

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4476355号
(P4476355)

(45) 発行日 平成22年6月9日(2010.6.9)

(24) 登録日 平成22年3月19日(2010.3.19)

(51) Int. Cl.		F I	
HO4R	3/02	(2006.01)	HO4R 3/02
HO4B	3/23	(2006.01)	HO4B 3/23
HO4M	1/58	(2006.01)	HO4M 1/58

請求項の数 30 (全 21 頁)

(21) 出願番号	特願2009-509908 (P2009-509908)	(73) 特許権者	395015319
(86) (22) 出願日	平成19年3月30日(2007.3.30)		株式会社ソニー・コンピュータエンタテインメント
(65) 公表番号	特表2009-535996 (P2009-535996A)		東京都港区南青山二丁目6番21号
(43) 公表日	平成21年10月1日(2009.10.1)	(74) 代理人	100105924
(86) 国際出願番号	PCT/US2007/065686		弁理士 森下 賢樹
(87) 国際公開番号	W02007/130765	(72) 発明者	シャドン マオ
(87) 国際公開日	平成19年11月15日(2007.11.15)		アメリカ合衆国、カリフォルニア州 94404、フォスター・シティー、セカンド・フロアー、イースト・ヒルスデイル・ブルバード 919
審査請求日	平成20年12月26日(2008.12.26)		
(31) 優先権主張番号	11/381,728	審査官	境 周一
(32) 優先日	平成18年5月4日(2006.5.4)		
(33) 優先権主張国	米国 (US)		
(31) 優先権主張番号	11/381,729		
(32) 優先日	平成18年5月4日(2006.5.4)		
(33) 優先権主張国	米国 (US)		

最終頁に続く

(54) 【発明の名称】 エコー及びノイズキャンセレーション

(57) 【特許請求の範囲】

【請求項1】

スピーカ信号 $x(t)$ を受け取るスピーカと、ローカル信号 $s(t)$ とエコー信号 $x_1(t)$ とを含むマイクロフォン信号 $d(t)$ を受け取るマイクロフォンとを有するシステムにおけるエコーキャンセレーション方法であって、エコー信号 $x_1(t)$ は、スピーカ信号 $x(t)$ に依存しており、

前記マイクロフォン信号 $d(t)$ を、第2適応フィルタに対して相補的なキャンセレーション特性を有する第1適応フィルタと、前記第2適応フィルタとで並行にフィルタリング処理するステップとを備え、

前記第1および第2適応フィルタのフィルタリング特性は、該第1および第2適応フィルタのうち的一方がその入力にうまく適応していないときに、他方はその入力にうまく適応するような特性であり、

前記第1および第2適応フィルタのうち的一方がうまく適応するのは、そのフィルタ関数 $h(n)$ が安定しており、真のエコーパスフィルタに収束し、過大予測も過小予測もされていないときであり、

前記第1適応フィルタは、音声アクティビティ検出フィルタであり、

前記第2適応フィルタは、相互相関解析フィルタであり、

本方法はさらに、

前記第1適応フィルタからの出力 $e_1(t)$ と、前記第2適応フィルタからの出力 $e_2(t)$ とから、前記スピーカ信号 $x(t)$ との相関がより小さい最小エコー出力 $e_3(t)$

10

20

)を決定するステップと、

前記最小エコー出力 $e_3(t)$ を用いてマイクロフォン出力を生成するステップと、
を備える方法。

【請求項2】

前記マイクロフォン信号 $d(t)$ を前記第1適応フィルタと前記第2適応フィルタとで
並行にフィルタリング処理するステップは、

$x(t)$ の強度が閾値を超えた場合に、前記第1適応フィルタのフィルタ係数の組を適
応させるステップと、

前記第2適応フィルタで、 $e_2(t)$ と $x(t)$ の相互相関を解析するステップを含む
請求項1に記載の方法。

10

【請求項3】

前記最小エコー出力 $e_3(t)$ を決定するステップは、前記第2適応フィルタが前記ロー
カル信号 $s(t)$ をフィルタリング処理により過度に除去していないか決定するステッ
プと、

前記第2適応フィルタが前記ローカル信号 $s(t)$ をフィルタリング処理により過度に
除去している場合に、前記第1適応フィルタの出力を前記最小エコー出力として用いるス
テップとを含む請求項2に記載の方法。

【請求項4】

前記第2適応フィルタが前記ローカル信号 $s(t)$ をフィルタリング処理により過度に
除去していないか決定するステップは、

前記第2適応フィルタの出力 $e_2(t)$ と、前記スピーカ信号 $x(t)$ との間の相互相
関をとるステップと、

前記第2適応フィルタの出力 $e_2(t)$ と前記スピーカ信号 $x(t)$ との間の相互相関
の期待値が所定の閾値未満であるか判定するステップと、

前記相互相関の期待値が前記閾値未満である場合に、前記第2適応フィルタが前記ロー
カル信号 $s(t)$ をフィルタリング処理により過度に除去していることを決定するステッ
プとを含む請求項3に記載の方法。

20

【請求項5】

前記第2適応フィルタが前記ローカル信号 $s(t)$ をフィルタリング処理により過度に
除去している場合に、前記第2適応フィルタのフィルタ係数の組を前記第1適応フィルタ
のフィルタ係数の組で置き換えるステップをさらに含む請求項3に記載の方法。

30

【請求項6】

前記最小エコー出力 $e_3(t)$ を用いて前記マイクロフォン出力を生成するステップは
、前記最小エコー出力 $e_3(t)$ と前記スピーカ信号 $x(t)$ との間の相互相関の解析を
含む第1残差予測 $ER_1(t)$ を並行に決定するステップと、前記最小エコー出力 $e_3(t)$
 と前記スピーカ信号 $x(t)$ との間のエコー距離 mismatch を含む第2残差予測 $ER_2(t)$
 を決定するステップとを含む請求項1に記載の方法。

【請求項7】

前記相互相関解析は、 $e_3(t)$ と $x(t)$ との間の相互相関の期待値を計算するステ
ップを含む請求項6に記載の方法。

40

【請求項8】

前記エコー距離 mismatch は、 $\{e_3^2(t) / x^2(t)\}$ の期待値を計算するステ
ップを含む請求項6に記載の方法。

【請求項9】

前記第1残差予測 $ER_1(t)$ を決定するステップは、第1残差予測部 $ER(1)$ の2
次ノルム $N(1)$ を計算するステップを含み、

前記第2残差予測 $ER_2(t)$ を決定するステップは、第2残差予測部 $ER(2)$ の2
次ノルム $N(2)$ を計算するステップを含み、

本方法はさらに、ダブルトークの間、 $e_3(t)$ に、対応するノルム $N(1)$ または $N(2)$
 がより小さい前記エコー残差予測部 $ER(1)$ または $ER(2)$ を適用するステッ

50

プを含む請求項 6 に記載の方法。

【請求項 10】

前記第 1 残差予測 $ER_1(t)$ を決定するステップは、第 1 残差予測部 $ER(1)$ の 2 次ノルム $N(1)$ を計算するステップを含み、

前記第 2 残差予測 $ER_2(t)$ を決定するステップは、第 2 残差予測部 $ER(2)$ の 2 次ノルム $N(2)$ を計算するステップを含み、

本方法はさらに、シングルトークの間、 $e_3(t)$ に、対応するノルム $N(1)$ または $N(2)$ がより大きい前記エコー残差予測部 $ER(1)$ または $ER(2)$ を適用するステップを含む請求項 6 に記載の方法。

【請求項 11】

前記最小エコー出力 $e_3(t)$ を用いてマイクロフォン出力を生成するステップはさらに、最小残差エコー予測 $ER_3(t)$ を決定するステップを含み、

前記最小残差エコー予測 $ER_3(t)$ は、 $ER_1(t)$ と $ER_2(t)$ のうちの一で、エネルギーが最小であり $x(t)$ との相関が最小である請求項 6 に記載の方法。

【請求項 12】

前記最小エコー出力 $e_3(t)$ から $ER_3(t)$ を選択的に除去するステップをさらに含む請求項 11 に記載の方法。

【請求項 13】

前記マイクロフォン出力から $ER_3(t)$ を選択的に除去するステップは、遠端通話のみの期間に、前記マイクロフォン出力から $ER_3(t)$ を最大限に除去するステップを含む請求項 12 に記載の方法。

【請求項 14】

前記マイクロフォン出力から $ER_3(t)$ を選択的に除去するステップは、前記ローカル信号 $s(t)$ の強度が 0 でないときに、前記マイクロフォン出力から $ER_3(t)$ を最小限に除去するステップを含む請求項 12 に記載の方法。

【請求項 15】

前記マイクロフォン信号 $d(t)$ から予測ノイズ信号 $n'(t)$ を計算するステップと、
前記予測ノイズ信号 $n'(t)$ のレベルを低減させて、低減ノイズ信号 $n''(t)$ を形成するステップと、

前記マイクロフォン出力信号に、前記低減ノイズ信号 $n''(t)$ を組み込むステップとをさらに含む請求項 1 に記載の方法。

【請求項 16】

スピーカとマイクロフォンを有するシステムにおいて用いられるエコーキャンセレーション装置であって、スピーカは、スピーカ信号 $x(t)$ を受けるように適合され、マイクロフォンは、ローカル信号 $s(t)$ とエコー信号 $x_1(t)$ とを含むマイクロフォン信号 $d(t)$ を受け取るように適合され、エコー信号 $x_1(t)$ は、スピーカ信号 $x(t)$ に依存しており、

