

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4989967号  
(P4989967)

(45) 発行日 平成24年8月1日(2012.8.1)

(24) 登録日 平成24年5月11日(2012.5.11)

(51) Int.Cl.		F I	
HO4R	3/00 (2006.01)	HO4R	3/00 320
G10L	21/02 (2006.01)	G10L	21/02 101A
HO4R	25/00 (2006.01)	HO4R	25/00 L
HO4R	1/40 (2006.01)	HO4R	1/40 320A

請求項の数 21 (全 35 頁)

(21) 出願番号	特願2006-517910 (P2006-517910)	(73) 特許権者	501472630
(86) (22) 出願日	平成16年7月12日 (2004.7.12)		コクレア リミテッド
(65) 公表番号	特表2007-525865 (P2007-525865A)		Cochlear Limited
(43) 公表日	平成19年9月6日 (2007.9.6)		オーストラリア国 ニュー サウス ウェ
(86) 国際出願番号	PCT/BE2004/000103		ールズ 2109、マッコリー ユニヴァ
(87) 国際公開番号	W02005/006808		ーシティ、ユニヴァーシティ アヴェニュー
(87) 国際公開日	平成17年1月20日 (2005.1.20)		ー 1
審査請求日	平成19年7月10日 (2007.7.10)	(74) 代理人	110000246
(31) 優先権主張番号	2003903575		特許業務法人OFH特許事務所
(32) 優先日	平成15年7月11日 (2003.7.11)	(72) 発明者	ドクロ, シモン
(33) 優先権主張国	オーストラリア (AU)		ベルギー王国 シルデ ビー-2970,
(31) 優先権主張番号	2004901931		グレーネ バンデリング 19
(32) 優先日	平成16年4月8日 (2004.4.8)	(72) 発明者	スプリエット, アン
(33) 優先権主張国	オーストラリア (AU)		ベルギー王国 イゼゲン ビー-8870
			, ニューストラート 20

最終頁に続く

(54) 【発明の名称】 ノイズ低減のための方法および装置

(57) 【特許請求の範囲】

【請求項1】

音声信号における雑音を低減させる方法であって、  
 少なくとも2つのバージョンの前記音声信号を第1のフィルタで受け取るステップと、  
 前記第1のフィルタが、所望の信号およびノイズ寄与を含む音声参照信号と、音声漏れ寄与およびノイズ寄与を含む少なくとも一つのノイズ参照信号とを出力するステップと、  
 前記少なくとも一つのノイズ参照信号にフィルタリング操作を行うステップと、  
 前記音声参照信号から前記少なくとも一つのノイズ参照信号を差し引き、低減されたノイズを有する前記音声信号の出力バージョンを提供するステップと、を含み、  
 前記少なくとも一つのノイズ参照信号の前記フィルタリング操作は、前記音声信号の前記出力バージョンにおける音声歪みエネルギーおよび残留ノイズエネルギーの加重和を最小にするよう構成されたフィルタ係数をもつ一つまたは複数のフィルタで行われ、前記音声歪みエネルギーは前記音声漏れ寄与のエネルギーであり、前記残留ノイズエネルギーは前記音声参照信号および少なくとも一つのノイズ参照信号中のノイズ寄与のエネルギーである、ノイズを低減するための方法。

【請求項2】

少なくとも2つのマイクロホンが使用され、  
 前記少なくとも2つのマイクロホンで前記音声信号を受け取るステップと、  
 前記少なくとも2つのマイクロホンのそれぞれから前記音声信号の一つのバージョンを前記第1のフィルタに提供するステップと、

を含む、請求項 1 に記載のノイズを低減するための方法。

【請求項 3】

前記第 1 のフィルタが、ビームフォーマー・フィルタおよびブロッキング・マトリックス・フィルタを含む空間プリプロセッサ・フィルタである、請求項 1 または 2 に記載のノイズを低減するための方法。

【請求項 4】

前記音声参照信号が前記ビームフォーマー・フィルタにより出力されるステップと、前記少なくとも一つのノイズ参照信号が前記ブロッキング・マトリックス・フィルタにより出力されるステップと、

を含む、請求項 3 に記載のノイズを低減するための方法。

10

【請求項 5】

前記音声参照信号からフィルタされた少なくとも一つのノイズ参照信号を差し引くステップを実施する前に前記音声参照信号を遅延させるステップを含む、請求項 1 から 4 のいずれか 1 項に記載のノイズを低減するための方法。

【請求項 6】

前記音声参照信号にフィルタリング操作を行うステップを含み、

前記音声信号の前記出力バージョンは、前記音声参照信号から前記フィルタ操作された音声参照信号および前記少なくとも一つのノイズ参照信号を差し引いて提供される、請求項 1 から 5 のいずれか 1 項に記載のノイズを低減するための方法。

【請求項 7】

20

一つまたは複数の前記音声漏れ寄与および前記所望の信号を考慮するよう前記フィルタ係数を適合させるステップを含む、請求項 1 から 6 のいずれか 1 項に記載のノイズを低減するための方法。

【請求項 8】

発話を増大する応用における、請求項 1 から 7 のいずれか 1 項に記載のノイズを低減するための方法の使用。

【請求項 9】

2 つのバージョンの音声信号を受け取り、音声参照信号および少なくとも一つのノイズ参照信号を出力するよう構成された第 1 のフィルタを備え、

前記音声参照信号は、所望の信号およびノイズ寄与を含み、前記少なくとも一つのノイズ参照信号は、音声漏れ寄与およびノイズ寄与を含み、

30

前記ノイズ参照信号をフィルタするよう構成された第 2 のフィルタと、

前記音声参照信号から、前記少なくとも一つのフィルタされたノイズ参照信号を差し引き、ノイズが低減された前記音声信号の出力バージョンを提供するよう構成された加算器と、を備え、

前記第 2 のフィルタは、前記音声漏れ寄与のエネルギーおよび前記音声信号の前記出力バージョンにおける前記ノイズ寄与のエネルギーの加重和を最小にするよう構成されたフィルタ係数をもつ、音声信号におけるノイズを減少するための信号プロセッサ。

【請求項 10】

前記第 1 のフィルタが、ビームフォーマー・フィルタおよびブロッキング・マトリックス・フィルタを備える、空間プリプロセッサ・フィルタである、請求項 9 に記載の信号プロセッサ。

40

【請求項 11】

前記ビームフォーマー・フィルタが遅延 / 合計ビームフォーマーである、請求項 10 に記載の信号プロセッサ。

【請求項 12】

補聴器で使用するための請求項 9 に記載の信号プロセッサ。

【請求項 13】

前記第 2 のフィルタが前記音声参照信号をフィルタするよう構成され、前記加算器が前記音声参照信号から前記フィルタされた音声参照信号および前記少なくとも一つのフィル

50

タされたノイズ参照信号を差し引いて前記音声信号の前記出力バージョンを提供する、請求項 9 に記載の信号プロセッサ。

【請求項 14】

前記フィルタ係数を一つまたは複数の前記音声漏れ寄与および前記所望の信号を考慮するよう適合させる、請求項 9 に記載の信号プロセッサ。

【請求項 15】

少なくとも 2 つのバージョンの音声信号をフィルタするフィルタ手段を備え、  
該フィルタ手段は、所望の信号およびノイズ寄与を含む音声参照信号、および音声漏れ寄与およびノイズ寄与を含む少なくとも一つのノイズ参照信号を出力するよう構成されており、

前記少なくとも一つのノイズ参照信号をフィルタするためのフィルタ手段と、

前記音声参照信号から、前記少なくとも一つのフィルタされたノイズ参照信号を差し引き、ノイズが低減された前記音声信号の出力バージョンを提供するための手段と、を備え、

前記少なくとも一つのノイズ参照信号をフィルタするフィルタ手段は、前記音声漏れ寄与のエネルギーおよび前記音声信号の前記出力バージョンにおける前記ノイズ寄与のエネルギーの加重和を最小にするよう構成されている、音声信号におけるノイズを減少するよう構成された信号プロセッサ。

【請求項 16】

少なくとも 2 つのバージョンの音声信号をフィルタするフィルタ手段が、ビームフォーマー・フィルタおよびブロッキング・マトリックス・フィルタを備える、空間プリプロセッサ・フィルタである、請求項 15 に記載の信号プロセッサ。

【請求項 17】

前記音声参照信号が前記ビームフォーマー・フィルタにより出力され、前記少なくとも一つのノイズ参照信号が前記ブロッキング・マトリックス・フィルタにより出力される、請求項 16 に記載の信号プロセッサ。

【請求項 18】

前記音声参照信号からフィルタされた少なくとも一つのノイズ参照信号を差し引く前に前記音声参照信号を遅延させる手段を有する、請求項 15 に記載のプロセッサ。

【請求項 19】

前記音声参照信号をフィルタリングする手段と、  
フィルタリングされた前記音声参照信号および前記少なくとも一つのノイズ参照信号を前記音声参照信号から差し引き、前記音声信号の出力バージョンを提供する手段と、  
を備える、請求項 15 に記載の信号プロセッサ。

【請求項 20】

一つまたは複数の前記音声漏れ寄与および前記所望の信号を考慮するよう前記ノイズ参照信号のフィルタリングを適合させる手段を有する、請求項 15 に記載の信号プロセッサ。

【請求項 21】

請求項 9 から 20 のいずれかに記載の信号プロセッサを備えた聴取装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、発話コミュニケーションの応用において、ノイズを適応的に低減するための方法および装置に関するものである。

先行技術の開示

【0002】

テレビ電話会議、ハンズフリーの電話および補聴器のような、発話コミュニケーションの応用において、バックグラウンドノイズの存在は、望ましい発話信号の了解度を著しく低減する。そのため、ノイズリダクションアルゴリズムの使用が必要である。マルチマイク

10

20

30

40

50

ロフォンシステムは、望ましい信号およびノイズ信号の一時的でスペクトルの情報に加えて、空間的な情報を使用しており、そのため、単一のマイクロフォン手段より好ましい。美観に関する理由のため、例えば補聴器の応用に対するマルチマイクロフォンの技術は、小さいサイズのアレイの使用と関係する。かなりのノイズ低減をそのようなアレイで達成することができるが、増加する感度を犠牲にして、マイクロフォンの mismatch、残響などの仮定した信号モデル中でのエラーとなる（例えば、Stadler & Rabinowitz, 'On the potential of fixed arrays for hearing aids', J. Acoust. Soc. Amer., vol. 94, no. 3, pp. 1332-1342, Sep 1993を参照のこと）。補聴器において、複数のマイクロフォンは利得や位相においてほとんど一致しない。マイクロフォン特性間の利得および位相の相違は、それぞれ、6 dB および  $10^\circ$  にのぼる。

10

#### 【0003】

広く研究された複数チャンネル適合ノイズリダクションアルゴリズムは、一般化されたサイドローブ・キャンセラ (GSC) である（例えば、Griffiths & Jim, 'An alternative approach to linearly constrained adaptive beamforming', IEEE Trans. Antennas Propag., vol. 30, no. 1, pp. 27-34, Jan. 1982 および US-5473701 'Adaptive microphone array' を参照のこと）。GSC は、固定ビームフォーマーおよびブロッキングマトリックスを含む、固定された空間的なプリプロセッサと、適合ノイズキャンセラ (ANC) に基づく適合ステージと、から構成される。ANC は出力ノイズパワーを最少化するのに対し、ブロッキングマトリックスはノイズ参照信号中への発話のリークを避ける。一般的な GSC は、望ましいスピーカーの位置、知られているマイクロフォンの特性および位置、および、存在しない発話信号の反射、を仮定する。これらの仮定が満たされていれば、最も少ない残留ノイズで歪みの無い強調された発話信号を得ることができる。しかしながら、実際にはこれらの仮定はしばしば乱され、その結果、いわゆる発話リークとなり、そのため、発話の歪みとなる。発話歪みを制限するために、ANC は通常ノイズの期間のみに適合される。例えば、補聴器の応用において、小さいサイズのアレイと組み合わせて使用されるときは、付加的なローバスト性の制約 (Cox et al., 'Robust adaptive beamforming', IEEE Trans. Acoust. Speech and Signal Processing, vol. 35, no. 10, pp. 1365-1376, Oct. 1987 を参照のこと) が、マイクロフォンの mismatch のような仮定された信号モデルの小さなエラーの存在下でのパフォーマンスを保証するために要求される。広く適用される方法は、二次元の不等式の制約を ANC (QIC-GSC) に課すことから

20

30

#### 【0004】

受信したマイクロフォン信号の1つにおける望ましい信号部分のミニマム・ミーン・スクウェア・エラー (MMSE) を与える、複数チャンネル・ウィナー・フィルタリング (MWF) の技術が提案されている (Doclo & Moonen, 'GSVD-based optimal filtering for single and multimicrophone speech enhancement', IEEE Trans. Signal Processing, vol. 50, no. 9, pp. 2230-2244, Sep. 2002 を参照のこと)。GSC の ANC とは対照的に、MWF は、それ自身の最適化の基準に発話の歪みを考慮することができ、その結果、発話歪みで重み付けした複数チャンネル・ウィナー・フィルタリング (SDW-MWF) を得ることができる。(SDW-)MWF のテクニックは、珍しくも、記録された発話信号およびノイズ信号の2次の統計量に基づくものである。ローバスト性の発話検出が、そのため、再び必要となる。GSC とは対照的に、(SDW-)MWF は、信号モデルについていかなる推測的な仮定をも行わず、その結果、小さいサイズのアレイと組み合わせて使用されるときのパフォーマンスを保証するために、厳しいローバスト性の制約を全くあるいはほとんど必要としない。特に、複数のノイズ源や拡散したノイズのような複雑なノイズシナリオにおいて、GSC がローバスト性の制約を補完した場合でさえも、(SDW-)MWF は GSC を凌いでいる。

40

50

## 【 0 0 0 5 】

( S D W - ) M W F の可能性のある実施は、入力データマトリックスおよびノイズデータマトリックスの一般化されたシンギュラー・バリュー・デコンポジション ( G S V D ) に基づいている。Q R デコンポジション ( Q R D ) に基づく割安な代替品が、Rombouts & Moonen, 'QRD-based unconstrained optimal filtering for acoustic noise reduction', Signal Processing, vol. 83, no. 9, pp. 1889-1904, Sep. 2003において提案されている。また、サブバンドでの実施が、フルバンドのアプローチと比較して、かなり低コストで改良された了解度の結果となる。しかしながら、G S C および Q I C - G S C と対照的に、( S D W - ) M W F の割安な確率グラジエントに基づく実施は、いまだ利用できない。Nordholm et al., 'Adaptive microphone array employing calibration signals: an analytical evaluation', IEEE Trans. Speech, Audio Processing, vol. 7, no. 3, pp. 241-252, May 1999において、M W F に対する L M S に基づくアルゴリズムが明らかになった。しかしながら、そのアルゴリズムはキャリブレーション信号の記録を必要とする。室内の音響、マイクロフォン特性、および、望ましいスピーカ位置は時間が経てば変化するため、頻繁な再キャリブレーションが必要となり、このアプローチを面倒で高価なものとする。また、キャリブレーション信号の必要を避ける S D W - M W F に基づく L M S が、提案されている ( Florencio & Malvar, 'Multichannel filtering for optimum noise reduction in microphone arrays', Int. Conf. on Acoust., Speech, and signal Proc., Salt lake City, USA, pp.197-200, May 2001を参照のこと )。このアルゴリズムは、しかしながら、必ずしも満足させることのできない、いくつかの独立した仮定を用いており、結果として低グレードのパフォーマンスとなる。

10

20

## 【 0 0 0 6 】

G S C および M W F の技術をこれからより詳細に説明する。

一般化されたサイドローブ・キャンセラー ( G S C )

## 【 0 0 0 7 】

図 1 は、固定された空間プリ・プロセッサ、すなわち、固定ビームフォーマー  $A(z)$  とブロッキング・マトリックス  $B(z)$ 、および、A N C から構成される一般化されたサイドローブ・キャンセラー G S C の概念を記載する。望ましい発話の寄与  $u_i^s[k]$  およびノイズ寄与  $u_i^n[k]$  を有するマイクロフォン信号  $M$  が与えられると、

