

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7667247号
(P7667247)

(45)発行日 令和7年4月22日(2025.4.22)

(24)登録日 令和7年4月14日(2025.4.14)

(51)国際特許分類 F I
G 1 0 L 21/0208(2013.01) G 1 0 L 21/0208 1 0 0 B
G 1 0 L 25/84 (2013.01) G 1 0 L 25/84

請求項の数 15 (全18頁)

(21)出願番号	特願2023-505851(P2023-505851)	(73)特許権者	507236292 ドルビー ラボラトリーズ ライセンシング コーポレーション アメリカ合衆国 9 4 1 0 3 カリフォル ニア州 サンフランシスコ マーケット ストリート 1 2 7 5
(86)(22)出願日	令和3年8月2日(2021.8.2)	(74)代理人	100107766 弁理士 伊東 忠重
(65)公表番号	特表2023-536104(P2023-536104 A)	(74)代理人	100070150 弁理士 伊東 忠彦
(43)公表日	令和5年8月23日(2023.8.23)	(74)代理人	100135079 弁理士 宮崎 修
(86)国際出願番号	PCT/US2021/044166	(72)発明者	シュアン,ズーウェイ アメリカ合衆国 9 4 1 0 3 カリフォル ニア州 サンフランシスコ マーケット 最終頁に続く
(87)国際公開番号	WO2022/026948		
(87)国際公開日	令和4年2月3日(2022.2.3)		
審査請求日	令和5年1月27日(2023.1.27)		
(31)優先権主張番号	PCT/CN2020/106270		
(32)優先日	令和2年7月31日(2020.7.31)		
(33)優先権主張国・地域又は機関	中国(CN)		
(31)優先権主張番号	63/068,227		
(32)優先日	令和2年8月20日(2020.8.20)		
	最終頁に続く		

(54)【発明の名称】 機械学習を用いたノイズ削減

(57)【特許請求の範囲】

【請求項1】

コンピュータ実装されるオーディオ処理方法であって、当該方法は：
第1帯域利得およびオーディオ信号の音声活動検出値を、前記オーディオ信号の表現を
入力として取る第1の機械学習モデルを使用して生成することであって、前記オーディオ
信号の前記表現は前記オーディオ信号のスペクトル解析の出力に基づく、ことを実行し；
前記第1帯域利得および前記音声活動検出値に基づいて背景ノイズ推定値を生成し；
前記背景ノイズ推定値によって制御されるウィナー・フィルタを使用して前記オーデ
ィオ信号を処理することによって、第2帯域利得を生成し；
前記第1帯域利得と前記第2帯域利得を組み合わせることによって、組み合わせられた利得
を生成し；
前記組み合わせられた利得を使用して前記オーディオ信号を修正することによって、修正
されたオーディオ信号を生成することを含む、
方法。

【請求項2】

前記第1の機械学習モデルが、トレーニング・データの多様性を増すようデータ増強を
使用して生成される、請求項1に記載の方法。

【請求項3】

前記第1帯域利得を生成することは、少なくとも2つの異なる帯域について少なくとも2
つの異なる制限を使用して前記第1帯域利得を制限することを含む、請求項1または2に

記載の方法。

【請求項 4】

前記背景ノイズ推定値を生成することは、特定の帯域についての閾値を超える、いくつかのノイズ・フレームに基づく、請求項 1 ないし 3 のうちいずれか一項に記載の方法。

【請求項 5】

前記第 2 帯域利得を生成することは、特定の帯域についての定常ノイズ・レベルに基づいて前記ウィナー・フィルタを使用することを含む、請求項 1 ないし 4 のうちいずれか一項に記載の方法。

【請求項 6】

前記第 2 帯域利得を生成することが、少なくとも 2 つの異なる帯域について少なくとも 2 つの異なる制限を使用して前記第 2 帯域利得を制限することを含む、請求項 1 ないし 5 のうちいずれか一項に記載の方法。

10

【請求項 7】

前記組み合わせられた利得を生成することは：

前記第 1 帯域利得と前記第 2 帯域利得を乗算し；

少なくとも 2 つの異なる帯域について少なくとも 2 つの異なる制限を使用して、前記組み合わせられた帯域利得を制限することを含む、請求項 1 ないし 6 のうちいずれか一項に記載の方法。

【請求項 8】

前記修正されたオーディオ信号を生成することは、前記組み合わせられた帯域利得を使用して前記オーディオ信号の振幅スペクトルを修正することを含む、請求項 1 ないし 7 のうちいずれか一項に記載の方法。

20

【請求項 9】

入力オーディオ信号に重複窓を適用して複数のフレームを生成することをさらに含み、前記オーディオ信号が該複数のフレームに対応する、請求項 1 ないし 8 のうちいずれか一項に記載の方法。

【請求項 10】

前記オーディオ信号に対してスペクトル解析を実行し、前記オーディオ信号の複数のピン特徴および基本周波数を生成することをさらに含み、

前記第 1 帯域利得および前記音声活動検出値は、前記複数のピン特徴および前記基本周波数に基づく、請求項 1 ないし 9 のうちいずれか一項に記載の方法。

30

【請求項 11】

前記複数のピン特徴に基づいて複数の帯域特徴を生成し、前記複数の帯域特徴は、メル周波数ケプストラム係数およびバーク周波数ケプストラム係数の一方を使用して生成され、

前記第 1 帯域利得および前記音声活動検出値は、前記複数の帯域特徴および前記基本周波数に基づく、請求項 10 に記載の方法。

【請求項 12】

前記組み合わせられた利得は、前記オーディオ信号の複数の帯域に関連する組み合わせられた帯域利得であり、当該方法は、さらに：

前記組み合わせられた帯域利得を組み合わせられたピン利得に変換することを含み、前記組み合わせられたピン利得は複数のピンに関連する、請求項 1 ないし 11 のうちいずれか一項に記載の方法。

40

【請求項 13】

プロセッサによって実行されたときに、請求項 1 ないし 12 のうちいずれか一項に記載の方法を含む処理を実行するよう装置を制御するコンピュータ・プログラムを記憶している、非一時的なコンピュータ読み取り可能な媒体。

【請求項 14】

オーディオ処理のための装置であって、当該装置は：

50

プロセッサ；および
メモリを有しており、

前記プロセッサは、第1帯域利得およびオーディオ信号の音声活動検出値を、前記オーディオ信号の表現を入力として取る第1の機械学習モデルを使用して生成することによって、前記オーディオ信号の前記表現は前記オーディオ信号のスペクトル解析の出力に基づく、ことを実行するよう当該装置を制御するように構成されており；

前記プロセッサは、前記第1帯域利得および前記音声活動検出値に基づいて背景ノイズ推定値を生成するよう当該装置を制御するように構成されており；

前記プロセッサは、前記背景ノイズ推定値によって制御されるウィナー・フィルタを使用して前記オーディオ信号を処理することによって、第2帯域利得を生成するよう当該装置を制御するよう構成されており；

