

(19)日本国特許庁(JP)

(12)公開特許公報(A)

(11)公開番号

特開2023-24295  
(P2023-24295A)

(43)公開日 令和5年2月16日(2023.2.16)

(51)国際特許分類	F I	テーマコード(参考)
G 1 0 L 21/0332(2013.01)	G 1 0 L 21/0332	5 D 1 6 2
G 1 0 L 25/78 (2013.01)	G 1 0 L 25/78	5 D 2 2 0
H 0 4 R 3/00 (2006.01)	H 0 4 R 3/00 3 1 0	
H 0 4 S 7/00 (2006.01)	H 0 4 S 7/00 3 0 0	

審査請求 未請求 請求項の数 10 O L (全16頁)

(21)出願番号 特願2022-110199(P2022-110199)	(71)出願人 592051453
(22)出願日 令和4年7月8日(2022.7.8)	ハーマン インターナショナル インダストリーズ インコーポレイテッド
(31)優先権主張番号 202110895493.X	アメリカ合衆国 コネティカット 0 6 9 0 1 , スタムフォード , アトランティック ストリート 4 0 0 , 1 5 階
(32)優先日 令和3年8月5日(2021.8.5)	エイチ フロア
(33)優先権主張国・地域又は機関 中国(CN)	(74)代理人 100078282 弁理士 山本 秀策
	(74)代理人 100113413 弁理士 森下 夏樹
	(72)発明者 シャオ - フー シー
	アメリカ合衆国 カリフォルニア 9 4 0 4 3 , マウンテン ビュー , エリス ストリート 6 3 6

最終頁に続く

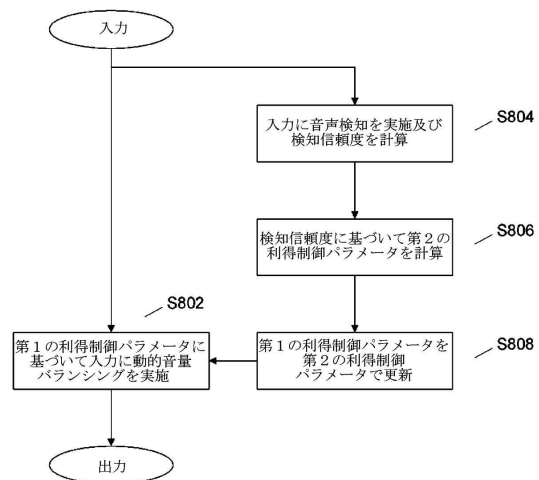
(54)【発明の名称】 動的音声強調のための方法及びシステム

(57)【要約】 (修正有)

【課題】オーディオソースの動的音声強調のための方法及びシステムを提供する。

【解決手段】オーディオソース入力信号を処理する方法は、オーディオソース入力を受信すること及び第1の利得制御パラメータに基づきオーディオソース入力に動的音量バランスングを実施する第1の経路信号処理と、オーディオソース入力に音声検知を実施すること及び検知信頼度を計算することならびに検知信頼度に基づき第2の利得制御パラメータを計算する第2の経路信号処理と、を含み、第1の利得制御パラメータを第2の処理経路によって計算された第2の利得制御パラメータで更新する。第1の経路信号処理及び第2の経路信号処理は、同期的または非同期的に実施され得る。

【選択図】図8



10

**【特許請求の範囲】****【請求項 1】**

動的音声強調の方法であって、

第 1 の経路信号処理を実施することであって、前記第 1 の経路信号処理が、オーディオソース入力を受信すること及び第 1 の利得制御パラメータに基づいて前記オーディオソース入力に動的音量バランスングを実施することを含む、前記第 1 の経路信号処理を実施することと、

第 2 の経路信号処理を実施することであって、前記第 2 の経路信号処理が、前記オーディオソース入力に音声検知を実施すること及び検知信頼度を計算することを含み、前記検知信頼度が、前記オーディオソース入力における前記音声の可能性を示す、前記第 2 の経路信号処理を実施することと、

前記検知信頼度に基づいて第 2 の利得制御パラメータを計算することと、

前記第 1 の利得制御パラメータを前記第 2 の利得制御パラメータで更新することと、前記更新された第 1 の利得制御パラメータに基づいて前記第 1 の経路信号処理を実施することと、

を含む、前記方法。

**【請求項 2】**

前記オーディオソース入力、マルチチャンネルソース入力を含み、前記オーディオソース入力に前記音声検知を実施すること及び検知信頼度を計算することが、

前記マルチチャンネルソース入力からセンタチャンネル信号を抽出することと、

前記センタチャンネル信号に正規化を実施することと、

前記正規化されたセンタチャンネル信号に高速自己相関を実施することであって、前記高速自己相関の結果が、前記検知信頼度を表す、前記高速自己相関を実施することと、を含む、請求項 1 に記載の方法。

**【請求項 3】**

前記検知信頼度に基づいて第 2 の利得制御パラメータを計算することが、

前記検知信頼度の対数関数に基づいて前記第 2 の利得制御パラメータを計算することと、

前記計算された第 2 の利得制御パラメータを平滑化することと、

前記平滑化された第 2 の利得制御パラメータを制限することと、を含む、請求項 1 に記載の方法。

**【請求項 4】**

前記オーディオソース入力、マルチチャンネルソース入力を含み、前記オーディオソース入力に前記動的音量バランスングを実施することが、

前記マルチチャンネルソース入力からセンタチャンネル信号を抽出することと、

前記センタチャンネル信号の音量を増強させること及び前記第 1 の利得制御パラメータまたは前記更新された第 1 の利得制御パラメータに基づいて他のチャンネル信号の音量を低減させることと、

前記増強されたセンタチャンネル信号及び前記低減された他のチャンネル信号を結合及び混合して、出力信号を生成することと、を含む、請求項 1 に記載の方法。

**【請求項 5】**

前記動的音量バランスングを実施する前に、前記オーディオソース入力にクロスオーバーフィルタリングを実施することをさらに含む、請求項 1 ~ 4 のいずれか一項に記載の方法。

**【請求項 6】**

前記オーディオソース入力の中間周波数域の信号にのみ前記動的音量バランスングを実施することと、

前記動的音量バランスング後に、前記オーディオソース入力の低周波数域及び高周波数域の信号と前記オーディオソース入力の前記中間周波数域の信号を結合及び混合して前記出力信号を生成することと、をさらに含む、請求項 5 に記載の方法。

