

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6422884号  
(P6422884)

(45) 発行日 平成30年11月14日(2018.11.14)

(24) 登録日 平成30年10月26日(2018.10.26)

(51) Int.Cl.	F I	
<b>H04B 3/23 (2006.01)</b>	H04B 3/23	
<b>H04M 1/00 (2006.01)</b>	H04M 1/00	R
	H04M 1/00	H

請求項の数 10 (全 15 頁)

(21) 出願番号	特願2015-549850 (P2015-549850)	(73) 特許権者	314015767
(86) (22) 出願日	平成25年12月21日(2013.12.21)		マイクロソフト テクノロジー ライセン
(65) 公表番号	特表2016-503262 (P2016-503262A)		シング, エルエルシー
(43) 公表日	平成28年2月1日(2016.2.1)		アメリカ合衆国 ワシントン州 9805
(86) 国際出願番号	PCT/US2013/077353		2 レッドモンド ワン マイクロソフト
(87) 国際公開番号	W02014/100792		ウェイ
(87) 国際公開日	平成26年6月26日(2014.6.26)	(74) 代理人	100107766
審査請求日	平成28年11月24日(2016.11.24)		弁理士 伊東 忠重
(31) 優先権主張番号	1223246.8	(74) 代理人	100070150
(32) 優先日	平成24年12月21日(2012.12.21)		弁理士 伊東 忠彦
(33) 優先権主張国	英国 (GB)	(74) 代理人	100091214
(31) 優先権主張番号	13/889, 229		弁理士 大貫 進介
(32) 優先日	平成25年5月7日(2013.5.7)		
(33) 優先権主張国	米国 (US)		

最終頁に続く

(54) 【発明の名称】 エコー抑圧

(57) 【特許請求の範囲】

【請求項 1】

エコーを抑圧する方法であって、  
 オーディオ出力信号を出力するステップと、  
 オーディオ入力信号を受信するステップであって、前記の受信されたオーディオ入力信号は、前記オーディオ出力信号の出力から生じるエコーを含む、ステップと、  
 前記の受信されたオーディオ入力信号における前記エコーのエコー・パス  $h(n)$  をモデル化するために、前記オーディオ出力信号及び前記の受信されたオーディオ入力信号に基づいて、時間領域で有限インパルス応答フィルタ推定値

【数 1】

$$\hat{h}(n)$$

を動的に適応させるステップと、  
 前記有限インパルス応答フィルタ推定値

【数 2】

$$\hat{h}(n)$$

及び前記オーディオ出力信号を使用して、前記の受信されたオーディオ入力信号におけるエコー・パワーを推定するステップと、

前記の推定されたエコー・パワーに基づいて、前記の受信されたオーディオ入力信号に

10

20

おける前記エコーを抑圧するステップと、  
を含む、方法。

【請求項 2】

前記エコーを抑圧する前にエコー・キャンセルを前記の受信されたオーディオ入力信号に適用することなく、前記エコーを抑圧することは、前記の受信されたオーディオ入力信号に対して実行される、請求項 1 記載の方法。

【請求項 3】

前記エコーを抑圧する前にエコー・キャンセルを前記の受信されたオーディオ入力信号に適用するステップをさらに含み、前記エコーを抑圧することは、前記の適用されたエコー・キャンセルの後、前記の受信されたオーディオ入力信号に残っている残差エコーを抑圧する、請求項 1 記載の方法。

10

【請求項 4】

前記エコーを抑圧することは、時間及び周波数に応じて変わる信号依存の抑圧である、請求項 1 乃至 3 いずれか一項記載の方法。

【請求項 5】

前記方法は、通信イベントにおける使用のために、ユーザ・デバイスにおいて実行され、前記の受信されたオーディオ入力信号は、前記通信イベントにおいて前記ユーザ・デバイスから送信されるユーザの音声を含む、請求項 1 乃至 4 いずれか一項記載の方法。

【請求項 6】

前記通信イベントは、ボイス・オーバー・インターネット・プロトコル (VoIP) 通話である、請求項 5 記載の方法。

20

【請求項 7】

前記オーディオ出力信号は、前記の受信されたオーディオ入力信号における前記エコーに含まれる、前記 VoIP 通話の遠端音声信号を含む、請求項 6 記載の方法。

【請求項 8】

エコー抑圧を実施するよう構成されるデバイスであって、  
オーディオ出力信号を出力するよう構成されるオーディオ出力装置と、  
オーディオ入力信号を受信するよう構成されるオーディオ入力装置であって、前記の受信されたオーディオ入力信号は、前記オーディオ出力信号の出力から生じるエコーを含む、オーディオ入力装置と、  
前記の受信されたオーディオ入力信号における前記エコーに関連するエコー・パス  $h(n)$  をモデル化するために、前記オーディオ出力信号及び前記の受信されたオーディオ入力信号に基づいて、時間領域で有限インパルス応答フィルタ推定値

30

【数 3】

$$\hat{h}(n)$$

を動的に適応させるよう構成されるフィルタ推定モジュールと、

前記有限インパルス応答フィルタ推定値

【数 4】

$$\hat{h}(n)$$

40

及び前記オーディオ出力信号に基づいて、前記の受信されたオーディオ入力信号におけるエコー・パワーを推定するよう構成されるパワー推定モジュールと、

前記の推定されたエコー・パワーに基づいて、前記の受信されたオーディオ入力信号における前記エコーを抑圧するよう構成されるエコー抑圧モジュールと、

を備えた、デバイス。

【請求項 9】

前記オーディオ出力装置は、前記オーディオ出力信号を出力するよう構成されるスピーカを含み、前記オーディオ入力装置は、前記オーディオ入力信号を受信するよう構成されるマイクロフォンを含む、請求項 8 記載のデバイス。