10

前記プロセッサは、前記第1帯域利得と前記第2帯域利得を組み合わせることによって、組み合わせられた利得を生成するよう当該装置を制御するように構成されており；

前記プロセッサは、前記組み合わせられた利得を使用して前記オーディオ信号を修正することによって、修正されたオーディオ信号を生成するよう当該装置を制御するように構成されている、
装置。

【請求項15】

前記第1帯域利得および前記第2帯域利得のうちの少なくとも1つを生成するときに、少なくとも1つの制限が適用される、請求項14に記載の装置。

20

【発明の詳細な説明】

【技術分野】

【0001】

関連出願への相互参照

本願は、2020年11月11日出願の欧州特許出願第20206921.7号、2020年11月5日出願の米国仮特許出願第63/110,114号、2020年8月20日出願の米国仮特許出願第63/068,227号および2020年7月31日出願の国際特許出願第PCT/CN2020/106270号の優先権の利益を主張するものであり、これらはすべて、ここにその全体が参照により組み込まれる。

【0002】

30

分野

本開示は、オーディオ処理、特にノイズ削減に関する。

【背景技術】

【0003】

本稿に別段の記載がない限り、本節に記載されているアプローチは、本願の請求項に対する先行技術ではなく、本節に含まれることによって先行技術であると自認されるものではない。

【0004】

ノイズ削減は、モバイル装置で実装するのが困難である。モバイル装置は、音声通信、ユーザー生成コンテンツの開発などを含む、多様な使用事例において定常的および非定常的ノイズの両方を捕捉する可能性がある。モバイル装置は電力消費および処理能力に制約がある可能性があるため、モバイル装置によって実装された場合に効果的であるノイズ削減プロセスを開発することは困難である。

40

【発明の概要】

【発明が解決しようとする課題】

【0005】

以上のことから、モバイル装置においてうまく機能するノイズ削減システムを開発する必要がある。

【課題を解決するための手段】

【0006】

50

ある実施形態によれば、コンピュータ実装されるオーディオ処理方法は、機械学習モデルを使用して、第1帯域利得およびオーディオ信号の音声活動検出値を生成することを含む。この方法は、さらに、第1帯域利得および音声活動検出値に基づいて背景ノイズ推定値を生成することを含む。この方法は、さらに、背景ノイズ推定値によって制御されるウィナー・フィルタを使用してオーディオ信号を処理することによって、第2帯域利得を生成することを含む。この方法はさらに、第1帯域利得と第2帯域利得を組み合わせることによって、組み合わせられた利得を生成することを含む。この方法はさらに、組み合わせられた利得を使用してオーディオ信号を修正することによって、修正オーディオ信号を生成することを含む。

【0007】

別の実施形態によれば、装置がプロセッサとメモリを含む。プロセッサは、本願に記載される方法の一つまたは複数を実装するよう当該装置を制御するように構成される。装置は、さらに、本願に記載される方法の一つまたは複数と同様の詳細を含んでいてもよい。

【0008】

別の実施形態によれば、非一時的なコンピュータ可読媒体が、プロセッサによって実行されると、本願に記載される方法の一つまたは複数を含む処理を実行するように装置を制御するコンピュータ・プログラムを記憶する。

【0009】

以下の詳細な説明と付属の図面は、さまざまな実装の性質および利点のさらなる理解を提供する。

【図面の簡単な説明】

【0010】

【図1】ノイズ削減システム100のブロック図である。

【0011】

【図2】本開示の例示の実施形態を実装するのに好適なシステム200の例のブロック図である。

【0012】

【図3】オーディオ処理の方法300のフロー図である。

【発明を実施するための形態】

【0013】

本願では、ノイズ削減に関する技法が記載される。以下の記述では、説明の目的で、本開示の十全な理解を提供するために、多数の例および個別的な詳細が記載される。しかしながら、請求項によって定義される本開示は、これらの例の特徴の一部または全部を単独で、または以下に記載される他の特徴との組み合わせで含むことができ、さらに、本願に記載される特徴および概念の修正および等価物を含むことができることは、当業者には明らかであろう。

【0014】

以下の記述では、さまざまな方法、プロセスおよび手順が詳述されている。具体的なステップがある順序で記述されていることがあるが、そのような順序は主に簡便のためである。特定のステップが複数回繰り返されてもよく、他のステップの前または後に行われてもよく（たとえそれらのステップが別の順序で記述されている場合でも）、他のステップと並列に行われてもよい。第2のステップは、第2のステップが開始される前に第1のステップが完了される必要がある場合にのみ、第1のステップの後になることが要求される。そのような状況は、文脈から明らかでない場合には、具体的に指摘される。

【0015】

本稿では、「および」、「または」および「および/または」という用語が使用される。そのような用語は包含的な意味をもつものと読むべきである。たとえば、「AおよびB」は、少なくとも以下を意味することがありうる：「AとBの両方」、「少なくともAとBの両方」。別の例として、「AまたはB」は少なくとも以下を意味することがありうる：「少なくともA」、「少なくともB」、「AとBの両方」、「少なくともAとBの両方」。別の例

10

20

30

40

50

として、「Aおよび/またはB」は少なくとも以下を意味することがありうる：「AおよびB」、「AまたはB」。排他的離接が意図されている場合、そのことが具体的に記載される（たとえば、「AかBのどちらか」、「高々AとBの一方」）。

【0016】

本稿は、ブロック、要素、コンポーネント、回路などの構造に関連するさまざまな処理機能を記述する。一般に、これらの構造は一つまたは複数のコンピュータ・プログラムによって制御されるプロセッサによって実装されうる。

【0017】

図1は、ノイズ削減システム100のブロック図である。ノイズ削減システム100は、携帯電話、マイクロフォン付きビデオカメラなどのモバイル装置（たとえば、図2参照）において実装されてもよい。ノイズ削減システム100のコンポーネントは、たとえば一つまたは複数のコンピュータ・プログラムに従って制御されるプロセッサによって実装されてもよい。ノイズ削減システム100は、窓掛けブロック102、変換ブロック104、帯域特徴解析ブロック106、ニューラルネットワーク108、ウィナー・フィルタ110、利得組み合わせブロック112、帯域利得対ピン利得ブロック114、信号修正ブロック116、逆変換ブロック118、逆窓掛けブロック120を含む。ノイズ削減システム100は、（簡潔のため）詳細に説明されていない他のコンポーネントを含んでいてもよい。