**【請求項 7】**

10

20

30

40

50

前記オーディオソース入力が、デュアルチャンネルソース入力をさらに含み、前記方法が、前記デュアルチャンネルソース入力に基づいてマルチチャンネルソース入力を生成することをさらに含む、請求項 1 に記載の方法。

【請求項 8】

前記デュアルチャンネルソース入力に基づいて前記マルチチャンネルソース入力を生成することが、

前記デュアルチャンネルソース入力からの左チャンネル信号と右チャンネル信号との間の相互相関を実施することと、

組み合わせ比に従って、前記マルチチャンネルソース入力を生成することと、を含み、

前記組み合わせ比が、前記相互相関の結果による、請求項 7 に記載の方法。

10

【請求項 9】

前記第 1 の経路信号処理及び前記第 2 の経路信号処理が、同期または非同期である、請求項 1 に記載の方法。

【請求項 10】

動的音声強調のシステムであって、

コンピュータ実行可能な命令を格納するように構成されたメモリと、

請求項 1 ~ 9 のいずれか一項に記載の方法を実装するため、前記コンピュータ実行可能な命令を実行するように構成されたプロセッサと、を備える、前記システム。

【発明の詳細な説明】

【技術分野】

20

【0001】

本開示は、一般にオーディオ信号処理の分野に関し、より詳細には、オーディオソースの動的音声強調のための方法及びシステムに関する。

【背景技術】

【0002】

高解像度ケーブルテレビ及びオンラインストリーミングなどのメディアの消費の新しい方法、ならびに大画面テレビ及びディスプレイの出現により、映画体験が、消費者市場で人気を得ている。これらのメディアソースは、マルチチャンネルオーディオ技術または一般にサラウンド技術と呼ばれるものを伴うことが多い。Dolby、THX、及びDTSなどのサラウンドの提供者は、ソースコンテンツ用のより良好な空間オーディオ解決策を提供する固有のマルチチャンネルオーディオ符号化技術を有する。ムービー形式のコンテンツの主目的は、没入型サラウンド体験を提供することなので、多くの場合、音声了解度を犠牲にしてサラウンド体験を選ぶことが好まれる。これは、没入及び空間的解決策という点では大きな利点を与えるが、音質の劣化をもたらすことが多く、動画コンテンツの理解さえ困難になる場合がある。動画コンテンツのソースの音声の質を改善して了解度及び可聴度を向上するため、音声強調の方法が動画コンテンツにしばしば適用される。

30

【0003】

既存の音声強調のための共通の方法は、静的等化を利用することである。この方法は、静的等化を約 200 Hz ~ 4 kHz のオーディオチャンネルにのみ適用して、音声帯域の音量を増加させる。この実装は、わずかなシステムリソースを必要とするが、この方法で発生するひずみは、すぐにわかる。この実装方法は、クリップ内に音声または対話がない場合でさえ常に機能するので、音の調子の不均衡が引き起こされ、バックグラウンドが増幅される。より進んだ方法は、まず各時間フレーム内の音声を検知し、その後オーディオ信号を検知結果に基づいて自動的に処理する。この一方向の実行方法では、正確な音声の検知及びシステム処理の高速応答が要求される。しかし、いくつかの既存の方法では、音声を迅速かつ正確に検知できず、しばしば信号をゆがめるので、聞き苦しい。

40

【0004】

したがって、既存の解決策における上述の欠点を克服する改善された技術解決策が必要である。

【発明の概要】

50

## 【課題を解決するための手段】

## 【0005】

本開示の態様によれば、動的音声強調の方法が提供される。方法は、オーディオソース入力を受信すること及び第1の利得制御パラメータに基づいてオーディオソース入力に動的音量バランスングを実施することを含む第1の経路信号処理を実施することを含み得る。方法はまた、オーディオソース入力に音声検知を実施すること及びオーディオソース入力における音声の可能性を示す検知信頼度を計算することを含む第2の経路信号処理を実施することと、検知信頼度に基づいて第2の利得制御パラメータを計算することとを含み得る。方法はさらに、第1の利得制御パラメータを第2の利得制御パラメータで更新することと、及び更新された第1の利得制御パラメータに基づいて第1の経路信号処理を実施することとを含み得る。

10

## 【0006】

1つまたは複数の実施形態によれば、オーディオソース入力は、マルチチャンネルソース入力を含むことができ、オーディオソース入力に音声検知を実施すること及び検知信頼度を計算することには、マルチチャンネルソース入力からセンタチャンネル信号を抽出することと、センタチャンネル信号に正規化を実施することと、正規化されたセンタチャンネル信号にその結果が検知信頼度を表す高速自己相関を実施することとが含まれ得る。

## 【0007】

1つまたは複数の実施形態によれば、検知信頼度に基づいて第2の利得制御パラメータを計算することには、検知信頼度の対数関数に基づいて第2の利得制御パラメータを計算することと、計算された第2の利得制御パラメータを平滑化することと、平滑化された第2の利得制御パラメータを制限することとが含まれ得る。

20

## 【0008】

1つまたは複数の実施形態によれば、オーディオソース入力は、マルチチャンネルソース入力を含むことができ、オーディオソース入力に動的音量バランスングを実施することには、マルチチャンネルソース入力からセンタチャンネル信号を抽出することと、センタチャンネル信号の音量を増強させることと、第1の利得制御パラメータまたは更新された第1の利得制御パラメータに基づいて他のチャンネル信号の音量を低減させることと、増強されたセンタチャンネル信号と低減された他のチャンネル信号を結合及び混合して、出力信号を生成することと、が含まれる。

30

## 【0009】

1つまたは複数の実施形態によれば、方法はまた、動的音量バランスングを実施する前に、オーディオソース入力にクロスオーバーフィルタリングを実施することを含み得る。

## 【0010】

1つまたは複数の実施形態によれば、方法はまた、オーディオソース入力の中間周波数域の信号にのみ動的音量バランスングを実施することと、動的音量バランスング後にオーディオソース入力の低周波数域及び高周波数域の信号とオーディオソース入力の中間周波数域の信号を結合及び混合して出力信号を生成することと、を含み得る。