本装置は、

前記スピーカと前記マイクロフォンに接続された第 1 適応フィルタと、

前記第 1 適応フィルタと並行に前記スピーカと前記マイクロフォンに接続された第 2 適応フィルタとを備え、

前記第 2 適応フィルタは、前記第 1 適応フィルタに対して相補的なエコーキャンセレーション特性を有し、

前記第 1 および第 2 適応フィルタのフィルタリング特性は、該第 1 および第 2 適応フィルタのうち的一方がその入力にうまく適応していないときに、他方はうまく適応するような特性であり、

前記第 1 および第 2 適応フィルタのうち的一方がうまく適応するのは、そのフィルタ関数 $h(n)$ が安定しており、真のエコーパスフィルタに収束し、過大予測も過小予測もされていないときであり、

10

20

30

40

50

前記第 1 適応フィルタは、音声アクティビティ検出フィルタであり、

前記第 2 適応フィルタは、相互相関解析フィルタであり、

本装置はさらに、

前記第 1 適応フィルタと前記第 2 適応フィルタに接続されたインテグレータとを備え、

前記インテグレータは、前記第 1 適応フィルタからの出力 $e_1(t)$ と、前記第 2 適応フィルタからの出力 $e_2(t)$ から、最小エコー出力 $e_3(t)$ を決定するように構成されており、

前記最小エコー出力 $e_3(t)$ と前記スピーカ信号 $x(t)$ との相関はより小さい装置。

【請求項 17】

10

前記インテグレータは、前記第 2 エコーキャンセレーションフィルタが前記ローカル信号 $s(t)$ をフィルタリング処理により過度に除去していないか決定するように構成され、前記第 2 適応フィルタが前記ローカル信号 $s(t)$ をフィルタリング処理により過度に除去している場合に、インテグレータは、最小エコー出力 $e_3(t)$ として前記出力 $e_1(t)$ を選択する請求項 16 に記載の装置。

【請求項 18】

前記インテグレータに接続された第 1 エコー残差予測部 $ER(1)$ と、

前記インテグレータに接続された第 2 エコー残差予測部 $ER(2)$ とをさらに備える請求項 16 に記載の装置。

【請求項 19】

20

前記第 1 エコー残差予測部 $ER(1)$ は、前記最小エコー出力 $e_3(t)$ と前記スピーカ信号 $x(t)$ との相互相関の解析を含む第 1 残差予測 $ER_1(t)$ を生成するように構成され、

前記第 2 エコー残差予測部 $ER(2)$ は、前記最小エコー出力 $e_3(t)$ と前記スピーカ信号 $x(t)$ との間のエコー距離 mismatch を含む第 2 残差予測 $ER_2(t)$ を決定するように構成されている請求項 18 に記載の装置。

【請求項 20】

前記相互相関解析は、 $e_3(t)$ と $x(t)$ との相互相関の期待値の計算を含む請求項 19 に記載の装置。

【請求項 21】

30

前記エコー距離 mismatch は、 $\{e_3^2(t) / x^2(t)\}$ の期待値の計算を含む請求項 19 に記載の装置。

【請求項 22】

前記第 1 および第 2 エコー残差予測部に接続された残差エコーキャンセレーションモジュールをさらに備える装置であって、

前記残差エコーキャンセレーションモジュールは、

第 1 残差予測部 $ER(1)$ の 2 次ノルム $N(1)$ を計算し、第 2 残差予測 $ER(2)$ の 2 次ノルム $N(2)$ を計算するように構成され、

前記残差エコーキャンセレーションモジュールは、ダブルトークの間、対応するノルム $N(1)$ または $N(2)$ がより小さい前記エコー残差予測部 $ER(1)$ または $ER(2)$ を、 $e_3(t)$ に適用するように構成されている請求項 19 に記載の装置。

40

【請求項 23】

前記第 1 および第 2 エコー残差予測部に接続された残差エコーキャンセレーションモジュールをさらに備える装置であって、

前記残差エコーキャンセレーションモジュールは、

第 1 残差予測部 $ER(1)$ の 2 次ノルム $N(1)$ を計算し、第 2 残差予測 $ER(2)$ の 2 次ノルム $N(2)$ を計算するように構成され、

前記残差エコーキャンセレーションモジュールは、シングルトークの間、対応するノルム $N(1)$ または $N(2)$ がより大きい前記エコー残差予測部 $ER(1)$ または $ER(2)$ を、 $e_3(t)$ に適用するように構成されている請求項 19 に記載の装置。

50

【請求項 2 4】

前記第 1 および第 2 エコー残差予測部に接続された残差エコーキャンセレーションモジュールをさらに備える装置であって、

前記残差エコーキャンセレーションモジュールは、最小残差エコー予測 $ER_3(t)$ を決定するように構成されており、

前記最小残差エコー予測 $ER_3(t)$ は、 $ER_1(t)$ と $ER_2(t)$ のうちの一で、エネルギーが最小であり $x(t)$ との相関が最小である請求項 1 9 に記載の装置。

【請求項 2 5】

前記残差エコーキャンセレーションモジュールは、前記最小エコー出力 $e_3(t)$ から $ER_3(t)$ を選択的に除去するように構成されている請求項 2 4 に記載の装置。

10

【請求項 2 6】

前記残差エコーキャンセレーションモジュールは、遠端通話のみの期間に、前記マイクロフォン出力から $ER_3(t)$ を最大限に除去するように構成されている請求項 2 5 に記載の装置。

【請求項 2 7】

前記残差エコーキャンセレーションモジュールは、前記ローカル信号 $s(t)$ の強度が 0 でない場合に、前記マイクロフォン出力から $ER_3(t)$ を最小限に除去するように構成されている請求項 2 5 に記載の装置。

【請求項 2 8】

前記マイクロフォンに接続されたノイズキャンセレーションモジュールをさらに備え、
前記ノイズキャンセレーションモジュールは、
マイクロフォン信号 $d(t)$ から予測ノイズ信号 $n'(t)$ を計算し、
予測ノイズ信号 $n'(t)$ レベルを減少させて、低減ノイズ信号 $n''(t)$ を形成し、
マイクロフォン出力信号に、その低減ノイズ信号 $n''(t)$ を取り入れるように構成されている請求項 1 6 に記載の装置。

20

【請求項 2 9】

マイクロフォンと、

スピーカと、

前記マイクロフォンとスピーカとに接続されているプロセッサと、

前記プロセッサに接続されているメモリとを備える音響信号処理システムであって、

30

前記メモリは、

スピーカ信号 $x(t)$ を受け取るスピーカと、

ローカル信号 $s(t)$ と、前記スピーカ信号 $x(t)$ に依存するエコー信号 $x_1(t)$ とを含むマイクロフォン信号 $d(t)$ を受け取るマイクロフォンを有するシステムにおけるエコーキャンセレーション方法を実装するためのプロセッサ可読命令の組を格納し、

前記プロセッサ可読命令は、

前記マイクロフォン信号 $d(t)$ を、第 2 適応フィルタに対して相補的なキャンセレーション特性を有する第 1 適応フィルタと、前記第 2 適応フィルタとで並行にフィルタリング処理するための命令を含み、

前記第 1 および第 2 適応フィルタのフィルタリング特性は、該第 1 および第 2 適応フィルタのうち的一方がその入力にうまく適応していないときに、他方はうまく適応するような特性であり、

40

前記第 1 および第 2 適応フィルタのうち的一方がうまく適応するのは、そのフィルタ関数 $h(n)$ が安定しており、真のエコーパスフィルタに収束し、過大予測も過小予測もされていないときであり、

前記第 1 適応フィルタは、音声アクティビティ検出フィルタであり、

前記第 2 適応フィルタは、相互相関解析フィルタであり、

前記プロセッサ可読命令はさらに、

前記第 1 適応フィルタからの出力 $e_1(t)$ と前記第 2 適応フィルタからの出力 $e_2(t)$ から、スピーカ信号 $x(t)$ との相関がより少ない最小エコー出力 $e_3(t)$ を決定

50

するための命令と、

前記最小エコー出力 $e_3(t)$ を用いて、マイクロフォン出力を生成するための命令と

を含む音響信号処理システム。

【請求項 30】

プロセッサに接続されているメモリを備えるプロセッサ可読媒体であって、

前記メモリは、

スピーカ信号 $x(t)$ を受け取るスピーカと、

ローカル信号 $s(t)$ と、スピーカ信号 $x(t)$ に依存するエコー信号 $x_1(t)$ とを含むマイクロフォン信号 $d(t)$ を受け取るマイクロフォンを有するシステムにおけるエコーキャンセレーション方法を実装するためのプロセッサ可読命令の組を格納し、

前記プロセッサ可読命令は、

前記マイクロフォン信号 $d(t)$ を、第 2 適応フィルタに対して相補的なキャンセレーション特性を有する第 1 適応フィルタと、前記第 2 適応フィルタとで並行にフィルタリング処理するための命令を含み、

