

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2017-538151

(P2017-538151A)

(43) 公表日 平成29年12月21日(2017.12.21)

(51) Int.Cl.		F I	テーマコード (参考)	
G 1 O L	21/0232 (2013.01)	G 1 O L	21/0232	5 D 2 2 0
G 1 O L	21/0364 (2013.01)	G 1 O L	21/0364	
H O 4 R	3/00 (2006.01)	H O 4 R	3/00 3 2 0	

審査請求 未請求 予備審査請求 未請求 (全 23 頁)

(21) 出願番号 特願2017-525347 (P2017-525347)
 (86) (22) 出願日 平成27年11月12日 (2015.11.12)
 (85) 翻訳文提出日 平成29年6月27日 (2017.6.27)
 (86) 国際出願番号 PCT/US2015/060337
 (87) 国際公開番号 W02016/077557
 (87) 国際公開日 平成28年5月19日 (2016.5.19)
 (31) 優先権主張番号 62/078,844
 (32) 優先日 平成26年11月12日 (2014.11.12)
 (33) 優先権主張国 米国 (US)
 (31) 優先権主張番号 14/938,816
 (32) 優先日 平成27年11月11日 (2015.11.11)
 (33) 優先権主張国 米国 (US)

(71) 出願人 504371240
 シラス ロジック、インコーポレイテッド
 アメリカ合衆国 テキサス 78701,
 オースティン, ダブリュー, 6ティ
 ーエイチ ストリート 800
 (74) 代理人 100107489
 弁理士 大塩 竹志
 (72) 発明者 シャーウッド, エリク
 アメリカ合衆国 ユタ 84020, ド
 ライパー, イー. 12200 サウス
 282, スイート 200
 (72) 発明者 グルンドストルム, カール
 アメリカ合衆国 ユタ 84020, ド
 ライパー, イー. 12200 サウス
 282, スイート 200
 最終頁に続く

(54) 【発明の名称】 適応チャネル間弁別的リスケーリングフィルタ

(57) 【要約】

オーディオ信号をフィルタ処理する方法は、一次および基準チャネルの高速フーリエ変換 (FFT) 係数の確率密度関数 (PDF) をモデル化することと、基準チャネルの雑音の大きさの推定値と一次チャネルの雑音の大きさの推定値との間の弁別的関連性差 (DRD) を提供するために、PDFを最大化することを含む。一次チャネルは、一次チャネルのスペクトルの大きさが基準チャネルのものよりも強いとき、強調され、基準チャネルのスペクトルの大きさが一次チャネルのものよりも強いとき、強調は抑えられる。乗算リスケーリング係数が、音声増強フィルタチェーンの事前段階において算出された利得に適用され、いかなる事前段階も存在しない場合、利得は、直接適用される。

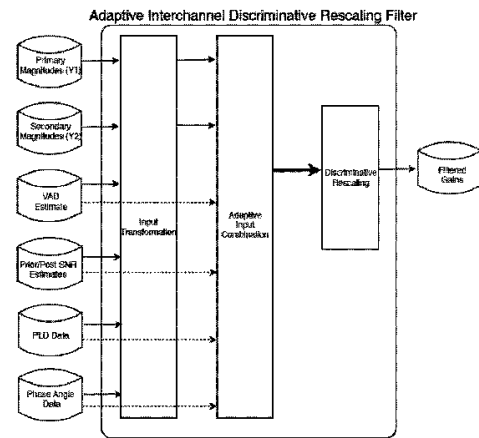


FIGURE 1

【特許請求の範囲】

【請求項 1】

オーディオ信号を変換する方法であって、

オーディオデバイスの一次マイクロホンを用いて、オーディオ信号の一次チャンネルを取得することと、

前記オーディオデバイスの基準マイクロホンを用いて、前記オーディオ信号の基準チャンネルを取得することと、

複数の周波数ビンに対する前記オーディオ信号の前記一次チャンネルのスペクトルの大きさを推定することと、

複数の周波数ビンに対する前記オーディオ信号の前記基準チャンネルのスペクトルの大きさを推定することと、

一次分数変換および高次有理関数変換のうちの少なくとも1つを適用することによって、1つ以上の周波数ビンに対する前記スペクトルの大きさのうちの1つ以上のものを変換することと、

前記スペクトルの大きさのうちの1つ以上のものを再正規化すること、

前記スペクトルの大きさのうちの1つ以上のものを累乗すること、

前記スペクトルの大きさのうちの1つ以上のものを時間平滑化すること、

前記スペクトルの大きさのうちの1つ以上のものを周波数平滑化すること、

前記スペクトルの大きさのうちの1つ以上のものをVADベースで平滑化すること、

前記スペクトルの大きさのうちの1つ以上のものを心理音響的に平滑化すること、

位相差の推定値を前記変換されたスペクトルの大きさのうちの1つ以上のものと組み合わせること、および、

VAD推定値を前記変換されたスペクトルの大きさのうちの1つ以上のものと組み合わせること

のうちの1つ以上によって、1つ以上の周波数ビンに対する前記スペクトルの大きさのうちの1つ以上のものを変換することと

を含む、方法。

【請求項 2】

増加的入力に基づいて、前記一次分数変換および前記高次有理関数変換のうちの少なくとも1つをビン毎に更新することをさらに含む、請求項 1 に記載の方法。

【請求項 3】

事前SNR推定値および事後SNR推定値のうちの少なくとも1つを、前記変換されたスペクトルの大きさのうちの1つ以上のものと組み合わせることをさらに含む、請求項 1 に記載の方法。

【請求項 4】

信号電力レベル差(SPLD)データを、前記変換されたスペクトルの大きさのうちの1つ以上のものと組み合わせることをさらに含む、請求項 1 に記載の方法。

【請求項 5】

雑音の大きさの推定値および雑音電力レベル差(NPLD)に基づいて、前記基準チャンネルの補正されたスペクトルの大きさを計算することをさらに含む、請求項 1 に記載の方法。

【請求項 6】

前記雑音の大きさの推定値および前記NPLDに基づいて、前記一次チャンネルの補正されたスペクトルの大きさを計算することをさらに含む、請求項 5 に記載の方法。

【請求項 7】

前記スペクトルの大きさのうちの1つ以上のものをフレーム内の近傍の周波数ビンにわたりとられる加重平均に置き換えることと、前記スペクトルの大きさのうちの1つ以上のものを前のフレームからの対応する周波数ビンにわたりとられる加重平均に置き換えることとのうちの少なくとも1つをさらに含む、請求項 1 に記載の方法。

【請求項 8】

10

20

30

40

50

オーディオ信号に適用されるフィルタ処理の程度を調節する方法であって、
オーディオデバイスの一次マイクロホンを用いて、オーディオ信号の一次チャンネルを取得することと、

前記オーディオデバイスの基準マイクロホンを用いて、前記オーディオ信号の基準チャンネルを取得することと、

前記オーディオ信号の前記一次チャンネルのスペクトルの大きさを推定することと、

前記オーディオ信号の前記基準チャンネルのスペクトルの大きさを推定することと、

前記オーディオ信号の前記一次チャンネルの高速フーリエ変換 (F F T) 係数の確率密度関数 (P D F) をモデル化することと、

前記オーディオ信号の前記基準チャンネルの高速フーリエ変換 (F F T) 係数の確率密度関数 (P D F) をモデル化することと、

単一チャンネル P D F および結合チャンネル P D F のうちの少なくとも 1 つを最大化し、前記基準チャンネルの雑音の大きさの推定値と前記一次チャンネルの雑音の大きさの推定値との間の弁別的関連性差 (D R D) を提供することと、

所与の周波数に対してどのスペクトルの大きさがより大きいかを決定することと、

前記一次チャンネルのスペクトルの大きさが前記基準チャンネルのスペクトルの大きさよりも強いとき、前記一次チャンネルを強調することと、

前記基準チャンネルのスペクトルの大きさが前記一次チャンネルのスペクトルの大きさよりも強いとき、前記一次チャンネルの強調を抑えることと

を含み、

前記強調することおよび強調を抑えることは、事前段階が存在する場合、乗算リスケーリング係数を算出し、音声増強フィルタチェーンの事前段階において算出された利得に前記乗算リスケーリング係数を適用することと、いかなる事前段階も存在しない場合、利得を直接適用することとを含む、

方法。

【請求項 9】

前記乗算リスケーリング係数は、利得として使用される、請求項 8 に記載の方法。

【請求項 10】

前記一次および基準オーディオチャンネルのうちの少なくとも 1 つの各スペクトルフレームに増加的入力を含めることをさらに含む、請求項 8 に記載の方法。

【請求項 11】

前記増加的入力は、前記一次チャンネルに対するスペクトルフレームの各ピンにおける事前 S N R および事後 S N R の推定値を含む、請求項 10 に記載の方法。