10

【0018】

窓掛けブロック102は、オーディオ信号150を受領し、オーディオ信号150に対して窓掛けを実行し、オーディオ・フレーム152を生成する。オーディオ信号150は、ノイズ削減システム100を実装するモバイル装置のマイクロフォンによって捕捉されうる。一般に、オーディオ信号150は、オーディオ・サンプルのシーケンスを含む時間領域信号である。たとえば、オーディオ信号150は48kHzのサンプリング・レートで捕捉され、各サンプルは16ビットのビットレートで量子化されるのでよい。他の例示的なサンプリング・レートは44.1kHz、96kHz、192kHzなどを含んでいてもよく、他のビットレートには24ビット、32ビットなどを含みうる。

20

【0019】

一般に、窓掛けブロック102は、オーディオ信号150のサンプルに重複窓を適用して、オーディオ・フレーム152を生成する。窓掛けブロック102は、長方形窓、三角形窓、台形窓、正弦窓などを含むさまざまな形の窓掛けを実装することができる。

30

【0020】

変換ブロック104は、オーディオ・フレーム152を受領し、オーディオ・フレーム152に対して変換を実行し、変換特徴154を生成する。変換は周波数領域変換であってもよく、変換特徴154は各オーディオ・フレームのピン特徴および基本周波数パラメータを含むことができる。（変換特徴154はピン特徴154と呼ばれることもある。）基本周波数パラメータは、F0と呼ばれる音声基本周波数を含んでいてもよい。変換ブロック104は、フーリエ変換（たとえば、高速フーリエ変換（FFT））、直交ミラーフィルタ（QMF）領域変換などを含むさまざまな変換を実装することができる。たとえば、変換ブロック104は、960ポイントの分解窓と480ポイントのフレーム・シフトをもつFFTを実装してもよい；あるいはまた、1024ポイントの分解窓と512ポイントのフレーム・シフトが実装されてもよい。変換特徴154におけるピンの数は、一般に変換分解のポイントの数に関係している。たとえば、960ポイントのFFTは481ピンになる。

40

【0021】

変換ブロック104は、各オーディオ・フレームの基本周波数パラメータを決定するためのさまざまなプロセスを実装することができる。たとえば、変換がFFTである場合、変換ブロック104はFFTパラメータから基本周波数パラメータを抽出することができる。別の例として、変換ブロック104は、時間領域信号（たとえば、オーディオフレーム152）の自己相関に基づいて基本周波数パラメータを抽出してもよい。

【0022】

帯域特徴解析ブロック106は、変換特徴154を受領し、変換特徴154に対して帯域解析

50

を実行し、帯域特徴156を生成する。帯域特徴156は、メル(Mel)スケール、バーク(Bark)スケールなどを含む、さまざまなスケールに応じて生成される。帯域特徴156における帯域の数は、異なるスケールを使用する場合には異なる場合があり、たとえば、Barkスケールについては24個の帯域、Melスケールについては80個の帯域などである。帯域特徴解析ブロック106は、帯域特徴156を基本周波数パラメータ(たとえばF0)と組み合わせてもよい。

【0023】

帯域特徴解析ブロック106は、長方形の帯域を使用することができる。帯域特徴解析ブロック106は、ピーク応答が帯域間の境界にある三角形の帯域を使用することもできる。

【0024】

帯域特徴156は、Mel帯域エネルギー、Bark帯域エネルギーなどの帯域エネルギーであってもよい。帯域特徴解析ブロック106は、Mel帯域エネルギーとBark帯域エネルギーの対数値を計算してもよい。帯域特徴解析ブロック106は、帯域エネルギーの離散コサイン変換(DCT)変換を適用して、新しい帯域特徴を生成して、新しい帯域特徴がもとの帯域特徴よりも相関の低いものになるようにしてもよい。たとえば、帯域特徴解析ブロック106は、メル周波数ケプストラム係数(Mel-frequency cepstral coefficient、MFCC)、バーク周波数ケプストラム係数(Bark-frequency cepstral coefficient、BFCC)などとして帯域特徴156を生成してもよい。

【0025】

帯域特徴解析ブロック106は、平滑化値(smoothing value)に従って、現在のフレームと前の諸フレームの平滑化を実行してもよい。帯域特徴解析ブロック106は、現在のフレームと前の諸フレームの間の一階の差分と二階の差分を計算することによって、差分解析を実行することもできる。

【0026】

帯域特徴解析ブロック106は、現在の帯域のどれだけが周期的な信号で構成されているかを示す帯域調和性特徴(band harmonicity feature)を計算してもよい。たとえば、帯域特徴解析ブロック106は、現在のフレームのFFT周波数バインド(FFT frequency bind)に基づいて帯域調和性特徴を計算してもよい。別の例として、帯域特徴解析ブロック106は、現在のフレームと直前のフレームとの相関に基づいて帯域調和性特徴を計算してもよい。

【0027】

一般に、帯域特徴156はピン特徴154よりも数が少なく、よって、ニューラルネットワーク108に入力されるデータの次元性を下げる。たとえば、ピン特徴は513または481個のピンのオーダーであってもよく、帯域特徴156は24または80個の帯域のオーダーであってもよい。

【0028】

ニューラルネットワーク108は帯域特徴156を受け取り、モデルに従って帯域特徴156を処理し、利得158と音声活動判断(voice activity decision、VAD)160を生成する。利得158は、たとえばニューラルネットワークの出力であることを示すために、DGainと呼ばれることもある。モデルはオフラインでトレーニングされている。トレーニング・データ・セットの準備を含むモデルのトレーニングについては、後のセクションで説明する。

【0029】

ニューラルネットワーク108は、このモデルを使用して、帯域特徴156(たとえば、基本周波数F0を含む)に基づいて各帯域についての利得および音声活動を推定し、利得158およびVAD 160を出力する。ニューラルネットワーク108は、全結合型ニューラルネットワーク(FCNN)、リカレントニューラルネットワーク(RNN)、畳み込みニューラルネットワーク(CNN)、別のタイプの機械学習システムなど、またはそれらの組み合わせでありうる。

【0030】

10

20

30

40

50

ノイズ削減システム100は、ニューラルネットワーク108のDGains出力に平滑化〔スムージング〕または制限〔リミッティング〕を適用してもよい。たとえば、ノイズ削減システム100は、時間軸、周波数軸などに沿って、平均平滑化またはメジアン・フィルタリングを利得158に適用してもよい。別の例として、ノイズ削減システム100は、最大の利得を1.0、最小の利得は異なる帯域については異なるものとして、利得158にリミッティングを適用してもよい。ある実装では、ノイズ削減システム100は、最も低い4つの帯域についての最小利得として0.1（たとえば - 20dB）の利得を設定し、中間帯域についての最小利得として0.18（たとえば - 15dB）の利得を設定する。最小利得を設定することは、DGainsの不連続性を緩和する。最小利得値は所望に応じて調整されうる。たとえば、 - 12dB、 - 15dB、 - 18dB、 - 20dBなどの最小利得がさまざまな帯域について設定されうる。

