

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

# 修士論文

## 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

指導教官 赤木 正人 教授

審査委員主査 赤木 正人 教授

審査委員 党 建武 助教授

審査委員 下平 博 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

910028 籠 貴司

2001年8月15日

## 要旨

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。減算型アレーの場合、これは (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 などの影響が原因であると考えられる。(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。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。

# 目次

<b>1</b>	<b>序論</b>	<b>1</b>
1.1	本研究の背景	1
1.2	本研究の目的	2
1.3	本論文の構成	3
<b>2</b>	<b>雑音除去アルゴリズム</b>	<b>4</b>
2.1	マイクロホンアレーを用いた雑音除去法	4
2.1.1	遅延和アレー	5
2.1.2	減算型アレー (適応型アレー)	6
2.1.3	水町らの提案法	6
2.2	雑音除去アルゴリズムの定式化	7
2.2.1	周波数帯域分割型アルゴリズム	7
2.2.2	雑音の到来方向推定	8
2.2.3	雑音スペクトルの推定	11
2.2.4	雑音スペクトルの除去	12
<b>3</b>	<b>計算機シミュレーションによる基礎的検証</b>	<b>13</b>
3.1	雑音除去アルゴリズムのパラメータ設定	13
3.2	複数雑音に対する雑音除去実験	15
3.2.1	シミュレーション条件	15
3.2.2	方向推定結果	15
3.2.3	雑音除去結果	17
3.3	周波数帯域が重複している複数雑音に対する雑音除去実験	21
3.3.1	シミュレーション条件	21
3.3.2	方向推定結果	21

3.3.3	雑音除去結果	23
3.4	音声が多音存在する場合に対する雑音除去実験	26
3.4.1	シミュレーション条件	26
3.4.2	雑音除去結果	26
4	実環境下における有効性検証	30
4.1	実環境下における複数雑音除去実験	30
4.2	近接音源に対する雑音除去実験	35
5	結論	37
5.1	本論文の要約	37
5.2	今後の課題	38

# 第 1 章

## 序論

### 1.1 本研究の背景

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、近年ではマイクロホンアレー (microphone array : 複数のマイクロホン素子から構成される受音器) を用いた雑音除去法も数多く提案されている。

単一マイクロホンを用いた雑音除去法として最も基本的な手法は、1979 年に Boll が提案した単点受音によるスペクトルサブトラクション (SS) であり、その改良法もいくつか発表されている。しかし、これらは処理が簡潔で実用化に適しているが、雑音の定常性を前提とするなど制約条件が実環境と大きく掛け離れている。これに対して多点受音であるマイクロホンアレーは、単点受音では得ることのできない信号の到来方向という空間情報を得ることにより、指向特性の設計 (ビームフォーミング) が可能である。

マイクロホンアレーは、遅延和アレーと減算型アレーの大きく 2 つに分けられる。遅延和アレーは目的信号方向に鋭い指向特性を形成することができ、雑音の性質や数に制約をうけることがなく、残響環境下においても有効である。しかしながら、そのために莫大な素子数を必要とし、アレーのサイズも大きくなってしまふ。一方、減算型アレーは、素子数の少ない小規模なアレーを用いて雑音方向に死角を形成することができるが、突発的な非定常雑音には適応できない。加えて、一般に形成できる死角数は (素子数 - 1) であるため、素子数以上の方向から到来する複数雑音は除去することができない [1]。

実環境下での応用を考えた場合、ハンズフリー型の自動音声認識器 (ASR) や補聴器など、マイクロホンアレーの素子数やサイズに制約がかかるため、小規模なマイクロホンアレーを用いた手法が求められる。遅延和アレーはマイクロホンアレーの素子数、サイズが大規模になるため、実際に応用するのは困難である。また減算型アレーを用いた場合、計

算機シミュレーションや制御された環境における実験では有効な雑音除去法であっても、実環境下においてはその制約条件が成立しないため、雑音除去精度が著しく低下してしまう。この原因として、次の3つの影響が考えられる。

- (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

素子数が少ないマイクロホンアレーほどこれらの影響を受け、雑音除去を強行すれば目的信号の歪みを引き起こす。よって実環境下で用いられ、かつ素子数の限られるハンズフリー型 ASR や補聴器に応用した場合、誤認識したり聴きづらくなるといった障害が生じる。

(1) の問題に関して水町らは、3ch. 解析型アレーを用いた雑音除去法を提案した。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、突発的な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、目的音が1つ、雑音が1つと仮定しているため、同時刻に到来する複数雑音には対応していない。

## 1.2 本研究の目的

本研究では、周波数帯域を分割し、各帯域で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を構築することを目的とする。この方法により、別々の方向から到来する複数雑音を同時に除去することができる。また基本となる雑音除去アルゴリズムには水町らの提案法を用いる。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。さらにハンズフリー型 ASR のフロントエンドなどに应用すべく、それを小規模マイクロホンアレーで実現する。

## 1.3 本論文の構成

本論文は全 5 章により構成されている。以下に各章の概要を述べる。

### 第 1 章：序論

マイクロホンアレーを用いた雑音除去研究の現状および問題点を明確にし、本研究の目的を示す。

### 第 2 章：雑音除去アルゴリズム

本雑音除去法のアルゴリズムを説明し、定式化を行う。

まず、序論で概説した従来の雑音除去法である、遅延和アレー、減算型アレーおよび解析型アレーを用いた水町らの提案法の概要を説明し、各手法の問題点を検討する。次にこれらの問題点をふまえ、本手法の特徴である周波数帯域分割型アルゴリズムをどのように水町らの提案法に組み込むかを説明する。最後に本手法の大まかな手順となる、雑音の到来方向推定、雑音スペクトルの推定、雑音スペクトルの除去について定式化を行い、詳細な説明を述べる。

### 第 3 章：計算機シミュレーションによる基礎的検証

第 2 章で定式化した雑音除去アルゴリズムの基礎的な雑音除去能力の検証を行う。

本研究の主目的である複数雑音の除去に関して、計算機シミュレーションによる雑音除去実験を行い、方向推定および雑音除去の有効性を調査する。

### 第 4 章：実環境下における有効性検証

残響や暗騒音が存在する、実環境下における本雑音除去アルゴリズムの有効性について検証する。

ここでも複数雑音の除去能力に着目し、本手法の有効性を示す。また、実環境下での応用を想定し、近接音源に対する雑音除去能力についても調査する。

### 第 5 章：結論

本研究で得られた結果を要約するとともに、今後検討すべき課題について述べる。

## 第 2 章

# 雑音除去アルゴリズム

### 2.1 マイクロホンアレーを用いた雑音除去法

本節では、マイクロホンアレーを用いた従来の雑音除去法の概要を説明し、その利点および欠点を明らかにする。序論でも述べた通り、マイクロホンアレーは遅延和アレーと減算型アレーの大きく 2 つの種類に分けられる。

- 遅延和アレー (図 2.1 左): 目的信号の到来方向に鋭い指向性を形成する。
- 減算型アレー (図 2.1 右): 雑音の到来方向に指向特性の死角を形成する。

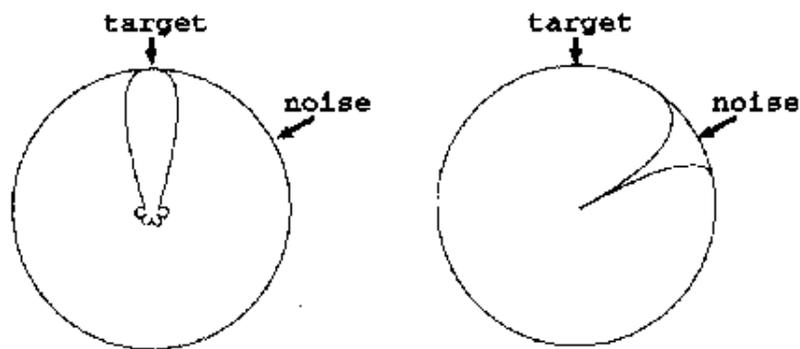


図 2.1: 遅延和アレーと減算型アレーの指向特性

### 2.1.1 遅延和アレー

遅延和アレーは図 2.2に示すような形で構成され、各マイクロホンの受信信号に遅延を与え同相化した後、係数を乗じ加算することにより目的信号を強調する。

図 2.2では、 $\theta$  方向から目的信号  $s(t)$  が 3ch. 遅延和アレーに到来する場合を考える。音波が平面波で伝搬すると仮定したとき、音速を  $c$ 、隣接マイクロホン間隔を  $d$  とすると、受信信号に与えるべき遅延量  $\tau_1$  は

$$\tau_1 = \frac{d \sin \theta}{c} \quad (2.1)$$

となり、目的信号  $s(t)$  は

$$\hat{s}(t) = \sum_{i=1}^3 w_i x_i(t - (i-1)\tau_1) \quad (2.2)$$

と推定される。ここで

$$\sum_{i=1}^3 w_i = 1 \quad (2.3)$$

より、 $w_1 = \frac{1}{3}$  とする。以上の手順により、目的信号の到来方向に鋭い指向性を形成する。また、目的信号の到来方向と異なった方向から到来する信号は同相化されないため、信号が相殺されて大きくなるしない。

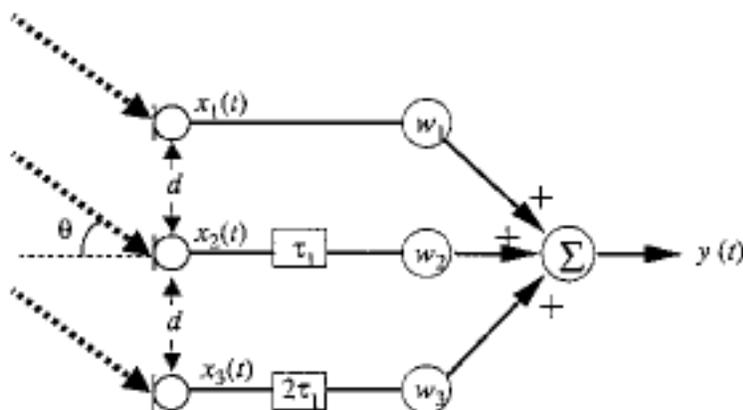


図 2.2: 3ch. 遅延和アレーのブロックダイアグラム

遅延和アレーは、雑音の到来方向に無関係にビームフォーミングを行うため、雑音に制約を受けることなく、突発的な非定常雑音や残響のようなあらゆる方向から到来する雑音にも有効である。しかしながら、高精度の雑音除去を実現するには莫大な素子数を必要とし、小規模なマイクロホンアレーほど特に低周波数帯域における雑音除去能力が低い。よって素子数やサイズに制約を受けるマイクロホンアレーへの応用は困難である [1]。

### 2.1.2 減算型アレー (適応型アレー)

減算型アレーは、各マイクロホンの受音信号に遅延を与え同相化した後、減算することにより雑音を消去する。この処理は、周囲から到来する雑音の方向を検出し、その方向に指向特性の死角 (感度が極めて低い方向) を形成することにより行われる。周囲の雑音環境に応じた指向特性を自動的に形成することから、このような機能を持つアレーは適応型アレーとも呼ばれている。

以上の操作において、遅延量  $\tau_N$  (雑音の到来方向  $\theta_N$ ) を知っている必要はない。つまり減算出力のパワー  $P'$  を監視しながら遅延量を変化させ、 $P'$  が最小となったときの遅延量を  $\tau_N$  とすることによって、自動的に決定される。遅延を付加することはデジタルフィルターを用いて実現でき、雑音成分のパワーが最小となるようにフィルターの係数が制御される (これを適応フィルターと呼ぶ)。また、目的信号が雑音とは異なった方向から到来するとすると、このような遅延操作では目的信号は同相化されないので、減算によって消去されることはない。

しかしながら適応型アレーは急激な音響的变化に追従することが困難であり、実環境下に多く存在する突発的な非定常雑音に対する除去能力が低い。また目的信号と雑音との相関が高い場合には、適応フィルタの収束が困難になると言われており、適応フィルタの係数が不適切に設定されると雑音が除去されないばかりか目的信号に新たな歪みを生じさせてしまう。加えて適応型アレーが形成できる死角数は一般に最大 (素子数  $-1$ ) であるため、素子数以上の複数雑音は同時に除去することができない。よって素子数の少ない減算型アレーを実環境下で用いた場合、非定常雑音および複数雑音の除去能力低下が大きな問題となる [2]。