30

$$u_i[k] = u_i^s[k] + u_i^n[k], \quad i = 1, \dots, M \quad (\text{式 1})$$

ビームを望ましい信号の方向へ導くことにより、固定ビームフォーマー  $A(z)$  (例えば遅延および合計) は、発話寄与  $y_0^s[k]$  およびノイズ寄与  $y_0^n[k]$  を含む、いわゆる発話参照

$$y_0[k] = y_0^s[k] + y_0^n[k], \quad (\text{式 2})$$

40

を作成する。ブロッキング・マトリックス  $B(z)$  は、ノイズ寄与  $y_i^n[k]$  が発話リーク寄与  $y_i^s[k]$  と比較して支配的となるように望ましい信号ソースの方向をゼロに導くことにより、いわゆるノイズ参照  $M - 1$

$$y_i[k] = y_i^s[k] + y_i^n[k], \quad i=1, \dots, M-1 \quad (\text{式 3})$$

を作成する。のちほど、サブスクリプトの  $s$  および  $n$  は、発話および信号のノイズ寄与を参照するために使用される。発話 + ノイズの期間において、ノイズ参照  $y_i[k]$ ,  $i=0 \dots M-1$  は発話 + ノイズを含む。ノイズのみの期間において、ノイズ参照はノイズ成分のみから

50

なり、すなわち、 $y_i[k] = y_i^n[k]$ である。ノイズ信号の二次統計は、それらがノイズのみの期間において推定できるように極めて固定されているとみなされる。

【0008】

固定された空間プリ・プロセッサを設計するために、マイクロフォン特性について、スピーカーの位置およびマイクロフォンの位置さらには反響が無いものと仮定する。これらの仮定が満たされれば、ノイズ参照は、いかなる発話をも含まず、すなわち、 $y_i^s[k]=0$ ,  $i=1, \dots, M-1$ である。しかしながら、実際には、これらの仮定はしばしば破られ(例えばマイクロフォンの mismatch や反響)、発話がノイズ参照に漏れる。そのような発話リークの影響を限定するために、フィルタ長  $L$  を有する ANC フィルタ

$$\mathbf{w}_{1:M-1} \in C^{(M-1)L \times 1}$$

10

$$\mathbf{w}_{1:M-1}^H = [\mathbf{w}_1^H \quad \mathbf{w}_2^H \quad \dots \quad \mathbf{w}_{M-1}^H] \quad (式4)$$

ここで、

$$\mathbf{w}_i = [w_i[0] \quad w_i[1] \quad \dots \quad w_i[L-1]]^T, \quad (式5)$$

が、ノイズのみの期間において適合される。(時間が支配要素となる実施において、適合フィルタ  $W_{1:M-1}$  およびフィルタ  $W_{1:M-1}$  の入力信号は実数となる。のちほど、式が複素数の入力信号に一般化され、それらはサブバンドの実施にもまた適用することができる。) そのため、ANC フィルタ  $W_{1:M-1}$  は出力ノイズパワーを最少化する、例えば、

20

$$\mathbf{w}_{1:M-1} = \arg \min_{\mathbf{w}_{1:M-1}} E \{ |y_0^n[k-\Delta] - \mathbf{w}_{1:M-1}^H[k] \mathbf{y}_{1:M-1}^n[k]|^2 \} \quad (式6)$$

続いて

$$\mathbf{w}_{1:M-1} = E \{ \mathbf{y}_{1:M-1}^n[k] \mathbf{y}_{1:M-1}^{n,H}[k] \}^{-1} E \{ \mathbf{y}_{1:M-1}^n[k] y_0^{n,*}[k-\Delta] \}, \quad (式7)$$

ここで

$$\mathbf{y}_{1:M-1}^{n,H}[k] = [\mathbf{y}_1^{n,H}[k] \quad \mathbf{y}_2^{n,H}[k] \quad \dots \quad \mathbf{y}_{M-1}^{n,H}[k]] \quad (式8)$$

$$\mathbf{y}_i^n[k] = [y_i^n[k] \quad y_i^n[k-1] \quad \dots \quad y_i^n[k-L+1]]^T \quad (式9)$$

30

であり、 $\Delta$  は発話参照がフィルタ  $W_{1:M-1}$  における一般的でない盗聴を許すよう発話参照に適用される遅延である。遅延  $\Delta$  は通常  $[L/2]$  に設定されており、ここで、 $[x]$  は  $x$  と同じまたはそれ以上の最も小さい整数を示す。 $W_{1:M-1}$  および  $\mathbf{Y}_{1:M-1}$  中のサブスクリプト  $1:M-1$  は、適合フィルタおよび入力ベクトルの最初と最後のチャンネル成分のサブスクリプトをそれぞれ参照する。

【0009】

理想的な状態 ( $y_i^s[k]=0$ ,  $i=1, \dots, M-1$ ) において、GSC は、残存ノイズを最小化する一方望ましい発話信号を歪めることがない、すなわち、 $Z^s[k]=y_0^s[k-\Delta]$  となる。しかしながら、小さいサイズのアレイと組み合わせて使用されると、仮定された信号モデル中の小さなエラー(結果として  $y_i^s[k] \neq 0$ ,  $i=1, \dots, M-1$  となる)は、ノイズのみの領域において適用した時でさえも、かなり歪んだ出力発話信号  $z^s[k]$  を作成するのにすでに十分となり、 $W_{1:M-1}$  におけるローバスト性の制約が要求される。

40

$$z^s[k] = y_0^s[k-\Delta] - \mathbf{w}_{1:M-1}^H \mathbf{y}_{1:M-1}^s[k], \quad (式10)$$

また、固定ビームフォーマー  $A(Z)$  は、全ての可能性のあるモデルエラーに対し、発話参照  $y_0^s[k]$  中の歪みが最小となるよう設計されるべきである。結局、遅延/合計ビームフォーマーが使用される。小さいサイズのアレイにおいて、このビームフォーマーは信号モデルエラーに対する十分なローバスト性を提供し、それはノイズ感度を最小化する。ノイズ感度は、望ましい信号の利得に対する空間的なホワイトノイズの利得の比として定義され

50

、しばしば、仮定された信号モデル中のエラーに対するアルゴリズムの感度を定量化するために使用される。実際に起きる信号モデルのエラーについて統計的な知識が与えられると、固定ビームフォーマーおよびブロック・マトリックスをさらに最適化することができる。

【 0 0 1 0 】

G S Cのローバスト性を増すための一般的なアプローチは、G S Cの最適化の基準(式6)を以下のように変形することで、二次元の不等式の制約(Q I C)をA N Cフィルタ $W_{1:M-1}$ に適用することである:

$$w_{1:M-1} = \arg \min_{w_{1:M-1}} E \{ |y_0^n[k-\Delta] - w_{1:M-1}^H y_{1:M-1}^n[k]|^2 \} \quad (式 11)$$

条件として  $w_{1:M-1}^H w_{1:M-1} \leq \beta^2$ .

Q I Cはフィルタの係数 $W_{1:M-1}$ の過度の成長を阻害する。そのため、発話がノイズ参照中へ漏れた場合に望ましくない発話歪みを減少させる。Q I C - G S Cは、それぞれのアップデートステップで、 $w_{1:M-1}^H w_{1:M-1}$ が $\beta^2$ を超えた時に、二次元の制限を

$$\frac{\beta}{\|w_{1:M-1}\|}$$

でフィルタの係数をスケールすることで新たに得られたA N Cフィルタに適用する: 適合スケール・プロジェクション・アルゴリズム(S P A)を使用することで、実行することができる。最近、Tian et al.は、バリエブル・ローディングを使用して二次元の制限を実行した('Recursive quadratic implementation for LCMP Beamforming under quadratic constraint', IEEE Trans. Signal Processing, vol. 49, no. 6, pp. 1138-1145, June 2001)。帰納的な最小2乗法(R L S)に対し、この技術は、スケールド・プロジェクション・アルゴリズムよりも最適解(式11)に対しより良い近似を与える。

#### 複数チャンネル・ウィナー・フィルタリング(M W F)

【 0 0 1 1 】

複数チャンネル・ウィナー・フィルタリング(M W F)の技術は、受信したマイクロフォン信号の1つにおける望ましい信号部の最小平均2乗誤差(M M S E)の推定量を与える。G S Cとは異なり、このフィルタリング技術は、信号モデルについて何らの推測的な仮定を作らずによりローバスト性を有している。特に、複数のノイズ発生源や拡散ノイズのような複雑なノイズのシナリオにおいて、M W Fは、G S Cがローバスト性の制約を受けている時でさえも、G S Cより性能が優れている。

【 0 0 1 2 】

M W F

$$\bar{w}_{1:M} \in C^{M \times 1}$$

は、i番目(例えば1番目)のマイクロフォンでの(未知の)発話信号 $u_i^s[k-\Delta]$ の遅延させたバージョンとMフィルタリング処理したマイクロフォン信号の合計

$$\bar{w}_{1:M}^H \mathbf{u}_{1:M}[k] \quad (40)$$

との間の平均2乗誤差(M S E)を最小化し、すなわち、以下のように表すことができる:

$$\bar{w}_{1:M} = \arg \min_{\bar{w}_{1:M}} E \{ |u_i^s[k-\Delta] - \bar{w}_{1:M}^H \mathbf{u}_{1:M}[k]|^2 \}, \quad (式 12)$$

この式から以下の式となる;

$$\bar{w}_{1:M} = E \{ \mathbf{u}_{1:M}[k] \mathbf{u}_{1:M}^H[k] \}^{-1} E \{ \mathbf{u}_{1:M}[k] u_i^{s,*}[k-\Delta] \}, \quad (式 13)$$

ここで、

$$\bar{\mathbf{w}}_{1:M}^H = [\bar{\mathbf{w}}_1^H \quad \bar{\mathbf{w}}_2^H \quad \text{L} \quad \bar{\mathbf{w}}_M^H], \quad (\text{式 14})$$

$$\mathbf{u}_{1:M}^H[k] = [\mathbf{u}_1^H[k] \quad \mathbf{u}_2^H[k] \quad \text{L} \quad \mathbf{u}_M^H[k]], \quad (\text{式 15})$$

$$\mathbf{u}_i[k] = [u_i[k] \quad u_i[k-1] \quad \text{L} \quad u_i[k-L+1]]^T. \quad (\text{式 16})$$

であり、また、 $u_i[k]$ は発話成分とノイズ成分とを備える。

【0013】

同様のアプローチでは、 $i$ 番目のマイクロフォンでの(未知の)ノイズ信号 $u_i^n[k-]$ の遅延させたバージョンを推定することから構成され、その結果、以下の式となる：

10

$$\mathbf{w}_{1:M} = \arg \min_{\mathbf{w}_{1:M}} E \left\{ \left| u_i^n[k-\Delta] - \mathbf{w}_{1:M}^H \mathbf{u}_{1:M}[k] \right|^2 \right\}, \quad (\text{式 17})$$

および

$$\mathbf{w}_{1:M} = E \{ \mathbf{u}_{1:M}[k] \mathbf{u}_{1:M}^H[k] \}^{-1} E \{ \mathbf{u}_{1:M}[k] u_i^{n,*}[k-\Delta] \}, \quad (\text{式 18})$$

ここで

$$\mathbf{w}_{1:M}^H = [\mathbf{w}_1^H \quad \mathbf{w}_2^H \quad \text{L} \quad \mathbf{w}_M^H]. \quad (\text{式 19})$$

である。発話成分 $u_i^s[k-]$ の推定値 $z[k]$ は、遅延させた $i$ 番目のマイクロフォン信号 $u_i[k-]$ から $u_i^n[k-]$ の推定値 $\mathbf{w}_{1:M}^H \mathbf{u}_{1:M}[k]$ を引くことによって以下のように得られる：

20

すなわち、

$$z[k] = u_i[k-\Delta] - \mathbf{w}_{1:M}^H \mathbf{u}_{1:M}[k]. \quad (\text{式 20})$$

これは図2において $u_i^n[k-] = u_i^n[k-]$ として記載される。

【0014】

MWFの残余誤差エネルギーは以下の式と同じであり、

$$E \{ |e[k]|^2 \} = E \left\{ \left| u_i^s[k-\Delta] - \bar{\mathbf{w}}_{1:M}^H \mathbf{u}_{1:M}[k] \right|^2 \right\}, \quad (\text{式 21})$$

30

以下の式に分解される：

$$E \left\{ \left| u_i^s[k-\Delta] - \bar{\mathbf{w}}_{1:M}^H \mathbf{u}_{1:M}[k] \right|^2 \right\} + E \left\{ \left| \bar{\mathbf{w}}_{1:M}^H \mathbf{u}_{1:M}[k] \right|^2 \right\} \quad (\text{式 22})$$

ここで、 $e_d^2$ は発話歪みエネルギーと同じであり、 $e_n^2$ は残余ノイズエネルギーと同じである。MWFの設計基準は、重み付け因子 $\mu$ を $\mu \in [0, 1]$ で内挿することによって、以下の式のように、発話歪みとノイズ・リダクションとの間のトレードオフを許すように一般化されることができる：

40

$$\bar{\mathbf{w}}_{1:M} = \arg \min_{\bar{\mathbf{w}}_{1:M}} E \left\{ \left| u_i^s[k-\Delta] - \bar{\mathbf{w}}_{1:M}^H \mathbf{u}_{1:M}[k] \right|^2 \right\} + \mu E \left\{ \left| \bar{\mathbf{w}}_{1:M}^H \mathbf{u}_{1:M}[k] \right|^2 \right\}. \quad (\text{式 23})$$

(式23)の解は以下の式によって与えられる：

$$\bar{\mathbf{w}}_{1:M} = E \left\{ \mathbf{u}_{1:M}^s[k] \mathbf{u}_{1:M}^{s,H}[k] + \mu \mathbf{u}_{1:M}^n[k] \mathbf{u}_{1:M}^{n,H}[k] \right\}^{-1} E \left\{ \mathbf{u}_{1:M}^s[k] u_i^{s,*}[k-\Delta] \right\}. \quad (\text{式 24})$$

【0015】

同様に、(式17)中における $\mathbf{w}_{1:M-1}$ に対する最適化の基準は、以下の式に変形でき、



$$\mathbf{w}_{1:M} = \arg \min_{\mathbf{w}_{1:M}} E \left\{ \left\| \mathbf{w}_{1:M}^H \mathbf{u}_{1:M}^s[k] \right\|^2 + \mu E \left\{ \left| u_i^n[k-\Delta] - \mathbf{w}_{1:M}^H \mathbf{u}_{1:M}^n[k] \right|^2 \right\} \right\}, \quad (\text{式 } 25)$$

結果として、以下の式が得られる：

$$\mathbf{w}_{1:M} = E \left\{ \mathbf{u}_{1:M}^n[k] \mathbf{u}_{1:M}^{n,H}[k] + \frac{1}{\mu} \mathbf{u}_{1:M}^s[k] \mathbf{u}_{1:M}^{s,H}[k] \right\}^{-1} E \left\{ \mathbf{u}_{1:M}^n[k] u_i^{n,*}[k-\Delta] \right\}. \quad (\text{式 } 26)$$

結局、(式26)は発話歪みで重み付けした複数チャンネル・ウィナー・フィルタ(SDW-MWF)として参照される。要素 $\mu$  [0, ]は発話歪みとノイズ・リダクションとをトレードオフする。 $\mu = 1$ のとき、MMSEの基準となる(式12)または(式17)が得られる。 $\mu > 1$ のとき、残存ノイズのレベルは、増加した発話歪みにより減少する。 $\mu$ を にセットすることで、ノイズ・リダクションが強調され、発話歪みは完全に無視される。これに対し $\mu$ を0にセットすることで、ノイズ・リダクションが全く行われない結果となる。

【0016】

実際、相関マトリックス

$$E \left\{ \mathbf{u}_{1:M}^s[k] \mathbf{u}_{1:M}^{s,H}[k] \right\}$$

は知られていない。発話期間の間、入力である $u_i[k]$ は、発話 + ノイズすなわち

$$u_i[k] = u_i^s[k] + u_i^n[k], i=1, \dots, M. \quad 20$$

から構成される。ノイズ期間間、ノイズ成分である $u_i^n[k]$ のみが観察される。発話信号とノイズ信号とが相関しないと仮定すると、

$$E \left\{ \mathbf{u}_{1:M}^s[k] \mathbf{u}_{1:M}^{s,H}[k] \right\}$$

を以下の式のように推定できる：

$$E \left\{ \mathbf{u}_{1:M}^s[k] \mathbf{u}_{1:M}^{s,H}[k] \right\} = E \left\{ \mathbf{u}_{1:M}[k] \mathbf{u}_{1:M}^H[k] \right\} - E \left\{ \mathbf{u}_{1:M}^n[k] \mathbf{u}_{1:M}^{n,H}[k] \right\}, \quad (\text{式 } 27)$$