50

## 【請求項 10】

受信されたオーディオ信号におけるエコーを抑圧するよう構成されるシステムであって

プロセッサと、

複数の動作を実行するための、前記プロセッサにより実行可能な命令を含むコンピュータ読み取り可能記憶メモリであって、前記複数の動作は、

オーディオ出力信号を出力する動作と、

オーディオ入力信号を受信する動作であって、前記の受信されたオーディオ入力信号は、前記オーディオ出力信号の出力から生じるエコーを含む、動作と、

前記の受信されたオーディオ入力信号における前記エコーのエコー・パス  $h(n)$  をモデル化するために、前記オーディオ出力信号及び前記の受信されたオーディオ入力信号に基づいて、時間領域で有限インパルス応答フィルタ推定値

【数 5】

$$\hat{h}(n)$$

を動的に適応させる動作と、

前記有限インパルス応答フィルタ推定値

【数 6】

$$\hat{h}(n)$$

及び前記オーディオ出力信号を使用して、前記の受信されたオーディオ入力信号におけるエコー・パワーを推定する動作と、

前記の推定されたエコー・パワーに基づいて、前記の受信されたオーディオ入力信号における前記エコーを抑圧する動作と、

を含む、コンピュータ読み取り可能記憶メモリと、  
を備えた、システム。

【発明の詳細な説明】

【背景技術】

【0001】

デバイスは、周囲環境からオーディオ信号を受信するために使用することができるオーディオ入力装置を有することができる。このデバイスはまた、周囲環境にオーディオ信号を出力するために使用することができるオーディオ出力装置を有することができる。例えば、デバイスは、オーディオ信号を出力するための1以上のスピーカと、オーディオ信号を受信するための1以上のマイクロフォンと、を有することができる。デバイスの1以上のスピーカから出力されるオーディオ信号は、1以上のマイクロフォンにより受信されるオーディオ信号における「エコー」として受信され得る。受信オーディオ信号におけるこのエコーは望ましくない場合があり得る。例えば、デバイスは、音声通話又はビデオ通話等の、ネットワークを介した別のユーザ・デバイスとの通信イベントにおいて使用されるユーザ・デバイス（携帯電話機、タブレット、ラップトップ、PC等）であり得る。通話の遠端信号は、ユーザ・デバイスにおいて、スピーカから出力され得、そのデバイスにおいてマイクロフォンにより受信されるオーディオ信号におけるエコーとして受信され得る。そのようなエコーは、通話のユーザを乱すことがあり、通話の知覚品質は、エコーのため低減され得る。詳細には、エコーは、マイクロフォンにより受信され、通話における遠端に送信されることが意図されている近端オーディオ信号の干渉を生じさせることがある。したがって、エコー・キャンセル及び/又はエコー抑圧を受信オーディオ信号に適用して、それにより、受信オーディオ信号におけるエコーを抑圧することができる。受信オーディオ信号におけるエコーのパワーは、ユーザ・デバイスの構成に応じて変わり得る。例えば、ユーザ・デバイスは携帯電話機であり得、その場合、受信オーディオ信号におけるエコーのパワーは、通常、携帯電話機が「ハンズ・フリー」モードで動作していないときと比べて、携帯電話機が「ハンズ・フリー」モードで動作しているときの方が、高いであ

10

20

30

40

50

ろう。

【 0 0 0 2 】

エコー・キャンセル（又は、「エコー減算（echo subtraction）」）技術は、スピーカから出力されるオーディオ信号の知識に基づいて、マイクロフォンにおいて受信されるオーディオ信号に含まれるエコー信号を推定することを目的としている。次いで、エコー信号の推定値を、受信オーディオ信号から減算することができ、それにより、受信オーディオ信号からエコーの少なくとも一部を取り除くことができる。エコー抑圧を使用して、周波数依存の抑圧を受信オーディオ信号に適用し、それにより、受信オーディオ信号におけるエコーを抑圧する。エコー抑圧が効果的に実施されるために、エコー抑圧部は、受信オーディオ信号におけるエコーのパワーの正確な推定値を有する必要がある。

10

【発明の概要】

【 0 0 0 3 】

この概要は、発明を実施するための形態において以下でさらに説明されるコンセプトのうち選択したものを簡略化した形で紹介するために提供される。この概要は、特許請求される主題の主要な特徴又は必要不可欠な特徴を特定することを意図するものではないし、特許請求される主題の範囲を限定するために使用されることを意図するものでもない。

【 0 0 0 4 】

受信オーディオ信号におけるエコーを抑圧する方法が提供される。エコー抑圧の一部として、エコーのエコー・パワーの推定値が、エコー・パス（echo path）のインパルス応答を近似するために適応される有限インパルス応答（FIR）フィルタを用いて決定される。すなわち、出力オーディオ信号及び受信オーディオ信号に基づいて、時間領域で有限インパルス応答フィルタ推定値（Finite Impulse Response filter estimate）

20

【数 1】

$$\hat{h}(n)$$

を動的に適応させ、それにより、受信オーディオ信号におけるエコーのエコー・パス  $h(n)$  のインパルス応答をモデル化する。フィルタ推定値

【数 2】

$$\hat{h}(n)$$

30

を使用して、受信オーディオ信号におけるエコーのエコー・パワーを推定し、推定したエコー・パワーを使用して、エコー抑圧を受信オーディオ信号に適用し、それにより、受信オーディオ信号におけるエコーを抑圧する。

【 0 0 0 5 】

当該方法は、通話（例えば、ユーザ・デバイス間でオーディオ・データを送信するためにボイス・オーバー・インターネット・プロトコル（VoIP）を実施する通話）において使用することができる。この場合、出力オーディオ信号は、通話の遠端から受信される遠端信号であり得、受信信号は、結果として生じるエコーと、通話の遠端に送信するための近端信号と、を含む。

