

(19)日本国特許庁(JP)

(12)公表特許公報(A)

(11)公表番号

特表2023-506354

(P2023-506354A)

(43)公表日 令和5年2月16日(2023.2.16)

(51)国際特許分類 F I
 G 1 0 L 21/0272(2013.01) G 1 0 L 21/0272 1 0 0 A
 G 1 0 L 21/0308(2013.01) G 1 0 L 21/0308 A

審査請求 有 予備審査請求 有 (全68頁)

(21)出願番号	特願2022-520412(P2022-520412)	(71)出願人	500341779
(86)(22)出願日	令和2年10月2日(2020.10.2)		フラウンホーファー・ゲゼルシャフト・
(85)翻訳文提出日	令和4年5月31日(2022.5.31)		ツール・フェルデルング・デル・アンゲ
(86)国際出願番号	PCT/EP2020/077716		ヴァンテン・フォルシュング・アインゲ
(87)国際公開番号	WO2021/064204		トラージェネル・フェライン
(87)国際公開日	令和3年4月8日(2021.4.8)		ドイツ連邦共和国, 8 0 6 8 6 ミュン
(31)優先権主張番号	19201575.8		ヘン, ハンザシュトラッセ 2 7 ツェー
(32)優先日	令和1年10月4日(2019.10.4)	(71)出願人	509347479
(33)優先権主張国・地域又は機関	欧州特許庁(EP)		テヒニツシェ ウニフェルジテート イル
(81)指定国・地域	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,A T,BE,BG,CH,CY,CZ,DE,DK,EE,ES,FI,FR ,GB,GR,HR,HU,IE,IS,IT,LT,LU,LV,MC, 最終頁に続く		ドイツ連邦共和国 9 8 6 9 3 イルメナ ウ エレンベルクシュトラッセ 2 9
		(74)代理人	100134119
			弁理士 奥町 哲行
		(72)発明者	シュラー・ゲラルト
			最終頁に続く

(54)【発明の名称】 ソース分離

(57)【要約】

本実施例は、異なるソース（例えば、オーディオソース）に関連する複数の出力信号を取得するための方法、装置、および技術に関する。一例では、第1の入力信号（502、M₀）またはその処理されたバージョンを第2の入力信号（M₁）の遅延されスケールングされたバージョン（503₁）と組み合わせ、第1の出力信号（504、S₀）を取得する、および

第2の入力信号（502、M₁）またはその処理されたバージョンを、前記第1の入力信号（M₀）の遅延されスケールングされたバージョン（503₀）と組み合わせ、第2の出力信号（504、S₁）を取得する、ことを行うことができる。

例えば、ランダム方向最適化（560）を使用して、

スケールング値（564、a₀、a₁）を決定し、第1および第2の入力信号の遅延およびスケールングされたバージョン（503₀）を取得することと、

遅延値（564、d₀、d₁）を決定し、第1および第2の入力信号の遅延およびスケールングされたバージョンを取得することと、が可能である。

10

【特許請求の範囲】

【請求項 1】

複数の入力信号 (502、 M_0 、 M_1) に基づいて、異なる音源 (source₀、source₁) に関連付けられた複数の出力信号 (504、 S'_0 、 S'_1) を取得するための装置 (500) であって、前記音源 (source₀、source₁) からの信号 (S_0 、 S_1) が結合され (501)、

前記装置は、第 1 の入力信号 (502、 M_0) またはその処理されたバージョンを第 2 の入力信号 (M_1) の遅延およびスケールリングされたバージョン (503₁) と組み合わせて (510)、第 1 の出力信号 (504、 S'_0) を取得するように構成され、

前記装置は、第 2 の入力信号 (502、 M_1)、またはその処理されたバージョンを前記第 1 の入力信号 (M_0) の遅延およびスケールリングされたバージョン (503₀) と組み合わせて (510)、第 2 の出力信号 (504、 S'_1) を取得するように構成され、

前記装置は、ランダム方向最適化 (560) を使用して、

前記第 1 の入力信号 (502、 M_0) の前記遅延されスケールリングされたバージョン (503₀) を取得するために使用される第 1 のスケールリング値 (564、 a_0) と、

前記第 1 の入力信号 (502、 M_0) の前記遅延されスケールリングされたバージョン (503₀) を取得するために使用される第 1 の遅延値 (564、 d_0) と、

前記第 2 の入力信号 (502、 M_1) の前記遅延されスケールリングされたバージョン (503₁) を取得するために使用される第 2 のスケールリング値 (564、 a_1) と、

前記第 2 の入力信号の前記遅延およびスケールリングされたバージョン (503₁) を取得するために使用される第 2 の遅延値 (564、 d_1) と、

を決定するように構成されている、

装置 (500)。

【請求項 2】

前記第 1 の入力信号 (502、 M_0) と結合されるべき前記第 2 の入力信号 (502、 M_1) の前記遅延されスケールリングされたバージョン (503₁) は、前記第 2 の入力信号 (502、 M_1) に部分遅延を適用することによって得られる、請求項 1 に記載の装置。

【請求項 3】

前記第 2 の入力信号 (502、 M_1) と結合されるべき前記第 1 の入力信号 (502、 M_0) の前記遅延されスケールリングされたバージョン (503₀) は、前記第 1 の入力信号 (502、 M_0) に部分遅延を適用することによって得られる、請求項 1 または 2 に記載の装置。

【請求項 4】

正規化された大きさの値の第 1 のセットのそれぞれの要素 (P_0) と、

正規化された大きさの値の第 1 のセット (522) のそれぞれの要素 (P_0) および

正規化された大きさの値の第 2 のセット (522) のそれぞれの要素 (P_1)

に基づいて形成された商 (702') の対数 (706') と、

の間の複数の積 (712'、710') を合計する (714、716) ように構成されており、

その結果、前記第 1 の正規化された大きさの値のセット (P_0) によって記述される信号部分 (s'_0) と前記第 2 の正規化された大きさの値のセット (P_1) によって記述される信号部分 (s'_1) との間の類似性または非類似性を記述する値 (D_{KL} 、 D 、532) を取得する (530)、請求項 1 から 3 のいずれか一項に記載の装置。

【請求項 5】

前記ランダム方向最適化 (560) は、候補パラメータが候補のベクトル (564、 a_0 、 a_1 、 d_0 、 d_1) を形成するようなものであり、前記候補のベクトルは、前記候補のベクトルをランダムな方向に修正することによって反復的に洗練される、請求項 1 から 4 のいずれか一項に記載の装置。

【請求項 6】

10

20

30

40

50

前記ランダム方向最適化 (560) は、候補パラメータが候補のベクトルを形成するようなものであり、前記候補のベクトルは、前記候補のベクトルをランダムな方向に修正することによって反復的に洗練される、請求項 1 から 5 のいずれか一項に記載の装置。

【請求項 7】

前記ランダム方向最適化 (560) が、前記第 1 の出力信号と前記第 2 の出力信号との間の前記類似性または非類似性を示すメトリックおよび / または値 (D_{KL} 、 D 、532) が測定され、前記第 1 の出力測定値および前記第 2 の出力測定値が、前記値に関連付けられた前記候補パラメータに関連付けられた測定値、または最も低い類似性または最も高い非類似性を示すメトリックであるように選択される、請求項 1 から 6 のいずれか一項に記載の装置。

10

【請求項 8】

複数の入力信号 (502、 M_0 、 M_1) に基づいて、異なる音源 ($source_1$ 、 $source_2$) に関連付けられた複数の出力信号 (504、 S'_0 、 S'_1) を取得するための装置 (500) であって、前記音源 ($source_1$ 、 $source_2$) からの信号が結合され (501)、

前記装置は、第 1 の出力信号 (504、 S'_0) を得るために、第 1 の入力信号 (502、 M_0)、またはその処理されたバージョンを、第 2 の入力信号 (502、 M_1) の遅延およびスケールされたバージョン (503₁) と結合するように構成され (510)、前記装置は、前記第 2 の入力信号 (502、 M_1) に部分遅延 (d_1) を適用するように構成され、

20

前記装置は、第 2 の入力信号 (502、 M_1)、またはその処理バージョンを前記第 1 の入力信号 (502、 M_0) の遅延およびスケールバージョン (503₀) と組み合わせて (510)、第 2 の出力信号 (504、 S'_1) を取得するように構成され、前記装置は、前記第 1 の入力信号 (502、 M_0) に部分遅延 (d_0) を適用するように構成され、

前記装置は、最適化 (560) を使用して、

前記第 1 の入力信号 (502、 M_0) の前記遅延されスケールされたバージョン (503₀) を取得するために使用される第 1 のスケール値 (564、 a_0) と、

前記第 1 の入力信号 (502、 M_0) の前記遅延されスケールされたバージョン (503₀) を取得するために使用される第 1 の部分遅延値 (564、 d_0) と、

30

前記第 2 の入力信号 (502、 M_1) の前記遅延されスケールされたバージョン (503₁) を取得するために使用される第 2 のスケール値 (564、 a_1) と、

前記第 2 の入力信号 (502、 M_1) の前記遅延されスケールされたバージョン (503₁) を取得するために使用される第 2 の部分遅延値 (564、 d_1) と、

を決定するように構成されている、

装置 (500)。

【請求項 9】

前記最適化がランダム方向最適化 (560) である、請求項 8 に記載の装置。

【請求項 10】

正規化された大きさの値の第 1 のセットのそれぞれの要素 (P_0) と、

正規化された大きさの値の第 1 のセットのそれぞれの要素 (P_0) および

正規化された大きさの値の第 2 のセットのそれぞれの要素 (P_1)

に基づいて形成された商 (702') の対数 (706') と、

の間の複数の積 (710、712) を合計 (714、716) するように構成され、

その結果、前記第 1 の正規化された大きさの値のセット (P_0) によって記述される信号部分 (s_0') と前記第 2 の正規化された大きさの値のセット (P_1) によって記述される信号部分 (s_1') との間の類似性または非類似性を記述する値 (D_{KL} 、 D 、532) を取得する (530)、請求項 8 または 9 のいずれか一項に記載の装置。

40

【請求項 11】

複数の入力信号 (M_0 、 M_1) に基づいて、異なる音源 ($source_0$ 、 $source_1$)

50

e_1)に関連付けられた複数の出力信号(504、 S'_0 、 S'_1)を取得するための装置(500)であって、前記音源からの信号が結合され、

前記装置が、第1の入力信号(502、 M_0)またはその処理バージョンを第2の入力信号(502、 M_1)の遅延およびスケールバージョン(503₁)と組み合わせて(510)、第1の出力信号(504、 S'_0)を取得するように構成され、

前記装置は、第2の入力信号(502、 M_1)またはその処理バージョンを前記第1の入力信号(502、 M_0)の遅延およびスケールバージョン(503₀)と組み合わせて(510)、第2の出力信号(504、 S'_1)を取得するように構成され、

前記装置は、

正規化された大きさの値の第1のセットのそれぞれの要素(P_0)と、

正規化された大きさの値の第1のセットのそれぞれの要素(P_0)および

正規化された大きさの値の第2のセットのそれぞれの要素(P_1)

に基づいて形成された商(702')の対数(706')と、

の間の複数の積を合算(714、716)するように構成され、

その結果、前記第1の正規化された大きさの値のセット(P_0)によって記述される信号部分(s_0')と前記第2の正規化された大きさの値のセットによって記述される信号部分(s_1')との間の類似性または非類似性を記述する値(D_{KL} 、 D 、532)を取得する(530)、

装置(500)。

【請求項12】

前記第1の入力信号(502、 M_0)の前記遅延されスケールされたバージョンを取得するために使用される第1のスケール値(564、 a_1)と、

前記第1の入力信号の前記遅延およびスケールされたバージョンを取得するために使用される第1の遅延値(564、 d_0)と、

前記第2の入力信号の前記遅延およびスケールされたバージョンを取得するために使用される第2のスケール値(564、 a_1)と、

最適化を使用して、前記第2の入力信号の前記遅延およびスケールされたバージョンを取得するために使用される第2の遅延値(564、 d_1)と、

のうちの少なくとも1つを決定するように構成されている、請求項11に記載の装置。

【請求項13】

前記第1の遅延値(d_0)は部分遅延である、請求項12に記載の装置。

【請求項14】

前記第2の遅延値(d_1)は部分遅延である、請求項11から13のいずれか一項に記載の装置。

【請求項15】

前記最適化はランダム方向最適化(560)である、請求項12から14のいずれか一項に記載の装置。

【請求項16】

前記第1のスケール値および前記第2のスケール値ならびに前記第1の遅延値および前記第2の遅延値のうちの少なくとも一方は、前記出力信号の前記相互情報または関連する尺度を最小化することによって得られる、請求項1から15のいずれか一項に記載の装置。

【請求項17】

前記最適化を反復的に実行するためのオブティマイザ(560)をさらに備え、前記オブティマイザは、各反復において、現在の候補ベクトルが現在の最良の候補ベクトルよりも良好に機能するかどうかを評価するための前記現在の候補ベクトルをランダムに生成するように構成され、

前記オブティマイザは、前記現在の候補ベクトルに関連して、物理信号間の類似度または非類似度に関連する目的関数を評価するように構成され、

前記オブティマイザは、前記現在の候補ベクトルが前記目的関数を前記現在の最良の候

10

20

30

40

50

補ベクトルに対して減少させる場合に、前記現在の候補ベクトルを前記新しい現在の最良の候補ベクトルとしてレンダリングするように構成されている、請求項 1 から 16 のいずれか一項に記載の装置。

【請求項 18】

前記第 1 の入力信号 (502、 M_0)、またはその処理されたバージョンを、前記時間領域および/または前記 z 変換もしくは周波数領域において、前記第 2 の入力信号 (502、 M_1) の前記遅延されスケーリングされたバージョン (503₁) と結合するようにさらに構成され、

前記第 2 の入力信号 (502、 M_1)、またはその処理されたバージョンを、前記時間領域および/または前記 z 変換もしくは周波数領域において前記第 1 の入力信号 (502、 M_0) の前記遅延およびスケーリングされたバージョン (503₀) と組み合わせるようにさらに構成されている、

10

請求項 1 から 17 のいずれか一項に記載の装置。

【請求項 19】

前記最適化は、前記時間領域および/または前記 z 変換もしくは周波数領域で実行される、

請求項 1 から 18 のいずれか一項に記載の装置。

【請求項 20】

前記第 2 の入力信号 (502、 M_1) に適用される前記遅延または部分遅延 (d_0) は、

20

前記第 1 のマイクロフォン (mic_0) によって受信された前記第 1 の音源 ($source_0$) からの前記信号、および

前記第 2 のマイクロフォン (mic_1) によって受信された前記第 1 の音源 ($source_0$) からの前記信号、

間の関係、および/または差もしくは到着を示す、

請求項 1 から 19 のいずれか一項に記載の装置。

【請求項 21】

前記第 1 の入力信号 (502、 M_0) に適用される前記遅延または部分遅延 (d_1) は、

前記第 2 のマイクロフォン (mic_1) によって受信された前記第 2 の音源 ($source_1$) からの前記信号、および

30

前記第 1 のマイクロフォン (mic_0) によって受信される前記第 2 の音源 ($source_1$) からの前記信号、

間の関係、および/または差もしくは到着を示す、

請求項 1 から 20 のいずれか一項に記載の装置。

【請求項 22】

異なる候補パラメータ (a_0 、 a_1 、 d_0 、 d_1) が反復的に選択されて処理され、前記候補パラメータの各々についてメトリック (532) が測定されるように最適化 (560) を実行するように構成され、前記メトリック (532) は類似度メトリックまたは非類似度メトリックであり、前記出力信号の最も低い類似度、または最大の非類似度を示す前記メトリックに関連付けられた前記候補パラメータ (a_0 、 a_1 、 d_0 、 d_1) を使用して前記第 1 の入力信号 (502、 M_0) および前記第 2 の入力信号 (502、 M_1) を処理して結合する、請求項 1 から 21 のいずれか一項に記載の装置。