### 2.1.3 水町らの提案法

これまで従来の小規模マイクロホンアレーを用いた手法では、実環境下における雑音に対応できないことを示してきた。この問題に対して水町らは、時々刻々推定した雑音成分を受音信号から減算する、突発的な非定常雑音に対してロバストな 3ch. 解析型アレーによる雑音除去法を提案した。

水町らの提案法の大まかな手順を以下に示す。

1. 目的信号と最も優勢な雑音の到来方向を推定する。
2. 目的信号を完全に抑圧し、雑音のスペクトルを推定する。
3. スペクトル領域において、受音信号から雑音成分を除去する。

信号の到来方向推定および雑音スペクトルの推定を行う際、2つのマイクロホン(マイクロホン対)の受信信号を利用し、目的信号を完全に抑圧するような減算型ビームフォーマを解析的に設計することが、この手法の特徴である。つまり減算型アレーでは通常雑音の到来方向に死角を設計するのに対して、目的信号方向に死角を設計し、目的音以外の信号を雑音と見なすことにより、少数のマイクロホンしか使用しなくても、高精度の雑音スペクトルの推定が可能となっている。また適応フィルタは使用せず、時々刻々解析的に雑音スペクトルを推定することにより、非定常な突発雑音に対して雑音除去精度が高い[3]。

しかしながら、この雑音除去アルゴリズムでは基本的に目的音源と雑音源とが1つずつ存在する状況を想定している。よって雑音除去が必要とされる実環境下においては、複数の雑音源や残響が存在するため十分に対応できない。より実環境に対応させるためには、複数雑音や残響をも除去できるようなアルゴリズムの改善が必要となってくる。そこで次節では水町らの提案法を改良し、特に複数雑音を除去できる雑音除去アルゴリズムを確立する。

## 2.2 雑音除去アルゴリズムの定式化

ここでは本雑音除去アルゴリズムの定式化を行う。複数雑音に対応させるための周波数帯域分割型アルゴリズムを、どのように水町らの提案法に組み込むのかを説明し、理論的に複数雑音および非定常雑音に対応できることを示す。

### 2.2.1 周波数帯域分割型アルゴリズム

減算型アレーにおいて指向特性の死角と共に重要となるのが、焦点(信号を必ず抽出する方向)である。雑音の到来方向と焦点が一致していない場合、雑音スペクトルを十分に推定することはできない。水町らの提案法では、全周波数帯域で1度しか雑音の方向推定を行わないため、決定できる焦点も1方向のみである。そのため雑音スペクトルの推定において、焦点ではなく指向特性の谷(感度が低い方向)から到来するような別の雑音は十分に推定することができなかった。

そこで本手法では、周波数帯域を分割し各帯域で雑音の到来方向の推定を行う。そして、雑音方向に焦点を合わせ、雑音スペクトルを推定する。これによって、各周波数帯域で最も優勢な雑音のスペクトルを推定することができる。つまり、方向推定を分割された周波数帯域で行うことができ、様々な方向に焦点を合わせた雑音スペクトルの推定が可能となる。また、各帯域において雑音方向に焦点が合っているので、十分に雑音スペクト

ルを推定することができる。この操作により、特に雑音の周波数がそれぞれ異なっていれば、複数雑音を全帯域の1つの雑音スペクトルとして推定することができ、複数雑音を同時に除去することが可能となる [4]。

以上をふまえ、本雑音除去アルゴリズムの大まかな手順を以下に示す。

1. 白色化相互相関関数により雑音の到来時間差を推定する。
2. 目的信号を完全に抑圧した減算型ビームフォーマにより雑音のスペクトルを推定する。  
(1、2を分割された周波数帯域ごとに繰り返し行い、全帯域の雑音スペクトルを求める。)
3. スペクトルサブトラクションにより雑音成分を除去する。

次節より、雑音の到来方向推定、雑音スペクトルの推定、雑音スペクトルの除去の順に定式化を行い、詳しい説明を述べる。

## 2.2.2 雑音の到来方向推定

本節では、短時間フレームごとに雑音の到来方向を推定するアルゴリズムを構築する。本手法では、受信信号の相互相関に基づく方向推定法を用いるが、方向推定にも積極的に減算型ビームフォーマを導入することにより、雑音環境下での信号到来方向の推定精度を向上させる。

本雑音除去法で使用するマイクロホンアレーは、図 2.3に示すような3本の無指向性マイクロホンを等間隔に直線配列したものである。信号が平面波としてマイクロホンアレーに到来する場合、信号の到来方向推定は、2つのマイクロホン間における信号の到来時間差を推定することに相当する。図 2.3に示す 3ch. 等間隔直線配列マイクロホンアレーにおいて、左側、中央、右側のマイクロホンでの受信信号をそれぞれ  $l(t)$ 、 $c(t)$ 、 $r(t)$  とする。目的信号  $s(t)$  以外に雑音  $n_m(t)$  が  $M$  個あるとすれば、各受信信号は次式のように表される。

$$l(t) = s(t - \zeta) + \sum_{m=1}^M n_m(t - \delta_m) \quad (2.4)$$

$$c(t) = s(t) + \sum_{m=1}^M n_m(t) \quad (2.5)$$

$$r(t) = s(t + \zeta) + \sum_{m=1}^M n_m(t + \delta_m) \quad (2.6)$$

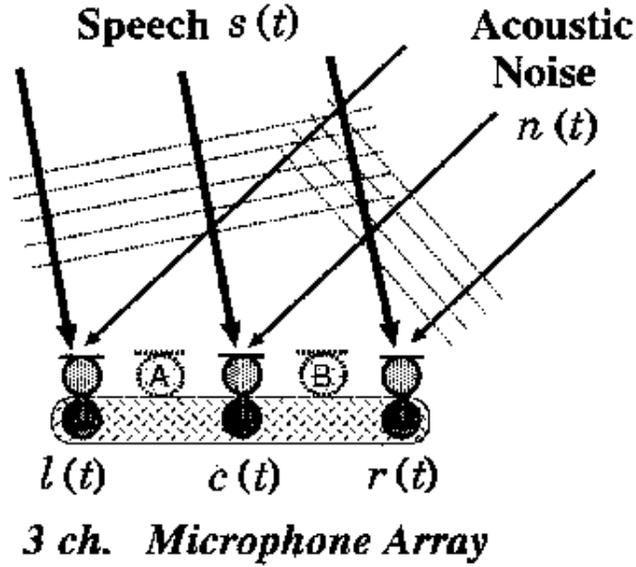


図 2.3: 3ch. 等間隔直線配列マイクロホンアレー

ここで  $\zeta$  は目的信号  $s(t)$ 、 $\delta_m$  は雑音  $n_m(t)$  の隣接マイクロホン間における到来時間差である。また  $l(t)$ 、 $c(t)$ 、 $r(t)$ 、 $s(t)$ 、 $n_m(t)$  のフーリエ変換をそれぞれ  $L(\omega)$ 、 $C(\omega)$ 、 $R(\omega)$ 、 $S(\omega)$ 、 $N_m(\omega)$  とすれば、上式は以下ようになる。

$$L(\omega) = S(\omega)e^{-j\omega\zeta} + \sum_{m=1}^M N_m(\omega)e^{-j\omega\delta_m} \quad (2.7)$$

$$C(\omega) = S(\omega) + \sum_{m=1}^M N_m(\omega) \quad (2.8)$$

$$R(\omega) = S(\omega)e^{j\omega\zeta} + \sum_{m=1}^M N_m(\omega)e^{j\omega\delta_m} \quad (2.9)$$

受信信号  $l(t)$  と  $r(t)$ 、 $l(t)$  と  $c(t)$ 、 $c(t)$  と  $r(t)$  を用いて、目的信号を完全に抑圧する信号、つまり到来時間差が  $\zeta$  である方向に指向特性の死角を形成するような減算型ビームフォーマ  $g_{lr}(t)$ 、 $g_{lc}(t)$ 、 $g_{cr}(t)$  を考える。

$$g_{lr}(t) = \frac{\{l(t+\zeta+\tau) - l(t+\zeta-\tau)\} - \{r(t-\zeta+\tau) - r(t-\zeta-\tau)\}}{4} \quad (2.10)$$

$$g_{lc}(t) = \frac{\{l(t+\zeta+\tau) - l(t+\zeta-\tau)\} - \{c(t+\tau) - c(t-\tau)\}}{4} \quad (2.11)$$

$$g_{cr}(t) = \frac{\{c(t+\tau) - c(t-\tau)\} - \{r(t-\zeta+\tau) - r(t-\zeta-\tau)\}}{4} \quad (2.12)$$

ここで  $\tau$  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。また  $g_{lr}(t)$ 、 $g_{lc}(t)$ 、 $g_{cr}(t)$  のフーリエ変換をそれぞれ  $G_{lr}(\omega)$ 、 $G_{lc}(\omega)$ 、 $G_{cr}(\omega)$  とす

れば、上式は以下ようになる。

$$G_{lr}(\omega) = \sum_{m=1}^M N_m(\omega) \sin \omega(\delta_m - \zeta) \sin \omega \tau \quad (2.13)$$

$$G_{lc}(\omega) = \sum_{m=1}^M N_m(\omega) e^{-j\omega \frac{\zeta - \delta_m}{2}} \sin \omega \frac{\zeta - \delta_m}{2} \sin \omega \tau \quad (2.14)$$

$$G_{cr}(\omega) = \sum_{m=1}^M N_m(\omega) e^{j\omega \frac{\delta_m - \zeta}{2}} \sin \omega \frac{\delta_m - \zeta}{2} \sin \omega \tau \quad (2.15)$$

これらの式に  $S(\omega)$  に関する項は一切存在せず、時間差  $\zeta$  で到来する目的信号を完全に抑圧した信号であることがわかる。

簡単化のため目的信号の到来方向は正面 ( $\zeta = 0$ ) とし、十分に狭い周波数帯域  $\omega_{k-1} < |\tilde{\omega}| \leq \omega_k$  ( $k = 1 \sim K$ ) において到来時間差が  $\delta_k$  となる雑音  $N_k(\tilde{\omega})$  を想定する。帯域を十分に狭くした場合、信号を純音として扱うことができ、純音の和は純音となるので

$$\sum_{m=1}^M N_m(\tilde{\omega}) \sin \tilde{\omega} \delta_m \sin \tilde{\omega} \tau = N_k(\tilde{\omega}) \sin \tilde{\omega} \delta_k \sin \tilde{\omega} \tau \quad (2.16)$$

と書ける。よって  $G_{lr}(\omega)$ 、 $G_{lc}(\omega)$ 、 $G_{cr}(\omega)$  を周波数帯域の和として表せば

$$G_{lr}(\omega) = \sum_{k=1}^K N_k(\tilde{\omega}) \sin \tilde{\omega} \delta_k \sin \tilde{\omega} \tau \quad \omega_{k-1} < |\tilde{\omega}| \leq \omega_k \quad (2.17)$$

$$G_{lc}(\omega) = \sum_{k=1}^K N_k(\tilde{\omega}) e^{j\tilde{\omega} \frac{\delta_k}{2}} \sin \tilde{\omega} \frac{-\delta_k}{2} \sin \tilde{\omega} \tau \quad \omega_{k-1} < |\tilde{\omega}| \leq \omega_k \quad (2.18)$$

$$G_{cr}(\omega) = \sum_{k=1}^K N_k(\tilde{\omega}) e^{j\tilde{\omega} \frac{\delta_k}{2}} \sin \tilde{\omega} \frac{\delta_k}{2} \sin \tilde{\omega} \tau \quad \omega_{k-1} < |\tilde{\omega}| \leq \omega_k \quad (2.19)$$

となる。 $G_{lc}(\omega)$ 、 $G_{cr}(\omega)$  より

$$F_{lc}(\tilde{\omega}) = \begin{cases} G_{lc}(\tilde{\omega}) / \sin \tilde{\omega} \tau & \omega_{k-1} < |\tilde{\omega}| \leq \omega_k \\ 0 & \text{otherwise} \end{cases} \quad (2.20)$$

$$F_{cr}(\tilde{\omega}) = \begin{cases} G_{cr}(\tilde{\omega}) / \sin \tilde{\omega} \tau & \omega_{k-1} < |\tilde{\omega}| \leq \omega_k \\ 0 & \text{otherwise} \end{cases} \quad (2.21)$$

