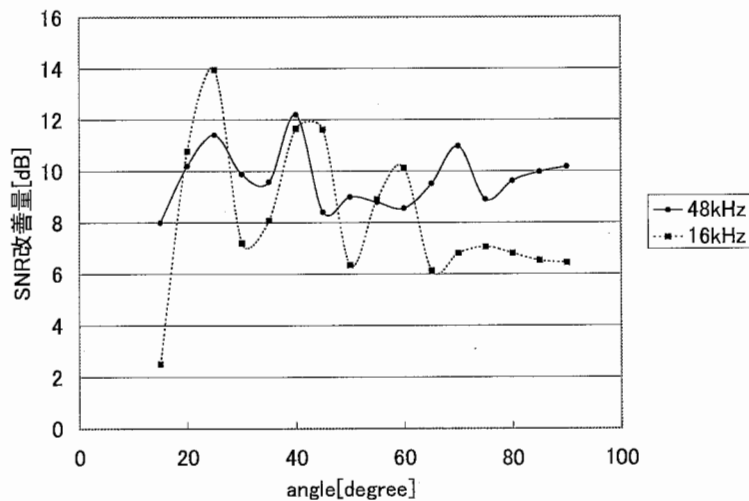


図 3: サンプリング周波数と SNR 改善量との関係 (バンドパスノイズ 3kHz ~ 5kHz の場合)

である。なお、本評価条件では、目的信号は正中面方向から到来するように設定しているため、 τ は式 (15) の雑音到来方向 δ に相当する。

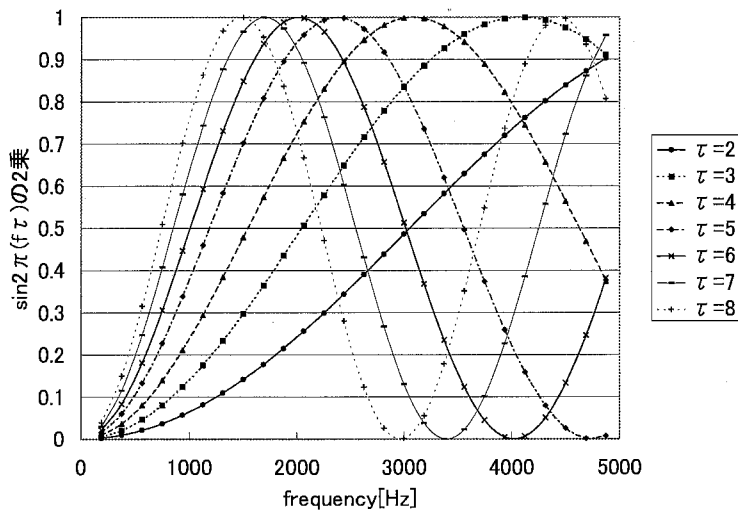


図 4: 信号到来時間差 τ [pt] と周波数 f [Hz] と $\sin^2 2\pi f\tau$ との関係

図 4 より、雑音到来方向が正中面方向に近い (τ の値が小さい) ほど、 $\sin^2 2\pi f\tau$ の値も小さくなることがわかる。つまり、式 (21) より雑音スペクトルを推定する際に、125Hz ~ 1kHz においては正確に推定ができない。逆に、図 4 の 125Hz ~ 1kHz においては、雑音到来方向が 90° に近づく (τ の値が大きくなる) につれ、 $\sin^2 2\pi f\tau$ の値が大きくなる。これより、図 1 の SNR 改善量が、雑音の到来方向が 90° に近づくにつれ大きくなっていることが納得できる。

(b) 1kHz ~ 3kHz バンドパスノイズ (図 2) について

本周波数帯域に関しては、雑音到来方向が $30^\circ \sim 90^\circ$ の範囲では、サンプリング周波数の差異による SNR 改善量の変化は 1dB 以内と非常に小さい。一方、雑音到来方向が $15^\circ \sim 25^\circ$ の範囲では、サンプリング周波数を 16kHz に変更すると、SNR 改善量が 2dB 程度大きくなることがわかる。

この要因の一つは、表 1 も示したように、空間分解能の差異である。サンプリング周波数が 16kHz と 48kHz との場合では、マイクロホン対への信号到来時間差の最小値が異なり、後者の方がより小さな値を取りうる。従って、雑音が目的信号方向へ近づくにつれ、図 4 より、サンプリング周波数が 48kHz の場合の方が雑音スペクトルの推定誤差が生じやすいと考えられる。