40

【請求項 23】

反復ごとに、前記候補パラメータは、前記第 2 の入力信号 (502、 M_1) に適用される候補遅延 (d_0) を含み、前記候補遅延 (d_0) は、

前記第 1 のマイクロフォン (mic_0) によって受信された前記第 1 の音源 ($source_0$) からの前記信号、および

前記第 2 のマイクロフォン (mic_1) によって受信された前記第 1 の音源 ($source_0$) からの前記信号、

50

の間の関係、および/または差もしくは到着の候補に関連付けられ得る、
請求項 22 に記載の装置。

【請求項 24】

反復ごとに、前記候補パラメータは、前記第 1 の入力信号 (502、 M_0) に適用される候補遅延 (d_1) を含み、前記候補遅延 (d_1) は、

前記第 2 のマイクロフォン (mic_1) によって受信された前記第 2 の音源 ($source_1$) からの前記信号、および

前記第 1 のマイクロフォン (mic_0) によって受信される前記第 2 の音源 ($source_1$) からの前記信号、

の間の関係、および/または差もしくは到着の候補に関連付け得る、

請求項 22 または 23 に記載の装置。

10

【請求項 25】

反復ごとに、前記候補パラメータは、前記第 2 の入力信号 (502、 M_1) に適用される候補相対減衰値 (564、 a_0) を含み、前記候補相対減衰値 (564、 a_0) は、

前記第 1 の音源 ($source_0$) から前記第 1 のマイクロフォン (mic_0) によって受信された前記信号の振幅、および

前記第 1 の音源 ($source_0$) から前記第 2 のマイクロフォン (mic_1) によって受信された前記信号の振幅

の間の候補関係および/または候補差を示す、

請求項 22 または 23 または 24 に記載の装置。

20

【請求項 26】

反復ごとに、前記候補パラメータ (564) は、前記第 1 の入力信号 (502、 M_0) に適用される候補相対減衰値 (a_1) を含み、前記候補相対減衰値 (a_1) は、

前記第 2 の音源 ($source_1$) から前記第 2 のマイクロフォン (mic_1) によって受信された前記信号の振幅、および

前記第 2 の音源 ($source_1$) から前記第 1 のマイクロフォン (mic_0) によって受信された前記信号の振幅

の間の候補関係および/または候補差を示す、

請求項 22 または 23 または 24 または 25 に記載の装置。

30

【請求項 27】

異なる反復のために少なくとも 1 つの候補パラメータを変更するように構成される、請求項 22 から 26 のいずれか一項に記載の装置。

【請求項 28】

先行する反復のための少なくとも 1 つの候補パラメータから後続の反復のための少なくとも 1 つの候補パラメータまでの少なくとも 1 つのステップをランダムに選択することによって、異なる反復のための少なくとも 1 つの候補パラメータを変更するように構成される、請求項 22 から 27 のいずれか一項に記載の装置。

【請求項 29】

前記少なくとも 1 つのステップをランダムに選択するように構成される、請求項 28 に記載の装置。

40

【請求項 30】

少なくとも 1 つのステップが、予め選択された重みによって重み付けされる、請求項 29 に記載の装置。

【請求項 31】

前記少なくとも 1 つのステップは、予め選択された重みによって制限される、請求項 29 または 30 に記載の装置。

【請求項 32】

候補パラメータ (a_0 、 a_1 、 d_0 、 d_1) が候補のベクトルを形成し、反復ごとに、前記候補のベクトルは、乱数のベクトルを適用することによって撓動され、前記乱数のベクトルは、前記候補のベクトルの前記要素によって要素ごとに乗算されるか、または前記

50

候補のベクトルの前記要素に加算される、請求項 22 から 31 のいずれか一項に記載の装置。

【請求項 33】

反復ごとに、前記候補のベクトルがステップに対して修正される、請求項 32 に記載の装置。

【請求項 34】

前記反復回数が所定の最大回数に制限される、請求項 22 から 33 のいずれか一項に記載の装置。

【請求項 35】

前記メトリック (532) は、カルバックライブラー発散として処理される、請求項 7 および 22 から 34 のいずれか一項に記載の装置。 10

【請求項 36】

前記メトリック (532) は、カルバックライブラー発散を用いて処理される、請求項 7 および 22 から 35 のいずれか一項に記載の装置。

【請求項 37】

前記メトリックは、前記クロストークが少ないほど、その値が高くなるようなものである、請求項 7 および 22 から 36 のいずれか一項に記載の装置。

【請求項 38】

前記メトリック (532) は、前記第 1 の信号および前記第 2 の信号 (M_0 、 M_1) の各々について、正規化された大きさの値のセットのそれぞれの要素を含む、請求項 22 から 37 のいずれか一項に記載の方法。 20

【請求項 39】

前記第 1 および第 2 の入力信号 (M_0 、 M_1) の少なくとも一方について、前記それぞれの要素は、前記候補パラメータから得られる前記候補第 1 または第 2 の出力信号 (S'_0 、 S'_1) に基づく、請求項 38 に記載の装置。

【請求項 40】

前記第 1 および第 2 の入力信号 (M_0 、 M_1) の少なくとも一方について、前記それぞれの要素は、

候補の第 1 または第 2 の出力信号 (S'_0 、 S'_1) に関連付けられた値、および

前記第 1 の出力信号または前記第 2 の出力信号の前記以前に取得された値に関連付けられたノルムの間の分数として取得される、請求項 39 に記載の装置。 30

【請求項 41】

前記第 1 および第 2 の入力信号 (M_0 、 M_1) のうちの少なくとも一方について、前記それぞれの要素は、

$$P_i(n) = \frac{|s'_i(n)|}{\|s'_i\|_1} \quad (7)$$

40

によって取得される、請求項 39 または 40 に記載の装置。

【請求項 42】

$P_i(n)$ が 0 と 1 との間に含まれる、請求項 41 に記載の装置。

【請求項 43】

前記メトリックは、

正規化された大きさの値の前記第 1 のセットの前記それぞれの要素および

正規化された大きさの値の第 2 のセットのそれぞれの要素

に基づいて形成される商の対数を含み、

その結果、前記第 1 の正規化された大きさの値のセット (P_0) によって記述される信 50

号部分 (s₀') と前記第 1 の正規化された大きさの値のセット (P₀) によって記述される信号部分 (s₁') との間の値 (5 3 2) を取得する (5 3 0) 、

請求項 2 2 から 4 2 のいずれかに記載の装置。

【請求項 4 4】

前記メトリックは、

$$D_{KL}(P||Q) = \sum_n P(n) \log \left(\frac{P(n)}{Q(n)} \right) \quad (6)$$

10

の形態で取得され、式中、P (n) は前記第 1 の入力信号に関連付けられた要素であり、Q (n) は前記第 2 の入力信号に関連付けられた要素である、

請求項 7 または 2 2 から 4 3 のいずれか一項に記載の装置。

【請求項 4 5】

前記メトリックは、

$$D(P_0, P_1) = - \sum_n \left[P_0(n) \log \left(\frac{P_0(n)}{P_1(n)} \right) + P_1(n) \log \left(\frac{P_1(n)}{P_0(n)} \right) \right] \quad (8)$$

20

の形態で取得され、

式中、P₁ (n) は、前記第 1 の入力信号に関連付けられた要素であり、P₂ (n) は、前記第 2 の入力信号に関連付けられた要素である、

請求項 2 2 から 4 3 のいずれか一項に記載の装置。

【請求項 4 6】

スライディングウィンドウを使用して前記最適化を実行するように構成される、請求項 2 2 から 4 5 のいずれか一項に記載の装置。

【請求項 4 7】

前記取得された第 1 および第 2 の出力信号 (S₀'、S₁') に関連する情報を周波数領域に変換するようにさらに構成される、請求項 1 から 4 6 のいずれか一項に記載の装置。

30

【請求項 4 8】

前記取得された第 1 および第 2 の出力信号 (S₀'、S₁') に関連する情報を符号化するようにさらに構成される、請求項 1 から 4 7 のいずれか一項に記載の装置。

【請求項 4 9】

前記取得された第 1 および第 2 の出力信号 (S₀'、S₁') に関連する情報を格納するようにさらに構成された、請求項 1 から 4 8 のいずれか一項に記載の装置。

【請求項 5 0】

前記取得された第 1 および第 2 の出力信号 (S₀'、S₁') に関連する情報を送信するようにさらに構成される、請求項 1 から 4 9 のいずれか一項に記載の装置。

40

【請求項 5 1】

前記第 1 の入力信号 (5 0 2、M₀) を取得するための第 1 のマイクロフォン (m i c₀) 、および前記第 2 の入力信号 (5 0 2、M₁) を取得するための第 2 のマイクロフォン (m i c₁) のうちの少なくとも 1 つと、を含む、請求項 1 から 5 0 のいずれか一項に記載の装置。

【請求項 5 2】

前記取得された第 1 および第 2 の出力信号 (S₀'、S₁') に関連する情報を送信するための機器と、請求項 1 から 5 1 のいずれか一項に記載の装置とを含む、テレビ会議のための装置。

50

【請求項 5 3】

請求項 1 から 5 2 のいずれか一項に記載の装置を含むバイノーラルシステム。

【請求項 5 4】

物理信号に関連する物理パラメータを反復的に最適化するためのオブティマイザ(560)であって、前記オブティマイザは、各反復において、現在の候補ベクトルが現在の最良の候補ベクトルよりも良好に機能するかどうかを評価するための前記現在の候補ベクトルをランダムに生成するように構成され、

前記オブティマイザは、前記現在の候補ベクトルに関連して、物理信号間の類似度または非類似度に関連する目的関数を評価するように構成され、

前記オブティマイザは、前記現在の候補ベクトルが前記目的関数を前記現在の最良の候補ベクトルに対して減少させる場合に、前記現在の候補ベクトルを前記新しい現在の最良の候補ベクトルとしてレンダリングするように構成される、オブティマイザ(560)。

【請求項 5 5】

前記物理信号は、異なるマイクロフォンによって得られたオーディオ信号を含む、請求項 5 4 に記載のオブティマイザ。

【請求項 5 6】

前記パラメータは、特定のマイクロフォンで得られたオーディオ信号の遅延および/またはスケール係数を含む、請求項 5 4 または 5 5 に記載のオブティマイザ。

【請求項 5 7】

前記目的関数は、カルバックライブラー発散である、請求項 5 4 から 5 6 のいずれか一項に記載のオブティマイザ。

【請求項 5 8】

前記カルバックライブラー発散が、正規化された大きさの値の第 1 および第 2 のセットに適用される、請求項 5 7 に記載のオブティマイザ。

【請求項 5 9】

前記目的関数は、

正規化された大きさの値の第 1 のセットのそれぞれの要素 (P_0) と、

正規化された大きさの値の前記第 1 のセット (522) の記それぞれの要素 (P_0) および

正規化された大きさの値の第 2 のセット (522) のそれぞれの要素 (P_1)

に基づいて形成された商 (702') の対数 (706') と、

の間の複数の積 (712'、710') を合計 (714、716) することによって得られ、

その結果、前記第 1 の正規化された大きさの値のセット (P_0) によって記述される信号部分 (s'_0) と前記第 2 の正規化された大きさの値のセット (P_1) によって記述される信号部分 (s'_1) との間の類似性または非類似性を記述する値 (D_{KL} 、 D 、532) を取得する (530)、

請求項 5 7 または 5 8 に記載のオブティマイザ。

【請求項 6 0】

前記目的関数は、

$$D(P_0, P_1) = - \sum_n \left[P_0(n) \log \left(\frac{P_0(n)}{P_1(n)} \right) + P_1(n) \log \left(\frac{P_1(n)}{P_0(n)} \right) \right] \quad (8)$$

または

10

20

30

40

50

$$D_{KL}(P||Q) = \sum_n P(n) \log \left(\frac{P(n)}{Q(n)} \right) \quad (6)$$

のように得られ、式中、 $P_1(n)$ または $P(n)$ は、前記第1の入力信号に関連付けられた要素であり、 $P_2(n)$ または $Q(n)$ は、前記第2の入力信号に関連付けられた要素である、

請求項57または58または59に記載のオプティマイザ。

10

【請求項61】

複数の入力信号(502、 M_0 、 M_1)に基づいて、異なる音源($source_0$ 、 $source_1$)に関連付けられた複数の出力信号(504、 S'_0 、 S'_1)を取得するための方法であって、前記音源($source_0$ 、 $source_1$)からの信号が結合され、

前記方法は、

第1の出力信号(504、 S'_0)を得るために、第1の入力信号(502、 M_0)またはその処理されたバージョンを第2の入力信号(502、 M_1)の遅延されスケールされたバージョン(503₁)と結合するステップと、

第2の入力信号(502、 M_1)またはその処理されたバージョンを、前記第1の入力信号(502、 M_0)の遅延されスケールされたバージョン(503₀)と結合して、第2の出力信号(504、 S'_1)を取得するステップと、

20

ランダム方向最適化(560)を使用して、

前記第1の入力信号(502、 M_0)の前記遅延されスケールされたバージョン(503₀)を取得するために使用される第1のスケール値(564、 a_0)と、

前記第1の入力信号(502、 M_0)の前記遅延されスケールされたバージョン(503₀)を取得するために使用される第1の遅延値(564、 d_0)と、

前記第2の入力信号(502、 M_1)の前記遅延されスケールされたバージョン(503₁)を取得するために使用される第2のスケール値(564、 a_1)と、

前記第2の入力信号の前記遅延およびスケールされたバージョン(503₁)を取得するために使用される第2の遅延値(564、 d_1)と、

30

のうちの少なくとも1つを決定するステップと、を含む、方法。

【請求項62】

複数の入力信号(502、 M_0 、 M_1)に基づいて、異なる音源($source_1$ 、 $source_2$)に関連付けられた複数の出力信号(504、 S'_0 、 S'_1)を取得するための方法であって、前記音源($source_1$ 、 $source_2$)からの信号が結合され、

前記方法は、

第1の出力信号(504、 S'_0)を得るために、第1の入力信号(502、 M_0)またはその処理されたバージョンを第2の入力信号(502、 M_1)の遅延およびスケールされたバージョン(503₁)と結合するステップ(510)であって、前記方法が、部分遅延(d_1)を第2の入力信号(502、 M_1)に適用するように構成される、ステップ、

40

第2の出力信号(504、 S'_1)を得るために、第2の入力信号(502、 M_1)、またはその処理されたバージョンを、前記第1の入力信号(502、 M_0)の遅延およびスケールされたバージョン(503₀)と結合するステップ(510)であって、前記方法が、部分遅延(d_0)を前記第1の入力信号(502、 M_0)に適用するように構成される、ステップ、

最適化を使用して、

前記第1の入力信号(502、 M_0)の前記遅延されスケールされたバージョン

50

(5030)を取得するために使用される第1のスケーリング値(564、 a_0)と、
前記第1の入力信号(502、 M_0)の前記遅延されスケーリングされたバージョン
(5030)を取得するために使用される第1の部分遅延値(564、 d_0)と、
前記第2の入力信号(502、 M_1)の前記遅延されスケーリングされたバージョン
(5031)を取得するために使用される第2のスケーリング値(564、 a_1)と、
前記第2の入力信号(502、 M_1)の前記遅延されスケーリングされたバージョン
(5031)を取得するために使用される第2の部分遅延値(564、 d_1)と、
のうちの少なくとも1つを決定するステップ、を含む、方法。