を計算し、白色化相互相関関数  $\phi_{F_{lc}F_{cr}}(t)$  を用いて到来時間差  $\delta_k$  を推定する。

$$\phi_{F_{lc}F_{cr}}(t) = \text{IFFT} \left[ \frac{F_{lc}(\tilde{\omega}) F_{cr}^*(\tilde{\omega})}{|F_{lc}(\tilde{\omega})| |F_{cr}(\tilde{\omega})|} \right] \quad (2.22)$$

$$\delta_k = \arg \max_t [\phi_{F_{lc}F_{cr}}(t)] \quad (2.23)$$

このように、クロススペクトルをパワースペクトルで正規化することにより、信号の振幅スペクトル形状に依存せず、より高精度の時間差推定が可能となる。

以上の操作により、ある周波数帯域で最も優勢な雑音の到来時間差を、目的信号を完全に抑圧した信号から推定することができる。

### 2.2.3 雑音スペクトルの推定

前節において推定した到来時間差  $\delta_k$  を利用し、目的信号を完全に抑圧した信号より雑音のスペクトル  $N_k(\tilde{\omega})$  ( $\omega_{k-1} < |\tilde{\omega}| \leq \omega_k$ ) の推定を行う。

目的信号を完全に抑圧した減算型ビームフォーマ  $G_{lr}(\tilde{\omega})$ 、 $G_{cr}(\tilde{\omega})$  において焦点を雑音  $N_k(\tilde{\omega})$  の到来方向に合わせる。マイクロホンアレーの焦点は各周波数において感度が等しく、雑音スペクトルを正確に推定することができる。つまり  $G_{lr}(\tilde{\omega})$  において  $\tau = \delta_k$ 、 $G_{cr}(\tilde{\omega})$  において  $\tau = \frac{\delta_k}{2}$  と代入すれば、周波数帯域  $\omega_{k-1} < |\tilde{\omega}| \leq \omega_k$  において

$$G_{lr}(\tilde{\omega}) = N_k(\tilde{\omega}) \sin^2 \tilde{\omega} \delta_k \quad \omega_{k-1} < |\tilde{\omega}| \leq \omega_k \quad (2.24)$$

$$G_{cr}(\tilde{\omega}) = N_k(\tilde{\omega}) e^{j\tilde{\omega} \frac{\delta_k}{2}} \sin^2 \tilde{\omega} \frac{\delta_k}{2} \quad \omega_{k-1} < |\tilde{\omega}| \leq \omega_k \quad (2.25)$$

となる。しかしながら、 $G_{lr}(\tilde{\omega})$  においては  $\tilde{\omega} = \frac{n\pi}{\delta_k}$  ( $n$ :整数)、 $G_{cr}(\tilde{\omega})$  においても  $n$  が偶数となる場合は正確に雑音のスペクトルを推定することができない。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。

$$\hat{N}_k(\tilde{\omega}) = \begin{cases} G_{lr}(\tilde{\omega}) / \sin^2 \tilde{\omega} \delta_k & \sin^2 \tilde{\omega} \delta_k > \varepsilon_1 \\ G_{cr}(\tilde{\omega}) e^{-j\tilde{\omega} \frac{\delta_k}{2}} / \sin^2 \tilde{\omega} \frac{\delta_k}{2} & (\sin^2 \tilde{\omega} \delta_k \leq \varepsilon_1) \wedge (\sin^2 \tilde{\omega} \frac{\delta_k}{2} > \varepsilon_2) \\ G_{lr}(\tilde{\omega}) & \sin^2 \tilde{\omega} \frac{\delta_k}{2} \leq \varepsilon_2 \end{cases} \quad (2.26)$$

つまり、 $\sin \tilde{\omega} \delta_k$  の値が十分大きい場合は、マイクロホン間隔がより大きい  $G_{lr}(\tilde{\omega})$  を用いて雑音スペクトルの推定を行う。これはマイクロホン間隔が大きいほど、より高精度の雑音スペクトルの推定が可能になるためである。 $\sin \tilde{\omega} \delta_k$  が0に近づくに従って計算誤差が大きくなるため、そのような場合は  $G_{cr}(\tilde{\omega})$  を用いて雑音スペクトルを推定する。 $G_{cr}(\tilde{\omega})$  を用いてもなお雑音スペクトルを推定できない場合に限って、雑音スペクトルを近似的に推定する。

以上の雑音の到来方向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  $\hat{N}_1(\tilde{\omega}) \sim \hat{N}_K(\tilde{\omega})$  ( $\omega_{k-1} < |\tilde{\omega}| \leq \omega_k$ ,  $k = 1 \sim K$ ) を求める。最後にそれらを統合することにより、全帯域の雑音スペクトル  $\hat{N}(\omega)$  を求める。

## 2.2.4 雑音スペクトルの除去

本雑音除去アルゴリズムでは、周波数領域における減算であるスペクトルサブトラクション (SS) によって雑音の除去を行う [5]。

雑音成分の除去は、マイクロホンアレーの中央のマイクロホンで受音した信号  $c(t)$  のフーリエ変換  $C(\omega)$  から求めた振幅スペクトル  $|C(\omega)|$  から、先に推定した雑音の振幅スペクトル  $|\hat{N}(\omega)|$  を引き去ることにより実現できる。このとき目的信号の振幅スペクトルの推定値  $|\hat{S}(\omega)|$  を、Boll の SS を改良した非線形 SS

$$|\hat{S}(\omega)| = \begin{cases} |C(\omega)| - \alpha|\hat{N}(\omega)| & |C(\omega)| \geq \alpha|\hat{N}(\omega)| \\ \beta|C(\omega)| & \text{otherwise} \end{cases} \quad (2.27)$$

$$\hat{s}(t) = \text{IFFT} \left[ |\hat{S}(\omega)| e^{j \cdot \arg C(\omega)} \right] \quad (2.28)$$

により推定する。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。雑音スペクトルの推定誤差により、受音信号の振幅スペクトル  $|C(\omega)|$  よりも、雑音の振幅スペクトル  $|\hat{N}(\omega)|$  の方が大きくなる可能性も考えられる。その際、目的信号の振幅スペクトルの推定値  $|\hat{S}(\omega)|$  が負の値とならないように、SS に非線形機構を導入した。

## 第 3 章

# 計算機シミュレーションによる基礎的検証

本章では、計算機シミュレーションにより、本雑音除去アルゴリズムの基礎的な能力を検証する。

本雑音除去アルゴリズムの特徴は、同時刻に到来する複数雑音の除去を実現するために、周波数帯域分割型アルゴリズムを備えていることである。従って本章では、複数雑音のそれぞれの到来方向推定精度および雑音スペクトルの推定精度について重点的に検証する。

### 3.1 雑音除去アルゴリズムのパラメータ設定

本雑音除去アルゴリズムの受信系は、隣接マイクロホン間隔 10 cm、両端のマイクロホン間隔 20 cm の 3ch. 等間隔直線配列マイクロホンアレーとする。用いるデータのサンプリング周波数は 48 kHz とし、雑音除去アルゴリズムの各パラメータは、特に断りのない限り、表 3.1 に示す値に設定する。

ここでバンド幅とは、分割された各周波数帯域の帯域幅を表す。つまりバンド幅 1 pt とは周波数帯域を最も細かく分割して解析することに相当し、表 3.1 の設定においては、 $48000/1024 = \text{約 } 46.9 \text{ Hz}$  毎に方向推定および雑音スペクトルの推定を行うことを意味する。また窓関数は、小さな周波数成分の検出に適している Hanning 窓を用い、閾値  $\epsilon_1$ 、 $\epsilon_2$  は水町らの研究において定められた値とした。

次に本雑音除去アルゴリズムにおける空間分解能を示す。サンプリング周波数を 48 kHz とした場合、信号の隣接マイクロホン間の到来時間差と音源方向との間には表 3.2 のような関係が成り立つ。定式化において、右方向からの到来時間差を正としているので、負の時間差は信号が左方向から到来していることを示す。

表 3.1: 雑音除去アルゴリズムのパラメータ設定値

パラメータ	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 窓
バンド幅	1 pt
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

表 3.2: 信号の到来時間差と音源方向との関係

時間差 [pt]	方向 [°]	時間差 [pt]	方向 [°]
1	4	8	35
2	8	9	40
3	12	10	45
4	16	11	51
5	21	12	58
6	25	13	67
7	30	14	83

## 3.2 複数雑音に対する雑音除去実験

目的：

異なる方向から複数の雑音が同時に到来するような雑音環境下において、本雑音除去アルゴリズムの方向推定精度および雑音除去精度を検証する。ここでは、それぞれの雑音の周波数帯域は重複していないものとする。また、水町らの提案法(周波数帯域分割アルゴリズム導入前)と本手法(周波数帯域分割アルゴリズム導入後)の雑音除去結果の比較を行い、本手法の優位性を確認する。

### 3.2.1 シミュレーション条件

目的信号として、ATR 音声データベースに収録されている男性話者 mht のクリーンな日本語単母音/a/を用い、到来方向は正中面とした。また雑音として、帯域幅が 0.1 kHz、中心周波数が約 1.8 kHz、2.8 kHz、3.7 kHz の 3 つのランダム帯域雑音を計算機上で作成したものを用い、到来方向はそれぞれ左 21 °(到来時間差 -5 pt)、左 30 °(-7 pt)、右 45 °(10 pt) とした。データのサンプリング周波数は 48 kHz であり、16 bit で量子化されている。データ長は 9600 pt (0.2 sec) で、3 つの受信信号  $l(t)$ 、 $c(t)$ 、 $r(t)$  は雑音信号にそれぞれ適切な時間シフトを与え、計算機上で目的信号と加算して作成した。

### 3.2.2 方向推定結果

図 3.1 に方向推定結果を示す。上図 (a) は各雑音の周波数帯域における方向推定結果であり、下図 (b) は全ての周波数帯域で方向推定を行った結果である。周波数帯域を分割し、各帯域で方向推定を行うことにより、3 つの雑音の到来方向がほぼ正確に推定されていることがわかる。これに対して全周波数帯域で方向推定を行った場合、雑音の方向を 1 つしか決定することができず、推定した方向も必ずしも雑音方向になっていない。

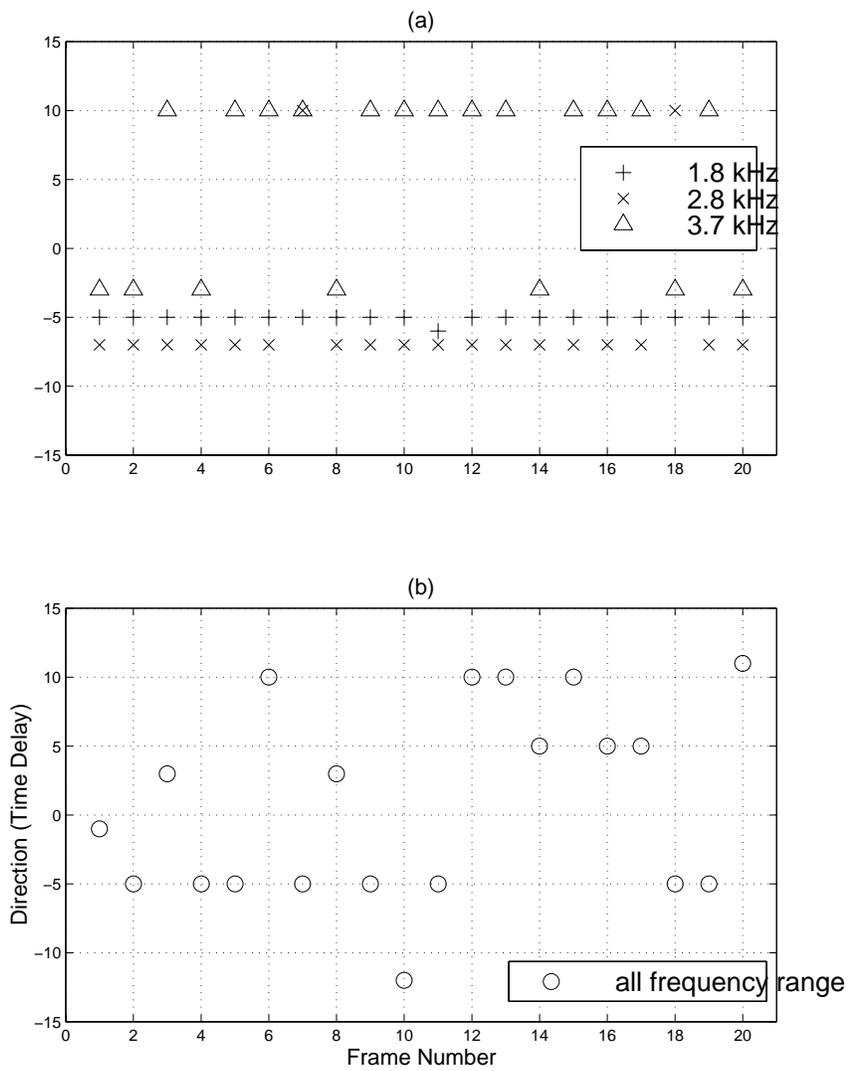


図 3.1: 方向推定結果 (a) 周波数帯域を分割し、各雑音の帯域で方向推定を行った場合 (b) 全周波数帯域で方向推定を行った場合

### 3.2.3 雑音除去結果

雑音除去結果を図 3.2 (スペクトル)、図 3.3 (波形) に示す。本手法で雑音除去を行った場合には、3 つの雑音がそれぞれ十分に推定され (図 3.2(c))、ほぼ完全に雑音が除去されたといえる (図 3.2(d)、図 3.3(c))。しかしながら水町らの提案法の場合では、全周波数帯域で 1 度しか方向推定を行っていないため、焦点の合っていない雑音は十分に推定することができず、雑音も除去しきれていない (図 3.2(e)、図 3.3(d))。