## 【0011】

1つまたは複数の実施形態によれば、オーディオソース入力はまた、デュアルチャンネルソース入力を含み、方法はまた、デュアルチャンネルソース入力に基づいてマルチチャンネルソース入力を生成することを含む。

40

## 【0012】

1つまたは複数の実施形態によれば、デュアルチャンネルソース入力に基づいてマルチチャンネルソース入力を生成することには、デュアルチャンネルソース入力からの左チャンネル信号と右チャンネル信号との間の相互相関を実施することと、組み合わせ比に従って、マルチチャンネルソース入力を生成することとが含まれ得る。組み合わせ比は、相互相関の結果による。

## 【0013】

1つまたは複数の実施形態によれば、第1の経路信号処理及び第2の経路信号処理は、

50

同期または非同期である。

【 0 0 1 4 】

本開示の別の態様によれば、メモリ及びプロセッサを含む、音声強調のためのシステムが提供される。メモリは、コンピュータ実行可能な命令を格納するように構成される。プロセッサは、命令を実行して、上述の方法を実装するように構成される。

【 0 0 1 5 】

本開示は、以下の非限定的実装形態の説明を添付の図面を参照して読むことによってより良く理解できる。

例えば、本願は以下の項目を提供する。

( 項目 1 )

10

動的音声強調の方法であって、

第 1 の経路信号処理を実施することであって、上記第 1 の経路信号処理が、オーディオソース入力を受信すること及び第 1 の利得制御パラメータに基づいて上記オーディオソース入力に動的音量ランシングを実施することを含む、上記第 1 の経路信号処理を実施することと、

第 2 の経路信号処理を実施することであって、上記第 2 の経路信号処理が、上記オーディオソース入力に音声検知を実施すること及び検知信頼度を計算することを含み、上記検知信頼度が、上記オーディオソース入力における上記音声の可能性を示す、上記第 2 の経路信号処理を実施することと、

上記検知信頼度に基づいて第 2 の利得制御パラメータを計算することと、

20

上記第 1 の利得制御パラメータを上記第 2 の利得制御パラメータで更新することと、上記更新された第 1 の利得制御パラメータに基づいて上記第 1 の経路信号処理を実施することと、

を含む、上記方法。

( 項目 2 )

上記オーディオソース入力、マルチチャンネルソース入力を含み、上記オーディオソース入力に上記音声検知を実施すること及び検知信頼度を計算することが、

上記マルチチャンネルソース入力からセンタチャンネル信号を抽出することと、

上記センタチャンネル信号に正規化を実施することと、

上記正規化されたセンタチャンネル信号に高速自己相関を実施することであって、上記高速自己相関の結果が、上記検知信頼度を表す、上記高速自己相関を実施することと、を含む、上記項目に記載の方法。

30

( 項目 3 )

上記検知信頼度に基づいて第 2 の利得制御パラメータを計算することが、上記検知信頼度の対数関数に基づいて上記第 2 の利得制御パラメータを計算することと、

上記計算された第 2 の利得制御パラメータを平滑化することと、

上記平滑化された第 2 の利得制御パラメータを制限することと、を含む、上記項目のいずれか一項に記載の方法。

( 項目 4 )

上記オーディオソース入力、マルチチャンネルソース入力を含み、上記オーディオソース入力に上記動的音量ランシングを実施することが、

40

上記マルチチャンネルソース入力からセンタチャンネル信号を抽出することと、

上記センタチャンネル信号の音量を増強させること及び上記第 1 の利得制御パラメータまたは上記更新された第 1 の利得制御パラメータに基づいて他のチャンネル信号の音量を低減させることと、

上記増強されたセンタチャンネル信号及び上記低減された他のチャンネル信号を結合及び混合して、出力信号を生成することと、を含む、上記項目のいずれか一項に記載の方法。

( 項目 5 )

上記動的音量ランシングを実施する前に、上記オーディオソース入力にクロスオーバーフィルタリングを実施することをさらに含む、上記項目のいずれか一項に記載の方法。

50

## (項目 6)

上記オーディオソース入力の中間周波数域の信号にのみ上記動的音量バランスングを実施することと、

上記動的音量バランスング後に、上記オーディオソース入力の低周波数域及び高周波数域の信号と上記オーディオソース入力の上記中間周波数域の信号を結合及び混合して上記出力信号を生成することと、をさらに含む、上記項目のいずれか一項に記載の方法。

## (項目 7)

上記オーディオソース入力、デュアルチャンネルソース入力をさらに含み、上記方法が、上記デュアルチャンネルソース入力に基づいてマルチチャンネルソース入力を生成することをさらに含む、上記項目のいずれか一項に記載の方法。

10

## (項目 8)

上記デュアルチャンネルソース入力に基づいて上記マルチチャンネルソース入力を生成することが、

上記デュアルチャンネルソース入力からの左チャンネル信号と右チャンネル信号との間の相互相関を実施することと、

組み合わせ比に従って、上記マルチチャンネルソース入力を生成することと、を含み、上記組み合わせ比が、上記相互相関の結果による、上記項目のいずれか一項に記載の方法。

## (項目 9)

上記第 1 の経路信号処理及び上記第 2 の経路信号処理が、同期または非同期である、上記項目のいずれか一項に記載の方法。

20

## (項目 10)

動的音声強調のシステムであって、

コンピュータ実行可能な命令を格納するように構成されたメモリと、

上記項目のいずれか一項に記載の方法を実装するため、上記コンピュータ実行可能な命令を実行するように構成されたプロセッサと、を備える、上記システム。

## (摘要)

本開示は、音声強調のための方法及びシステムを提供する。本開示の方法及びシステムは、同時に入力信号に関する 2 つの経路の信号処理を実施できる。第 1 の経路信号処理は、オーディオソース入力を受信すること及び第 1 の利得制御パラメータに基づきオーディオソース入力に動的音量バランスングを実施することを含む。第 2 の経路信号処理は、オーディオソース入力に音声検知を実施すること及び検知信頼度を計算することならびに検知信頼度に基づき第 2 の利得制御パラメータを計算することを含む。第 1 の経路信号処理及び第 2 の経路信号処理は、同期または非同期であってよい。本開示の方法は第 1 の利得制御パラメータを第 2 の処理経路によって計算された第 2 の利得制御パラメータで更新すること及び更新された第 1 の利得制御パラメータに基づき第 1 の経路信号処理を実施することを含む。