【請求項 12】

前記増加的入力は、前記一次チャンネルおよび前記基準チャンネルに対するスペクトルフレームの対応するピン間のピンあたり N P L D の推定値を含む、請求項 10 に記載の方法。

【請求項 13】

前記増加的入力は、前記一次チャンネルおよび基準チャンネルに対するスペクトルフレームの対応するピン間のピンあたり S P L D の推定値を含む、請求項 10 に記載の方法。

【請求項 14】

前記増加的入力は、前記一次チャンネルと前記基準チャンネルとの間のフレームあたり位相差の推定値を含む、請求項 10 に記載の方法。

【請求項 15】

オーディオデバイスであって、

オーディオ信号を受信し、前記オーディオ信号の一次チャンネルを通信するための一次マイクロホンと、

前記オーディオ信号を前記一次マイクロホンとは異なる状況で受信し、前記オーディオ信号の基準チャンネルを通信するための基準マイクロホンと、

前記オーディオ信号をフィルタ処理および / または明瞭化するために前記オーディオ信号を処理する少なくとも 1 つの処理要素と

10

20

30

40

50

を備え、
前記少なくとも1つの処理要素は、方法を実施するためのプログラムを実行するように構成され、

前記方法は、

オーディオデバイスの一次マイクロホンを用いて、オーディオ信号の一次チャネルを取得することと、

前記オーディオデバイスの基準マイクロホンを用いて、前記オーディオ信号の基準チャネルを取得することと、

前記オーディオ信号の前記一次チャネルのスペクトルの大きさを推定することと、

前記オーディオ信号の前記基準チャネルのスペクトルの大きさを推定することと、

前記オーディオ信号の前記一次チャネルの高速フーリエ変換 (FFT) 係数の確率密度関数 (PDF) をモデル化することと、

前記オーディオ信号の前記基準チャネルの高速フーリエ変換 (FFT) 係数の確率密度関数 (PDF) をモデル化することと、

単一チャネルPDFおよび結合チャネルPDFのうちの少なくとも1つを最大化し、前記基準チャネルの雑音の大きさの推定値と前記一次チャネルの雑音の大きさの推定値との間の弁別性差 (DRD) を提供することと、

所与の周波数に対してどのスペクトルの大きさがより大きいかを決定することと、

前記一次チャネルのスペクトルの大きさが前記基準チャネルのスペクトルの大きさよりも強い場合、前記一次チャネルを強調することと、

前記基準チャネルのスペクトルの大きさが前記一次チャネルのスペクトルの大きさよりも強い場合、前記一次チャネルの強調を抑えることと

を含み、

前記強調することおよび強調を抑えることは、事前段階が存在する場合、乗算リスケーリング係数を算出し、音声増強フィルタチェーンの事前段階において算出された利得に前記乗算リスケーリング係数を適用することと、いかなる事前段階も存在しない場合、利得を直接適用することとを含む、オーディオデバイス。

【請求項16】

オーディオデバイスであって、

オーディオ信号を受信し、前記オーディオ信号の一次チャネルを通信するための一次マイクロホンと、

前記オーディオ信号を前記一次マイクロホンとは異なる状況で受信し、前記オーディオ信号の基準チャネルを通信するための基準マイクロホンと、

前記オーディオ信号をフィルタ処理および/または明瞭化するために前記オーディオ信号を処理する少なくとも1つの処理要素であって、

前記少なくとも1つの処理要素は、方法を実施するためのプログラムを実行するように構成され、

前記方法は、

オーディオデバイスの一次マイクロホンを用いて、オーディオ信号の一次チャネルを取得することと、

前記オーディオデバイスの基準マイクロホンを用いて、前記オーディオ信号の基準チャネルを取得することと、

複数の周波数ビンに対する前記オーディオ信号の前記一次チャネルのスペクトルの大きさを推定することと、

複数の周波数ビンに対する前記オーディオ信号の前記基準チャネルのスペクトルの大きさを推定することと、

一次分数変換および高次有理関数変換のうちの少なくとも1つを適用することによって、1つ以上の周波数ビンに対する前記スペクトルの大きさのうちの1つ以上のものを変換することと、

前記スペクトルの大きさのうちの1つ以上のものを再正規化すること、

10

20

30

40

50

前記スペクトルの大きさのうちの1つ以上のものを累乗すること、
 前記スペクトルの大きさのうちの1つ以上のものを時間平滑化すること、
 前記スペクトルの大きさのうちの1つ以上のものを周波数平滑化すること、
 前記スペクトルの大きさのうちの1つ以上のものをVADベースで平滑化すること、
 前記スペクトルの大きさのうちの1つ以上のものを心理音響的に平滑化すること、
 位相差の推定値を前記変換されたスペクトルの大きさのうちの1つ以上のものと組み
 合わせることを、および、

VAD推定値を前記変換されたスペクトルの大きさのうちの1つ以上のものと組み合
 わせること

のうちの1つ以上によって、1つ以上の周波数ビンに対する前記スペクトルの大きさの
 うちの1つ以上のものを変換することと

を含む、オーディオデバイス。

【発明の詳細な説明】

【技術分野】

【0001】

(関連出願の引用)

本願は、米国仮出願第62/078,844号(2014年11月12日出願、名称「
 Adaptive Interchannel Discriminative Rescaling Filter」)に対する優先権を主張し、上記出願は、その全体が参照
 により本明細書に引用される。

【0002】

(技術分野)

本開示は、概して、音声データを隔離すること、オーディオ信号から雑音を除去するこ
 と、または別様にオーディオ信号を出力することに先立ってオーディオ信号を増強するこ
 とを行うための技法を含むオーディオ信号を処理する技法に関する。オーディオ信号を処
 理するための装置およびシステムも、開示される。

【背景技術】

【0003】

最新技術のモバイル電話を含む種々のオーディオデバイスは、意図された源からオーデ
 ィオを受信するように位置付けられ、向けられる一次マイクロホンと、意図された源から
 背景雑音を受信する一方、オーディオを殆どまたは全く受信しないように位置付けられ、
 向けられる基準マイクロホンとを含む。多くの使用シナリオでは、基準マイクロホンは、
 一次マイクロホンによって取得されるオーディオ信号の一次チャンネルに存在する可能性が
 高い雑音の量のインジケータを提供する。特に、一次チャンネルと基準チャンネルとの間の所
 与の周波数帯域に対する相対スペクトル電力レベルは、その周波数帯域が一次チャンネルに
 おいて雑音によって支配されているか、または信号によって支配されているかを示し得る。
 次いで、その周波数帯域における一次チャンネルオーディオは、適宜、選択的に抑制また
 は増強され得る。

【0004】

しかしながら、一次チャンネルと基準チャンネルとの間の修正されていない相対スペクトル
 電力レベルの関数と考えられる一次チャンネルにおける音声(それぞれ、雑音)支配の確率
 は、周波数ビンによって変動し得、経時的に固定されていない場合があることが事実であ
 る。したがって、チャンネル間の比較ベースのフィルタ処理における生の電力比、固定され
 た閾値、および/または固定されたリスケーリング係数の使用は、一次チャンネルオーデ
 ィオにおける望ましくない音声抑制および/または雑音増幅をよくもたらし得る。

【0005】

故に、入力チャンネル間の雑音支配/音声支配電力レベルにおける差異を推定すること、
 一次入力チャンネルにおいて雑音を抑制し、音声存在を増強することにおける改良が、追求
 される。

【発明の概要】

10

20

30

40

50

【課題を解決するための手段】

【0006】

本発明の一側面は、いくつかの実施形態では、オーディオ信号を変換する方法を特徴とする。方法は、オーディオデバイスの一次マイクロホンを用いて、オーディオ信号の一次チャンネルを取得することと、オーディオデバイスの基準マイクロホンを用いて、オーディオ信号の基準チャンネルを取得することと、複数の周波数ビンに対するオーディオ信号の一次チャンネルのスペクトルの大きさを推定することと、複数の周波数ビンに対するオーディオ信号の基準チャンネルのスペクトルの大きさを推定することを含む。方法はさらに、一次分数変換および高次有理関数変換のうち少なくとも1つを適用することによって、1つ以上の周波数ビンに対するスペクトルの大きさのうち1つ以上のものを変換することと、1つ以上の周波数ビンに対するスペクトルの大きさのうち1つ以上のものをさらに変換することを含む。さらなる変換は、スペクトルの大きさのうち1つ以上のものを再正規化することと、スペクトルの大きさのうち1つ以上のものを累乗することと、スペクトルの大きさのうち1つ以上のものを時間平滑化することと、スペクトルの大きさのうち1つ以上のものを周波数平滑化することと、スペクトルの大きさのうち1つ以上のものをVADベースで平滑化することと、スペクトルの大きさのうち1つ以上のものを心理音響的に平滑化することと、位相差の推定値を変換されたスペクトルの大きさのうち1つ以上のものと組み合わせることと、VAD推定値を変換されたスペクトルの大きさのうち1つ以上のものと組み合わせることとを含むことができる。

