

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

修 士 論 文

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

指導教官 赤木 正人 助教授

北陸先端科学技術大学院大学
情報科学研究科情報処理学専攻

森井 大貴

1998 年 2 月 13 日

目次

1	序論	1
1.1	本研究の背景・目的	1
1.2	本研究の特徴	2
2	雑音除去アルゴリズムの概要	3
2.1	雑音除去におけるマイクロホンと信号・雑音の関係	3
2.2	雑音除去アルゴリズムの概要	4
3	雑音除去アルゴリズムの定式化	5
3.1	定式化	5
3.2	雑音方向推定	7
4	雑音除去実験	10
4.1	シミュレーション実験	10
4.1.1	雑音の到来時間差	10
4.1.2	実験条件	11
4.1.3	評価法	13
4.1.4	シミュレーション実験の結果	13
4.1.5	考察	22
4.2	実環境下での実験	23
4.2.1	収音を行なった環境・機材	23
4.2.2	実験条件	26
4.2.3	評価	27
4.2.4	実験結果	27

4.2.5	考察	34
5	結論	35
5.1	本研究についてのまとめ	35
5.2	今後の課題	35
	謝辞	37
	参考文献	38

目 次

2.1	マイクロホンの構成	3
3.1	マイクロホンと信号・雑音の関係	5
3.2	本雑音除去アルゴリズムのブロック図	9
4.1	クリーンな音声波形 (bunri)	14
4.2	スイープ音を付加した音声波形 (方位角右 30°、仰角上 30°の方向から到来) (データ 1)	14
4.3	主対のみで雑音除去を行なった場合の波形 (データ 1)	15
4.4	副対のみで雑音除去を行なった場合の波形 (データ 1)	15
4.5	本手法による雑音除去後の波形 (データ 1)	15
4.6	パワースペクトル (データ 1) (上段は全波形区間、下段は DC ~ 10 kHz を拡大) 実線: クリーンな音声、点線: 雑音を付加した音声、破線: 本手法による雑音除去後	16
4.7	パワースペクトル (データ 1) (上段は、実線: 雑音を付加した音声、破線: 本手法による雑音除去後、下段は、鎖線: 主対のみによる雑音除去後、点線: 副対のみによる雑音除去後)	17
4.8	主対のみで雑音除去を行なった場合の波形 (データ 2)	17
4.9	副対のみで雑音除去を行なった場合の波形 (データ 2)	18
4.10	本手法による雑音除去後の波形 (データ 2)	18
4.11	パワースペクトル (データ 2) (上段は全波形区間、下段は 1 Hz ~ 10 kHz を拡大) 実線: クリーンな音声、点線: 雑音を付加した音声、破線: 本手法による雑音除去後	19

4.12	パワースペクトル(データ 2)(上段は、実線：雑音を付加した音声、破線： 本手法による雑音除去後、下段は、鎖線：主対のみによる雑音除去後、点 線：副対のみによる雑音除去後)	20
4.13	帯域雑音を付加した音声波形(データ 3)	20
4.14	主対のみで雑音除去を行なった場合の波形(データ 3)	21
4.15	副対のみで雑音除去を行なった場合の波形(データ 3)	21
4.16	本手法による雑音除去後の波形(データ 3)	21
4.17	防音室内のマイクロホンとスピーカの配置平面図	24
4.18	システム構成図	25
4.19	スピーカで収録した音声波形(bunri)	28
4.20	スイープ音を付加した音声波形(データ 4)(方位角右 45°、仰角上 20°の 方向から到来)	28
4.21	主対のみで雑音除去を行なった場合の波形(データ 4)	29
4.22	副対のみで雑音除去を行なった場合の波形(データ 4)	29
4.23	本手法による雑音除去後の波形(データ 4)	29
4.24	パワースペクトル(データ 4)(上段は全波形区間、下段は DC ~ 10 kHz を 拡大) 実線：クリーンな音声、点線：雑音を付加した音声、破線：本手法 による雑音除去後	30
4.25	パワースペクトル(データ 4)(上段は、実線：雑音を付加した音声、破線： 本手法による雑音除去後、下段は、鎖線：主対のみによる雑音除去後、点 線：副対のみによる雑音除去後)	31
4.26	帯域雑音を付加した音声波形(データ 5)(方位角右 45°、仰角上 20°の方 向から到来)	31
4.27	主対のみで雑音除去を行なった場合の波形(データ 5)	32
4.28	副対のみで雑音除去を行なった場合の波形(データ 5)	32
4.29	本手法による雑音除去後の波形(データ 5)	32
4.30	パワースペクトル(データ 5)(上段は全波形区間、下段は DC ~ 10 kHz を 拡大) 実線：クリーンな音声、点線：雑音を付加した音声、破線：本手法 による雑音除去後	33

表 目 次

4.1	雑音の到来方向 (シミュレーション実験)	12
4.2	到来時間差、計算できない帯域 (シミュレーション実験)	12
4.3	パラメータ設定	13
4.4	SN 比 (シミュレーション実験)	22
4.5	使用機材一覧	24
4.6	雑音の到来方向 (実環境下での実験)	26
4.7	到来時間差、計算できない帯域 (実環境下での実験)	27
4.8	パラメータ設定 (実環境)	27
4.9	SN 比 (実環境下での実験)	30

第 1 章

序論

1.1 本研究の背景・目的

現在、機械による音声認識は、雑音・残響が含まれていないクリーンな音声ではほぼ実用化レベルに達している。しかし、周囲に雑音が存在する環境では、目的音が収音される際に付加する雑音・残響により認識率が著しく低下する。このため、雑音の混ざった音声から雑音を除去し取り出すことは認識率の向上に非常に有効であり、認識処理の前処理として大きな期待が寄せられている。

その一方、人間は雑音・残響のある環境でも目的音を選択し聴取することができるいわゆる「カクテルパーティー効果」といわれる能力を持っている。この能力は、音環境にほとんど影響を受けず頑健である。

音声認識におけるこの問題を解決するため、適応型アレーや複数のマイクロホンを用いたビームフォーマなど様々な雑音除去に関する方法が研究されている [2][5]。しかし、これらの方法は適応フィルタの収束が遅い、信号音に歪みが生ずる、大がかりな装置が必要であるなどといった問題があり実用化するまでには至っていない。また、赤木らによるマイクロホン対を用いた雑音除去法 [1] は、これらの問題を解析的手法を用いることで回避し、時間・周波数が局在した雑音を除去できることが報告されている。

しかし、これらの手法ではマイクロホンと音源の位置において平面だけを考えているため、音源が 3 次元空間に存在している実際の環境下では平面だけを考慮していれば精度が悪くなる。

そこで本研究では、自動車内などでの音声認識の実用化を目指したプリプロセッサとしての雑音除去法を提案する。そして、従来法において平面だけで考えている点について、3次元空間を考慮に入れ、空間に存在する雑音の除去を目的とした、多数マイクロホンを用いた雑音除去法を提案する。その際、音声認識アルゴリズムに依存しないプリプロセッサとともに、マイクロホン対を水平・垂直の2方向に用いることで3次元空間に存在する雑音の除去を目指す。

1.2 本研究の特徴

本研究では、雑音のある環境においては、現行の音声認識システムの性能が低下するという現状を鑑み、雑音のある実環境下での音声認識システムを広く一般的にできるようにするための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ことが目的である。

手法としては、赤木らによる1組のマイクロホン対を用いた雑音除去法を拡張して、マイクロホン対を水平・垂直方向に2組、中央に補助マイクロホンを1本の計5本のマイクロホンを用いてることで、3次元空間に存在しているある1点から到来する雑音の除去を行う。そして、従来法において平面だけで考えていた点について、3次元空間を考慮に入れることで雑音除去の精度を向上させるとともに、音声認識のプリプロセッサとして適用できるような雑音除去法を提案する。この際、位相情報を考慮に入れた手法を提案する。

第 2 章

雑音除去アルゴリズムの概要

2.1 雑音除去におけるマイクロホンと信号・雑音の関係

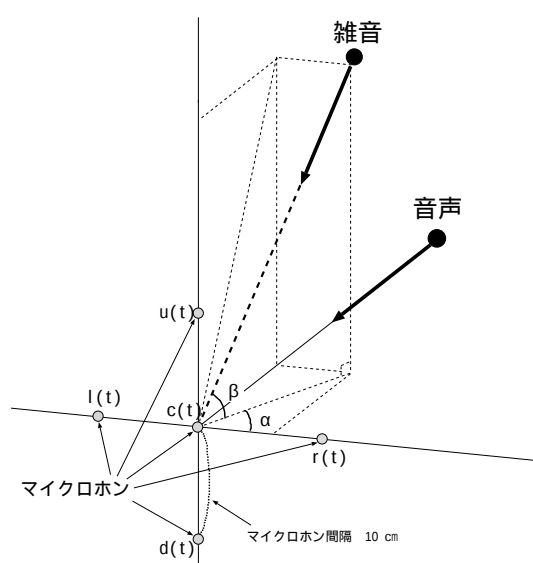


図 2.1: マイクロホンの構成

本研究における雑音除去アルゴリズムでは、図 2.1 に示すような 5 本のマイクロホンを等間隔に直線配置したマイクロホンアレーを使用する。このとき、上下のマイクロホン対を主対、左右のマイクロホン対を副対とし、中央のマイクロホンを補助マイクロホンとした構造で考える。

また、マイクロホンに到来する信号・雑音はすべて平面波とし、信号は全てのマイクロ

ホンへの到来時間差が等しい方向（正中面）から到来し、雑音は3次元空間のある一方向（方位角 α 、仰角 β ）から到来するものとして考える。

2.2 雑音除去アルゴリズムの概要

本研究で提案した雑音除去アルゴリズムについて概説する。

第1ステップ 雑音の推定 第1ステップとして雑音の推定を行なう。まず、主対、副対のマイクロホン対の各々の受音信号を利用して、各々のマイクロホン対での雑音の到来方向の推定を行なう。そして、その得られた到来方向の推定の時間差情報を用いることで雑音の推定を行なう。