また雑音の評価尺度として、SNR (信号対雑音比) および LSD (対数スペクトル距離) を用い、以下の式を用いて求めた。

$$\text{SNR} = 10 \log_{10} \frac{\sum_n s^2(t_n)}{\sum_n \{s(t_n) - \hat{s}(t_n)\}^2} \quad (3.1)$$

$$\text{LSD} = \sqrt{\frac{1}{W} \sum_{\omega} \left(20 \log_{10} |S(\omega)| - 20 \log_{10} |\hat{S}(\omega)|\right)^2} \quad (3.2)$$

ここで  $s(t_n)$  はクリーンな音声、 $\hat{s}(t_n)$  は評価対象の音声とし、それぞれの振幅スペクトルを  $|S(\omega)|$ 、 $|\hat{S}(\omega)|$  とする。単位は共に dB であり、 $W = 6$  kHz とした。表 3.3 に雑音除去前および雑音除去後の SNR、LSD の値を示す。

表 3.3: SNR、LSD による雑音除去評価 (カッコ内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]	-11.1	15.0	0.4
LSD [dB]	27.4 (13.6)	18.2 (9.7)	21.5 (10.1)

上記の結果より、本手法が SNR を約 26 dB 増加させ、LSD を約 9 dB 減少させたことがわかる。また、水町らの提案法の値と比較することにより、周波数帯域分割型アルゴリズムの導入によって、複数雑音に対する雑音除去精度が向上したことが確認できる。

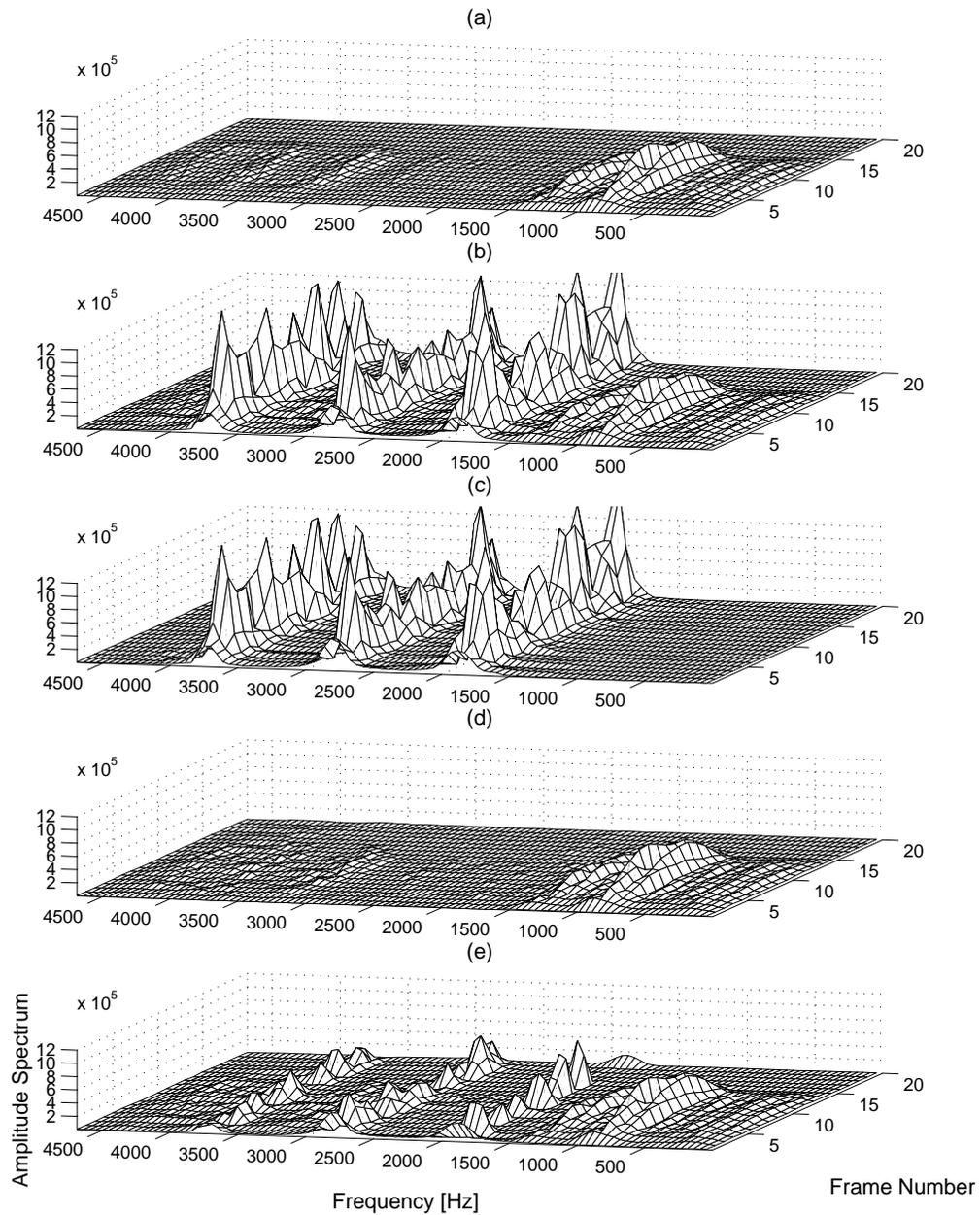


図 3.2: 雑音除去結果〈スペクトル〉(a) クリーンな音声 (b) 雑音を付加した音声 (c) 推定した雑音 (本手法) (d) 雑音除去後の音声 (本手法) (e) 雑音除去後の音声 (水町らの提案法)

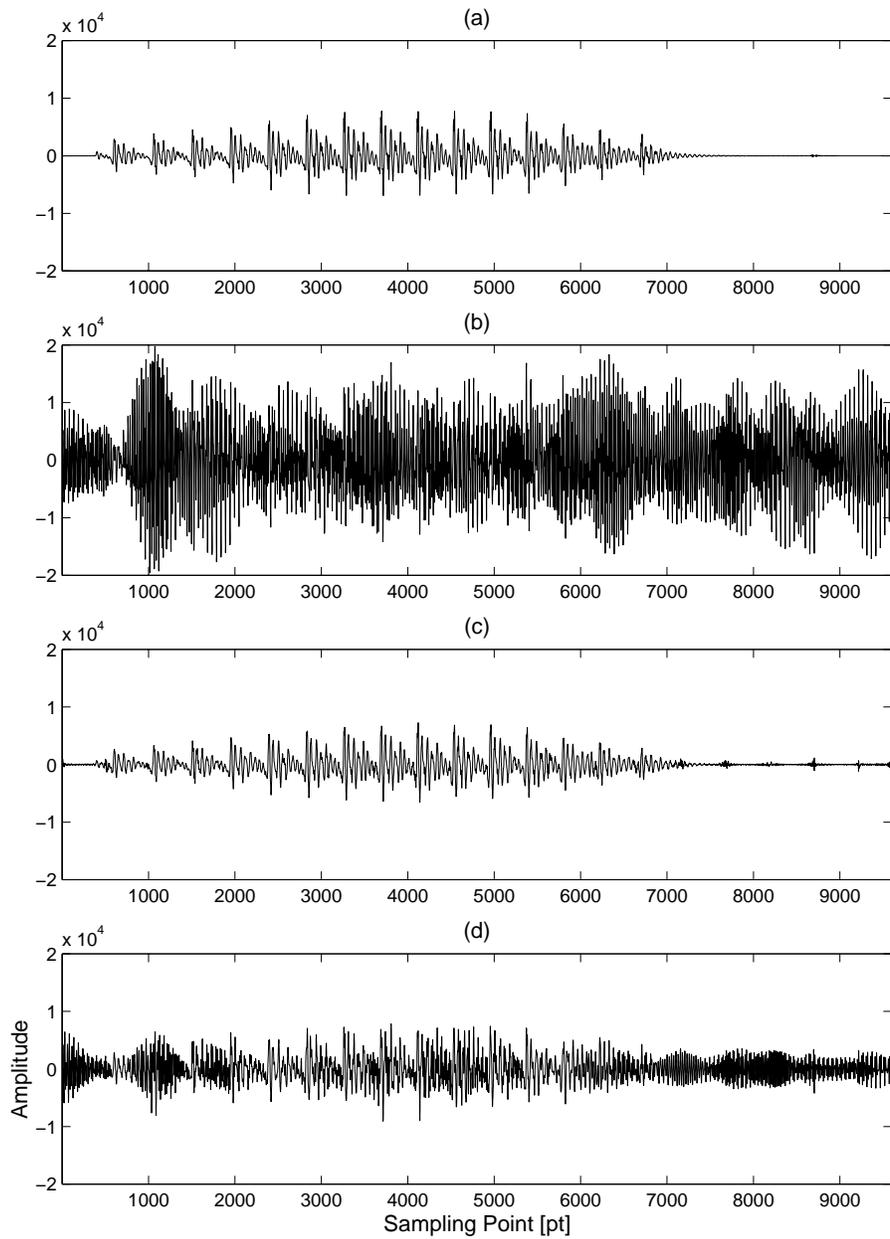


図 3.3: 雑音除去結果〈波形〉(a) クリーンな音声 (b) 雑音を付加した音声 (c) 雑音除去後の音声 (本手法) (d) 雑音除去後の音声 (水町らの提案法)

また、3つの雑音の周波数帯域および到来方向は変化させず、クリーンな音声に対してSNRが約0 dB、11 dBとなるように雑音を付加した音声を作成し、同様に雑音除去を行った。直前の結果(雑音除去前のSNRの値:約-11 dB)も含め、雑音除去後のSNR、LSDの値をグラフにしたものを図3.4に示す。実線は本手法(周波数帯域分割型アルゴリズム導入後)、点線は水町らの提案法(周波数帯域分割型アルゴリズム導入前)の値を表し、SNR(図3.4(a))では実線の値が高く、LSD(図3.4(b))では実線の値が低いことから、本手法の優位性が確認できる。水町らの提案法に比べ、SNRを平均して約8.1 dB、LSDを平均して約2.8 dB改善させ、雑音除去精度をより向上させたことがわかる。