【図面の簡単な説明】

40

【 0 0 0 6 】

【図 1】 1 以上の実施形態に従った通信システムの概略図。

【図 2】 1 以上の実施形態に従ったユーザ・デバイスの概略ブロック図。

【図 3】 1 以上の実施形態に従ったエコー抑圧において使用するためのユーザ・デバイスのモジュールを示す機能図。

【図 4】 1 以上の実施形態に従った、エコーを抑圧するプロセスのフローチャート。

【発明を実施するための形態】

【 0 0 0 7 】

エコー抑圧が効果的に実施されるために、エコー抑圧部は、受信オーディオ信号におけるエコーのパワーの正確な推定値を有する必要がある。本明細書で説明するように、オー

50

ディオ信号を出力するラウドスピーカと、出力オーディオ信号から生じるエコーを含むオーディオ信号を受信するマイクロフォンと、の間のエコー・パスのインパルス応答を近似するために適応されるFIRフィルタからの出力を用いて、エコー・パワーを推定することができる。

#### 【0008】

エコー・パワーを推定するために、FIRフィルタが使用され得るだけであり、実際のエコー信号を推定するわけではない。実際のエコー信号を推定するためにFIRフィルタが使用される場合と比べて、エコー・パワーを推定するためにFIRフィルタが使用される場合の方が、FIRフィルタにおける正確性の要件がずっと低いので、これは効果的であり得る。したがって、FIRフィルタから、(エコー信号ではなく)エコー・パワーを推定することにより、エコー抑圧は、VoIPクライアントにおける(例えば、ラウドスピーカからの)プレイアウト(playout)と(例えば、マイクロフォンにおける)記録サイドとの間のクロック・ドリフト、エコー・パスにおける非線形性、及びエコー・パスにおける変化等の問題に対して、よりロバストである。本明細書で説明する実施形態において、FIRフィルタは、位相情報を含む時間領域データを用いて適応される。

#### 【0009】

図1は、第1のユーザ・デバイス104に関連付けられている第1のユーザ102(「ユーザA」)と、第2のユーザ・デバイス110に関連付けられている第2のユーザ108(「ユーザB」)と、を含む通信システム100を示している。他の実施形態では、通信システム100は、任意の数のユーザ及び関連ユーザ・デバイスを含んでもよい。ユーザ・デバイス104及びユーザ・デバイス110は、通信システム100において、ネットワーク106を介して通信することができ、それにより、ユーザ102及びユーザ108は、ネットワーク106を介して、互いと通信することが可能になる。図1に示す通信システム100は、パケットベースの通信システムであるが、他のタイプの通信システムが使用されてもよい。ネットワーク106は、例えば、インターネットであり得る。ユーザ・デバイス104及びユーザ・デバイス110の各々は、例えば、携帯電話機、タブレット、ラップトップ、パーソナル・コンピュータ(「PC」)(例えば、Windows(登録商標)PC、Mac OS(登録商標)PC、及びLinux(登録商標)PCを含む)、ゲーム・デバイス、テレビジョン、携帯情報端末(「PDA」)、又は、ネットワーク106に接続することができる他の組み込みデバイスであり得る。ユーザ・デバイス104は、ユーザ・デバイス104のユーザ102から情報を受け取り、ユーザ102に情報を出力するよう構成されている。ユーザ・デバイス104は、ディスプレイ及びスピーカ等の出力手段を備える。ユーザ・デバイス104はまた、キーボード、タッチ・スクリーン、オーディオ信号を受信するためのマイクロフォン、及び/又はビデオ信号の画像をキャプチャするためのカメラ等の入力手段を備える。ユーザ・デバイス104は、ネットワーク106に接続される。

#### 【0010】

ユーザ・デバイス104は、通信システム100に関連付けられたソフトウェア・プロバイダにより提供される通信クライアントのインスタンス(instance)を実行する。通信クライアントは、ユーザ・デバイス104内のローカル・プロセッサ上で実行されるソフトウェア・プログラムである。クライアントは、ユーザ・デバイス104が通信システム100を介してデータを送受信するために、ユーザ・デバイス104において必要とされる処理を実行する。

#### 【0011】

ユーザ・デバイス110は、ユーザ・デバイス104に対応し、ユーザ・デバイス104において実行される通信クライアントに対応する通信クライアントを、ローカル・プロセッサ上で実行する。ユーザ・デバイス110におけるクライアントは、ユーザ・デバイス104におけるクライアントが、ユーザ102がネットワーク106を介して通信することを可能にするのに必要とされる処理を実行するのと同様に、ユーザ108がネットワーク106を介して通信することを可能にするのに必要とされる処理を実行する。ユーザ

10

20

30

40

50

・デバイス 104 及びユーザ・デバイス 110 は、通信システム 100 におけるエンドポイントである。図 1 は、明瞭さのために、2 人のユーザ (102 及び 108) 及び 2 つのユーザ・デバイス (104 及び 110) しか示していないが、より多くのユーザ及びより多くのユーザ・デバイスが、通信システム 100 に含まれてもよく、それぞれのユーザ・デバイス上で実行されるそれぞれの通信クライアントを用いて、通信システム 100 を介して通信してもよい。

