

Title	3次元空間を考慮した多数マイクロホンによる雑音除去法に関する研究
Author(s)	森井, 大貴
Citation	
Issue Date	1998-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/1140
Rights	
Description	Supervisor:赤木 正人, 情報科学研究科, 修士

3次元空間を考慮した多数マイクロホンによる 雑音除去法に関する研究

森井 大貴

北陸先端科学技術大学院大学 情報科学研究科

1998 年 2 月 13 日

キーワード: 雑音除去、マイクロホンアレー、3次元空間、計算できない帯域.

1 はじめに

現在音声認識システムの実用化には大きな関心が寄せられているが、雑音などが存在する実環境下での使用においては、音声認識システムの性能が大幅に低下することが報告されている。この問題を解決するべく、これまでに様々な雑音除去に関する研究 [2][3] がされているが、実用化するには多くの問題を抱えている。

そこで、本研究では、実環境下における音声認識システムの実用化を支援できるような、音声認識のプリプロセッサとしての雑音除去法を提案する。手法として、赤木らによる雑音除去法 [1] を拡張し、マイクロホン対を水平・垂直方向に用いることで、3次元空間のある1点から到来する雑音の除去を行なった。

2 雑音除去アルゴリズム

本研究では、図1のように、上下、左右のマイクロホン対と中央の補助マイクロホンを全て等間隔 10 cm に配置した構成 (計5本) のマイクロホンアレーを使用する。ここで、上下、左右のマイクロホンを各々、主対、副対とし、この2組のマイクロホン対を組み合わせる雑音除去を行なう。

また、マイクロホンに到来する信号、雑音は全て平面波であるとし、信号は正中面方向から、雑音は3次元空間のある1点から到来するものとする。

2.1 定式化

まず、1組のマイクロホン対での状況を考える。

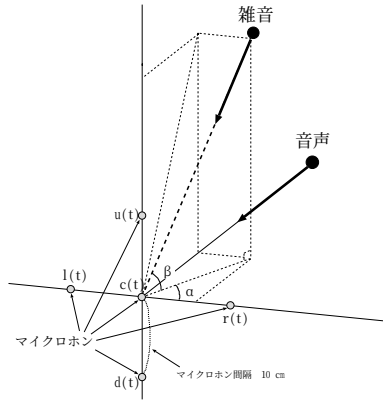


図 1: マイクロホンアレーの構成

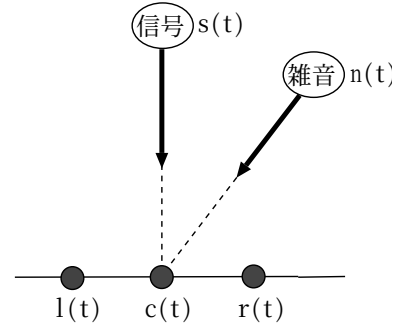


図 2: マイクロホン対と信号、雑音の関係

図2に示すように、左、中央、右のマイクロホンで受信した信号を各々 $l(t)$ 、 $c(t)$ 、 $r(t)$ 、目的信号を $s(t)$ 、雑音を $n(t)$ とすると、

$$\text{左: } l(t) = s(t) + n(t - \delta) \quad (1)$$

$$\text{中央: } c(t) = s(t) + n(t) \quad (2)$$

$$\text{右: } r(t) = s(t) + n(t + \delta) \quad (3)$$

と表すことができる。ここで、 2δ は、雑音が両端のマイクロホンへ到来する際の時間差である。

いま、雑音検出方向の時間差が 2τ であるとき、式 (1)、(3) を $\pm\tau$ ずらした信号から次式を定めると、

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

となり、 $g_{lr}(t)$ のフーリエ変換 $G_{lr}(\omega)$ は

$$G_{lr}(\omega) = \sin \omega \tau N(\omega) \sin \omega \delta \quad (5)$$

となる。そして、雑音の到来方向を推定し、 $\tau = \delta$ と定め、雑音の予測値を

$$\tilde{N}(\omega) = \begin{cases} G_{lr}(\omega) / \sin^2 \omega \delta & , \quad |\sin \omega \delta| > \varepsilon \\ G_{lr}(\omega) & , \quad |\sin \omega \delta| \leq \varepsilon \end{cases} \quad (6)$$