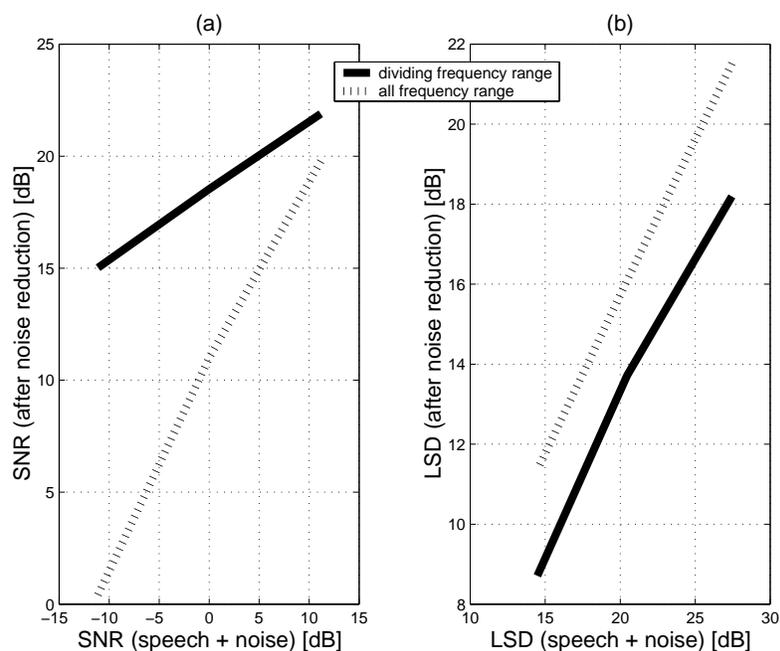


図 3.4: SNR、LSD の比較 : 実線は本手法 (周波数帯域分割型アルゴリズム導入後)、点線は水町らの手法 (周波数帯域分割型アルゴリズム導入前) の値

### 3.3 周波数帯域が重複している複数雑音に対する雑音除去実験

目的：

複数雑音の周波数帯域が重複している場合において、本雑音除去アルゴリズムの雑音除去精度を検証する。

#### 3.3.1 シミュレーション条件

目的信号として、ATR 音声データベースに収録されている男性話者 mht のクリーンな日本語単母音/a/を用い、到来方向は正中面とした。また雑音として、帯域幅が1 kHz、中心周波数が約2 kHz、2.5 kHz の2つのランダム帯域雑音を計算機上で作成したのを用い、到来方向はそれぞれ左右45°(到来時間差10 pt)、左30°(-7 pt)とした。つまりこれらの雑音は2~2.5 kHzにおいて周波数帯域が重複していることになる。データ作成の条件は前節と同様である。

#### 3.3.2 方向推定結果

図3.5に方向推定結果を示す。上図(a)は周波数帯域を分割し、約1.7、2.2、2.7 kHz成分で方向推定を行った結果であり、それぞれ+、×、△で示してある。1.7 kHz成分は中心周波数2 kHzの雑音方向を示すはずであり、ほとんど正しく10 ptを示している。同様に2.7 kHz成分は中心周波数2.5 kHzの雑音方向を示すはずであり、すべてのフレームで正しく-7 ptを示している。そして2.2 kHz成分は2つの雑音が重複する帯域であるが、どの雑音の方向でもない12 ptを示した。これはその周波数帯域で重畳された雑音が最大となる焦点の方向であると考えられる。つまり、片方の雑音に焦点を合わせ、もう一方の雑音が十分に抽出できない指向特性よりも、2つの雑音が共に十分に抽出できるような指向特性となる焦点が選ばれた、といえる。これに対して、全周波数帯域で方向推定を行った場合(下図(b)○で示す)、2つの雑音のどちらかの方向が示されていることがわかる。

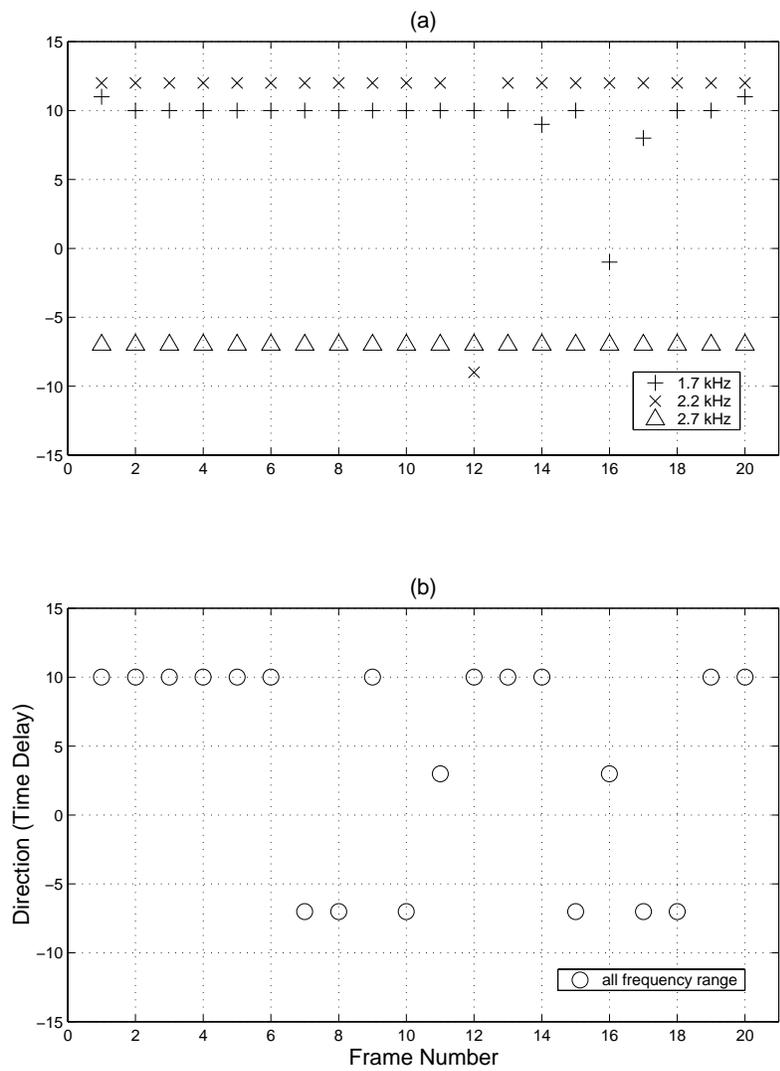


図 3.5: 方向推定結果 (a) 周波数帯域を分割し、各雑音の帯域で方向推定を行った場合 (b) 全周波数帯域で方向推定を行った場合

### 3.3.3 雑音除去結果

雑音除去結果を図 3.6(スペクトル)、図 3.7(波形) に示す。図 3.6(d) からわかるように、全体的にほぼ雑音除去が為された。また重複した帯域の雑音 (2 ~ 2.5 kHz) もほとんどは抑圧されたといえる。これは焦点以外の方向から到来する雑音も、減算型ビームフォーマの指向特性よりある程度除去できること、また周波数帯域を分割し各帯域において雑音がトータルで最大となる焦点が決定できたためである。全周波数帯域で方向推定を行う水町らの提案法では、後者の特性がないため推定される雑音も十分でなく、本手法ほど雑音除去されなかった (図 3.6(e))。

また表 3.4 に雑音除去前および雑音除去後の SNR、LSD の値を示す。この結果からも、周波数帯域分割型アルゴリズムを導入した本手法が、水町らの提案法よりも雑音除去精度が高いことが確認できる。

表 3.4: SNR、LSD による雑音除去評価 (カッコ内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]	-8.7	10.2	0.6
LSD [dB]	26.1 (14.5)	16.1 (6.9)	18.8 (8.1)

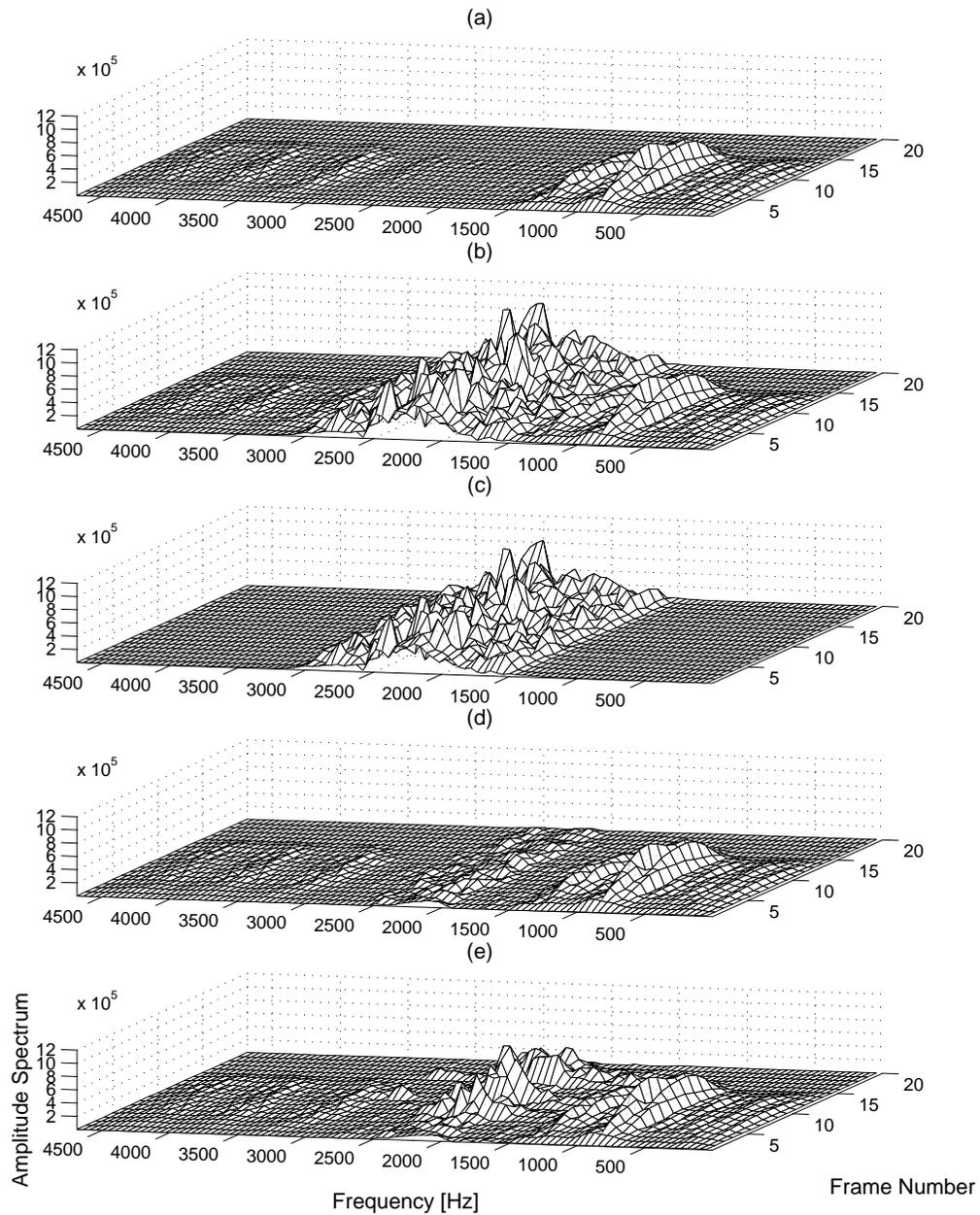


図 3.6: 雑音除去結果〈スペクトル〉(a) クリーンな音声 (b) 雑音を付加した音声 (c) 推定した雑音 (本手法) (d) 雑音除去後の音声 (本手法) (e) 雑音除去後の音声 (水町らの提案法)