この際、主対のマイクロホン対で正確に推定できない帯域に対しては、副対のマイクロホン対で推定したものを用いて雑音の推定を行なう。

第2ステップ 雑音の除去 第2ステップでは、主対のマイクロホン対で受信した信号の加算平均をして求めた信号から第1ステップで推定した雑音を引きさることによって目的信号を求める。

本雑音除去アルゴリズムは、この2ステップの処理を行なうことで、雑音除去を行なっている。

第 3 章

雑音除去アルゴリズムの定式化

3.1 定式化

本研究で提案する雑音除去アルゴリズムは、主対、副対で各々独自に雑音を推定し、主対において計算できない帯域の周辺を、副対のもので補って雑音の推定を行なう。

そこでまず、1組のマイクロホン対での状況を考える。

マイクロホン対と信号・雑音の関係は、図 3.1に示すような関係になっている。

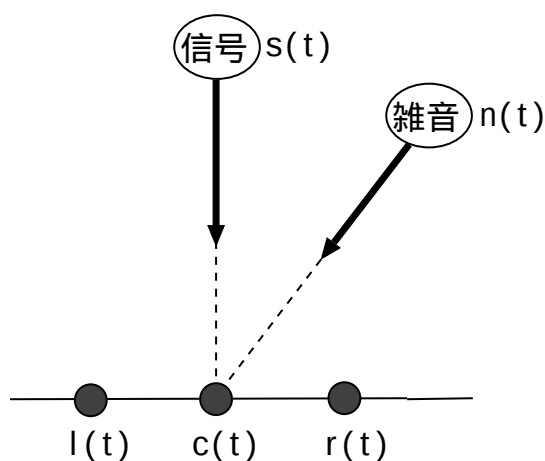


図 3.1: マイクロホンと信号・雑音の関係

図 3.1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 $l(t)$ 、 $c(t)$ 、 $r(t)$

とし、目的信号（音声）を $s(t)$ 、雑音を $n(t)$ とすると、

$$\text{左} : l(t) = s(t) + n(t - \delta) \quad (3.1)$$

$$\text{中央} : c(t) = s(t) + n(t) \quad (3.2)$$

$$\text{右} : r(t) = s(t) + n(t + \delta) \quad (3.3)$$

と表すことができる。ただし、 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。

そして、式 (3.1)、(3.3) のフーリエ変換を、各々 $L(\omega)$ 、 $R(\omega)$ とすると、

$$L(\omega) = S(\omega) + N(\omega)e^{-j\omega\delta} \quad (3.4)$$

$$R(\omega) = S(\omega) + N(\omega)e^{j\omega\delta} \quad (3.5)$$

となる。

ここで、雑音検出方向の時間差が 2τ であるとき、式 (3.1)、(3.3) を $\pm\tau$ ずらした信号から次式を定めると、

$$g_{lr}(t) = \frac{\{l(t + \tau) - l(t - \tau)\} - \{r(t + \tau) - r(t - \tau)\}}{4} \quad (3.6)$$

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

$$G_{lr}(\omega) = \sin \omega \tau N(\omega) \sin \omega \delta \quad (3.7)$$

となる。そして、雑音の到来方向を推定し $\tau = \delta$ と定めると、式 (3.7) は

$$G_{lr}(\omega) = N(\omega) \sin^2 \omega \delta \quad (3.8)$$

となり、雑音の予測値は

$$\tilde{N}(\omega) = \frac{G_{lr}(\omega)}{\sin^2 \omega \delta} \quad , \quad \omega \delta \neq n\pi \quad (3.9)$$

と求まる。

雑音の到来方向推定については次の 3.2 節で述べる。

ここまでの過程を、各々のマイクロホン対で各々独自に行なう。

そして、各々の雑音の予測値 $\tilde{N}(\omega)$ は、 $\omega\delta = n\pi$ のときに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存在する。このため、閾値として ε_1 、 ε_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うことで、雑音の推定を行なっている。

$$\tilde{N}(\omega) = \begin{cases} G_{du}(\omega) / \sin^2 \omega \delta_{du}, & |\sin \omega \delta_{du}| > \varepsilon_1 \\ G_{lr}(\omega) / \sin^2 \omega \delta_{lr}, & |\sin \omega \delta_{du}| \leq \varepsilon_1, \\ & \text{and } |\sin \omega \delta_{lr}| > \varepsilon_2 \\ G_{du}(\omega), & |\sin \omega \delta_{lr}| \leq \varepsilon_2 \end{cases} \quad (3.10)$$

一方、式 (3.4)、(3.5) から

$$\begin{aligned} U_{lr}(\omega) &= F \left[\frac{l(t) + r(t)}{2} \right] \\ &= S(\omega) + N(\omega) \cos \omega \delta \end{aligned} \quad (3.11)$$

であるので、目的信号は式 (3.12) の第 2 項を式 (3.10) を用いて消すことによって求まるので、式 (3.12) と表すことができる。

$$s(t) = F^{-1} [U_{lr}(\omega) - \tilde{N}(\omega) \cos \omega \delta] \quad (3.12)$$

3.2 雑音方向推定

雑音の到来方向を推定する際、ある周波数範囲 ($\omega_0 < |\omega| \leq \omega_1$)、ある方向 (時間差 $2t_r$) からの雑音が他の雑音成分に比べて極めて大きいとする。

雑音方向を推定する際、目的信号と雑音の共存する区間では、各々が相互に影響して推定誤りが生じる。そこで、信号音の影響を除去し雑音のみで雑音方向を推定する方法とし

て、左と中央のマイクロホン対、中央と右のマイクロホン対で、それぞれ式 (3.6)、(3.7)、(3.8) と同様の計算を行ない、

$$\hat{L}(\omega) = \begin{cases} F_{lc}(\omega) = N(\omega)e^{-j\frac{\delta}{2}\omega \sin \omega \delta}, & \omega_0 < |\omega| \leq \omega_1 \\ 0, & other \end{cases} \quad (3.13)$$

$$\hat{R}(\omega) = \begin{cases} F_{cr}(\omega) = N(\omega)e^{j\frac{\delta}{2}\omega \sin \omega \tau}, & \omega_0 < |\omega| \leq \omega_1 \\ 0, & other \end{cases}$$

とする。こうすることで、 G_{lc} 、 G_{cr} には雑音成分しか含まれていなく、時間差成分 $e^{\pm j\frac{\delta}{2}\omega}$ を含んでいる。そこで、

$$F^{-1} \left[\frac{\hat{L}(\omega)\hat{R}^*(\omega)}{|\hat{L}(\omega)||\hat{R}(\omega)|} \right] \equiv d_{\omega_0, \omega_1}(t) \quad (3.14)$$

を計算することで、目的信号に影響されずに雑音の到来方向を推定することができ、時間差 2δ は

$$\delta = \arg \max_t [d_{\omega_0, \omega_1}(t)] \quad (3.15)$$

から得ることができる。

最後に、雑音除去アルゴリズム全体の流れのブロック図を図 3.2 に示す。

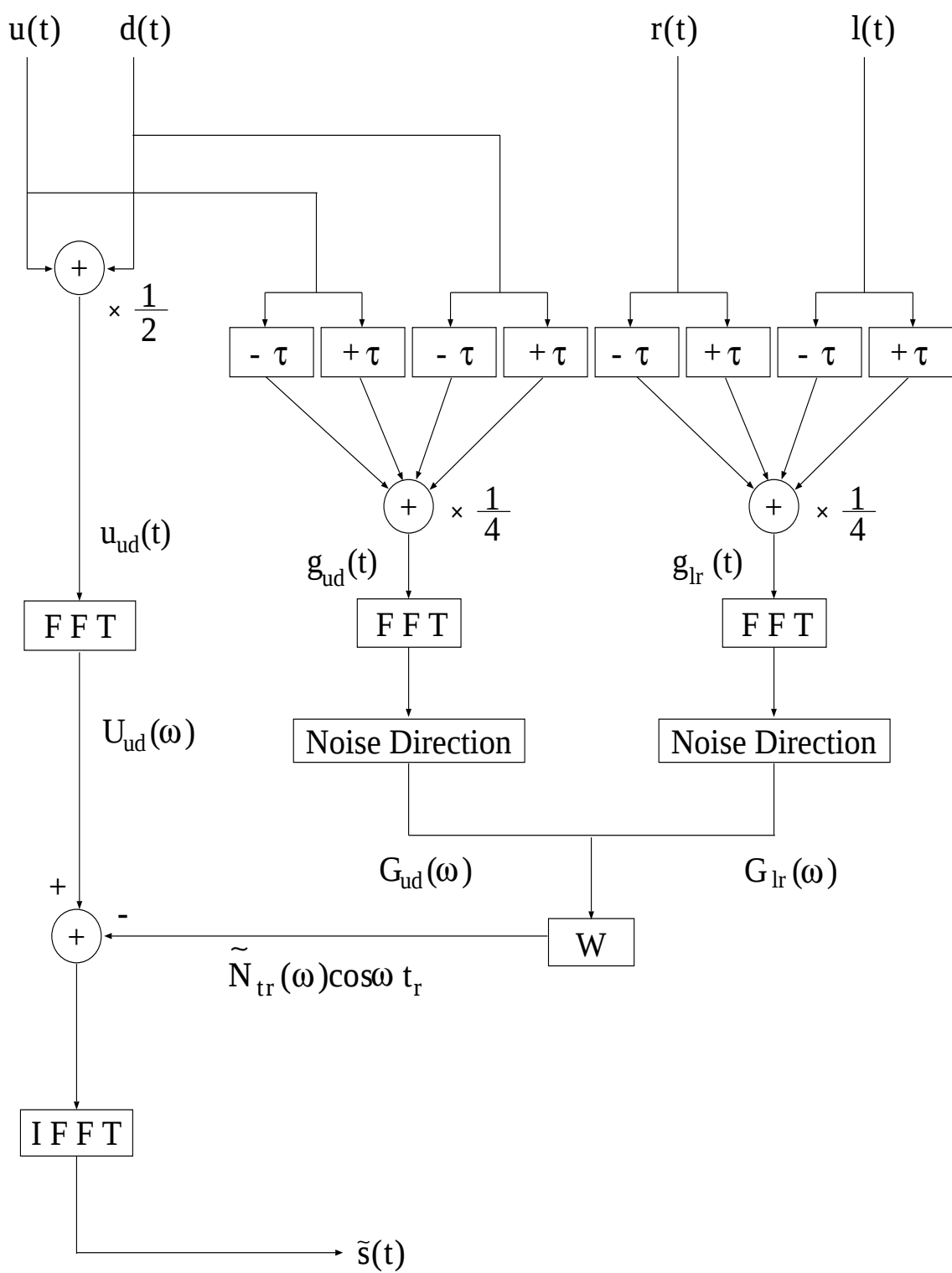


図 3.2: 本雑音除去アルゴリズムのブロック図

第 4 章

雑音除去実験

雑音除去実験を行なう目的としては、本研究で提案した雑音除去アルゴリズムがどれほど有効であるかを調査・検討するためである。

本章では、まず初めに計算機上で作成した信号データを用いてのシミュレーション実験を行ない、従来法に比べて本手法がどれだけの精度の向上が得られるかを評価・検討する。次に、本手法の実環境下での有効性を調査するために、実際に音声・雑音信号をスピーカを用いて室内に放射し、その信号を実際に作成したマイクロホンアレーで收音し、收音した信号を用いて本手法の有効性を評価・検討をう。