(c) 3kHz ~ 5kHz バンドパスノイズ (図 3) について

本周波数帯域では、サンプリング周波数を 16kHz へと変更すると、雑音到来方向に対する

雑音除去性能のばらつきが大きくなっている。雑音到来方向に対する SNR 改善量の変化量に着目すると、サンプリング周波数が 48kHz の場合は 4dB 以内であったが、サンプリング周波数が 16kHz の場合には 12dB と増大する。

本アルゴリズムは、高周波数帯域における空間的エイリアジング [3] の影響を回避するために、式 (21) において閾値 ϵ_1 、 ϵ_2 を導入している。そして、これらの設定値は、サンプリング周波数が 48kHz の場合の最適値であり、サンプリング周波数を 16kHz と変更した場合には最適値である保証はない。これが、雑音到来方向により SNR 改善量にばらつきが生じた要因であろうと考えられる。

まとめ:

一般に、サンプリング周波数を 48kHz から 16kHz へと変更すると、空間分解能は粗くなり、それぞれのビームフォーマの空間特性は平均的な特性となる。大局的には、これが低周波数帯域での SNR 改善量の増大、高周波数帯域での減少につながる事がわかった。本評価実験の結果、音声のパワーが集中する低周波数帯域 (3kHz 程度まで) に着目し、しかも雑音到来方向と目的信号到来方向が 30° 以上離れていれば、サンプリング周波数を 48kHz から 16kHz へと変更しても問題ないと考えられる。

なお、サンプリング周波数を 48kHz から 16kHz へと変更することにより、演算量は約 1/3 倍に削減できる。

3.2 方向推定精度と雑音除去性能との関係

第 2 章で概説した雑音除去アルゴリズムは、各フレームごとに 2 つの信号到来方向を推定していた。しかし、ラップトップ PC や PDA などをプラットフォームとする携帯型音声認識システムへの応用を考えた場合には、目的信号の到来方向を限定することが可能である。

方向推定アルゴリズムの簡略化は、演算量の削減が見込まれるため、目的信号の到来方向を固定した場合、さらに雑音の到来方向も固定した場合の雑音除去性能について検討する。後者に関しては、実際の雑音到来方向は未知であるため、設定した方向と異なる方向から雑音が到来する場合のミスマッチ条件における雑音除去性能の低減量を詳細に調査する。

(3.2.1) 目的信号の到来方向のみ固定した場合

評価条件:

- 目的信号: 日本語母音 /a/、マイクロホンアレーの正中面方向から到来
- 雑音信号: 3kHz ~ 5kHz バンドパスノイズ、 15° から 5° 刻みで 90° のいずれかから到来

- 信号継続時間長: 0.2s
- 入力信号の SNR: 0dB
- 入力信号のサンプリング周波数: 16kHz

評価結果:

図 5 に、各雑音到来方向に対する SNR 改善量を示す。○印 (実線) は目的信号の到来方向を既知として雑音の到来方向のみを推定した場合、□印 (点線) は目的信号と雑音の到来方向を推定した場合の SNR 改善量を表す。なお、図 5 では、両条件の SNR 改善量は完全に一致している。

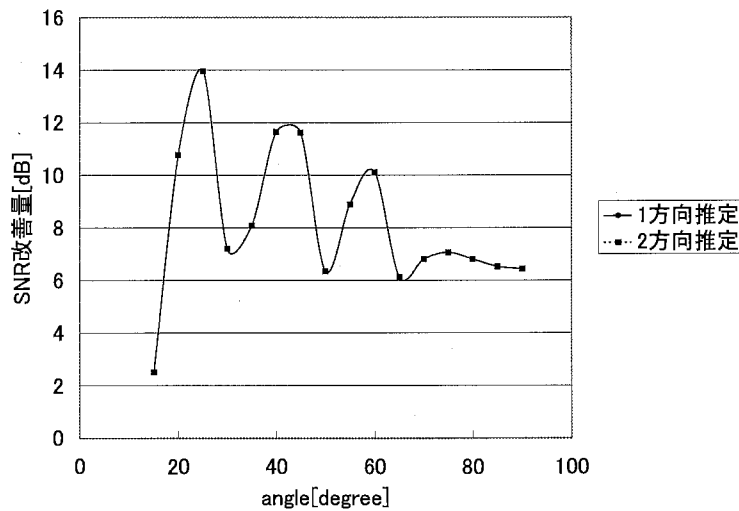


図 5: 方向推定対象音源数の変化による SNR 改善量の差異 (1 方向推定と 2 方向推定との比較)