10

【0031】

ウィーナー・フィルタ110は、帯域特徴156、利得158、VAD 160を受け取り、ウィーナー・フィルタリングを実行し、利得162を生成する。利得162は、たとえばそれがウィーナー・フィルタの出力であることを示すために、WGainsと呼ばれてもよい。一般に、ウィーナー・フィルタ110は、帯域特徴156に従って、入力信号150の各帯域における背景ノイズを推定する。（背景ノイズは定常ノイズと呼ばれることもある。）ウィーナー・フィルタ110は、ニューラルネットワークによって推定された利得158とVAD 160を使用して、そのフィルタリング・プロセスを制御する。ある実装では、音声活動のない（たとえば、VAD 160が0.5未満である）所与の入力フレーム（対応する帯域特徴156をもつ）について、ウィーナー・フィルタ110は、所与の入力フレームについての帯域利得を（利得158（DGains）に従って）チェックする。DGainsが0.5未満の帯域については、ウィーナー・フィルタ110はこれらの帯域をノイズ・フレームと見なし、これらのフレームの帯域エネルギーを平滑化して背景ノイズの推定値を得る。

20

【0032】

ウィーナー・フィルタ110は、各帯域についての帯域エネルギーを計算してノイズ推定値を得るために使用される平均フレーム数を追跡してもよい。所与の帯域についての平均数がフレーム数の閾値より大きい場合、所与の帯域についてのウィーナー帯域利得を計算するために、ウィーナー・フィルタ110が適用される。所与の帯域についての平均数がフレーム数の閾値より小さい場合、ウィーナー帯域利得は所与の帯域について1.0となる。各帯域についてのウィーナー帯域利得は、ウィーナー利得（またはWGains）とも呼ばれる利得162として出力される。

30

【0033】

事実上、ウィーナー・フィルタ110は、信号履歴（たとえば、入力信号150のいくつかのフレーム）に基づいて各帯域における背景ノイズを推定する。フレーム数の閾値は、ウィーナー・フィルタ110に、背景ノイズの信頼性のある推定につながる十分な数のフレームを与える。ある実装では、フレーム数の閾値は50である。あるフレームが10msである場合、これは入力信号150の0.5秒に相当する。フレーム数が閾値より小さい場合、事実上、ウィーナー・フィルタ110はバイパスされる（たとえば、WGainsは1.0）。

【0034】

ノイズ削減システム100は、ウィーナー・フィルタ110のWGains出力にリミッティングを適用してもよく、最大利得は1.0であり、最小利得は異なる帯域については異なる。ある実装では、ノイズ削減システム100は、最も低い4つの帯域についての最小利得として0.1（たとえば - 20dB）の利得を設定し、中間帯域についての最小利得として0.18（たとえば - 15dB）の利得を設定する。最小利得を設定することは、WGainsの不連続性を緩和する。最小利得値は所望に応じて調整されうる。たとえば、 - 12dB、 - 15dB、 - 18dB、 - 20dBなどの最小利得がさまざまな帯域について設定されうる。

40

【0035】

利得組み合わせブロック112は、利得158（DGains）と利得162（WGains）を受け取り、それらの利得を組み合わせ、利得164を生成する。利得164は、たとえばそれがDG

50

ainsとWGainsの組み合わせであることを示すために、帯域利得、組み合わせられた帯域利得〔組み合わせられた帯域利得〕、またはCGainsと呼ばれることもある。例として、利得組み合わせブロック112は、DGainsとWGainsを乗算してCGainsを帯域ごとに生成してもよい。

【0036】

ノイズ削減システム100は、利得組み合わせブロック112のCGains出力にリミッティングを適用してもよく、最大利得は1.0であり、最小利得は異なる帯域については異なる。ある実装では、ノイズ削減システム100は、最も低い4つの帯域についての最小利得として0.1（たとえば - 20dB）の利得を設定し、中間帯域についての最小利得として0.18（たとえば - 15dB）の利得を設定する。最小利得を設定することは、CGainsの不連続性を緩和する。最小利得値は所望に応じて調整されうる。たとえば、- 12dB、- 15dB、- 18dB、- 20dBなどの最小利得がさまざまな帯域について設定されうる。

10

【0037】

帯域利得からピン利得ブロック114は、利得164を受け取り、帯域利得をピン利得に変換して、利得166（ピン利得とも呼ばれる）を生成する。事実上、帯域利得からピン利得ブロック114は、利得164を帯域利得からピン利得に変換するために、帯域特徴解析ブロック106によって実行される処理の逆を実行する。たとえば、帯域特徴解析ブロック106が1024ポイントのFFTピンを24個のバーク・スケール帯域に処理した場合、帯域利得からピン利得ブロック114は、利得164の24個のバーク・スケール帯域を利得166の1024個のFFTピンに変換する。

20

【0038】

帯域利得からピン利得ブロック114は、帯域利得をピン利得に変換するさまざまな技術を実装することができる。たとえば、帯域利得からピン利得ブロック114は、補間、たとえば線形補間を使用することができる。

【0039】

信号修正ブロック116は、変換特徴154（ピン特徴と基本周波数F0を含む）と利得166を受け取り、利得166に従って変換特徴154を修正し、修正された変換特徴168（修正されたピン特徴と基本周波数F0を含む）を生成する。（修正された変換特徴168は、修正されたピン特徴168と呼ばれることもある。）信号修正ブロック116は、利得166に基づいてピン特徴154の振幅スペクトルを修正してもよい。ある実装では、信号修正ブロック116は、修正されたピン特徴168を生成するときに、ピン特徴154の位相スペクトルを変更しないままにする。別の実装では、信号修正ブロック116は、修正されたピン特徴168を生成するときに、たとえば修正されたピン特徴168に基づいて推定を実行することによって、ピン特徴154の位相スペクトルを調整する。例として、信号修正ブロック116は、たとえばグリフィン・リム（Griffin-Lim）プロセスを実装することによって、位相スペクトルを調整するために、短時間フーリエ変換を使用することができる。

30

【0040】

逆変換ブロック118は、修正された変換特徴168を受け取り、修正された変換特徴168に対して逆変換を実行し、オーディオ・フレーム170を生成する。一般に、実行される逆変換は、変換ブロック104によって実行される変換の逆である。たとえば、逆変換ブロック118は、逆フーリエ変換（たとえば、逆FFT）、逆QMF変換などを実装することができる。

40

【0041】

逆窓掛けブロック120は、オーディオ・フレーム170を受領し、オーディオ・フレーム170に対して逆窓掛けを実行し、オーディオ信号172を生成する。一般に、実行される逆窓掛けは、窓掛けブロック102によって実行される窓掛けの逆である。たとえば、逆窓掛けブロック120は、オーディオ信号172を生成するために、オーディオ・フレーム170に対して重複加算を実行してもよい。

【0042】

結果として、ニューラルネットワーク108の出力を使用してウィナー・フィルタ110

50

を制御するという組み合わせは、単にニューラルネットワークのみを使用してノイズ削減を実行するよりも、改善された結果を提供する可能性がある。多くのニューラルネットワークが単に短いメモリを使用して動作するからである。