【請求項63】

複数の入力信号(M_0 、 M_1)に基づいて、異なる音源($source_0$ 、 $source_1$)に対応する複数の出力信号(504、 S'_0 、 S'_1)を取得するための方法であって、前記音源からの信号を結合し、前記方法は、

第1の入力信号(502、 M_0)またはその処理されたバージョンを第2の入力信号(502、 M_1)の遅延およびスケーリングされたバージョン(5031)と結合して、第1の出力信号(504、 S'_0)を取得するためのもの、

第2の入力信号(502、 M_1)またはその処理されたバージョンを、前記第1の入力信号(502、 M_0)の遅延されスケーリングされたバージョン(5030)と合成して、第2の出力信号(504、 S'_1)を取得するためのもの、

正規化された大きさの値の第1のセットのそれぞれの要素(P_0)と、

正規化された大きさの値の前記第1のセットの前記それぞれの要素(P_0)および

正規化された大きさの値の第2のセットのそれぞれの要素(P_1 、 Q)

に基づいて形成された商(702')の対数(706')と、の間の複数の積を合計し(714、716)、

その結果、前記第1の正規化された大きさの値のセット(P_0)によって記述される信号部分(s_0')と、前記第2の正規化された大きさの値のセット(P_0)によって記述される信号部分(s_1' (n))との間の類似性または非類似性を記述する値(532)を取得する(530)ためのもの、である、方法。

【請求項64】

前記製品請求項のいずれか一項に記載の機器を使用するように構成された、請求項61から63のいずれか一項に記載の方法。

【請求項65】

前記部分遅延(d_1)は、前記第2のソース($source_1$)から前記第1のマイクロフォン(mic_0)に到達する前記信号(M_0)の前記遅延と、前記第2のソース($source_1$)から前記第2のマイクロフォン(mic_1)に到達する前記信号(M_1)の前記遅延(H_1 、 1)との間の関係および/または差を示す、請求項61から64のいずれか一項に記載の方法。

【請求項66】

物理信号に関連する物理パラメータを反復的に最適化するための最適化方法であって、前記方法は、反復ごとに、現在の候補ベクトルが現在の最良の候補ベクトルよりも良好に機能するかどうかを評価するための前記現在の候補ベクトルをランダムに生成することを含み、

前記オブティマイザは、前記現在の候補ベクトルに関連して、物理信号間の類似度または非類似度に関連する目的関数を評価するように構成され、

前記オブティマイザは、前記現在の候補ベクトルが前記目的関数を前記現在の最良の候補ベクトルに対して減少させる場合に、前記現在の候補ベクトルを前記新しい現在の最良の候補ベクトルとしてレンダリングするように構成される、最適化方法。

【請求項67】

プロセッサによって実行されると、前記プロセッサに請求項61から66のいずれか一項に記載の方法を実行させる命令を格納する非一時的記憶ユニット。

【発明の詳細な説明】

10

20

30

40

50

【技術分野】

【0001】

本実施例は、異なるソース（例えば、オーディオソース）に関連する複数の出力信号を取得するための方法および装置に関する。本実施例はまた、信号分離のための方法および装置に関する。本実施例は、テレビ会議のための方法および装置にも言及する。分離のための技術（例えば、音源分離）も開示される。高速時間領域ステレオ音源分離（例えば、部分遅延フィルタを使用する）のための技法も論じる。

【背景技術】

【0002】

図1は、50aで示されるマイクロフォンの設定を示す。マイクロフォン50aは、ここでは互いに5cm（50mm）の距離で示されている第1のマイクロフォンmic₀および第2のマイクロフォンmic₁を含むことができる。他の距離も可能である。2つの異なるソース（source₀およびsource₁）がここに示されている。角度 θ_0 および θ_1 によって識別されるように、それらは異なる位置に（ここでは互いに対して異なる向きにも）配置される。

【0003】

source₀およびsource₁から複数の入力信号M₀およびM₁（マルチチャンネルまたはステレオとしてまとめて示されるマイクロフォンからの入力信号502）が得られる。source₀はS₀としてインデックス付けされたオーディオ音を生成するが、source₁はS₁としてインデックス付けされたオーディオ音を生成する。

【0004】

マイクロフォン信号M₀、M₁は、例えば入力信号と考えるてもよい。ステレオ信号502の代わりに、3つ以上のチャンネルを有するマルチチャンネルを考えることができる。

【0005】

いくつかの例（例えば、M₀およびM₁以外の他の追加の入力チャンネル）では、2つのチャンネルのみが主に説明されているが、入力信号は3つ以上であってもよい。それにもかかわらず、本例は、任意のマルチチャンネル入力信号に対して有効である。例では、信号M₀およびM₁は、例えば格納されたオーディオファイルから取得され得るので、マイクロフォンによって直接取得される必要もない。

【0006】

図2aおよび図4は、音源source₀およびsource₁とマイクロフォンmic₀およびmic₁との間の相互作用を示す。例えば、source₀は、主にマイクロフォンmic₀に到達し、かつマイクロフォンmic₁にも到達するオーディオ音S₀を生成する。生成された音源S₁が主にマイクロフォンmic₁に到達し、マイクロフォンmic₀にも到達するsource₁についても同様である。音S₀がマイクロフォンmic₀に到達するのに必要な時間は、マイクロフォンmic₁に到達するのに必要な時間よりも短いことが、図2aおよび図4から分かる。同様に、音S₁は、mic₀に到達するのに要する時間よりも、mic₁に到達するのに要する時間が短い。マイクロフォンmic₁に到達するときの信号S₀の強度は、一般に、mic₀に到達するときに対して減衰され、逆もまた同様である。

【0007】

したがって、マルチチャンネル入力信号502において、チャンネル信号M₀、M₁は、source₀、source₁からの信号S₀、S₁が互いに組み合わせられたものとなる。したがって、分離技術が追求される。

【発明の概要】

【0008】

以下、角括弧および丸括弧内のテキストは非限定的な例を示す。

一態様によれば、複数の入力信号〔例えば、マイクロフォン信号〕に基づいて異なる音源〔source₀、source₁〕に関連する複数の出力信号〔S'₀、S'₁〕を取得するための装置〔例えば、マルチチャンネルまたはステレオオーディオソース分離装置〕

10

20

30

40

50

であって、音源 [$source_0$ 、 $source_1$] からの信号が結合され、

装置は、第 1 の入力信号 [M_0]、またはその処理 [例えば、遅延および / またはスケールリングされる] バージョンを、第 2 の入力信号 [例えば、 M_1] の遅延およびスケールリングされたバージョン [$a_1 \cdot z^{-d_1} \cdot M_1$] [例えば、第 1 の入力信号から第 2 の入力信号の遅延されスケールリングされたバージョンを減算することによって、例えば、 $S'_0 = M_0(z) - a_1 \cdot z^{-d_1} \cdot M_1(z)$] と組み合わせ、第 1 の出力信号 [S'_0] を取得するように構成され、

装置は、第 2 の入力信号 [M_1]、またはその処理 [例えば、遅延および / またはスケールリングされる] バージョンを、第 1 の入力信号 [M_0] [例えば、第 2 の入力信号から第 1 の入力信号の遅延されスケールリングされたバージョンを減算することによって、例えば、 $S'_1 = M_1(z) - a_0 \cdot z^{-d_0} \cdot M_0(z)$] の遅延およびスケールリングされたバージョン [$a_0 \cdot z^{-d_0} \cdot M_0$] と組み合わせ、第 2 の出力信号 [S'_1] を取得するように構成され、

装置は、ランダム方向最適化 [例えば、他の請求項で定義されている動作のうちの 1 つを実行することによって、例えば、および / または、例えば式 (6) および / または (8) のものであり得る目的関数を最小化する遅延値および減衰値を見つけることによって] を使用して、

第 1 の入力信号 [M_0] の遅延およびスケールリングされたバージョン [$a_0 \cdot z^{-d_0} \cdot M_0$] を取得するために使用される第 1 のスケールリング値 [a_0]、

第 1 の入力信号 [M_0] の遅延およびスケールリングされたバージョン [$a_0 \cdot z^{-d_0} \cdot M_0$] を取得するために使用される第 1 の遅延値 [d_0]、

第 2 の入力信号 [M_1] の遅延されスケールリングされたバージョン [$a_1 \cdot z^{-d_1} \cdot M_1$] を取得するために使用される第 2 のスケールリング値 [a_1] と、

第 2 の入力信号 [$a_1 \cdot z^{-d_1} \cdot M_1$] の遅延されスケールリングされたバージョンを取得するために使用される第 2 の遅延値 [d_1] と、を決定するように構成される、装置を提供する。

【 0 0 0 9 】

第 2 の入力信号 [M_1] の遅延およびスケールリングされたバージョン [$a_1 \cdot z^{-d_1} \cdot M_1$] は、第 1 の入力信号 [M_0] と組み合わせることができ、第 2 の入力信号 [M_1] に部分遅延を適用することによって得られる。

【 0 0 1 0 】

第 1 の入力信号 [M_0] の遅延およびスケールリングされたバージョン [$a_0 \cdot z^{-d_0} \cdot M_0$] は、第 2 の入力信号 [M_1] と組み合わせることができ、第 1 の入力信号 [M_0] に部分遅延を適用することによって得られる。

【 0 0 1 1 】

装置は、

正規化された大きさの値の第 1 のセット [例えば、式 (7) のように] のそれぞれの要素 [$P_i(n)$ 、ここで、 i は 0 または 1 である] と、

正規化された大きさの値の第 1 のセットのそれぞれの要素 [$P(n)$ または $P_1(n)$] および正規化された大きさの値の第 2 のセットのそれぞれの要素 [$Q(n)$ または $Q_1(n)$] に基づいて形成された商の対数と、

の間の複数の積 [例えば、式 (6) または (8) のように] を合計することができ、

その結果、第 1 の正規化された大きさの値のセット [$P_0(n)$ 、 $n = 1 \sim \dots$ の場合] によって記述される信号部分 [$s_0'(n)$] と第 2 の正規化された大きさの値のセット [$P_1(n)$ 、 $n = 1 \sim \dots$ の場合] によって記述される信号部分 [$s_1'(n)$] との間の類似性 [または非類似性] を記述する値 [式 (6) または式 (8) 中の $DKL(P || Q)$ または $D(P_0, P_1)$] を取得する。

【 0 0 1 2 】

ランダム方向最適化は、候補パラメータが候補のベクトル [例えば、4 つのエントリを有し、例えば、 a_0 、 a_1 、 d_0 、 d_1 に対応する] を形成するようなものであってもよ

く、ベクトルは、ランダム方向にベクトルを修正することによって反復的に洗練される [例えば、異なる反復において、請求項 5 0 7 f f も参照されたい]。

【 0 0 1 3 】

ランダム方向最適化は、候補パラメータが候補のベクトル [例えば、4つのエントリを有し、例えば、 a_0 、 a_1 、 d_0 、 d_1 に対応する] を形成するようなものであってもよく、ベクトルは、ランダム方向にベクトルを修正することによって反復的に洗練される [例えば、異なる反復では、以下も参照されたい]。

【 0 0 1 4 】

ランダム方向最適化は、第 1 および第 2 の出力信号間の類似性 (または非類似性) を示すメトリックおよび / または値が測定され、第 1 および第 2 の出力測定値が、値に関連する候補パラメータに関連する測定値、または最も低い類似性 (または最も高い非類似性) を示すメトリックであるように選択されるようなものであってもよい。

【 0 0 1 5 】

第 1 のスケーリング値および第 2 のスケーリング値ならびに第 1 の遅延値および第 2 の遅延値の少なくとも一方は、出力信号の相互情報または関連する尺度を最小化することによって得ることができる。

【 0 0 1 6 】

一態様によれば、複数の入力信号 [例えば、マイクロフォン信号] [M_0 、 M_1] に基づいて、異なる音源 [$source_1$ 、 $source_2$] に関連する複数の出力信号 [S'_0 、 S'_1] を取得するための装置が提供され、音源 [$source_1$ 、 $source_2$] からの信号は結合され、

装置が、第 1 の入力信号 [M_0] またはその処理 [例えば、遅延および / またはスケーリングされる] バージョンを、第 2 の入力信号 [M_1] の遅延およびスケーリングされたバージョン [$a_1 \cdot z^{-d_1} \cdot M_1$] と組み合わせて第 1 の出力信号 [S'_0] を得るように構成され、装置が、部分遅延 [d_1] を第 2 の入力信号 [M_1] に適用するように構成され [部分遅延 (d_1) は、第 2 のソース ($source_1$) から第 1 のマイクロフォン (mic_0) に到達する信号 ($H_{1,0} \cdot S_1$) の遅延 (例えば、 $H_{1,0}$ で表される遅延) と、第 2 のソース ($source_1$) から第 2 のマイクロフォン (mic_1) に到達する信号 ($H_{1,1} \cdot S_1$) の遅延 (例えば、 $H_{1,1}$ で表される遅延) との関係および / または差を示すことができる] [例では、部分遅延 d_1 は、分数 $H_{1,0}(z) / H_{1,1}(z)$ の結果の z 項の指数を近似するものとして理解することができる]、

装置は、第 2 の入力信号 [M_1]、またはその処理 [例えば、遅延および / またはスケーリングされる] バージョンを、第 1 の入力信号 [M_0] の遅延およびスケーリングされたバージョン [$a_0 \cdot z^{-d_0} \cdot M_0$] と組み合わせて第 2 の出力信号 [S'_1] を取得するように構成され、装置は、第 1 の入力信号 [M_0] に部分遅延 [d_0] を適用するように構成され [部分遅延 (d_0) は、第 1 の音源 ($source_0$) から第 1 のマイクロフォン (mic_0) に到達する信号 ($H_{0,0} \cdot S_0$) の遅延 (例えば、 $H_{0,0}$ で表される遅延) と、第 1 の音源 ($source_0$) から第 2 のマイクロフォン (mic_1) に到達する信号 ($H_{0,1} \cdot S_0$) の遅延 (例えば、 $H_{0,1}$ で表される遅延) との間関係および / または差を示すことができる] [例では、部分遅延 d_0 は、分数 $H_{0,1}(z) / H_{0,0}(z)$ の結果の z 項の指数を近似するものとして理解することができる]、

装置は、最適化を使用して、

第 1 の入力信号 [M_0] の遅延およびスケーリングされたバージョン [$a_0 \cdot z^{-d_0} \cdot M_0$] を取得するために使用される第 1 のスケーリング値 [a_0]、

第 1 の入力信号 [M_0] の遅延およびスケーリングされたバージョン [$a_0 \cdot z^{-d_0} \cdot M_0$] を取得するために使用される第 1 の部分遅延値 [d_0]、

第 2 の入力信号 [M_1] の遅延されスケーリングされたバージョン [$a_1 \cdot z^{-d_1} \cdot M_1$] を取得するために使用される第 2 のスケーリング値 [a_1] と、

第 2 の入力信号 [M_1] の遅延およびスケーリングされたバージョン [$a_1 \cdot z^{-d_1} \cdot M_1$] を取得するために使用される第 2 の部分遅延値 [d_1] と、を決定するように構