30

## 【図面の簡単な説明】

## 【0016】

【図 1】本開示の実装形態の 1 つまたは複数の実施形態による、音声強調の概略ブロック図を概略的に示す。

40

【図 2】本開示の 1 つまたは複数の実施形態による、音声検知の概略ブロック図を例示的に示す。

【図 3】本開示の 1 つまたは複数の実施形態による音声検知に基づく利得推定の概略ブロック図を例示的に示す。

【図 4】本開示の 1 つまたは複数の実施形態による、動的音量バランスングプロセスの概略図を例示的に示す。

【図 5】本開示の別の実装形態の 1 つまたは複数の実施形態による、音声強調の概略図を示す。

【図 6】図 5 における実装形態の 1 つまたは複数の実施形態による、動的音量バランス

50

グプロセスの概略図を示す。

【図 7】本開示の 1 つまたは複数の実施形態による、ソース入力が、デュアルチャンネルソース入力である場合のデュアルチャンネルソース入力に基づいてマルチチャンネルソース入力を生成するプロセスを概略的に示す。

【図 8】本開示の 1 つまたは複数の実施形態による、動的音声強調のための方法を概略的に示す。

【発明を実施するための形態】

【0017】

本実施形態の以下の説明は、例示のみを目的とし、制限するものではないことを理解されたい。図に示された機能的ブロック、モジュールまたはユニットの例の分割は、これらの機能的ブロック、モジュールまたはユニットが、物理的に分離されたユニットとして実装されなければならないことを意味すると解釈されるべきではない。図示されたまたは記載された機能的ブロック、モジュールまたはユニットは、別個のユニット、回路、チップ、機能的ブロック、モジュールまたは回路素子として実装され得る。1 つまたは複数の機能的ブロックまたはユニットはまた、共通の回路、チップ、回路素子、またはユニットに実装され得る。

【0018】

単数形の使用（例えば、「a」であるが、これに制限されるものではない）は、品目の数を制限することを意図しない。関係語、例えば、「頂部（top）」、「底部（bottom）」、「左（left）」、「右（right）」、「上方（upper）」、「下方（lower）」、「下（down）」、「上（up）」、「側面（side）」、「第 1 の」、「第 2 の」（「第 3 の」など）、「入口」、「出口」などであるがこれらに制限されない語は、特に記載がない限り、図面に対する特定の参照において説明を明瞭にするため記述された説明で使用され、本開示または添付の特許請求の範囲の範囲を制限することを意図しない。「結合する（couple）」、「結合している（coupling）」、「結合される（being coupled）」、「結合された（coupled）」、「結合器（coupler）」という用語及び同様の用語は、本明細書で広く使用され、固定する、接合する、接着する、締結する、取り付け、関連付ける、挿入する、その上に又はその中に形成する、通信する、またはそうでなければ、直接または間接的に、機械的に、磁気によって、電氣的に、化学的に、及び中間要素に及び 1 つまたは複数の部材に動作可能に関連付けられるための任意の方法またはデバイスを含み得るか、または、統一された方法で別の部材と一体成形される 1 つの部材を含み得るが、それに限定されない。結合していることは、回転を含む、任意の方向で起こり得る。「含む（including）」及び「など（such as）」という用語は、制限的ではなく例示的であり、「であってよい（may）」という単語は、別段の記載がない限り、「であってよい、が必然的ではない」を意味する。任意の他の用語が本開示で使用されるが、図示された実施形態は、例示及び説明目的で与えられた例であり、そのみが本明細書の主題の実施形態ではない。

【0019】

既存の技術的解決策の欠陥を克服し、ユーザにより良好な体験をもたらすため音声出力の品質を改善するために、本開示は、オーディオソース入力での声の可能性を示す検知信頼度に基づいてオーディオソースで（例えば、シアターオーディオソース）、人の声を動的に検知し、声の音量を動的に強化する解決策を提案する。本開示の方法及びシステムは、同時に入力信号に関する 2 つの経路の信号処理を実施することができる。第 1 の経路信号処理は、オーディオソース入力を受信すること及び第 1 の利得制御パラメータに基づいてオーディオソース入力に動的音量ランシングを実施することを含む。第 2 の経路信号処理は、オーディオソース入力に音声検知を実施すること及び検知信頼度を計算することならびに検知信頼度に基づいて第 2 の利得制御パラメータを計算することを含む。第 1 の経路信号処理及び第 2 の経路信号処理は、同期または非同期であってよい。本開示の方法はまた、第 1 の利得制御パラメータを第 2 の処理経路によって計算された第 2 の利得制

10

20

30

40

50

御パラメータで更新すること及び更新された第1の利得制御パラメータに基づいて第1の経路信号処理を実施することを含む。本開示の方法及びシステムは、より良好に音声の了解度を強化し、オーディオ製品を使用するユーザの体験を改善できる。

#### 【0020】

本開示の様々な実装形態の様々な実施形態による動的音声強調の方法及びシステムは、添付の図面を参照して以下に詳細に記載される。図1は、本開示の実装形態の1つまたは複数の実施形態による音声の方法及びシステムの概略ブロック図を示す。理解を容易にするため、本開示を、本方法及びシステムの主要処理手順によるいくつかのモジュールを参照して説明する。モジュールの記述を参照することは、解決策をより明確にするためであり、制限の目的ではないことを当業者は理解するであろう。

10

#### 【0021】

図1は、本開示の実装形態の1つまたは複数の実施形態による概略図を示す。図1に示す1つまたは複数の実施形態では、本開示におけるオーディオソース入力信号を処理する方法及びシステムは、ソース入力モジュール102、動的音量バランシングモジュール104、信号出力モジュール106、音声検知モジュール108、及び利得制御モジュール110を含む。図1から分かるように、本開示の方法及びシステムは、同時に入力信号に関する2つの経路の信号処理を実施することができる。第1の経路信号処理を主に使用して、受信したソース入力信号に対して動的音量バランシングを実施する。第2の経路信号処理を使用して、受信したソース入力信号に音声検知を実施し、利得を推定する。第1の経路信号処理及び第2の経路信号処理は、同期的または非同期的に実施され得る。これは、処理能力及び実際のシステムのレイテンシ要求次第である。このソース入力信号のための二重経路処理設計は、遅延を最小化し、オーディオの歪みを防止する。例えば、第1の経路信号処理及び第2の信号処理が非同期的に実施される場合、一方で、信号は、システム全体を迅速に低遅延で通過することができる。他方で、利得は、比較的低率であると推定され得る。それにより、推定される利得は、より正確で平滑であり、オーディオの歪みを防止する上で大きな一助となる。