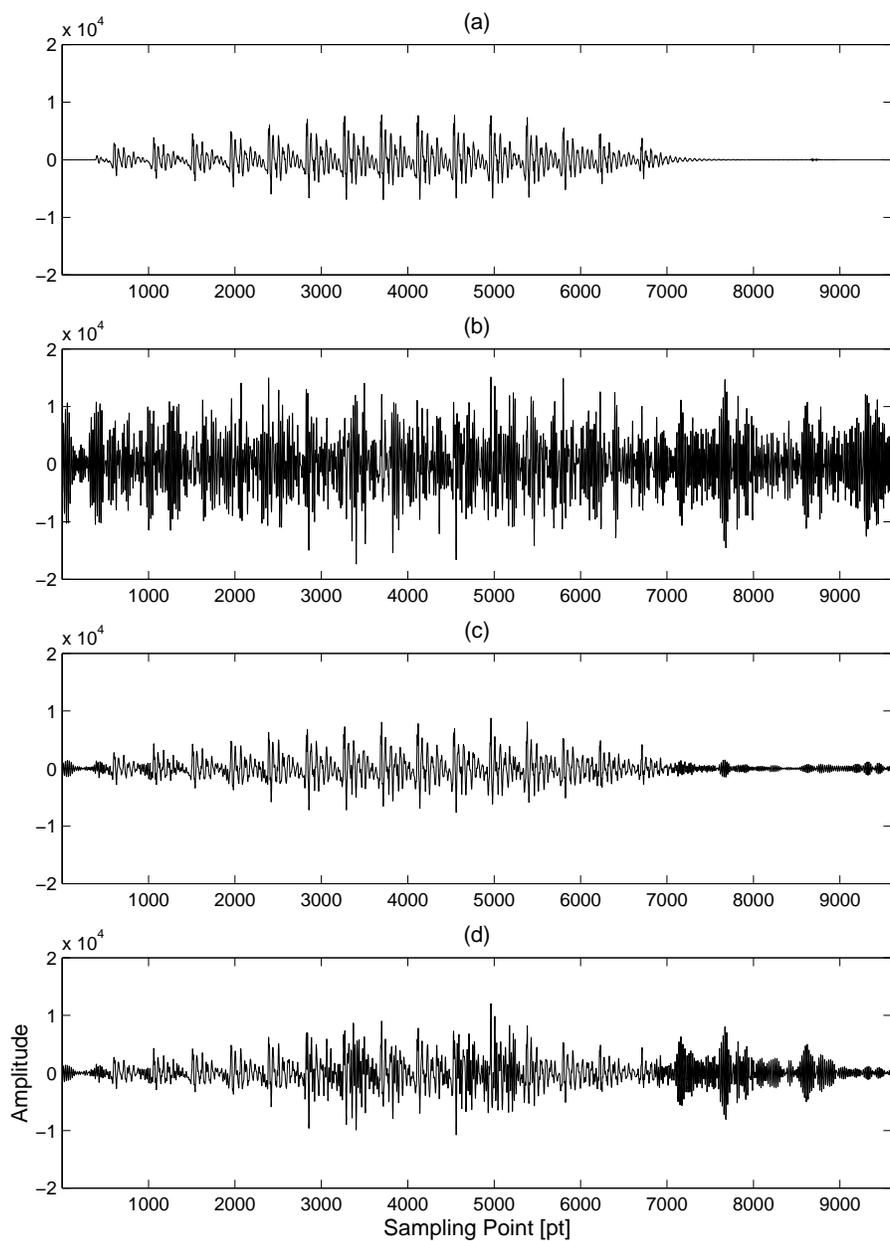


図 3.7: 雑音除去結果〈波形〉(a) クリーンな音声 (b) 雑音を付加した音声 (c) 雑音除去後の音声 (本手法) (d) 雑音除去後の音声 (水町らの提案法)

## 3.4 音声が多音存在する場合に対する雑音除去実験

目的：

目的信号以外にも音声が多音存在する場合、つまり雑音が音声である場合に対する本雑音除去アルゴリズムの雑音除去精度を検証する。音声の周波数帯域はほぼ同じ場所に集中するので、これにより目的音と雑音の周波数が重複した場合における雑音除去精度の調査を行う。

### 3.4.1 シミュレーション条件

目的信号として、ATR 音声データベースに収録されている男性話者 mht の日本語文章「通訳電話国際会議事務局です」を用い、到来方向は正中面とした。また雑音として、女性話者 faf の「それではぜひ送っていただきたいのですが」、女性話者 fkn の「ご住所をおしえていただけますでしょうか」、男性話者 mau の「題目の締め切りを教えてください」を用い、到来方向はそれぞれ左 67°、左 30°、右 45°とした。これは同時に発話された4つの文章から、1つの文章だけを抽出することに相当する。データ作成の条件は前節と同様である。

### 3.4.2 雑音除去結果

雑音除去結果を図 3.8(スペクトル)、図 3.9(波形) に示す。図 3.8(c) からわかるように、500 Hz 以下の低周波数帯域においては雑音のスペクトルが十分に推定されていない。これはマイクロホンアレーの指向特性が、低周波数帯域においては鋭さを欠くためである。しかしながら、スペクトルおよび波形を見ても、雑音除去後の音声は元のクリーンな音声に近づいたとは言い難い。

そこで図 3.10 にスペクトログラムを示す。雑音を付加した音声 (b) には様々な周波数に成分が存在することがわかるが、雑音除去後の音声 (c) では余分な成分が除去され、見た目にも元のクリーンな音声 (a) に近づいたことがわかる。また表 3.5 からわかるように LSD の値も約 7 dB 改善され、目的音と雑音の周波数帯域が重複している場合でも本手法が十分に働くことが確認された。

表 3.5: SNR、LSD による雑音除去評価 (カッコ内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)
SNR [dB]	-3.4	-1.0
LSD [dB]	21.4	14.3

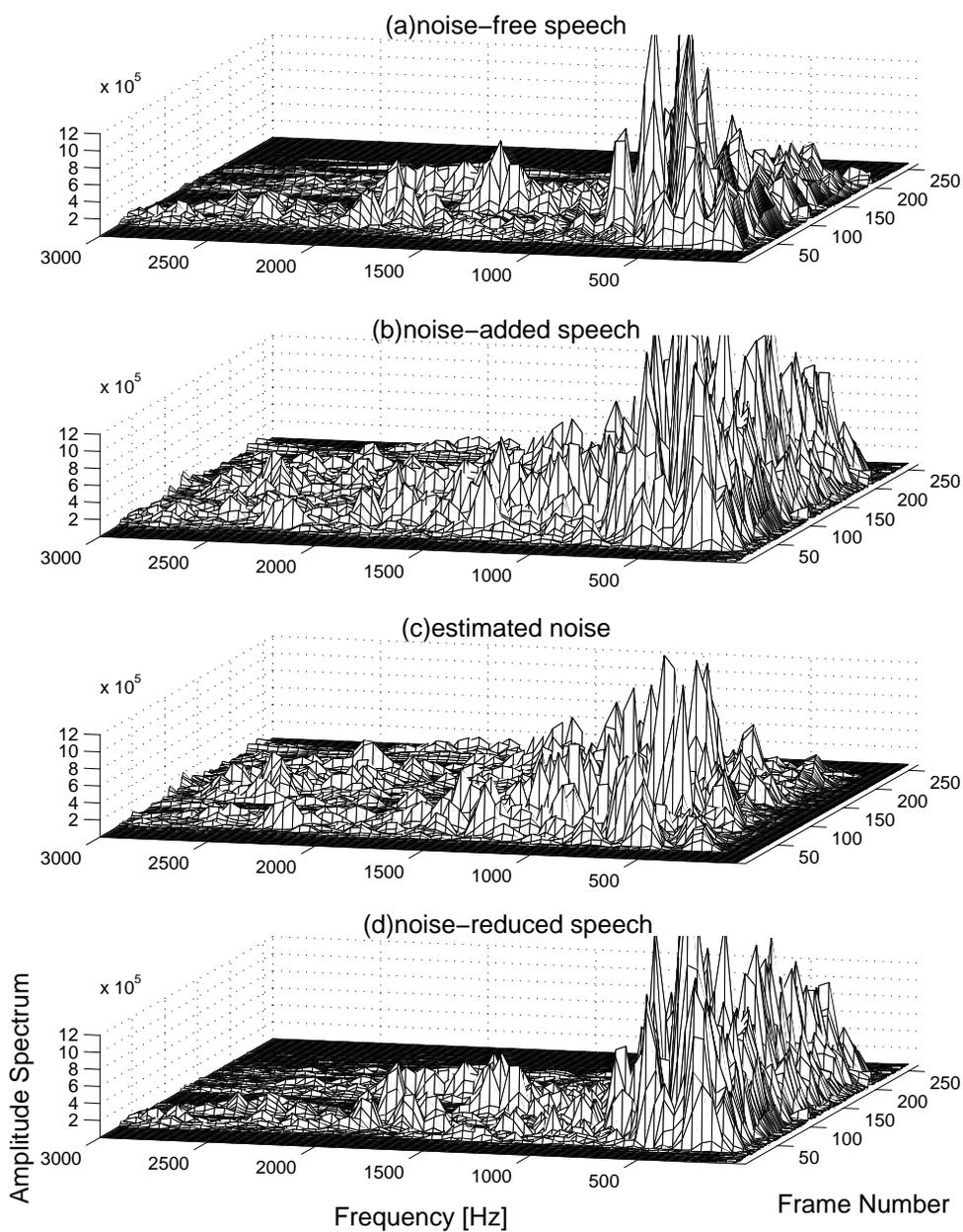


図 3.8: 雑音除去結果〈スペクトル〉 (a) クリーンな音声 (b) 雑音を付加した音声 (c) 推定した雑音 (本手法) (d) 雑音除去後の音声 (本手法)

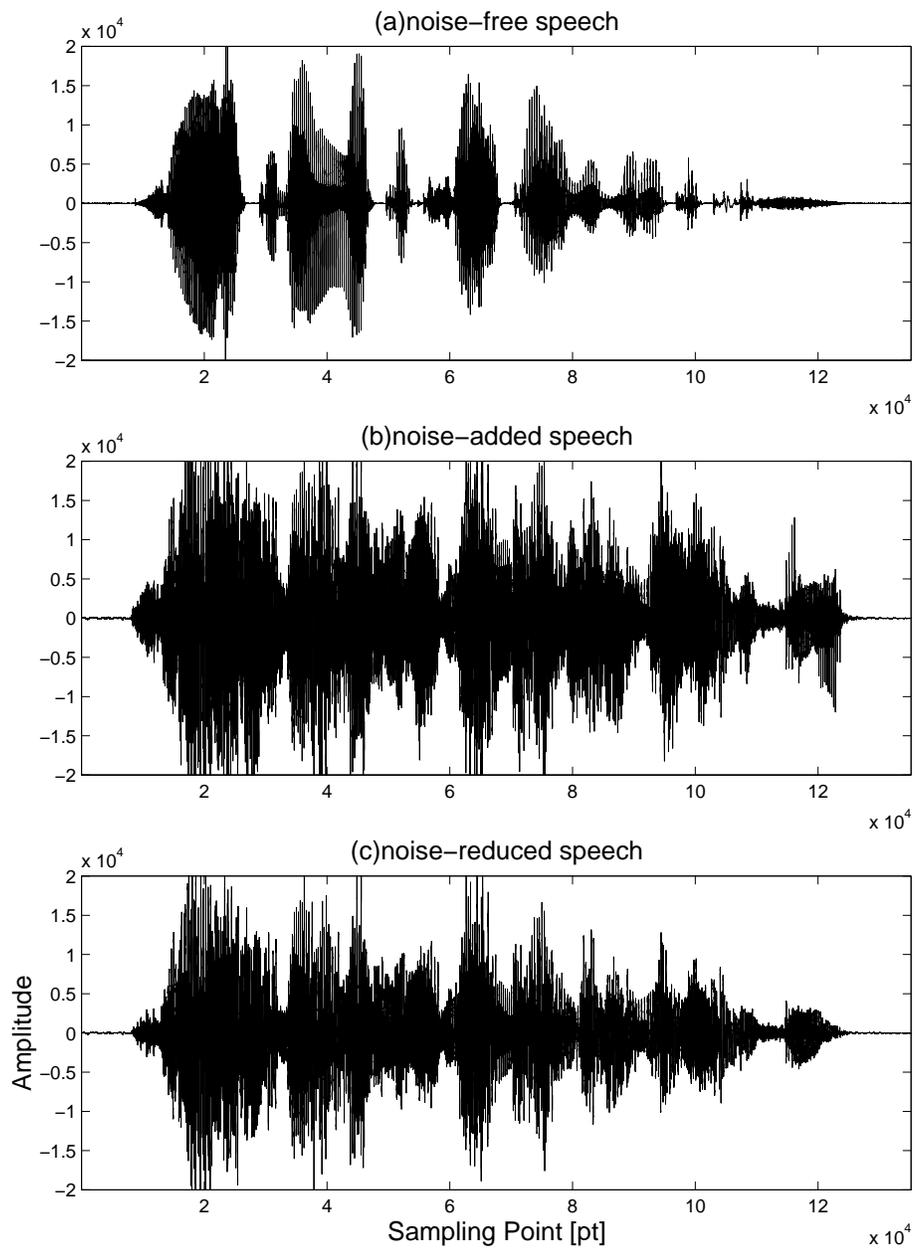


図 3.9: 雑音除去結果〈波形〉 (a) クリーンな音声 (b) 雑音を付加した音声 (c) 雑音除去後の音声 (本手法)