4.1 シミュレーション実験

シミュレーション実験を行なう際、音声は正中面方向から、雑音は正中面方向以外の 1 方向からのみ到来すると仮定し、計算機上で作成した信号データを用いて実験を行なった。

なお、各マイクロホンの間隔は全て等しく 10 cm と仮定して実験を行なった。

4.1.1 雑音の到来時間差

ここで、各マイクロホン対に対しての各々の時間差について述べる。

各マイクロホン対における雑音の到来時間差 δ は

$$\text{主対のマイクロホン対} : \delta_{ud} = \left(\frac{2d}{c} \right) \cdot \sin \beta \quad (4.1)$$

$$\text{副対のマイクロホン対} : \delta_{lr} = \left(\frac{2d}{c} \right) \cdot \cos \alpha \cdot \cos \beta \quad (4.2)$$

但し、 α :方位角、 β :仰角、 c :音速、 $2d$:各マイクロホン対の間隔

のようにして表すことができ、これらの式を計算することで求めることができる。(図??参照)

したがって、方位角右 30° 、仰角上 30° 方向から雑音が到来すると仮定した場合、主対のマイクロホン対と副対のマイクロホン対の各々の雑音の到来時間差は、サンプリング周波数が 48kHz の場合に、7point、11point と求めることができる。

そして、これに伴い主対、副対のマイクロホン対の各々における計算できない帯域は、

$$\begin{aligned} \text{主対} : f &= \frac{n}{2\delta_{ud}} = \frac{48000}{14} * n[\text{Hz}] \approx 3.4 * n[\text{kHz}] \\ \text{副対} : f &= \frac{n}{2\delta_{lr}} = \frac{48000}{22} * n[\text{Hz}] \approx 2.2 * n[\text{kHz}] \end{aligned} \quad (4.3)$$

として求めることができる。

本雑音除去アルゴリズムでは、この主対のマイクロホン対で計算できない帯域の周辺を副対のマイクロホン対の物を使い補うことで雑音除去の精度向上をはかる。

4.1.2 実験条件

ここで、シミュレーション実験を行なう際に用いた信号データおよび雑音除去アルゴリズムの各種パラメータについて説明する。

1. 実験データ

正中面から到来するクリーンな音声データとして、ATR 音声データベース [6] に収録されている単語 (mht14348/bunri/) を用い、正中面以外のある 1 方向から到来する雑音のデータとして、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)を計算機で作成したものをうい、それらを計算機上で加算して信号データを作成した。

雑音の到来方向としては、表 4.1に示す方向から到来するものとした。

表 4.1: 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

なお、スイープ音での方位角右 60°、仰角上 30°のデータ 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。

またここで、実験に用いた雑音の到来方向からの到来時間差、および、計算できない帯域を表 4.2に示す。

表 4.2: 到来時間差、計算できない帯域 (シミュレーション実験)

雑音の到来方向	到来時間差 (主対、副対)	推定できない帯域 (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]

2. パラメータ設定

本雑音除去アルゴリズムの各種パラメータは、以下の表 4.3 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 4.3: パラメータ設定

パラメータ	スイープ音	帯域雑音
サンプリング周波数	48 kHz	左同
フレーム長	2048 point	1024 point
フレーム周期	1024 point	512 point
窓関数	Hamming 窓 (方向推定)	左同
	三角窓 (雑音推定・除去)	左同
閾値 ε_1	0.7	0.05
閾値 ε_2	0.5	0.05

4.1.3 評価法

評価法としては、全波形区間での信号対雑音比 (SNR:Signal to Noise Ration) の計算を行なった (式 4.4)

$$SNR = 10 \log_{10} \frac{\sum_n s^2(t_n)}{\sum_n \{s(t_n) - \tilde{s}(t_n)\}^2} \quad (dB) \quad (4.4)$$

但し、 $s(t_n)$ は原波形、 $\tilde{s}(t_n)$ は評価対象波形である。

4.1.4 シミュレーション実験の結果

本手法を用いて雑音除去を行なった実験結果を示す。

まず、図 4.1 ~ 図 4.7 に、方位角右 30° 、仰角上 30° の方向から到来する SN 比 -10 dB のスイープ音を付加した信号データを用いた実験結果を示す。

次に、図 4.8 ~ 図 4.12 に、方位角右 60° 、仰角上 30° の方向から到来する SN 比 -10 dB のスイープ音を付加した信号データを用いた実験結果を示す。なお、図 4.12 下段では鎖線、点線が同じになり重なっている。

最後に、図 4.13 ~ 図 4.16 に、方位角右 30° 、仰角上 30° の方向から到来する SN 比 -3.6 dB の帯域雑音を付加した信号データを用いた実験結果を示す。

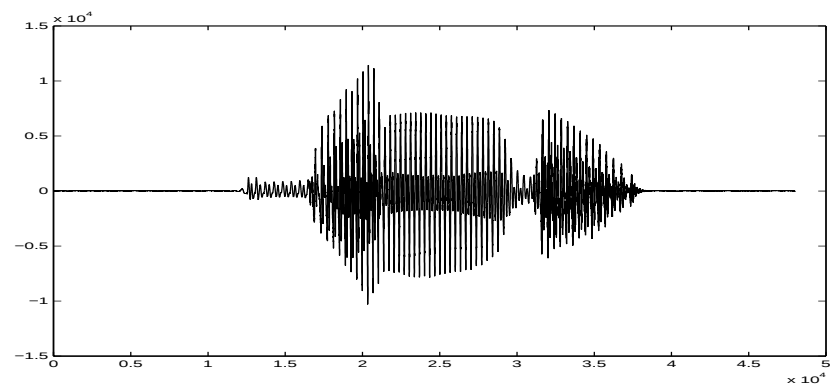


図 4.1: クリーンな音声波形 (bunri)

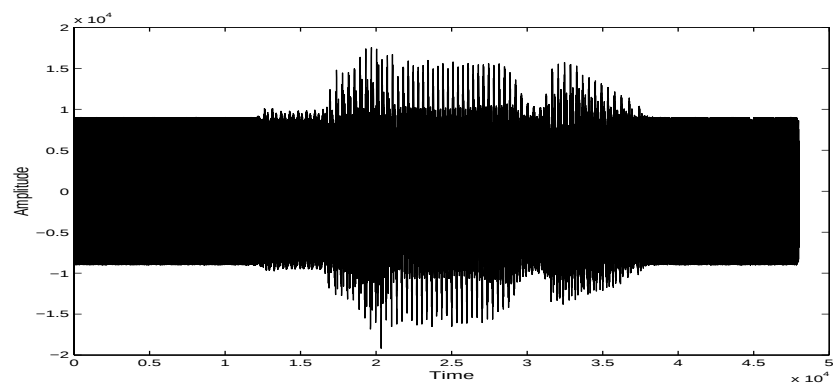


図 4.2: スイープ音を付加した音声波形 (方位角右 30°、仰角上 30° の方向から到来) (データ 1)

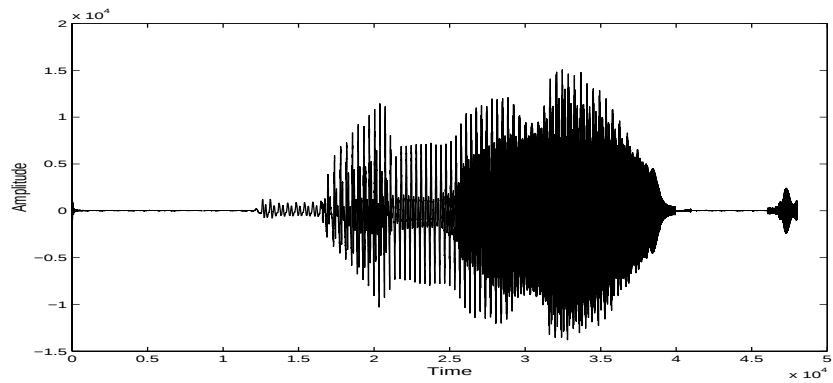


図 4.3: 主対のみで雑音除去を行なった場合の波形 (データ1)

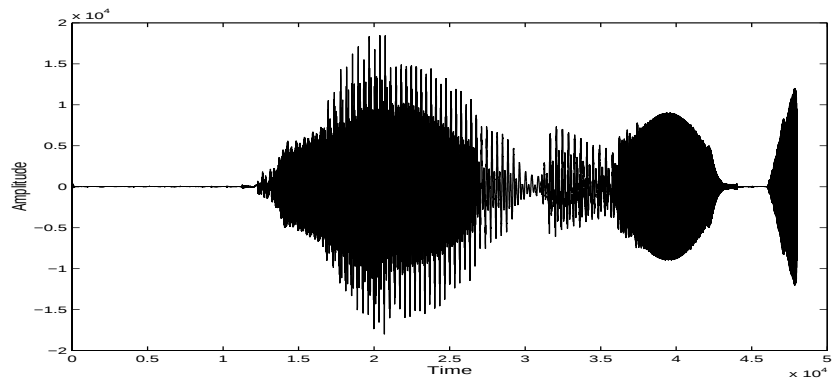


図 4.4: 副対のみで雑音除去を行なった場合の波形 (データ1)