#### 【0012】

図 2 は、通信システム 100 を介して通信するための通信クライアント・インスタンス 206 が実行されるユーザ・デバイス 104 の詳細な図を示している。ユーザ・デバイス 104 は、中央処理装置 (「CPU」) 又は「処理モジュール」202 を備える。CPU 202 には、タッチ・スクリーンとして実装することができるディスプレイ 208、及びオーディオ信号を出力するためのスピーカ (又は、「ラウドスピーカ」) 210 等の出力デバイス; オーディオ信号を受信するためのマイクロフォン 212、画像データを受信するためのカメラ 216、及びキーパッド 218 等の入力デバイス; データを記憶するためのメモリ 214; 及び、ネットワーク 106 と通信するためのモデム等のネットワーク・インタフェース 220 が接続される。ユーザ・デバイス 104 は、図 2 に示す要素以外の他の要素を備えてもよい。ディスプレイ 208、スピーカ 210、マイクロフォン 212、メモリ 214、カメラ 216、キーパッド 218、及びネットワーク・インタフェース 220 は、図 2 に示すユーザ・デバイス 104 に一体化され得る。代替的なユーザ・デバイスでは、ディスプレイ 208、スピーカ 210、マイクロフォン 212、メモリ 214、カメラ 216、キーパッド 218、及びネットワーク・インタフェース 220 のうちの 1 以上が、ユーザ・デバイス 104 に一体化されなくてもよく、それぞれのインタフェースを介して CPU 202 に接続されてもよい。そのようなインタフェースの一例は、USB インタフェースである。ネットワーク・インタフェース 220 を介したネットワーク 106 へのユーザ・デバイス 104 の接続が、無線接続である場合、ネットワーク・インタフェース 220 は、ネットワーク 106 に信号を無線で送信し、ネットワーク 106 から信号を無線で受信するためのアンテナを含み得る。

#### 【0013】

図 2 はまた、CPU 202 上で実行されるオペレーティング・システム (「OS」) 204 を示している。通信システム 100 のクライアント・インスタンス 206 のソフトウェアが、OS 204 の上で実行されている。オペレーティング・システム 204 は、コンピュータのハードウェア・リソースを管理し、ネットワーク・インタフェース 220 を介してネットワーク 106 に / から送信されているデータを処理する。クライアント 206 は、オペレーティング・システム 204 と通信し、通信システムを介した接続を管理する。クライアント 206 は、ユーザ 102 に情報を提示しユーザ 104 から情報を受け取るために使用されるクライアント・ユーザ・インタフェースを有する。このように、クライアント 206 は、ユーザ 102 が通信システム 100 を介して通信することを可能にするのに必要とされる処理を実行する。

#### 【0014】

図 3 及び図 4 を参照して、エコーを抑圧する方法について説明する。図 3 は、エコー抑圧プロセスがどのように実施されるかを示す、ユーザ・デバイス 104 の一部の機能図であり、図 4 は、エコーを抑圧するプロセスのフローチャートである。

#### 【0015】

図 3 に示されるように、ユーザ・デバイス 104 は、スピーカ 210、マイクロフォン 212、FIR フィルタ・モジュール 302、パワー推定モジュール 304、及びエコー抑圧モジュール 306 を備える。スピーカ 210 から出力されるべき信号  $x(t)$  は、スピーカ 210 の入力に結合される。本明細書で説明する実施形態では、(図において参照番号 210 により示される) 1 つのラウドスピーカしか存在しないが、他の実施形態では、出力されるべき信号が (ラウドスピーカから出力するために) 結合される 2 以上のラウドスピーカが存在してもよいことに留意すべきである。同様に、本明細書で説明する実施

形態では、（図において参照番号 212 により示される）1つのマイクロフォンしか存在しないが、他の実施形態では、周囲環境からオーディオ信号を受信する2以上のマイクロフォンが存在してもよい。ラウドスピーカ 210 から出力されるべき信号はまた、FIR フィルタ・モジュール 302 の第1の入力に結合され、パワー推定モジュール 304 の第1の入力に結合される。マイクロフォン 212 の出力は、FIR フィルタ・モジュール 302 の第2の入力に結合され、エコー抑圧モジュール 306 の第1の入力に結合される。FIR フィルタ・モジュール 302 の出力は、パワー推定モジュール 304 の第2の入力に結合される。パワー推定モジュール 304 の出力は、エコー抑圧モジュール 306 の第2の入力に結合される。エコー抑圧モジュール 306 の出力を使用して、ユーザ・デバイス 104 におけるさらなる処理のために、（エコー抑圧が適用された）受信信号を供給する。

10

#### 【0016】

ステップ S 402 において、スピーカ 210 から出力されるべき信号が受信される。例えば、出力されるべき信号は、ユーザ 102 とユーザ 108 との間の通話中に通信システム 100 を介してユーザ・デバイス 110 からユーザ・デバイス 104 で受信される遠端信号であり得る。受信信号に対して実行されるのに必要とされる任意の処理（例えば、音声コーデックを用いた復号化、脱パケット化等）は、当技術分野では知られているように、スピーカ 210 から出力されるべき適切な信号  $x(t)$  に達するために、（例えば、クライアント 206 により）実行される。信号  $x(t)$  はデジタル信号である。スピーカ 210 から信号を出力する前の、ユーザ・デバイス 104 における信号の処理の少なくとも一部は、デジタル領域で実行される。当技術分野では知られているように、ラウドスピーカ 210 からのプレイアウトの前に、デジタル・アナログ・コンバータ（DAC）が、デジタル信号  $x(t)$  に適用される。同様に、デジタル信号  $y(t)$  に達するために、アナログ・デジタル・コンバータ（ADC）が、マイクロフォン 212 によりキャプチャされた信号に適用される。

20

#### 【0017】

他の実施形態において、出力されるべき信号は、通話において通信システム 100 を介してではなく、どこかから受信されることもある。例えば、出力されるべき信号は、メモリ 214 に記憶されているかもしれず、ステップ S 402 は、メモリ 214 から信号を取り出すことを含み得る。

30

#### 【0018】

ステップ S 404 において、オーディオ信号  $x(t)$  が、スピーカ 210 から出力される。このように、オーディオ信号  $x(t)$  が、ユーザ 102 に出力される。