前記第 1 および第 2 適応フィルタのフィルタリング特性は、該第 1 および第 2 適応フィルタのうち的一方がその入力にうまく適応していないときに、他方はうまく適応するような特性であり、

前記第 1 および第 2 適応フィルタのうち的一方がうまく適応するのは、そのフィルタ関数 $h(n)$ が安定しており、真のエコーパスフィルタに収束し、過大予測も過小予測もされていないときであり、

前記第 1 適応フィルタは、音声アクティビティ検出フィルタであり、

前記第 2 適応フィルタは、相互相関解析フィルタであり、

前記プロセッサ可読命令はさらに、

前記第 1 適応フィルタからの出力 $e_1(t)$ と前記第 2 適応フィルタからの出力 $e_2(t)$ から、スピーカ信号 $x(t)$ との相関がより小さい最小エコー出力 $e_3(t)$ を決定するための命令と、

前記最小エコー出力 $e_3(t)$ を用いてマイクロフォン出力を生成するための命令と、

を含む媒体。

【発明の詳細な説明】

【技術分野】

【0001】

[優先権の主張]

本出願は、本出願と譲受人が共通であって本出願と同時に係属する特許文献 1 の恩恵を主張し、その開示内容全体をここに援用する。本出願は、本出願と譲受人が共通であって本出願と同時に係属する特許文献 2 の恩恵を主張し、その開示内容全体をここに援用する。本出願はまた、本出願と譲受人が共通であって本出願と同時に係属する特許文献 3 の恩恵を主張し、その開示内容全体をここに援用する。本出願はまた、本出願と譲受人が共通であって本出願と同時に係属する特許文献 4 の恩恵を主張し、その開示内容全体をここに援用する。本出願はまた、本出願と譲受人が共通であって本出願と同時に係属する特許文献 5 の恩恵を主張し、その開示内容全体をここに援用する。本出願はまた、本出願と譲受人が共通であって本出願と同時に係属する特許文献 6 の恩恵を主張し、その開示内容全体をここに援用する。本出願はまた、本出願と譲受人が共通であって本出願と同時に係属する特許文献 7 の恩恵を主張し、その開示内容全体をここに援用する。本出願はまた、本出願と譲受人が共通であって本出願と同時に係属する特許文献 8 の恩恵を主張し、その開示内容全体をここに援用する。本出願はまた、本出願と譲受人が共通であって本出願と同時に係属する特許文献 9 の恩恵を主張し、その開示内容全体をここに援用する。本出願はまた、本出願と譲受人が共通であって本出願と同時に係属する特許文献 10 の恩恵を主張し、その開示内容全体をここに援用する。

【特許文献 1】米国特許出願 第 11/381,728 号, シャドン マオ, "ECHO AND NOISE CANCE

10

20

30

40

50

LATION", 2006年5月4日出願, (代理人整理番号SCEA05064US00)

【特許文献2】米国特許出願 第11/381,729号, シャドン マオ, "ULTRA SMALL MICROPHONE ARRAY", 2006年5月4日出願, (代理人整理番号SCEA05062US00)

【特許文献3】米国特許出願 第11/381,725号, シャドン マオ, "METHODS AND APPARATUS FOR TARGETED SOUND DETECTION", 2006年5月4日出願, (代理人整理番号SCEA05072US00)

【特許文献4】米国特許出願 第11/381,727号, シャドン マオ, "NOISE REMOVAL FOR ELECTRONIC DEVICE WITH FAR FIELD MICROPHONE ON CONSOLE", 2006年5月4日出願, (代理人整理番号SCEA05073US00)

【特許文献5】米国特許出願 第11/381,724号, シャドン マオ, "METHODS AND APPARATUS FOR TARGETED SOUND DETECTION AND CHARACTERIZATION", 2006年5月4日出願, (代理人整理番号SCEA05079US00) 10

【特許文献6】米国特許出願 第11/381,721号, シャドン マオ, "SELECTIVE SOUND SOURCE LISTENING IN CONJUNCTION WITH COMPUTER INTERACTIVE PROCESSING", 2006年5月4日出願, (代理人整理番号SCEA04005 JUMBOUS)

【特許文献7】PCT出願 PCT/US06/17483号, シャドン マオ, "SELECTIVE SOUND SOURCE LISTENING IN CONJUNCTION WITH COMPUTER INTERACTIVE PROCESSING", 2006年5月4日出願, (代理人整理番号SCEA04005 JUMBOPCT)

【特許文献8】米国特許出願 第11/418,988号, シャドン マオ, "METHODS AND APPARATUS FOR ADJUSTING A LISTENING AREA FOR CAPTURING SOUNDS", 2006年5月4日出願, (代理人整理番号SCEA-00300) 20

【特許文献9】米国特許出願 第11/418,989号, シャドン マオ, "METHODS AND APPARATUS FOR CAPTURING AN AUDIO SIGNAL BASED ON VISUAL IMAGE", 2006年5月4日出願, (代理人整理番号SCEA-00400)

【特許文献10】米国特許出願 第11/429,047号, シャドン マオ, "METHODS AND APPARATUS FOR CAPTURING AN AUDIO SIGNAL BASED ON A LOCATION OF THE SIGNAL", 2006年5月4日出願, (代理人整理番号SCEA-00500)

【0002】

[本発明の技術分野]

本発明は、音響信号処理に関し、とくに、音響信号処理におけるエコーおよびキャンセリングに関する。 30

【背景技術】

【0003】

インタラクティブテレビゲームコントローラなどのような多くの携帯電子装置は、双方向音響信号を扱うことができる。このような装置は、典型的にはその装置のユーザからのローカルスピーチ信号 $s(t)$ を受けるマイクロフォンと、ユーザが聞くことができるスピーカ信号 $x(t)$ を発信するスピーカとを備える。テレビゲームコントローラをより小型化するために、マイクロフォンとスピーカは、比較的近く(例えば20cm以内など)に設置することが望ましい。これに対してユーザは、マイクロフォンからより離れたところ(例えば3メートルから5メートルなど)に位置するかもしれない。マイクロフォンはローカルスピーチ信号 $s(t)$ とスピーカエコー信号 $x_1(t)$ との両方を含む信号 $d(t)$ を生成する。これに加えて、マイクロフォンはバックグラウンドノイズ $n(t)$ を受けるかもしれない。そのため、全体のマイクロフォン信号は、 $d(t) = s(t) + x_1(t) + n(t)$ となる。比較的スピーカの近傍にあるため、マイクロフォン信号 $d(t)$ は、スピーカエコー信号 $x_1(t)$ によって、占められるかもしれない。 40

【0004】

電気通信の応用例において、スピーカエコーは広くみられる現象であり、エコーサプレッションとエコーキャンセレーションは比較的成熟した手法である。エコーサプレッサは、回線において1方向に向かう音声信号の存在を検出した場合に作動し、他の方向に大きな損失を挿入する。通常、回線の遠端にあるエコーサプレッサが回線の近端からの音声を 50

検出した場合に、そのエコーサプレッサがこの損失を加える。この加えられた損失により、スピーカ信号 $x(t)$ が、ローカルスピーチ信号 $d(t)$ へと再送されることを阻止することができる。

【発明の開示】

【発明が解決しようとする課題】

【0005】

エコーサプレッションは効果的ではあるが、多くの場合、いくつかの問題につながる。例えば、ローカルスピーチ信号 $s(t)$ とリモートスピーカ信号 $x(t)$ は、少なくとも短時間に限れば、同時に生ずることがよくある。この状況はダブルトークとも呼ばれる。リモートスピーカ信号のみが存在するような状況は、リモートシングルトークとも呼ばれる。各エコーサプレッサが回路の遠端 (far-end) からの音声エネルギーを検出するため、その結果、通常、同時に双方向に損失が挿入されることとなり、両側の通話がブロックされる。これを防止するため、エコーサプレッサを、近端のスピーカからの音声アクティビティのみを検出するように設定することができる。これにより、近端話者と遠端話者が同時に話しているときには、損失が挿入されなくなる (または、より小さい損失のみ挿入される)。残念ながら、これは、当初のエコーサプレッサの効果まで、一時的にうち消してしまう。

【0006】

さらに、エコーサプレッサは、交互に、損失を挿入し、除去するため、新たな話者が話し始めたときにしばしば小さな遅延が生じ、その話者のスピーチの初めの方の音がクリッピングされてしまう。さらに、遠端の相手方の周囲がうるさいときには、遠端話者が話しているときには、近端話者にそのバックグラウンド音が聞こえるが、近端話者が話し始めるとエコーサプレッサがそのバックグラウンド音を抑制する。これにより、バックグラウンド音が突然無くなるため、近端のユーザは回線が切れたかのような印象を受けることになる。

【0007】

上述の問題に対処するため、エコーキャンセレーション手法が開発された。エコーキャンセレーションは、アナログまたはデジタルフィルタを用いて、望ましくないノイズやエコーを入力信号から取り除き、フィルタリング処理された信号 $e(t)$ を生成する。エコーキャンセレーションにおいては、スピーチモデルを計算するために複雑なアルゴリズム手順が用いられる。この手順は、マイクロフォン信号 $d(t)$ と、リモート信号 $x(t)$ の一部を、エコーキャンセレーションプロセッサに入力するステップと、スピーカエコー信号 $x_1(t)$ を予測するステップと、そしてこれをマイクロフォン信号 $d(t)$ から差し引くステップとを含む。エコー予測方式は、適用 (adaptation) として知られるプロセスにおいて、エコーキャンセレーションプロセッサにより学習されなければならない。