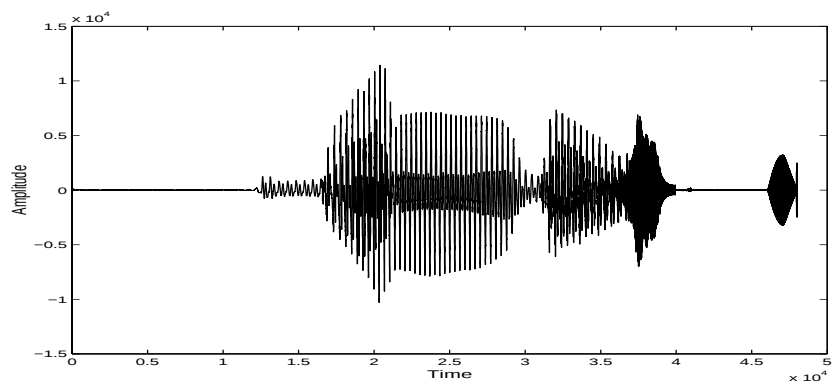


図 4.5: 本手法による雑音除去後の波形 (データ1)

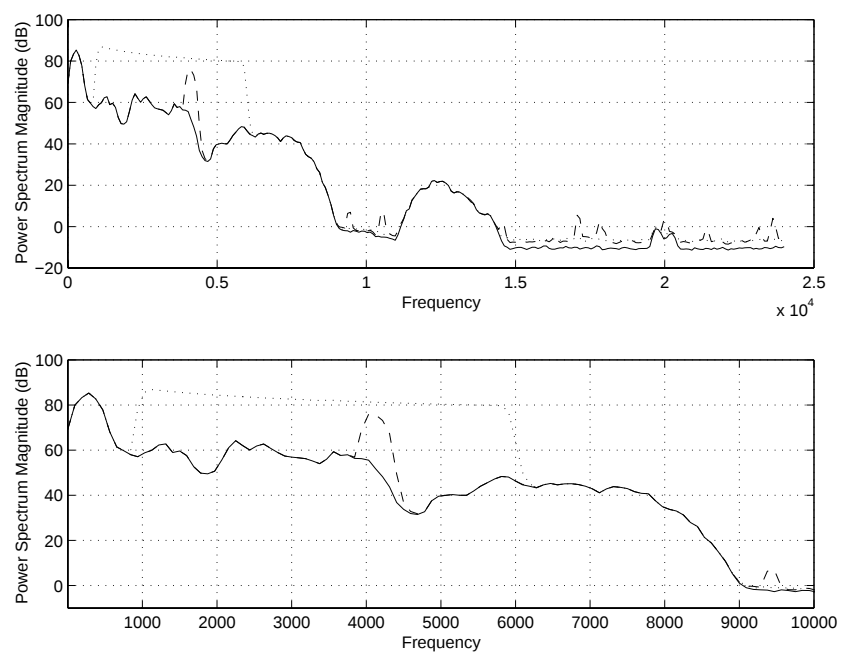


図 4.6: パワースペクトル(データ1)(上段は全波形区間、下段は DC ~ 10 kHz を拡大)
 実線: クリーンな音声、点線: 雑音を付加した音声、破線: 本手法による雑音除去後

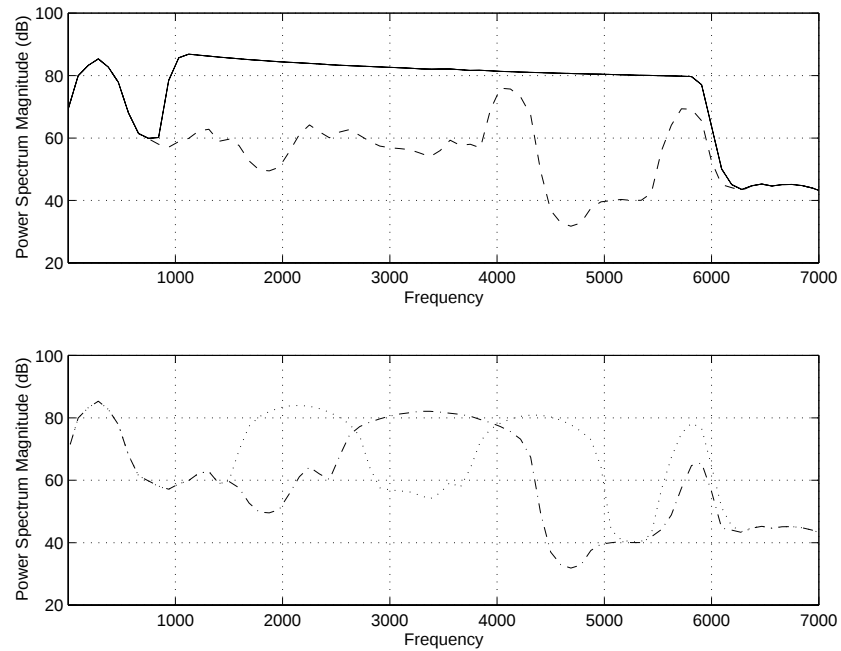


図 4.7: パワースペクトル(データ1)(上段は、実線:雑音を付加した音声、破線:本手法による雑音除去後、下段は、鎖線:主対のみによる雑音除去後、点線:副対のみによる雑音除去後)

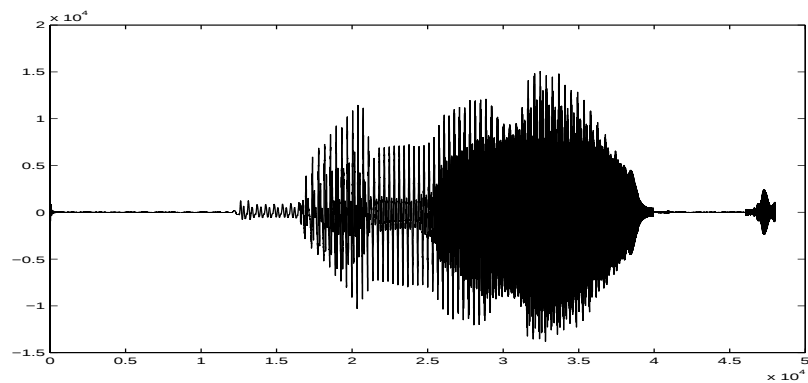


図 4.8: 主対のみで雑音除去を行なった場合の波形(データ2)

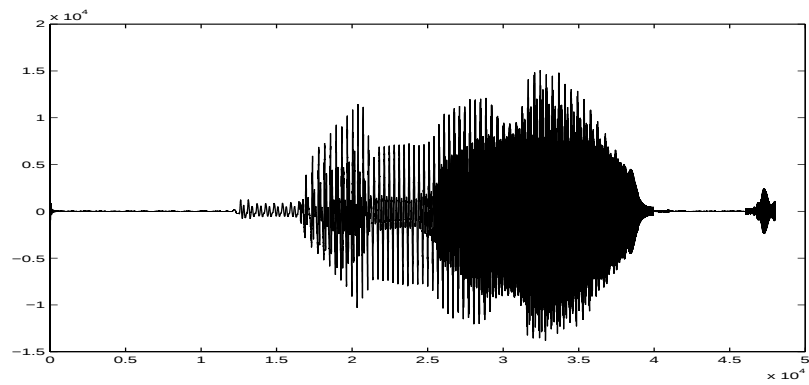


図 4.9: 副対のみで雑音除去を行なった場合の波形 (データ 2)

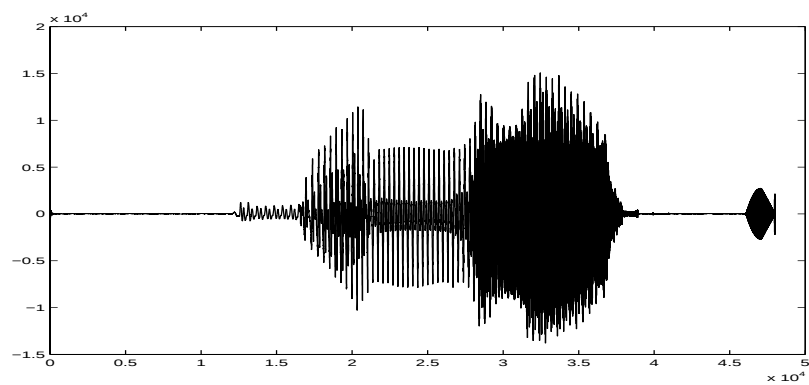


図 4.10: 本手法による雑音除去後の波形 (データ 2)