【 0 0 4 3 】

図2は、本開示の例示的な実施形態を実装するのに適した例示的なシステム200のブロック図を示す。システム200は、一つまたは複数のサーバー・コンピュータまたは任意のクライアント装置を含む。システム200は、スマートフォン、メディアプレーヤー、タブレットコンピュータ、ラップトップ、ウェアラブルコンピュータ、車両コンピュータ、ゲームコンソール、サラウンドシステム、キオスクなどを含むがこれらに限定されない、任意の消費者装置を含む。

10

【 0 0 4 4 】

示されているように、システム200は、たとえばリードオンリーメモリ (ROM) 202に格納されたプログラム、またはたとえば記憶ユニット208からランダムアクセスメモリ (RAM) 203にロードされたプログラムに従って、さまざまな処理を実行することができる中央処理装置 (CPU) 201を含む。RAM 203では、CPU 201がさまざまなプロセスを実行する際に必要になるデータも必要に応じて格納される。CPU 201、ROM 202、RAM 203はバス204を介して互いに接続される。入出力 (I/O) インターフェース205もバス204に接続されている。

【 0 0 4 5 】

以下のコンポーネントがI/Oインターフェース205に接続されている：キーボード、マウス、タッチスクリーン、モーションセンサー、カメラなどを含みうる入力ユニット206；液晶ディスプレイ (LCD) などのディスプレイと一つまたは複数のスピーカーを含みうる出力ユニット207；ハードディスクまたは他の好適な記憶装置を含む記憶ユニット208；ネットワークカード (たとえば有線または無線) などのネットワークインターフェースカードを含む通信ユニット209。通信ユニット209は、たとえばワイヤレスマイクロフォン、ワイヤレスイヤホン、ワイヤレススピーカーなどのワイヤレス入出力コンポーネントと通信することもできる。

20

【 0 0 4 6 】

いくつかの実装では、入力ユニット206は、さまざまなフォーマット (たとえば、モノラル、ステレオ、空間的、没入的、その他の好適なフォーマット) のオーディオ信号の捕捉を可能にする、異なる位置 (ホスト装置に依存する) にある一つまたは複数のマイクロフォンを含む。

30

【 0 0 4 7 】

いくつかの実装では、出力ユニット207は、さまざまな数のスピーカーをもつシステムを含む。図2に示されるように、出力ユニット207は (ホスト装置の機能に依存して) さまざまなフォーマット (たとえば、モノラル、ステレオ、没入的、バイノーラル、その他の好適なフォーマット) のオーディオ信号をレンダリングすることができる。

【 0 0 4 8 】

通信ユニット209は、他の装置と (たとえばネットワークを介して) 通信するように構成される。必要に応じて、ドライブ210もI/Oインターフェース205に接続される。ドライブ210には、磁気ディスク、光ディスク、光磁気ディスク、フラッシュドライブ、または他の好適なリムーバブルメディアなどのリムーバブルメディア211がマウントされ、必要に応じて、そこから読み取られたコンピュータ・プログラムが記憶ユニット208にインストールされる。システム200は上記の構成要素を含むものとして説明されているが、実際の適用では、これらの構成要素のいくつかを追加、除去、および/または置換することが可能であり、これらのすべての修正または変更は、みな本開示の範囲に含まれることを当業者は理解するであろう。

40

【 0 0 4 9 】

たとえば、システム200は、たとえばCPU 201上で一つまたは複数のコンピュータ・プログラムを実行することによって、ノイズ削減システム100 (図1参照) の一つまたは

50

複数の構成要素を実装することができる。ROM 802、RAM 803、記憶ユニット808などは、ニューラルネットワーク108が使用するモデルを記憶してもよい。入力装置206に接続されたマイクロフォンがオーディオ信号150を捕捉してもよく、出力装置207に接続されたスピーカがオーディオ信号172に対応する音を出力することができる。

【0050】

図3はオーディオ処理の方法300のフロー図である。方法300は、一つまたは複数のコンピュータ・プログラムの実行によって制御されるように、装置（たとえば、図2のシステム200）によって実装されうる。

【0051】

302では、機械学習モデルを使用して、オーディオ信号の第1帯域利得および音声活動検出値が生成される。たとえば、CPU 201は、モデルに従って帯域特徴156を処理することによって、利得158およびVAD 160を生成するニューラルネットワーク108（図1参照）を実装してもよい。

10

【0052】

304では、第1帯域利得および音声活動検出値に基づいて背景ノイズ推定値が生成される。たとえば、CPU 201は、ウィナー・フィルタ110を動作させることの一部として、利得158およびVAD 160に基づいて背景ノイズ推定値を生成してもよい。

【0053】

306では、背景ノイズ推定値によって制御されるウィナー・フィルタを使用してオーディオ信号を処理することによって、第2帯域利得が生成される。たとえば、CPU 201は、背景ノイズ推定値（304を参照）によって制御される帯域特徴156を処理することによって利得162を生成するよう、ウィナー・フィルタ110を実装してもよい。たとえば、ノイズ・フレームの数が特定の帯域について閾値（たとえば50個のノイズ・フレーム）を超えると、ウィナー・フィルタはその特定の帯域について第2帯域利得を生成する。

20

【0054】

308では、第1帯域利得と第2帯域利得を組み合わせることによって、組み合わせられた利得が生成される。たとえば、CPU 201は、利得158（ニューラルネットワーク108から）と利得162（ウィナーフィルタ110から）を組み合わせることによって利得164を生成する利得組み合わせブロック112を実装してもよい。第1帯域利得と第2帯域利得は、乗算によって組み合わせられてもよい。第1帯域利得と第2帯域利得は、各帯域について第1帯域利得と第2帯域利得のうちの最大値を選択することによって組み合わせられてもよい。組み合わせられた利得にリミットングが適用されてもよい。第1帯域利得と第2帯域利得は乗算によって、または各帯域についての最大値を選択することによって組み合わせられてもよく、組み合わせられた利得にリミットングが適用されてもよい。

30

【0055】

310では、組み合わせられた利得を使用してオーディオ信号を修正することによって、修正されたオーディオ信号が生成される。たとえば、CPU 201は、利得166を使用してピン特徴154を修正することによって、修正されたピン特徴168を生成するために、信号修正ブロック116を実装することができる。