20

#### 【0022】

例えば、図1を参照すると、第1の経路信号処理は、ソース入力モジュール102を介してオーディオソース入力信号を受信すること、及び動的音量バランシングモジュール104により、現在の利得制御パラメータに基づいて受信したオーディオソース入力信号に動的バランシングを実施することを含み得る。第2の経路処理は、音声検知モジュール108で入力モジュール102から受信されたオーディオソース入力信号を検知すること、及び検知信頼度を計算することを含み得る。例えば、第2の経路処理はまた、利得制御モジュール110が、計算された検知信頼度に基づいて新規の利得制御パラメータを推定できることを含む。

30

#### 【0023】

利得制御モジュール110によって推定された新規の利得制御パラメータを使用して、動的音量バランシングモジュール104によって現在使用されている利得制御パラメータを更新できる。したがって、動的音量バランシングモジュール104は、更新された利得制御パラメータに基づいて第1の経路信号処理を実施し得る。すなわち、動的音量バランシングモジュール104は、更新された利得制御パラメータに基づいて受信されたオーディオソース入力信号に動的音量バランシングを実施し得る。動的音量バランシングの後のオーディオ信号は、信号出力モジュール106を介して出力され得る。

40

#### 【0024】

オーディオソース入力は、マルチチャンネルソース入力、デュアルチャンネルソース入力、及びシングルチャンネルソース入力を含み得る。異なるソース入力の処理手順を、添付の図面を参照して以下に説明する。図2は、音声入力源がマルチチャンネルソース入力を含む、本開示の1つまたは複数の実施形態による、音声検知の概略ブロック図を例示的に示す。図2に示す音声検知プロセスは、例えば、図1の音声検知モジュール108によって実施され得る。図2に示すように、センタチャンネル抽出がまず実施される。すなわち、センタチ

50

チャンネル信号が、マルチチャンネルソース入力から抽出される。通常、音声信号のほとんどは、センタチャンネルに存在する。次に、入力信号が、同様の水準に調整されるように、抽出されたセンタチャンネル信号に正規化を実施する。例えば、正規化された信号は、以下の式(1)で表される。

$$x_{i\_norm}(n) = (x_i(n) - \mu_i) / \sigma_i \quad (1)$$

式中、 $x_i(n)$ は、第*i*のタイムフレームの第*n*のサンプリング点における入力信号を表し、 $x_{i\_norm}(n)$ は、第*i*のタイムフレームの第*n*のサンプリング点における出力信号を表す。すなわち、正規化された信号 $\mu_i$ 及び $\sigma_i$ は、第*i*のタイムフレームに対応する入力信号の平均及び分散である。

【0025】

次に、高速自己相関処理を、正規化された信号に実施し、自己相関の結果を出力する。例えば、高速自己相関処理は、まず短時間フーリエ変換(STFT)法を使用することによって正規化された入力信号にフーリエ変換を実施し、高速自己相関をフーリエ変換信号に実施し得る。例えば、高速自己相関処理手順を以下の式(2)~(4)に示す。

【数1】

$$X_i(z) = \text{STFT}(x_{i\_norm}(n)) \quad (2)$$

$$c_i(n) = \text{iSTFT}(X_i(z) * \overline{X_i(z)}) \quad (3)$$

$$C_i = \text{norm}(c_i(n)) \quad (4)$$

式中、 $X_i(z)$ は、フーリエ変換信号であり、

【数2】

$$\overline{X_i(z)}$$

は、 $X_i(z)$ の共役を表し、 $\text{iSTFT}$ は、逆短時間フーリエ変換であり、 $c_i(n)$ は、第*i*のタイムフレームの信号の自己相関である。次に、 $c_i(n)$ の規範を計算して、 $C_i$ を得る。例えば、最終自己相関の結果の出力 $C_i$ は、ユークリッドノルムに基づいて取得される。自己相関の結果の出力 $C_i$ は、センタチャンネル信号内の音声の可能性を表し得る検知信頼度を表す。

【0026】

図3は、本開示の1つまたは複数の実施形態による、音声検知に基づいて動的利得を推定する方法及びシステムの概略ブロック図を例示的に示す。図3に示す音声検知に基づいて動的利得を推定するプロセスは、例えば、図1の利得制御モジュール110によって実施され得る。例えば、図2に示すプロセスを参照すると、音声検知モジュール108を介して生成された検知信頼度 $C_i$ は、利得制御モジュール110に対する入力として作用する。入力に基づいて、音声に対する利得(これ以後、利得制御パラメータとも呼ばれてよい)は、利得制御モジュール110内で処理後、動的音量パラランシングモジュール104への入力として出力される。いくつかの例では、利得のダイナミックレンジは、以下の式(5)で計算される。

$$G_i = D_0 * \ln(C_i + D_1) \quad (5)$$

式中 $G_i$ は、動的制御モジュールの出力を表し、 $D_0$ 及び $D_1$ は、動的利得変動範囲の制御パラメータであり、0より大きい実数であり得、 $\ln(\ )$ は、自然対数関数である。いくつかの例では、 $G_i$ は、利得制御モジュール110からの出力として動的音量パラランシングモジュール104に提供され得る。

10

20

30

40

50

## 【 0 0 2 7 】

いくつかのその他の例では、 $G_i$ は、さらに処理されてもよく、その後、利得制御モジュール104からの出力として作用し得る。例えば、 $G_i$ は、平滑化されて、オーディオの歪みを低減する。加えて、ソフトリミッタを使用して、利得 $G_{i\_lim}$ が、妥当な大きさの範囲内に確実にあるようにすることもできる。例えば、以下の式(6)のタンジェント関数は、ソフトリミッタとして使用され得る。

$$G_{i\_lim} = \tanh(G_i + \quad) + \quad \quad (6)$$

式中、 $\quad$ 、 $\quad$ 及び $\quad$ は、リミッタのパラメータであり、システム構成に依存し、 $\quad$ は、0より大きい実数であってよく、 $\quad$ 及び $\quad$ は、0でない実数であってよい。現時点では、 $G_{i\_lim}$ は、利得制御モジュール110からの出力として作用し得る。