【0008】

このような手法の効果は、エコー抑制比 (ESR: echo suppression ratio) によって測定される。これは単に、マイクロフォンが受ける真のエコーエネルギーと、フィルタリング処理された信号 $x_1(t)$ に残る残余エコーエネルギーとの比である (典型的にはデシベルで表される)。国際電気通信ユニオン (ITU) が定めた基準によると、リモートシングルトークの場合、エコーレベルについて、少なくとも 45 デシベルの減衰が必要である。ダブルトークの最中 (または強いバックグラウンドノイズの最中) には、この減衰レベルは 30 デシベルまで低くなってもよい。しかしながら、これらの推奨基準は、ローカルスピーチ信号を発生するユーザが、マイクロフォンに、より近いようなシステムにおいて開発されたものである。したがって録音された SN 比 (ターゲット音声エネルギーのエコーノイズエネルギーに対する比) は、大抵、5 デシベルよりも良い。例えばテレビゲームコントローラのような、ユーザが 3 メートルから 5 メートルも離れており、オープンマイクロフォンから 0.5 メートルよりも近傍にあるラウドスピーカが大きなエコーを発生するようなアプリケーションにおいては、これらの推奨基準はあてはまらない。このようなアプリ

10

20

30

40

50

ケーションにおいては、SN比は - 15 デシベルから - 30 デシベル未満であろう。リモートシングルトークにおいては60デシベル以上のESR、ダブルトークについては35デシベル以上ESRが要求されるかもしれない。既存のエコーキャンセレーション手法ではこのような高いESRレベルを達成することができない。

【0009】

したがって、当該技術分野においては前述の不利な点を克服するエコーキャンセレーションシステムおよび方法が必要とされている。

【課題を解決するための手段】

【0010】

[発明の概要]

前述の不利な点を克服するため、本発明の実施形態は、スピーカとマイクロフォンを有するシステムにおけるエコーキャンセレーション方法および装置に照準を合わせる。スピーカはスピーカ信号 $x(t)$ を受信する。マイクロフォンは、ローカル信号 $s(t)$ とエコー信号 $x_1(t)$ を含むマイクロフォン信号 $d(t)$ を受け取る。エコー信号 $x_1(t)$ は、スピーカ信号 $x(t)$ に依存する。マイクロフォン信号 $d(t)$ は、互いに相補的なエコーキャンセレーション特性を有する第1適応フィルタ、および第2適応フィルタによって、パラレルにフィルタリング処理される。最小エコー出力 $e_3(t)$ は、第1適応フィルタからの出力 $e_1(t)$ と、第2適応フィルタからの出力 $e_2(t)$ が決定される。最小エコー出力のエネルギーはより小さく、最小エコー出力とスピーカ信号 $x(t)$ との間の相関はより小さい。そして、マイクロフォン出力が、最小エコー出力 $e_3(t)$ を用いて生成される。オプションとして、残差エコーキャンセレーション、かつ/または、ノイズキャンセレーションが、最小エコー出力に適用されてもよい。

【図面の簡単な説明】

【0011】

【図1A】本発明の実施形態にかかるエコーキャンセレーション装置の概略図である。

【図1B】図1Aのエコーキャンセレーション装置において用いられうる音声アクティビティ検出適応フィルタの概略図である。

【図1C】図1Aのエコーキャンセレーション装置において用いられうる相互相関解析を伴う適応フィルタの概要図である。

【図2A】本発明の実施形態にかかるエコーキャンセレーション方法を説明するフローチャートである。

【図2B】本発明の実施形態にかかるエコーキャンセレーションのための別の方法を説明するフローチャートである。

【図3】本発明の別の実施形態にかかるエコーキャンセレーション装置の概略図である。

【実施例】

【0012】

[具体的な実施形態の説明]

以下の詳細な説明は、説明の目的のため、具体的な細部を含むが、本発明の範囲内において、後述の細部について多くの変形や変更が可能であることは、当該技術分野において通常の知識を有する者に理解されるところである。したがって、以下に記述される本発明の実施例の説明により、特許請求の範囲に記載されている発明が一般性を失うことなく、また、以下の説明は、特許請求の範囲に記載されている発明について制限を課すものではない。

【0013】

本発明の実施形態によると、機能的に同一である二つのフィルタを有する一体型のエコーおよびノイズキャンセラの新しい構成が提案される。これらのフィルタは、直交制御と表現(orthogonal controls and representations)を伴う。このような構成においては、雑音のあるハンドフリー音声通信において、システム全体のロバスト性(robustness)を引き上げるように、二つの直交フィルタは互いに補完し合う。

10

20

30

40

50

【 0 0 1 4 】

特に、一体型のエコーノイズキャンセラは、別個に制御される二つのサブシステムを並行に用いる。これらのサブシステムはそれぞれ、直行制御メカニズムを伴う。エコーノイズキャンセラは、フロント・エコーキャンセラと、バックアップ・エコーキャンセラとを含む。フロント・エコーキャンセラは、ダブルトーク検出を用いる。ローカル音声に対して確実にロバストであるようにするために、フロント・エコーキャンセラは、保守的な適応アプローチをとりながらも、提供するエコーサプレッションはより小さく、スピーチ、エコーの変化への適応は遅い。バックアップ・エコーキャンセラは、相互相関を用いて、エラー信号とエコー信号との間の類似性を測定する。バックアップ・エコーキャンセラは、フィルタが迅速に更新されるように、積極的な戦略をとる。バックアップ・エコーキャンセラは、大きなエコーサプレッションを提供しながらも、過剰に適応してしまう可能性があるため、ローカル音声/ノイズに対して不安定である。これらの二つのエコーキャンセラの出力の統合は、どちらのエコーキャンセラとエコー信号との差が大きいかを測定する相互相関解析に基づいて実行される。この統合においてはまた、両方のエコーキャンセラのフィルタ安定性がチェックされる。一のフィルタが過大予測または過小予測されている場合、そのフィルタは他方のフィルタによって補完される。このようなシステムは、いかなるときでも確実に一のフィルタが正しく動作するように設計される。

10

【 0 0 1 5 】

本システムはオプションで、同様のアプローチをとるエコー残差ノイズ予測部を含んでもよい。エコー残差ノイズ予測部は、直行制御を伴う二つの独立なサブ予測部を並行に用いる。第1予測部は、ロバストなダブルトーク検出部に依存するエコー距離ミスマッチ (e c h o - d i s t a n c e - m i s m a t c h) に基づく。第1予測部は、比較的正確でありながら、ダブルトーク検出エラーのために不安定である。第2予測部は相互スペクトル解析 (c r o s s - s p e c t r u m - a n a l y s i s) に基づく。第2予測部の予測にはバイアスがかかっているが安定であり、ローカル音声検出に依存せず、一貫性がある。これらの二つの残差エコーの予測の統合においては、遠端通話のみの場合、またはダブルトークの場合にそれぞれ、最小/最大アプローチがとられる。

20

【 0 0 1 6 】

図1Aは、本発明の一実施形態にかかるエコーキャンセレーション装置100を用いたオーディオシステム99を示す図である。装置100の動作は、図2Aに示される方法200のフローチャート、および図2Bに示される方法220を参照することによって理解されるであろう。オーディオシステム99は一般的に、リモート信号 $x(t)$ を受け取るスピーカ102とマイクロフォン104とを含む。ローカル音源101は、ローカルスピーチ信号 $s(t)$ を発する。マイクロフォン104は、ローカルスピーチ信号 $s(t)$ と、スピーカ信号 $x(t)$ に関連するエコー信号 $x_1(t)$ の両方を受け取る。マイクロフォン104はまた、マイクロフォン104が位置する環境から発生するノイズ $n(t)$ をも受け取る。そして、マイクロフォン104は、マイクロフォン信号 $d(t)$ を生成する。マイクロフォン信号 $d(t)$ は、 $d(t) = s(t) + x_1(t) + n(t)$ によって与えられるだろう。

30

【 0 0 1 7 】

エコーキャンセレーション装置100は、一般的に、第1適応エコーキャンセレーションフィルタ $EC(1)$ と第2適応エコーキャンセレーションフィルタ $EC(2)$ とを含む。それぞれの適応フィルタは、マイクロフォン信号 $d(t)$ とスピーカ信号 $x(t)$ とを受け取る。図2A-2Bに示されるように、フィルタ $EC(1)$ はステップ202に示されるようにマイクロフォン信号 $d(t)$ を適応フィルタリング処理し、フィルタ $EC(2)$ は、ステップ204に示されるように、第1フィルタ $EC(1)$ と並行してマイクロフォン信号 $d(t)$ を適応フィルタリング処理する。ここで用いられているように、フィルタが「並行にオペレーションする」とは、実質的に同じ入力 $d(t)$ を受け取ることをいう。並行オペレーションは、一のフィルタの出力が、他方のフィルタの入力となるシリアルオペレーションとは、区別される。二つのフィルタ $EC(1)$ 、 $EC(2)$ の状態によ