【0056】

方法300は、ノイズ削減システム100に関して上述したものと同様の他のステップを含むことができる。例示的なステップの網羅的でない議論は下記を含む。窓掛けステップ（窓掛けブロック102参照）が、ニューラルネットワーク108への入力を生成することの一部として、オーディオ信号に対して実行されてもよい。変換ステップ（変換ブロック104参照）は、ニューラルネットワーク108への入力を生成することの一部として、時間領域情報を周波数領域情報に変換するために、オーディオ信号に対して実行されてもよい。ピンから帯域への変換ステップ（帯域特徴解析ブロック106参照）は、ニューラルネットワーク108への入力の次元を減らすために、オーディオ信号に対して実行されてもよい。帯域からピンへの変換ステップ（帯域利得からピン利得ブロック114参照）が、帯域利得（たとえば利得164）をピン利得（たとえば利得166）に変換するために実行されてもよい

40

50

。逆変換ステップ（逆変換ブロック118参照）が、修正されたピン特徴168を周波数領域情報から時間領域情報（たとえば、オーディオフレーム170）に変換するために実行されてもよい。逆窓掛けステップ（逆窓掛けブロック120参照）が、オーディオ信号172を窓掛けステップの逆として再構成するために実行されてもよい。

【0057】

モデルの作成

【0058】

前述のように、ニューラルネットワーク108（図1参照）で使用されるモデルは、オフラインでトレーニングされ、次いでノイズ削減システム100によって記憶され、使用される。たとえば、コンピュータシステムは、たとえば一つまたは複数のコンピュータ・プログラムを実行することによって、モデルをトレーニングするモデル・トレーニング・システムを実装してもよい。モデルをトレーニングすることの一部は、入力特徴およびターゲット特徴を生成するためにトレーニング・データを準備することを含む。入力特徴は、ノイズのあるデータ（X）の帯域特徴計算によって計算される。ターゲット特徴は、理想的な帯域利得とVAD判定で構成される。

10

【0059】

ノイズのあるデータ（X）は、クリーンな発話（S）とノイズのあるデータ（N）を組み合わせることによって生成される。

【0060】

$$X = S + N$$

20

VAD判定は、クリーンな発話Sの解析に基づいていてもよい。ある実装では、VAD判定は、現在のフレームのエネルギーの絶対閾値によって決定される。他の実装では、他のVAD方法が使用される。たとえば、VADは手動でラベルを付けされることができる。

【0061】

理想的な帯域利得gは次式によって計算される。

【0062】

$$g_b = (E_s(b)/E_x(b))$$

上式で、 $E_s(b)$ はクリーンな発話の帯域bのエネルギーであり、 $E_x(b)$ ノイズのある発話の帯域bのエネルギーである。

【0063】

異なる使用事例に対してモデルを堅牢にするために、モデル・トレーニング・システムはトレーニング・データに対してデータ増強を実行してもよい。 S_i および N_i をもつ入力発話ファイルが与えられると、モデル・トレーニング・システムは、ノイズのあるデータを混合する前に S_i および N_i を変更する。データ増強は、3つの一般的なステップを含む。

30

【0064】

第1のステップは、クリーンな発話の振幅を制御することである。ノイズ削減モデルにとっての一般的な問題は、低音量の発話を抑制することである。このように、モデル・トレーニング・システムは、さまざまな振幅の発話を含むトレーニング・データを準備することによって、データ増強を実行する。

【0065】

モデル・トレーニング・システムは、-45dBから0dBの範囲のランダムなターゲット平均振幅を設定する（たとえば、-45, -40, -35, -30, -25, -20, -15, -10, -5, 0）。モデル・トレーニング・システムは、ターゲット平均振幅に一致するように、値aによって入力発話ファイルを修正する。

40

$$S_m = a * S_i$$

【0066】

2番目のステップは、信号対雑音比（SNR）を制御することである。発話ファイルとノイズ・ファイルのそれぞれの組み合わせについて、モデル・トレーニング・システムはランダムなターゲットSNRを設定する。ある実装では、ターゲットSNRは等しい確率でSNRの集合[-5, -3, 0, 3, 5, 10, 15, 18, 20, 30]からランダムに選択される。次

50

に、モデル・トレーニング・システムは、入力ノイズ・ファイルを値bによって修正して、 S_m の N_m の間のSNRをターゲットSNRに一致させる。

$$N_m = b \cdot N_i$$

【0067】

3番目のステップは、混合されたデータを制限することである。モデル・トレーニング・システムは、まず次式によって混合信号 X_m を計算する。

$$X_m = (S_m + N_m)$$

【0068】

クリッピングする場合（たとえば、16ビット量子化で.wavファイルとして X_m を保存する場合）、モデル・トレーニング・システムは、 A_{max} と記される X_m の最大絶対値を計算する。

【0069】

次に、修正比cが次式によって計算できる。

$$c = 32767 / A_{max}$$

【0070】

上記の式で、値32767は16ビット量子化からくる；この値は、他のビット量子化精度のために、必要に応じて調整されうる。

【0071】

次いで、

$$S = c \cdot S_m$$

$$N = c \cdot N_m$$

【0072】

SとNはノイズのある発話Xに混合される。

$$X = S + N$$

【0073】

平均振幅とSNRの計算は、所望に応じてさまざまなプロセスに従って実行されうる。モデル・トレーニング・システムは、平均振幅を計算する前に、最小閾値を使用して無音セグメントを除去してもよい。

【0074】

このように、多様なターゲット平均振幅とターゲットSNRを使用してトレーニング・データのセグメントを調整することによって、トレーニング・データの多様性を増やすために、データ増強が使用される。たとえば、ターゲット平均振幅の10個の変形とターゲットSNRの10個の変形を使用すると、トレーニング・データの単一セグメントの100通りの変形が得られる。データ増強は、トレーニング・データのサイズを増やす必要はない。トレーニング・データがデータ増強の前に100時間である場合、増強されたトレーニング・データの1万時間のフルセットがモデルをトレーニングするために使用される必要はない；増強されたトレーニング・データ・セットは、より小さいサイズ、たとえば100時間に制限されてもよい。さらに重要なことに、データ増強により、トレーニング・データにおける振幅とSNRの変動性が大きくなる。

【0075】

実装の詳細

【0076】

実施形態は、ハードウェア、コンピュータ可読媒体に格納された実行可能モジュール、またはその両方の組み合わせ（たとえばプログラマブルロジックアレイ）で実装されうる。特に断りのない限り、実施形態によって実行されるステップは、本来的にいかなる特定のコンピュータまたは他の装置にも関連する必要はない。ただし、ある種の実施形態ではそうであってもよい。特に、さまざまな汎用マシンが、本稿での教示に従って書かれたプログラムと一緒に使用されてもよく、あるいは必要とされる方法ステップを実行するために、より特化した装置（たとえば集積回路）を構築するほうが便利な場合もある。よって、それぞれが少なくとも1つのプロセッサ、少なくとも1つのデータ記憶システム（揮発性

10

20

30

40

50

および不揮発性メモリおよび/または記憶素子を含む)、少なくとも1つの入力装置またはポート、および少なくとも1つの出力装置またはポートを含む、一つまたは複数のプログラム可能なコンピュータシステム上で実行される一つまたは複数のコンピュータ・プログラムにおいて実装されてもよい。プログラムコードは、本稿で説明される機能を実行し、出力情報を生成するために入力データに適用される。出力情報は、既知の仕方の一つまたは複数の出力装置に適用される。