#### 【0019】

ステップ S 406 において、マイクロフォン 212 は、オーディオ信号を受信する。図 3 に示されるように、受信オーディオ信号は、所望信号又は「プライマリ信号」である近端信号を含み得る。近端信号は、マイクロフォン 212 が受信するものとユーザ 102 が意図している信号である。しかしながら、受信オーディオ信号はまた、ステップ S 404 においてスピーカ 210 から出力されたオーディオ信号から生じるエコー信号も含む。受信オーディオ信号はまた、背景雑音等の雑音も含む。したがって、総受信オーディオ信号  $y(t)$  は、近端信号、エコー、及び雑音の合計により与えられ得る。エコー及び雑音は、近端信号の干渉として働く。

40

#### 【0020】

FIR フィルタ・モジュール 302 は、出力オーディオ信号  $x(t)$  及び受信オーディオ信号  $y(t)$  を、入力として受け取る。ステップ S 408 において、FIR フィルタ・モジュール 302 は、受信オーディオ信号  $y(t)$  におけるエコーのエコー・パス  $h(n)$  をモデル化するために、出力オーディオ信号  $x(t)$  及び受信オーディオ信号  $y(t)$  に基づいて、時間領域で FIR フィルタ推定値

【数 3】

$$\hat{h}(n)$$

を動的に適応させる。「エコー・パス  $h(n)$  のインパルス応答」は、本明細書において、「エコー・パス  $h(n)$ 」とも呼ばれる。

【0021】

長さ  $N_{true} + 1$  の線形エコー・パス・インパルス応答に関して、エコー・パス  $h(n)$  は、例えば、式

【数 4】

$$y^{echo}(t) = \sum_{n=0}^{N_{true}} h(n)x(t-n)$$

10

に従って、受信オーディオ信号におけるエコーが、スピーカ 210 から出力されたオーディオ信号  $x(t)$  にどのように関連するかを記述するものである。ここで、 $y^{echo}(t)$  は、受信オーディオ信号  $y(t)$  におけるエコーである。エコー・パス  $h(n)$  は、時間とともに変わり得る。エコー・パス  $h(n)$  は、(i) スピーカ 210 及びマイクロフォン 212 周囲の現在の環境状態（例えば、スピーカ 210 からマイクロフォン 212 へのオーディオ信号の通路に物理的障害物があるかどうか、空気圧、温度、風等）、及び (ii) 信号が出力及び／又は受信されるときに信号を変え得るスピーカ 210 及び／又はマイクロフォン 212 の特性、に依存し得る。

【0022】

20

FIR フィルタ・モジュール 302 は、出力オーディオ信号  $x(t)$  の現在の値と、出力オーディオ信号  $x(t)$  の有限個 ( $N$  個) の以前の値と、の重み付け合計を決定することにより、受信オーディオ信号におけるエコーのエコー・パス  $h(n)$  をモデル化する。FIR フィルタ・モジュール 302 は、したがって、エコー・パスの推定値

【数 5】

$$\hat{h}(n)$$

を決定する際に出力オーディオ信号  $x(t)$  の値を考慮に入れる有限長（時間）を有する  $N$  次 FIR フィルタを実装する。このように、FIR フィルタ・モジュール 302 は、FIR フィルタ推定値

30

【数 6】

$$\hat{h}(n)$$

を動的に適応させる。出力オーディオ信号  $x(t)$  の観点で、受信オーディオ信号  $y(t)$  におけるエコーを定義する以下の式により、オペレーションが記述される。

【数 7】

$$\hat{y}^{echo}(t) = \sum_{n=0}^N \hat{h}(n)x(t-n)$$

【0023】

したがって、出力オーディオ信号  $x(t)$  の  $N + 1$  個のサンプルとともに、それぞれの  $N + 1$  個の重み

40

【数 8】

$$\hat{h}(n)$$

が使用される。  $N + 1$  個の重み

【数 9】

$$\hat{h}(n)$$

のセットは、本明細書において、単に、エコー・パスの推定値



【数 1 0】

$$\hat{h}(n)$$

と呼ばれる。すなわち、エコー・パスの推定値

【数 1 1】

$$\hat{h}(n)$$

は、 $N + 1$  個の値からなる。ここで、FIR フィルタ・モジュール 3 0 2 は、信号  $x(t)$  の  $N + 1$  個の値（例えば、 $N + 1$  個のサンプル）を考慮に入れる  $N$  次 FIR フィルタを実装する。

10

【0 0 2 4】

エコーが受信オーディオ信号の支配部であるとき、すなわち、

【数 1 2】

$$y(t) \cong y^{echo}(t)$$

であるとき、FIR フィルタ推定値

【数 1 3】

$$\hat{h}(n)$$

をより適応させやすいことが理解できるであろう。例えば、いくつかの実施形態において、近端信号のパワーがエコーのパワーよりも大きいとき（例えば、ユーザ 1 0 2 が喋っているとき）を検出することが可能であり得るが、これは、FIR 推定値

20

【数 1 4】

$$\hat{h}(n)$$

が適応されない場合である。しかしながら、近端信号のパワーが受信オーディオ信号  $y(t)$  におけるエコーのパワーよりも小さいとき（例えば、ユーザ 1 0 2 が喋っていないとき）、FIR 推定値

【数 1 5】

$$\hat{h}(n)$$

30

が適応される。

【0 0 2 5】

しかしながら、エコーが受信オーディオ信号の支配部でないときでも、FIR フィルタ推定値

【数 1 6】

$$\hat{h}(n)$$

を適応させることが可能であり得る。例えば、適応がより遅い場合、外乱に対するロバスト性におけるゲインが存在する。

40

【0 0 2 6】

FIR フィルタ推定値

【数 1 7】

$$\hat{h}(n)$$

が、FIR フィルタ・モジュール 3 0 2 からパワー推定モジュール 3 0 4 に渡される。ステップ S 4 1 0 において、パワー推定モジュール 3 0 4 は、ステップ S 4 0 8 において決定されたフィルタ推定値