40

50

って、一のフィルタが、主要な「フロント」フィルタの役目を果たし、他方のフィルタが「バックアップ」フィルタの役目を果たす。一のフィルタは、エコーキャンセレーションに対して慎重なアプローチをとる一方、他方のフィルタはより積極的なアプローチをとる。

【0018】

フィルタEC(1)、EC(2)の状態は、以下の信号モデルに関連して理解されるであろう。

$$y(t) = x(t) * h(n)$$

$$d(t) = y_0(t) + s(t)$$

$$e(t) = d(t) - y(t)$$

10

ここで、 $y(t)$ は、エコーキャンセラフィルタによって合成されたエコーである。

$x(t)$ は、ラウドスピーカにおいてプレイするエコーである。

$h(n)$ は、エコーキャンセラフィルタの適応フィルタ関数である。

$d(t)$ は、マイクロフォンが受けた雑音の多い信号である。

$y_0(t)$ は、マイクロフォンにおいて現れる、真のエコーである。

$s(t)$ は、ローカル音声である。

そして、 $e(t)$ は、エコーキャンセラフィルタによって生成されたエコーキャンセル済み残差信号である。

【0019】

二つのフィルタEC(1)、EC(2)は、相補的なエコーキャンセレーション特質を有する。ここで用いられるように、「相補的なエコーキャンセレーションを有する」とは、同じ入力を受け取る二つの適応フィルタにおいて、一のフィルタが入力にうまく適応していないときに、他方のフィルタが入力にうまく適応しているような場合をいう。本アプリケーションの文脈において、フィルタ関数 $h(n)$ が、「うまく適応している」とは、そのフィルタ関数 $h(n)$ が安定であり、真のエコーパスフィルタ(echo-path-filter)に収束しており、過大予測でもなく過小予測でもないときをいう。

20

【0020】

$h(n)$ が真のエコーパスフィルタに収束している($y(t) \sim y_0(t)$)場合、すなわち、予測されたエコーが真のエコーと近似的に等しい場合、コヒーレンス関数を用いて、エコーキャンセラフィルタEC(1)、EC(2)の状態が定量化されるだろう。

30

は、 $y(t)$ と $e(t)$ の間の相互相関に関連し、式1が成り立つ。

(式1)

$$\alpha = E(e(t) * y(t)) / E(y(t) * y(t)),$$

ここで“E”は、統計的期待値である。

式2に示す演算子は、相互相関演算を表す。

(式2)

*

40

離散的な関数 f_i と g_i について、相互相関は式3で定義される。

(式3)

$$(f * g)_i \equiv \sum_j f_j^* g_{i+j}$$

ここで、和は適切な値の整数 j についてとられており、アスタリスクは、複共役を表す。連続関数 $f(x)$ と $g(x)$ について、相互相関は式4で定義される。

(式4)

50

$$(f * g)(x) \equiv \int f^*(t)g(x+t) dt$$

ここで積分は適切な t の値についてとられる。

【 0 0 2 1 】

コヒーレンス関数 において、分子は、 $e(t)$ と $y(t)$ の相互相関を表す。分母は、 $y(t)$ の自己相関を表し、正規化項の役目を果たす。

【 0 0 2 2 】

理想的には、 $h(n)$ が収束するならば、 ρ は「0」に近いはずである（残差信号 $e(t)$ は $y(t)$ を含まないからである）。 $h(n)$ が収束しないならば、 ρ は「1」に近いはずである（ $e(t)$ は $y(t)$ の強いエコーを含むからである）。 $h(n)$ がおかしな挙動をし、または発散するならば、 ρ は負であるはずである（フィルタの発散のため、 $e(t)$ は、位相が 180 度シフトした強いエコーを含むからである）。

【 0 0 2 3 】

したがって、例えば、コヒーレンス関数 の値は、フィルタ EC(1)、EC(2) の状態について、四つの可能な状態を定義するために用いられてもよい。ただしこれに制限されるものではない。

(1) フィルタ $h(n)$ が安定であり、収束し、過大予測でも過小予測でもない場合には、 $0 < \rho < 1$

(2) フィルタ $h(n)$ が安定ではあるが、過小予測されているときには（まだ収束していない） $\rho > 0$

(3) フィルタ $h(n)$ が過大予測されているときは、 $\rho < -0$

(4) フィルタ $h(n)$ が発散するときには、 $\rho < -0$

これらの異なる状態について、異なる ρ の値の範囲が決定されうることは当業者には理解されるであろう。

【 0 0 2 4 】

フィルタの状態がよいならば（例えば状態(1)）、その後発散したときのリカバリのために、その設定が保存されてもよい。フィルタが発散し、または過小予測され、または過大予測されている場合には、フロントおよびバックアップ・エコーキャンセラはその役割を交換する。フロントフィルタがバックアップとなる一方、バックアップフィルタがフロントフィルタの役割を担う。一のフィルタが慎重な適応アプローチをとり、他方が積極的な適応アプローチをとるため、この交換により、最終的には、両方のフィルタがより早く収束し、よりダイナミックに安定する。

【 0 0 2 5 】

さらにフィルタが過小予測または過大予測されている場合、より早い収束、またはトラッキングのよりよい安定のために、適応スピードを加速させ、または減速させるように、適応ステップサイズが小さなデルタ値で増加または減少されてもよい。通常、収束を速くするためにはより大きなステップサイズが必要である。これにより、細部に関するよいトラッキングは犠牲となり、エコーサプレッション比 ESR は、低くなる。小さなステップサイズを用いてよりゆっくりと収束させる場合は、より安定的であり、わずかな変化もトラッキングする機能を有するが、エコーディスロケーションを速くトラッキングするには適さない。

【 0 0 2 6 】

動的なステップサイズと、フロント/バックアップでのフィルタ交換を組み合わせることにより、速いトラッキング対詳細なトラッキング、安定性対収束の観点において、システム全体のバランスが良くなる。この二つが、適応システム設計において本当に重要な双子の課題である。

【 0 0 2 7 】

フィルタの一方が発散した場合において、他方のフィルタがよい状態にあるならば、発散

したフィルタを再初期化するために、その他方のフィルタの設定が複製されてもよい。別の方法では、発散したフィルタは、以前に保存された、よい状態のフィルタ設定を用いて復旧されて (r e c o v e r e d) もよい。

【 0 0 2 8 】

例えば、エコーキャンセリング適応フィルタ E C (1) と E C (2) は、周波数領域正規化最小二乗適応フィルタに基づいてもよい。ただし、これに制限されるものではない。各フィルタは、ハードウェア、ソフトウェア、またはハードウェアとソフトウェアの組み合わせとして実装されうる。

【 0 0 2 9 】

図 1 B と図 1 C は、適切な相補適応フィルタの例を示す。具体的には、図 1 B は、音声アクティビティ検出を伴う適応エコーキャンセレーションフィルタ 1 2 0 を示す。フィルタ 1 2 0 は、第 1 適応フィルタ E C (1) として用いることができる。フィルタ 1 2 0 は、フィルタ係数 w_t によって特徴付けられる有限インパルス応答 (F I R) フィルタを有する可変フィルタ 1 2 2 を含む。可変フィルタ 1 2 2 は、マイクロフォン信号 $d(t)$ を受け取り、フィルタ係数 w_t の値に従ってフィルタリング処理し、フィルタリング処理された信号 $d'(t)$ を生成する。可変フィルタ 1 2 4 は、入力信号を、係数 w_{t-1} によって決定されるインパルス応答で畳み込むことにより、望ましい信号を予測する。各フィルタ係数 w_{t-1} は、更新アルゴリズム 1 2 4 にしたがって、量 w_t の規則的な間隔で更新される。一例として、フィルタ信号 $d'(t)$ が、望ましい信号としてスピーカエコー信号 $x_1(t)$ を予測しようとするように、フィルタ係数 w_t が選択されてもよい。差分ユニット 1 2 6 は、マイクロフォン信号 $d(t)$ からフィルタリング処理された信号 $d'(t)$ を差し引いて、予測信号 $e_1(t)$ を供給する。予測信号 $e_1(t)$ は、ローカルスピーチ信号 $s(t)$ を予測する。フィルタリング処理された信号 $d'(t)$ をリモート信号 $x(t)$ から差し引いて、誤差信号 $e(t)$ を生成してもよい。誤差信号 $e(t)$ は、更新アルゴリズム 1 2 4 によってフィルタ係数 w_t を調整するために用いられる。適応アルゴリズム 1 2 4 は、リモート信号 $x(t)$ と誤差信号に基づいて、補正因子 (c o r r e c t i o n f a c t o r) を生成する。係数更新アルゴリズムの例には、最小 2 乗法 (L M S) と再帰最小 2 乗法 (R L S : r e c u r s i v e l e a s t s q u a r e s) が含まれる。L M S 更新アルゴリズムにおいては、例えば、フィルタ係数は、式 $w_{t-1+1} = w_{t-1} + \mu e(t) x(t)$ に基づいて更新される。ここで、 μ はステップサイズである。初めは、すべての w_{t-1} について、 $w_{t-1} = 0$ である。この例において、量 $\mu e(t) x(t)$ は、量 w_t であることに注意されたい。上述のように、ステップサイズ μ は、適応フィルタの状態によって、動的に調整されてもよい。具体的には、フィルタが過小予測されている場合には、早く収束するように適応スピードを加速するために、ステップサイズ μ を、小さなデルタ量増加させてもよい。フィルタが過大予測されている場合には、この代わりに、トラッキングがよりよく安定するように適応スピードを減速させるために、適応ステップサイズ μ が、小さなデルタ量でそれぞれ引き下げられてもよい。