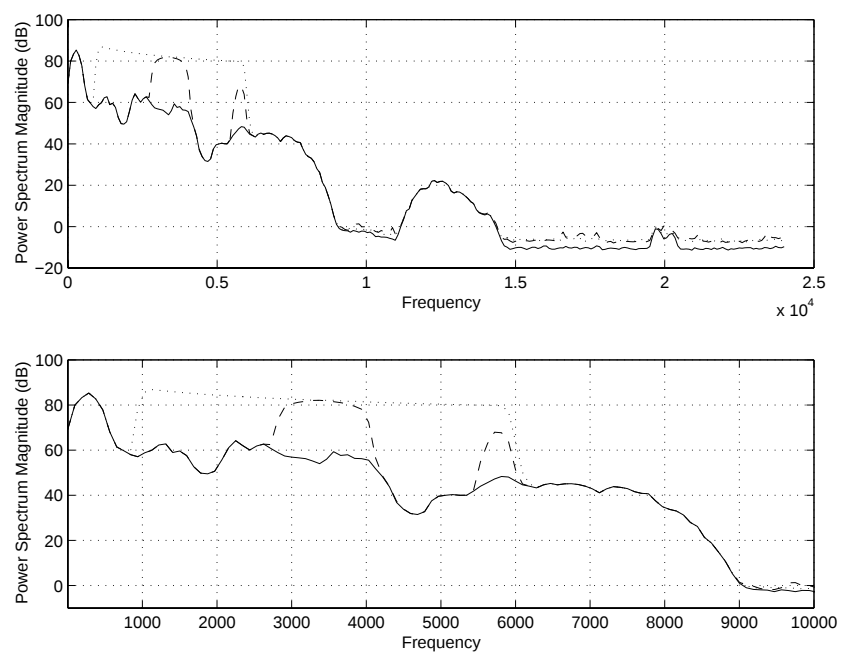


図 4.11: パワースペクトル(データ2)(上段は全波形区間、下段は 1 Hz ~ 10 kHz を拡大)実線: クリーンな音声、点線: 雑音を付加した音声、破線: 本手法による雑音除去後

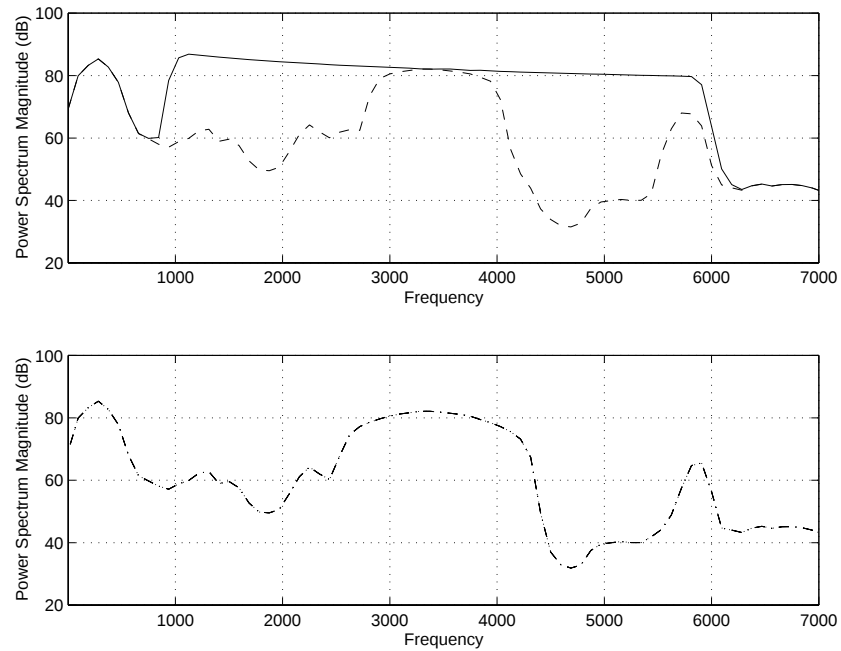


図 4.12: パワースペクトル (データ 2) (上段は、実線：雑音を付加した音声、破線：本手法による雑音除去後、下段は、鎖線：主対のみによる雑音除去後、点線：副対のみによる雑音除去後)

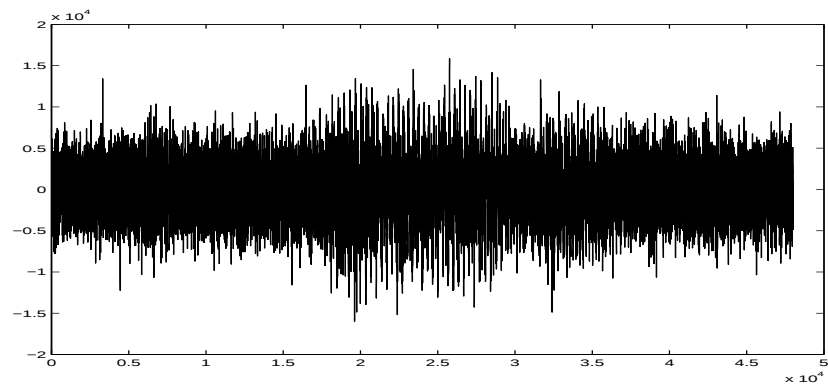


図 4.13: 帯域雑音を付加した音声波形 (データ 3)

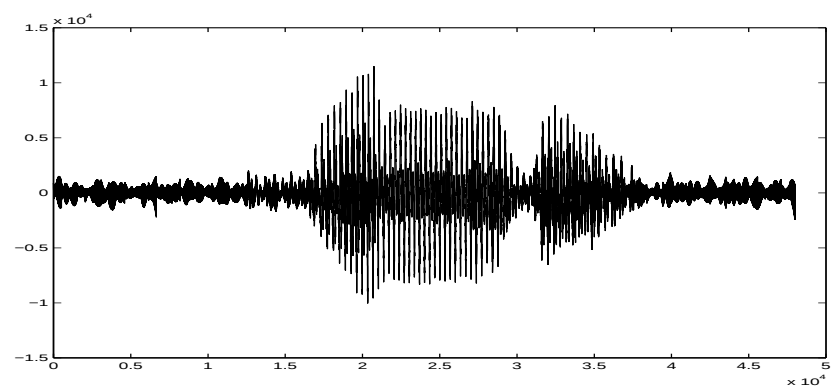


図 4.14: 主対のみで雑音除去を行なった場合の波形 (データ 3)

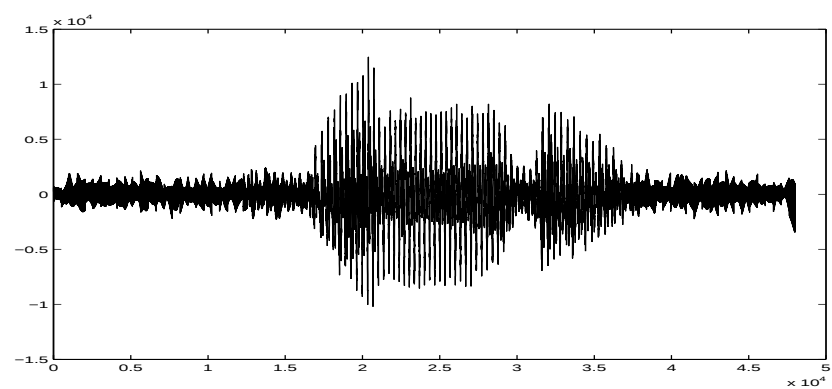


図 4.15: 副対のみで雑音除去を行なった場合の波形 (データ 3)

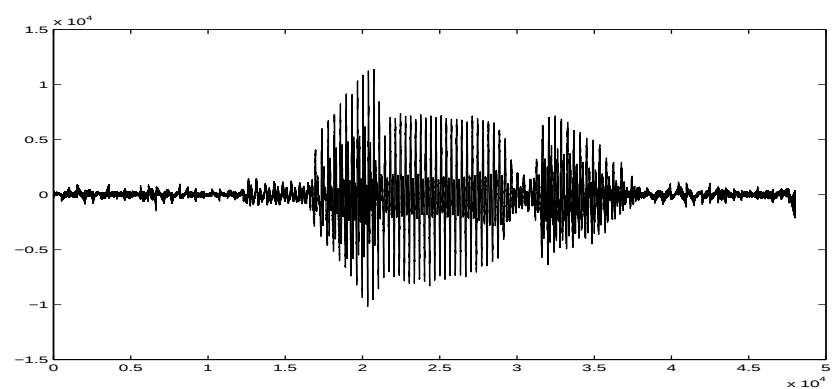


図 4.16: 本手法による雑音除去後の波形 (データ 3)

この得られたデータ結果から、式 4.4を計算して SN 比を計算した結果を表 4.4に示す。

表 4.4: 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)

この結果、本手法により方位角右 30°、仰角上 30°から雑音が到来する場合、スイープ音 (データ 1) では SN 比が約 19 dB、帯域雑音 (データ 3) では SN 比が約 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データ 2 (スイープ音 : 位角右 60°、仰角上 30°) の場合、SN 比が約 8 dB 向上し、主対で行なった場合よりも僅かではあるが改善量が向上した。

4.1.5 考察

シミュレーション実験の結果に基づき、本雑音除去アルゴリズムについて考察する。

まず、今回行なった実験条件においては、雑音の到来時間差が違う場合においては、スイープ音、帯域雑音ともに本手法により SN 比が約 20 dB 改善する結果を得ることができ、また、主対のみで雑音除去した結果と比較しても数 dB ~ 数十 dB と改善量が向上する結果を得ることができた。この結果は、雑音を推定する際に、主対のマイクロホン対での計算できない帯域周辺を副対のマイクロホン対のもので補った、本手法の特徴によるものであると考えることができる。

図 4.7の下段に示した 1 kHz ~ 6 kHz 辺りを見ると、主対のみで雑音除去した結果では約 3.1 ~ 3.5 kHz を中心とした周辺で改善量が少なく、副対のみで雑音除去した結果では約 2.2、4.4 kHz を中心とした周辺で改善量が少ない。これは、雑音の到来方向が方位角右 30°、仰角上 30° の場合においての、主対、副対の各々の計算できない帯域 $3.4 \cdot n$ kHz、 $2.2 \cdot n$ kHz を中心とした周辺の周波数に相等するものであり、図の上段に示した本手法の

結果を見ると明らかに、主対で計算できない 3.4 kHz 周辺を副対で補っていることがわかる。

一方、図 4.12の結果を見ると、主対、副対ともに計算できない帯域が同じになる時間差方向から到来する雑音（方向位角右 60°、仰角上 30°）の場合には 3.4 kHz 周辺の周波数においてほとんど雑音成分の改善されていないことがわかる。これは、SN 比の改善量を見ても明らかである。

また、計算できない帯域において、図 4.7の結果からもわかるように、主対、副対において帯域が違っている場合でも、その周辺に影響を及ぼすために、それらが重なる周波数帯域においては、パワースペクトルの改善があまりされていないことがわかる。図 4.5において、37000 ~ 40000、46000 ~ 48000 point において雑音が残っているが、このフレーム区間においては、雑音方向の推定も正確にできていた。しかし、この区間の成分を見ると、主対、副対での計算できない帯域は違うものの、その周辺の周波数成分が残ってしまい雑音がうまく消えなかったことがわかった。