ここで、二次統計量

$$E \left\{ \mathbf{u}_{1:M}[k] \mathbf{u}_{1:M}^H[k] \right\} \quad 30$$

は発話 + ノイズの間評価され、二次統計量

$$E \left\{ \mathbf{u}_{1:M}^n[k] \mathbf{u}_{1:M}^{n,H}[k] \right\}$$

はノイズの期間の間でのみ評価される。GSCでは、そのためローバスタ性の発話検知が必要となる。(式27)、(式24)および(式26)を使用して、以下のように書き直すことができる：

$$\begin{aligned} \bar{\mathbf{w}}_{1:M} = & \left( E \left\{ \mathbf{u}_{1:M}[k] \mathbf{u}_{1:M}^H[k] \right\} + (\mu - 1) E \left\{ \mathbf{u}_{1:M}^n[k] \mathbf{u}_{1:M}^{n,H}[k] \right\} \right)^{-1} \\ & \times \left( E \left\{ \mathbf{u}_{1:M}[k] u_i^{n,*}[k-\Delta] \right\} - E \left\{ \mathbf{u}_{1:M}^n[k] u_i^{n,*}[k-\Delta] \right\} \right) \end{aligned} \quad (\text{式 } 28) \quad 40$$

および

$$\mathbf{w}_{1:M} = \left( \frac{1}{\mu} E \left\{ \mathbf{u}_{1:M}[k] \mathbf{u}_{1:M}^H[k] \right\} + \left( 1 - \frac{1}{\mu} \right) E \left\{ \mathbf{u}_{1:M}^n[k] \mathbf{u}_{1:M}^{n,H}[k] \right\} \right)^{-1} E \left\{ \mathbf{u}_{1:M}^n[k] u_i^{n,*}[k-\Delta] \right\}. \quad (\text{式 } 29)$$

ウィナー・フィルタは、発話 + ノイズおよびノイズデータマトリックスの一般化されたシンギュラー・バリュー・デコンポジション(GSVD)によって、各時間点kで計算することができる。QRデコンポジションに基づくより安価な再帰的な手段を用いることもできる。また、サブバンドでの実行は、結果としての発話理解度を増加して複雑さを減少し、補聴器の応用としてそれを好適とする。

発明の目的

【0017】

本発明は、発話を強調する応用において、ノイズ特にバックグラウンド・ノイズを適応させて減少し、それにより、先行技術の解法における問題および課題を解消する、方法および装置を提供することを目的とする。

発明の要約

【0018】

本発明は、  
ノイズを有する少なくとも2種類（バージョン）の発話信号（音声信号）を第1のフィルタに適用する工程であって、第1のフィルタが、発話参照信号と少なくとも1つのノイズ参照信号とを出力する工程と、

少なくとも1つのノイズ参照信号の各々にフィルタリング操作を適用する工程と、  
発話参照信号からフィルタリング処理したノイズ参照信号の各々を差し引く工程と、を備える方法であって、

フィルタリング操作が、少なくとも1つのノイズ参照信号中の発話リークの寄与を考慮して決定されるフィルタ係数を有するフィルタを用いて実行されることを特徴とするノイズを有する発話信号のノイズを低減するための方法に関するものである。

【0019】

典型的な実施例において、ノイズを有する少なくとも2種類（バージョン）の発話信号は、ノイズを有する発話信号を取り出す少なくとも2種類のマイクロフォンからの信号である。

【0020】

好ましくは、第1のフィルタは、ビームフォーマー・フィルタおよびブロッキング・マトリックス・フィルタを備える空間プリプロセッサである。

【0021】

好ましい実施例において、発話参照信号はビームフォーマー・フィルタによって出力され、少なくとも1つのノイズ参照信号はブロッキング・マトリックス・フィルタによって出力される。

【0022】

好ましい実施例において、発話参照信号は、差し引き工程を実施する前に遅延される。

【0023】

好適には、フィルタリング操作は発話参照信号に付加的に適用され、ここで、フィルタリングされた発話参照信号は、また、発話参照信号から差し引かれる。

【0024】

他の好ましい実施例において、方法はさらにフィルタの係数を定期的に適合させる工程を備える。それにより、少なくとも1つのノイズ参照信号における発話リークの寄与が考慮され、あるいは、少なくとも1つのノイズ参照信号中の発話リークの寄与および発話参照信号中の発話の寄与の両者が考慮される。

【0025】

本発明は、また、発話を増加させる応用において前述のようにノイズを減少するための方法の使用に関する。

【0026】

第2の目的では、本発明は、

- ・少なくとも2つの入力を有し、発話参照信号および少なくとも1つのノイズ参照信号を出力するために設けられた第1のフィルタと、

- ・そこへ発話参照信号を適用するフィルタおよびそこへ少なくとも1つのノイズ参照信号を適用するフィルタと、

- ・発話参照信号から、フィルタリングされた発話参照信号およびフィルタリングされたノイズ参照信号の各々を差し引くための加算手段と、を備え、

ノイズを有する発話信号におけるノイズを減少するための信号処理回路に関する。

【0027】

好適には、第1のフィルタは、ビームフォーマー・フィルタおよびブロッキング・マトリックス・フィルタを備える空間プリプロセッサである。

【0028】

他の実施例において、ビームフォーマー・フィルタは遅延/合計ビームフォーマーである。

【0029】

本発明は、また、上述した信号処理回路を備える聴取のための装置に関する。聴取のための装置とは、補聴器（外部に装着するものまたは移植するもののいずれか）または人工耳を意味する。

10

【図面の簡単な説明】

【0030】

【図1】図1は一般化されたサイドローブ・キャンセラの概念を示す。

【0031】

【図2】図2は複数チャンネル・ウィナー・フィルタリングの等価なアプローチを示す。

【0032】

【図3】図3は空間的に前処理されたSDW-MWFを示す。

【0033】

【図4】図4は、複数チャンネル・フィルタ $w_d$ およびシングルチャンネル・フィルタ $e_{1-w_0}$ での $w_0$ を有するSP-SDW-MWFの分解を示す。

20

【0034】

【図5】図5は試験のためのセットアップを示す。

【0035】

【図6】図6は、第2マイクロフォンでの異なる利得ミスマッチ $\mu_2$ に対するSDR-GSCの実行における $1/\mu$ の影響を示す。

【0036】

【図7】図7は、第2マイクロフォンでの異なる利得ミスマッチ $\mu_2$ に対する $w_0$ を有するSP-SDW-MWFの実行における $1/\mu$ の影響を示す。

【0037】

30

【図8】図8は、第2マイクロフォンでの異なる利得ミスマッチ $\mu_2$ に対する $\mu_2$ の関数としてのQIC-GSCに対する $SNR_{intellig}$ および $SD_{intellig}$ を示す。

【0038】

【図9】図9は、チャンネルあたりのフィルタ長Lの関数としてのLPフィルタを有するTDおよびFD確率勾配(SG)アルゴリズムの複合を示す；M=3（比較のため、標準的なNLMSANCおよびSPAの複合も示される）。

【0039】

【図10】図10は、異なるFD確率勾配(FD-SG)アルゴリズムの実行を示す；(a)は90°における固定発話類似ノイズであり；(b)は90°におけるマルチトーカー・パブル・ノイズである。

40

【0040】

【図11】図11は、 $w_0$ を備えないものと $w_0$ を備えるものに対するFD確率勾配SP-SDW-MWF( $1/\mu=0.5$ )の実行におけるLPフィルタの影響を示す。90°のパブル・ノイズ。

【0041】

【図12】図12は、 $\theta=0$ および $\theta=0.9998$ に対するFD-SGの収束挙動を示す。ノイズ源の位置は90°から180°へあるいはその逆に突然変化する。

【0042】

【図13】図13は、複数ノイズ源のシナリオにおいてLPフィルタ( $\theta=0.9998$ )を有するSP-SDW-MWFのFD確率勾配の実施能力を示す。

50

【 0 0 4 3 】

【 図 1 4 】 図 1 4 は、複数ノイズ源のシナリオにおける F D S P A の能力を示す。

【 0 0 4 4 】

【 図 1 5 】 図 1 5 は、複数ノイズ源のシナリオにおける周波数領域 S P - S D W - M W F ( アルゴリズム 2 およびアルゴリズム 4 ) の S N R により改良を示す。

【 0 0 4 5 】

【 図 1 6 】 図 1 6 は、複数ノイズ源のシナリオにおける周波数領域 S P - S D W - M W F ( アルゴリズム 2 およびアルゴリズム 4 ) の発話歪みを示す。

【 0 0 4 6 】

#### 発明の詳細な説明

本発明を、以下詳細に記載する。まず第 1 に、本発明で提案される適合複数チャンネル・リダクション技術を、空間的に前処理された発話歪みで重み付けされた複数チャンネル・ウィナー・フィルタを参照して、記載する。

【 0 0 4 7 】

本発明の第 1 実施例は、発話歪み一般化 G S C ( S D R - G S C ) として参照される。新しい設計基準が G S C の適合ステージに対し発展し： A N C 設計基準が、信号モデルエラーに基づく発話歪みを限定する一般化条件で補足される。S D R - G S C において、パラメータ  $\mu$  は、発話歪みとノイズ・リダクションとの間のトレードオフを許すように組み込まれる。ノイズ・リダクションに全ての注意をはらうと、一般的な G S C の結果となり、一方、発話歪みに全ての注意をはらうと、固定ビームフォーマーの出力の結果となる。低い S N R を有するノイズシナリオにおいて、S D R - G S C における適合性を、発話歪みに対する注意を増加されることによって、言い換えると、パラメータを  $\mu$  から 0 に減少させることによって、簡単に減少あるいは排除することができる。S D R - G S C は、マイクロフォンの mismatch、残響音のような信号モデルエラーに対し G S C の感度を減少させる Q I C - G S C に代替可能である。Q I C - G S C と比べて、S D R - G S C は、発話歪みの量が大きくなる時発話歪みに重要性を置くようになる。信号モデルエラーがなければ、G S C の能力は維持される。その結果、大きなモデルエラーに対しロバスト性を保証する一方、より良いノイズ・リダクション能力が小さいモデルエラーに対し得られる。

【 0 0 4 8 】

次のステップにおいて、S D R - G S C のノイズ・リダクション能力は、発話参照信号に特別な適合フィルタリング・オペレーション  $w_0$  を加えることによって、さらに改良される。この一般化された計画は、空間的に前処理された発話歪みで重み付けされた複数チャンネル・ウィナー・フィルタ ( S P - S D W - M W F ) として参照される。この S P - S D W - M W F は、図 3 に記載され、その特別な場合として M W F を含む。再び、パラメータ  $\mu$  が、発話歪みとノイズ・リダクションとの間のトレードオフを許すように組み込まれる。発話歪みに全ての注意をはらうと、固定ビームフォーマーの出力となる。また、ここで、適合性を、 $\mu$  から 0 へ減少させることによって、簡単に減少あるいは排除することができる。(発話の漏れがなく、無限に長いフィルタ長に対し)、S P - S D W - M W F は、発話歪みで重み付けされた複数チャンネル・ウィナー・フィルタ ( S D W - S W F ) を備える S D R - G S C のカスケード表示に対応する。発話漏れが存在する場合、 $w_0$  を有する S P - S D W - M W F はその能力を維持しようと試み：S P - S D W - M W F は、その後、発話漏れに基づく能力の低下を補償する特別なフィルタリング動作を含む。そのため、S D R - G S C ( およびまた G S C ) と比較して、能力はマイクロフォン・mismatch により低下する。G S V D または Q R 分離に基づく、( S D W - ) M W F の繰り返し実施が存在する。また、サブバンド実施、フルバンド・アプローチと比べてかなり低い複雑性で改良された理解度の結果となる。これらの技術は、S D R - G S C を実施することで、より一般的には、S P - S D W - M W F を実施することで、拡張される。

【 0 0 4 9 】

本発明において、S D R - G S C および S P - S D W - M W F の安価な時間領域および周

10

20

30

40

50

波数領域の確率勾配での実施が、同様に提案されている。SDR-GSCの設計基準、より一般的には、SP-SDW-MWFの設計基準から始めると、時間領域での確率勾配アルゴリズムが引き出される。収束速度を増加させて計算の複雑性を減少させるために、アルゴリズムは周波数領域で実施される。高いレベルの非定常ノイズ中で使用されたとき、そこから確率的勾配アルゴリズムが受ける大きな過度のエラーを減少するために、ローパスフィルタが、発話歪みを限定する勾配の推定の部分に適用される。ローパスフィルタは、望ましい発話要素の高いレベルでの時間-変化歪みを避ける一方、時間-変化シナリオで必要なトラッキング能力を劣化させることはない。また、実験は、提案された確率的勾配アルゴリズムはQIC-GSCを超えるSP-SDW-MWFの高価を保持する一方、その計算の複雑性はQICを実施するNLMSに基づき縮尺されたプロジェクショナル  
 アルゴリズムに匹敵する。ローパスフィルタを備える空間的に前処理された発話歪みで重み付けされた複数チャンネル・ウィナー・フィルタ(SP-SDW-MWF)確率的勾配アルゴリズムは、しかしながら、データバッファを必要とし、結果として大きなメモリーコストがかかる。メモリーコストは、(対角)相関マトリックスを使用して周波数領域で正規化項を求めることによって減少させることができ、複雑さおよびメモリーコストの両面で、市販の補聴器でSP-SDW-MWFの実行を可能とする。試験結果は、相関マトリックスを使用した確率的勾配アルゴリズムが、ローパスフィルタを備える確率的勾配アルゴリズムと同じ性能を有することを、示している。

10

空間的に前処理されたSDW複数チャンネル・ウィナー・フィルタ  
 概念

20

【0050】

図3は空間的に前処理された発話歪みで重み付けされた複数チャンネル・ウィナー・フィルタ(SP-SDW-MWF)を示す。SP-SDW-MWFは、固定された空間プリプロセッサすなわち固定ビームフォーマーA(z)およびブロックングマトリックスB(z)と、適合された発話歪み重み付け複数チャンネル・ウィナー・フィルタ(SDW-MWF)と、から構成されている。u<sub>i</sub><sup>s</sup>[k]を望ましい発話寄与とし、u<sub>i</sub><sup>n</sup>[k]をノイズ寄与としたとき、M個のマイクロフォン信号

$$u_i[k] = u_i^s[k] + u_i^n[k], i = 1, \dots, M \quad (式30)$$

30

を与えると、固定ビームフォーマーA(z)は、いわゆる発話参照

$$y_0[k] = y_0^s[k] + y_0^n[k], \quad (式31)$$

を、ビームを望ましい信号の方向に向けることで作り出し、上記発話参照は発話寄与y<sub>0</sub><sup>s</sup>[k]およびノイズ寄与y<sub>0</sub><sup>n</sup>[k]を備えている。MWFのローバスタ性の効果を維持するために、固定ビームフォーマーA(z)は、発話参照y<sub>0</sub><sup>s</sup>[k]中の歪みが、マイクロフォンミスマッチのように、仮定した信号モデルにおけるすべての可能なエラーが最小となるよう設計すべきである。その後、遅延/合計ビームフォーマーが使用される。小さいサイズのアレイに対し、このビームフォーマーは、ノイズ感度を最小にすることで、信号モデルエラーに対し十分なローバスタ性を提供する。実際に起きる信号モデルエラーについての統計的な知識を与えることで、より最適化されたフィルタ/合計ビームフォーマーA(z)を設計することができる。ブロックングマトリックスB(z)は、M-1個のいわゆるノイズ参照

40

$$y_i[k] = y_i^s[k] + y_i^n[k], i = 1, \dots, M-1 \quad (式32)$$

を、ノイズ寄与y<sub>i</sub><sup>n</sup>[k]が発話漏れ寄与y<sub>i</sub><sup>s</sup>[k]と比べて優性となるように、ゼロを興味のある方向に向けることによって、作り出す。ノイズ参照を作り出すシンプルな技術は、時間列のマイクロフォン信号をペアで差し引くことから構成される。さらに最適化されたノイズ参照は、例えば、興味のある方向のみに代えて興味のある方向の近くの特性の角度領域(例えば、興味のある方向近くの-20 から20 の角度領域)に対する発話漏れを最

50

小にすることによって、作り出される。また、実際に起きる信号モデルエラーについての統計的な知識を与えることで、発話漏れを考え得る信号モデルエラーのすべてに対し最小化することができる。

【0051】

以後、スーパースクリプト（上側の添え字） $s$  および  $n$  は、信号の発話およびノイズ寄与を参照するものとして使用される。発話 + ノイズの期間では、参照  $y_i[k]$ ,  $i=0, \dots, M-1$  は発話 + ノイズを含む。ノイズのみの期間では、 $y_i[k]$ ,  $i=0, \dots, M-1$  はノイズ要素のみから構成され、すなわち、 $y_i[k]=y_i^n[k]$  である。ノイズ信号の2次統計量は非常に静止した状態と仮定でき、それらはノイズのみの期間で推定できる。