【 0 0 3 0 】

時間領域表現 $e(t) x(t)$ は、乗算である。この計算は、以下のように周波数領域において実装されてもよい。初めに、 $e(t)$ 、 $x(t)$ 、および $h(n)$ は、時間領域から周波数領域に、例えば高速フーリエ変換 (F F T) によって変換されてもよい。

$$E(k) = f f t (e (t))$$

$$X(k) = f f t (x (t))$$

$$H(k) = f f t (h (n))$$

【 0 0 3 1 】

実際の周波数領域における L M S 更新アルゴリズムは、以下ようになる。

$$H(k) = H(k) + (\mu * c o n j (X(k)) . * E(k)) / (+ X(k) * c o n j (X(k))$$

ここで、 μ はフィルタ適応ステップサイズであり、動的である。

$\text{conj}(a)$ は、複素数 a の複素共役を示す。

$*$ は、複素乗算 (complex multiplication) を示す。

そして、 β は、分母が数量的に不安定になるのを防ぐレギュレータ (regulator) である。

【0032】

上の方程式において、「 $\text{conj}(X(k)) \cdot * E(k)$ 」は、「 $e(t) \times (t)$ 」タスクを実行する。分母において、「 $X(k) * \text{conj}(X(k))$ 」は、安定性を高める目的で正規化する役割を果たす。

【0033】

音声アクティブ化された検出 VAD は、更新アルゴリズム 124 を調整して、リモート信号 $x(t)$ が存在するとき (例えば所定の閾値以上であるならば)、可変フィルタ 122 が、マイクロフォン信号 $d(t)$ のみを適応的にフィルタリング処理するようにしてもよい。図 1B に示される音声アクティブ化された検出 (ダブルトーク検出と呼ばれることもある) を用いる適応フィルタは、比較的ゆっくりと適応するフィルタである。しかし、このフィルタはまた、擬陽性をほとんど生じないという点において、非常に正確である。フィルタ 120 に対する相補適応フィルタは、例えば、比較的早く適応するが、しばしば擬陽性を生じる傾向があるフィルタであるかもしれない。

【0034】

一例として、図 1C は、図 1B のフィルタ 120 に対して相補的な適応フィルタ 130 を示す。適応フィルタ 130 は、フィルタ係数 w_{t2} と更新アルゴリズム 134 (例えば上述の LMS 更新アルゴリズム) によって特徴づけられる可変フィルタを含む。フィルタ 132 は、スピーカエコー信号 $x_1(t)$ を望ましい信号として予測しようとする。差分ユニット 136 は、フィルタリング処理された信号 $d'(t)$ をマイクロフォン信号 $d(t)$ から差し引いて、ローカルスピーチ信号 $s(t)$ を予測する予測信号 $e_2(t)$ を提供する。フィルタリング処理された信号 $d'(t)$ をリモート信号 $x(t)$ から差し引いて、誤差信号 $e(t)$ を発生させてもよい。誤差信号 $e(t)$ はフィルタ係数 w_{t2} を調整するために更新アルゴリズム 134 によって用いられる。フィルタ 130 において相互相関解析 CCA は、可変フィルタ 132 が、予測信号 $e_2(t)$ とスピーカエコー信号 $x(t)$ との間の相互相関を低減させようとするように、更新アルゴリズム 134 を調整する。

【0035】

$e_2(t)$ と $x(t)$ が非常に強く相関しているとき、フィルタリング処理は過小予測されているといわれ、更新アルゴリズム 134 は、 w_{t2} を増加させるように調整される。 $e_2(t)$ と $x(t)$ との間の相互相関が閾値未満であるとき、フィルタリング処理は過大予測されているといわれ、更新アルゴリズム 134 は、 w_{t2} を減少させるように調整される。

【0036】

図 1C に示されるタイプの相互相関解析 (クロススペクトラム解析ともいわれる) を用いる適応フィルタは、比較的速くフィルタを適応させる。しかし、このフィルタはまた、しばしば擬陽性を生じるという点において、不安定である。したがって、フィルタ 120 とフィルタ 130 は、相補フィルタの例となる。

【0037】

再び図 1A を参照する。インテグレータ 106 は、第 1 適応フィルタ $EC(1)$ と第 2 適応フィルタ $EC(2)$ に接続される。インテグレータ 106 は、第 1 および第 2 適応フィルタのそれぞれの出力 $e_1(t)$ 、 $e_2(t)$ から、最小エコー出力 $e_3(t)$ を決定するように構成されている。最小エコー出力 $e_3(t)$ は、 $e_1(t)$ と $e_2(t)$ のいずれかであり、エネルギーがより小さく、スピーチ信号 $x(t)$ との相関がより小さい方である。 $e_1(t)$ と $e_2(t)$ のうちの一方のエネルギーの方がより小さいが、 $x(t)$ との相関は、他方がより小さい場合には、相関がより小さい方を最小エコー出力 $e_3(t)$ として用いる。例えば、フィルタのうちの一が過大予測されている (すなわち目標音声を

10

20

30

40

50

キャンセルしがちであるためにエネルギー出力が小さい)とき、エネルギーにかかわらず相関が小さいほうがよい。最小エネルギーは、 $E\{e_1(t)\}$ と $E\{e_2(t)\}$ との最小値を決定することにより決定されてもよい。ここで、 $E\{\}$ はカッコ内の量の期待値を決定する演算を示す。再び図2A-2Bを参照する。ステップ206において、 $e_1(t)$ と $e_2(t)$ のどちらがスピーカ信号 $x_1(t)$ との相互相関が小さいか決定するために、 $e_1(t)$ と $e_2(t)$ について相互相関解析が実行されてもよい。相互相関解析は、下記の式5と式6の最小値を決定するステップを含んでもよい。

(式5)

$$E\{e_1(t) \star x(t)\}$$

(式6)

$$E\{e_2(t) \star x(t)\}$$

ここで、式7の演算子

(式7)

\star

は、例えば、上で定義されたように演算子の両側の量について、その間の相互相関をとる演算を表現する。最小エコー出力 $e_3(t)$ は、マイクロフォン104のフィルタリング処理された出力として用いられてもよい。

【0038】

いくつかの状況においては、フィルタ $EC(1)$ 、 $EC(2)$ のうちの一が、ローカル信号を過度にフィルタリング処理するかもしれない。そのような状況においては、そのフィルタは「発散した」といわれる。これは、特に $EC(2)$ が、例えば図1Cに示されるようなタイプの相互相関フィルタであるときに実際に起こりうる。この可能性に対処するために、ステップ208において $EC(2)$ が発散するかどうか、決定される。一例としてインテグレータ106は、第2適応エコーキャンセレーションフィルタが、過度にフィルタリング処理することにより、ローカル信号 $s(t)$ を除去していないか、決定するように構成されてもよい。これは $e_2(t)$ とスピーカエコー信号 $x_1(t)$ との間の相互相関の期待値を調べることにより実行することができる。すなわち、式8で表される。

(式8)

$$E\{e_2(t) \star x(t)\}$$

典型的には、式9が成り立つ。

(式9)

$$E\{e_2(t) \star x(t)\} > 0$$

しかしながら、式10が、ある閾値(例えば約0.2)未満であるときには、 $EC(2)$ が過度にフィルタリング処理することにより、ローカル信号 $s(t)$ が除去されている。

(式10)

$$E\{e_2(t) \star x(t)\}$$

このような状況において、インテグレータ106は、 $e_1(t)$ を最小エコー出力 $e_3(t)$ として選択してもよい。適応フィルタリング処理を安定させるために、ステップ212において、 $EC(2)$ のフィルタ係数 w_{t2} が、 $EC(1)$ のフィルタ係数 w_{t1} として設定されてもよい。そしてステップ215において、 $EC(2)$ は、0、または、以前のうまく適応したことが知られている状態に、再初期化されてもよい。例えば、フィルタ係数は、規則的な間隔で(例えば約10秒から20秒ごとに)保存されて、 $EC(2)$ が発散したときにこれを再初期化するために用いられてもよい。

10

20

30

40

50

【 0 0 3 9 】

通常、相互相関フィルタが発散しないときに、そのフィルタはうまく適応していると言われる。EC(2)とEC(1)は相補的なフィルタリング特性を有するため、EC(2)がうまく適応しているとき、EC(1)は過小予測されていることになる。適応フィルタリング処理を安定化させるため、ステップ214に示されるように、第1適応フィルタEC(1)のフィルタ係数 w_{t_1} が、第2適応フィルタEC(2)のフィルタ係数 w_{t_2} と交換される。フィルタをソフトウェアに実装する際には、係数 w_{t_1} 、 w_{t_2} は、メモリにおいてポインタによって特定される位置に格納されてもよい。係数 w_{t_1} 、 w_{t_2} は、例えば、 w_{t_1} および w_{t_2} へのポインタを切り替えることによって、交換されてもよい。