10

20

【0007】

いくつかの実施形態では、方法は、増加的入力に基づいて、一次分数変換および高次有理関数変換のうち少なくとも1つをピン毎に更新することを含む。

【0008】

いくつかの実施形態では、方法は、事前SNR推定値および事後SNR推定値のうち少なくとも1つを、変換されたスペクトルの大きさのうち1つ以上のものと組み合わせることを含む。

【0009】

いくつかの実施形態では、方法は、信号電力レベル差(SPLD)データを、変換されたスペクトルの大きさのうち1つ以上のものと組み合わせることを含む。

30

【0010】

いくつかの実施形態では、方法は、雑音の大きさの推定値および雑音電力レベル差(NPLD)に基づいて、基準チャンネルの補正されたスペクトルの大きさを計算することを含む。いくつかの実施形態では、方法は、雑音の大きさの推定値およびNPLDに基づいて、一次チャンネルの補正されたスペクトルの大きさを計算することを含む。

【0011】

いくつかの実施形態では、方法は、スペクトルの大きさのうち1つ以上のものをフレーム内の近傍の周波数ビンにわたりとられる加重平均に置き換えることと、スペクトルの大きさのうち1つ以上のものを前のフレームからの対応する周波数ビンにわたりとられる加重平均に置き換えることとを含む。

40

【0012】

本発明の別の側面は、いくつかの実施形態では、オーディオ信号に適用されるフィルタ処理の程度を調節する方法を特徴とする。方法は、オーディオデバイスの一次マイクロホンを用いて、オーディオ信号の一次チャンネルを取得することと、オーディオデバイスの基準マイクロホンを用いて、オーディオ信号の基準チャンネルを取得することと、オーディオ信号の一次チャンネルのスペクトルの大きさを推定することと、オーディオ信号の基準チャンネルのスペクトルの大きさを推定することを含む。方法はさらに、オーディオ信号の一次チャンネルの高速フーリエ変換(FFT)係数の確率密度関数(PDF)をモデル化することと、オーディオ信号の基準チャンネルの高速フーリエ変換(FFT)係数の確率密度関数(PDF)をモデル化することと、基準チャンネルの雑音の大きさの推定値と一次チャネ

50

ルの雑音の大きさの推定値との間の弁別的関連性差 (DRD) を提供するために、単一チャンネル P D F および結合チャンネル P D F のうちの少なくとも 1 つを最大化することと、所与の周波数に対してどのスペクトルの大きさがより大きいかを決定することを含む。方法はさらに、一次チャンネルのスペクトルの大きさが基準チャンネルのスペクトルの大きさよりも強いとき、一次チャンネルを強調することと、基準チャンネルのスペクトルの大きさが一次チャンネルのスペクトルの大きさよりも強いとき、一次チャンネルの強調を抑えることとを含み、強調することおよび強調を抑えることは、事前段階が存在する場合、乗算リスケーリング係数を算出し、音声増強フィルタチェーンの事前段階において算出された利得に乗算リスケーリング係数を適用することと、いかなる事前段階も存在しない場合、利得を直接適用することを含む。

10

【0013】

いくつかの実施形態では、乗算リスケーリング係数は、利得として使用される。

【0014】

いくつかの実施形態では、方法は、一次および基準オーディオチャンネルのうちの少なくとも 1 つの各スペクトルフレームに増加的入力を含めることを含む。

【0015】

いくつかの実施形態では、増加的入力は、一次チャンネルに対するスペクトルフレームの各ピンにおける事前 S N R および事後 S N R の推定値を含む。いくつかの実施形態では、増加的入力は、一次チャンネルおよび基準チャンネルに対するスペクトルフレームの対応するピン間のピンあたり N P L D の推定値を含む。いくつかの実施形態では、増加的入力は、一次チャンネルおよび基準チャンネルに対するスペクトルフレームの対応するピン間のピンあたり S P L D の推定値を含む。いくつかの実施形態では、増加的入力は、一次チャンネルと基準チャンネルとの間のフレームあたり位相差の推定値を含む。

20

【0016】

本発明の別の側面は、いくつかの実施形態では、オーディオ信号を受信し、オーディオ信号の一次チャンネルを通信するための一次マイクロホンと、オーディオ信号を一次マイクロホンとは異なる状況で受信し、オーディオ信号の基準チャンネルを通信するための基準マイクロホンと、オーディオ信号を処理し、オーディオ信号をフィルタ処理および/または明瞭化するための少なくとも 1 つの処理要素であって、本明細書に説明される方法のいずれかを行うためのプログラムを実行するように構成される、少なくとも 1 つの処理要素と

30

【0017】

本発明のより完全な理解が、図と併せて考慮されるとき、発明を実施するための形態を参照することによってもたらされ得る。

【図面の簡単な説明】

【0018】

【図1】図1は、一実施形態による、適応チャンネル間弁別的リスケーリングフィルタプロセスを例証する。

【図2】図2は、一実施形態による、適応チャンネル間弁別的リスケーリングフィルタプロセスにおいて使用するための入力変換を例証する。

40

【図3】図3は、一実施形態による、雑音および音声電力レベルの比較を例証する。

【図4】図4は、一実施形態による、雑音および音声電力レベル確率分布関数の推定を例証する。

【図5】図5は、一実施形態による、雑音および音声電力レベルの比較を例証する。

【図6】図6は、一実施形態による、雑音および音声電力レベル確率分布関数の推定を例証する。

【図7】図7は、一実施形態による、雑音および音声電力レベルと弁別的利得関数の推定値との比較を例証する。

【図8】図8は、デジタルオーディオデータを分析するためのコンピュータアーキテクチャを例証する。

50

【発明を実施するための形態】

【0019】

以下の説明は、本発明の例示的实施形態にすぎず、本発明の範囲、可用性、または構成を限定することは意図されない。むしろ、以下の説明は、本発明の種々の実施形態を実装するための便宜な例証を提供することが意図される。明白になるであろうように、種々の変更が、本明細書に記載されるような本発明の範囲から逸脱することなく、これらの実施形態に説明される要素の機能および配列において成され得る。したがって、本明細書における発明を実施するための形態は、限定ではなく、例証のみを目的として提示される。

【0020】

本明細書における「一実施形態」または「ある実施形態」の言及は、実施形態と関連して説明される特定の特徵、構造、または特性が、本発明の少なくともある実施形態に含まれること示すことが意図される。本明細書の種々の箇所における語句「一実施形態では」または「ある実施形態」の出現は、必ずしも、全てが同一の実施形態を指すわけではない。

10

【0021】

本発明は、デジタルデータを分析する方法、システム、およびコンピュータプログラム製品に及ぶ。分析されるデジタルデータは、例えば、デジタルオーディオファイル、デジタルビデオファイル、リアルタイムオーディオストリーム、およびリアルタイムビデオストリーム等の形態であり得る。本発明は、デジタルデータのソースにおけるパターンを識別し、識別されたパターンを使用し、デジタルデータを分析、分類、およびフィルタ処理し、例えば、音声データを隔離または増強する。本発明の特定の实施形態は、デジタルオーディオに関する。任意のオーディオソースからの非破壊オーディオ隔離および分離を実施するための実施形態が、設計される。

20

【0022】

適応チャンネル間弁別リスケーリング(AIDR)フィルタの目的は、一次スペクトル Y_1 および基準スペクトル Y_2 の関連性調節相対電力レベルに基づいて、雑音からの電力よりも所望される信号からの電力をより多く含むと推測される一次マイクロホンからの入力のスペクトル表現のフィルタ処理の程度を調節することである。基準マイクロホンからの入力、所望される信号からよりも交絡雑音からの関連性調節電力をより多く含むと推測される。

30

【0023】

二次マイクロホン入力が一次マイクロホン入力よりも多くの音声を含む傾向がある(例えば、ユーザが電話を逆転された向きにおいて保持している)ことが検出される場合、 Y_1 および Y_2 の相対的大きさに関する期待値も、逆転されるであろう。次いで、以下の説明では、 Y_1 および Y_2 等の役割は、利得修正が Y_1 に適用され続け得ることを除いて、単純に置換される。