【0052】

$$\mathbf{w}_{0:M-1} = \left( \frac{1}{\mu} E\{y_{0:M-1}^s[k] y_{0:M-1}^{s,H}[k]\} + E\{y_{0:M-1}^n[k] y_{0:M-1}^{n,H}[k]\} \right)^{-1} E\{y_{0:M-1}^n[k] y_0^{n*}[k-\Delta]\}, \quad (\text{式33})$$

ここで

$$\mathbf{w}_{0:M-1}^H[k] = [\mathbf{w}_0^H[k] \quad \mathbf{w}_1^H[k] \quad \dots \quad \mathbf{w}_{M-1}^H[k]], \quad (\text{式34})$$

$$\mathbf{w}_i[k] = [w_i[0] \quad w_i[1] \quad \dots \quad w_i[L-1]]^T \quad (\text{式35})$$

$$\mathbf{y}_{0:M-1}^H[k] = [y_0^H[k] \quad y_1^H[k] \quad \dots \quad y_{M-1}^H[k]], \quad (\text{式36})$$

$$\mathbf{y}_i[k] = [y_i[k] \quad y_i[k-1] \quad \dots \quad y_i[k-L+1]]^T, \quad (\text{式37})$$

である、SDW-MWFフィルタ  $\mathbf{w}_{0:M-1}$  は、コスト関数  $J(\mathbf{w}_{0:M-1})$

$$J(\mathbf{w}_{0:M-1}) = \frac{1}{\mu} E\left\{ \sum_{i=0}^{M-1} \left\| \mathbf{w}_{0:M-1}^H y_{0:M-1}^s[k] \right\|^2 \right\} + E\left\{ \left\| y_0^n[k-\Delta] - \sum_{i=0}^{M-1} \mathbf{w}_{0:M-1}^H y_{0:M-1}^n[k] \right\|^2 \right\}. \quad (\text{式38})$$

を最小化することによって、発話参照中のノイズ寄与  $y_0^n[k-\Delta]$  の推定値  $\mathbf{w}_{0:M-1}^H y_{0:M-1}^n[k]$  を提供する。

$\mathbf{w}_{0:M-1}$  および  $y_{0:M-1}$  中のサブスクリプト（下側の添え字） $0:M-1$  は、適合フィルタおよび入力ベクトルのそれぞれの最初と最後のチャンネル構成要素のサブスクリプトを参照する。 $\sigma_d^2$  の項は発話歪みエネルギーを表し、 $\sigma_n^2$  は残留ノイズエネルギーを示す。コスト関数（式38）の  $(1/\mu) \sigma_d^2$  の項は、SP-SDW-MWFの出力での発話歪みの取り得る量を制限する。そのため、SP-SDW-MWFは、適合ステージでの設計基準において発話歪みを明示的に考慮することで、信号モデルエラーに対するロバスト性をDSCに付加する。パラメータ  $(1/\mu) \sigma_d^2$  はノイズ・リダクションおよび発話歪みをトレードオフし、大きい場合は  $1/\mu$  であり、小さい場合は取り得る発話歪みの量である。 $\mu = 0$  において、サンプルによって遅延された固定ビームフォーマー  $A(z)$  の出力が得られる。適合性は、 $\mu$  を0まで減少することで、SP-SDW-MWF中で簡単に減少または排除させることができる（例えば、-10dBの大変に低い信号/ノイズ比(SNR)を有するノイズシナリオにおいては、固定ビームフォーマーが好適である）。また、適合性はQICを  $\mathbf{w}_{0:M-1}$  に適用することによって制限することができる。

【0053】

固定ビームフォーマー  $A(z)$  およびブロッキングマトリックス  $B(z)$  が以下のようにセットされたとき、

$$A(z) = [1 \quad 0 \quad \dots \quad 0]^H \quad (\text{式39})$$

$$B(z) = \begin{bmatrix} 0 & 1 & 0 & L & 0 \\ 0 & 0 & 0 & 0 & M \\ M & 0 & 0 & 1 & 0 \\ 0 & L & 0 & 0 & 1 \end{bmatrix}^H, \quad (\text{式40})$$

受信したマイクロフォン信号  $u_i[k]$ ,  $i=1, \dots, M$  上で動作するオリジナルのSDW-MWF

10

20

30

40

50

が得られることに注意のこと。

【 0 0 5 4 】

以下に、S P - S D W - M W F の異なるパラメータのセッティングを議論する。パラメータ  $\mu$  のセッティングおよびフィルタ  $w_0$  の存在あるいは不存在により、発話歪みで規則化した G S C ( S D R - G S C ) のような中間の解法と同様に、G S C、( S D W - ) M W F が得られる。2つのケース、すなわち、フィルタ  $w_0$  が発話参照に適用されないケース ( フィルタ長さ  $L_0=0$  )、および、付加的なフィルタ  $w_0$  が使用されるケース (  $L_0 > 0$  ) を区別できる。

S D R - G S C、すなわち、 $w_0$ なしの S P - S D W - M W F

10

【 0 0 5 5 】

第 1 に、 $w_0$ なしのケース、すなわち、 $L_0=0$ のケースを考える。(式 3 3)における  $w_{1:M-1}$  に対する解は、以下の式に換算され、

$$\arg \min_{w_{1:M-1}} \frac{1}{\mu} E \left\{ \left\| w_{1:M-1}^H y_{1:M-1}^s[k] \right\|_{\varepsilon_d^2}^2 \right\} + E \left\{ \left\| y_0^n[k-\Delta] - w_{1:M-1}^H y_{1:M-1}^n[k] \right\|_{\varepsilon_n^2}^2 \right\}, \quad (\text{式 4 1})$$

以下の式となる：

$$w_{1:M-1} = \left( \frac{1}{\mu} E \left\{ y_{1:M-1}^s[k] y_{1:M-1}^{s,H}[k] \right\} + E \left\{ y_{1:M-1}^n[k] y_{1:M-1}^{n,H}[k] \right\} \right)^{-1} E \left\{ y_{1:M-1}^n[k] y_0^{n,*}[k-\Delta] \right\} \quad (\text{式 4 2})$$

20

ここで、 $\varepsilon_d^2$ は発話歪みエネルギーであり、 $\varepsilon_n^2$ は残留ノイズエネルギーである。

【 0 0 5 6 】

G S C の最適化基準 (式 6) と比べて、正規化項

$$\frac{1}{\mu} E \left\{ \left\| w_{1:M-1}^H y_{1:M-1}^s[k] \right\|^2 \right\} \quad (\text{式 4 3})$$

が付加されている。この正規化項は、発話がノイズ参照に漏れる、すなわち、 $y_i^s[k] \neq 0$ ,  $i=1, \dots, M-1$  のとき、フィルタ  $w_{1:M-1}$  によって発生する発話歪みの量を制限する。その後、 $L_0=0$ を備える S P - S D W - M W F は、そのため、発話歪みで規則化した G S C ( S D R - G S C ) として参照される。 $\mu$  が小さくなると、発話歪みの起こる量は少なくなる。 $\mu = 0$  においては、 $z[k]$  が サンプルで遅延された固定ビームフォーマー A ( z ) の出力であるように、発話歪みにすべての重点を置く。 $\mu = \infty$  において、ノイズ歪みにすべての重点を置き、発話歪みは考慮されない。これは基準 G S C に対応する。そのため、S D R - G S C は特別なケースとして G S C を包含する。

30

【 0 0 5 7 】

$1/\mu \rightarrow 0$  を有する正規化項 (式 4 3) は G S C にローバスタ性を付加するが、一方、発話漏れのない場合のノイズリダクション能力に悪影響を与える：

- ・発話漏れのない場合、すなわち、 $y_i^s[k]=0$ ,  $i=1, \dots, M-1$  の場合、正規化項はすべての  $w_{1:M-1}$  に対し 0 と等しくなり、そのため、残留ノイズエネルギー  $\varepsilon_n^2$  は効果的に最小化される。言い換えると、発話漏れのない場合、G S C の解が得られる。

40

- ・発話漏れが存在する場合、すなわち、 $y_i^s[k] \neq 0$ ,  $i=1, \dots, M-1$  の場合、適合フィルタ  $w_{1:M-1}$  に対する最適化基準 (式 4 1) において、発話歪みが明示的に考慮され、発話歪みを限定し、一方、ノイズを減少する。

交互に発話歪みを限定するために、Q I C は、しばしば、フィルタ  $w_{1:M-1}$  に義務を負わす。S D R - G S C とは異なり、Q I C は、存在する発話漏れ  $y^s[k]$  の量とは関係なく動作する。(式 1 1) 中の制限値  $\varepsilon_d^2$  は、発生するであろう最大化モデルエラーに基づき選択されなければならない。結果として、ノイズリダクション性能は、モデルエラーが全く存在しないあるいは大変少ししか存在しないときでさえ、妥協される。そのため、Q I C は、実験結果で示されるように、S D R - G S C よりもより堅実である。

50

フィルタ  $w_0$  を備える SP - SDW - MWF

【 0 0 5 8 】

SDW - MWF (式 3 3) は、その最適化基準において、発話歪みを明示的に考慮しているため、発話参照  $y_0[k]$  における付加フィルタ  $w_0$  を付加することができる。SDW - MWF (式 3 3) は以下のより一般的な最適化基準として解ける：

$$w_{0:M-1} = \arg \min_{w_{0:M-1}} E \left\{ \left[ y_0^n[k-\Delta] - \begin{bmatrix} w_0^H & w_{1:M-1}^H \end{bmatrix} \begin{bmatrix} y_0^n[k] \\ y_{1:M-1}^n[k] \end{bmatrix} \right]^2 \right\} + \frac{1}{\mu} E \left\{ \left[ \begin{bmatrix} w_0^H & w_{1:M-1}^H \end{bmatrix} \begin{bmatrix} y_0^s[k] \\ y_{1:M-1}^s[k] \end{bmatrix} \right]^2 \right\}, \quad (式 4 4)$$

ここで、

$$w_{0:M-1}^H = [w_0^H \ w_{1:M-1}^H]$$

が (式 3 3) によって与えられる。

【 0 0 5 9 】

再び、 $\mu$  は発話歪みおよびノイズリダクションをトレードオフする。 $\mu = \infty$  に対し、発話歪み  $\sigma_d^2$  は完全に無視され、結果として 0 の出力信号となる。 $\mu = 0$  に対し、出力信号がサンプルによって遅延された固定ビームフォーマーの出力に等しくなるように、発話歪みにすべての重点を置く。

また、発話漏れの存在しない場合、すなわち、 $y_i^s[k]=0, i=1, \dots, M-1$  の場合、そして、無限に長いフィルタ  $w_i, i=1, \dots, M-1$  に対し、SP - SDW - MWF ( $w_0$  を備える) が、SDR - GSC および SDW シングルチャンネル WF (SDW - SWF) のカスケード表示に対応することが、観察される。発話漏れが存在する場合、SP - SDW - MWF ( $w_0$  を備える) は、その性能を維持しようと試みる：SP - SDW - MWF は、そのため、発話漏れに基づく性能の劣化を補償する特別のフィルタリング操作を含んでいる。これは図 4 に記載されている。無限のフィルタ長に対し、SP - SDW - MWF ( $w_0$  を備える) の性能が、固定ビームフォーマー A (z) の出力で望ましい発話要素が変化せず残っている限り、マイクロフォン・ミスマッチによって影響を受けないことが照明される。

実験結果

【 0 0 6 0 】

理論的な結果が補聴器の応用に対する実験結果によって示される。第 1 に、セットアップおよび使用された性能測定器を記載する。次に、性能における SP - SDW - MWF の異なるパラメータでのセッティングの影響およびシングル・モデル・エラーに対する感度を評価する。比較は QIC - GSC に対し行われる。

【 0 0 6 1 】

図 5 は実験に対するセットアップを示す。3つの無指向性マイクロフォン (ノーレス FG - 3452) を備える 3 マイクロフォンのビハインド・ジ・エアタイプの補聴器が、部屋内のダミーヘッド上に搭載される。第 1 のマイクロフォンと第 2 のマイクロフォンとの間の空間は約 1 cm であり、第 2 のマイクロフォンと第 3 のマイクロフォンとの間の空間は約 1.5 cm である。部屋の残響時間  $T_{60dB}$  は発話で重み付けられたノイズに対し約 700 ms である。望ましい発話信号およびノイズ信号は無相関である。発話およびノイズ信号の両者とも、ヘッドの中央で 70 dB SPL のレベルを有している。望ましい発話源およびノイズ源は、ヘッドから 1 メートル離れて位置決めされており：発話源はヘッドの正面であり ( $0^\circ$ )、ノイズ源は発話現に対して角度  $\theta$  である (図 5 参照)。指向性のみに基づく平均性能についての発想を得るために、静止した発話およびノイズ信号、平均長期パワースペクトル密度が使用される。入力信号の総持続時間は 10 秒で、そのうち 5



秒はノイズのみを含み、他の5秒は発話およびノイズの両者の信号を含む。評価の目的で、発話およびノイズ信号は別々に記録される。

【0062】

マイクロフォン信号は、認識性を高めるために、処理に先立って予め白色とされ、出力は従って白色以外の色とする。実験において、マイクロフォンは、0°の位置における無エコー性の発話で重み付けられたノイズ信号の記録によってキャリブレーションされる。遅延/合計ビームフォーマーは、(小さいマイクロフォンの空間の場合)モデルエラーに対して大変ローバストであることが知られているため、固定ビームフォーマーとして使用される。ブロックングマトリクスBは、ペアで時間列に並べられたキャリブレーションされたマイクロフォン信号を差し引く。

10

【0063】

異なるパラメータセッティング(例えば $\mu$ 、 $w_0$ )の性能上の影響を調査するために、フィルタ係数が(式33)を用いて計算され、ここで、

$$E\{y_{0:M-1}^s y_{0:M-1}^{s,H}\}$$

はマイクロフォン信号のクリーンな発話寄与によって推定される。実際に、

$$E\{y_{0:M-1}^s y_{0:M-1}^{s,H}\}$$

は(式27)を使用して近似される。性能上の近似(式27)の影響は、与えられたデータセットに対し小さかった(言い換えると、認識性に重きを置いたSNRの改良では最大でも0.5dBの相違である)。QIC-GSCは種々のロード用RLSを使用して実行される。チャンネル当たりのフィルタ長Lは96と同じである。

20

【0064】

異なるアプローチの性能を評価するために、ブロードバンドの認識性で重み付けられたSNRの改良が用いられ、以下の式で定義される：

$$\Delta \text{SNR}_{\text{intellig}} = \sum_i I_i (\text{SNR}_{i,\text{out}} - \text{SNR}_{i,\text{in}}), \quad (\text{式45})$$

ここで、バンド重要性関数 $I_i$ は、認識性に対し中央周波数 $f_i^c$ を備えるi番目の3分の1オクターブバンドの従量制を表現し、 $\text{SNR}_{i,\text{out}}$ は出力SNR(in dB)であり、 $\text{SNR}_{i,\text{in}}$ はi番目の3分の1オクターブバンドにおける入力SNR(in dB)である('ANSI S3.5-1997, American National Standard Methods for Calculation of the Speech Intelligibility Index')。認識性で重み付けされたSNRは、発話歪みを考慮しないノイズ・リダクション・アルゴリズムによってどれだけ認識性が改良されたかをしめしている。

30

【0065】

発話歪みの量を測定するために、以下の認識性で重み付けられたスペクトル歪みの測定量を定義する：

$$\text{SD}_{\text{intellig}} = \sum_i I_i \text{SD}_i \quad (\text{式46})$$

測定量は、以下の式により測定された $\text{SD}_i$ 、i番目の3分の1バンドにおける平均スペクトル歪み(dB)を備えており；

40

$$\text{SD}_i = \int_{2^{-1/6} f_i^c}^{2^{1/6} f_i^c} |10 \log_{10} G^s(f)| df / \left[ (2^{1/6} - 2^{-1/6}) f_i^c \right], \quad (\text{式47})$$

ここで、 $G^s(f)$ は、ノイズリダクション・アルゴリズムの入力から出力までの発話のパワー・トランスファー関数である。空間的なプリプロセッサの影響を排除するために、性能の測定は、固定ビームフォーマーの出力に対して計算される。