により求める。但し、 $\varepsilon \approx 0$ とする。

ここまでの過程を主対、副対でそれぞれ行なう。

そして、各々の対での雑音予測値 $\tilde{N}(\omega)$ は、 $\omega \delta = n\pi$ のときに $\varepsilon \approx 0$ としているため無限大になってしまい正確に推定できない周波数帯域が

$$f = n/2\delta, \quad n = 1, 2, \dots [Hz] \quad (7)$$

に存在する。

このため、新たに閾値として $\varepsilon_1, \varepsilon_2$ を設定し、主対で推定できない帯域のものを副対のものを
利用し補って、

$$\tilde{N}(\omega) = \begin{cases} G_{du}(\omega)/\sin^2 \omega \delta_{du}, & |\sin \omega \delta_{du}| > \varepsilon_1 \\ G_{lr}(\omega)/\sin^2 \omega \delta_{lr}, & |\sin \omega \delta_{du}| \leq \varepsilon_1, \\ & \text{and } |\sin \omega \delta_{lr}| > \varepsilon_2 \\ G_{du}(\omega), & |\sin \omega \delta_{lr}| \leq \varepsilon_2 \end{cases} \quad (8)$$

として推定する。

そして、この推定した雑音を、主対の受信信号の加算平均したものから引きさることで目的信号を求める。

雑音の到来方向推定に関しては、主対、副対ともに、相互相関に基づいた方向推定法を利用している。

3 雑音除去実験

本雑音除去アルゴリズムを評価をするために、計算機上で作成した信号を用いてのシミュレーション実験と、実環境下でスピーカから音を放射し、作成したマイクロホンアレーで収音したデータを用いての実験の2種類を行なった。

3.1 実験データ・条件

音声 (ATR,mht14348/bunri/) は正中面方向から、雑音 (スイープ音：始端周波数 1 kHz、終端周波数 6 kHz、継続時間 48000 point、帯域雑音：中心周波数 3.4kHz、帯域幅 3kHz、継続時間 48000point) は表 1 に示す方向から到来するものを用い、計算機上で加算した。なお、サンプリング周波数は 48 kHz とした。

表 1: 雑音の到来方向

シミュレーション実験	スイープ音：方位角右 30°、仰角上 30°	データ 1
	スイープ音：方位角右 60°、仰角上 30°	データ 2
	帯域雑音：方位角右 30°、仰角上 30°	データ 3
実環境下での実験	スイープ音：方位角右 45°、仰角上 20°	データ 4
	帯域雑音：方位角右 45°、仰角上 20°	データ 5

データ 2 は、主対、副対ともに雑音の到来時間差が同じものであり、計算できない帯域が同じ場合においての本雑音除去アルゴリズムの精度を調査するために作成した。

なお、実環境下での実験では、使用機材の関係上、24kHz で収音したデータを計算機上で 96kHz にアップサンプリングし、時間成分を修正したのち 48kHz にダウンサンプリングしたデータを用いた。

次に、雑音除去アルゴリズムの各種パラメータを表 2 のように設定した。閾値 $\varepsilon_{1,2}$ については実験を繰り返し、雑音誤差が最小になるよう経験的に定めた。

表 2: パラメータ設定

パラメータ	設定値（シミュレーション）	設定値（実環境下）
サンプリング周波数	48kHz	同左
フレーム長	2048 point（スイープ音） 1024 point（帯域雑音）	4096 point
フレーム周期	1024 point（スイープ音） 512 point（帯域雑音）	2048 point
窓関数	Hamming 窓（方向推定） 三角窓（雑音推定・除去）	同左
閾値 ε_1	0.7（スイープ音） 0.05（帯域雑音）	0.9
閾値 ε_2	0.5（スイープ音） 0.05（帯域雑音）	0.7

3.2 雑音の到来時間差

各マイクロホン対における雑音の到来時間差 δ は

$$\begin{aligned} \text{主対} &: \delta_{ud} = (2d/c) \cdot \sin \beta \\ \text{副対} &: \delta_{lr} = (2d/c) \cdot \cos \alpha \cdot \cos \beta \end{aligned} \quad (9)$$

α : 方位角、 β : 仰角、 c : 音速、 $2d$: 各マイクロホン対の間隔（図 1 参照）

として求まる。

そこで、表 3 に実験に用いた雑音の到来方向に対しての到来時間差および、計算できない帯域を示す。

表 3: 到来時間差、計算できない帯域

雑音の到来方向	到来時間差（主対、副対）	計算できない帯域（主対、副対）
方位角右 30°、仰角上 30°	7point、11point	3.4*n[kHz]、2.2*n[kHz]
方位角右 60°、仰角上 30°	7point、7point	3.4*n[kHz]、3.4*n[kHz]
方位角右 45°、仰角上 20°	5point、9point	4.8*n[kHz]、2.6*n[kHz]

3.3 実験結果

評価法としては、全波形区間で SN 比の計算を行なった。

$$SNR = 10 \log_{10} \frac{\sum_n s^2(t_n)}{\sum_n \{s(t_n) - \tilde{s}(t_n)\}^2} \quad (dB) \quad (10)$$