しかし、それ以外の周波数帯域では、図 4.6の結果からもわかるように、ほぼクリーンな音声の成分と同じ結果が得られた。

4.2 実環境下での実験

本実験では、提案した雑音除去アルゴリズムの実環境下での有効性を調べるため、実際にスピーカから音を放射させ、作成したマイクロホンアレーによって收音した信号データを用いて雑音除去実験を行なった。

音声データはマイクロホンアレーの正中面方向に設置したスピーカから、雑音データについては正中面方向以外のある 1 方向からのスピーカから各々単独で放射し收音したものを計算機上で修正を加えた後に加算したものを利用した。

また、マイクロホンアレー作成において、各マイクロホン間隔は全て等しく 10 cm となるように作成した。

4.2.1 收音を行なった環境・機材

信号データの收音の際に使用した環境は、室容積が約 20 m^3 で、500 Hz における残響時間が約 50 ms の防音室を利用した。そして、マイクロホンアレーと正中面のスピーカの

距離は 2.7 m で、もう一方のスピーカとの距離は 1.9 m である。また、マイクロホンと正中面のスピーカは床上 1.2 m で、もう一方のスピーカは床上 1.7 m でに設置した。防音室内のマイクロホンアレーとスピーカの位置関係の平面図を図 4.17に示す。

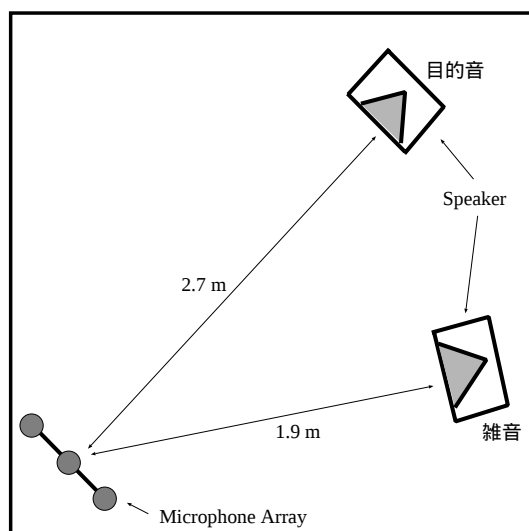


図 4.17: 防音室内のマイクロホンとスピーカの配置平面図

次に、音声データを収録する際に使用した機材を表 4.5に、収録システムの構成を図 4.18に示す。

表 4.5: 使用機材一覧

使用機材	名称
マイクロホンアレー	自作（使用マイクロホン SONY ECM-44B を使用）
スピーカ	自作（使用ユニット FOSTEX FE204 を使用）
レコーダ	TEAC RD-T145T
パワーアンプ	SANAUI AU-α907MR
DAT リンク	T.C.T. DAT-Link+
DA コンバーター	STAX DAC-TALENT BD

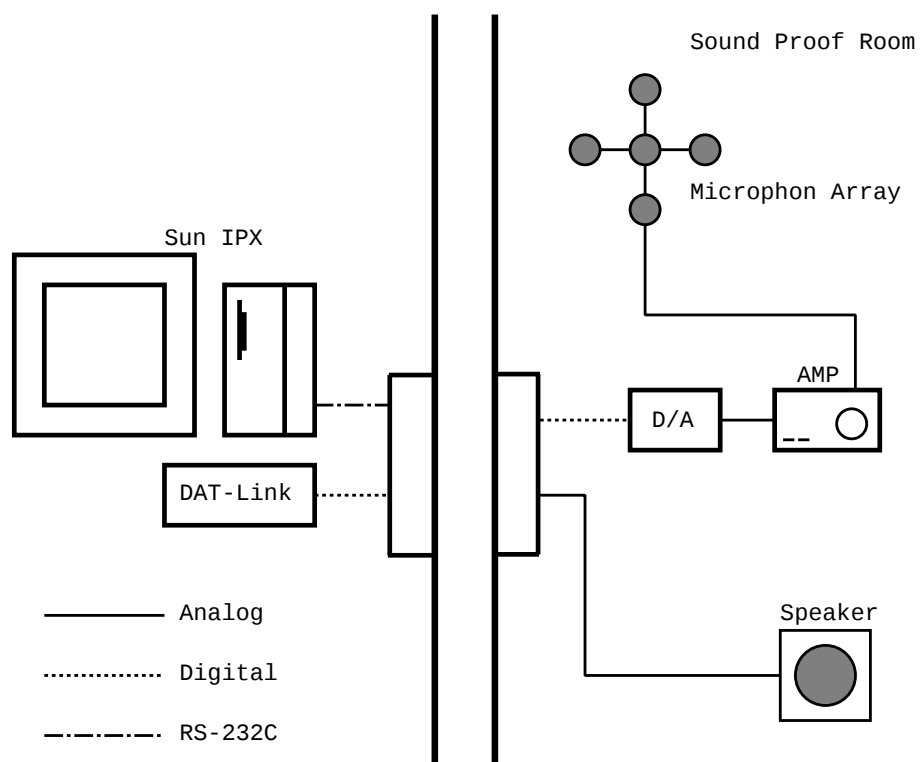


図 4.18: システム構成図

4.2.2 実験条件

ここで、本実験を行なう際に用いた信号データおよび雑音除去アルゴリズムの各種パラメータについて説明する。

1. 実験データ

スピーカから放射するデータとして、正中面から放射する音声データは、ATR 音声データベース [6] に収録されている単語 (mht14348/bunri/) を用い、正中面以外の 1 方向から放射する雑音データとして、帯域雑音 (中心周波数 3.4 kHz、帯域幅 3 kHz、継続時間 48000 point) 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) を使用した。そして、スピーカから放射された信号を、作成したマイクロホンアレーを用いて收音した。

この收音したデータは、使用した実験機材 (データレコーダ) の性能により、5 本のマイクロホンを用いて收音する場合には、サンプリング周波数が 24 kHz でしか收音することが出来ない。

そこで今回は、計算機上で 48 kHz に変換したものを本雑音除去アルゴリズムに用いた。まず、24 kHz で收音したデータを 96 kHz までアップサンプリングし、そのデータの時間差成分の修正を行ない、そして、48 kHz にダウンサンプリングして後に計算機上で音声データと雑音データの加算をした。

また、雑音の到来方向としては、表 4.6 に示す方向から到来するものとした。

表 4.6: 雑音の到来方向 (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

そして、シミュレーション実験でも前述したように (4.1.1 節参照) 雑音の到来方向による時間差から、主対・副対のマイクロホン対で各々計算できない帯域が存在する。そこで、実験に用いた雑音の到来方向からの時間差、および、計算できない帯域について表 4.7 に示す。

表 4.7: 到来時間差、計算できない帯域（実環境下での実験）

雑音の到来方向	到来時間差（主対、副対）	推定できない帯域（主対、副対）
方位角右 45°、仰角上 20°	5point、9point	4.8*n[kHz]、2.6*n[kHz]
方位角右 70°、仰角上 20°	5point、5point	4.8*n[kHz]、4.8*n[kHz]

2. パラメータ設定

雑音除去アルゴリズムのパラメータは表 4.8のように設定した。閾値については、シミュレーション実験同様、繰り返し実験し雑音誤差が最小になるよう経験的に定めた。

表 4.8: パラメータ設定

パラメータ	設定値
サンプリング周波数	48kHz 24 → 96 → 48 [kHz]
フレーム長	4096point
フレーム周期	2048point
窓関数	Hamming 窓（方向推定） 三角窓（雑音推定・除去）
閾値 ε_1	0.9
閾値 ε_2	0.7

4.2.3 評価

評価法としては、シミュレーション実験と同様に全波形区間で SN 比の計算を行なった（式 4.4）。

4.2.4 実験結果

本手法を用いて雑音除去を行なった実験結果を示す。

まず、図 4.19 ~ 図 4.25 に、方位角右 45° 、仰角上 20° の方向から到来する SN 比 -1.8 dB のスイープ音を付加した信号データを用いた実験結果を示す。

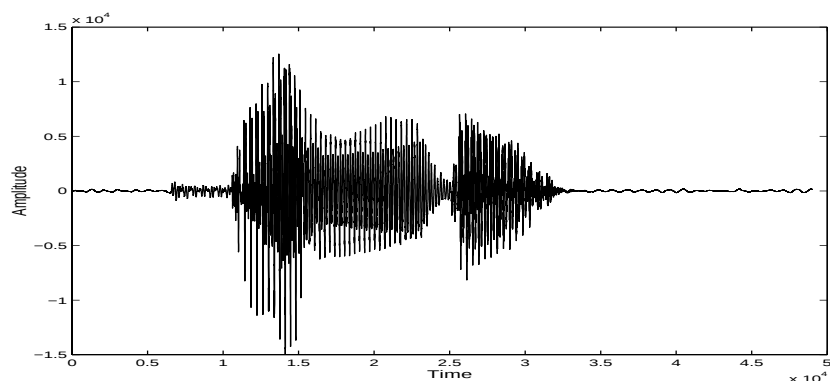


図 4.19: スピーカで収音した音声波形 (bunri)

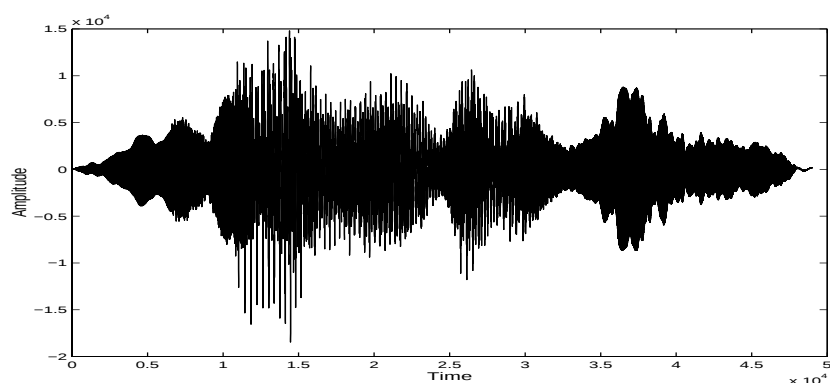


図 4.20: スイープ音を付加した音声波形 (データ 4) (方位角右 45° 、仰角上 20° の方向から到来)