【0066】

SP-SDW-MWFの性能上の $\mu$ および $w_0$ に対する異なるパラメータセッティングの影響は、5つのノイズ源シナリオで示される。5つのノイズ源は、望ましい0°のソースに対し角度75°、120°、180°、240°、285°に位置する。仮定した信号モデルにおけるエラーに対するアルゴリズムの感度を評価するために、性能上の、マイク

50

ロフォン・ミスマッチの影響、例えば、第2のマイクロフォンの利得のミスマッチの影響が評価される。異なる可能性のある信号モデルエラーの中で、マイクロフォン・ミスマッチが、補聴器の応用においてGSCの性能に大変有害であることがわかった。補聴器において、マイクロフォンは利得およびフェーズにおいてほとんどマッチしない。それぞれ6 dBおよび10°のマイクロフォン特性間の利得およびフェーズの相違が報告されている。

#### $w_0$ 無しのSP - SDW - MWF (SDR - GSC)

【0067】

図6は、第2のマイクロフォンにおける異なる利得ミスマッチ $\mu_2$ に対しSDR - GSC (すなわち、フィルタ $w_0$ 無しのSP - SDW - MWF)によって得られた $1/\mu$ の関数としての改良SNR<sub>intelligible</sub>および発話歪みSD<sub>intelligible</sub>をプロットする。マイクロフォンミスマッチが無い場合、ノイズ参照への発話漏れの量は限定される。そのため、発話歪みの量は、全ての $\mu$ に対し低い。残響による少量の発話歪みがなお存在するため、 $1/\mu$ を増加することで、特に、 $1/\mu > 1$ とすることで、ノイズリダクションの量および発話漏れは若干減少する。マイクロフォンミスマッチが存在する場合、ノイズ参照への発話漏れの量は増加する。 $1/\mu = 0$  (GSC)において、発話はかなりの歪みを得る。望ましい信号のキャンセル操作により、改良SNR<sub>intelligible</sub>は劣化する。 $1/\mu > 0$ と設定することが、信号モデルエラーの存在しない状態での性能を妥協せずに、モデルエラーの存在する状態におけるGSCの性能を改良する。与えられたセットアップに対し、0.5近辺の $1/\mu$ の値が、4 dBまでの利得ミスマッチに対し良好な性能を保証するために適当に思える。

10

20

#### フィルタ $w_0$ を備えるSP - SDW - MWF

【0068】

図7は、フィルタ $w_0$ を備えるSP - SDW - MWFの性能測定量SNR<sub>intelligible</sub>およびSD<sub>intelligible</sub>をプロットする。一般的に、 $1/\mu$ を減少させることで、発話歪みの量およびノイズリダクションが増加する。 $1/\mu = 0$ において、ノイズリダクションにすべての重点を置く。図7に示されているように、これは、発話およびノイズ信号の全体としてのキャンセルの結果となり、そのため、性能が劣化する。モデルエラーがない場合、 $L_0 = 0$ および $L_0 = 0$ のセッティングは、( $1/\mu = 0$ 以外で)SNR<sub>intelligible</sub>と同じ結果となり、一方、 $w_0$ を備えるSP - SDW - MWFに対する歪みは、付加的なシングルチャンネルSDW - MWFにより、より高くなる。 $L_0 = 0$ において、性能は( $L_0 = 0$ の場合と比べて)マイクロフォンミスマッチにより劣化しない。

30

【0069】

図8は、 $QIC$ の関数として、QIC - GSCの改良SNR<sub>intelligible</sub>および発話歪みSD<sub>intelligible</sub>をそれぞれ示す。SDR - GSCのように、QICはGSCのローバスト性を増加する。QICは発話漏れの量と独立である。その結果、歪みは、利得ミスマッチを増加させると、より早く増加する。制限値 $\mu_{max}$ は、最大で可能な発話歪みのレベルが最も大きい可能なモデルエラーを超えないように、選択されるべきである。明らかに、これは、小さいモデルエラーに対し減少させたノイズリダクションを犠牲にしている。一方SDR - GSCは、全てのモデルエラーに対し限定された発話歪みを保持する(図6参照)。発話漏れの量が増加するならば、発話歪みに対する重要性がより増加する。その結果、小さいモデルエラーに対しよりよいノイズリダクション性能が得られ、一方、大きなモデルエラーに対して十分なローバスト性を保証する。図7は、付加フィルタ $w_0$ が信号モデルエラーの存在する状態から改良されていることを示している。

40

【0070】

先に議論した実施例において、一般化されたノイズリダクションの計画が達成され、SDW - MWFに基づく固定空間プリプロセッサおよび適合ステージを備える、空間的にプリプロセッサされた発話歪みに重み付けした複数チャンネルウィナー・フィルタ(SP - SD

50

W - MWF)として参照される。新しい計画は特別なケースとしてGSCおよびMWFを包含する。また、中間の回答を発話歪みで一般化されたGSC(SDR-GSC)として呼ぶ。トレードオフパラメータ $\mu$ のセッティングおよび発話参照のフィルタ $w_0$ の存在または不存在に従って、GSC、(SDW-)MWFが得られる。SP-SDW-MWFの異なるパラメータのセッティングが以下のように表される：

- ・ $w_0$ なしで、SP-SDW-MWFはSDR-GSCに対応する：ANC設計基準が、信号モデルエラーによる発話歪みを限定する一般化項で実行される。 $1/\mu$ が大きくなると、歪み量は少なくなる。 $1/\mu = 0$ において、GSCの解に対応する歪みは完全に無視される。SDR-GSCはGSCの感度を信号モデルエラーまで減少させるQIC-GSCに対する他の技術となる。QIC-GSCと比べて、SDR-GSCは、発話漏れの量が  
10  
増えると、発話歪みに重点をシフトする。信号モデルエラーが存在する場合、GSCの性能は保持される。その結果、小さなモデルエラーに対してより良いノイズリダクション性能が得られ、一方大きなモデルエラーに対するローバスト性も保証される。

- ・SP-SDW-MWFは発話歪みを明らかに考慮しているため、発話参照上のフィルタ $w_0$ が付加される。(発話漏れの存在しない状態で、無限の長さのフィルタ長に対し)、SP-SDW-MWFは、SDW-SWFポストフィルタを備えるSDR-GSCのカスケード表示に対応する。発話漏れの存在する状態で、 $w_0$ を備えるSP-SDW-MWFはその性能を維持するよう試み：SP-SDW-MWFは、発話漏れによる性能劣化を補償する特別なフィルタリング操作を含む。SDR-GSC(およびそのためGSC)と  
20  
比べて、性能はマイクロフォンミスマッチにより劣化しない。

補聴器の応用に対する実験結果は、理論結果を確認する。SP-SDW-MWFは、実際に、信号モデルエラーに対しGSCのローバスト性を増加する。広く研究されたQIC-GSCとの比較は、SP-SDW-MWFが与えられた最大許容発話歪みに対するより良いノイズリダクション性能を達成することを、示している。

### 確率的勾配の実行

#### 【0071】

(SDW-)MWFの機能的な実施が、GSVDまたはQRの分離に基づいて提案されている。また、サブバンドでの実施が、フルバンドアプローチと比べてかなり低いコストで改良された了解度の結果となる。これらの技術はSP-SDW-MWFを実施するために  
30  
拡張することができる。しかしながら、GSCおよびQIC-GSCと比べて、安価な確率的勾配に基づくSP-SDW-MWFの実行を行うことはできない。本発明では、QIC-GSCを超えるマトリックスに基づくSP-SDW-MWFの高価を保存する、SP-SDW-MWFの時間領域および周波数領域の確率的勾配の実施が提案されている。実験結果は、SP-SDW-MWFの提案された確率的勾配に基づく実行がSPAより性能が優れており、一方、それらの計算コストは限定されることを示している。

#### 【0072】

SP-SDW-MWFのコスト関数から、時間領域での確率的勾配アルゴリズムが引き出される。収束スピードを増やし、計算の複雑性を減少するために、確率的勾配アルゴリズムが周波数領域で実施される。確率的勾配アルゴリズムは、非常に時間で変化するノイズ  
40  
シナリオに用いられたとき、大きな過度のエラーの損害を得るため、性能は、発話歪みを限定する勾配の推定の一部にローパスフィルタを適用することによって、改良される。ローパスフィルタは、非常に時間で変化する望ましい発話要素の歪みを避ける、一方、時間で変化するノイズシナリオで必要なトラッキング性能を劣化しない。次に、異なる周波数領域の確率的勾配アルゴリズムの性能を比較する。実験結果は、提案された確率的勾配アルゴリズムが、QIC-GSCよりSP-SDW-MWFの高価を保持することを示している。最後に、ローパスフィルタを備える周波数領域の確率的勾配アルゴリズムのメモリーコストが、データバッファの代わりに(対角線の)相関マトリックスを使用する周波数領域での正規化項を近似することによって、減少することが示される。実験は、相関マトリックスを使用する確率的勾配アルゴリズムが、ローパスフィルタを備える確率的勾配ア  
50

ルゴリズムと同等の性能を有していることを示す。

確率的勾配アルゴリズム

誘導

【 0 0 7 3 】

確率的勾配アルゴリズムは、瞬間の傾きの推定値を使用して、最も急な下降アルゴリズムを近似する。コスト関数（式 3 8）が与えられると、最も急な下降アルゴリズムが以下のように繰り返される（今後、適合フィルタ  $w_{0:M-1}$  中のサブスクリプト  $0:M-1$  および入力ベクトル  $y_{0:M-1}$  は簡単にするため除外されることに注意のこと）：

$$w[n+1] = w[n] + \frac{\rho}{2} \left( -\frac{\partial J(w)}{\partial w} \right)_{w=w[n]} \tag{式 4 8}$$

$$= w[n] + \rho \left( E\{y^n[k]y_0^{n*}[k-\Delta]\} - E\{y^n[k]y^{n,H}[k]\}w[n] - \frac{1}{\mu} E\{y^s[k]y^{s,H}[k]\}w[n] \right),$$

10

$$w[k], y[k] \in C^{NL \times 1}$$

を備え、ここで、Nは適合フィルタに対する入力チャンネルの数を示し、Lはチャンネル毎にフィルタタップの数を示す。繰り返し指数 n を時間指数 k に置き換え、予想値 E { } を除外することで、以下の更新式が得られる。

$$w[k+1] = w[k] + \rho \left\{ y^n[k](y_0^{n*}[k-\Delta] - y^{n,H}[k]w[k]) - \frac{1}{\mu} y^s[k]y^{s,H}[k]w[k] \right\} + r[k] \tag{式 4 9}$$

20

発話参照において  $1 / \mu = 0$  およびフィルタ  $w_0$  がない場合、（式 4 9）は、ノイズのみの期間（すなわち、 $y_i[k] = y_i^n[k]$ ,  $i=1, \dots, M-1$  のとき）G S C で使用される更新公式まで減少する。傾きの推定値における付加項  $r[k]$  は、考え得る信号モデルエラーにより発話歪みを限定する。

【 0 0 7 4 】

式（ 4 9 ）はきれいな発話の相関マトリックス  $y^S[k] y^{S,H}[k]$  または  $E \{ y^S[k] y^{S,H}[k] \}$  の知識を必要とする。実際には、この情報を利用することはできない。キャリブレーションの必要を避けるために、発話 + ノイズ信号ベクトル  $y_{buf_1}$  は、処理の最中に、循環バッファ  $B_{1 \times R_{N \times L_{buf_1}}}$  に記憶される。ノイズのみの期間（すなわち、 $y_i[k] = y_i^n[k]$ ,  $i=1, \dots, M-1$  のとき）中、フィルタ  $w$  は、（式 4 9）中の

30

$$r[k] = \frac{1}{\mu} y^s[k] y^{s,H}[k] w[k]$$

項の以下の近似を使用して、更新され、

$$\frac{1}{\mu} y^s y^{s,H}[k] w[k] \approx \frac{1}{\mu} (y_{buf_1} y_{buf_1}^H[k] - y y^H[k]) w[k], \tag{式 5 0}$$

結果として更新公式は以下ようになる。

40

$$w[k+1] = w[k] + \rho \left\{ y[k](y_0^*[k-\Delta] - y^H[k]w[k]) - \frac{1}{\mu} (y_{buf_1}[k]y_{buf_1}^H[k] - y[k]y^H[k])w[k] \right\} + r[k] \tag{式 5 1}$$

その後、以下の式の一般化ステップサイズ  $\rho'$  が使用され、

$$\rho = \frac{\rho'}{\frac{1}{\mu} |y_{buf_1}^H[k]y_{buf_1}[k] - y^H[k]y[k]| + y^H[k]y[k] + \delta}, \tag{式 5 2}$$

ここで、 $\delta$  は小さい正の定数である。絶対値

50

$$|\mathbf{y}_{buf_1}^H \mathbf{y}_{buf_1} - \mathbf{y}^H \mathbf{y}|$$

が、きれいな発話エネルギー  $\mathbf{y}^{S,H}[k] \mathbf{y}^S[k]$  の正の値の推定値を保証するために、挿入される。第2バッファ  $B_2$   $R_{M \times L_{buf_2}}$  におけるノイズのみのベクトル  $\mathbf{y}_{buf_2}$  の付加的な記憶は、以下の式を使用して、発話 + ノイズの期間中もまた  $\mathbf{w}$  を適合させることを許す。

$$\mathbf{w}[k+1] = \mathbf{w}[k] + \rho \left\{ \mathbf{y}_{buf_2}[k] (\mathbf{y}_{0,buf_2}^*[k-\Delta] - \mathbf{y}_{buf_2}^H[k] \mathbf{w}[k]) + \frac{1}{\mu} (\mathbf{y}_{buf_2}[k] \mathbf{y}_{buf_2}^H[k] - \mathbf{y}[k] \mathbf{y}^H[k]) \mathbf{w}[k] \right\} \quad (式53)$$

10

ここで

$$\rho = \frac{\rho'}{\frac{1}{\mu} |\mathbf{y}^H[k] \mathbf{y}[k] - \mathbf{y}_{buf_2}^H[k] \mathbf{y}_{buf_2}[k]| + \mathbf{y}_{buf_2}^H[k] \mathbf{y}_{buf_2}[k] + \delta} \quad (式54)$$

である。

簡潔さのみの理由で、ノイズのみの期間における時間領域の確率的勾配アルゴリズムの更新手順がその後考慮され、ここでは  $\mathbf{y}[k] = \mathbf{y}^n[k]$  である。第2のノイズのみのバッファ  $B_2$  を使用した発話 + ノイズ期間中の更新操作に対する継続は複雑でなく：ノイズのみの入力ベクトル  $\mathbf{y}[k]$  を  $\mathbf{y}_{buf_2}[k]$  で置き換え、発話 + ノイズベクトル  $\mathbf{y}_{buf_1}[k]$  を入力発話 + ノイズベクトル  $\mathbf{y}[k]$  で置き換えることで、式を得ることができる。

20

アルゴリズム (式51) - (式52) は、ステップサイズ  $\rho$  が

$$E\left\{\frac{1}{\mu} \mathbf{y}_{buf_1} \mathbf{y}_{buf_1}^H + (1 - \frac{1}{\mu}) \mathbf{y} \mathbf{y}^H\right\}$$

の最大固有値である  $\lambda_{\max}$  を有する  $2 / \lambda_{\max}$  より小さい場合に、収束する。標準的な NLS を備える (式51) の類似性から、 $\lambda_i, i=1, \dots, NL$  を有する

$$\rho < \frac{2}{\sum_i \lambda_i}$$

のセッティング、

$$E\left\{\frac{1}{\mu} \mathbf{y}_{buf_1} \mathbf{y}_{buf_1}^H + (1 - \frac{1}{\mu}) \mathbf{y} \mathbf{y}^H\right\} \in R^{NL \times NL}$$

30

の固有値、または、- FIRフィルタの場合 - 以下の式のセッティングは、

$$\rho < \frac{2}{\frac{1}{\mu} L \sum_{i=M-N}^{M-1} E\{y_{i,buf_1}^2[k]\} + (1 - \frac{1}{\mu}) L \sum_{i=M-N}^{M-1} E\{y_i^2[k]\}} \quad (式55)$$

平均平方における収束を保証する。式(55)はステップサイズ  $\rho$  に対する(式52)および(式54)の正規化を説明する。

【0075】

しかしながら、一般的に以下の関係にあるため、

$$\mathbf{y}[k] \mathbf{y}^H[k] \neq \mathbf{y}_{buf_1}^n[k] \mathbf{y}_{buf_1}^{n,H}[k], \quad (式56)$$

40

(式51)における瞬間の傾き推定は、- (式49)と比較して - 以下の式によって付加的に乱される、

$$\frac{1}{\mu} (\mathbf{y}[k] \mathbf{y}^H[k] - \mathbf{y}_{buf_1}^n[k] \mathbf{y}_{buf_1}^{n,H}[k]) \mathbf{w}[k], \quad (式57)$$