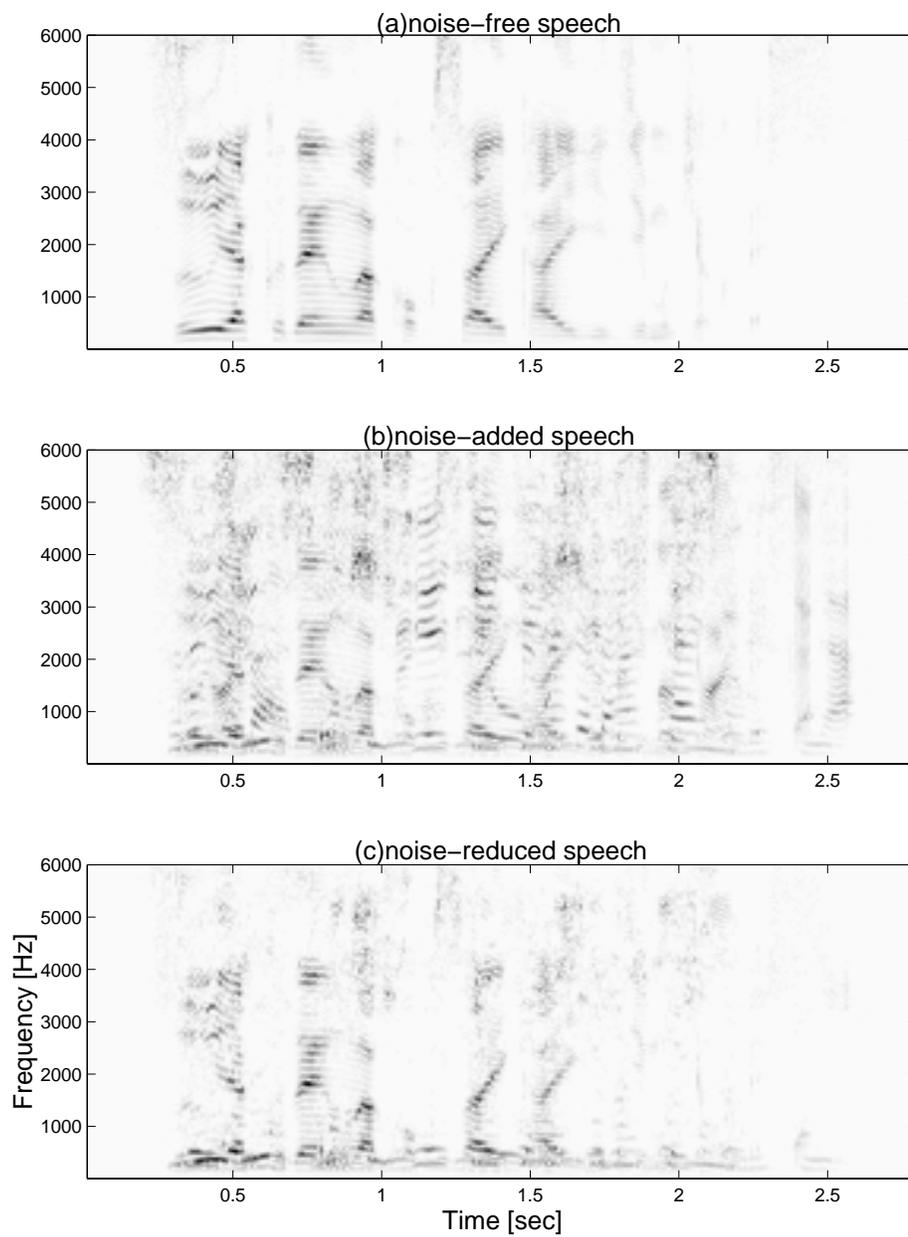


図 3.10: 雑音除去結果〈スペクトログラム〉(a) クリーンな音声 (b) 雑音を付加した音声 (c) 雑音除去後の音声 (本手法)

## 第 4 章

# 実環境下における有効性検証

第 2 章において本雑音除去アルゴリズムの定式化を行い、第 3 章では本手法の基礎的な性能評価を行った。その結果、本手法は高精度の雑音除去が可能であり、特に複数雑音に対して水町らの提案法よりも雑音除去能力が高いことを確認した。しかし第 3 章の評価実験は、実環境下において最も問題となる暗騒音や残響が存在しない環境を想定した、計算機上での実験である。そこで本章では、本手法の実環境適応性を調査するため、実環境下における雑音除去実験を行う。

### 4.1 実環境下における複数雑音除去実験

目的：

実環境下における雑音除去実験により、本手法の実環境適応性について検討する。また、周波数帯域分割型アルゴリズムが実環境下においても有効に機能するかどうかを調査する。

実験条件：

雑音除去実験は暗騒音がほとんどなく、残響の少ない防音室で行った。防音室は日常生活で使用する一般的な部屋と比較すると、理想的環境に近い部屋であるといえる。

防音室は  $4.0 \text{ m} \times 4.5 \text{ m} \times 2.4 \text{ m}$  の大きさで、遮音処理、防音処理が施された部屋であるが、無響室とは大きく異なり反射音が存在する。防音室の残響特性は、 $500 \text{ Hz}$  以下の帯域において残響時間が  $100 \sim 300 \text{ msec}$ 、 $500 \text{ Hz}$  以上の帯域において  $50 \text{ msec}$  となり、低周波数帯域における残響時間はかなり長い。よって音声の基本周波数付近に反射音が存在することがわかる。

実環境下における雑音除去実験では、クリーンな音声と雑音をそれぞれ別のスピーカー

から実音場に放射した。それをアナログターミナルを通してサンプリング周波数 48 kHz、16 bit 量子化でデータ収録ソフトウェアを用いてディスクレコーディングし、そのデータに対して雑音除去を行った。雑音除去アルゴリズムにおけるパラメーター設定値は、前章と同様である。また、受信系は隣接マイクロホン間隔 10 cm の 3ch. 等間隔直線配列マイクロホンアレーを用いた。データ収録に用いた機器を表 4.1 に示す。

表 4.1: データ収録に用いた機器一覧

機器	メーカー	型番
DAT Deck	SONY	DTC-57ES
Power Amp. (speech)	SANSUI	AU- $\alpha$ 907MR
Power Amp. (noise)	SONY	TA-F333ESA
Loud Speaker (speech)	AURATONE	5PSC
Loud Speaker (noise)	audio-technica	AT-NF602
Speaker Stand	ULTIMATE	TS-40
DAT (noise)	SONY	DT-60RA
Microphone	SONY	ECM-44B
Mic. Stand	SONY	B-305B
Mic. Amp.	TASCAM	MX-4
Analog Terminal	Thinknet	DF-2021Z
Software	Thinknet	MDLOG-UT

目的信号を放射するスピーカー (直径 10 cm) をマイクロホンアレーに対して正中面方向、雑音 1 を放射するスピーカーを左 45 °方向、雑音 2 を放射するスピーカーを右 30 °方向に設置した。マイクロホンアレーと目的信号のスピーカーとの距離を 0.5 m、雑音のスピーカーとの距離を 2.0 m とした。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「ガススプレーの噴射音」、雑音 2 として「携帯電話の着信音」を用いた。「ガススプレーの噴射音」は 4~8 kHz にわたる帯域雑音であり、「携帯電話の着信音」は約 2.8、5.6 kHz 付近に大きなパワーを持つ断続音である。

## 雑音除去結果：

雑音除去結果を図 4.1(スペクトル)、図 4.2(波形) に示す。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。図 4.1(b) より、約 2.8、5.6 kHz 付近に携帯電話の着信音の成分が、また 4~8 kHz にわたってガススプレーの噴射音の成分が付加されていることがわかる。本手法においては、これらの雑音スペクトルがほぼ推定されていることがわかる (図 4.1(c))。また正中面方向から到来しなかった、目的信号の反射音や残響もある程度 雑音成分として推定されていることが確認できる。しかし、計算機シミュレーションほど十分に雑音は除去されず、雑音除去後の音声にも特にガススプレーの噴射音の影響が残ってしまった (図 4.1(d)、図 4.2(c))。これは正中面方向から到来する、雑音の反射音や残響が目的音と見なされたためであると考えられる。水町らの提案法においては、2 方向から到来する雑音を同時に推定できず、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 4.1(e))。

また表 4.2 に雑音除去前および雑音除去後の SNR、LSD の値を示す。この結果から、実環境下における雑音除去実験においても、本手法の優位性が認められる。

表 4.2: 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

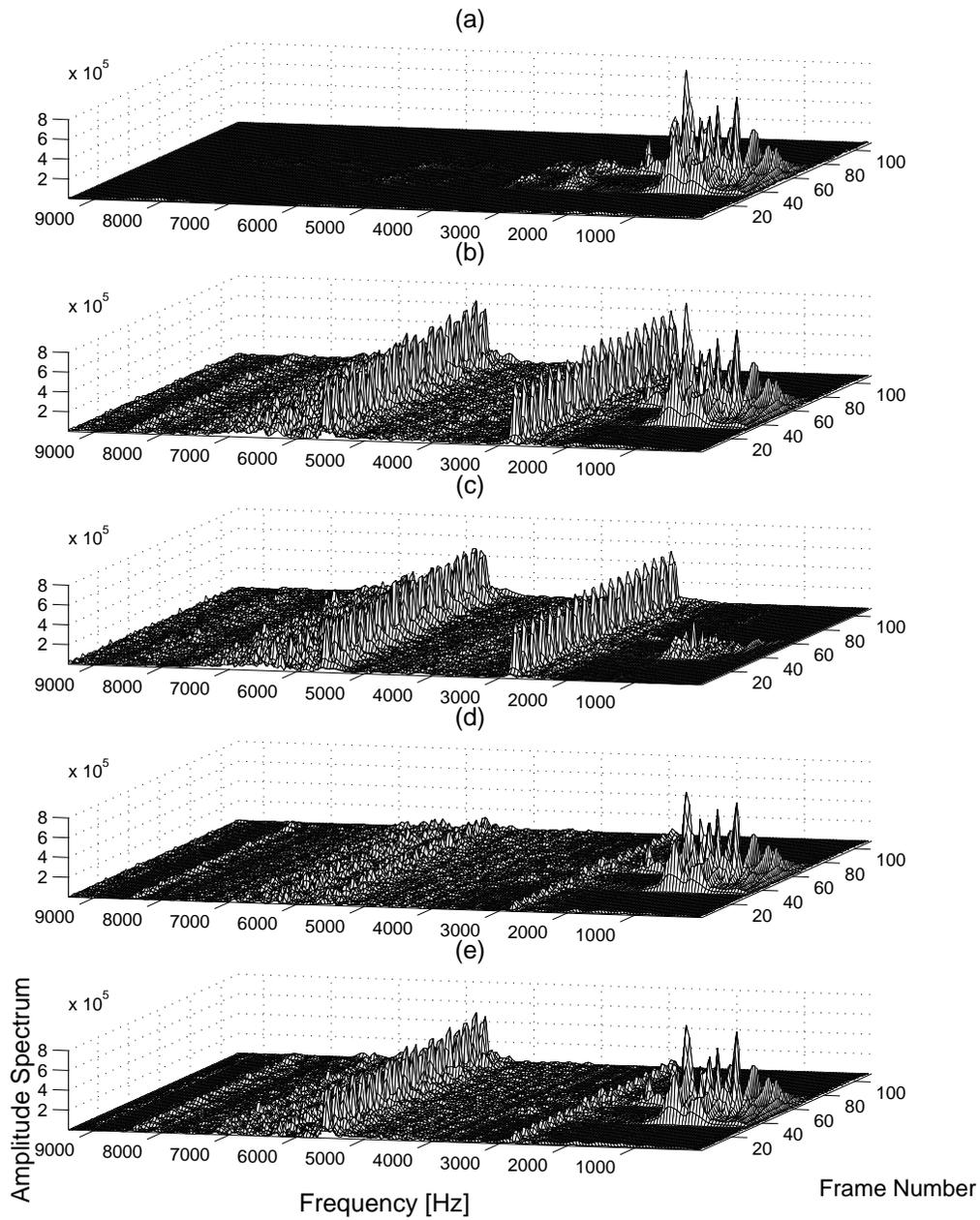


図 4.1: 雑音除去結果〈スペクトル〉(a) クリーンな音声 (b) 雑音を付加した音声 (c) 推定した雑音 (本手法) (d) 雑音除去後の音声 (本手法) (e) 雑音除去後の音声 (水町らの提案法)