$s(t_n)$ は原波形、 $\tilde{s}(t_n)$ は評価対象波形

SN 比を計算した結果を表 4、5に示す。

また、データ 1、データ 2を用いて行なった実験結果のパワースペクトルを図 3、??に示す。

この結果、本手法によりデータ 1を用いた結果では SN 比が約 19 dB、データ 3を用いた結果では約 21 dB 向上し、主対のみで雑音除去した結果より各々約 11 dB、8 dB 改善量が向上した。また、同じ時間差の方向から到来するスイープ音（データ 2）の場合には、主対、副対の各々の計算できない帯域が同じになってしまうので、改善量が僅かであった。

表 4: SN 比（シミュレーション実験）

	データ 1	データ 2	データ 3
雑音を付加した音声波形	-10.45 (dB)	-10.45 (dB)	-3.61 (dB)
主対のみでの雑音除去後波形	-2.77 (dB)	-2.77 (dB)	11.27 (dB)
副対のみでの雑音除去後波形	-4.97 (dB)	-2.77 (dB)	9.54 (dB)
本手法による雑音除去後波形	8.61 (dB)	-2.15 (dB)	18.84 (dB)

表 5: SN 比（実環境下での実験）

	データ 4	データ 5
雑音を付加した音声波形	-1.83 (dB)	-1.86 (dB)
主対のみでの雑音除去後波形	3.07 (dB)	2.05 (dB)
副対のみでの雑音除去後波形	3.54 (dB)	3.68 (dB)
本手法による雑音除去後波形	4.64 (dB)	3.68 (dB)

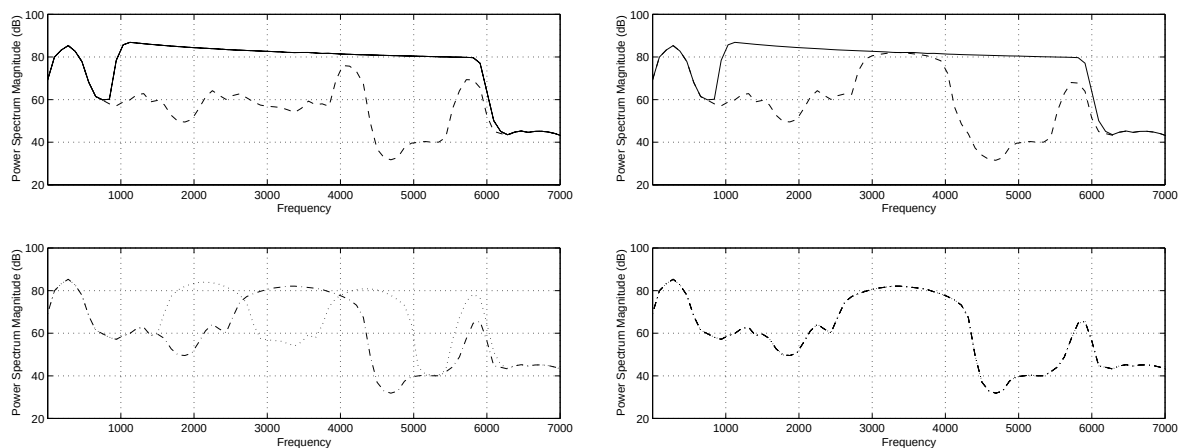


図 3: パワースペクトル（左図：データ 1 の結果、右図：データ 2 の結果）実線：雑音を付加した音声、破線：本手法による雑音除去後波形、鎖線：主対のみによる雑音除去後波形、点線：副対のみによる雑音除去後波形

4 まとめ

本手法は、3次元空間のある1点から到来する雑音を除去する行なうとともに、2組のマイクロホン対を組み合わせて用いることで、1組のマイクロホン対だけでは正確に推定できない帯域を補うことができた（図3左参照）。結果、雑音の到来時間差が違うものについては、SN比が主対のマイクロホン対で雑音除去した結果よりも数dB改善量が向上した。しかし、到来時間差が同じものについては今後検討が必要である（図3右参照）。

参考文献

- [1] 赤木、水町:“ マイクロホン対を用いた雑音除去法 (NORPAM), ”信学技SP,(1997.7)
- [2] L.J.Griffiths and C.W.Jim:“ An alternative approach to linearly constrained adaptive beam-forming, ”IEEE Trans. AP-30,1,(1982)
- [3] 大賀、山崎、金田:“ 音響システムとディジタル処理, ”電子情報通信学会,(1995)
- [4] M.Tobita,N.Sugamura:“ Effects of enviromental conditions on speech recognition, ” J.Acoust.Soc.Jpn(J)., vol.51, no.4, pp.331-335,(1995)