ここで  $1 / \mu \neq 0$  である。そのため、 $1 / \mu \neq 0$  において、更新式(式51) - (式54)は、(式49)より大きい残留過剰エラーを受ける。この付加的な過剰エラーは、 $\mu$  の減少、ステップサイズ  $\rho$  の増加、ベクトル  $\mathbf{y}$  のベクトル長  $L$   $N$  の増加に対し、増加する。高い非静止ノイズ、例えば複数話者バブルノイズに対し特に大きくなることが予想される。

$\mu > 1$  において、いくつかの独立的な仮定を行うことによって、アルゴリズム(式51)

50

- (式54) から引き出されることに注意のこと。同時に、しかしながら、これらの独立的な仮定がかなりの性能の劣化の結果となり、一方、計算の複雑性をほとんど減少させないことがわかる。

周波数領域での実行

【0076】

上述したように、確率的勾配アルゴリズム(式51) - (式54) は、異なる時間点  $k$  で測定されたランク 1 の相関マトリックス  $y^n[k] y^{n,H}[k]$  の間の大きな相違により、大きな  $\sigma^2 / \mu$ 、および  $\sigma^2$  または、高い時間で変化するノイズに対し、大きな過剰エラーを受けることが予想される。傾きの推定は、(式51) 中の以下の式

$$y_{buf_1}[k] y_{buf_1}^H[k] - y[k] y^H[k] \quad (\text{式58})$$

を、以下の式で置き換えることで改良することができる：

$$\frac{1}{K} \sum_{l=k-K+1}^k y_{buf_1}[l] y_{buf_1}^H[l] - \frac{1}{K} \sum_{l=k-K+1}^k y[l] y^H[l], \quad (\text{式59})$$

ここで、

$$\frac{1}{K} \sum_{l=k-K+1}^k y_{buf_1}[l] y_{buf_1}^H[l]$$

は発話 + ノイズの期間において更新され、

$$\frac{1}{K} \sum_{l=k-K+1}^k y[l] y^H[l]$$

はノイズのみの期間において更新される。しかしながら、これは高価なマトリクス演算を必要とする。ブロックに基づく実行は、元来、この平均化を行う：

$$w[(k+1)K] = w[kK] + \frac{\rho}{K} \left[ \sum_{i=0}^{K-1} y[kK+i] (y_0^*[kK+i-\Delta] - y^H[kK+i]) w[kK] \right. \\ \left. - \frac{1}{\mu} \sum_{i=0}^{K-1} (y_{buf_1}[kK+i] y_{buf_1}^H[kK+i] - y[kK+i] y^H[kK+i]) w[kK] \right]. \quad (\text{式60})$$

勾配そしてそのため

$$y_{buf_1}[k] y_{buf_1}^H[k] - y[k] y^H[k]$$

は、 $w$  に対する調整を行う前に  $k$  回以上の繰り返しで、平均化される。これは減少 (すなわち係数  $k$  によって) された収束速度の犠牲にして達成される。

【0077】

ブロックに基づく実行は、それが周波数領域で特に大きなフィルタ長に対し実行されるとき、計算上より効率がよく：線型のたたみ込みあるいは相関関係は、オーバーラップセーブまたはオーバーラップアッドに基づく FFT アルゴリズムによって効率的に実現化可能である。また、周波数領域での実行では、各周波数ピンがそれ自身のステップサイズを有し、そのため、時間領域の実行と比較してより早く収束し、一方、定常状態の過剰 MSE を劣化しない結果となる。

【0078】

アルゴリズム 1 は、(式51) - (式54) のオーバーラップセーブに基づく周波数領域での実行を合計する。アルゴリズム 1 は、長さ  $2L$  の  $(3N+4)$  の FFT を必要とする。FFT 変換された発話 + ノイズおよびバッファ  $B_1$   $C^{N \times L_{buf1}}$  および  $B_2$   $C^{N \times L_{buf2}}$  のそれぞれにおけるノイズのみのベクトルを記憶することによって、周波数領域のベクトルを記憶する代わりに、 $N$  回の FFT 操作を省くことができる。入力信号は実数であるため、FFT 要素の半分は共役の複素数であることに注意のこと。そのため、実際には、複素 FFT 要素の半分のみがメモリに記憶される。発話 + ノイズの期間に適合かするときは、以下の時間領域ベクトル

10

20

30

40

50

$$[y_0[kL-\Delta] \quad L \quad y_0[kL-\Delta+L-1]]^T \quad (\text{式61})$$

が、ノイズのみの期間において、付加的なバッファ

$$\mathbf{B}_{2,0} \in R^{L \times \frac{L_{buf2}}{2}}$$

に記憶されるべきであり、このことは、(N = Mにおいて)、時間領域のベクトルがバッファ B<sub>1</sub> および B<sub>2</sub> に記憶されるときと比較して、L<sub>buf2</sub> / 2ワードの付加的な記憶量の結果となる。

アルゴリズム 1 において、共通のトレードオフパラメータ μ がすべての周波数ビンにおいて使用されていることに注意のこと。あるいは、μ に対する異なるセッティングを異なる周波数ビンにおいて使用することができる。例えば、w<sub>0</sub> = 0 を有する SP - SDW - MWF に対し、1 / μ は、GSC が十分なローパス性を有する周波数で、例えば、高い周波数の小さいサイズのアレイに対し、0 にセットすることができる。これに対し、正規化項 R<sub>i</sub>[k], i=M-N, ..., M-1 の極まれな周波数要素は計算されることが必要で、計算の複雑性を減少させる。

10

### アルゴリズム 1 : オーバーラップセーブに基づく周波数領域での確率的勾配 SP - SDW - MWF

初期化 :

$$\mathbf{W}_i[0] = [0 \quad L \quad 0]^T, \quad i = M-N, \dots, M-1$$

20

$$P_m[0] = \delta_m, \quad m = 0, \dots, 2L-1$$

マトリックスの定義 :

$$\mathbf{g} = \begin{bmatrix} \mathbf{I}_L & \mathbf{0}_L \\ \mathbf{0}_L & \mathbf{I}_L \end{bmatrix}; \mathbf{k} = [\mathbf{0}_L \quad \mathbf{I}_L]; \mathbf{F} = 2L \times 2L \text{ DFT matrix};$$

NL 入力サンプルの各新規ブロックに対し :

ノイズが検出された場合は :

$$1. \mathbf{F}[y_i[kL-L] \quad \dots \quad y_i[kL+L-1]]^T, \quad i = M-N, \dots, M-1 \rightarrow \text{ノイズバッファ } \mathbf{B}_2$$

30

$$[y_0[kL-\Delta] \quad \dots \quad y_0[kL-\Delta+L-1]]^T \rightarrow \text{ノイズバッファ } \mathbf{B}_{2,0}$$

$$2. \mathbf{Y}_i^n[k] = \text{diag}\{\mathbf{F}[y_i[kL-L] \quad \dots \quad y_i[kL+L-1]]^T\}, \quad i = M-N, \dots, M-1$$

$$\mathbf{d}[k] = [y_0[kL-\Delta] \quad L \quad y_0[kL-\Delta+L-1]]^T$$

発話 + ノイズバッファ B<sub>1</sub> 中のデータから Y<sub>i</sub>[k] を作成する。

発話が検出された場合は :

$$1. \mathbf{F}[y_i[kL-L] \quad \dots \quad y_i[kL+L-1]]^T, \quad i = M-N, \dots, M-1 \rightarrow \text{発話 + ノイズバッファ } \mathbf{B}_1$$

$$2. \mathbf{Y}_i[k] = \text{diag}\{\mathbf{F}[y_i[kL-L] \quad \dots \quad y_i[kL+L-1]]^T\}, \quad i = M-N, \dots, M-1$$

40

ノイズバッファ B<sub>2,0</sub> および B<sub>2</sub> から d[k] および Y<sub>i</sub><sup>n</sup>[k] を作成する。

更新のための公式 :

$$\begin{aligned}
 1. \quad & \mathbf{e}_1[k] = \mathbf{kF}^{-1} \sum_{j=M-N}^{M-1} \mathbf{Y}_j^n[k] \mathbf{W}_j[k] = \mathbf{y}_{out,1} \\
 & \mathbf{e}[k] = \mathbf{d}[k] - \mathbf{e}_1[k] \\
 & \mathbf{e}_2[k] = \mathbf{kF}^{-1} \sum_{j=M-N}^{M-1} \mathbf{Y}_j[k] \mathbf{W}_j[k] = \mathbf{y}_{out,2} \\
 & \mathbf{E}_1[k] = \mathbf{Fk}^T \mathbf{e}_1[k]; \mathbf{E}_2[k] = \mathbf{Fk}^T \mathbf{e}_2[k]; \quad \mathbf{E}[k] = \mathbf{Fk}^T \mathbf{e}[k] \\
 2. \quad & \Lambda[k] = \frac{2\mu}{L} \text{diag}\{P_0^{-1}[k], \dots, P_{2L-1}^{-1}[k]\} \\
 & P_m[k] = \gamma P_m[k-1] + (1-\gamma) \left( \sum_{j=M-N}^{M-1} |\mathbf{Y}_{j,m}^n|^2 + \frac{1}{\mu} \left| \sum_{j=M-N}^{M-1} (\mathbf{Y}_{j,m} - \mathbf{Y}_{j,m}^n) \right|^2 \right) \\
 3. \quad & \mathbf{W}_i[k+1] = \mathbf{W}_i[k] + \mathbf{FgF}^{-1} \Lambda[k] \left\{ \mathbf{Y}_i^{n,H}[k] \mathbf{E}[k] - \frac{1}{\mu} (\mathbf{Y}_i^H \mathbf{E}_2[k] - \mathbf{Y}_i^{n,H} \mathbf{E}_1[k]) \right\}, \\
 & \hspace{15em} (i=M-N, \dots, M-1)
 \end{aligned}$$

10

出力：

$$y_0[k] = [y_0[kL-\Delta] \quad L \quad y_0[kL-\Delta+L-1]]^T$$

- ・ ノイズが検出された場合：  $y_{out}[k] = y_0[k] - y_{out,1}[k]$
- ・ 発話が検出された場合：  $y_{out}[k] = y_0[k] - y_{out,2}[k]$

改良1：ローパスフィルタを有する確率的勾配アルゴリズム

【0079】

20

スペクトル的に静止したノイズに対し、ブロックに基づく周波数領域の確率的勾配の実行による(式59)の限定された(すなわち  $K=L$ )平均化は、短時間の発話相関マトリックス  $E\{y^s y^{s,H}\}$  の妥当な推定を提供することができる。しかしながら、実際のシナリオでは、発話信号およびノイズ信号は、しばしば、スペクトル的に非常に非静止状態であり(例えば、複数話者のバブルノイズ)、一方、それらの長時間のスペクトルおよび空間特性(例えば、ソースの位置)は、通常、時間的に非常にゆっくりと変化する。これらのシナリオに対し、短時間のスペクトルより空間的な特性を得る長時間の発話相関マトリックス  $E\{y^s y^{s,H}\}$  の信頼性のある推定は、なお、 $K \gg L$  のサンプルにわたって(式59)を平均化することによって、得られる。スペクトル的に非常に非静止のノイズは、正規化項  $r[k]$  における長時間の発話相関マトリックスの推定を使用することによって、なお、空間的に抑制される。発話歪みを考慮する(すなわち、(式51)中の  $r[k]$  項)勾配推定の部分をローパスフィルタでフィルタリングすることによる、確率的勾配アルゴリズムにおける(式59)の長時間の平均化( $K \gg L$ )を組み込むことによる安価な方法を、以下、提案する。平均化方法を、まず第1に、時間領域のアルゴリズム(式51) - (式54)に対し説明し、次に、周波数領域の実行に翻訳する。

30

長時間のスペクトルおよび空間特性が、少なくとも  $K$  の発話 + ノイズサンプルと  $K$  のノイズサンプルにおいて擬似静止状態であると仮定する。長時間の発話相関マトリックス  $E\{y^s y^{s,H}\}$  の信頼性のある推定が、 $K \gg L$  において(式59)によって得られる。高価なマトリックスの計算を避けるため、 $r[k]$  を以下の式に近似することができる。

$$\frac{1}{K} \sum_{l=k-K+1}^{l=k} \left( y_{buf_1}[l] y_{buf_1}^H[l] - y[l] y^H[l] \right) w[l], \quad (式62)$$

40

確率的勾配アルゴリズムのフィルタ係数  $w$  は時間的にゆっくり変化するため、(式62)は、特に小さいステップサイズ  $\mu$  に対し、 $r[k]$  と良い近似を示す。平均化操作(式62)は、(式51)においてローパスフィルタを  $r[k]$  に適用することによって行われる：

$$r[k] = \mathcal{L}\{r[k-1]\} + (1 - \mathcal{L}\{\frac{1}{\mu}\}) \left( y_{buf_1}[k] y_{buf_1}^H[k] - y[k] y^H[k] \right) w[k], \quad (式63)$$

ここで、

$$\mathcal{L}\{x\}$$

50



である。これは、約

$$\frac{1}{1-\beta}$$

のサンプルの平均化ウィンドウ  $K$  に対応する。この一般化されたステップサイズ  $\beta$  は、以下の式に変更される、

$$\rho = \frac{\rho'}{r_{\text{avg}}[k] + \mathbf{y}^H[k]\mathbf{y}[k] + \delta} \quad (\text{式 6 4})$$

$$r_{\text{avg}}[k] = \beta r_{\text{avg}}[k-1] + (1-\beta) \frac{1}{\mu} |\mathbf{y}_{\text{buf}_1}^H[k]\mathbf{y}_{\text{buf}_1}[k] - \mathbf{y}^H[k]\mathbf{y}[k]|. \quad (\text{式 6 5})$$

10

(式 5 1) と比較して、(式 6 3) は、3NL-1 の付加的な MAC および NLx1 ベクトル  $\mathbf{r}[k]$  の特別な記録装置を必要とする。

【 0 0 8 0 】

(式 6 3) は周波数領域に簡単に拡張できる。アルゴリズム 1 における  $w_i[k+1]$  に対する更新式は、(アルゴリズム 2) となる：

$$\begin{aligned} \mathbf{W}_i[k+1] &= \mathbf{W}_i[k] + \mathbf{F}\mathbf{g}\mathbf{F}^{-1}\Delta[k](\mathbf{Y}_i^{n,H}[k]\mathbf{E}[k] - \mathbf{R}_i[k]); \\ \mathbf{R}_i[k] &= \lambda\mathbf{R}_i[k-1] + (1-\lambda) \frac{1}{\mu} (\mathbf{Y}_i^H[k]\mathbf{E}_2[k] - \mathbf{Y}_i^{n,H}[k]\mathbf{E}_1[k]) \end{aligned} \quad (\text{式 6 6})$$

20

ここで、

$$\mathbf{E}[k] = \mathbf{F}\mathbf{k}^T \left( \mathbf{y}_0^n[k] - \mathbf{k}\mathbf{F}^{-1} \sum_{j=M-N}^{M-1} \mathbf{Y}_j^n[k]\mathbf{W}_j[k] \right); \quad (\text{式 6 7})$$

$$\mathbf{E}_1[k] = \mathbf{F}\mathbf{k}^T \mathbf{k}\mathbf{F}^{-1} \sum_{j=M-N}^{M-1} \mathbf{Y}_j^n[k]\mathbf{W}_j[k]; \quad (\text{式 6 8})$$

$$\mathbf{E}_2[k] = \mathbf{F}\mathbf{k}^T \mathbf{k}\mathbf{F}^{-1} \sum_{j=M-N}^{M-1} \mathbf{Y}_j[k]\mathbf{W}_j[k]. \quad (\text{式 6 9})$$

であり、また、 $\Delta[k]$  は以下のように計算される：

$$\Delta[k] = \frac{2\rho'}{L} \text{diag}\{P_0^{-1}[k], \dots, P_{2L-1}^{-1}[k]\} \quad (\text{式 7 0})$$

30

$$P_m[k] = \gamma P_m[k-1] + (1-\gamma)(P_{1,m}[k] + P_{2,m}[k]) \quad (\text{式 7 1})$$

$$P_{1,m}[k] = \sum_{j=M-N}^{M-1} |\mathbf{Y}_{j,m}^n[k]|^2 \quad (\text{式 7 2})$$

$$P_{2,m}[k] = \lambda P_{2,m}[k-1] + (1-\lambda) \frac{1}{\mu} \left| \sum_{j=M-N}^{M-1} (|\mathbf{Y}_{j,m}[k]|^2 - |\mathbf{Y}_{j,m}^n[k]|^2) \right|. \quad (\text{式 7 3})$$