【数 1 8】

$$\hat{h}(n)$$

に基づいて、受信オーディオ信号におけるエコーのエコー・パワーを推定する。ステップ S 4 1 0 は、受信オーディオ信号  $y(t)$  におけるエコー信号

【数 1 9】

$$\hat{y}^{echo}(t)$$

を推定することを含まなくてよい。エコー・パワーは、時間及び周波数の関数として推定される。エコー抑圧部のためにエコー・パワーを推定するために、推定エコー・パス

10

【数 2 0】

$$\hat{h}(n)$$

及び入力信号  $x(t)$  を用いることは、エコー減算部 / キャンセル部のためにエコー信号の推定値

【数 2 1】

$$\hat{y}^{echo}(t)$$

を用いるよりも、エコー・パスの F I R フィルタ近似における不正確性に対してずっと影響を受けにくい。例えば、エコー・パワー推定のための、F I R フィルタ・モジュール 3 0 2 により提供されるエコー・パスの推定値

20

【数 2 2】

$$\hat{h}(n)$$

及び入力信号  $x(t)$  の使用は、プレイアウトとローカル・エンドポイント上の記録サイドとの間のクロック・ドリフト、準最適な F I R フィルタ・アップデート、エコー・パスにおける非線形性、及びエコー・パスにおける時間変動等の問題に対してずっと影響を受けにくい。エコー抑圧においては、良好なエコー抑圧を実現するために、エコー・パワー推定の正確性がかなり低くても十分である。このように、エコー抑圧のために使用されるパワー・スペクトルを推定するための線形フィルタが使用される。これは、エコー減算部のために時間領域のエコー信号を推定するための線形フィルタを用いるよりもロバストである。

30

【0 0 2 7】

ステップ S 4 1 0 は、時間  $t$  及び周波数  $f$  のスカラー指数ビン (scalar power bin) である、エコー・パワー

【数 2 3】

$$\hat{P}_s(t, f)$$

を推定することを含み得る。これは、当業者により知られているように、フーリエ変換を用いて計算することができる。

40

【0 0 2 8】

このように、時間変動する V o I P のエコー・パス  $h(n)$  を近似するためにスピーカ信号  $x(t)$  及びマイクロフォン信号  $y(t)$  を用いて適応されている F I R フィルタ推定値

【数 2 4】

$$\hat{h}(n)$$

を、出力オーディオ信号サンプル  $x(t)$  とともに使用して、時間  $t$  及び周波数  $f$  におけるエコー信号のパワー

【数 2 5】

$$\hat{P}_s(t, f)$$

を推定する。

【0 0 2 9】

エコー・パワーの推定値

【数 2 6】

$$\hat{P}_s(t, f)$$

が、パワー推定モジュール 3 0 4 から出力され、エコー抑圧モジュール 3 0 6 により受信される。エコー抑圧モジュール 3 0 6 はまた、マイクロフォン 2 1 2 からオーディオ信号  $y(t)$  も受信する。ステップ S 4 1 2 において、エコー抑圧モジュール 3 0 6 は、エコー・パワーの推定値

10

【数 2 7】

$$\hat{P}_s(t, f)$$

を使用して、エコー抑圧を受信オーディオ信号  $y(t)$  に適用し、それにより、受信オーディオ信号におけるエコーを抑圧する。エコー・パワーの推定値

【数 2 8】

$$\hat{P}_s(t, f)$$

20

は、周波数に依存し、エコー抑圧モジュール 3 0 6 により適用される抑圧もまた、周波数に依存する。

【0 0 3 0】

エコー抑圧部の目的は、例えば、VoIPクライアントにおけるマイクロフォン信号内に存在するラウドスピーカ・エコーを、マイクロフォン 2 1 2 により受信される近端音（非エコー音）の存在内で目立たない／乱さないよう十分低いレベルに抑圧することである。適切な量のエコー抑圧を選択できるようにするために、（例えば、周波数及び時間の関数としての）エコー・パワーの良好な推定値が必要とされ、上述したように、これが、パワー推定モジュール 3 0 4 により、エコー抑圧モジュール 3 0 6 に提供される。エコー抑圧モジュール 3 0 6 は、時間及び周波数の両方に応じて変わる信号依存の抑圧を、受信オーディオ信号  $y(t)$  に適用するよう設計される。エコー抑圧方法は、当技術分野において知られている。さらに、エコー抑圧モジュール 3 0 6 により適用されるエコー抑圧方法は、様々な方法で実装することができる。そのようなものとして、エコー抑圧方法の正確な詳細については、したがって、本明細書において詳細に説明しない。

30

【0 0 3 1】

エコー抑圧モジュール 3 0 6 は、ユーザ・デバイス 1 0 4 におけるさらなる処理のために、エコーが抑圧された受信信号を出力する。例えば、エコー抑圧モジュール 3 0 6 から出力される信号は、クライアント 2 0 6 により処理（例えば、符号化及びパケット化）することができ、次いで、ユーザ 1 0 2 とユーザ 1 0 8 との間の通話において、ネットワーク 1 0 6 を介して、ユーザ・デバイス 1 1 0 に送信することができる。追加的に又は代替的に、エコー抑圧モジュール 3 0 6 から出力される信号は、他の目的のために、ユーザ・デバイス 1 0 4 により使用されてもよい。例えば、信号は、メモリ 2 1 4 に記憶されてもよいし、ユーザ・デバイス 1 0 4 において実行されているアプリケーションへの入力として使用されてもよい。

40

【0 0 3 2】

したがって、（例えば、VoIPクライアント 2 0 6 による使用のために）エコー抑圧エフェクト／フィルタを計算及び適用する目的で、マイクロフォン 2 1 2 により受信されたラウドスピーカのエコー信号のパワーを推定するためにエコー・パスをモデル化する FIR フィルタ・モジュール 3 0 2 の使用について、本明細書で説明した。