【0077】

そのような各コンピュータ・プログラムは、記憶媒体またはデバイスがコンピュータシステムによって読み取られるときに、本稿で説明する手順を実行するようコンピュータを構成し、動作させるための、汎用または特殊目的のプログラム可能なコンピュータによって読み取り可能な記憶媒体またはデバイス(たとえば、ソリッドステートメモリもしくは媒体、磁気もしくは光媒体)に記憶またはダウンロードされることが望ましい。また、本発明のシステムは、コンピュータ・プログラムをもって構成された、コンピュータ読み取り可能な記憶媒体として実装されると考えられる。そのように構成された記憶媒体は、コンピュータシステムに、本稿で記載される機能を実行するよう、特定の、事前に定義された仕方で動作させる。(ソフトウェア自体、および無形または一時的な信号は、特許を受けることができない主題である限りにおいて、除外される。)

10

【0078】

上記の記述は、本開示の諸側面がどのように実装されうるかの例とともに、本開示のさまざまな実施形態を例示している。上記の例および実施形態は、唯一の実施形態とみなされるべきではなく、以下の請求項によって定義される本開示の柔軟性および利点を説明するために提示されている。上記の開示および以下の請求項に基づき、他の配置、実施形態、実装および等価物が、当業者には明らかとなり、請求項によって定義される本開示の精神および範囲から逸脱することなく採用されうる。

20

【0079】

本発明のさまざまな側面は、以下の箇条書き例示的实施形態(enumerated example embodiment、EEE)から理解されうる。

〔EEE 1〕

コンピュータ実装されるオーディオ処理方法であって、当該方法は:

機械学習モデルを使用して、第1帯域利得およびオーディオ信号の音声活動検出値を生成し;

30

前記第1帯域利得および前記音声活動検出値に基づいて背景ノイズ推定値を生成し;

前記背景ノイズ推定値によって制御されるウィナー・フィルタを使用して前記オーディオ信号を処理することによって、第2帯域利得を生成し;

前記第1帯域利得と前記第2帯域利得を組み合わせることによって、組み合わせられた利得を生成し;

前記組み合わせられた利得を使用して前記オーディオ信号を修正することによって、修正されたオーディオ信号を生成することを含む、

方法。

〔EEE 2〕

40

前記機械学習モデルが、トレーニング・データの多様性を増すようデータ増強を使用して生成される、EEE 1に記載の方法。

〔EEE 3〕

前記第1帯域利得および前記音声活動検出値を生成することは、全結合型ニューラルネットワーク、リカレントニューラルネットワーク、および畳み込みニューラルネットワークのいずれかを使用して実行される、EEE 1または2に記載の方法。

〔EEE 4〕

前記第1帯域利得を生成することは、少なくとも2つの異なる帯域について少なくとも2つの異なる制限を使用して前記第1帯域利得を制限することを含む、EEE 1ないし3のうちいずれか一項に記載の方法。

50

〔 EEE 5 〕

前記背景ノイズ推定値を生成することは、特定の帯域についての閾値を超える、いくつかのノイズ・フレームに基づく、EEE 1 ないし 4 のうちいずれか一項に記載の方法。

〔 EEE 6 〕

前記第2帯域利得を生成することは、特定の帯域についての定常ノイズ・レベルに基づいて前記ウィナー・フィルタを使用することを含み、EEE 1 ないし 5 のうちいずれか一項に記載の方法。

〔 EEE 7 〕

前記第2帯域利得を生成することが、少なくとも2つの異なる帯域について少なくとも2つの異なる制限を使用して前記第2帯域利得を制限することを含み、EEE 1 ないし 6 のうちいずれか一項に記載の方法。

10

〔 EEE 8 〕

前記組み合わせられた利得を生成することは：

前記第1帯域利得と前記第2帯域利得を乗算し；

少なくとも2つの異なる帯域について少なくとも2つの異なる制限を使用して、前記組み合わせられた帯域利得を制限することを含み、
EEE 1 ないし 7 のうちいずれか一項に記載の方法。

〔 EEE 9 〕

前記修正されたオーディオ信号を生成することは、前記組み合わせられた帯域利得を使用して前記オーディオ信号の振幅スペクトルを修正することを含み、EEE 1 ないし 8 のうちいずれか一項に記載の方法。

20

〔 EEE 10 〕

入力オーディオ信号に重複窓を適用して複数のフレームを生成することをさらに含み、前記オーディオ信号が該複数のフレームに対応する、EEE 1 ないし 9 のうちいずれか一項に記載の方法。

〔 EEE 11 〕

前記オーディオ信号に対してスペクトル解析を実行し、前記オーディオ信号の複数のピン特徴および基本周波数を生成することをさらに含み、

前記第1帯域利得および前記音声活動検出値は、前記複数のピン特徴および前記基本周波数に基づく、

EEE 1 ないし 10 のうちいずれか一項に記載の方法。

30

〔 EEE 12 〕

前記複数のピン特徴に基づいて複数の帯域特徴を生成し、前記複数の帯域特徴は、メル周波数ケプストラム係数およびバーク周波数ケプストラム係数の一方を使用して生成され、

前記第1帯域利得および前記音声活動検出値は、前記複数の帯域特徴および前記基本周波数に基づく、

EEE 11 に記載の方法。

〔 EEE 13 〕

前記組み合わせられた利得は、前記オーディオ信号の複数の帯域に関連する組み合わせられた帯域利得であり、当該方法は、さらに：

前記組み合わせられた帯域利得を組み合わせられたピン利得に変換することを含み、前記組み合わせられたピン利得は複数のピンに関連する、

EEE 1 ないし 12 のうちいずれか一項に記載の方法。

40

〔 EEE 14 〕

プロセッサによって実行されたときに、EEE 1 ないし 13 のうちいずれか一項に記載の方法を含む処理を実行するよう装置を制御するコンピュータ・プログラムを記憶している、非一時的なコンピュータ読み取り可能な媒体。

〔 EEE 15 〕

オーディオ処理のための装置であって、当該装置は：

プロセッサ；および

50

メモリを有しており、

前記プロセッサは、機械学習モデルを使用して、第1帯域利得およびオーディオ信号の音声活動検出値を生成するよう当該装置を制御するように構成されており；

前記プロセッサは、前記第1帯域利得および前記音声活動検出値に基づいて背景ノイズ推定値を生成するよう当該装置を制御するように構成されており；

前記プロセッサは、前記背景ノイズ推定値によって制御されるウィナー・フィルタを使用して前記オーディオ信号を処理することによって、第2帯域利得を生成するよう当該装置を制御するよう構成されており；