10

## 【 0 0 2 8 】

図4は、本開示の1つまたは複数の実施形態による、各チャンネルの動的音量バランスングプロセスの概略図を例示的に示す。図4の動的音量バランスング処理は、動的音量バランスングモジュール104によって実施され得る。例えば、マルチチャンネルソース入力の受信後、動的音量バランスングモジュール104はまず、チャンネル抽出を実施して、センタチャンネル信号を抽出する。次に、センタチャンネル信号の音量を増強して、その他のチャンネル信号の音量を利得制御パラメータに基づいて低減する。次いで、増強されたセンタチャンネル信号及び低減された他のチャンネル信号を結合及び混合して、出力信号を生成する。利得制御パラメータは、現在の利得制御パラメータまたは更新された利得パラメータであってよい。例えば、第1の経路信号処理と第2の経路信号処理とが同期である場合、現在のタイムフレーム(例えば、第*i*のタイムフレーム)の信号の動的音量バランスングに使用される利得制御パラメータは、リアルタイムで更新された計算された利得制御パラメータ、例えば、リアルタイムで更新された $G_i$ または $G_{i\_lim}$ である。第1の経路信号処理と第2の経路信号処理とが、非同期である場合、音声検知及び利得推定を含む第2の経路信号処理の速度が比較的遅いので、現在のタイムフレーム(例えば、第*i*のタイムフレーム)の信号の動的音量バランスングに使用される利得制御パラメータは、 $G_{i-n}$ または $G_{i-n\_lim}$ などの、以前のタイムフレームの信号の動的音量バランスングに使用される利得制御パラメータであってよく、式中、 $n$ は、0より大きい整数であり、その値は、システムの実際の処理能力または技術者の実務経験によって変化し得る。さらに、現在の/更新された利得制御パラメータに基づいて、センタチャンネルの信号及びその他のチャンネルの信号は、それぞれ異なる比で増強及び低減され得る。すなわち、センタチャンネル信号の音量を増強するための強調制御パラメータ及びセンタチャンネル信号の音量を低減するための減衰制御パラメータは、それぞれ現在/更新された利得制御パラメータに基づいてさらに決定され得る。例えば、強調制御パラメータ及び減衰制御パラメータは、システム要件または体験にしたがって、技術者によって設定された比例計算、関数計算、またはその他の計算方法によって決定され得る。結果として、全体的なシステムの音量は変わらないままであるが、各チャンネルの音量は動的にバランスがとれている。

20

30

## 【 0 0 2 9 】

図5は、本開示の別の実装形態の1つまたは複数の実施形態による、方法及びシステムの概略図を示す。図5に示す1つまたは複数の実施形態では、本開示におけるオーディオソース入力信号を処理する方法及びシステムは、ソース入力モジュール502、動的音量バランスングモジュール504、信号出力モジュール506、音声検知モジュール508、及び利得制御モジュール510を含む。これらのモジュールは、図1の102~110の対応するモジュールと実質的に同じ原理で動作する。その上、図5に示す方法及びシステムは、クロスオーバーフィルタリングモジュール512をさらに含み得る。図5に示す処理手順と図1~4を参照して上述された処理手順との間の差が、クロスオーバーフィルタリングが第1の信号経路に追加されることであることが理解されよう。したがって、入力モジュール502から受信されたソース入力信号は、まずクロスオーバーフィルタリングモジュール512によって処理され、その後、動的音量バランスングのため動的音量バランスングモジュール504によって処理される。人の声の周波数域は、基本的に中間周

40

50

波数域にあるので、クロスオーバーフィルタは、入力信号を処理するために選択されて、異なる周波数域の信号を区別することができる。したがって、利得制御は、入力信号の中間周波数域内の信号にのみ適用されるが、入力信号のその他の周波数域の信号は、変化しないままである。追加されたクロスオーバーフィルタリングを介して、非音声周波数域でできるだけ歪みを回避するため、ソース入力信号の中間周波数域内の信号にのみ動的音量バランスを実施することができる。スペースを節約するため、図5及び図1に示す実施形態の異なる部品のみを以下に説明する。他の同一の部品については、図1～4及び関連する説明を参照されたい。

#### 【0030】

図6は、図5の実装形態の1つまたは複数の実施形態による、動的音量バランスプロセスの概略図を示す。図6に示すように、クロスオーバーフィルタリングの後のソース入力信号は、中間周波数、高周波数、及び低周波数域に信号を含み得る。次に、動的音量バランスを中間周波数域内の信号にのみ実施する。動的音量バランスは、センタチャンネル信号を抽出するためチャンネル抽出を含む。そして、センタチャンネル信号の音量を増強して、その他のチャンネル信号の音量を現在の/更新された利得制御パラメータに基づいて低減する。マルチチャンネルソース入力信号の低周波数域及び高周波数域の信号は、動的音量バランスの影響を受けないが、動的音量バランスの後、中間周波数域の信号に直接結合及び混合されて、出力信号を生成する。したがって、非音声信号によって引き起こされた歪みは、回避する方がよい可能性がある。

#### 【0031】

ソース入力があるセンタチャンネルを有するマルチチャンネルソース入力である場合に実施されるいくつかの処理手順を、図1～6と共に説明する。当業者であれば、ソース入力があるシングルチャンネル入力である場合、図1～6に示す処理手順も実行することができ、センタチャンネル抽出のプロセスが省略されることを本開示から理解し得る。すなわち、上述の2つの経路の信号処理は、シングルチャンネルソース入力に直接実施される。

#### 【0032】

ソース入力があるデュアルチャンネルソース入力であるケースでは、マルチチャンネルソース入力があるデュアルチャンネルソース入力に基づいて生成されるように、上述の方法及びシステムを実装する前に、事前にセンタ抽出プロセスを追加することが必要である。図7は、本開示の1つまたは複数の実施形態による、ソース入力があるデュアルチャンネルソース入力である場合の、デュアルチャンネルソース入力に基づいてマルチチャンネルソース入力を生成するプロセスを概略的に示す。