アルゴリズム 1 と比較して、(式 6 6) - (式 6 9) は、特別の 2L 点の FFT と 8NL-2N-2L の特別の L に対する MAC サンプルと、2NLx1 の実数データベクトルの付加的な記憶装置とを必要とする。K=1 の時間領域でのバージョンと同様の平均化操作における時間定数を 40 得るために、 $\beta$  は

$$\beta^k.$$

と等しくなるべきである。

以下の実験結果は、確率的勾配アルゴリズムの性能が特に大きな  $\beta$  に対しローパスフィルタによってかなり改良されることを示している。

【 0 0 8 1 】

次に、異なる確率的勾配アルゴリズムの計算の複雑性を議論する。表 1 は、時間領域 (TD) および周波数領域 (FD) における確率的勾配 (SG) に基づくアルゴリズムの、計算の複雑性を要約して示す (実数の乗加算 (MAC)、分割 (D)、平方根 (Sq) およ 50

び絶対値 (Abs) の数として表現される)。比較は、標準化された NLMS および NLMS に基づく SPA と行う。1つの複素数の乗算は、4つの実数の乗算および2つの実数の加算と同等と仮定する。実数の入力ベクトルの 2L 点での FFT は、 $2L \log_2 2L$  の実数の MAC (2 基数の FFT アルゴリズムと仮定した場合) を必要とする。

表 1 は、フィルタ  $w_0$  を備えていない TD - SG アルゴリズムおよび SPA が、標準の ANC より、約 2 倍複雑であることを示している。正規化項にローパスフィルタ (LP) を適用したときは、TD - SG アルゴリズムは ANC の約 3 倍の複雑性を有している。周波数領域での実行の複雑性の増加はほとんどない。

【表 1】

アルゴリズム	更新の公式	ステップサイズの適合	
TD	NLMS ANC	$(2M-2)L+1$ MAC	10
	NLMS based SPA	$(4(M-1)L+1)MAC+1D+1Sq$	$1D+(M-1)LMAC$
	SG	$(4NL+5)MAC$	$1D+1Abs+(2NL+2)MAC$
	SG with LP	$(7NL+4)MAC$	$1D+1Abs+(2NL+4)MAC$
FD	NLMS ANC	$(10M-7-\frac{4(M-1)}{L})+$ $(6M-2)\log_2 2L MAC$	$1D+(2M+2)MAC$
	NLMS based SPA	$14M-11-\frac{4(M-1)}{L}+$ $(6M-2)\log_2 2L MAC$ $+1/L Sq+1/LD$	$1D+(2M+2)MAC$
	SG	$(18N+6-\frac{8N}{L})+$	$1D+1Abs+(4N+4)MAC$
	(Algorithm 1)	$(6N+8)\log_2 2L MAC$	
	SG with LP	$(26N+4-\frac{10N}{L})$	$1D+1Abs+(4N+6)MAC$
(Algorithm 2)	$+(6N+10)\log_2 2L MAC$		20

【0082】

表示したように、図 9 は、 $M = 3$  に対し  $L$  の関数でサンプリング周波数が  $f_s = 16 \text{ kHz}$  の LP フィルタを有する時間領域および周波数領域の確率的勾配アルゴリズムの複雑性をプロットする (秒あたりメガの操作の数 (Mops) として表現される)。比較は、GSC の NLMS に基づく ANC および SPA と行う。小さい  $M$  に対し、FD - NLMS ANC のコストと比較されるため、FD SPA の複雑性は記載されていない。 $L > 8$  において、周波数領域の実行は、それらの時間領域の同等物と比べてかなり低い複雑性の結果となる。LP を備える FD 確率論的な勾配アルゴリズムの計算の複雑性は限定され、補聴器での実行に対し SPA の良い代替となる。

表 1 および図 9 において、時間領域および周波数領域での NLMS ANC および NLMS に基づく SPA の複雑性は、適合フィルタのみがノイズのみの期間に更新されるときに複雑性を表す。適合フィルタがノイズバッファからのデータを使用して発話 + ノイズ期間にも更新される場合は、時間領域の実行は付加的にサンプル毎の  $NL MAC$  を必要とし、周波数領域の実行は付加的に  $2 FFT$  および  $L$  サンプル毎の  $(4L(M-1) - 2(M-1) + L)$  の MAC を必要とする。

【0083】

SP - SDW - MWF の異なる FD の確率的勾配の実行性能は、補聴器の応用に対する実験結果に基づいて評価される。比較は、FD - NLMS に基づく SPA と行う。比較を正当にするために、FD - NLMS に基づく SPA は、(確率的勾配アルゴリズムのように)、ノイズバッファからのデータを使用して発話 + ノイズ期間中に適用される。

【0084】

セットアップは、上述した例と同等である (図 5 を参照)。FD の確率的勾配アルゴリズムの性能は、チャンネル毎のフィルタ長  $L = 32$  タップ、 $\gamma = 0.8$  および  $\beta = 0$  に対し、評価される。空間プリプロセッサの影響を除外するため、性能の測定は、固定ビームフォーマーの出力に対して計算される。仮定した信号モデルにおけるエラーに対するアル

10

20

30

40

50

ゴリズムの感度は、マイクロフォン mismatch 例えば第2マイクロフォンの利得 mismatch  $\mu_2 = 4 \text{ dB}$  に対し、説明される。

【0085】

図10(a)および10(b)は、 $90^\circ$ で静止および非静止(例えば複数話者のバブル)のそれぞれの状態のノイズソースに対するトレードオフパラメータ $\mu$ の関数として、 $w_0$ なしの異なるFDの確率的勾配(SG)SP-SDW-MWFアルゴリズム(すなわち、SDR-GSC)の性能を比較している。性能上における近似式(式50)の影響を分析するために、きれいな発話を使用する、(式49)のFDの実行結果も示される。このアルゴリズムは最適FD-SGアルゴリズムと参照される。ローパス(LP)フィルタなしで、確率的勾配アルゴリズムは、特に大きな $1/\mu$ に対し、最適FD-SGアルゴリズム(式49)より悪い性能となる。静止状態の発話のようなノイズソースに対し、FD-SGアルゴリズムは、近似式(式50)からそれほど多くの影響を受けない。複数話者のバブルのような、非常に時間に対する変化の激しいノイズのシナリオでは、FDの実行における $r[k]$ の限定された平均化は、(式49)により達成される大きなノイズリダクションを維持するのに十分でない。ノイズリダクション性能におけるロス、減少された収束速度を犠牲にして、ステップサイズ $\mu$ を少なくすることによって、減少することができる。例えば $\mu = 0.999$ のローパスフィルタ(式66)を適用することは、すべての $1/\mu$ に対する性能をかなり改良するが、ノイズシナリオ中の変化がなおトラックされる。

10

【0086】

図11は、 $\mu$ がLPフィルタの指数重み付け要素である場合(式66を参照)の $1/(1-\mu)$ を関数とするバブルノイズシナリオに対し、フィルタ $w_0$ を有する場合と有さない場合において、SNRで改良した $\text{SNR}_{\text{intelligible}}$ およびSP-SDW-MWF( $1/\mu = 0.5$ )の発話歪み $\text{SD}_{\text{intelligible}}$ をプロットする。性能は $\mu$ を増加することで明らかに向上する。小さい $\mu$ に対し、 $w_0$ を有するSP-SDW-MWFは、 $w_0$ を有さないSP-SDW-MWFと比較して、より大きな過剰エラー、(そのためより悪い $\text{SNR}_{\text{intelligible}}$ )、を受ける。これは、 $E\{y^s y^{s,H}\}$ の大きな規模による。

20

【0087】

LPフィルタは、短時間の発話相関マトリックス $E\{y^s y^{s,H}\}$ の程度の悪い推定によって、および/または、より高い非静止状態の短時間の発話スペクトルによって、生じたフィルタの重み付け $w_i[k]$ における変動を減少する。ステップサイズ $\mu$ における減少と比較して、LPフィルタは、ノイズシナリオにおける変化のトラッキングに妥協しない。記載しているように、図12は、ノイズソースの位置が突然 $90^\circ$ から $180^\circ$ に変化したときの、 $\mu = 0$ および $\mu = 0.998$ のそれぞれに対する、 $w_0$ なしのFDの確率的勾配アルゴリズム(すなわちSDR-GSC)の収束挙動をプロットしている。 $4 \text{ dB}$ の利得の mismatch  $\mu_2$ が第2マイクロフォンに適用された。残留ノイズエネルギー $\sigma_n^2$ および発話歪みエネルギー $\sigma_d^2$ における速い変動を避けるために、この試験において望ましいおよび干渉するノイズソースは静止した発話のようなものである。上図は、入力サンプルの数の関数としての残留ノイズエネルギー $\sigma_n^2$ を示し、下図は、発話+ノイズサンプルの数の関数としての発話+ノイズ期間における残留発話歪み $\sigma_d^2$ をプロットする。両者のアルゴリズム(すなわち、 $\mu = 0$ および $\mu = 0.998$ )とも、ほぼ同じ収束速度を有する。位置の変化が起こると、 $\mu = 0.998$ のアルゴリズムがより速く収束する。 $\mu = 0$ に対し、バッファ中のノイズベクトルが最新のものにならないため、しばらくの間、近似エラー(式50)は大きく残る。 $\mu = 0.998$ に対し、瞬時の大きな近似エラーの影響は、ローパスフィルタのおかげで、減少する。

30

40

【0088】

図13および図14は、複数ノイズソースのシナリオにおいて、LPフィルタ( $\mu = 0.998$ )を有するFDの確率的勾配アルゴリズムおよびFD-NLMSに基づくSPAの性能を比較する。ノイズシナリオは、 $0^\circ$ における望ましい位置に対し、 $75^\circ$ 、 $120^\circ$ 、 $180^\circ$ 、 $240^\circ$ 、 $285^\circ$ の角度に位置する5つの複数話者バブルソースから構

50

成される。仮定された信号モデルにおけるエラーに対するアルゴリズムの感度を評価するため、性能上の、マイクロフォン mismatches すなわち第2マイクロフォンの利得 mismatches  $\gamma_2 = 4 \text{ dB}$  の影響も記載した。図13において、フィルタ  $w_0$  を有する場合と有さない場合において、SNRで改良した  $\text{SNR}_{\text{intelligible}}$  および  $\text{SP-SDW-MWF}$  の発話歪み  $\text{SD}_{\text{intelligible}}$  を、トレードオフパラメータ  $1/\mu$  の関数として記載する。図14は、FD-NLMSに基づくSPAを使用して実行された、異なる拘束値  $\beta^2$  に対する、QIC-GSCの性能

$$w^H w \leq \beta^2 \tag{式74}$$

を示す。

SPAおよび確率的勾配に基づくSP-SDW-MWFは両者とも、GSCのローバスタ性を増加する(すなわち、 $w_0$  なしで  $1/\mu = 0$  のSP-SDW-MWF)。与えられた最大値として取り得る発話歪み  $\text{SD}_{\text{intelligible}}$  に対し、 $w_0$  を有するおよび有さないSP-SDW-MWFは、SPAよりもより良好なノイズリダクション性能を達成する。 $w_0$  を有するSP-SDW-MWFの性能は、( $w_0$  を有さないSP-SDW-MWFと比べて)、マイクロフォン mismatches の影響を受けない。モデルエラーがない場合は、 $w_0$  を有するSP-SDW-MWFは、 $w_0$  を有さないSP-SDW-MWFより幾分悪い性能となる。これは、 $w_0$  を有する場合、 $(1/\mu) E\{y^s y^{s,H}\}$  の推定が、 $(1/\mu) E\{y^s y^{s,H}\}$  の大きな規模により、正確でなくなるとの事実によって、説明できる(図11を参照)。最後に、SP-SDW-MWFの提案された確率的勾配の実行が、QIC-GSCを超えるSP-SDW-MWFの効果

10

20

改良2：相関マトリックスを使用する周波数領域での確率的勾配アルゴリズム

【0089】

周波数領域における正規化項を近似することによって、(対角上の)発話およびノイズの相関マトリックスをデータバッファの代わりに使用でき、その結果、メモリーの使用量を劇的に減少でき、一方、計算上の複雑性をさらに減少する。実験結果は、この近似が、ローパスフィルタを有する確率的勾配アルゴリズムと比較して小さい(正または負の)性能の相違の結果となり、提案されたアルゴリズムはQIC-GSCを超えるSP-SDW-MWFのローバスタ性の効果を維持し、一方、その計算上の複雑性およびメモリー使用量の両者は、QIC-GSCを実行するためのNLMSに基づくSPAと匹敵する、ことを示している。

30

【0090】

(式51)中の  $r[k]$  の推定が非常に悪いことが示されており、大きな過剰エラーの結果となるため、(式59)において、平均クリーン発話相関マトリックスの推定を使用することが示唆されている。これは、 $r[k]$  を以下のように計算することを認める、

$$r[k] = \frac{1}{\mu} (1 - \tilde{\lambda}) \sum_{l=0}^k \tilde{\lambda}^{k-l} (\mathbf{y}_{buf_l} [l] \mathbf{y}_{buf_l}^H [l] - \mathbf{y}^n [l] \mathbf{y}^{n,H} [l]) \cdot \mathbf{w}[k], \tag{式75}$$

ここで、

$$\tilde{\lambda}$$

40

は指数重み付け要素である。静止状態のノイズに対し、小さい

$$\tilde{\lambda}$$

すなわち

$$1/(1 - \tilde{\lambda}) \sim NL$$

が満たされる。しかしながら、実際には、発話およびノイズ信号はスペクトル的に非常に非静止状態(例えば、複数話者バブルノイズ)であり、そのため、それらの長時間のスペクトルおよび空間特性は、通常、時間的に遅く変化する。スペクトル的に非常に非静

50

止状態のノイズは、なお、 $r[k]$ における長時間相関マトリックスの推定すなわち

$$1/(1-\tilde{\lambda}) \gg NL$$

を使用することで、空間的に抑制される。

計算のための高価なマトリックス操作(式75)を避けるために、予め、 $w[k]$ が時間的に遅く変化するすなわち  $w[k] = w[k]$  であると仮定し、その結果、(式75)は、直接的にローパスフィルタを正規化項  $r[k]$  に適用することで、マトリックスの操作の代わりにベクトルで近似することができる、(式63)を参照のこと、

$$r[k] = \frac{1}{\mu} (1-\tilde{\lambda}) \sum_{l=0}^k \tilde{\lambda}^{k-l} (\mathbf{y}_{buf_1}[l] \mathbf{y}_{buf_1}^H[l] - \mathbf{y}^n[l] \mathbf{y}^{n,H}[l]) \cdot \mathbf{w}[l] \quad (\text{式76})$$

10

$$= \lambda r[k-1] + (1-\lambda) \frac{1}{\mu} (\mathbf{y}_{buf_1}[k] \mathbf{y}_{buf_1}^H[k] - \mathbf{y}^n[k] \mathbf{y}^{n,H}[k]) \mathbf{w}[k]. \quad (\text{式77})$$

しかしながら、この仮定は、これから示すように、周波数領域での実行を実際には必要としない。

【0091】

アルゴリズム2と呼ばれる周波数領域でのアルゴリズムは大きなデータバッファを必要とし、そのため、大きなデータ量の記憶装置を必要とする(良好な性能を発揮するためには、循環バッファ  $B_1$  および  $B_2$  のバッファ長が 10000...20000 であることに注意のこと)。実質的なメモリー(および計算の複雑性)の減少は以下の2つのステップにより達成できる:

20

- ・正規化項を計算するために(式77)の代わりに(式75)を使用するとき、データサンプルの代わりに相関マトリックスを記憶する必要がある。結果としてのアルゴリズムの周波数領域の実行は、アルゴリズム3に要約され、ここでは、 $2L \times 2L$ の大きさの発話およびノイズ相関マトリックス  $S_{ij}[k]$  および  $S_{ij}^n[k]$ ,  $i, j = M-N...M-1$  が、正規化項  $R_i[k]$  およびステップサイズ  $\mu[k]$  (の部分) を計算するために使用される。これらの相関マトリックスは、発話+ノイズ期間では望ましい信号がノイズバッファ  $B_2$  からこれ以上構成されないため、発話+ノイズ期間およびノイズのみの期間のそれぞれにおいて更新される。