10

【 0 0 4 0 】

最小エコー出力 $e_3(t)$ は、いくばくかの、スピーカ信号 $x(t)$ からの残差エコー $x_e(t)$ を含むかもしれない。装置100は、オプションで、インテグレータ106に接続された第1および第2エコー残差予測部ER(1)とER(2)、および、エコー残差予測部ER(1)とER(2)に接続された残差エコーキャンセレーションモジュール108を含んでもよい。

【 0 0 4 1 】

第1エコー残差予測部ER(1)は、最小エコー出力 $e_3(t)$ とスピーカ信号 $x(t)$ との間の相互相関解析を含む第1残差エコー予測 $ER_1(t)$ を生成するように構成されてもよい。図2Bのステップ222に示されるように、最小エコー出力 $e_3(t)$ とスピーカ信号 $x(t)$ との間の相互相関解析から、例えば、式11の値を決定することにより、第1残差エコー予測 $ER_1(t)$ が決定されてもよい。

20

(式11)

$$\operatorname{argmin}(E\{e_3(t) \star x(t)\}),$$

ここで、式11の値は、 $e_3(t)$ が式12の相互相関の期待値を最小化するとき、真である。

(式12)

$$e_3(t) \star x(t)$$

30

この最小化問題は、本質的に、適応により実現されるであろう。例えば、エコー残差予測部ER(1)が、初期状態においては単位フィルタ(すべて値"1")であると仮定されたい。すべてのフレームにおいて、サーチサーフェス(search surface)の接線方向(tangent direction)に向かうにつれて、第1残差エコー予測 $ER_1(t)$ は、増加するかもしれない。これは、ニュートンソルバ(Newton solver)アルゴリズムによって実現されてもよい。第2残差エコー予測部ER(2)は、最小エコー出力 $e_3(t)$ とスピーカ信号 $x(t)$ との間のエコー距離 mismatch(echo-distance mismatch)を含む第2残差エコー予測 $ER_2(t)$ を決定するように構成されてもよい。図2Bのステップ224に示されるように、最小エコー出力 $e_3(t)$ とスピーカ信号 $x(t)$ との間のエコー距離 mismatchから、例えば、 $\operatorname{argmin}(E\{(e_3(t))^2 / (x(t))^2\})$ を決定することにより、第2残差エコー予測 $ER_2(t)$ が決定されてもよい。ここで、 $e_3(t)$ が商 $(e_3(t))^2 / (x(t))^2$ の期待値を最小化するとき、 $\operatorname{argmin}(E\{(e_3(t))^2 / (x(t))^2\})$ は真である。ここでも再び、最小化は、ニュートンソルバアルゴリズムを用いて実現されてもよい。

40

【 0 0 4 2 】

残差エコーキャンセレーションモジュール108は、二つの残差エコー予測 $ER_1(t)$ と $ER_2(t)$ の最小残差エコー予測 $ER_3(t)$ を決定して、その最小値 $ER_3(t)$ に従ってフィルタリング処理された信号 $e_3(t)$ を調整してもよい。一例として、最小残差エコー予測 $ER_3(t)$ は、 $ER_1(t)$ と $ER_2(t)$ のうち、エネルギーが最小

50

であり、 $x(t)$ に対する相関が最小であるものであってもよい。例えば図2Bのステップ226に示されるように、 $ER_1(t)$ と $ER_2(t)$ のうちの最小値に設定され、ステップ228に示されるように、その結果である ER_3 の値が $e_3(t)$ から差しひかれて、残差エコーキャンセルフィルタリング処理された信号 $e_3'(t)$ が生成される。 ER_3 が $ER_1(t)$ に等しいならば、残差エコー $x_e(t)$ は、ローカルスピーチ信号 $s(t)$ の強度が0でないときに、最小限に除去される。 $ER_3(t)$ が $ER_2(t)$ に等しいならば、残差エコー $x_e(t)$ は、遠端のエコー $x(t)$ のみが存在するとき(遠端発話のみの期間)最大限に除去される。

【0043】

一例として、2次のノルム $N(1)$ と $N(2)$ が、二つのエコー残差予測部 $ER(1)$ と $ER(2)$ のためにそれぞれ計算されてもよい。

$$N(1) = ER(1)$$

$$N(2) = ER(2)$$

【0044】

ダブルトーク状況下においては、より小さいノルムを有するエコー残差予測部が、エコー残差ノイズを取り除くために、 $e_3(t)$ に適用されてもよい。シングルトーク状況下においては、より大きいノルムを有するエコー残差予測部が、エコー残差ノイズを取り除くために、 $e_3(t)$ に適用されてもよい。

【0045】

エコーキャンセレーションにおいては、フィルタリング処理された信号 $e_3(t)$ 、または、残差エコーキャンセルフィルタリング処理された信号 $e_3'(t)$ から、ノイズ $n(t)$ が除去されてもよい。ただし、このようなノイズキャンセレーションは、望ましくないかもしれない。なぜならば、信号 $e_3(t)$ または $e_3'(t)$ のリモート受信者は、ノイズがない状態を、マイクロフォン104からのすべての通信が失われた徴候であると解釈するかもしれないからである。この問題に対処するために、装置100はオプションで、ノイズキャンセラユニット110を含んでもよい。ノイズキャンセレーションモジュール110は、例えば図2A-2Bのステップ217に示されるように、マイクロフォン信号 $d(t)$ から予測ノイズ信号 $n'(t)$ を計算するように構成されてもよい。予測ノイズ信号 $n'(t)$ は、減衰係数で減衰されて、低減されたノイズ信号 $n''(t) = n'(t)$ を形成してもよい。減衰されたノイズ信号 $n''(t)$ は、図2Aのステップ218に示されるように $e_3(t)$ に加算されることにより、または、図2Bのステップ230に示されるように $e_3'(t)$ に加算されることにより、マイクロフォン出力信号 $s'(t)$ に組み込まれてもよい。

【0046】

本発明の実施形態においては、図1A-1Cに関連して説明された装置、および図2A-2Cに関連して説明された方法は、プログラマブルなプロセッサとメモリを有するシステム上のソフトウェアとして実装されてもよい。

【0047】

本発明の実施形態によると、図1および図2A-Bに関連して説明されたタイプの、前述のように動作する信号処理方法は、図3に示されるように、信号処理装置300の一部として実装されてもよい。システム300は、プロセッサ301とメモリ302(例えば、RAM、DRAM、ROMなど)を含んでもよい。信号処理装置300はさらに、並行処理が実装される場合には、複数のプロセッサ301を有してもよい。メモリ302は前述のように構成されたデータおよびコードを含む。具体的には、メモリ302には、プログラムコード304と信号データ306が格納されてもよい。コード304は、上述の、エコーキャンセリング適応フィルタ $EC(1)$ 、 $ER(2)$ 、インテグレータ106、エコー残差フィルタ $ER(1)$ 、 $ER(2)$ 、残差エコーキャンセレーションモジュール108、ノイズキャンセラ110を実装してもよい。信号データ306は、マイクロフォン信号 $d(t)$ 、かつ/または、スピーカ信号 $x(t)$ のデジタル表現を含んでもよい。

【0048】

10

20

30

40

50

装置300はまた、入出力(I/O)エレメント311、電源(P/S)312、クロック(CLK)313、キャッシュメモリ314といった、周知のサポート機能310を含んでもよい。装置300は、プログラム、かつ/または、データを格納するためのディスクドライブ、CD-ROMドライブ、テープドライブといった大容量記憶装置315をオプションで含んでもよい。コントローラは、また、オプションで、コントローラ300とユーザの間の対話を手助けするためのディスプレイユニット316と、ユーザインタフェイスユニット318を含んでもよい。ディスプレイユニット316は、ブラウン管型でもよく、またフラットパネルスクリーンでもよい。これらはテキスト、数値、グラフィックシンボル、画像を表示する。ユーザインタフェイス318は、キーボード、マウス、ジョイスティック、ライトペン(light pen)やそのほかの装置を含んでもよい。さらに、スピーカ322とマイクロフォン324は、入出力構成エレメント311を介してプロセッサ301に接続されていてもよい。プロセッサ301、メモリ302、そしてシステム300のほかの構成要素は、図3に示されるようにシステムバス320を介して互いに信号(例えば、コード・インストラクションとデータ)を交換してもよい。

【0049】

ここで用いられるように、入出力という言葉は、一般的に、システム300への、またはシステム300からの、および周辺装置への、または周辺装置からのデータを転送する任意のプログラム、オペレーション、または装置を指す。すべてのデータ転送が、一の装置からの出力であり、他の一の装置への入力であると見なすことができるであろう。周辺装置は、キーボードやマウスなどの入力のための装置や、プリンタなどの出力のための装置、そして上書き可能CD-ROMなどの入力および出力装置として動作する装置を含む。周辺装置という言葉には、マウス、キーボード、プリンタ、モニタ、マイクロフォン、ゲームコントローラ、カメラ、外部Zipドライブ、スキャナなどの外部装置と、CD-ROMドライブ、CD-Rドライブ、内部モデムなどの内部装置、および、フラッシュメモリ用リーダー/ライター、ハードドライブなどのそのほかの周辺装置を含む。