50

## 【 0 0 3 3 】

上述した実施形態において、エコー抑圧は、V o I Pシステムにおいて実施される（例えば、受信オーディオ信号は、ユーザ 1 0 2 とユーザ 1 0 8 との間の通話中に通信システム 1 0 0 を介してユーザ・デバイス 1 1 0 に送信されるユーザ 1 0 2 の音声を含み得る）。しかしながら、本明細書で説明したエコー抑圧方法は、エコー抑圧が適用されるべき任意の適切なシステムにおいて適用することができる。

## 【 0 0 3 4 】

上述した実施形態において、図に示されるように、エコー・キャンセル（又は、「エコー減算」）は、受信オーディオ信号  $y(t)$  に適用されない。すなわち、ユーザ・デバイス 1 0 4 内にはエコー・キャンセル・モジュールは存在せず、エコー抑圧は、エコー・キャンセルを受信オーディオ信号  $y(t)$  に適用する従来のステップなく、受信オーディオ信号  $y(t)$  に適用される。

## 【 0 0 3 5 】

しかしながら、他の実施形態では、エコー・キャンセル・モジュールにより、エコー・キャンセルが、受信オーディオ信号  $y(t)$  に適用されてもよい。詳細には、エコー抑圧モジュール 3 0 6 により適用されるエコー抑圧が、受信オーディオ信号  $y(t)$  の処理において、エコー・キャンセルのダウンストリームで（すなわち、後で）、適用されてもよい。エコー・キャンセル・モジュールは、受信オーディオ信号からエコー信号の推定値を減算するが、エコー信号の推定における不正確性に起因して、残差エコーが、受信オーディオ信号に残っている可能性が非常に高い。次いで、エコー抑圧モジュール 3 0 6 により抑圧されるのが、残差エコーである。このエコー抑圧は、エコー・キャンセルが適用されない実施形態において本明細書で説明したのと同じ方法で適用することができる。エコー減算が用いられる場合、その結果が、エコー抑圧において考慮に入れられる。

## 【 0 0 3 6 】

本明細書で説明した方法は、ユーザ・デバイス 1 0 4 においてコンピュータ・プログラム製品（例えば、クライアント 2 0 6 ）を実行することにより、実施することができる。すなわち、コンピュータ・プログラム製品は、受信オーディオ信号  $y(t)$  におけるエコーを抑圧するよう構成され得る。ここで、コンピュータ・プログラム製品は、コンピュータ読み取り可能記憶媒体上で具現化され（例えば、メモリ 2 1 4 に記憶され）、C P U 2 0 2 上で実行されたときに、本明細書で説明した方法のうちいずれかのオペレーションを実行するよう構成される。用語「コンピュータ読み取り可能記憶媒体」は、全ての固定形態の媒体をカバーするよう意図されるものであり、したがって、非固定形態の媒体を含まない。

## 【 0 0 3 7 】

一般に、本明細書で説明した機能（例えば、図 3 に示した機能モジュール及び図 4 に示した機能ステップ）のいずれも、ソフトウェア、ファームウェア、ハードウェア（例えば、固定論理回路）、又はそれらの実装の組合せを用いて実装することができる。図 3 及び図 4 に別々に示したモジュール及びステップは、別々のモジュール又はステップとして実装されてもよいし、別々のモジュール又はステップとして実装されなくてもよい。例えば、エコー抑圧モジュール 3 0 6 は、パワー推定モジュール 3 0 4 の機能を実行してもよいので、別のパワー推定モジュール 3 0 4 は、エコー抑圧モジュール 3 0 6 に加えて必要とされるものではない。本明細書で使用される「モジュール」、「機能」、「コンポーネント」、及び「ロジック」という用語は、一般に、ソフトウェア、ファームウェア、ハードウェア、又はそれらの組合せを表す。ソフトウェア実装の場合、モジュール、機能、又はロジックは、プロセッサ（例えば、1 以上の C P U ）上で実行されたときに特定のタスクを実行するプログラム・コードを表す。プログラム・コードは、1 以上のコンピュータ読み取り可能メモリ・デバイスに記憶することができる。本明細書で説明した技術の特徴は、プラットフォームに依存しない、すなわち、本明細書で説明した技術は、様々なプロセッサを有する様々な商用コンピューティング・プラットフォーム上で実施できることを意味する。

## 【 0 0 3 8 】

例えば、ユーザ・デバイスはまた、ユーザ・デバイスのハードウェアに、例えば、プロセッサの機能ブロック等のオペレーションを実行させるエンティティ（例えば、ソフトウェア）を含んでもよい。例えば、ユーザ・デバイスは、ユーザ・デバイス、より具体的には、ユーザ・デバイスのオペレーティング・システム及び関連ハードウェアにオペレーションを実行させる命令を保持するよう構成され得るコンピュータ読み取り可能媒体を含んでもよい。したがって、そのような命令は、オペレーションを実行するオペレーティング・システム及び関連ハードウェアを構成するよう機能し、このようにして、機能を実行するオペレーティング・システム及び関連ハードウェアの変形（transformation）をもたらす。命令は、様々な異なる構成を通して、コンピュータ読み取り可能媒体により、ユーザ・デバイスに提供されてもよい。