【0024】

AIDRフィルタの論理は、大まかに言えば、所与の周波数に対して、基準入力が一次入力よりも強いとき、一次入力における対応するスペクトルの大きさは、信号よりも雑音を表し、抑制されるべきである(または少なくとも強調されない)。基準および一次入力の相対強度が逆転されると、一次入力における対応するスペクトルの大きさは、雑音よりも信号を表し、強調されるべきである(または少なくとも抑制されない)。

40

【0025】

しかしながら、雑音抑制/音声増強文脈に関連する様式において、一次入力の所与のスペクトル成分が、実際には基準チャンネルにおけるその対応物よりも「強い」かどうかを正確に決定することは、典型的には、一次および基準スペクトル入力の一方または両方が、好適な形態にアルゴリズム的に変換されることを要求する。変換に続いて、フィルタ処理および雑音抑制が、一次入力チャンネルのスペクトル成分の弁別リスケーリングを介して行われる。この抑制/増強は、典型的には、音声増強フィルタチェーンの事前段階において算出された利得に適用されるべき乗算リスケーリング係数を算出することによって達成

50

されるが、リスケーリング係数は、パラメータの適切な選定によって利得自体としても使用され得る。

【0026】

(1 フィルタ入力)

A I D Rフィルタの多段階推定および弁別プロセスの図式的概観が、図1に提示される。一次および二次(基準)マイクロホンからの時間領域信号 y_1 、 y_2 が、A I D Rフィルタの上流でサンプルの等しい長さのフレーム $y_i(s, t)$ に処理されていると仮定され、 $i \in \{1, 2\}$ であり、 $s = 0, 1, \dots$ は、フレーム内のサンプル指数であり、 $t = 0, 1, \dots$ は、フレーム指数である。これらのサンプルは、フーリエ変換を介してスペクトル領域にさらに変換されており、したがって、 $y_i \rightarrow Y_i$ であり、 $Y_i(k, m)$ は、 m 番目のスペクトルフレームの k 番目の離散周波数成分(「ビン」)を示し、 $k = 1, 2, \dots, K$ であり、 $m = 0, 1, \dots$ である。スペクトルフレームあたりの周波数ビンの数 K は、典型的には、時間領域におけるサンプリング率に従って決定され、例えば、16 kHzのサンプリング率に対して512ビンであることに留意されたい。 $Y_1(k, m)$ および $Y_2(k, m)$ は、A I D Rフィルタに必要な入力であると見なされる。

10

【0027】

A I D Rフィルタが、他の処理構成要素に続く音声増強フィルタチェーンに組み込まれる場合、追加の情報を伝える増加的入力が、各スペクトルフレームに加わり得る。(異なるフィルタ変形において使用される)特定の例示的着目入力は、以下を含む。

20

1. 一次信号に対するスペクトルフレームの各ビンにおける事前 $SNR(k, m)$ および事後 $SNR(k, m)$ の推定値。これらの値は、典型的には、前の統計的フィルタ処理段階、例えば、M M S E、電力レベル差(P L D)等によって算出されているであろう。これらは、 Y_i と同一の長さのベクトル入力である。

2. 一次および二次信号に対するスペクトルフレームの対応するビン間のビンあたり雑音電力レベル差(N P L D)である $N P L D(k, m)$ の推定値。これらの値は、P L Dフィルタによって算出されているであろう。これらは、 Y_i と同一の長さのベクトル入力である。

3. 一次および二次信号に対するスペクトルフレームの対応するビン間のビンあたり音声電力レベル差(S P L D)である、 $S P L D(k, m)$ の推定値。これらの値は、P L Dフィルタによって算出されるであろう。これらは、 Y_i と同一の長さのベクトル入力である。

30

4. 前の音声活動検出(V A D)段階によって算出される、一次および二次信号における音声存在の確率である、 S_1 および/または S_2 の推定値。スカラー $S_i \in [0, 1]$ であると仮定される。

5. 好適な事前処理段階、例えば、P H A T(位相変換)、G C C - P H A T(位相変換との一般化交差相関)等によって提供されるような、 m 番目のフレームにおける一次および基準入力のスペクトル間の位相角分離である、 $\phi(k, m)$ の推定値。

【0028】

(2 段階1 a: 入力変換)

必要な入力 Y_i は、まもなく説明されるであろうような弁別リスケーリング(段階2)において使用するための単一ベクトルに組み合わせられる。A I D Rフィルタの入力変換および組み合わせプロセスの拡大図が、図2に提示される。この組み合わせプロセスは、必ずしも、大きさ $Y_i(k, m)$ に直接作用するわけではなく、むしろ、生の大きさは、最初に、より好適な表現

40

【数1】

$$\bar{Y}_i(k, m)$$

に変換され得、それは、例えば、時間および周波数間変動を平滑化すること、または周波数依存性様式において大きさを再重みづけ/リスケールを行うように作用する。

50

【 0 0 2 9 】

プロトタイプの変換（「段階 1 事前処理」）は、以下を含む。

1 . 大きさの再正規化、例えば、

【数 1 0】

$$\bar{Y}_i(k, m) = \frac{Y(k, m)}{\sum_{i=1}^K Y(k, m)}$$

2 . ある電力への大きさの引き上げ、すなわち

【数 1 0 - 1】

10

$$\bar{Y}_i(k, m) = Y_i(k, m)^{p_i}$$

である。 p_i は、負数であり得、必ずしも、整数値ではない場合があり、 p_1 は、 p_2 に等しくない場合があることに留意されたい。適切に選定された p_i に対して、そのような変換の 1 つの効果は、所与のフレーム内のスペクトルピークを引き上げ、かつスペクトルトラフを平坦にすることによって、差異を強調することであり得る。

3 . フレーム内の近傍の周波数ビンにわたりとられる加重平均による大きさの置き換え。この変換は、周波数における局所平滑化を提供し、すでに FFT の大きさを編集している場合がある事前処理ステップにおいて導入されている場合がある音楽雑音の悪影響を低減させることに役立ち得る。例として、大きさ $Y(k, m)$ は、

20

【数 1 1】

$$\bar{Y}(k, m) = \begin{cases} \frac{\sum_{i=-1}^1 w_k(i+1)Y(k+i, m)}{\sum_{j=1}^3 w_k(j)} & k = 2, \dots, K-1 \text{ の場合} \\ \frac{\sum_{i=0}^1 w_k(i+1)Y(k+i, m)}{\sum_{j=2}^3 w_k(j)} & k = 1 \text{ の場合} \\ \frac{\sum_{i=-1}^0 w_k(i+1)Y(k+i, m)}{\sum_{j=1}^2 w_k(j)} & k = K \text{ の場合} \end{cases}$$

を介して、その値および隣接する周波数ビンの大きさの値の加重平均に置き換えられ得、式中、 $w_k = (1, 2, 1)$ は、周波数ビン重みのベクトルである。下付き文字 k は、局所平均に対する重みベクトルが異なる周波数に対して異なり得る（例えば、低周波数に対してより狭く、高周波数に対してより広い）可能性を表すために、 w に対して含まれる。重みベクトルは、 k 番目の（中央の）ビンに対して対称的である必要はない。例えば、それは、中央のビンの（ビン指数および対応する周波数の両方の）上方のビンをより重く重みづけするために非対称にされ得る。これは、有声音声中、基本周波数およびその高調波の近傍のビンに重点を置くために、有用であり得る。

30

40

4 . 前のフレームからの対応するビンにわたりとられる加重平均による大きさの置き換え。この変換は、各周波数ビン内の時間平滑化を提供し、すでに FFT の大きさを編集している場合がある事前処理ステップにおいて導入されている場合がある音楽雑音の悪影響を低減させることに役立ち得る。時間平滑化は、種々の方法において実装され得る。例えば、

a) 単純な加重平均化

【数 1 2】

$$\bar{Y}(k, m) = \begin{cases} \frac{\sum_{i=-2}^0 w_k(i+3)Y(k, m+i)}{\sum_{j=1}^3 w_k(j)}, w = (1, 2, 3) & m \geq 2 \text{ の場合} \\ \frac{\sum_{i=-1}^0 w_k(i+2)Y(k, m+i)}{\sum_{j=2}^3 w_k(j)} & m = 1 \text{ の場合} \\ Y(k, m) & m = 0 \text{ の場合} \end{cases}$$

b) 指数平滑化

10

【数 1 3】

$$\bar{Y}(k, m) = \begin{cases} Y(k, m) & m = 0 \text{ の場合} \\ \beta Y(k, m) + (1 - \beta)\bar{Y}(k, m - 1) & m > 0 \text{ の場合} \end{cases}$$

である。ここは、 $[0, 1]$ は、前のフレームに対する現在のフレームからのピンの大きさの相対的重みづけを決定する平滑化パラメータである。

5. VADベースの重みづけを用いた指数平滑化。音声情報を含む/含まないそれらの前のフレームのみからのピンの大きさが含まれる時間平滑化を実施することも、有用であり得る。これは、事前信号処理段階によって算出される十分に正確なVAD情報(増加的入力)を要求する。VAD情報は、以下のように指数平滑化に組み込まれ得る。