次に、図 4.26 ~ 図 4.30 に、方位角右 45° 、仰角上 20° の方向から到来する SN 比 -1.8 dB のスイープ音を付加した信号データを用いた実験結果を示す。

この得られたデータ結果から、式 4.4 を計算して SN 比を計算した結果を表 4.9 に示す。

この結果、本手法によりデータ 4、5 でそれぞれ SN 比が約 6 dB、 5.5 dB 向上し、主対のみの場合と比べて、それぞれともに約 1.5 dB 改善量が向上した。

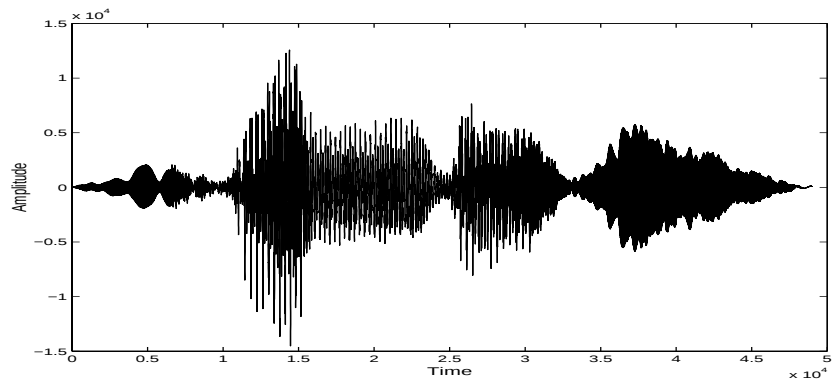


図 4.21: 主対のみで雑音除去を行なった場合の波形 (データ 4)

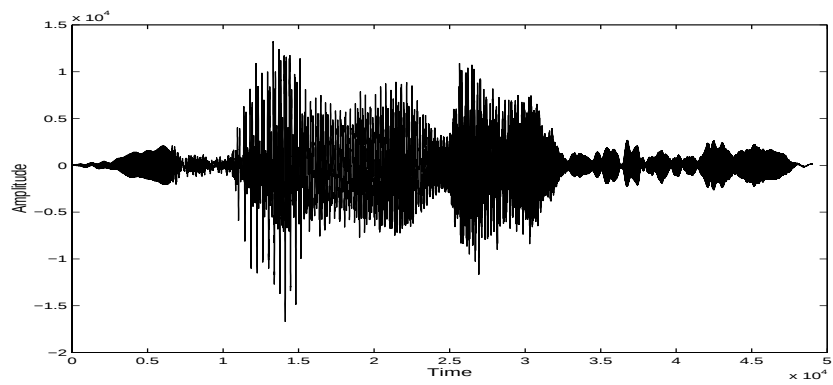


図 4.22: 副対のみで雑音除去を行なった場合の波形 (データ 4)

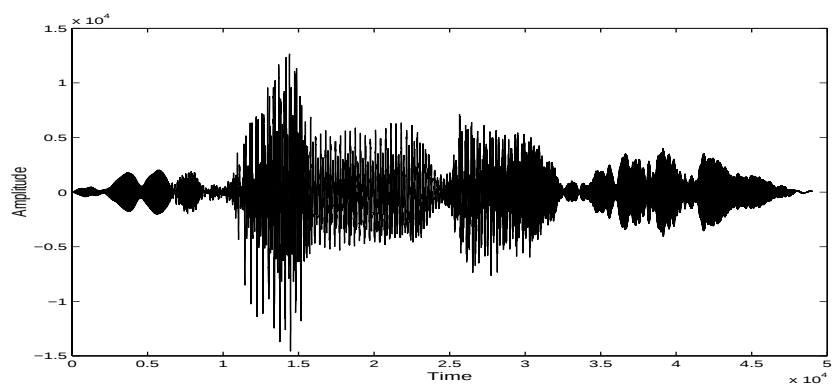


図 4.23: 本手法による雑音除去後の波形 (データ 4)

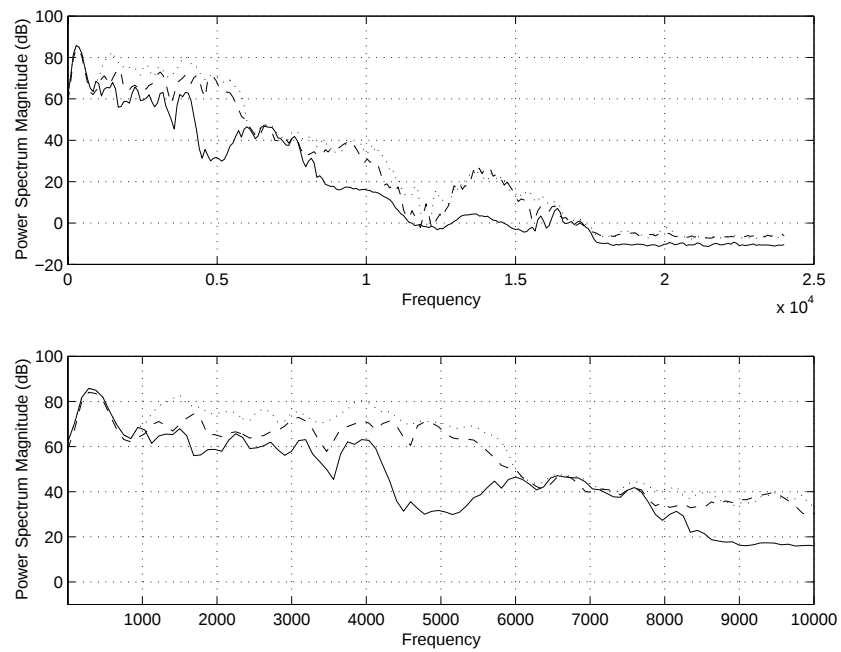


図 4.24: パワースペクトル(データ 4)(上段は全波形区間、下段は DC ~ 10 kHz を拡大)
 実線 : クリーンな音声、点線 : 雑音を付加した音声、破線 : 本手法による雑音除去後

表 4.9: 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)

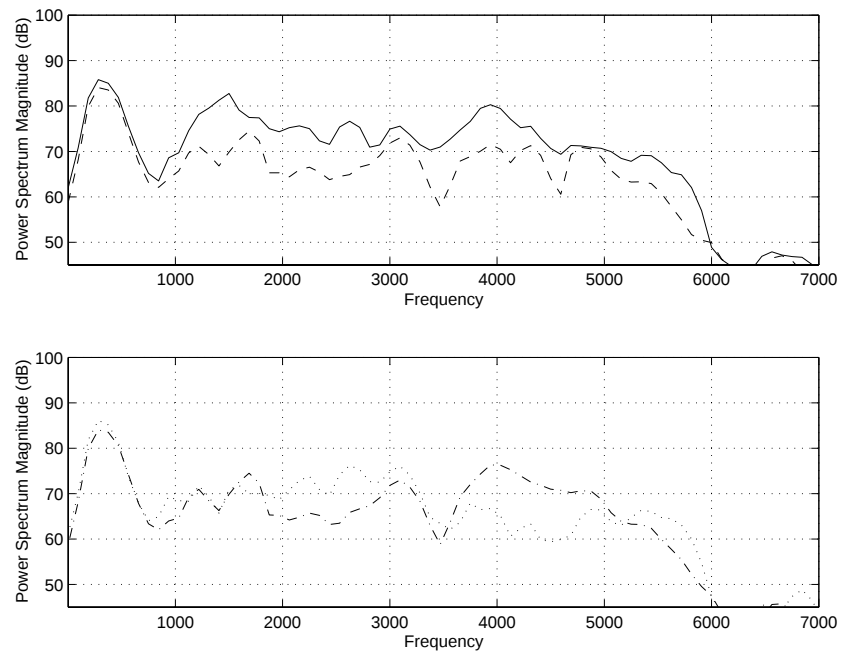


図 4.25: パワースペクトル (データ 4) (上段は、実線：雑音を付加した音声、破線：本手法による雑音除去後、下段は、鎖線：主対のみによる雑音除去後、点線：副対のみによる雑音除去後)

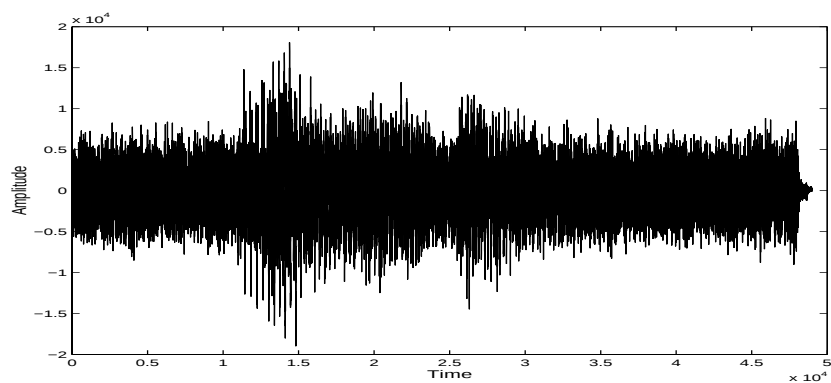


図 4.26: 帯域雑音を付加した音声波形 (データ 5) (方位角右 45° 、仰角上 20° の方向から到来)