考察及びまとめ:

図 5 より、目的信号の到来方向を既知として雑音到来方向のみを推定した場合と、目的信号及び雑音の到来方向を推定した場合では、雑音除去性能に全く差がないことがわかる。この結果は、2 信号の方向推定に全く誤差が生じていないことを意味する。但し、目的信号の到来方向を既知と、目的信号及び雑音の到来方向を推定する場合と比較して、演算量は約 5/6 倍に削減される。

(3.2.2) 目的信号及び雑音の到来方向を固定した場合

評価条件:

- 目的信号: 日本語母音 /a/、マイクロホンアレーの正中面方向から到来

- 雑音信号: 125Hz ~ 1kHz、1kHz ~ 3kHz、3kHz ~ 5kHz のバンドパスノイズ、
15° から 5° 刻みで 90° のいずれかから到来
- 信号継続時間長: 0.2s
- 入力信号の SNR: 0dB
- 入力信号のサンプリング周波数: 16kHz

表 1 より、本評価条件では雑音の到来時間差は 4 通り得られるが、ここでは最も正中面方向に近い方向を除外した 3 通りについて検討する。

評価結果:

目的信号の到来方向を正中面方向に固定し、雑音の到来方向 (図中の τ) を 3 方向のいずれかに固定した場合の雑音到来方向に対する SNR 改善量を図 6 ~ 図 8 に示す。各図の横軸は雑音の到来方向、縦軸は SNR 改善量を示し、○印 (実線) は雑音到来方向を $\tau = 2$ と固定した場合、□印 (点線) は $\tau = 3$ と固定した場合、△印 (破線) は $\tau = 4$ と固定した場合の SNR 改善量を表す。

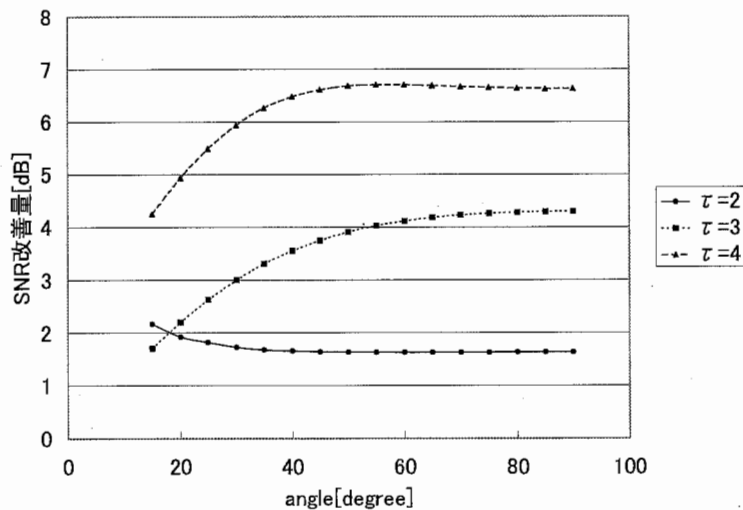


図 6: 目的信号及び雑音の到来方向を固定した場合の SNR 改善量 (バンドパスノイズ 125Hz ~ 1kHz の場合)

考察:

(a) 125Hz ~ 1kHz バンドパスノイズ (図 6) について

本周波数帯域における SNR 改善量は全般的に小さく、固定した雑音方向と実際の雑音到来方向との一致するか否かと SNR 改善量との関係に関しては、直接的な相関は見られない。こ