20

a)

【数 1 4】

$$\bar{Y}(k, m) = \begin{cases} Y(k, m) & m = 0 \text{ の場合} \\ \beta Y(k, m) + (1 - \beta)\bar{Y}(k, m^*) & m > 0 \text{ の場合} \end{cases}$$

この変形では、 $m^* < m$ は、 $S_i(m^*)$ が音声存在/不在を示す規定された閾値を上回る(または下回る)ような最も近い前のフレームの指数である。

30

b) 代替として、音声存在の確率は、平滑化率を直接修正するために使用され得る。

【数 1 5】

$$\bar{Y}(k, m) = \begin{cases} Y(k, m) & m = 0 \text{ の場合} \\ \beta(S_i)Y(k, m) + (1 - \beta(S_i))\bar{Y}(k, m^*) & m > 0 \text{ の場合} \end{cases}$$

この変形では、 S_i は、 S_i の関数、例えば、シグモイド関数であり、パラメータは、 S_i が所与の閾値を下回って(上回って)移動するにつれて、 (S_i) が固定値 a (b) に接近するように選定される。

40

6. 心理音響的重要性による再重みづけ：メル周波数およびERBスケール重みづけ。

【0030】

上記の段階のいずれかおよび/または全ては、組み合わせられ得る、またはいくつかの段階は、省略され得、それらのそれぞれのパラメータは、用途(例えば、モバイル電話ではなく、自動音声認識のために使用されるメルスケール再重みづけ)に従って調節されることに留意されたい。

【0031】

(3 段階 1 b : 適応入力組み合わせ)

フレーム指数 m に対する入力変換段階の最終出力は、 $u(m)$ と指定される。 $u(m)$

50

は、 Y_i と同一の長さ K を有するベクトルであり、 $u(k, m)$ は、 m 番目のスペクトルフレームの k 番目の離散周波数成分に関連付けられる u の成分を示すことに留意されたい。 $u(m)$ の算出は、修正された必要な入力

【数 15 - 1】

$$\bar{Y}_1, \bar{Y}_2$$

を要求し、一般的形態では、これは、ベクトル値関数

【数 15 - 2】

$$f: \mathbb{R}^{2K} \rightarrow \mathbb{R}^K, f(\bar{Y}_1(m), \bar{Y}_2(m)) = \mathbf{u}(m).$$

10

によって遂行される。

【0032】

その最も単純な実装では、

【数 15 - 3】

$$\bar{Y}_1(k, m), \bar{Y}_2(k, m)$$

に対する f のピンあたり作用は、一次分数変換として表され得る：

【数 16】

$$f_k(\bar{Y}_1(k, m), \bar{Y}_2(k, m)) = u(k, m) = \frac{A_k \bar{Y}_1(k, m) + B_k}{C_k \bar{Y}_2(k, m) + D_k}$$

20

【0033】

一般性を失うことなく、より大きい値の $u(k, m)$ が、 k 番目の周波数ピンにおいて、時間指数 m において交絡雑音からよりも所望される信号からより多くの電力があることを示すと推測され得る。

【0034】

より一般的には、 f_k の分子および分母が、代わりに、

【数 16 - 1】

$$\bar{Y}_1(k, m), \bar{Y}_2(k, m)$$

30

において高次有理式を伴い得る：

【数 17】

$$f_k(\bar{Y}_1(k, m), \bar{Y}_2(k, m)) = u(k, m) = \frac{\sum_{i=0}^Q A_{i,k} (\bar{Y}_1(k, m))^i}{\sum_{j=0}^R C_{j,k} (\bar{Y}_2(k, m))^j}$$

【0035】

さらに、任意の区分的平滑変換が、この一般的表現（チザム近似）を用いて任意の所望される正確度内で表され得る。加えて、変換パラメータ（これらの例における A_k 、 B_k 、 C_k 、 D_k 、または $A_{i,k}$ 、 $C_{j,k}$ ）は、周波数ピンによって変動し得る。例えば、予期される雑音電力特性がより低い周波数とより高い周波数とにおいて異なる場合、より低い周波数とより高い周波数とにおけるピンに対して異なるパラメータを使用することが有用であり得る。

40

【0036】

実践では、 f_k のパラメータは、固定されず、むしろ、増加的入力に基づいてフレーム毎に更新され、例えば、

【数 101】

$$B_k = B_k(\alpha_{NPLD}(k, m), \xi(k, m), \eta(k, m), S_1(m), \Delta\phi(m)), \quad (1)$$

50

【数 1 0 2】

$$D_k = D_k(\alpha_{NPLD}(k, m), S_1(m), \Delta\phi(m)) \quad (2)$$

または、

【数 1 0 3】

$$A_{i,k} = A_{i,k}(\alpha_{NPLD}(k, m), \xi(k, m), \eta(k, m), S_1(m), \Delta\phi(m)), \quad (3)$$

【数 1 0 4】

$$C_{j,k} = C_{j,k}(\alpha_{NPLD}(k, m), S_1(m), \Delta\phi(m)) \quad (4)$$

10

等である。

【0 0 3 7】

生の入力 $Y_1(k, m)$, $Y_2(k, m)$ に対する調節は、入力 $Y_1(k, m)$ のどの成分が所望される信号に主として関連するかを弁別する目的により関連する量への生のスペクトル電力推定値のビンあたり変換をもたらす。変換は、例えば、一次および/または基準スペクトルにおける相対ピークおよびトラフをリスケーリングすること、スペクトル過渡を平滑化（または鮮鋭化）すること、ならびに/または一次マイクロホンと基準マイクロホンとの間の向きもしくは空間的分離における差異を補正することを行うように作用し得る。そのような要因は経時的に変化し得るので、変換の関連パラメータは、典型的には、A I D R フィルタがアクティブである間、フレーム毎に 1 回更新される。

20

【0 0 3 8】

(4 段階 2 : 弁別リスケーリング)

第 2 段階の目標は、所望される音声よりも多く雑音を含むと推定されるそれらの $Y_1(k, m)$ の大きさを低減させることによって、一次信号から雑音成分をフィルタ処理することである。段階 1 の出力 $u(m)$ は、この推定値としての役割を果たす。段階 2 の出力を $Y_1(m)$ の各周波数成分に対する乗算利得のベクトルであるとする場合、 k 番目の利得は、 $u(k, m)$ が非常に低い S N R を示すとき、小さく（0 に近似する）、 $u(k, m)$ が非常に高い S N R を示す場合、大きく（1 に近似する、例えば、利得が非構成的であると制限される場合）すべきである。中間的な場合に対して、これらの極の間で漸進遷移であることが望ましい。

30

【0 0 3 9】

一般的に言って、フィルタの第 2 のステップでは、ベクトル u は、小さい値 u_k が小さい値 w_k にマッピングされ、大きい値 u_k がより大きい非負値 w_k にマッピングされるような方式で、ベクトル w に区分的に平滑に変換される。ここで、 k は、周波数ビン指数を示す。この変換は、 $g(u) = w$ を与えるベクトル値関数

【数 1 0 4 - 1】

$$g : \mathbb{R}^N \rightarrow \mathbb{R}^N$$

を介して達成される。要素毎の g は、非負区分的平滑関数

【数 1 0 4 - 2】

$$g_k : \mathbb{R} \rightarrow \mathbb{R}$$

40

によって説明される。ある有限 B_k に対して、 $0 \leq w_k \leq B_k$ であれば、 g は、有界でなくとも、非負でなくともよい。しかしながら、各 g_k は、妥当な範囲の入力 u_k にわたって有限かつ非負であるべきである。

【0 0 4 0】

g のプロトタイプの例は、各座標における単純なシグモイド関数

【数 1 0 4 - 3】

$$g_k(u(k, m)) = w_k = \alpha_k + \frac{\beta_k}{1 + \exp(-\delta_k u(k, m) - \epsilon_k)}$$

を特徴とする。

【0 0 4 1】

一般化ロジスティック関数は、より柔軟性がある：

【数 1 0 4 - 4】

$$g_k(u(k, m)) = w_k = \alpha_k + \frac{\beta_k - \alpha_k}{(1 + \gamma_k \exp(-\delta_k(u(k, m) - \mu_k)))^{\nu_k}}$$

10

【0 0 4 2】

パラメータ α_k は、 w_k に対する最小値を設定する。これは、典型的には、 $Y(k, m)$ の全体的抑制を回避するために、小さい正值、例えば、0.1 であるように選定される。