成される。

【0017】

最適化はランダム方向最適化であってもよい。

装置は、

正規化された大きさの値の第1のセット [例えば、式(7)のように] のそれぞれの要素 $[P_i(n)]$ 、ここで、 i は0または1である]と、

正規化された大きさの値の第1のセットのそれぞれの要素 $[P(n)]$ または $[P_1(n)]$ および正規化された大きさの値の第2のセットのそれぞれの要素 $[Q(n)]$ または $[Q_1(n)]$ に基づいて形成された商の対数と、

の間の複数の積 [例えば、式(6)または(8)のように] を合計することができ、

その結果、第1の正規化された大きさの値のセット $[P_0(n)]$ 、 $n = 1 \sim \dots$ の場合] によって記述される信号部分 $[s_0'(n)]$ と第2の正規化された大きさの値のセット $[P_1(n)]$ 、 $n = 1 \sim \dots$ の場合] によって記述される信号部分 $[s_1'(n)]$ との間の類似性 [または非類似性] を記述する値 [式(6)または式(8)中の $DKL(P||Q)$ または $D(P_0, P_1)$] を取得する。

10

【0018】

一態様によれば、複数の入力信号 [例えば、マイクロフォン信号] $[M_0, M_1]$ に基づいて、異なる音源 $[source_0, source_1]$ に関連する複数の出力信号 $[S_0, S_1]$ を取得するための装置 [例えば、マルチチャネルまたはステレオオーディオソース分離装置] が提供され、音源からの信号が結合され [例えば、第1の入力信号から第2の入力信号の遅延およびスケールされたバージョンを減算することによって、および/または第2の入力信号から第1の入力信号の遅延およびスケールされたバージョンを減算することによって]、

20

装置は、第1の入力信号 $[M_0]$ 、またはその処理 [例えば、遅延および/またはスケールされる] バージョンを、第2の入力信号 $[M_1]$ [例えば、第1の入力信号から第2の入力信号の遅延されスケールされたバージョンを減算することによって] の遅延およびスケールされたバージョン $[a_1 \cdot z^{-d_1} \cdot M_1]$ と組み合わせて、第1の出力信号 $[S_0]$ を取得するように構成され、

装置は、第2の入力信号 $[M_1]$ 、またはその処理 [例えば、遅延および/またはスケールされる] バージョンを、第1の入力信号 $[M_0]$ [例えば、第2の入力信号から第1の入力信号の遅延されスケールされたバージョンを減算することによって] の遅延およびスケールされたバージョン $[a_0 \cdot z^{-d_0} \cdot M_0]$ と組み合わせて、第2の出力信号 $[S_1]$ を取得するように構成され、

30

装置は、

正規化された大きさの値の第1のセット [例えば、式(7)のように] のそれぞれの要素 $[P_i(n)]$ 、ここで、 i は0または1である]と、

正規化された大きさの値の第1のセットのそれぞれの要素 $[P(n)]$ または $[P_1(n)]$ および正規化された大きさの値の第2のセットのそれぞれの要素 $[Q(n)]$ または $[Q_1(n)]$ に基づいて形成される商の対数と、

の間の複数の積 [例えば、式(6)または(8)のように] を合計するように構成され

40

、その結果、第1の正規化された大きさの値のセット $[P_0(n)]$ 、 $n = 1 \sim \dots$ の場合] によって記述される信号部分 $[s_0'(n)]$ と第2の正規化された大きさの値のセット $[P_1(n)]$ 、 $n = 1 \sim \dots$ の場合] によって記述される信号部分 $[s_1'(n)]$ との間の類似性 [または非類似性] を記述する値 [式(6)または式(8)中の $DKL(P||Q)$ または $D(P_0, P_1)$] を取得する。

【0019】

本装置は、

第1の入力信号 $[M_0]$ の遅延されスケールされたバージョンを取得するために使用される第1のスケール値 $[a_1]$ と、

50

第 1 の入力信号の遅延およびスケーリングされたバージョンを取得するために使用される第 1 の遅延値 $[d_0]$ と、

第 2 の入力信号の遅延されスケーリングされたバージョンを取得するために使用される第 2 のスケーリング値 $[a_1]$ と、

最適化 [例えば、「修正された KLD 計算」に基づいて、] を使用して、第 2 の入力信号の遅延およびスケーリングされたバージョンを取得するために使用される第 2 の遅延値 $[d_1]$ と、を決定し得る。

【0020】

第 1 の遅延値 $[d_0]$ は、部分遅延であってもよい。第 2 の遅延値 $[d_1]$ は、部分遅延である。

最適化はランダム方向最適化であってもよい。

【0021】

装置は、時間領域におけるプロセスの少なくともいくつかを実行することができる。装置は、 z 変換または周波数領域における処理のうちの少なくともいくつかを実行し得る。

【0022】

本装置は、

時間領域および / または z 変換もしくは周波数領域において、第 1 の入力信号 $[M_0]$ またはその処理 [例えば、遅延および / またはスケーリングされる] バージョンを、第 2 の入力信号 $[M_1]$ の遅延およびスケーリングされたバージョン $[a_1 \cdot z^{-d_1} \cdot M_1]$ と結合し、

第 2 の入力信号 $[M_1]$ 、またはその処理 [例えば、遅延および / またはスケーリングされる] バージョンを、時間領域および / または z 変換もしくは周波数領域において第 1 の入力信号 $[M_0]$ の遅延およびスケーリングされたバージョン $[a_0 \cdot z^{-d_0} \cdot M_0]$ と結合するように、構成され得る。

【0023】

最適化は、時間領域および / または z 変換もしくは周波数領域で実行することができる。

第 2 の入力信号 $[M_1]$ に印加される部分遅延 (d_0) は、

第 1 のマイクロフォン $[mic_0]$ によって受信された第 1 のソース $[source_0]$ からの信号 $[S_0 \cdot H_{0,0}(z)]$ 、および

第 2 のマイクロフォン $[mic_1]$ によって受信された第 1 のソース $[source_0]$ からの信号 $[S_0 \cdot H_{0,1}(z)]$ の間の関係、および / または差もしくは到着を示し得る。

【0024】

第 1 の入力信号 $[M_0]$ に適用される部分遅延 (d_1) は、

第 2 のマイクロフォン $[mic_1]$ によって受信された第 2 のソース $[source_1]$ からの信号 $[S_1 \cdot H_{1,1}(z)]$ 、および

第 1 のマイクロフォン $[mic_0]$ によって受信された第 2 のソース $[source_1]$ からの信号 $[S_1 \cdot H_{1,0}(z)]$ の間の関係、および / または差もしくは到着を示し得る。

【0025】

装置は、最適化 [例えば、異なる候補パラメータ $[a_0, a_1, d_0, d_1]$ が反復的に選択されて処理されるような最適化] を実行することができ、候補パラメータの各々についてメトリック [例えば、式 (6) または (8) のように] [例えば、「修正された KLD 計算」に基づいて] [例えば、目的関数] が測定され、メトリックは類似度メトリック (または非類似度メトリック) であり、最も低い類似度 (または最大の非類似度) を示すメトリックに関連付けられた候補パラメータ $[a_0, a_1, d_0, d_1]$ を使用して取得された第 1 の入力信号 $[M_0]$ および第 2 の入力信号 $[M_0]$ を選択する。[類似性は、第 1 の入力信号と第 2 の入力信号との間の統計的依存性 (またはそれに関連する値、例えば式 (7) の値) として推測することができ、および / または非類似性は、第 1 の入力

10

20

30

40

50

信号と第 2 の入力信号との間の統計的独立性（または、式（7）の値などの、それらに関連付けられた値）として推測することができる]

【0026】

反復ごとに、候補パラメータは、第 2 の入力信号 [M_1] に適用される候補遅延 (d_0) [例えば、候補部分遅延] を含むことができ、候補遅延 (d_0) は、以下の間の候補関係および / または候補差もしくは到着に関連付けることができる：

第 1 のマイクロフォン [mic_0] によって受信された第 1 のソース [$source_0$] からの信号 [$S_0 \cdot H_{0,0}(z)$]、および

第 2 のマイクロフォン [mic_1] によって受信された第 1 のソース [$source_0$] からの信号 [$S_0 \cdot H_{0,1}(z)$]。

10

【0027】

反復ごとに、候補パラメータは、第 1 の入力信号 [M_0] に適用される候補遅延 (d_1) [例えば、候補部分遅延] を含み、候補遅延 (d_1) は、以下の間の候補関係および / または候補差もしくは到着に関連付けることができる：

第 2 のマイクロフォン [mic_1] によって受信された第 2 のソース [$source_1$] からの信号 [$S_1 \cdot H_{1,1}(z)$]、および

第 1 のマイクロフォン [mic_0] によって受信された第 2 のソース [$source_1$] からの信号 [$S_1 \cdot H_{1,0}(z)$]。

【0028】

反復ごとに、候補パラメータは、第 2 の入力信号 [M_1] に適用される候補相対減衰値 [a_0] を含むことができ、候補相対減衰値 [a_0] は、以下の間の候補関係および / または候補差を示す：

20

第 1 のマイクロフォン [mic_0] によって第 1 のソース [$source_0$] から受信された信号 [$S_0 \cdot H_{0,0}(z)$] の振幅、および

第 1 の音源 [$source_0$] から第 2 のマイクロフォン [mic_1] によって受信された信号 [$S_0 \cdot H_{0,1}(z)$] の振幅。

【0029】

反復ごとに、候補パラメータは、第 1 の入力信号 [M_0] に適用される候補相対減衰値 [a_1] を含むことができ、候補相対減衰値 [a_1] は、以下の間の候補関係および / または候補差を示す：

30

第 2 のマイクロフォン [mic_1] によって第 2 のソース [$source_1$] から受信された信号 [$S_1 \cdot H_{1,1}(z)$] の振幅、および

第 2 の音源 [$source_1$] から第 1 のマイクロフォン [mic_0] によって受信された信号 [$S_1 \cdot H_{1,0}(z)$] の振幅。

【0030】

装置は、先行する反復の少なくとも 1 つの候補パラメータから後続の反復の少なくとも 1 つの候補パラメータまでの少なくとも 1 つのステップをランダムに選択することによって、異なる反復の少なくとも 1 つの候補パラメータを変更することができる [例えば、ランダム方向最適化]。

【0031】

40

装置は、少なくとも 1 つのステップ [例えば、アルゴリズム 1 の線 10 における $coefficient\ variation$] をランダムに選択することができる [例えば、ランダム方向最適化]。

少なくとも 1 つのステップは、予め選択された重み [例えば、アルゴリズム 1 のライン 5 における $coeff\ weight$] によって重み付けされてもよい。

少なくとも 1 つのステップは、予め選択された重み [例えば、アルゴリズム 1 のライン 5 における $coeff\ weight$] によって制限される。

【0032】

装置は、候補パラメータ [a_0 、 a_1 、 d_0 、 d_1] が候補のベクトルを形成するようになっていてもよく、反復ごとに、候補のベクトルは、候補のベクトルの要素によって要

50

素ごとに乗算される（または加算される）一様分布乱数 [例えば、それぞれ - 0 . 5 から + 0 . 5 の間] のベクトルを適用することによって摂動される [例えば、ランダムに] 。
 [勾配処理を回避することができる] [例えば、ランダム方向最適化] 。

反復ごとに、候補のベクトルは、ステップ [例えば、それぞれ - 0 . 5 から + 0 . 5 の間である] に対して修正される（例えば、摂動される）。

【 0 0 3 3 】

装置は、反復の数値が所定の最大数に制限されるようになっていてもよく、所定の最大数は 1 0 から 3 0 の間である（例えば、2 0、サブセクション 2 . 3 のように、最後の 3 行）。

メトリックは、カルバックライブラー発散として処理され得る。

10

メトリックは、以下に基づいてもよい：

【 0 0 3 4 】

第 1 の信号および第 2 の信号 [M_0 、 M_1] の各々に対して、正規化された大きさの値の第 1 のセット [例えば、式 (7) のように] のそれぞれの要素 [$P_i(n)$ 、ここで、 i は 0 または 1 である] 。 [錯覚は、時間領域サンプルの正規化された大きさの値を確率分布として考慮し、その後、メトリックを測定すること（例えば、カルバックライブラー発散として、例えば、式 (6) または (8) ）であつてもよい]

【 0 0 3 5 】

第 1 および第 2 の入力信号 [M_0 、 M_1] の少なくとも一方について、それぞれの要素 [$P_i(n)$] は、候補パラメータ [例えば、式 (7) のように] から得られる候補の第 1 または第 2 の出力信号 [S'_0 、 S'_1] に基づくことができる。

20

【 0 0 3 6 】

第 1 および第 2 の入力信号 [M_0 、 M_1] の少なくとも一方について、それぞれの要素 [$P_i(n)$] は、候補パラメータ [例えば、式 (7) のように] から得られる候補の第 1 または第 2 の出力信号 [S'_0 、 S'_1] に基づくことができる。

【 0 0 3 7 】

第 1 および第 2 の入力信号 [M_0 、 M_1] の少なくとも一方について、それぞれの要素 [$P_i(n)$] は、

候補の第 1 または第 2 の出力信号 [$S'_0(n)$ 、 $S'_1(n)$] [例えば、絶対値で] に関連付けられた値 [例えば、絶対値]、および

30

第 1 または第 2 の出力信号 [$S'_0(\dots n - 1)$ 、 $S'_1(\dots n - 1)$] の以前に取得された値に関連付けられたノルム [例えば、1 - ノルム] の間の分数として取得され得る。

【 0 0 3 8 】

第 1 および第 2 の入力信号 [M_0 、 M_1] の少なくとも一方について、それぞれの要素 [$P_i(n)$] は、

$$P_i(n) = \frac{|s'_i(n)|}{\|s'_i\|_1} \tag{7}$$

40

によって取得され得る。

（ここで、「 $s'_i(n)$ 」および「 s'_i 」は、 z 変換されないため、大文字なしで記述される）。

【 0 0 3 9 】

メトリックは、

正規化された大きさの値の第 1 のセットのそれぞれの要素 [$P(n)$ または $P_1(n)$]、および

正規化された大きさの値の第 2 のセットのそれぞれの要素 [$Q(n)$ または $Q_1(n)$] に基づいて形成される商の対数を含むことができ、

50

その結果、

第1の正規化された大きさの値のセット [$P_0(n)$ 、 $n = 1 \sim \dots$ の場合] によって記述される信号部分 [$s_0'(n)$] と第2の正規化された大きさの値のセット [$P_1(n)$ 、 $n = 1 \sim \dots$ の場合] によって記述される信号部分 [$s_1'(n)$] との間の類似性 [または非類似性] を記述する値 [式 (6) または式 (8) 中の $D_{KL}(P||Q)$ または $D(P_0, P_1)$] を取得する。

【0040】

メトリックは、以下の形式で取得することができ、

$$D_{KL}(P||Q) = \sum_n P(n) \log \left(\frac{P(n)}{Q(n)} \right) \quad (6)$$

10

$P(n)$ は、第1の入力信号に関連付けられた要素であり [例えば、正規化された大きさの値の第1のセットの $P_1(n)$ または要素]、 $Q(n)$ は、第2の入力信号に関連付けられた要素である [例えば、正規化された大きさの値の第2のセットの $P_2(n)$ または要素]。

【0041】

メトリックは、以下の形式で取得することができる。

$$D(P_0, P_1) = - \sum_n \left[P_0(n) \log \left(\frac{P_0(n)}{P_1(n)} \right) + P_1(n) \log \left(\frac{P_1(n)}{P_0(n)} \right) \right] \quad (8)$$

20

【0042】

$P_1(n)$ は、第1の入力信号に関連付けられた要素であり [例えば、正規化された大きさの値の第1のセットの $P_1(n)$ または要素]、 $P_2(n)$ は、第2の入力信号に関連付けられた要素である [例えば、正規化された大きさの値の第2のセットの要素]。