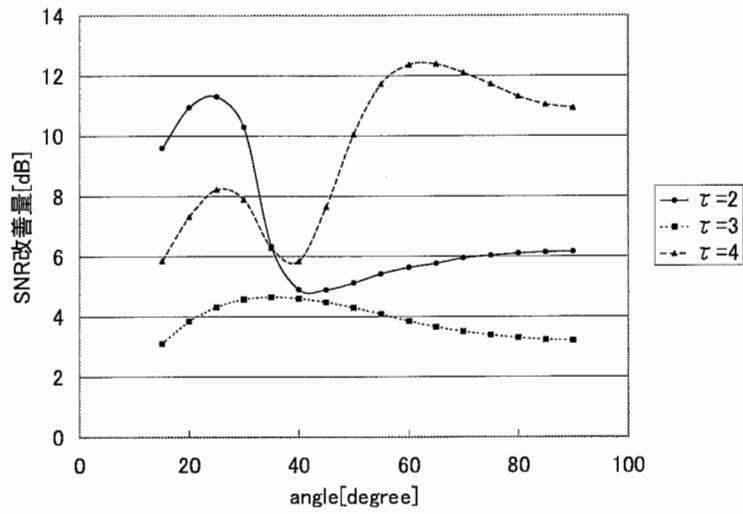


図 7: 目的信号及び雑音の到来方向を固定した場合の SNR 改善量 (バンドパスノイズ 1kHz ~ 3kHz の場合)

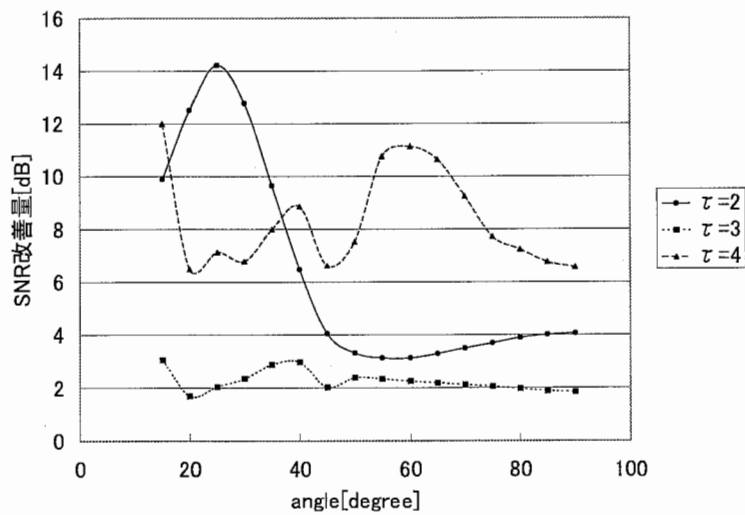


図 8: 目的信号及び雑音の到来方向を固定した場合の SNR 改善量 (バンドパスノイズ 3kHz ~ 5kHz の場合)

れは、図4からもわかるように、正確に雑音スペクトルを推定することが困難であるためである。また、雑音の到来方向が正中面方向から離れる (τ の値が大きい) ほど SNR 改善量も大きくなるが、これも図4により説明できる。

(b) 1kHz ~ 3kHz バンドパスノイズ (図7) について

本周波数帯域では、固定した雑音方向と実際の雑音到来方向とが一致した場合に SNR 改善量が最大となっている。但し、雑音方向として固定する方向 (τ の値) によって、SNR 改善量に大きな差が生じる。雑音方向を $\tau = 3$ と固定した場合、あらゆる雑音到来方向に対して雑音除去性能は最小となっている。これは普遍的な傾向ではなく、例えば式 (21) の ϵ_1 、 ϵ_2 などのパラメータに依存して変化するものであろうと考えられる。

(c) 3kHz ~ 5kHz バンドパスノイズ (図8) について

本周波数帯域においても、固定した雑音方向と実際の雑音到来方向に近い場合に SNR 改善量が大きくなっている。但し、 $\tau = 4$ の場合は、図4 及び式 (21) から、高周波数帯域において雑音スペクトルの推定精度が低下することが推測できる。従って、全周波数成分を評価対象とする SNR 改善量が安定しないのではないかと考えられる。

まとめ:

本評価実験の結果、雑音の到来方向を固定し、実際には異なった方向から到来する雑音に対しても、雑音除去の効果があることがわかった。なお、雑音の到来方向 (τ の値) を固定する場合には、すべての周波数帯域において $\tau = 4$ に固定した場合が最適であった。

ここで、評価対象周波数全域 (125Hz ~ 5kHz) における総合的な SNR 改善量についても調査し、雑音方向を $\tau = 4$ に固定した場合が最適であることを確認することが望ましい。この結果を図9に示しているが、全周波数帯域においても、各周波数帯域での結果と同様に雑音到来方向を $\tau = 4$ と固定した場合に SNR 改善量が最大となることがわかる。