【0 0 4 3】

パラメータ β_k は、 w_k に対する最大値の一次決定因子であり、それは、概して、1 に設定され、それによって、高 SNR 成分は、フィルタによって修正されない。しかしながら、いくつかの用途に対して、 β_k は、1 よりもわずかに大きくされ得る。AIDR が、例えば、より大きいフィルタ処理アルゴリズムにおける後処理構成要素として使用され、事前フィルタ処理段階が一次信号を（全体的に、または特定の周波数帯域において）減衰させる傾向にあるとき、 $\beta_k > 1$ が、前に抑制されたいいくつかの音声成分を復元するように作用し得る。

20

【0 0 4 4】

$u(k, m)$ 値の過渡的な中間範囲内の g_k の出力は、最大勾配の程度、横座標、および縦座標を制御する、パラメータ α_k 、 β_k 、および μ_k によって決定される。

【0 0 4 5】

これらのパラメータの初期値は、広い範囲の雑音条件下の種々の話者に対する $u(k, m)$ 値の分布を調べ、 $u(k, m)$ 値を雑音および音声の相対電力レベルと比較することによって決定される。これらの分布は、混合 SNR および雑音タイプによって実質的に変動し得、すなわち、話者間の変動は、少ない。（心理音響 / 周波数）帯域間にも明確な差異が存在する。種々の周波数帯域内の雑音対音声電力レベルに対する確率分布の実施例が、図 3 - 6 に示される。

30

【0 0 4 6】

そのように取得された経験的曲線は、一般化ロジスティック関数によって良好に合致される。一般化ロジスティック関数は、最良適合を提供するが、単純なシグモイドが、多くの場合、適切である。図 7 は、経験的確率データに対する基本シグモイド関数および一般化ロジスティック関数適合を示す。単一「最良」パラメータセットが、多くの話者および雑音タイプを集約することによって見出されることができ、またはパラメータセットが、具体的話者および雑音タイプに適合され得る。

40

【0 0 4 7】

（5 補記）

便宜上、

【数 1 0 4 - 5】

$$\bar{u}(k, m)$$

が、段階 2 の（一般化）ロジスティック関数において $u(k, m)$ に代入され得る。これは、数桁を上回って及び得る値をはるかに小さい間隔に集中させる効果を及ぼす。しかしながら、同一の最終結果が、対数を使用するパラメータ値のリスケーリングおよび代数再結合によって、関数入力の変換をとることに頼らずに達成され得る。

50

【0048】

段階2におけるパラメータ値は、固定された制限内で「デジジョンダイレクテッドベース」で調節され得る。

【0049】

ベクトル w は、一次入力のスベクトルの大きさに適用されるべき乗算利得のスタンドアロンベクトルとして使用され得るか、または、それは、事前フィルタ処理段階において算出された利得に対するスケーリングおよび/もしくはシフト係数として使用され得る。

【0050】

スタンドアロンフィルタが使用されるとき、AIDRフィルタは、事前SNRのアドホック推定値としてのスベクトル電力の修正された相対レベルと、利得関数としてのシグモイド関数とを使用して、基本雑音抑制を提供する。

10

【0051】

本発明の実施形態はまた、デジタルデータを分析するためのコンピュータプログラム製品にも及び得る。そのようなコンピュータプログラム製品は、デジタルデータを分析する方法を実施するために、コンピュータプロセッサ上でコンピュータ実行可能命令を実行することが意図され得る。そのようなコンピュータプログラム製品は、エンコードされたコンピュータ実行可能命令を有するコンピュータ読み取り可能な媒体を備え得、コンピュータ実行可能命令は、好適なコンピュータ環境内の好適なプロセッサ上で実行されると、本明細書にさらに説明されるようなデジタルデータを分析する方法を実施する。

【0052】

本発明の実施形態は、以下にさらに詳細に議論されるように、例えば、1つ以上のコンピュータプロセッサおよびデータ記憶装置もしくはシステムメモリ等のコンピュータハードウェアを含む専用または汎用コンピュータを備えているか、または利用し得る。本発明の範囲内の実施形態はまた、コンピュータ実行可能命令および/またはデータ構造を伝搬もしくは記憶するための物理的および他のコンピュータ読み取り可能な媒体を含む。そのようなコンピュータ読み取り可能な媒体は、汎用または専用コンピュータシステムによってアクセスされ得る任意の利用可能な媒体であり得る。コンピュータ実行可能命令を記憶するコンピュータ読み取り可能な媒体は、コンピュータ記憶媒体である。コンピュータ実行可能命令を伝搬するコンピュータ読み取り可能な媒体は、伝送媒体である。したがって、限定ではなく、例として、本発明の実施形態は、少なくとも2つの明確に異なる種類のコンピュータ読み取り可能な媒体、すなわち、コンピュータ記憶媒体と、伝送媒体とを備えていることができる。

20

30

【0053】

コンピュータ記憶媒体は、RAM、ROM、EEPROM、CD-ROM、もしくは他の光学ディスク記憶装置、磁気ディスク記憶装置もしくは他の磁気記憶デバイス、またはコンピュータ実行可能命令もしくはデータ構造の形態の所望されるプログラムコード手段を記憶するように使用され得、汎用もしくは専用コンピュータによってアクセスされ得る、任意の他の物理的媒体を含む。

【0054】

「ネットワーク」は、コンピュータシステムおよび/またはモジュールおよび/または他の電子デバイス間の電子データの転送を可能にする、1つ以上のデータリンクとして定義される。情報がネットワークまたは別の通信接続（有線、無線、または有線もしくは無線の組み合わせのいずれか）を経由してコンピュータに伝達もしくは提供されると、コンピュータは、適切に、接続を伝送媒体と見なす。伝送媒体は、汎用または専用コンピュータによって受信もしくはアクセスされ得る、コンピュータ実行可能命令および/もしくはデータ構造の形態の所望されるプログラムコード手段を伝搬もしくは伝送するように使用され得る、ネットワークおよび/もしくはデータリンクを含むことができる。上記の組み合わせもまた、コンピュータ読み取り可能な媒体の範囲内に含まれるべきである。

40

【0055】

さらに、種々のコンピュータシステム構成要素に到達すると、コンピュータ実行可能命

50

令またはデータ構造の形態のプログラムコード手段は、伝送媒体からコンピュータ記憶媒体に自動的に伝達されることができ（逆もまた同様である）。例えば、ネットワークまたはデータリンクを経由して受信されるコンピュータ実行可能命令もしくはデータ構造は、ネットワークインターフェースモジュール（例えば、「NIC」）内のRAMにおいてバッファリングされ、次いで、最終的に、コンピュータシステムRAMおよび/またはコンピュータシステムにおけるより揮発性の低いコンピュータ記憶媒体に伝達されることができ。したがって、コンピュータ記憶媒体は、また、（または場合によっては主として）伝送媒体を利用するコンピュータシステム構成要素に含まれ得ることを理解されたい。

【0056】

コンピュータ実行可能命令は、例えば、プロセッサにおいて実行されると、汎用コンピュータ、専用コンピュータ、または専用処理デバイスに、ある機能もしくは機能群を実施させる命令およびデータを含む。コンピュータ実行可能命令は、例えば、プロセッサ上で直接実行され得るバイナリ、アセンブリ言語等の中間フォーマット命令、または特定の機械もしくはプロセッサを標的とするコンパイラによるコンパイルを要求し得るさらに高レベルのソースコードであり得る。本主題は、構造的特徴および/または方法論的行為に特有の言語で説明されたが、添付される請求項に定義される主題は、必ずしも、上記に説明される、説明される特徴または行為に限定されないことを理解されたい。むしろ、説明される特徴および行為は、本請求項を実装する例示的形態として開示される。

【0057】

当業者は、本発明が、パーソナルコンピュータ、デスクトップコンピュータ、ラップトップコンピュータ、メッセージプロセッサ、ハンドヘルドデバイス、マルチプロセッサシステム、マイクロプロセッサベースまたはプログラマブル消費者用電子機器、ネットワークPC、ミニコンピュータ、メインフレームコンピュータ、モバイル電話、PDA、ページャ、ルータ、スイッチ等を含む、多くのタイプのコンピュータシステム構成を伴うネットワークコンピューティング環境において実践され得ることを理解するであろう。本発明はまた、ネットワークを通して（有線データリンク、無線データリンク、または有線および無線データリンクの組み合わせのいずれかによって）リンクされる、ローカルおよびリモートコンピュータシステムが両方ともタスクを実施する、分散システム環境において実践され得る。分散システム環境では、プログラムモジュールが、ローカルおよびリモート両方のメモリ記憶デバイスに位置し得る。