30

【0043】

装置は、スライディングウィンドウを使用して最適化を実行することができる [例えば、最適化は、最後の $0.1s \sim 1.0s$ のTDサンプルを考慮に入れることができる]。

装置は、取得された第1および第2の出力信号 (S'_0 、 S'_1) に関連する情報を周波数領域に変換することができる。

【0044】

装置は、取得された第1および第2の出力信号に関連する情報を符号化することができる (S'_0 、 S'_1)。

装置は、取得された第1および第2の出力信号 (S'_0 、 S'_1) に関連する情報を格納することができる。

40

装置は、取得された第1および第2の出力信号 (S'_0 、 S'_1) に関連する情報を送信することができる。

【0045】

前述の請求項のいずれか一項に記載の装置は、第1の入力信号 [M_0] を取得するための第1のマイクロフォン (mic_0)、および第2の入力信号 [M_1] を取得するための第2のマイクロフォン (mic_1) のうちの少なくとも1つを含むことができる [例えば、一定の距離で]。

【0046】

上記のような装置と、取得された第1および第2の出力信号 (S'_0 、 S'_1) に関連す

50

る情報を送信するための機器とを含む、テレビ会議のための装置を提供することができる。

上記のような装置を含むバイノーラルシステムが開示される。

【0047】

物理信号に関連する物理パラメータを反復的に最適化するためのオブティマイザが開示され、オブティマイザは、各反復において、現在の候補ベクトルが現在の最良の候補ベクトルよりも良好に機能するかどうかを評価するための現在の候補ベクトルをランダムに生成するように構成され、

オブティマイザは、現在の候補ベクトルに関連して、物理信号間の類似度または非類似度に関連する目的関数を評価するように構成され、

オブティマイザは、現在の候補ベクトルが目的関数を現在の最良の候補ベクトルに対して減少させる場合に、現在の候補ベクトルを新しい現在の最良の候補ベクトルとしてレンダリングするように構成される。

【0048】

物理信号は、異なるマイクロフォンによって取得されたオーディオ信号を含むことができる。

パラメータは、特定のマイクロフォンで取得されたオーディオ信号の遅延および/またはスケール係数を含むことができる。

【0049】

目的関数は、カルバックライブラー発散である。カルバックライブラー発散は、正規化された大きさの値の第1および第2のセットに適用されてもよい。

【0050】

目的関数は、

正規化された大きさの値の第1のセット [例えば、式(7)のように] のそれぞれの要素 [$P_i(n)$ 、ここで、 i は0または1である] と、

正規化された大きさの値の第1のセットのそれぞれの要素 [$P(n)$ または $P_1(n)$] および正規化された大きさの値の第2のセットのそれぞれの要素 [$Q(n)$ または $Q_1(n)$] に基づいて形成される商の対数と、

の間の複数の積 [例えば、式(6)または(8)のように] を合計することによって得ることができる。

その結果、第1の正規化された大きさの値のセット [$P_0(n)$ 、 $n = 1 \sim \dots$ の場合] によって記述される信号部分 [$s_0'(n)$] と第2の正規化された大きさの値のセット [$P_1(n)$ 、 $n = 1 \sim \dots$ の場合] によって記述される信号部分 [$s_1'(n)$] との間の類似性 [または非類似性] を記述する値 [式(6)または式(8)中の $D_{KL}(P||Q)$ または $D(P_0, P_1)$] を取得する。

【0051】

目的関数は、以下のように取得することができ、

$$D(P_0, P_1) = - \sum_n \left[P_0(n) \log \left(\frac{P_0(n)}{P_1(n)} \right) + P_1(n) \log \left(\frac{P_1(n)}{P_0(n)} \right) \right] \quad (8)$$

または

$$D_{KL}(P||Q) = \sum_n P(n) \log \left(\frac{P(n)}{Q(n)} \right) \quad (6)$$

10

20

30

40

50

$P_1(n)$ または $P(n)$ は、第 1 の入力信号に関連付けられた要素であり [例えば、正規化された大きさの値の第 1 のセットの $P_1(n)$ または要素]、 $P_2(n)$ または $Q(n)$ は、第 2 の入力信号に関連付けられた要素である。

【0052】

一例によれば、複数の入力信号 [例えば、マイクロフォン信号] [M_0 、 M_1] に基づいて異なる音源 [$source_0$ 、 $source_1$] に関連する複数の出力信号 [S'_0 、 S'_1] を取得するための方法が提供され、音源 [$source_0$ 、 $source_1$] からの信号は結合され、

方法は、

第 1 の出力信号 [S'_0] を得るために、第 1 の入力信号 [M_0]、またはその処理 [例えば、遅延および/またはスケーリングされる] バージョンを、第 2 の入力信号 [M_1] [例えば、第 1 の入力信号から第 2 の入力信号の遅延されスケーリングされたバージョンを減算することによって、例えば、 $S'_0 = M_0(z) - a_1 \cdot z^{-d_1} \cdot M_1(z)$] の遅延およびスケーリングされたバージョン [$a_1 \cdot z^{-d_1} \cdot M_1$] と結合するステップと、

第 2 の出力信号 [S'_1] を得るために、第 2 の入力信号 [M_1]、またはその処理 [例えば、遅延および/またはスケーリングされる] バージョンを、第 1 の入力信号 [M_0] [例えば、第 2 の入力信号から第 1 の入力信号の遅延されスケーリングされたバージョンを減算することによって、例えば、 $S'_1 = M_1(z) - a_0 \cdot z^{-d_0} \cdot M_0(z)$] の遅延およびスケーリングされたバージョン [$a_0 \cdot z^{-d_0} \cdot M_0$] と結合するステップと、

ランダム方向最適化 [例えば、他の請求項で定義されている動作のうちの一つを実行することによって、例えば、および/または、例えば式 (6) および/または (8) のものであり得る目的関数を最小化する遅延値および減衰値を見つけることによって] を使用して、

第 1 の入力信号 [M_0] の遅延およびスケーリングされたバージョン [$a_0 * z^{-d_0} * M_0$] を取得するために使用される第 1 のスケーリング値 [a_0]、

第 1 の入力信号 [M_0] の遅延およびスケーリングされたバージョン [$a_0 * z^{-d_0} * M_0$] を取得するために使用される第 1 の遅延値 [d_0]、

第 2 の入力信号 [M_1] の遅延およびスケーリングされたバージョン [$a_1 * z^{-d_1} * M_1$] を取得するために使用される第 2 のスケーリング値 [a_1]、および

第 2 の入力信号 [$a_1 * z^{-d_1} * M_1$] の遅延およびスケーリングされたバージョンを取得するために使用される第 2 の遅延値 [d_1] を決定するステップと、を含む。

【0053】

一例によれば、複数の入力信号 [例えば、マイクロフォン信号] [M_0 、 M_1] に基づいて異なる音源 [$source_1$ 、 $source_2$] に関連する複数の出力信号 [S'_0 、 S'_1] を取得するための方法が提供され、音源 [$source_1$ 、 $source_2$] からの信号は結合され、

方法は、

第 1 の出力信号 [S'_0] を得るために、第 1 の入力信号 [M_0]、またはその処理された [例えば、遅延および/またはスケーリングされる] バージョンを、第 2 の入力信号 [M_1] の遅延およびスケーリングされたバージョン [$a_1 * z^{-d_1} * M_1$] と結合するステップであって、方法は、部分遅延 [d_1] を第 2 の入力信号 [M_1] に適用するように構成される、ステップ [部分遅延 (d_1) は、第 2 のソース ($source_1$) から第 1 のマイクロフォン (mic_0) に到達する信号 ($H_{1,0} * S_1$) の遅延 (例えば、 $H_{1,0}$ で表される遅延) と、第 2 のソース ($source_1$) から第 2 のマイクロフォン (mic_1) に到達する信号 ($H_{1,1} * S_1$) の遅延 (例えば、 $H_{1,1}$ で表される遅延) との間の関係および/または差を示すことができる] [例では、部分遅延 d_1 は、分数 $H_{1,0}(z) / H_{1,1}(z)$ の結果の z 項の指数を近似するものとして理解することができる]、

【 0 0 5 4 】

第2の入力信号 $[M_1]$ 、またはその処理 [例えば、遅延および/またはスケーリングされる] バージョンを、第1の入力信号 $[M_0]$ の遅延およびスケーリングされたバージョン $[a_0 * z^{-d_0} * M_0]$ と結合して、第2の出力信号 $[S'_1]$ を取得するステップであって、方法は、部分遅延 $[d_0]$ を第1の入力信号 $[M_0]$ に適用するように構成される、ステップ [部分遅延 (d_0) は、第1の音源 ($source_0$) から第1のマイクロフォン (mic_0) に到達する信号 ($H_{0,0} * S_0$) の遅延 (例えば、 $H_{0,0}$ で表される遅延) と、第1の音源 ($source_0$) から第2のマイクロフォン (mic_1) に到達する信号 ($H_{0,1} * S_0$) の遅延 (例えば、 $H_{0,1}$ で表される遅延) との間の関係および/または差を示すことができる] [例では、部分遅延 d_0 は、分数 $H_{0,1}(z) / H_{0,0}(z)$ の結果の z 項の指数を近似するものとして理解することができる]、

10

【 0 0 5 5 】

最適化を使用して、

第1の入力信号 $[M_0]$ の遅延およびスケーリングされたバージョン $[a_0 * z^{-d_0} * M_0]$ を取得するために使用される第1のスケーリング値 $[a_0]$ 、

第1の入力信号 $[M_0]$ の遅延およびスケーリングされたバージョン $[a_0 * z^{-d_0} * M_0]$ を取得するために使用される第1の部分遅延値 $[d_0]$ 、

第2の入力信号 $[M_1]$ の遅延およびスケーリングされたバージョン $[a_1 * z^{-d_1} * M_1]$ を取得するために使用される第2のスケーリング値 $[a_1]$ 、および

第2の入力信号 $[M_1]$ の遅延およびスケーリングされたバージョン $[a_1 * z^{-d_1} * M_1]$ を取得するために使用される第2の部分遅延値 $[d_1]$ を決定するステップ、を含む。

20

【 0 0 5 6 】

一例によれば、方法が提供され、方法は、複数の入力信号 [例えば、マイクロフォン信号] $[M_0, M_1]$ に基づいて異なる音源 $[source_0, source_1]$ に関連付けられた複数の出力信号 $[S'_0, S'_1]$ を取得するためのものであり、ここでは音源からの信号が結合され [例えば、第1の入力信号から第2の入力信号の遅延およびスケーリングされたバージョンを減算することによって、および/または第2の入力信号から第1の入力信号の遅延およびスケーリングされたバージョンを減算することによって]、方法は、

30

第1の出力信号 $[S'_0]$ を得るために、第1の入力信号 $[M_0]$ 、またはその処理 [例えば、遅延および/またはスケーリングされる] バージョンを、第2の入力信号 $[M_1]$ [例えば、第1の入力信号から第2の入力信号の遅延されスケーリングされたバージョンを減算することによって、] の遅延およびスケーリングされたバージョン $[a_1 * z^{-d_1} * M_1]$ と結合するためのものであり、

第2の入力信号 $[M_1]$ 、またはその処理 [例えば、遅延および/またはスケーリングされる] バージョンを、第1の入力信号 $[M_0]$ [例えば、第2の入力信号から第1の入力信号の遅延されスケーリングされたバージョンを減算することによって] の遅延およびスケーリングされたバージョン $[a_0 * z^{-d_0} * M_0]$ と組み合わせて、第2の出力信号 $[S'_1]$ を取得するためのものであり、

40

正規化された大きさの値の第1のセット [例えば、式 (7) のように] のそれぞれの要素 $[P_i(n)]$ 、ここで、 i は0または1である] と、

正規化された大きさの値の第1のセットのそれぞれの要素 $[P(n)]$ または $P_1(n)$] および正規化された大きさの値の第2のセットのそれぞれの要素 $[Q(n)]$ または $Q_1(n)$] に基づいて形成される商の対数と、

の間の複数の積 [例えば、式 (6) または (8) のように] を合計するためのものであり、

その結果、第1の正規化された大きさの値のセット $[P_0(n), n = 1 \sim \dots]$ の場合] によって記述される信号部分 $[s_0'(n)]$ と第2の正規化された大きさの値のセット $[P_1(n), n = 1 \sim \dots]$ の場合] によって記述される信号部分 $[s_1'(n)]$

50

]との間の類似性[または非類似性]を記述する値[式(6)または式(8)中のDKL(P||Q)またはD(P0、P1)]を取得する。

【0057】

一例によれば、上記または下記のような機器を使用するように構成された、前述の方法請求項のいずれか一項に記載の方法が提供される。

プロセッサによって実行されると、プロセッサに、前述の方法請求項のいずれか一項に記載の方法を実行させる命令を格納する非一時的記憶ユニット。

【図面の簡単な説明】

【0058】

【図1】本発明を理解するのに有用なマイクロフォンおよび音源のレイアウトを示す図である。 10

【図2a】本発明による機能技術を示す図である。

【図2b】痙攣混合および混合プロセスの信号ブロック図を示す図である。

【図3】模擬データに適用したBSSアルゴリズムの性能評価を示す図である。

【図4】本発明を理解するのに有用なマイクロフォンおよび音源のレイアウトを示す図である。

【図5】本発明による装置を示す図である。

【図6a】本発明を用いて得られる結果を示す図である。

【図6b】本発明を用いて得られる結果を示す図である。

【図6c】本発明を用いて得られる結果を示す図である。 20

【図7】図5の装置の要素を示す図である。

【発明を実施するための形態】

【0059】

上述および後述するような技術を適用することによって、信号は、互いに分離された複数の信号

s_1^i

および

s_0^i

に到達するように処理され得ることが理解されている。したがって、出力信号

s_1^i

は、音 s_0 から影響を受けない(または無視できるほどもしくは無視できるほど影響を受けない)が、出力信号

s_0^i

は、マイクロフォン m_{ic0} への音 s_1 の影響による影響を受けない(または無視できるほどもしくは無視できるほど影響を受けない)という結果になる。

【0060】 40

一例が図2bによって提供され、マイクロフォン M_0 および M_1 から集合的に得られた生成音 s_0 および s_1 と信号502との間の物理的関係のモデルを示す。結果は、ここでは z 変換で表される(場合によっては簡潔にするために示されていない)。ブロック501から分かるように、音響信号 s_0 は、音響信号 s_1 に加算される伝達関数 $H_{0,0}(z)$ を受ける(伝達関数 $H_{1,0}(z)$ によって修正される)。したがって、信号 $M_0(z)$ はマイクロフォン m_{ic0} で得られ、望ましくないことに、音響信号 $s_1(z)$ を考慮に入れる。同様に、マイクロフォン m_{ic1} で得られる信号 $M_1(z)$ は、音響信号 $s_1(z)$ に関連する成分(伝達関数 $H_{1,1}(z)$ によって得られる)と、音響信号 $s_0(z)$ に起因する第2の不要成分(伝達関数 $H_{0,1}(z)$ を受けた後で)との両方を含む。この現象をクロストークと呼ぶ。 50