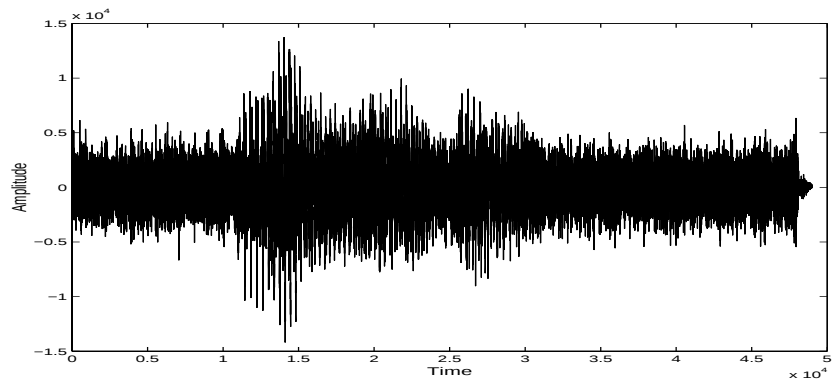


図 4.27: 主対のみで雑音除去を行なった場合の波形 (データ 5)

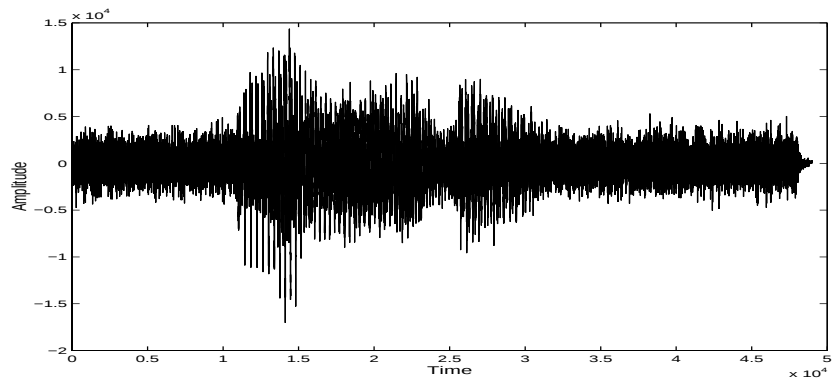


図 4.28: 副対のみで雑音除去を行なった場合の波形 (データ 5)

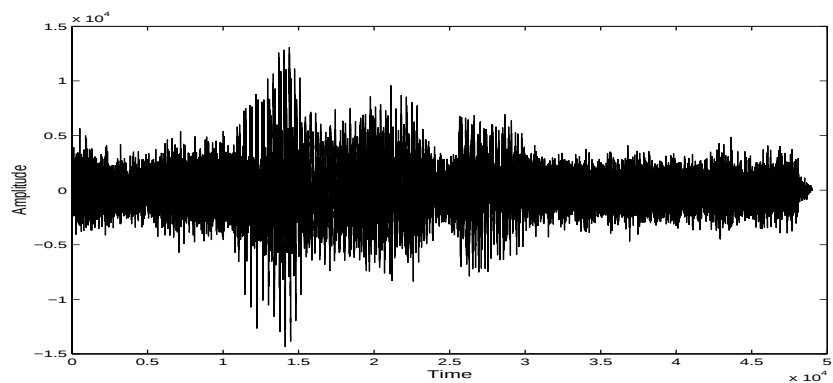


図 4.29: 本手法による雑音除去後の波形 (データ 5)

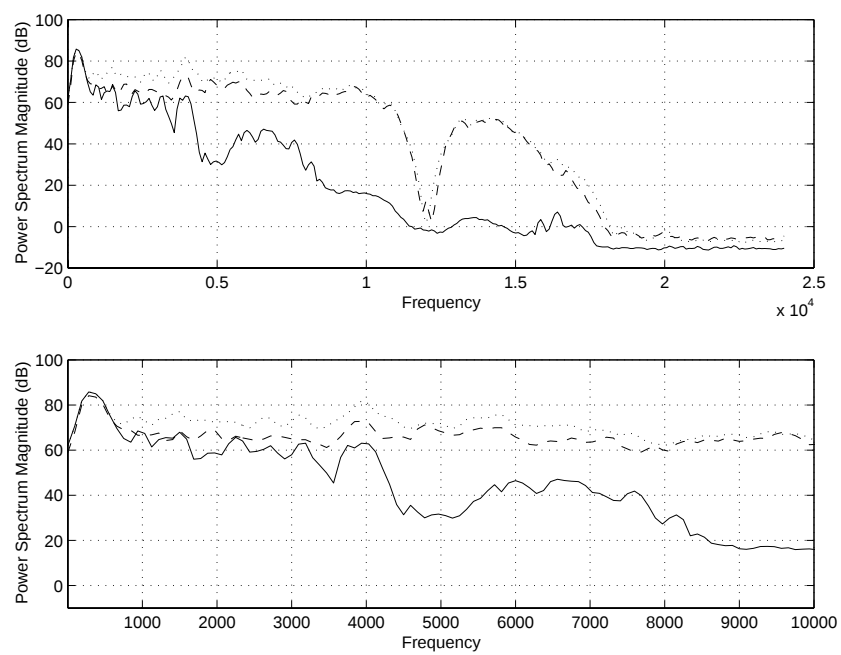


図 4.30: パワースペクトル(データ5)(上段は全波形区間、下段は DC ~ 10 kHz を拡大)
 実線 : クリーンな音声、点線 : 雑音を付加した音声、破線 : 本手法による雑音除去後

4.2.5 考察

実環境下での実験の結果、シミュレーション実験の結果と同様、SN 比で評価した場合、主対のみで雑音除去を行なった結果より、本手法の方が改善量は向上した。これは、シミュレーション実験の時と同様に、主対と副対を組み合わせているために、1 組のマイクロホン対で雑音除去した結果よりも SN 比の向上が見られた。だが、波形レベルでは、雑音成分が大きく残ってしまった。

これは、やはりシミュレーション実験と違って、実際にスピーカで音を出力し、マイクロホンで受音したデータを用いたためであり、音がマイクロホンに到着するまでの間に、反射音などによる影響が及ぼしたものと考えられる。図 4.24 のパワースペクトルの結果を見ると、音声、雑音ともにスピーカから放射する前のものとは違い歪んでいる（スピーカから放射する音の原波形は、シミュレーション実験に用いたものと同じである）。このため、収音したデータで本雑音除去アルゴリズムを行なうと、雑音の到来方向推定が正しく推定できないフレームが多くなり、雑音成分が残ってしまうので、今回のような結果となってしまったことがわかる。

また、データ 5 で行なった実験結果で、副対のものと改善量が変わらないという結果が得られたが、これは、雑音の到来方向と反射音による信号の歪みによるものと考えられる。データ 5 の場合に、主対、副対のマイクロホン対での到来方向による計算できない帯域は、それぞれ $4.8 \times n$ [kHz]、 $2.6 \times n$ [kHz] になり、5 [kHz] 周辺で、重なってしまうことも原因の 1 つと考えられる。

第 5 章

結論

5.1 本研究についてのまとめ

本研究では、音声認識のプリプロセッサとしての雑音除去を目指した手法を提案した。

そして、マイクロホン対を水平・垂直の 2 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を除去することができ、2 組のマイクロホン対を組み合わせることで、1 組のマイクロホン対のみで雑音除去した結果よりも良い結果を得ることができた。

しかし、シミュレーション実験だけを見れば、かなり実用化の期待を持てる結果を得られるものの、実環境下での実験結果では、反射音などの影響により、まだまだ実用化するには多くの問題を抱えていることがわかった。

5.2 今後の課題

今後の課題としては以下のことが挙げられる。

- 本手法のさらなる強化

本手法は、従来法の問題点であった計算できない帯域に対して、他のマイクロホン対を用いることで精度の向上を可能にした。しかし、主対、副対ともに同じ時間差で入射してくる場合には、計算できない帯域が同じになってしまい従来法と改善量がほとんど変わらない。そのため、現在使用している 5 本のマイクロホンの中から別のマイクロホン対の組を作るなどアルゴリズムの改良が必要である。

- 音源方向の推定

現在、音源方向については固定しているものと考えているが、実環境下においては、必ずしも音源が静止しているものだけではない。そのため、実際に実用化するためには、音源が移動する場合も当然考慮に入れる必要があると考えられる。

- 複数音源への対応

本手法では信号 1 つと雑音 1 つのみを考えているが、実際には様々な方向からの複数の雑音信号が存在する。よって、更に複数の雑音がある場合にも目的信号を抽出するためのアルゴリズムの考案が必要であると考えられる。

謝辞

本研究を進める上で、終始熱心に御指導をして頂きました赤木正人助教授、岩城護助手に心から御礼申し上げます。また、パターン関連研究室合同ゼミ等で熱心な御議論ならびに様々なアドバイスを下さいました諸先生方および研究室の皆様にも、この場を借りて御礼申し上げます。

そして、水町光徳氏、鵜木祐史氏を始めとする赤木研究室の先輩方、同期、後輩の皆様には貴重な助言を数多く頂き、心から感謝致します。

最後に、本研究に対するアドバイス、励まし等を頂いた友人をはじめとする皆様に感謝致します。

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 SP97-34,(1997.7)
- [2] L.J.Griffiths and C.W.Jim:“ An alternative approach to linearly constrained adaptive beamforming, ”IEEE Trans. AP-30,1,(1982)
- [3] Kaneda.Y. and Ohga.J: “ Adaptive microphonearray system for noise reduction, ” IEEE trans.ASSP,34,6,1391-1400,(1986)
- [4] Flanagan.J.L,et.al:“ Autodirective microphone systems, ”Acoustica,73,2,58-71,(1991)
- [5] 大賀、山崎、金田:“ 音響システムとデジタル処理, ”電子情報通信学会,(1995)
- [6] 武田、匂坂、片桐、阿部、桑原:“ 研究用日本語音声データベース利用解説書, ”ATR Technical Report,TR-I-0028,(1995)
- [7] 赤木 正人 : “ カクテルパーティー効果とそのモデル化, ”電子情報通信学会誌 78, 5, pp.450-453 (1995-5).
- [8] 飛田、菅村:“ 音声認識における周囲環境の影響, ”日本音響学会誌, vol.51, no.4, pp.331-335,(1995)
- [9] 日本音響材料協会 :“ 騒音・振動対策ハンドブック, ”技報堂,(1982)