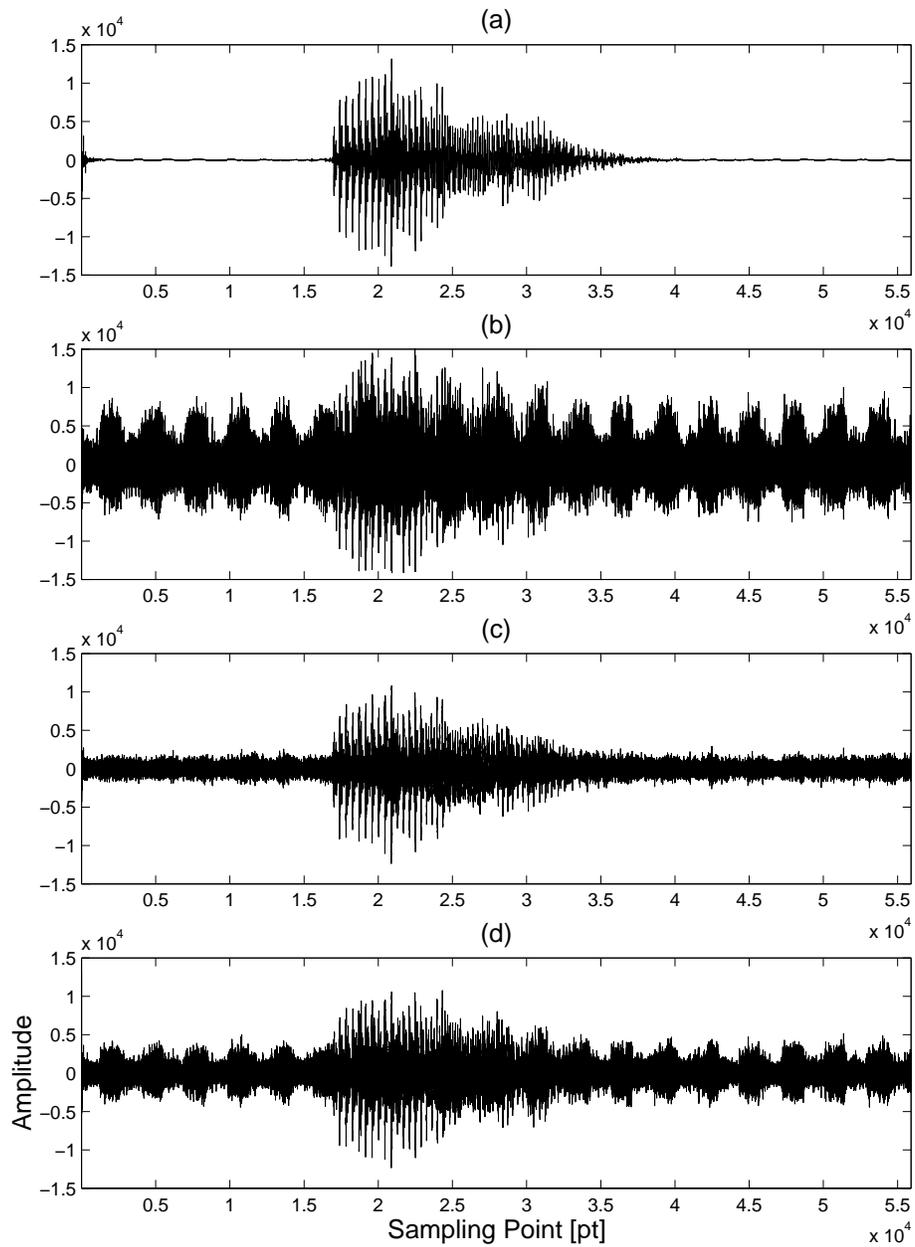


図 4.2: 雑音除去結果〈波形〉(a) クリーンな音声 (b) 雑音を付加した音声 (c) 雑音除去後の音声 (本手法) (d) 雑音除去後の音声 (水町らの提案法)

## 4.2 近接音源に対する雑音除去実験

### 目的：

水町らの研究においては、目的音源はマイクロホンアレーから十分離れていると仮定しており、近接音源に関する調査は行われていない。しかし、例えば ASR など実環境下での応用を考えた場合、マイクロホンアレーと話者(目的音源)との距離は非常に近いと考えられ、近接音源を考慮に入れることが重要となってくる。また本雑音除去アルゴリズムは、目的信号が正中面から到来すると仮定しているが、実際には必ずしも正中面から到来しないことが予想される。そこで本節では、目的信号の距離と到来方向誤差に関する雑音除去精度の調査を行う。

### 実験条件：

雑音除去実験は前節と同じ防音室内で行い、データ収録条件および用いた機器も前節と同様である。

目的信号を放射するスピーカーとマイクロホンアレーとの距離を 2 m、1 m、0.5 m と近づけ、各距離において 0°、2°、4°、8°、12°、16°、20° と正中面からスピーカーの位置をずらしてデータ収録を行った。なお雑音除去アルゴリズムにおいて、目的信号の到来方向は正中面と常に仮定したままである。また前節と同様に、雑音 1 を放射するスピーカーを左 45°、雑音 2 を放射するスピーカーを右 30° にそれぞれ 2 m の位置に設置し、雑音の位置、ボリュームは変化させず、受信信号の SN 比が 3 つの距離でほぼ等しくなるように、目的信号の入力データの振幅を調整した。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」、雑音 2 として「携帯電話の着信音」を用いた。

### 実験結果：

雑音除去後の LSD の値を図 4.3 に示す。実線は目的信号のスピーカーとマイクロホンアレーとの距離が 2 m、破線は 1 m、点線は 0.5 m の値であり、横軸は目的信号方向が正中面からどれだけ外れているかを表す。また基準となるクリーンな音声は、各距離において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。

この結果より、目的信号方向が正中面から外れるにつれ LSD の値が増加し、雑音除去精度が低下していることがわかる。これは正中面から到来しない信号は雑音として見なされるためであり、本手法の理論と一致する。しかしながら、3 本の線にそれほど大きな距離の差はないことから、目的信号のスピーカーとマイクロホンアレーとの距離が近づいて

も、雑音除去精度の著しい変化はないことがわかった。よって本手法が、目的音源が 2 m と十分離れている場合でも、0.5 m と近接音源である場合でも対応可能であることが確認できた。

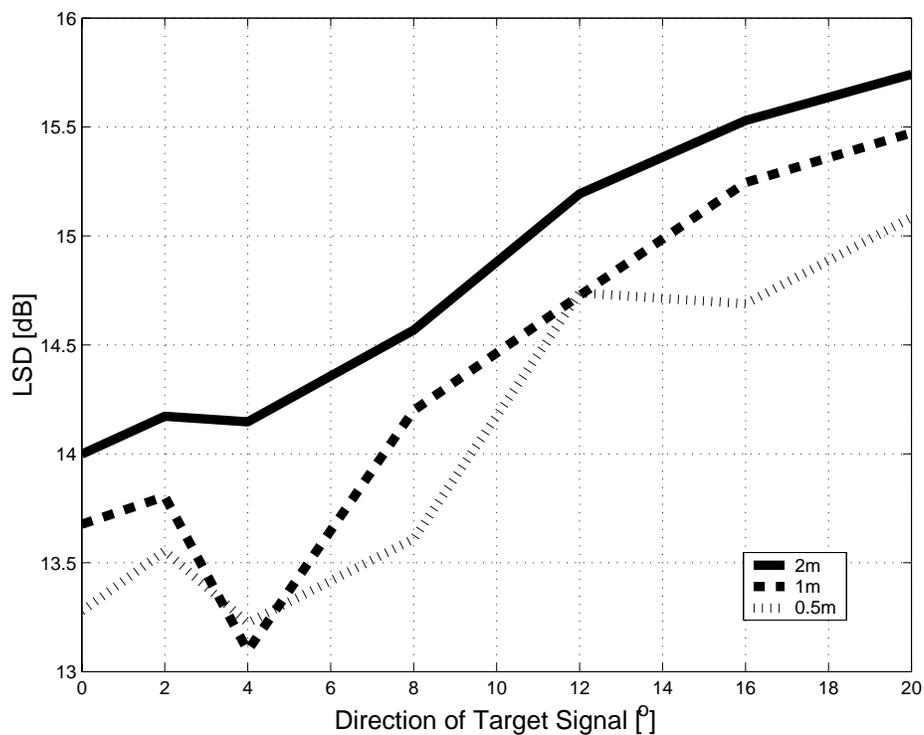


図 4.3: 目的信号の距離と方向誤差に関する雑音除去評価：実線は目的信号との距離 2 m、破線は 1 m、点線は 0.5 m を表す

# 第 5 章

## 結論

### 5.1 本論文の要約

本論文では、小規模マイクロホンアレーを用いた従来の雑音除去法では扱うことが困難であった、同時刻に異なる方向から到来する複数雑音に対してロバストな雑音除去法を構築した。

本研究では、複数雑音に対応させるための周波数帯域分割型アルゴリズムを、突発的な非定常雑音に対して精度の高い水町らの提案法に導入した。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、より実環境に対応できる雑音除去法の確立を目指した。周波数帯域分割型アルゴリズムは、分割した各周波数帯域で最も優勢な雑音の到来方向、スペクトルを推定することができ、複数雑音の周波数帯域が各々異なっている場合に特に有効に働く。水町らの提案法では雑音源は1つであると仮定し、全周波数帯域で雑音の到来方向、スペクトルの推定を行っていたため、雑音スペクトルが十分に推定できていなかった。

複数雑音に対する計算機シミュレーションおよび実環境下における雑音除去実験により、本雑音除去アルゴリズムの有効性、および水町らの提案法との比較によって優位性が確認できた。また目的音源が近接音源の場合でも、雑音除去精度に大きな変化はなく、話者とマイクの距離が近くなるような ASR などにも応用できる可能性も示した。

## 5.2 今後の課題

本雑音除去アルゴリズムの実用性を高めるためにも、今後、以下に挙げる課題を解決することが望ましい。それによって、本手法がハンズフリー型 ASR のフロントエンドや補聴器などへの実用化に、さらに近づくことが期待される。

### 残響環境下への対応

本研究では、残響時間の短い防音室内において雑音除去実験を行ったが、実際に雑音除去が必要とされる環境の残響時間は防音室内に比べて長い。本雑音除去アルゴリズムの残響への有効性を明確にするためにも、様々な残響時間および残響特性に対して、雑音除去能力を定量的に評価することが望ましい。

反射音や残響が存在する場合には、目的信号の反射音や残響が雑音として見なされたり、逆に雑音が目音として見なされたりすることは、本雑音除去アルゴリズムにおいて理論的にあり得ることであり、本論文においてもそれを明らかにした。また残響によって雑音の方向推定誤差が生じ、それに伴う推定した雑音スペクトルの誤差のために、雑音除去精度が低下することが水町らの研究において確認されている。よって、実環境下で本雑音除去アルゴリズムを有効に働かせるためには、残響に対してもロバストに改良することが必要である。

### リアルタイム雑音除去システムの実現

解析的にビームフォーミングを行う水町らの提案法は、適応フィルターを用いる従来の減算型ビームフォーミング手法よりも演算量は少ないとされる。しかしながら、本手法においては周波数帯域分割型アルゴリズムの導入により、分割した帯域の数だけ同様の演算を行わなければならない。単純に考えて、周波数帯域を 10 分割した場合、水町らの提案法の 10 倍の演算量が必要となる。

リアルタイム処理を前提に実装を行う場合には、演算量と雑音除去精度の関係をより詳細に調査し、雑音除去精度に大きな影響を与えないような処理は簡略化することが重要である。

# 謝辞

本研究を行うにあたり、終始多大なる御指導、御鞭撻を賜りました北陸先端科学技術大学院大学 情報科学研究科 赤木 正人 教授に深く感謝の意を表します。並びに本研究の過程において多くの御協力を賜りました、本学の教官の皆様にも心より感謝致します。

本研究を進める上で、有益な御議論、御助言をして頂いた ATR 音声言語通信研究所 水町 光徳 博士、赤木研究室の伊藤 一仁 氏、石本 祐一 氏に厚く御礼申し上げます。また日頃から有益な助言と激励を頂いた、赤木研究室の平成 13 年度修了生、並びに諸先輩方、後輩達に深く感謝致します。

最後に大学院における研究生生活を温かく見守ってくれた、両親、弟に心から感謝致します。

## 参考文献

- [1] 鹿野 清宏, 中村 哲, 伊勢 史朗, “音声・情報のデジタル信号処理”, 昭光堂 (1997)
- [2] 大賀 寿朗, 山崎 芳男, 金田 豊, “音響システムとデジタル信号処理”, 電子情報通信学会 (1995)
- [3] 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)
- [4] Masato Akagi and Mitsunori Mizumachi, “NOISE REDUCTION BY PAIRED MICROPHONES”, Proc.EUROSPEECH '97, Rhodes, Greece, Vol.1, pp.335-338 (1997)
- [5] S.F.Boll, “Suppression of Acoustic Noise in Speech Using Spectral Subtraction”, IEEE Trans.ASSP, Vol.27, No.2, pp.113-120 (1979)