【 0 0 6 1 】

クロストークを補償するために、ブロック 5 1 0 に示された解決策が利用され得る。ここで、マルチチャネル出力信号 5 0 4 は、

・入力信号 M_0 、および減算成分 5 0 3 1 (信号 M_1 の遅延および / またはスケーリングされたバージョンであり、信号 M_1 を伝達関数

$$-a_1 \cdot z^{-d_1}$$

にかけることによって取得され得る) の少なくとも 2 つの成分を含む、第 1 の出力信号 $S'_0(z)$

(マイクロフォン $m i c_0$ で収集されるがクロストークから研磨された音 S_0 を表す) と

・入力信号 M_1 、および減算成分 5 0 3 0 (マイクロフォン $m i c_0$ で得られた第 1 の入力信号 M_0 の遅延および / またはスケーリングされたバージョンであり、信号 M_0 を伝達関数にかけることによって得られ得る)

$$-a_0 \cdot z^{-d_0}$$

を含む、出力信号

$$S'_1(z)$$

(マイクロフォン $m i c_1$ で集音されたがクロストークから研磨された音 S_1 を表す) と、を含む。

【 0 0 6 2 】

数学的説明は以下に提供されるが、ブロック 5 1 0 における減算成分 5 0 3 1 および 5 0 3 0 は、ブロック 5 0 1 において生じる望ましくない成分を補償することが理解され得る。したがって、ブロック 5 1 0 が、複数 (5 0 2) の入力信号 [例えば、マイクロフォン信号] (M_0 、 M_1) に基づいて、異なる音源 ($s o u r c e_0$ 、 $s o u r c e_1$) に関連する複数 (5 0 4) の出力信号 (S'_0 、 S'_1) を取得することを可能にし、音源 ($s o u r c e_0$ 、 $s o u r c e_1$) からの信号 (S_0 、 S_1) が (望ましくなく) 結合される (5 0 1) ことは明らかである。ブロック 5 1 0 は、第 1 の入力信号 (M_0) またはその処理 [例えば、遅延および / またはスケーリングされる] バージョンを、第 2 の入力信号 (M_1) [例えば、第 1 の入力信号から第 2 の入力信号の遅延されスケーリングされたバージョンを減算することによって、例えば、 $S'_0(z) = M_0(z) - a_1 \cdot z^{-d_1} \cdot M_1(z)$] の遅延およびスケーリングされたバージョン (5 0 3 1) [$a_1 \cdot z^{-d_1} \cdot M_1$] と組み合わせて (5 1 0)、第 1 の出力信号 (S'_0) を取得するように構成することができ、ブロックが、第 2 の入力信号 (M_1) またはその処理 [例えば、遅延および / またはスケーリングされる] バージョンを、前記第 1 の入力信号 [M_0] [例えば、第 2 の入力信号から第 1 の入力信号の遅延されスケーリングされたバージョンを減算することによって、例えば、 $S'_1(z) = M_1(z) - a_0 \cdot z^{-d_0} \cdot M_0(z)$] の遅延およびスケーリングされたバージョン (5 0 3 0) [$a_0 \cdot z^{-d_0} \cdot M_0$] と組み合

わせて (5 1 0)、第 2 の出力信号 [S'_1] を取得するように構成される。
 z 変換はこの場合に特に有用であるが、他の種類の変換を利用することも、時間領域で直接動作することも可能である。

【 0 0 6 3 】

基本的に、一対のスケーリング値 a_0 および a_1 は、入力信号のスケーリングされたバージョンを取得するために減算成分 5 0 3 1 および 5 0 3 0 の振幅を修正することが理解され得、遅延 d_0 および d_1 は部分遅延として理解され得る。例では、部分遅延 d_0 は、分数 $H_{0,1}(z) / H_{0,0}(z)$ の結果の z 項の指数を近似するものとして理解することができる。部分遅延 d_1 は、第 2 のソース ($s o u r c e_1$) から第 1 のマイクロフォン ($m i c_0$) に到達する信号 ($H_{1,0} \cdot S_1$) の遅延 (例えば、 $H_{1,0}$ で表され

10

20

30

40

50

る遅延)と、第2のソース(s_{01})から第2のマイクロフォン(m_{11})に到達する信号($H_{1,1} \cdot S_{11}$)の遅延(例えば、 $H_{1,1}$ で表される遅延)との間の関係および/または差を示すことができる。例では、部分遅延 d_{11} は、分数 $H_{1,0}(z) / H_{1,1}(z)$ の結果の z 項の指数を近似するものとして理解することができる。部分遅延 d_{01} は、第1のソース(s_{00})から第1のマイクロフォン(m_{10})に到達する信号($H_{0,0} \cdot S_{00}$)の遅延(例えば、 $H_{0,0}$ で表される遅延)と、第1のソース(s_{00}) [例では、部分遅延 d_{01} は、分数 $H_{0,1}(z) / H_{0,0}(z)$ の結果の z 項の指数を近似するものとして理解することができる]から第2のマイクロフォン(m_{11})に到達する信号($H_{0,1} \cdot S_{00}$)の遅延(例えば、 $H_{0,1}$ で表される遅延)との間の関係および/または差を示すことができる。

10

【0064】

後に説明するように、特に、(参照符号564でまとめて示される)最も好ましい値、

- ・例えば、第1の入力信号[502、 M_0]の遅延されスケールされたバージョン5030 [$a_0 \cdot z^{-d_0} \cdot M_0$]を取得するために使用される第1のスケール値 [a_0]、

- ・例えば、第1の入力信号[502、 M_0]の遅延されスケールされたバージョン5030 [$a_0 \cdot z^{-d_0} \cdot M_0$]を取得するために使用される第1の部分遅延値 [d_0]、

- ・例えば、第2の入力信号[502、 M_1]の遅延されスケールされたバージョン5031 [$a_1 \cdot z^{-d_1} \cdot M_1$]を取得するために使用される第2のスケール値 [a_1]と、

20

- ・例えば、第2の入力信号[502、 M_1]の遅延されスケールされたバージョン5031 [$a_1 \cdot z^{-d_1} \cdot M_1$]を取得するために使用される第2の部分遅延値 [d_1]と、を見出すことが可能である。

【0065】

最も好ましいスケール値 a_0 および a_1 ならびに遅延値 d_0 および d_1 を得るための技術が、ここで、特に図5を参照して論じられる。図5から分かるように、ステレオまたはマルチチャネルの信号502(入力信号 $M_0(z)$ および $M_1(z)$ を含む)が得られる。図から分かるように、この方法は、採用されるべきスケール値および遅延値の最良値を得るために複数の反復に沿ってサイクルすることが可能であるという意味で、反復的であってもよい。

30

【0066】

図5は、信号

 $s_0'(z)$

および

 $s_1'(z)$

によって形成され、例えば複数の反復後に最適化される出力504を示す。図5は混合ブロック510を示し、これは図2bのブロック510であってもよい。

40

【0067】

したがって、マルチチャネル信号510(そのチャネル成分、すなわち複数の入力信号 $s_0'(z)$

および

 $s_1'(z)$

を含む)は、反復に沿ってますます最適化されるスケール値 a_0 および a_1 ならびに遅延値 d_0 および d_1 を利用することによって得られる。

【0068】

50

ブロック 5 2 0 において、信号

$s'_i(z)$

および

$s'_i(z)$

に対して正規化が実行される。正規化の例は、以下の商として表される式 (7) によって提供される。

$$P_i(n) = \frac{|s'_i(n)|}{\|s'_i\|_1} \quad (7) \quad 10$$

【 0 0 6 9 】

ここでは、

$i = 0, 1$

は、入力信号 M_0 に対する正規化値 $P_0(n)$ と、入力信号 M_1 に対する正規化値 $P_1(n)$ とがあることを示している。インデックス n は、時間領域入力信号の時間インデックスである。ここで、

$s'_i(n)$

20

は信号 M_i ($i = 0, 1$ として、) の時間領域サンプルインデックス (z 変換ではない) である。

$|s'_i(n)|$

は、

$s'_i(n)$

の大きさ (例えば、絶対値) が得られ、したがって正であるか、または悪い場合には 0 であることを示す。これは、式 (7) の分子が正であること、またはさらに悪いことに 0 であることを意味する。

30

$\|s'_i\|_1$

は、式 (7) の分母がベクトル

s'_i

の 1 - ノルムによって形成されることを示す。1 - ノルム

$\| \dots \|_1$

40

は、大きさ

$|s'_i(n)|$

の合計を示し、

\sum

は、信号サンプル、例えば現在のインデックス (例えば、信号サンプルは、過去のインデックスから現在のインデックスまでの所定のウィンドウ内で取得されてもよい) までである。したがって、

$\|s'_i\|_1$

50

(式(7)の分母)は正(場合によっては0)である。また、常に $|s'_i(n)| \leq \|s'_i\|_1$

であり、 $0 \leq P_i(n) \leq 1$ ($i = 0, 1$)であることを意味する。さらに、以下も検証される。

$$\sum_n \frac{s'_i(n)}{\|s'_i\|_1} = 1$$

【0070】

したがって、式(7)を採用することによって、 $P_0(n)$ および $P_1(n)$ が以下を検証するので、 $P_0(n)$ および $P_1(n)$ を確率として人為的に考えることができることに留意されたい。

1 .
 $P_i(n) \geq 0, \forall n$

2 .
 $\sum_{n=0}^{\infty} P_i(n) = 1$

ここで、
 $i = 0, 1$

(さらなる考察を以下に提供する)である。「 \cdot 」は数学的形式に使用されるが、考慮される信号にわたって近似することができる。

式(7)を介して得られたものだけでなく、他の種類の正規化が提供されてもよいことに留意されたい。

【0071】

図5は、正規化値522によって入力され、第1および第2の入力値 M_0 および M_1 の間の情報を与える類似度値(または類似度値)532を出力するブロック530を示す。ブロック530は、入力信号 M_0 および M_1 が互いにどの程度類似している(または類似していない)かを示すメトリックを測定するブロックとして理解することができる。

【0072】

第1の入力値と第2の入力値との間の類似性または非類似性を示すために選択されるメトリックは、いわゆるカルバックライブラー発散(KLD)であってもよいことが理解されている。これは、式(6)または(8)を使用して得ることができる。

$$D_{KL}(P||Q) = \sum_n P(n) \log \left(\frac{P(n)}{Q(n)} \right) \tag{6}$$

$$D(P_0, P_1) = -\sum_n \left[P_0(n) \cdot \log \left(\frac{P_0(n)}{P_1(n)} \right) + P_1(n) \cdot \log \left(\frac{P_1(n)}{P_0(n)} \right) \right]$$

(8)

【0073】

ここで、カルバックライブラー発散(KLD)をどのようにして得るかについての議論を提供する。図7は、図5のブロック520の下流のブロック530の例を示す。したがって、ブロック520は、例えば上述の式(7)を使用して、

$$P_0(n)$$

および

$$P_1(n)$$

(522)を提供する(他の技術が使用されてもよい)。ブロック530(カルバックライブラープロセッサまたはKLプロセッサとして理解されてもよい)は、メトリック532を取得するように適合されてもよく、これは、この場合、式(8)で計算されるカルバックライブラー発散である。

【0074】

図7を参照すると、第1の分岐700aにおいて、ブロック702で

$$P_0(n)$$

と

$$P_1(n)$$

10

との商702'が計算される。ブロック706において、商702'の対数が計算され、したがって、値706'を得る。次に、対数值706'は、スケーリングブロック710で正規化された値

$$P_0$$

をスケーリングするために使用されてもよく、したがって、積710'を得る。第2の分岐700bにおいて、商704'がブロック704において計算される。商704'の対数708'は、ブロック704において計算される。次に、対数值708'は、スケーリングブロック712で正規化された値をスケーリングするために使用され、したがって、積712'を得る。

20

【0075】

加算器ブロック714において、値710'および712'(それぞれ分岐700aおよび700bで得られる)が互いに組み合わせられる。ブロック716において、組み合わせられた値714'は、互いに、サンプル領域インデックスに沿って合計される。加算値716'は、ブロック718(例えば、-1でスケーリングされる)において反転されて反転値718'を得ることができる。値716'は類似値として理解することができるが、反転値718'は非類似値として理解することができることに留意されたい。値716'または値718'のいずれかは、上述したように(類似度を示す値716'、非類似度を示す値718')メトリック532としてオブティマイザ560に提供されてもよい。

30

【0076】

したがって、オブティマイザブロック530は、したがって、式(8)、すなわち、

$$D(P_0, P_1) = -\sum_n \left[P_0(n) \log \left(\frac{P_0(n)}{P_1(n)} \right) + P_1(n) \log \left(\frac{P_1(n)}{P_0(n)} \right) \right]$$

に到達することを可能にすることができる。式(6)、例えば、

$$D_{KL}$$

に到達するために、図7からブロック704、708、712および714を排除し、

$$P_0$$

40

を

$$P$$

と、

$$P_1$$

をQと置換することが単に可能であり得る。

【0077】

カルバックライブラー発散は、確率に関する測定値を与えるために自然に考えられたものであり、原則として、入力信号M0およびM1の物理的有意性とは無関係である。それ

50

にもかかわらず、本発明者らは、信号

s'_0

および

s'_1

を正規化し、

$P_0(i)$

および

$P_1(i)$

10

などの正規化された値を取得することによって、カルバックライブラー発散が、入力信号 M_0 と M_1 との間の類似性 / 非類似性を測定するための有効なメトリックを提供することを理解している。したがって、時間領域サンプルの正規化された大きさの値を確率分布と見なすことができ、その後、メトリック（例えば、カルバックライブラー発散として、例えば、式（6）または（8））を測定することができる。

【0078】

ここで再び図5を参照する。反復ごとに、メトリック532は、スケーリング値 a_0 および a_1 ならびに遅延値 d_0 および d_1 の有効性の良好な推定値を提供する。反復に沿って、スケーリング値 a_0 および a_1 ならびに遅延値 d_0 および d_1 の異なる候補値が、最も低い類似度または最も高い非類似度を提示するそれらの候補の中から選択される。

20

【0079】

ブロック560（オプティマイザ）は、メトリック532によって入力され、遅延値 d_0 、 d_1 およびスケーリング値 a_0 、 a_1 の候補564（ベクトル）を出力する。オプティマイザ560は、候補 a_0 、 a_1 、 d_0 、 d_1 の異なるグループに対して取得された異なるメトリックを測定し、それらを変更し、最も低い類似度（または最も高い非類似度）532に関連付けられた候補のグループを選択することができる。したがって、出力504（出力信号

$S'_0(z), S'_1(z)$

30

）は最良の近似を提供する。候補値564は、ベクトルにグループ化されてもよく、これは、例えばランダムな技法（図5は、ランダム入力542をオプティマイザ560に提供するランダム発生器540を示す）によってその後修正することができる。オプティマイザ560は、候補値564（ a_0 、 a_1 、 d_0 、 d_1 ）がスケーリング（例えば、ランダムに）される重みを利用することができる。初期係数重み562は、例えばデフォルトで提供されてもよい。オプティマイザ560の処理の一例が提供され、以下で広く説明される（「アルゴリズム1」）。アルゴリズムのラインと図5の要素との間の可能な対応関係も図5に示されている。

【0080】

理解され得るように、オプティマイザ564は、値 a_0 、 a_1 、 d_0 、 d_1 のベクトル564を出力し、これらは続いて、新しい値512、新しい正規化された値522、および新しいメトリック532を取得するために混合ブロック510で再使用される。特定の反復回数（例えば予め定義することができる）の後、最大反復回数は、例えば、10から20の間で選択された回数であってもよい。基本的に、オプティマイザ560は、例えば、ブロック530でおよび / または式（6）および（8）を使用して取得されたメトリック532であり得る目的関数を最小化する遅延および反復値を見つけるものとして理解され得る。