30

- ・周波数領域での相関マトリックスは、アルゴリズム3中の  $\mathbf{F} \mathbf{k}^T \mathbf{k} \mathbf{F}^{-1}$  が  $1_{2L}/2$  によって良く近似できるため、対角マトリックスによって近似することができる。そのため、発話およびノイズ相関マトリックスは以下のように更新され、

$$S_{ij}[k] = \lambda S_{ij}[k-1] + (1-\lambda) \mathbf{Y}_i^H[k] \mathbf{Y}_j[k] / 2, \quad (\text{式78})$$

$$S_{ij}^n[k] = \lambda S_{ij}^n[k-1] + (1-\lambda) \mathbf{Y}_i^{n,H}[k] \mathbf{Y}_j^n[k] / 2, \quad (\text{式79})$$

メモリー使用量および計算上の複雑性のかなりの減少となり、一方、性能およびローパス性について最小の影響しを与えない。このアルゴリズムは、アルゴリズム4として参照される。

40

アルゴリズム3 相関マトリックス(近似なし)を有する周波数領域での実行初期化およびマトリックスの定義

$$\mathbf{W}_i[0] = [0 \quad L \quad 0]^T, i = M-N...M-1$$

$$P_m[0] = \delta_m, m = 0...2L-1$$

$F = 2L \times 2L$ の大きさのDFTマトリックス

50

$$\mathbf{g} = \begin{bmatrix} \mathbf{I}_L & \mathbf{0}_L \\ \mathbf{0}_L & \mathbf{0}_L \end{bmatrix}, \quad \mathbf{k} = [\mathbf{0}_L \quad \mathbf{I}_L]$$

$\mathbf{0}_L = L \times L$  の大きさのゼロマトリックス、 $\mathbf{I}_L = L \times L$  の大きさの単位マトリックス  
 $L$  サンプル (チャンネル毎) の各新規ブロックに対し :

$$\mathbf{d}[k] = [y_0[kL - \Delta] \quad L \quad y_0[kL - \Delta + L - 1]]^T$$

$$\mathbf{Y}_i[k] = \text{diag}\left\{\mathbf{F}[y_i[kL - L] \quad L \quad y_i[kL + L - 1]]^T\right\}, i = M - N \dots M - 1$$

出力信号 :

$$\mathbf{e}[k] = \mathbf{d}[k] - \mathbf{k}\mathbf{F}^{-1} \sum_{j=M-N}^{M-1} \mathbf{Y}_j[k] \mathbf{W}_j[k], \quad \mathbf{E}[k] = \mathbf{F}\mathbf{k}^T \mathbf{e}[k]$$

10

発話が検出された場合 :

$$\mathbf{S}_{ij}[k] = (1 - \lambda) \sum_{l=0}^k \lambda^{k-l} \mathbf{Y}_i^{n,H}[l] \mathbf{F}\mathbf{k}^T \mathbf{k}\mathbf{F}^{-1} \mathbf{Y}_j[l] = \lambda \mathbf{S}_{ij}[k-1] + (1 - \lambda) \mathbf{Y}_i^{n,H}[k] \mathbf{F}\mathbf{k}^T \mathbf{k}\mathbf{F}^{-1} \mathbf{Y}_j[k]$$

ノイズが検出された場合 :  $\mathbf{Y}_i[k] = \mathbf{Y}_i^n[k]$

$$\mathbf{S}_{ij}^n[k] = (1 - \lambda) \sum_{l=0}^k \lambda^{k-l} \mathbf{Y}_i^{n,H}[l] \mathbf{F}\mathbf{k}^T \mathbf{k}\mathbf{F}^{-1} \mathbf{Y}_j^n[l] = \lambda \mathbf{S}_{ij}^n[k-1] + (1 - \lambda) \mathbf{Y}_i^{n,H}[k] \mathbf{F}\mathbf{k}^T \mathbf{k}\mathbf{F}^{-1} \mathbf{Y}_j^n[k]$$

20

更新のための公式 (ノイズのみの期間のみ) :

$$\mathbf{R}_i[k] = \frac{1}{\mu} \sum_{j=M-N}^{M-1} [\mathbf{S}_{ij}[k] - \mathbf{S}_{ij}^n[k]] \mathbf{W}_j[k], i = M - N \dots M - 1$$

$$\mathbf{W}_i[k+1] = \mathbf{W}_i[k] + \mathbf{F}\mathbf{g}\mathbf{F}^{-1} \mathbf{\Lambda}[k] \left\{ \mathbf{Y}_i^{n,H}[k] \mathbf{E}[k] - \mathbf{R}_i[k] \right\}, i = M - N \dots M - 1$$

ここで

$$\mathbf{\Lambda}[k] = \frac{2\rho'}{L} \text{diag}\{P_0^{-1}[k], \dots, P_{2L-1}^{-1}[k]\}$$

$$P_m[k] = \gamma P_m[k-1] + (1 - \gamma) (P_{1,m}[k] + P_{2,m}[k]), m = 0 \dots 2L - 1$$

$$P_{1,m}[k] = \sum_{j=M-N}^{M-1} |Y_{j,m}^n[k]|^2, P_{2,m}[k] = \frac{1}{\mu} \left| \sum_{j=M-N}^{M-1} S_{ij,m}[k] - S_{ij,m}^n[k] \right|, m = 0 \dots 2L - 1$$

30

## 【 0 0 9 2 】

表 2 は、Q I C - G S C を実行するための周波数領域での N L M S に基づく S P A および S P - S D W - M W F を実行するための周波数領域での確率的勾配アルゴリズム (アルゴリズム 2 およびアルゴリズム 4) の計算上の複雑性およびメモリー使用量を要約する。計算上の複雑性は、再び、秒当たりのメガ操作数 (M o p s) として表現され、一方、メモリー使用量は k ワードで表現される。以下のパラメータが使用された :  $M = 3$ 、 $L = 32$ 、 $f_s = 16 \text{ kHz}$ 、 $L_{\text{buf}1} = 10000$ 、(a)  $N = M - 1$ 、(b)  $N = M$ 。この表から、以下の結論を導き出せる :

40

- ・フィルタ  $w_0$  を有する S P - S D W - M W F (アルゴリズム 2) の計算上の複雑性は、Q I C - G S C の複雑性の約 2 倍である (そして、フィルタ  $w_0$  を使用しない場合はそれ以下である)。アルゴリズム 4 中の正規化項の近似は、さらに、計算上の複雑性を減少する。しかしながら、これは、近似は 2 次の項 0 ( $N^2$ ) を導くため、唯一、入力チャンネルの小さい数に対し正しく残る。

- ・循環発話 + ノイズバッファ  $B_1$  中のデータサンプルの記憶により、S P - S D W - M W F (アルゴリズム 2) のメモリー使用量は、Q I C - G S C (もちろんデータバッファ  $L_{\text{buf}1}$  のサイズによるが) と比較して、非常に大きくなる。アルゴリズム 4 の正規化項の近似を使用することによって、データバッファの代わりに対角相関マトリックスを記憶することとなるため、メモリー使用量を劇的に減少させることができる。しかしながら、メモ

50

リー使用量に対し、二次の項  $O(N^2)$  が存在することに注意のこと。

【表 2】

アルゴリズム	計算上の複雑性		Mops
	更新の公式	ステップサイズの適合	
NLMS based SPA	$(14M - 11 - \frac{4(M-1)}{L}) + (6M - 2) \log_2 2L \text{ MAC} + 1/L \text{ Sq} + 1/L \text{ D}$	$(2M + 2) \text{ MAC} + 1 \text{ D}$	2.16
SG with LP (Algorithm 2)	$(26N + 4 - \frac{10N}{L}) + (6N + 10) \log_2 2L \text{ MAC}$	$(4N + 6) \text{ MAC} + 1 \text{ D} + 1 \text{ Abs}$	3.22 <sup>(a)</sup> , 4.27 <sup>(b)</sup>
SG with correlation matrices (Algorithm 4)	$(10N^2 + 13N - \frac{4N^2 + 3N}{L}) + (6N + 4) \log_2 2L \text{ MAC}$	$(2N + 4) \text{ MAC} + 1 \text{ D} + 1 \text{ Abs}$	2.71 <sup>(a)</sup> , 4.31 <sup>(b)</sup>
	メモリ使用量	kワード	
NLMS based SPA	$4(M - 1)L + 6L$		0.45
SG with LP (Algorithm 2)	$2NL_{\text{buf}} + 6LN + 7L$		40.61 <sup>(a)</sup> , 60.80 <sup>(b)</sup>
SG with correlation matrices (Algorithm 4)	$4LN^2 + 6LN + 7L$		1.12 <sup>(a)</sup> , 1.95 <sup>(b)</sup>

【0093】

実質的には、アルゴリズム 2 とアルゴリズム 4 との間に何の性能の相違も存在しないことがわかり、その結果、(対角)相関マトリックスの実行を用いる SP - SDW - MWF は、なお、GSC (そして QIC - GSC) を超えるローバスト性の恩恵を維持する。同じセットアップが前述した実験においても使用される。

周波数領域での確率的勾配アルゴリズムの性能は、チャンネル毎にフィルタ長  $L = 32$ 、 $\epsilon = 0.8$ 、 $\mu = 0.95$  および  $\mu = 0.998$  に対し、評価される。考慮したすべてのアルゴリズムに対し、ノイズのみの期間においてフィルタの適合化のみ行われる。空間プリプロセッサの影響を排除するために、性能の測定は、固定ビームフォーマーの出力に対し計算される。仮定した信号モデルのエラーに対するアルゴリズムの感度は、マイクロフォン mismatches すなわち第 2 マイクロフォンでの利得 mismatches  $\epsilon_2 = 4 \text{ dB}$  に対し、記載される。

【0094】

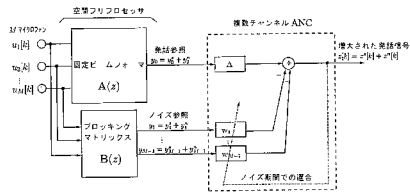
図 15 および図 16 は、トレードオフパラメータ  $1/\mu$  の関数として、アルゴリズム 2 (実線) およびアルゴリズム 4 (点線) を使用して実行した、SNR で改良した  $SNR_{\text{intellig}}$  および SP - SDW - MWF ( $w_0$  を有する) と SDR - GSC ( $w_0$  を有していない) の発話歪み  $SD_{\text{intellig}}$  を示す。これらの図は、また、第 2 マイクロフォンでの利得 mismatches  $\epsilon_2 = 4 \text{ dB}$  の影響を示す。これらの図から、周波数領域のみでの正規化項の近似は小さい性能の相違の結果となることが観察できる。多くのシナリオに対し、性能は、アルゴリズム 2 よりもアルゴリズム 4 の方がより良好である (すなわち、大きな SNR の改良と小さな発話歪み)。

【0095】

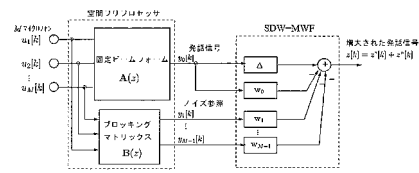
そのため、提案されたアルゴリズム 4 を使用して SP - SDW - MWF を実行するときでも、なお、GSC (および QIC - GSC) を超えるローバスト性の効果を維持する。例えば、GSC (すなわち、 $1/\mu = 0$  を有する SDR - GSC) は、マイクロフォン mismatches が起きるとき、大きな発話歪み (そして小さな SNR の改良) の結果となることが観察できる。SDR - GSC および SP - SDW - MWF の両者は、GSC にローバスト性を付与する、すなわち、歪みは  $1/\mu$  を増加させることで減少する。SP - SDW - MWF ( $w_0$  を有する) の性能は、繰り返しになるが、マイクロフォン mismatches によって

ほとんど影響を受けない。

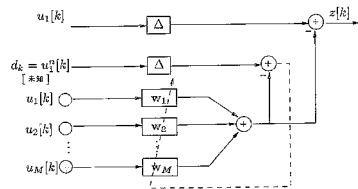
【図1】



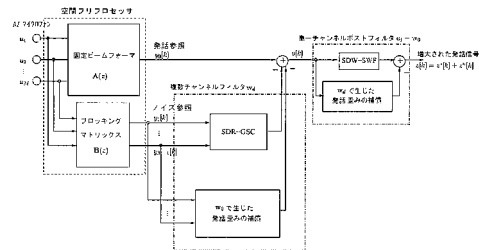
【図3】



【図2】

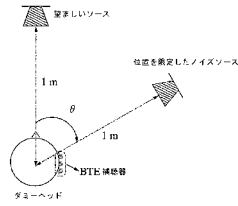


【図4】

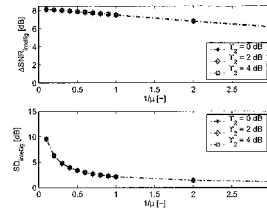




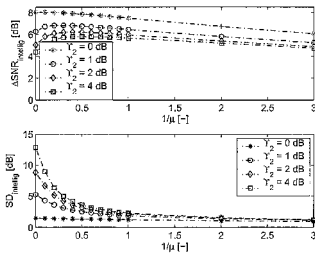
【図5】



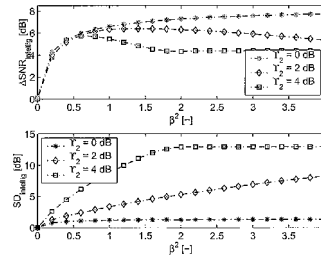
【図7】



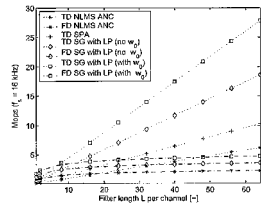
【図6】



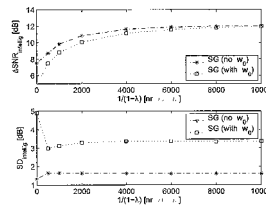
【図8】



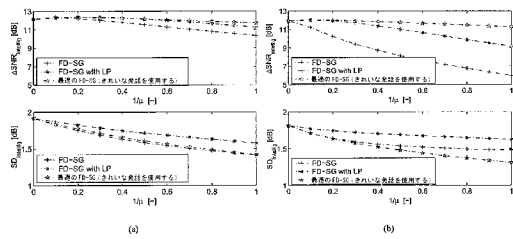
【図9】



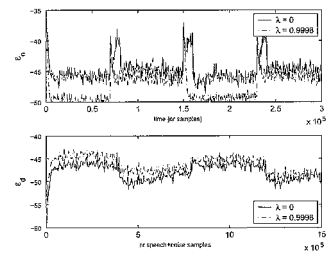
【図11】



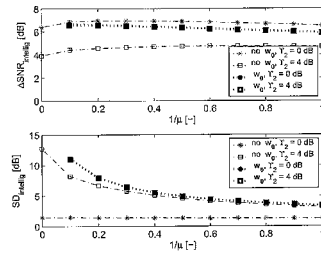
【図10】



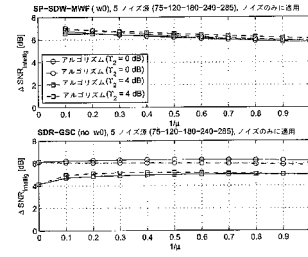
【図12】



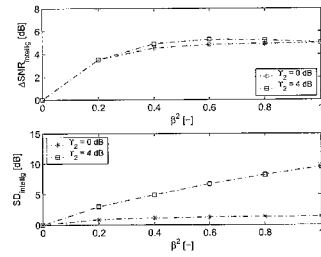
【図13】



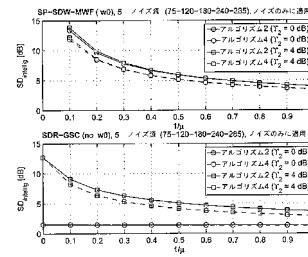
【図15】



【図14】



【図16】



---

フロントページの続き

- (72)発明者 ムーネン, マルク  
ベルギー王国 ヘレント/ウインクセレ ビィ - 3 0 2 0 , タルメルデルラン 7
- (72)発明者 ボウタース, ヤン  
ベルギー王国 ホルスピーク ビィ - 3 3 2 0 , レミング 1 2 5

審査官 大野 弘

- (56)参考文献 特開平08 - 1 2 2 4 2 4 ( J P , A )  
特表2003 - 5 2 7 0 1 2 ( J P , A )  
W . H . Neo and B . Farhang - Boroujeny , Robust micr  
ophone arrays using subband adaptive filter  
s , Vision , Image and Signal Processing , IEE Pr  
oceedings , 2002年12月 , vol . 1 4 9 , no . 1 , 1 7 - 2 5

(58)調査した分野(Int.Cl. , DB名)

H04R 3/00  
G10L 21/02  
H04R 1/40  
H04R 25/00