【0058】

図8を参照すると、デジタルオーディオデータを分析するための例示的コンピュータアーキテクチャ600が、例証される。本明細書ではコンピュータシステム600とも称されるコンピュータアーキテクチャ600は、1つ以上のコンピュータプロセッサ602と、データ記憶装置とを含む。データ記憶装置は、コンピューティングシステム600内のメモリ604であり得、揮発性または不揮発性メモリであり得る。コンピューティングシステム600はまた、データまたは他の情報の表示のためのディスプレイ612も備え得る。コンピューティングシステム600はまた、コンピューティングシステム600が、例えば、ネットワーク（おそらくインターネット610等）を経由して他のコンピューティングシステム、デバイス、またはデータソースと通信することを可能にする、通信チャンネル608も含み得る。コンピューティングシステム600はまた、デジタルまたはアナログデータのソースがアクセスされることを可能にする、マイクロホン606等の入力デバイスも備え得る。そのようなデジタルまたはアナログデータは、例えば、オーディオまたはビデオデータであり得る。デジタルまたはアナログデータは、ライブマイクロホン等からのリアルタイムストリーミングデータの形態であり得る、またはコンピューティングシステム600によって直接アクセス可能である、もしくは通信チャンネル608を通して、もしくはインターネット610等のネットワークを介してより遠隔でアクセスされ得る、データ記憶装置614からアクセスされる記憶されたデータであり得る。

【0059】

通信チャンネル608は、伝送媒体の例である。伝送媒体は、典型的には、コンピュータ

10

20

30

40

50

読み取り可能な命令、データ構造、プログラムモジュール、または他のデータを搬送波もしくは他の転送機構等の変調データ信号に具現化し、任意の情報送達媒体を含む。限定ではなく、例として、伝送媒体は、有線ネットワークおよび直接有線接続等の無線媒体、ならびに音響、高周波、赤外線、および他の無線媒体等の無線媒体を含む。本明細書で使用されるような用語「コンピュータ読み取り可能な媒体」は、コンピュータ記憶媒体および伝送媒体を両方とも含む。

【0060】

本発明の範囲内の実施形態はまた、その上に記憶されるコンピュータ実行可能命令またはデータ構造を伝搬もしくは有するためのコンピュータ読み取り可能な媒体を含む。「コンピュータ記憶媒体」と称される、そのような物理的コンピュータ読み取り可能な媒体は、汎用または専用コンピュータによってアクセスされ得る任意の利用可能な物理的媒体であり得る。限定ではなく、例として、そのようなコンピュータ読み取り可能な媒体は、RAM、ROM、EEPROM、CD-ROM、もしくは他の光学ディスク記憶装置、磁気ディスク記憶装置もしくは他の磁気記憶デバイス、またはコンピュータ実行可能命令もしくはデータ構造の形態の所望されるプログラムコード手段を記憶するように使用され得、汎用もしくは専用コンピュータによってアクセスされ得る、任意の他の物理的媒体等の物理的記憶装置および/またはメモリ媒体を含むことができる。

10

【0061】

コンピュータシステムは、例えば、ローカルエリアネットワーク(「LAN」)、広域ネットワーク(「WAN」)、無線広域ネットワーク(「WWAN」)、およびさらにはインターネット110等のネットワークを経由して互いに接続され得る(またはその一部である)。故に、描写されるコンピュータシステムならびに任意の他の接続されるコンピュータシステムおよびそれらの構成要素はそれぞれ、メッセージ関連データを作成し、ネットワークを経由してメッセージ関連データ(例えば、インターネットプロトコル(「IP」)データグラムおよびIPデータグラムを利用する、伝送制御プロトコル(「TCP」)、ハイパーテキスト転送プロトコル(「HTTP」)、または簡易メール転送プロトコル(「SMTP」)等の他の上位層プロトコル)を交換することができる。

20

【0062】

開示される主題の他の側面、ならびにその種々の側面の特徴および利点は、上記に提供される開示、付随の図面、および添付される請求項の考慮を通して、当業者に明白となるはずである。

30

【0063】

前述の開示は多くの詳細を提供しているが、これらは、続く請求項のいずれかの範囲を限定するものとして解釈されるべきではない。請求項の範囲から逸脱しない他の実施形態が、考案され得る。異なる実施形態からの特徴が、組み合わせて採用され得る。

【0064】

最後に、本発明は、種々の例示的实施形態に対して上記に説明されたが、多くの変更、組み合わせ、および修正が、本発明の範囲から逸脱することなく、実施形態に成され得る。例えば、本発明は、音声検出における使用に対して説明されたが、本発明の側面は、他のオーディオ、ビデオ、データ検出スキームに容易に適用され得る。さらに、種々の要素、構成要素、および/またはプロセスが、代替方法において実装され得る。これらの代替は、特定の用途に応じて、または方法もしくはシステムの実装もしくは動作と関連付けられる任意の数の要因を考慮して、好適に選択されることができる。加えて、本明細書に説明される技法は、他のタイプの用途およびシステムと併用するために拡張または修正され得る。これらおよび他の変更または修正は、本発明の範囲内に含まれることが意図される。