40

【0081】

したがって、オプティマイザ560は、候補パラメータが候補のベクトル [例えば、5

50

64、 a_0 、 a_1 、 d_0 、 d_1 に対応する4つのエントリを有する]を形成するように、ランダム方向最適化技術に基づくことができ、候補のベクトルは、候補のベクトルをランダム方向に修正することによって反復的に洗練されることが理解されている。

【0082】

基本的に、候補のベクトル(a_0 、 a_1 、 d_0 、 d_1 の後続の値を示す)は、ランダムな方向の候補ベクトルを修正することによって反復的に洗練され得る。例えば、ランダム入力542に続いて、ランダムに変化する異なる重みを使用して異なる候補値を修正することができる。ランダム方向は、例えば、所定の規則なしで、いくつかの候補値が増加している間、他の候補値が減少していること、またはその逆を意味し得る。また、最大閾値が事前定義されていても、重みの増分はランダムであってもよい。

10

【0083】

オブティマイザ460は、候補パラメータ[a_0 、 a_1 、 d_0 、 d_1]が候補のベクトルを形成するようにすることができ、反復ごとに、候補のベクトルの要素に要素ごとに乗算される[例えば、それぞれ-0.5から+0.5の間]一様分布乱数のベクトルを適用することによって、候補のベクトルが摂動される[例えば、ランダムに]。したがって、例えばランダム方向最適化を使用することによって、勾配処理を回避することが可能である。したがって、係数のベクトルをランダムに摂動させることによって、 a_0 、 a_1 、 d_0 、 d_1 の好ましい値、および合成音 S_0 および S_1 が適切に補償された出力信号504に段階的に到達することが可能である。アルゴリズムについては、以下で詳細に説明する。

20

【0084】

本例では、常に、2つの入力チャンネル(例えば、 M_0 、 M_1)によって形成されるマルチチャンネル入力信号502が参照される。とにかく、上記の同じ例は、3つ以上のチャンネルにも適用される。

例では、対数は任意の底であってもよい。上述のベースは10であると推測することができる。

【0085】

技術の詳細な説明

1つの目標は、利用可能なステレオウェブカメラのように、ステレオマイクロフォンからあまり遠くない、小さなオフィス環境において、2つのスピーカ、またはスピーカおよび楽器またはノイズ源を分離するためのテレビ会議のためのシステムである。スピーカまたは音源は、ステレオマイクロフォンの対向する(左右の)側にあると仮定される。リアルタイムのテレビ会議で有用であるために、できるだけ低い遅延でオンラインで動作することが望まれる。比較のために、この論文では、オフライン実装に焦点を当てている。提案された手法は、マイクロフォン信号間の減衰係数および部分遅延を使用してクロストークを最小化し、部分遅延および和ビームフォーマの原理を使用して、時間領域で機能する。他の手法と比較して、これは、最適化すべき変数の数がより少ないという利点を有し、周波数領域におけるICAのような手法の順列問題を有していない。分離を最適化するために、本発明者らは、得られた分離されたシグナル間の負のカルバックライブラー導出目的関数を最小化する。最適化のために、非常に高速で堅牢な勾配を必要とせずに、「ランダムな方向」の新規なアルゴリズムを使用する。本発明者らは、ルームインパルス応答シミュレータを使用して、および現実の記録を用いて、TIMITデータセットから得られた音声信号から生成された畳み込み混合に対する本発明者らの手法を評価する。結果は、提案されたシナリオについて、本発明者らの手法がその分離性能に関して競争力があり、従来技術の手法よりも計算複雑度およびシステム遅延が低いことを示している。

30

40

【0086】

インデックス用語 - ブラインド音源分離、時間領域、バイノーラル室内インパルス応答、最適化

1. 導入、以前の手法

本発明者らのシステムは、2つのマイクロフォンを有し、2つの音源を分離したい用途

50

向けである。これは、例えば、オフィスにステレオウェブカメラがあり、その周りに2人のスピーカがある、または計算の複雑さが低いことが重要である補聴器のためのテレビ会議シナリオであり得る。

【0087】

以前のアプローチ：以前の初期のアプローチは、独立成分分析（ICA）である。それは、混合物中で遅延なしに信号の混合を解くことができる。非ガウス性を最大化するか、またはカルバックライブラー発散 [1、2] を最大化することによって、非混合行列の係数を見つける。しかし、オーディオ信号およびステレオマイクロフォンペアの場合、ミックスにおいて、伝播遅延、一般的にはルームインパルス応答 [3] との畳み込みが常に存在する。これに対処するための手法は、信号 [4]、例えば $AuxIVA$ [5] および $ILRMA$ [6、7] に短時間フーリエ変換（STFT）を適用することが多い。これにより、信号遅延がSTFTサブバンド内の複素数値係数に変換され、得られたサブバンド（例えば、[8]）に（複素数値）ICAを適用することができる。

10

【0088】

問題：ここで発生する問題は、サブバンドにおける順序であり、分離されたソースは、異なるサブバンドにおいて異なる順序で現れる可能性がある。異なるサブバンド内の異なるソースの利得は異なり、修正されたスペクトル形状、スペクトル平坦化をもたらす可能性がある。また、STFTを適用することから生じる信号遅延もある。ブロックサイズ [9、10] に対応するシステム遅延を必要とするブロックへの信号の組み立てが必要である。

20

【0089】

TRINICON [11] のような時間領域アプローチ、または短いブロックおよびより多くのマイクロフォン [12、13] を有するSTFTを使用するアプローチは、STFTの大きなブロック遅延を有さないという利点を有するが、通常、より高い計算複雑度を有し、これにより、小型デバイスで使用する事が困難になる。

図1も参照されたい。図1は、シミュレーションにおけるスピーカおよびマイクロフォンの設定を示している。

2. 提案されたアプローチ

【0090】

周波数領域手法に関連する処理遅延を回避するために、時間領域手法を使用する。FIRフィルタを使用する代わりに、本発明者らは、減衰係数、部分遅延および和の原理または適応ビームフォーマ [16、17、18] を有する部分遅延オールパスフィルタ [14、15] として実装されるIIRフィルタを使用する。これは、そのような各フィルタが2つの係数、部分遅延および減衰のみを有するという利点を有する。2チャンネルステレオの場合、これは合計でわずか4つの係数をもたらす、最適化がより容易になる。簡単にするために、残響除去も行わず、クロストーク最小化に注目する。実際には、2つのマイクロフォン間の相対伝達関数を減衰および純粋な部分遅延によってモデル化する。次に、「差分進化」法と同様に、「ランダム方向」の新規最適化を適用する。

30

【0091】

2つのマイクロフォン（ s_0

40

および
 s_1

）で作られた2つの音源（ M_0

および
 M_1

50

）からの混合録音を想定する。しかしながら、3つ以上の音源についても同じ結果が有効である。音源は、図1に示すように固定された位置にあると仮定することができる。非因果的インパルス応答のモデル化の必要性を回避するために、音源はマイクロフォン対の異なる半平面（左右）になければならない。

【0092】

一般的に使用されるSTFTの代わりに、関連する遅延を伴う信号のブロックへの分解を必要としないため、数学的導出にz変換を使用することができる。これにより、アルゴリズム遅延のない時間領域実装に適している。サンプルインデックス

n

10

を用いた時間領域信号

$x(n)$

の（片側）z変換は、

$$X(z) = \sum_{n=0}^{\infty} x(n)z^{-n}$$

のように定義されることに留意されたい。z変換ドメイン信号を示すために大文字を使用する。

【0093】

瞬間の2つの時間領域音響信号（サンプルインデックス）

n

20

として

$s_0(n)$

および

$s_1(n)$

を定義し、それらのz変換を

$S_0(z)$

30

および

$S_1(z)$

として定義する。2つのマイクロフォン信号（まとめて502で示す）は、

$m_0(n)$

および

$m_1(n)$

40

であり、それらのz変換は、

$M_0(z)$

および

$M_1(z)$

である（図2）。

iのソースからjのマイクロフォンへのルームインパルス応答（RIR）は

$h_{i,j}(n)$

50

であり、それらの z 変換は

$$H_{i,j}(z)$$

である。したがって、本発明者らの畳み込み混合システムは、 z 領域において、以下のよう
に記述することができる。

$$\begin{bmatrix} M_0(z) \\ M_1(z) \end{bmatrix} = \begin{bmatrix} H_{0,0}(z) & H_{1,0}(z) \\ H_{0,1}(z) & H_{1,1}(z) \end{bmatrix} \cdot \begin{bmatrix} S_0(z) \\ S_1(z) \end{bmatrix}$$

(1)

【 0 0 9 4 】

10

単純化された行列乗算では、式 (1) を次のように書き換えることができる。

$$M(z) = H(z) \cdot S(z)$$

(2)

【 0 0 9 5 】

理想的な音源分離のためには、混合行列

$$H(z)$$

を反転する必要がある。したがって、音源は次のように計算できる。

$$S(z) = H^{-1}(z)M(z) \Rightarrow$$

20

$$\begin{bmatrix} S_0(z) \\ S_1(z) \end{bmatrix} = \begin{bmatrix} H_{1,1}(z) & -H_{1,0}(z) \\ -H_{0,1}(z) & H_{0,0}(z) \end{bmatrix} \cdot \frac{1}{\det(H(z))} \cdot \begin{bmatrix} M_0(z) \\ M_1(z) \end{bmatrix}$$

(3)

【 0 0 9 6 】

逆行列の

$$\det(H(z))$$

および対角要素は、混合解除に寄与しない線形フィルタであるため、分離のためにそれら
を無視し、それらを式 (3) の左側に持ってくることもできる。 30

この結果、

$$\begin{bmatrix} H_{1,1}^{-1}(z) & 0 \\ 0 & H_{0,0}^{-1}(z) \end{bmatrix} \cdot \begin{bmatrix} S_0(z) \\ S_1(z) \end{bmatrix} \cdot \det(H(z)) =$$

【 0 0 9 7 】

$$\begin{bmatrix} 1 & -H_{1,1}^{-1}(z)H_{1,0}(z) \\ -H_{0,0}^{-1}(z)H_{0,1}(z) & 1 \end{bmatrix} \cdot \begin{bmatrix} M_0(z) \\ M_1(z) \end{bmatrix}$$

(4)

40

であり、

$$H_{1,1}^{-1}(z) \cdot H_{1,0}(z)$$

および

$$H_{0,0}^{-1}(z) \cdot H_{0,1}(z)$$

は、相対ルーム伝達関数である。

【 0 0 9 8 】

次に、

$$d_i$$

50

サンプルによる部分遅延および減衰係数

a_i

によってこれらの相対ルーム伝達関数を近似する。

$$H_{i,i}^{-1}(z) \cdot H_{i,j}(z) \approx a_i z^{-d_i}$$

(5)

ここで、

$i, j \in \{0, 1\}$

である。

【 0 0 9 9 】

この近似は、残響またはエコーが強すぎない場合に特にうまく機能する。

d_i

サンプルによる部分遅延については、次のセクション (2 . 1) の部分遅延フィルタを使用する。簡単にするために、行列式および行列の対角線

$H_{i,i}(z)$

から得られる線形フィルタを左側に保持し、これは残響除去がないことを意味することに留意されたい。

【 0 1 0 0 】

ここでは、図 1 を参照して例が提供されている。ステレオマイクロフォンペアの周りの対向する両側に対称的に、反射のない自由音場の 2 つの音源を想定する。source 0 から近用マイクロフォンまでの (サンプル単位の) 距離は、

m_0

= 5 0 サンプル、遠用マイクロフォンまで

m_1

= 5 5 サンプルとする。音の振幅は、ある定数

k

の関数

$k/(m_i^2)$

に従って減衰するものとする。次いで、z 領域内のルーム伝達関数は、

$$H_{0,0}(z) = H_{1,1}(z) = k/50^2 \cdot z^{-50}$$

および

$$H_{0,1}(z) = H_{1,0}(z) = \frac{k}{55^2} \cdot z^{-55}$$

であり、一方の相対的なルーム伝達関数は

$$H_{0,0}^{-1}(z) \cdot H_{0,1}(z) = \frac{50^2}{55^2} \cdot z^{-5}$$

であり、他方の相対的なルーム伝達関数について同じである。この単純なケースでは、相対的なルーム伝達関数は実際に

$$0.825 \cdot z^{-5}$$

であり、正確には減衰および遅延であることが分かる。畳み込み混合およびデミックス処

10

20

30

40

50

理の信号フローチャートを図 2 b に示す。(図 2 b : 畳み込み混合および逆混合プロセスの信号ブロック図である) 。

【 0 1 0 1 】

2 . 1 . 部分遅延オールパスフィルタ式 (5) の遅延

$$z^{d_i}$$

を実現するための部分遅延オールパスフィルタは、ただ 1 つの係数から I I R フィルタを生成し、良好なクロストークキャンセルに必要な正確な部分遅延 (正確な遅延は整数値ではない) の実装を可能にするので、本発明者らの方式において重要な役割を果たす。I I R 特性はまた、効率的な実施を可能にする。[1 4 、 1 5] は、最大平坦群遅延を有する I I R フィルタに基づいて、部分遅延オールパスフィルタを設計するための実際的な方法を記載している [1 9] 。以下の式を使用して、部分遅延

10

$$\tau = d_i$$

のための部分遅延オールパスフィルタの係数を取得する。z ドメインでのその伝達関数は $A(z)$

のようになる。[1 4]

$$A(z) = \frac{z^{-L} D(z)}{D(z)}$$

20

ここで、

$$D(z)$$

は次のように定義される順序

$$L = \lceil \tau \rceil$$

である。

$$D(z) = 1 + \sum_{n=1}^L d(n) z^{-n}$$

30

フィルタ

$$d(n)$$

は、

$$0 \leq n \leq (L - 1)$$

について以下のように生成される。

$$d(0) = 1,$$

$$d(n+1) = d(n) \cdot \frac{(L-n)(L-n-\tau)}{(n+1)(n+1+\tau)}$$

40

【 0 1 0 2 】

2 . 2 . 目的関数

目的関数として、カルバックライブラー 発散 (K L D)

$$D_{KL}(P||Q) = \sum_n P(n) \cdot \log \left(\frac{P(n)}{Q(n)} \right)$$

(6)

から導出された関数

$$D(P_0, P_1)$$

50

を使用する。

【 0 1 0 3 】

ここで、

$$P(n)$$

および

$$Q(n)$$

は、本発明者らの（混合されていない）マイクロフォンチャンネルの確率分布であり、離散分布上で

n

10

が実行される。

【 0 1 0 4 】

計算をより高速にするために、ヒストグラムの計算を回避する。ヒストグラムの代わりに、時間領域信号自体の正規化された大きさを使用し、

$$P_i(n) = \frac{p_i(n)}{\sum_i p_i(n)}$$

(7)

ここで、

n

20

は時間領域サンプルインデックスである。

$$P_i(n)$$

は、確率の特性と同様の特性を有する、すなわち、

1 .

$$P_i(n) \geq 0, \forall n$$

.

2 .

$$\sum_{n=0}^{\infty} P_i(n) = 1$$

.