#### 【0033】

図7に示すアップミックスプロセスは、デュアルチャンネルソース入力に基づいてマルチチャンネルソース入力を出力するため、センタ抽出アルゴリズムを取り入れてもよい。センタ抽出アルゴリズムは、例えば、左チャンネルと右チャンネルの入力信号間の相互相関を計算すること、及び左及び右チャンネル入力信号をセンタチャンネル信号に組み合わせることを含んでもよく、組み合わせ比は、以下の式(7)を参照する、相互相関次第である。

$$\text{センタ}(n) = \text{corr}(\text{左}(n), \text{右}(n)) * (\text{左}(n) + \text{右}(n)) \quad (7)$$

式中、左(n)は、左チャンネル入力信号であり、右(n)は、右チャンネル入力信号であり、センタ(n)は、センタチャンネル信号であり、corr()は、相互相関関数を表し、は、実際には調整パラメータであり、は、0より大きく、1以下である。

#### 【0034】

図8は、本開示の1つまたは複数の実施形態による、動的音声強調のための方法を概略的に示す。図8に示すように、方法は、第1の経路信号処理を実施することを含む。第1の経路信号処理は、オーディオソース入力を受信すること及び第1の利得制御パラメータS802に基づいてオーディオソース入力に動的音量バランスを実施することを含む。方法はまた、第2の経路信号処理を実施することを含む。第2の経路信号処理は、オーディオソース入力に音声検知を実施すること及び検知信頼度S804を計算すること、ならびに検知信頼度S806に基づいて第2の利得制御パラメータを計算することを含む。

方法はまた、第1の利得制御パラメータを第2の利得制御パラメータS808で更新すること、及び更新された第1の利得制御パラメータS802に基づいて第1の経路信号処理を実施することを含み得る。図8に示す方法は、プロセッサによって実施され得る。

【0035】

本開示によって提供される方法及びシステムは、サウンドバー及びステレオスピーカなどの消費者向け製品だけでなく、劇場及びコンサートホールなどの、映画の用途における製品にも適用され得る。本開示によって提供される方法及びシステムは、より良好に音声の了解度を強化し、オーディオ製品及びアプリケーションを使用するユーザの体験を改善できる。添付の図面に関して本開示に記載の上述の方法及びシステムは、プロセッサによって実装され得る。

10

【0036】

条項1

動的音声強調のための方法であって、第1の経路信号処理を実施することであって、第1の経路信号処理が、オーディオソース入力を受信すること及び第1の利得制御パラメータに基づいてオーディオソース入力に動的音量ランシングを実施することを含む、第1の経路信号処理を実施することと、第2の経路信号処理を実施することであって、第2の経路信号処理が、オーディオソース入力に音声検知を実施すること及び検知信頼度を計算することを含み、検知信頼度が、オーディオソース入力における音声の可能性を示す、第2の経路信号処理を実施することと、検知信頼度に基づいて第2の利得制御パラメータを計算することと、第1の利得制御パラメータを第2の利得制御パラメータで更新することと、更新された第1の利得制御パラメータに基づいて第1の経路信号処理を実施することと、を含む、方法。

20

【0037】

条項2

オーディオソース入力、マルチチャンネルソース入力を含み、オーディオソース入力に音声検知を実施すること及び検知信頼度を計算することが、センタチャンネル信号をマルチチャンネルソース入力から抽出することと、センタチャンネル信号に正規化を実施することと、正規化されたセンタチャンネル信号に高速自己相関を実施することと、を含む、高速自己相関の結果が、検知信頼度を表す、条項1に記載の方法。

【0038】

30

条項3

検知信頼度に基づいて第2の利得制御パラメータを計算することは、検知信頼度の対数関数に基づいて第2の利得制御パラメータを計算することと、計算された第2の利得制御パラメータを平滑化することと、平滑化された第2の利得制御パラメータを制限することと、を含む、条項1～2のいずれか1つに記載の方法。

【0039】

条項4

オーディオソース入力、マルチチャンネルソース入力を含み、オーディオソース入力に動的音量ランシングを実施することが、マルチチャンネルソース入力からセンタチャンネル信号を抽出することと、センタチャンネル信号の音量を増強させること及び第1の利得制御パラメータまたは更新された第1の利得制御パラメータに基づいて他のチャンネル信号の音量を低減させることと、増強されたセンタチャンネル信号と低減された他のチャンネル信号とを結合及び混合して、出力信号を生成することと、を含む、条項1～3のいずれか1つに記載の方法。

40

【0040】

条項5

動的音量ランシングを実施する前に、オーディオソース入力にクロスオーバーフィルタリングを実施することをさらに含む、条項1～4のいずれか1つに記載の方法。

【0041】

条項6

50

オーディオソース入力の間周波数域の信号にのみ動的音量バランスングを実施することと、動的音量バランスング後にオーディオソース入力の低周波数域及び高周波数域の信号とオーディオソース入力の間周波数域の信号を結合及び混合して、出力信号を生成することと、をさらに含む、条項 1 ~ 5 のいずれか 1 つに記載の方法。

## 【 0 0 4 2 】

## 条項 7

オーディオソース入力、デュアルチャンネルソース入力をさらに含み、方法が、デュアルチャンネルソース入力に基づいてマルチチャンネルソース入力を生成することをさらに含む、条項 1 ~ 6 のいずれか 1 つに記載の方法。

## 【 0 0 4 3 】

## 条項 8

デュアルチャンネルソース入力に基づいてマルチチャンネルソース入力を生成することが、デュアルチャンネルソース入力からの左チャンネル信号と右チャンネル信号との間の相互相関を実施することと、組み合わせ比に従ってマルチチャンネルソース入力生成することと、を含み、組み合わせ比が相互相関の結果次第である、条項 1 ~ 7 のいずれか 1 つに記載の方法。

## 【 0 0 4 4 】

## 条項 9

第 1 の経路信号処理及び第 2 の経路信号処理が、同期または非同期である、条項 1 ~ 8 のいずれか 1 つに記載の方法。

## 【 0 0 4 5 】

## 条項 10

コンピュータ実行可能な命令を格納するように構成されたメモリと、条項 1 ~ 9 のいずれか 1 つに記載の方法を実装するため、コンピュータ実行可能な命令を実行するように構成されたプロセッサと、を備える、動的音声強調のシステム。