40

【 図 1 】

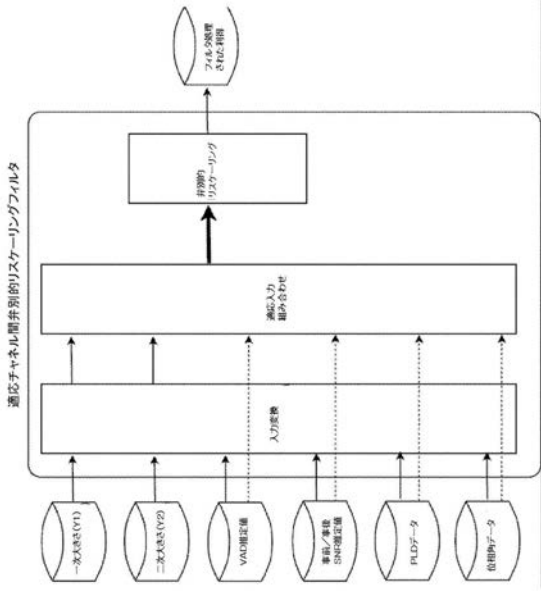


FIGURE 1

【 図 2 】

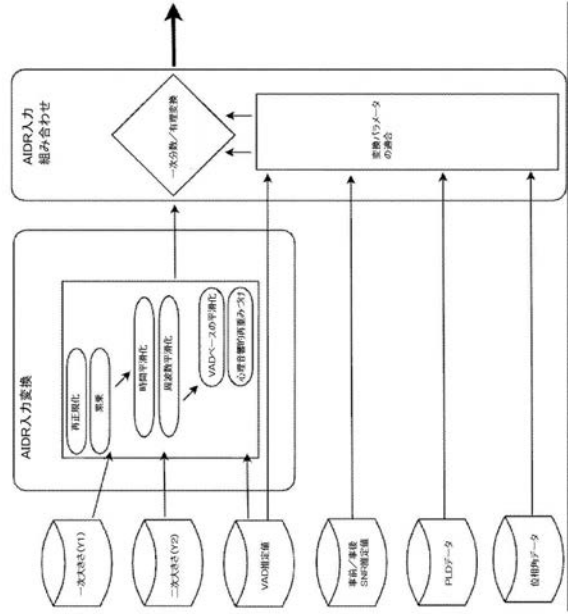


FIGURE 2

【 図 3 】

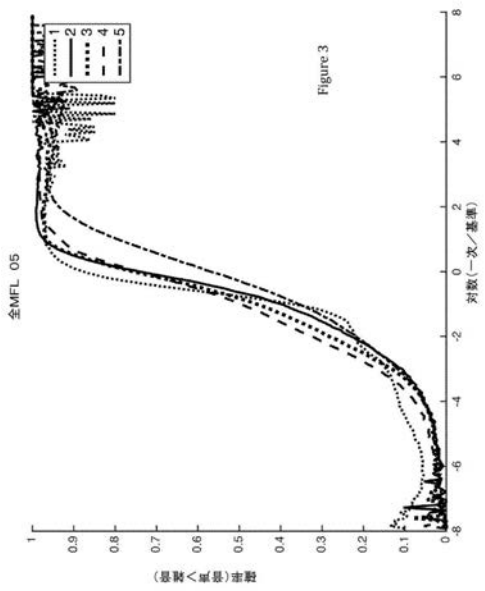


Figure 3

【 図 4 】

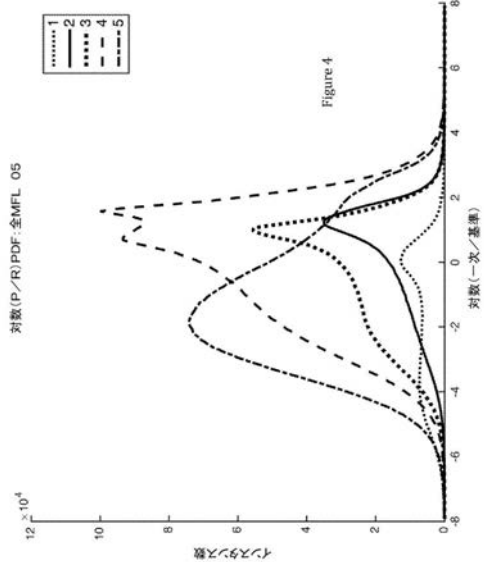
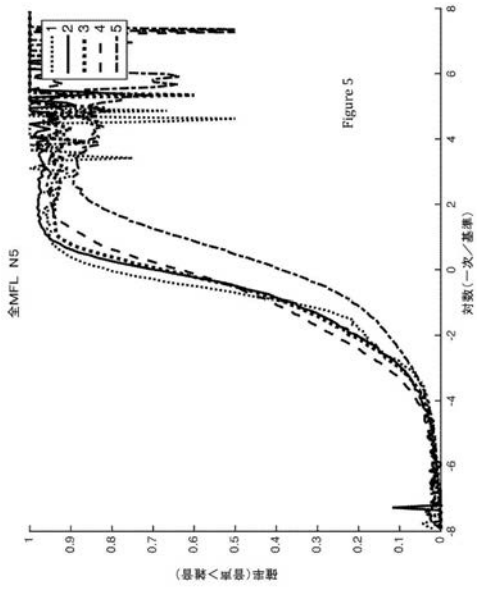
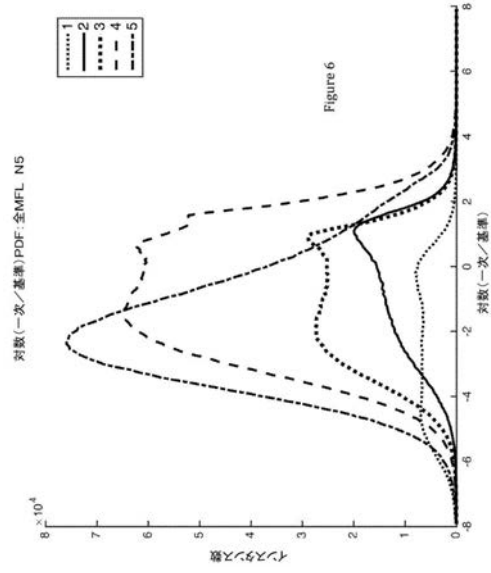


Figure 4

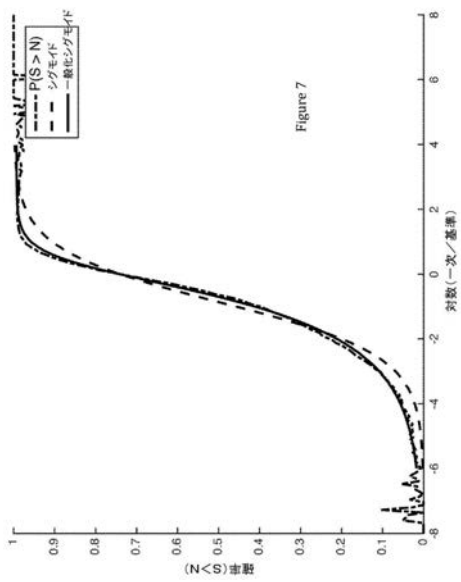
【図5】



【図6】



【図7】



【図8】

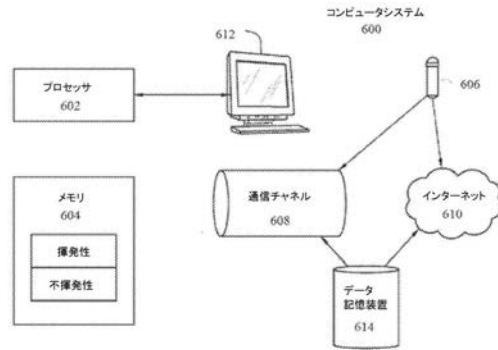


FIGURE 8

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		International application No. PCT/US2015/060337
A. CLASSIFICATION OF SUBJECT MATTER IPC(8) - H04B 15/00 (2015.01) CPC - H04B 15/00 (2015.12) According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC(8) - G10L 21/02; G10L 25/84; H04B 15/00; H04B 1/3827 (2015.01) CPC - G10L 21/02; G10L 25/84; H04B 15/00; H04B 1/3827 (2015.12)		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched US: 381/122, 150; 704/226 (keyword delimited)		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) Orbit, Google Patents, Google Scholar, Google Search terms used: Primary microphone, reference microphone, reference channel, estimating magnitude		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2014/0029762 A1 (XIE et al) 30 January 2014 (30.01.2014) entire document	1-7, 16
Y	Pashe et al. Closed form Filtering for Linear Fractional Transformation Models. Proceedings of the 17th World Congress. IFAC, July 6-11, 2008. [retrieved on 2015-02-17]. Retrieved from the Internet:<URL: http://www.nt.ntnu.no/users/skoge/prost/proceedings/ifac2008/data/papers/2777.pdf >. entire document.	1-7, 16
Y	Lu et al. SPEECH ENHANCEMENT BY COMBINING STATISTICAL ESTIMATORS OF SPEECH AND NOISE. ICASSP 2010. [retrieved on 2015-02-17]. Retrieved from the Internet: <URL: http://www.mirlab.org/conference_papers/International_Conference/ICASSP%202010/pdfs/0004754.pdf>. entire document.	3
Y	US 2003/0206674 A1 (ALTMAN et al) 06 November 2003 (06.11.2003) entire document	4
Y	US 2013/0054231 A1 (JEUB et al) 28 February 2013 (28.02.2013) entire document	5, 6
Y	US 7,171,003 B1 (VENKATESH et al) 30 January 2007 (30.01.2007) entire document	7
A	US 20120226691 A1 (EDWARDS) 06 September 2012 (06.09.2012) entire document	1-7, 16
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search 17 February 2016		Date of mailing of the international search report 07 MAR 2016
Name and mailing address of the ISA/ Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, VA 22313-1450 Facsimile No. 571-273-8300		Authorized officer Blaine R. Copenhaver PCT Helpdesk: 571-272-4300 PCT Q&P: 571-272-7774

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2015/060337

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

See supplemental page

1. As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-7, 16

Remark on Protest

- The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2015/060337

Continued from Box No. III Observations where unity of invention is lacking

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

Group I, claims 1-7, 16, drawn to transforming an audio signal.

Group II, claims 8-15, drawn to adjusting a degree of filtering applied to an audio signal.

The inventions listed as Groups I-II do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: the special technical feature of the Group I invention: exponentiating one or more of the spectral magnitudes; temporal smoothing of one or more of the spectral magnitudes; frequency smoothing of one or more of the spectral magnitudes; VAD-based smoothing of one or more of the spectral magnitudes; psychoacoustic smoothing of one or more of the spectral magnitudes as claimed therein is not present in the invention of Group II. The special technical feature of the Group II invention: modeling a probability density function (PDF) of a fast Fourier transform (FFT) coefficient of the reference channel of the audio signal; maximizing at least one of a single channel PDF and a joint channel PDF to provide a discriminative relevance difference (DRD) between a noise magnitude estimate of the reference channel and a noise magnitude estimate of the primary channel as claimed therein is not present in the invention of Group I.

Groups I and II lack unity of invention because even though the inventions of these groups require the technical feature of obtaining a primary channel of an audio signal with a primary microphone of an audio device; obtaining a reference channel of the audio signal with a reference microphone of the audio device; estimating a spectral magnitude of the primary channel of the audio signal; estimating a spectral magnitude of the reference channel of the audio signal, this technical feature is not a special technical feature as it does not make a contribution over the prior art.

Specifically, US 2012/0226691 A1 (EDWARDS) 06 September 2012 (06.09.2012) teaches obtaining a primary channel of an audio signal with a primary microphone of an audio device (Para. 58); obtaining a reference channel of the audio signal with a reference microphone of the audio device (Patterns within the data, the data matching a pattern, or other data related to the received and interpreted data may be stored or referenced in one or more tables managed by the tables module 252, Para. 65); estimating a spectral magnitude of the primary channel of the audio signal (Plotting or graphically illustrating the three or more dimensions of the data may be used for spectral analysis, para. 106); estimating a spectral magnitude of the reference channel of the audio signal (Para. 102).

Since none of the special technical features of the Group I or II inventions are found in more than one of the inventions, unity of invention is lacking.

フロントページの続き

(81)指定国 AP(BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, RU, TJ, TM), EP(AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US

Fターム(参考) 5D220 BA04 BB03