10

## 【 0 0 3 9 】

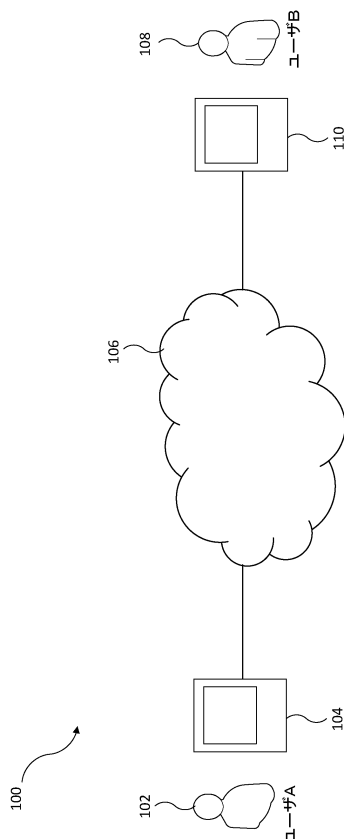
コンピュータ読み取り可能媒体の1つのそのような構成は、信号搬送媒体であり、したがって、例えば、ネットワークを介して、命令を（例えば、搬送波として）コンピューティング・デバイスに送信するよう構成される。コンピュータ読み取り可能媒体は、コンピュータ読み取り可能記憶媒体として構成されてもよく、したがって、信号搬送媒体ではない。コンピュータ読み取り可能記憶媒体の例は、ランダム・アクセス・メモリ（RAM）、読み取り専用メモリ（ROM）、光ディスク、フラッシュ・メモリ、ハード・ディスク・メモリ、及び、命令及び他のデータを記憶するために磁気技術、光技術、及び他の技術を使用することができる他のメモリ・デバイスを含む。

20

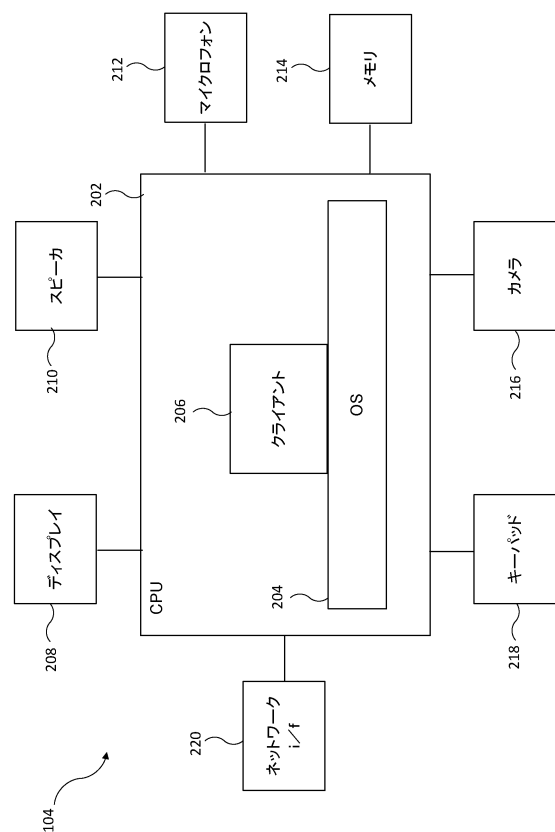
## 【 0 0 4 0 】

構造的特徴及び／又は方法論的動作に特有の言葉で主題を説明したが、添付の特許請求の範囲において定められる主題は、上述した特定の特徴又は動作に必ずしも限定されないことを理解すべきである。むしろ、上述した特定の特徴及び動作は、請求項を実施する例示的な形態として開示されたものである。

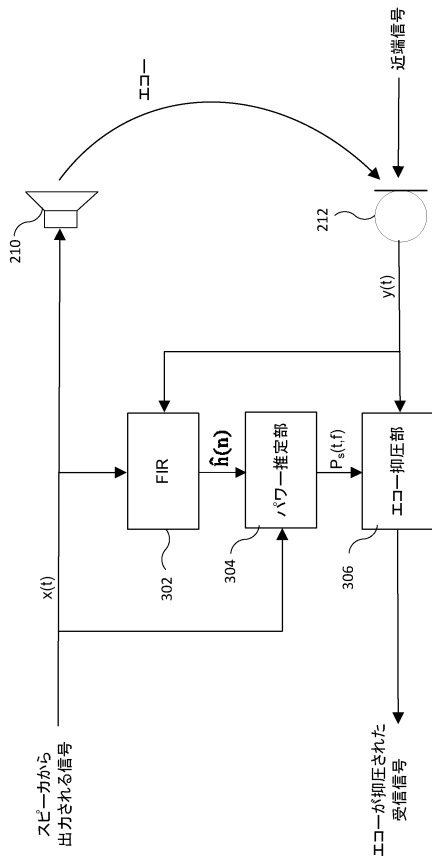
## 【 図 1 】



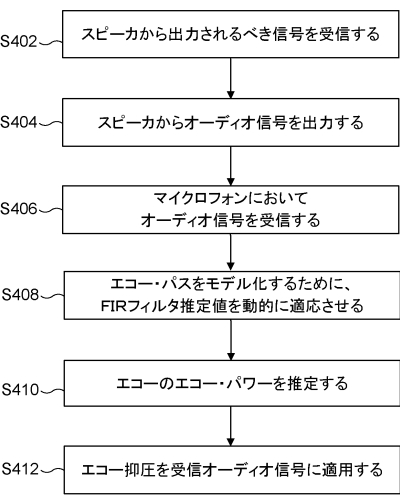
## 【 図 2 】



【図 3】



【図 4】



---

フロントページの続き

(72)発明者 ソレンセン, カルステン, ヴァンドボルグ

アメリカ合衆国 98052-6399 ワシントン州 レッドモンド ワン マイクロソフト  
ウェイ マイクロソフト コーポレーション エルシーイー - インターナショナル パテント 内

審査官 鴨川 学

(56)参考文献 米国特許出願公開第2011/0176687 (US, A1)

特開2009-021741 (JP, A)

特開2007-097050 (JP, A)

特表2006-513621 (JP, A)

特開2002-064617 (JP, A)

(58)調査した分野(Int.Cl., DB名)

H04B 1/76 - 3/44

H04B 3/50 - 3/60

H04B 7/005 - 7/015

H04M 1/00

H04M 1/24 - 1/82

H04M 99/00