## 【 0 0 4 6 】

実装形態の説明は、例示及び説明のために提示されてきた。実装形態は、上記説明に従って適切に修正及び変更されてもよく、またはこれらの修正及び変更は、方法を実践することによって取得されてもよい。例えば、別段の指示がない限り、記載された方法の 1 つまたは複数は、適切なデバイス及び/またはデバイスの組み合わせによって実施され得る。方法は、1 つまたは複数の追加のハードウェア要素（記憶装置、メモリ、ハードウェアネットワークインターフェース/アンテナ、スイッチ、アクチュエータ、クロック回路など）と組み合わせて 1 つまたは複数の論理デバイス（例えば、プロセッサ）を使用して、格納された命令を実施することによって実施され得る。記載された方法及び関連付けられた行為はまた、本出願に記載の順序以外の様々な順序で平行して及び/または同時に実行されてもよい。記載されたシステムは、本質的に例示であり、追加の要素を含む、及び/または要素を省いてもよい。本開示の主題は、すべての開示された様々なシステム及び構成ならびに他の特徴、機能及び/または特性の新規かつ非自明の組み合わせを含む。

## 【 0 0 4 7 】

システムは、追加のまたは異なる論理を含んでもよく、多くの異なる方法で実装されてもよい。プロセッサは、マイクロプロセッサ、マイクロコントローラ、特定用途向け集積回路（ASIC）、デジタル信号処理用プロセッサ DSP、個別論理、あるいは、回路または論理のこれら及び/もしくは他の種類の組み合わせとして実装されてもよい。同様に、メモリは、DRAM、SRAM、フラッシュメモリ、または他の種類のメモリであってもよい。パラメータ（例えば、条件及び閾値）及び他のデータ構造は、別個に格納及び管理されてもよく、単一メモリまたはデータベースに組み合わせられてもよく、あるいは、多くの異なる方法で、論理的及び物理的に組織化されてもよい。プログラム及び命令のセットは、単一のプログラムまたは別個プログラムの一部であってもよく、あるいは、複数のメモリ及びプロセッサにわたって分散されてもよい。

## 【 0 0 4 8 】

10

20

30

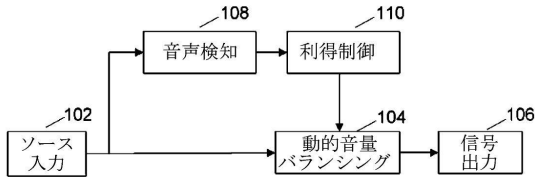
40

50

本出願に使用される、単数形で列記された及び単語「1つの(a)/ (one)」が先行する要素またはステップは、除外が指示されていない限り、複数の前述の要素またはステップを除外しないものと理解されるべきである。さらに、本開示の「1つの実装形態」または「1つの例」への言及は、列挙された特徴も組み入れる追加の実装形態の存在を除外すると解釈されることを意図しない。本発明は、特定の実装形態を参照して上述されてきた。しかし、当業者は、添付の特許請求の範囲に記載される、本発明のより広範な趣旨及び範囲から逸脱することなく、様々な修正及び変更が本明細書になされ得ることを理解するであろう。

【 図面 】

【 図 1 】



【 図 2 】

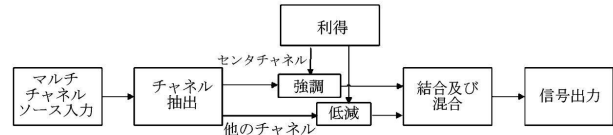
10



【 図 3 】

【 図 4 】

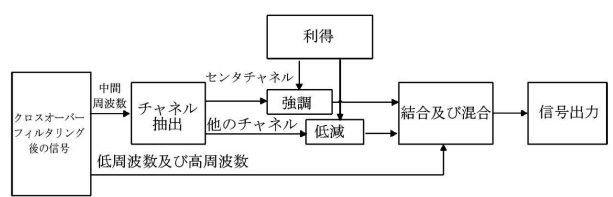
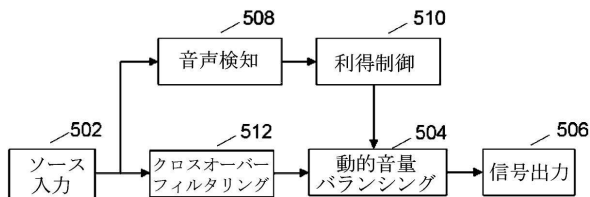
20



【 図 5 】

【 図 6 】

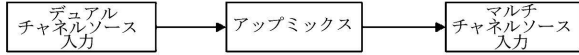
30



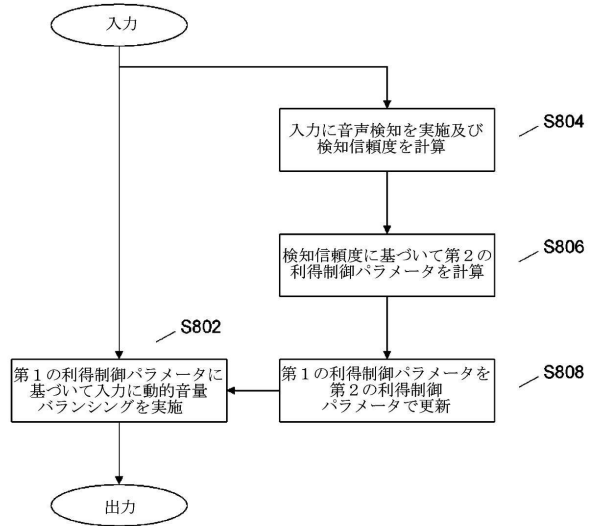
40

50

【 図 7 】



【 図 8 】



10

20

30

40

50

## フロントページの続き

(72)発明者 ジェン ジャンウェン

中華人民共和国 518067 シェンジェン, ナンシャン, シェコウ, ゴンイエ 3アールデー  
ロード, ナンバー1, チャイナ マーチャンツ ポート プラザ, 20エフ

(72)発明者 シャオ イー

中華人民共和国 518067 シェンジェン, ナンシャン, シェコウ, ゴンイエ 3アールデー  
ロード, ナンバー1, チャイナ マーチャンツ ポート プラザ, 20エフ

(72)発明者 ジャオ エヴィン

中華人民共和国 518052 グアン ドン プロピンス, シェンジェン シティ, ナンシャン  
ディストリクト, サウス ケユアン ロード, ナンバー 3331, アリ センター, ビルディ  
ング ティー2, ベースメント

Fターム(参考) 5D162 AA13 CA01 CA15 CB12 CD14 EG02

5D220 AB01