【0050】

プロセッサ301は、信号データ306およびメモリ302によって格納され、獲得されプロセッサモジュール301によって実行されるプログラム304のプログラムコード命令に応じて、信号データ306にデジタル信号処理を実行する。プログラム304のコードの一部はアセンブリ、C++、Java(登録商標)またはそのほかの多くの言語のような様々な異なるプログラミング言語のうちの一であってよい。プロセッサモジュール301は、プログラムコード304のようなプログラムを実行するときには特別な目的のコンピュータとなる汎用コンピュータを構成する。プログラムコード304は、ここでは、汎用コンピュータ上で実行されるソフトウェアとして実装されるものとして説明されたが、これに代えて、アプリケーション特定集積回路(ASIC)のようなハードウェアを用いて、タスク管理方法が実現されることは当業者には理解されるであろう。そのように、本発明の実施形態は、全体的にまたは部分的に、ソフトウェア、ハードウェア、またはこれらの組合せによって実現されることは理解されるであろう。

【0051】

ある実施形態においては、とりわけプログラムコード304は、図2Aの方法200や図2Bの方法220に共通な特徴を有する方法を実現するためのプロセッサ可読命令のセットを含んでもよい。プログラムコード304は、一般的に、以下のような命令を含んでもよい。すなわち、プロセッサ301に、相補的エコーキャンセレーション特性を有する第1および第2適応フィルタによって並行にマイクロフォン信号 $d(t)$ をフィルタリング処理させ、エコーキャンセル処理された出力 $e_1(t)$ と $e_2(t)$ を生成させる命令、 $e_1(t)$ と $e_2(t)$ から最小エコー出力 $e_3(t)$ を決定する命令、最小エコー出力を用いてマイクロフォン出力を生成する命令である。

【0052】

本発明の実施形態によると、相互相関解析のみ、または音声アクティビティ検出(ダブルトーク検出)のみで可能な、よりロバストでありながら正確なエコーキャンセレーショ

10

20

30

40

50

ンが可能となる。このような改良されたエコーキャンセレーションによると、スピーカエコー $x(t)$ に大部分を占められているマイクロフォン信号 $d(t)$ からローカルスピーチを $s(t)$ を抽出することが可能となる。

【0053】

本発明の実施形態は、ここで提示されたように用いられてもよく、また他のユーザ入力メカニズムと共に用いられてもよい。方位角方向や音声のボリュームを追跡したり測定したりするメカニズム、かつ/または、能動的または受動的にオブジェクトの位置を追跡するメカニズム、マシン・ビジョンを用いるメカニズム、これらの組み合わせなどである。追跡されるオブジェクトは、システムへのフィードバックを操作する補助的なコントロール装置やボタンを含んでもよい。そのようなフィードバックには、光源からの光の放射、音質の歪曲手段、その他の適切な送信機、変調器、コントロール装置、ボタン、圧力パッドなどが含まれてもよいが、これらに制限されるものではない。それは、同じ符号化状態の転送や変調に影響を及ぼしてもよく、かつ/または、システムによって追跡されている装置への命令や、その装置からの命令を転送してもよい。そのような装置は、本発明の実施形態に関連して用いられるシステムの一部であったり、またはシステムと相互作用したり、またはシステムに影響を与えたりする。

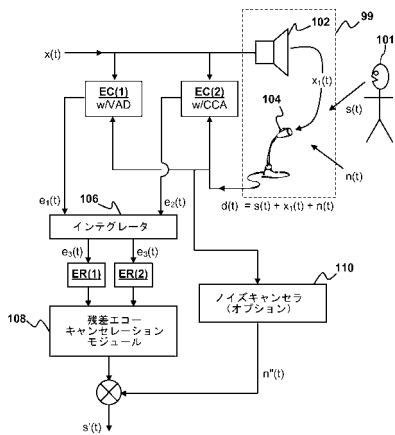
10

【0054】

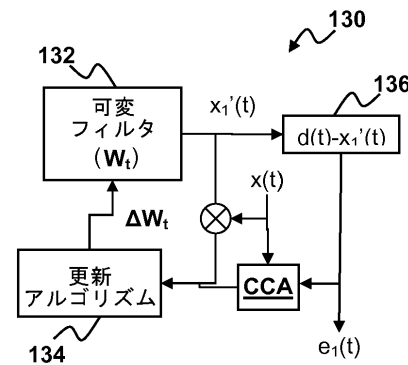
上記は、本発明の好ましい実施形態の完全な記述であるが、他の様々な変形、変更、等価物への置換が可能である。それゆえ、本発明の範囲は、上記の記述によって決定されるのではなく、以下の請求項によって決定されるべきであり、その完全な等価物もその範囲に含まれる。ここで記述された特徴は、好ましいものであるか否かに関わらず、ここで述べたいずれの特徴と組み合わせられてもよい。以下の請求項においては、特に明示的に断らない限りは、各要素の数量は一以上である。ここに、添付される請求項は、所与の請求項において、「～ための手段」との語句を用いて明示的に示される場合の他は、ミーンズ・プラス・ファンクションの制限を含むと解されてはならない。

20

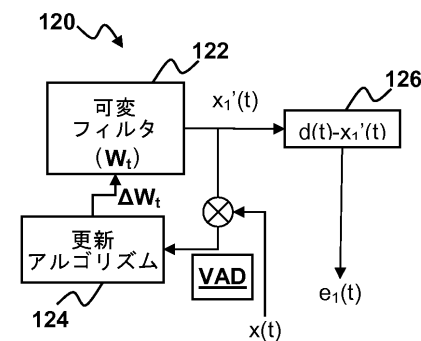
【図1A】



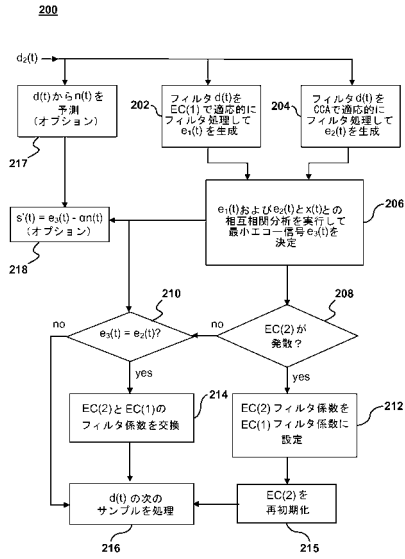
【図1C】



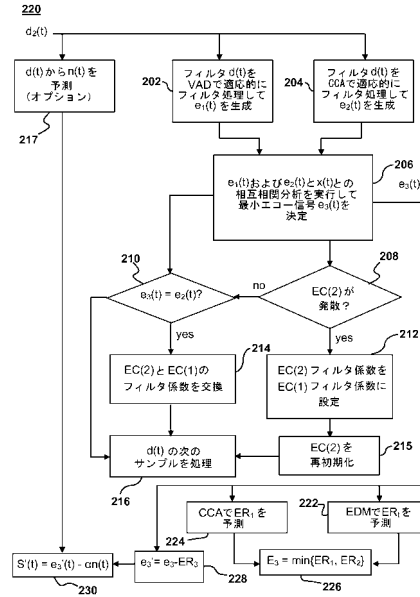
【図1B】



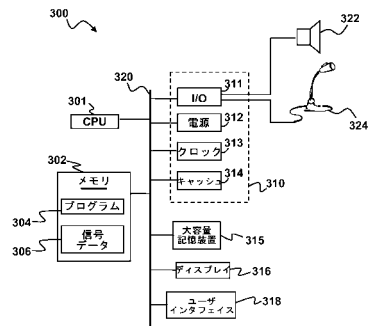
【図 2 A】



【図 2 B】



【図 3】



フロントページの続き

- (31)優先権主張番号 11/381,725
(32)優先日 平成18年5月4日(2006.5.4)
(33)優先権主張国 米国(US)
- (31)優先権主張番号 11/381,727
(32)優先日 平成18年5月4日(2006.5.4)
(33)優先権主張国 米国(US)
- (31)優先権主張番号 11/381,724
(32)優先日 平成18年5月4日(2006.5.4)
(33)優先権主張国 米国(US)
- (31)優先権主張番号 11/381,721
(32)優先日 平成18年5月4日(2006.5.4)
(33)優先権主張国 米国(US)
- (31)優先権主張番号 PCT/US2006/017483
(32)優先日 平成18年5月4日(2006.5.4)
(33)優先権主張国 米国(US)
- (31)優先権主張番号 11/418,988
(32)優先日 平成18年5月4日(2006.5.4)
(33)優先権主張国 米国(US)
- (31)優先権主張番号 11/418,989
(32)優先日 平成18年5月4日(2006.5.4)
(33)優先権主張国 米国(US)
- (31)優先権主張番号 11/429,047
(32)優先日 平成18年5月4日(2006.5.4)
(33)優先権主張国 米国(US)

早期審査対象出願

- (56)参考文献 米国特許出願公開第2004/0161101(US,A1)
国際公開第2005/079109(WO,A1)
特表平07-503109(JP,A)
特表平11-514516(JP,A)
特開2004-080647(JP,A)

(58)調査した分野(Int.Cl.,DB名)

H04R 1/00-31/00
H04B 3/23
H04M 1/00-1/82