なお、雑音の到来方向を固定する場合の演算量は、雑音の到来方向を推定する場合と比較して約 2/3 倍、目的信号及び雑音の到来方向を推定する場合と比較して約 1/2 倍に削減できる。

3.3 ビームフォーマ実装法の最適化

ここで、減算型ビームフォーマの実装手法について検討する。従来は、例えば式 (10) のように、時間領域において受信信号に適切な遅延を与えた信号を加減算し、ビームフォーマを実装していた。一方、初めに各受信信号のスペクトルを計算し、それらを用いて周波数領域でビームフォーマを実装することも可能であり、これにより演算量の低減が期待できる。目的信号及び雑音の到来方向を固定して方向推定を行わない場合、周波数領域でのビームフォーマ実装により、従来の時間領域での実装と比較して演算量を約 1/2 に削減できることを確認した。

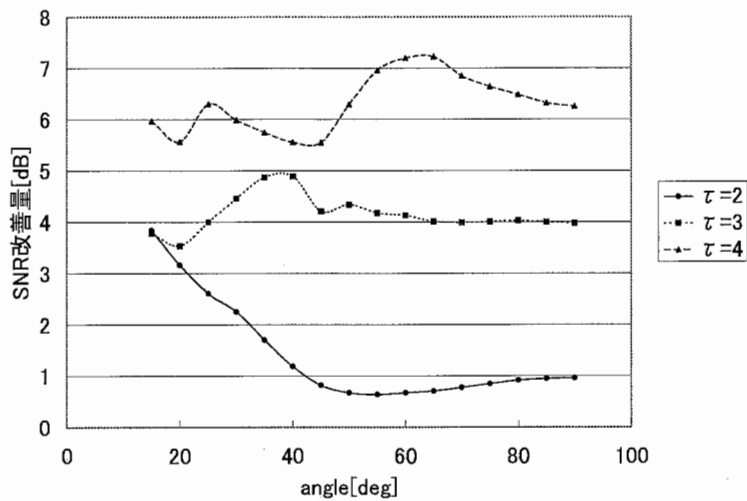


図 9: 目的信号及び雑音の到来方向を固定した場合の SNR 改善量 (全周波数帯域 125Hz ~ 5kHz の場合)

3.4 使用プログラミング言語の最適化

本雑音除去アルゴリズムは、研究に利用することを目的に、Mathworks 社の MATLAB を用いて実装していた。MATLAB は、行列演算に適したインタープリタ言語であるため、それに適した記述を心掛けることが重要である。同社は、MATLAB から C へのコード自動変換アプリケーションである MATLAB Compiler を提供しており、MATLAB Compiler を用いて MATLAB コードを C コードへと自動変換することにより、一般的にはプログラム実行時間の短縮が見込まれる。しかし、本雑音除去アルゴリズムの場合、MATLAB Compiler により自動生成された C コードを用いても実行時間を短縮することはできなかった。

そこで、本雑音除去アルゴリズムを自動変換ではなく、MATLAB コードを参考に C コードで再構築した結果、MATLAB での実行時間と比較し、ほぼ同等の処理を約 1/5 倍の時間で実行することが可能となった。

4 最適化アルゴリズムの雑音除去性能

第3章において、目的信号が正中面方向から到来するような状況を想定し、雑音除去アルゴリズムの最適化を行った。具体的には、目的信号の到来方向を正中面方向、雑音の到来方向を $\tau = 4$ に相当する方向に固定し、信号の到来方向推定を行わず、ビームフォーマの実装は周波数領域で行う。本章では、最適化した雑音除去アルゴリズムの性能評価を行う。

評価条件:

- 目的信号: 日本語母音 /a/、マイクロホンアレーの正中面方向から到来
- 雑音信号: 日本語母音 /u/、 15° から 5° 刻みで 90° のいずれかから到来
- 信号継続時間長: 0.2s
- 入力信号の SNR: 0dB
- 入力信号のサンプリング周波数: 16kHz

評価結果:

図10に、横軸を雑音の到来方向、縦軸を SNR 改善量として、最適化アルゴリズムの雑音除去性能を示す。

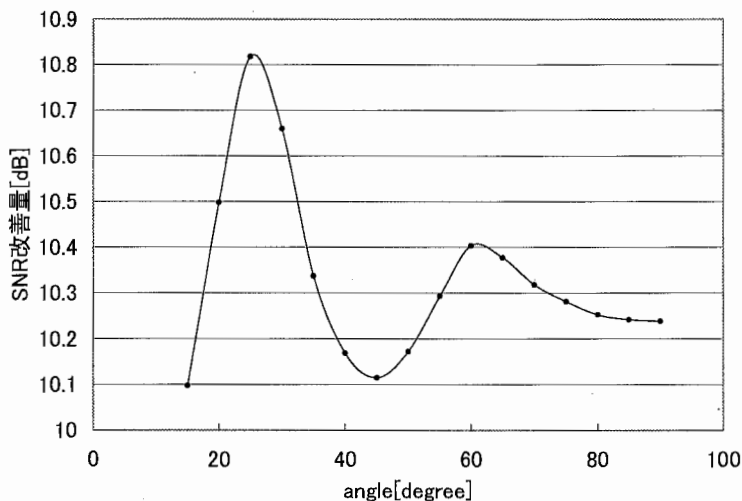


図 10: 最適化アルゴリズムの雑音除去性能

考察:

第3章では雑音として特定周波数帯域のランダム雑音を用いて評価を行ったが、本章では目的信号と相関の高い雑音を用いた。図10より、本評価条件においては、あらゆる雑音到来方向

に対して 10dB 以上の SNR 改善量が得られることがわかる。また、雑音到来方向に対する雑音除去性能のばらつきが 1dB 以下と非常に小さいが、これは雑音 (日本語母音 /u/) のパワーが低周波数帯域に集中しているためである。実環境での利用を想定した場合、音声妨害音としての雑音となる可能性は極めて大きいことから、図 10 は、本雑音除去アルゴリズムの実環境有効性を示唆している。

5 まとめ

本稿では、目的信号が正中面方向から到来する状況を想定し、小規模マイクロホンアレーを用いた雑音除去アルゴリズムの最適化を行い、雑音除去性能及び実行速度の評価を行った。まず、本雑音除去アルゴリズムを音声認識システムのフロントエンドとして利用することを考え、サンプリング周波数を 48kHz から 16kHz へと変更した。次に、従来は目的信号及び雑音の到来方向を推定していたが、その両者を固定した場合でも、雑音除去性能の低下は大きくないことを確認した。また、雑音除去処理に要する時間を短縮することを目的に、従来は時間領域で実装していたビームフォーマを周波数領域での実装へと変更し、プログラミング言語を MATLAB から C へと変更した。

ランダム雑音及び音声を妨害音である雑音とした場合、サンプリング周波数の変更あるいは方向推定を行わずビームパターンを固定した場合でも、あらゆる雑音到来方向に対して雑音除去性能の減衰量は許容範囲内であることを確認した。なお、最適化アルゴリズムの実行時間は、サンプリング周波数の変更により約 1/3、信号到来方向を固定することにより約 2/3、ビームフォーマ実装法の変更により約 1/2、プログラミング言語の変更により約 1/5、つまり全体として約 1/45 に短縮できた。

6 今後の課題

本稿では、雑音除去アルゴリズムの各種パラメータの最適化は行っていない。例えば、マイクロホン配置は、等間隔直線配列とし、隣接マイクロホン間隔は10cmとしていたが、実際には対象とする周波数に基づいて決定すべきである。一般に、低周波数帯域の信号を扱うためにはマイクロホン間隔を拡げ、高周波数帯域の信号を扱うためには狭めることが望ましい [3]。マイクロホン間隔を変更することにより空間分解能が変化し、それにより本雑音除去アルゴリズムの評価結果も異なることが予想される。

従って、雑音除去の対象を明確にし、それに適したマイクロホン配置を検討し、パラメータの自動最適化を行うことが望ましい。

参考文献

- [1] 水町光徳, “マイクロホンアレーを用いた雑音除去法に関する研究,” 北陸先端科学技術大学院大学 博士論文 (2000).
- [2] S. F. Boll, “Suppression of acoustic noise in speech using spectral subtraction,” IEEE Trans. ASSP, Vol. 27, No. 2, pp. 113-120 (1979).
- [3] 大賀 寿郎, 山崎 芳男, 金田 豊, “音響システムとデジタル処理,” (社) 電子情報通信学会 (1995).