前記プロセッサは、前記第1帯域利得と前記第2帯域利得を組み合わせることによって、組み合わせられた利得を生成するよう当該装置を制御するように構成されており；

前記プロセッサは、前記組み合わせられた利得を使用して前記オーディオ信号を修正することによって、修正されたオーディオ信号を生成するよう当該装置を制御するように構成されている、

装置。

〔EEE 16〕

前記機械学習モデルが、トレーニング・データの多様性を増すようデータ増強を使用して生成される、EEE 16に記載の装置。

〔EEE 17〕

前記第1帯域利得および前記第2帯域利得のうちの少なくとも1つを生成するときに、少なくとも1つの制限が適用される、EEE 15または16に記載の装置。

〔EEE 18〕

前記背景ノイズ推定値を生成することは、特定の帯域についての閾値を超える、いくつかのノイズ・フレームに基づく、EEE 15ないし17のうちいずれか一項に記載の装置。

〔EEE 19〕

前記プロセッサは、前記オーディオ信号に対してスペクトル解析を実行し、前記オーディオ信号の複数のピン特徴および基本周波数を生成するよう当該装置を制御するように構成されており、

前記第1帯域利得および前記音声活動検出値は、前記複数のピン特徴および前記基本周波数に基づく、

EEE 15ないし18のうちいずれか一項に記載の装置。

〔EEE 20〕

前記プロセッサは、前記複数のピン特徴に基づいて複数の帯域特徴を生成するよう当該装置を制御するように構成されており、前記複数の帯域特徴は、メル周波数ケプストラム係数およびパーク周波数ケプストラム係数の一方を使用して生成され、

前記第1帯域利得および前記音声活動検出値は、前記複数の帯域特徴および前記基本周波数に基づく、

EEE 19に記載の装置。

【先行技術文献】

【特許文献】

【0080】

【文献】米国特許出願公開第2019/0378531号

【文献】米国特許第10,546,593B2号

【文献】米国特許第10,224,053B2号

【文献】米国特許第9,053,697B2号

【文献】中国特許公開第105513605B号

【文献】中国特許公開第111192599A号

【文献】中国特許公開第110660407B号

【文献】中国特許公開第110211598A号

【文献】中国特許公開第110085249A号

【文献】中国特許公開第109378013A号

10

20

30

40

50

【文献】中国特許公開第109065067A号

【文献】中国特許公開第107863099A号

【非特許文献】

【0081】

【文献】Jean-Marc Valin, “A Hybrid DSP Deep Learning Approach to Real-Time Full-Band Speech Enhancement”, 2018 IEEE 20th International Workshop on Multimedia Signal Processing (MMSP), DOI: 10.1109/MMSP.2018.8547084.

【文献】Xia, Y., Stern, R., “A Priori SNR Estimation Based on a Recurrent Neural Network for Robust Speech Enhancement”, Proc. Interspeech 2018, 3274-3278, DOI: 10.21437/Interspeech.2018-2423. 10

【文献】Zhang, Q., Nicolson, A. M., Wang, M., Paliwal, K., & Wang, C.-X., “DeepMMSE: A Deep Learning Approach to MMSE-based Noise Power Spectral Density Estimation”, IEEE/ACM Transactions on Audio, Speech, and Language Processing, 1-1. DOI:10.1109/taslp.2020.2987441.

20

30

40

50

【図面】

【図 1】

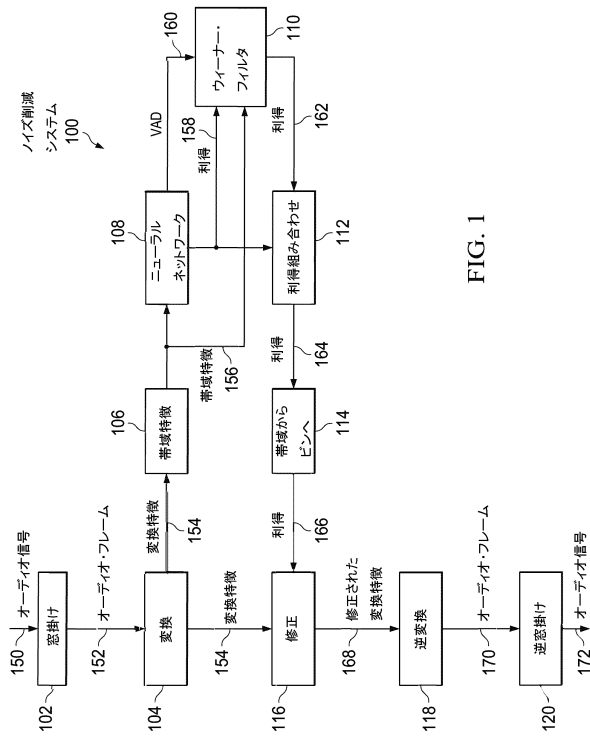


FIG. 1

【図 2】

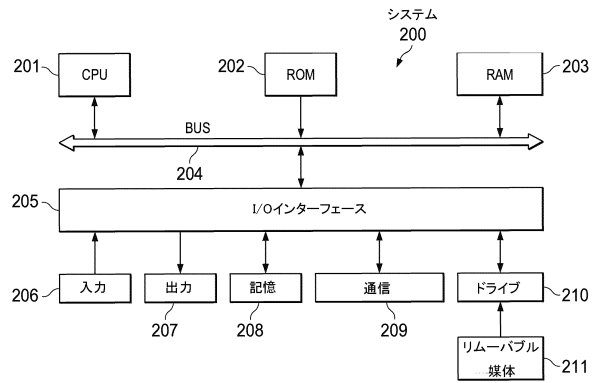


FIG. 2

【図 3】

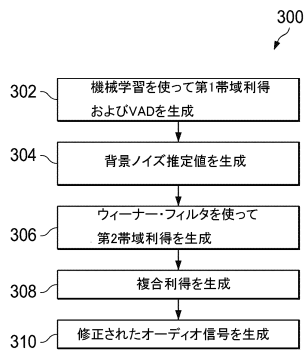


FIG. 3

10

20

30

40

50

フロントページの続き

(33)優先権主張国・地域又は機関

米国(US)

(31)優先権主張番号 63/110,114

(32)優先日 令和2年11月5日(2020.11.5)

(33)優先権主張国・地域又は機関

米国(US)

(31)優先権主張番号 20206921.7

(32)優先日 令和2年11月11日(2020.11.11)

(33)優先権主張国・地域又は機関

欧州特許庁(EP)

ストリート 1275 ドルビー ラボラトリーズ インコーポレイテッド 内

審査官 菊池 智紀

(56)参考文献 特開2018-014711(JP,A)

特開2020-115206(JP,A)

特開平05-232986(JP,A)

特表2009-503568(JP,A)

(58)調査した分野 (Int.Cl., DB名)

G10L 21/02 - 21/0364, 15/20

IEEE Xplore