式中、 $i = 0, 1$ であることに留意されたい。カルバックライブラー発散を直接使用する代わりに、和

$$D_{KL}(P||Q) + D_{KL}(Q||P)$$

を使用することによって目的関数を対称（距離）関数に変える。これにより、2つのチャンネル間の分離がより安定されるからである。最大化の代わりに最小化を適用するために、その負の値をとる。したがって、本発明者らの結果として得られる目的関数

$$D(P_0, P_1)$$

30

40

は、

$$D(P_0, P_1) = -\sum_n \left[P_0(n) \cdot \log \left(\frac{P_0(n)}{P_1(n)} \right) + P_1(n) \cdot \log \left(\frac{P_1(n)}{P_0(n)} \right) \right]$$

(8)

である。

【 0 1 0 5 】

2 . 3 . 最適化

50

BSSのための広範な最適化方法は勾配降下法である。これは、最適値への「最も急な」方法を見つけるという利点を有するが、勾配の計算を必要とし、目的関数の極小に容易に陥るか、または「狭い谷」によって減速される。したがって、係数の最適化のために、「差分進化」[20、21、22]と同様に、「ランダム方向」の新規な最適化を使用する。更新のための係数ベクトルの差の代わりに、重みベクトルを使用して、係数の予想される分散分布をモデル化する。これは、非常に単純であるが非常に高速な最適化アルゴリズムをもたらし、これはまた、リアルタイム通信アプリケーションにとって重要なリアルタイム処理に容易に適用することができる。アルゴリズムは、固定された開始点[1.0、1.0、1.0、1.0]から始まり、我々は、これがロバストな収束挙動をもたらすことを見出した。次に、重みベクトルと要素ごとに乗算された、-0.5から+0.5(ランダム方向)の間の一様分布乱数のベクトルで現在点を摂動させる(アルゴリズム1の線10)。この摂動された点がより低い目的関数値を有する場合、それを次の現在点として選択し、以下同様である。最適化アルゴリズムの擬似コードは、アルゴリズム1に見ることができる。ここで、`minabskl__i`(アルゴリズム1で`negabskl__i`として示されている)は、係数ベクトル`coeffs`およびアレイX内のマイクロフォン信号からKLDを計算する目的関数である。

10

【0106】

本発明者らは、本発明者らのテストファイルに対して(ファイル全体の毎回)20回の反復(したがって、20回の目的関数評価のみ)で既に十分であることを見出し、これを非常に高速なアルゴリズムにする。

20

【0107】

最適化は、ブロック560において実行されてもよく、または例えば実行されてもよい(上記参照)。

【0108】

アルゴリズム1を以下に示す。

アルゴリズム1 最適化アルゴリズム

```

1: 手順 分離係数を最適化
2:   初期化
3:   X                               畳み込み混合
4:   init_coeffs = [1.0, 1.0, 1.0, 1.0]           30
                                           分離係数の初期推定
5:   coeffweights                               ランダム検索の重み
6:   coeff = init_coeffs                           分離係数
7:   negabskl__0(coeffs, X)                       KLDの演算
8:   最適化ルーチン
9:   ループ:
10:    coeffvariation = (random(4) * coeffweights)
                                           分離係数のランダム変数
11:    negabskl__1(coeffs + coeffvariation, X)     40
                                           新たなKLDの演算
12:    if negabskl__1 < negabskl__0 then
13:      negabskl__0 = negabskl__1
14:      coeffs = coeffs + coeffvariation
                                           分離係数を更新

```

【0109】

3. 実験結果

このセクションでは、本発明者らは、シミュレーションされたルームインパルス応答およびTIMITデータセット[23]からの異なる音声信号を使用することによって、A

50

I R E S (t i m e d o m A i n f R a c t i o n a l d E l a y S e p a r a t i o n) と呼ばれる提案された時間領域分離方法を評価する。さらに、本発明者らは、現実の部屋環境で現実の記録を行った。本発明者らの提案する手法の性能を評価するために、最先端の B S S アルゴリズム、すなわち時間領域 T R I N I C O N [1 1]、周波数領域 A u x I V A [5] および I L R M A [6、7] との比較を行った。T R I N I C O N B S S の実装は、その著者から受領されている。A u x I V A および I L R M A B S S の実装は、それぞれ [2 4] および [2 5] から取得された。実験は、C P U コア i 7 第 8 世代および 1 6 G b の R A M を有するラップトップ上で M A T L A B R 2 0 1 7 a を使用して行われた。

【 0 1 1 0 】

3 . 1 . 合成 R I R による分離性能

画像モデル技術 [2 6、2 7] に基づくルームインパルス応答シミュレータを使用して、ルームインパルス応答を生成した。シミュレーション設定のために、部屋のサイズが
7m × 5m × 3m

であるように選択されている。マイクロフォンは、[3 . 4 7 5、2 . 0、1 . 5] m および [3 . 5 2 5、2 . 0、1 . 5] m で部屋の中央に配置され、サンプリング周波数は 1 6 k H z であった。全 T I M I T データセットから 1 0 対の音声信号をランダムに選択し、模擬 R I R と置み込んだ。各信号対について、マイクロフォンに対する音源のランダムな角度位置について、4 つの異なる距離および 3 つの残響時間 (R T 6 0) について、シミュレーションを 1 6 回繰り返した。すべてのシミュレーションで使用された共通のパラメータを表 1 に示し、設定の視覚化を図 1 に示す。分離性能の評価は、元の音源が利用可能であるため、信号対歪み比 (S D R) 測定値 [2 8] および計算時間を計算することによって客観的に行われた。結果を図 3 に示す。

【 0 1 1 1 】

得られた結果は、0 . 2 秒未満の残響時間に対する本発明者らの手法の良好な性能を示す。R T 6 0 = 0 . 0 5 s では、すべての距離にわたる平均 S D R 測定値は 1 5 . 6 4 d B であり、R T 6 0 = 0 . 1 s では 1 0 . 2 4 d B である。残響時間 R T 6 0 = 0 . 2 s について、本発明者らの提案する B S S アルゴリズムは、T R I N I C O N 後に I L R M A と 2 位を共有する。すべてのシミュレーションにわたる (本発明者らのコンピュータ上の) 平均計算時間を表 2 に示す。図から分かるように、A I R E S は、計算時間に関してすべての最先端のアルゴリズムよりも優れている。

【 0 1 1 2 】

結果を聞くことにより、本発明者らは、約 8 d B の S D R が良好な発話了解度をもたらし、本発明者らの手法は実際に、不自然なサウンディングアーチファクトを特徴としないことを見出した。

【 0 1 1 3 】

【 表 1 】

表 1 : シミュレーションで使用したパラメータ :

部屋の寸法	7 m × 5 m × 3 m
マイクロフォンの変位	0 . 0 5 m
残響時間 R T 6 0	0 . 0 5、0 . 1、0 . 2 s
混合量	1 0 ランダム混合
各混合の条件	1 6 ランダム角度 × 4 距離

【 0 1 1 4 】

図 3 の特徴付け (シミュレートされたデータに適用された B S S アルゴリズムの性能評価を示す) :

(a) $RT_{60} = 0.05 \text{ s}$

(b) $RT_{60} = 0.1 \text{ s}$

(c) $RT_{60} = 0.2 \text{ s}$

【 0 1 1 5 】

【 表 2 】

表 2 : 平均計算時間 (シミュレートされたデータセット) の比較

BSS	AIRES	TRINICON	ILRMA	AuxIVA
演算時間	0.05 s	23.4 s	4.4 s	0.6 s

10

【 0 1 1 6 】

3.2. 現実の実験

最後に、提案された音源分離方法を評価するために、現実の実験を行った。実際の記録は、3つの異なる部屋タイプでキャプチャされている。すなわち、小さなアパートの部屋 (3 m × 3 m)、オフィスの部屋 (7 m × 5 m)、大きな会議室 (15 m × 4 m) である。各部屋のタイプについて、2つのスピーカの10個のステレオ混合が記録されている。「グランドトゥルス」信号がないため、分離性能を評価するために、分離されたチャンネル間の相互情報測度 [29] が計算されている。結果を表3に示す。混合マイクロフォンチャンネルの平均相互情報は1.37であり、分離された信号間の相互情報が低いほど良好な分離であることに留意されたい。

20

【 0 1 1 7 】

【 表 3 】

表 3 : 平均相互情報量 (MI、実際の記録) を用いた分離性能の比較

BSS	AIRES	TRINICON	ILRMA	AuxIVA
平均MI	0.5	0.47	0.52	0.52
平均演算時間	0.22 s	120.2 s	7.86 s	1.5 s

30

【 0 1 1 8 】

比較表3から、実際に記録された畳み込み混合物の分離の性能傾向は、シミュレートされたデータの場合と同じままであったことが分かる。したがって、AIRESはその単純さにもかかわらず、従来技術のブラインドソース分離アルゴリズムと競合する可能性がある」と結論付けることができる。

【 0 1 1 9 】

4. 結論

この論文では、2つのオーディオチャンネル間のクロストークを最小限に抑えるために、IIR部分遅延フィルタの推定に基づく高速時間領域ブラインド音源分離技術を提示した。部分遅延および減衰係数の推定は、ステレオ畳み込み混合からのソース信号の迅速かつ効果的な分離をもたらすことを示した。このために、負のカルバックライブラー発散から導出された目的関数を導入した。最小化を堅牢かつ高速にするために、本発明者らは、「差分進化」の最適化に類似する新規な「ランダム方向」最適化方法を提示した。提案されたBSS技術を評価するために、一連の実験を行った。本発明者らは、本発明者らのシステムを、シミュレートされたデータおよび実際の部屋の記録に関する他の最先端の方法と評価および比較した。結果は、本発明者らのシステムが、その単純さにもかかわらず、その分離性能において競争力があるが、はるかに低い計算複雑度を有し、システム遅延がないことを示している。これはまた、リアルタイム最小遅延アプリケーションおよび移動する音源 (移動するスピーカなど) に対するオンライン適応を可能にする。これらの特性により、AIRESは、補聴器または小型テレビ会議装置などの小型デバイス上のリアルタ

40

50

イムアプリケーションによく適している。A I R E S B S S のテストプログラムは、当方の G i t H u b [3 0] で入手可能である。

【 0 1 2 0 】

さらなる態様（上記および／または下記の例も参照）

例えば、マルチチャンネルまたはステレオオーディオソース分離方法およびその更新方法に関して、さらなる態様がここで説明される。それは（相互情報のような）目的関数を最小化し、例えば、他のチャンネルから信号を取得することによってクロストーク低減を使用し、それに減衰係数および（可能な分数）遅延を適用し、それを現在のチャンネルから減算する。これは、例えば、遅延および減衰係数を更新するために「ランダムな方向」の方法を使用する。

【 0 1 2 1 】

以下の態様も参照されたい。

1：分離のための係数がランダム更新ベクトルを追加することによって反復的に更新され、この更新が改善された分離をもたらす場合、次の反復のためにこの更新を維持する、マルチチャンネルまたはステレオソース分離方法；そうでなければ廃棄する。

2：各係数のランダム更新に適した固定分散を有する態様 1 に記載の方法。

3：更新がライブオーディオストリーム上で行われ、更新が過去のオーディオサンプルに基づく、態様 1 または 2 に記載の方法。

4：更新の分散が経時的にゆっくりと減少する、態様 1、2 または 3 に記載の方法。

【 0 1 2 2 】

他の態様は、

1) 多次元入力ベクトル x を用いて目的関数を最小化するパラメータのオンライン最適化の方法であって、まず所与の近傍空間からテストベクトルを決定し、次いで探索ベクトルを用いて以前の最良ベクトルを更新し、目的関数のそのような取得された値が新しい最良値になる場合、この更新されたベクトルは新しい最良値になり、そうでなければ破棄される、方法。

【 0 1 2 3 】

2) N 個のチャンネルのオーディオストリーム内の N 個のソースを分離するタスクに適用され、係数ベクトルは、不要なソースをキャンセルするために必要な遅延および減衰値からなる、態様 1 に記載の方法。

3) 1 つまたは複数のサイトで 2 人以上の話者がいるテレビ会議のタスクに適用される、態様 2 に記載の方法。

4) オーディオソースをオーディオエンコーダで符号化する前にオーディオソースを分離するタスクに適用される、態様 2 に記載の方法。

5) 音楽または雑音源から音声源を分離するタスクに適用される、態様 2 に記載の方法。

【 0 1 2 4 】

さらなる態様

さらなる態様がここで論じられる。

導入ステレオ分離

目標は、複数のマイクロフォンで音源を分離することである（ここでは 2）。異なるマイクロフォンは、異なる振幅および遅延で音を拾う。以下の考察では、Python におけるプログラミング例を考慮する。これは、理解を容易にするため、アルゴリズムが機能するかどうかおよびどのように機能するかを試験するため、および結果の再現性のため、アルゴリズムを試験可能にし、他の研究者にとって有用にするためである。

【 0 1 2 5 】

その他の態様

以下、他の態様について説明する。それらは必ずしも互いに境界付けられているわけではなく、新しい実施形態を作成するために組み合わせることができる。各箇条書きは、他のものから独立していてもよく、単独で、または他の特徴（例えば、他の箇条書き）、ま

10

20

30

40

50

たは上記もしくは下記で論じられる他の特徴と組み合わせて、上記および/または下記の例の少なくともいくつか、および/または特許請求の範囲に開示される特徴のいくつかを補完またはさらに指定してもよい。

【0126】

空間知覚

耳は、音の特殊な方向を推定するために主に2つの効果を使用する。

両耳間レベル差 (ILD)

両耳間時間差 (ITD)

スタジオ録音からの音楽は、主に「パンニング」と呼ばれるレベル差 (時間差なし) のみを使用する。

例えばステレオウェブカメラからのステレオマイクロフォンからの信号は、主に時間差を示し、あまりレベル差を示さない。

ステレオマイクロフォンを用いたステレオ録音

時間差を有するステレオマイクロフォンを用いた設定。

観測：有限の音速および減衰からの音遅延の影響は、遅延を伴う混合行列によって記述することができる。

図2a：ステレオ遠隔会議設定を参照されたい。マイクロフォン間の信号遅延を観察する。

【0127】

ステレオ分離、ILD

- 時間差がない場合 (パンニングのみ) の方が簡単である
- 通常、独立成分分析を使用する
- それは、統計的に独立した信号を生成する「非混合」行列を計算する。
- 2つの信号XおよびYのジョイントエントロピーが $H(X, Y)$ であることを想起されたい
- 条件付きエントロピーは $H(X | Y)$ である
- 統計的に独立とは、 $H(X | Y) = H(X)$ 、 $H(X, Y) = H(X) + H(Y)$ も意味する

- 相互情報： $I(X, Y) = H(X, Y) - H(X | Y) - H(Y | X) = 0$ 、
独立のために、それは0になる

- 一般に、最適化はこの相互情報を最小化する。

ITD事前手法

- 時間差があるITDの場合は、解決がより困難である
- 「最小分散原理」は、マイクロフォンアレイからのノイズを抑制しようとする
- マイクロフォンチャンネルの相関行列を推定し、原理成分分析 (PCA) またはカルハウネン-レーベ変換 (KLT) を適用する。
- これにより、得られたチャンネルが脱相関する。
- 最も低い固有値を有するチャンネルはノイズであると仮定され、0に設定される

【0128】

ITD前手法例

- 我々のステレオの場合、2つのチャンネルしか得られず、より大きな固有値およびより小さな固有値で信号を計算および再生する：`python KLT_separation.py`

観測：より大きな固有値を有するチャンネルは、より低い周波数、他の (より弱い) より高い周波数を有するようであり、これらはしばしば単なるノイズである。

ITD、新たなアプローチ

理想：非混合行列は、混合行列を反転させるために、非混合のための信号遅延を必要とする

重要な単純化：2つの音源のそれぞれについて、マイクロフォン間の単一の信号遅延経路のみを仮定する。

10

20

30

40

50

実際には、部屋の反射から複数の遅延があるが、通常は直接経路よりもはるかに弱い。

【 0 1 2 9 】

立体分離、ITD、新しいアプローチ

減衰 a を有する混合行列、および d の遅延、サンプルを仮定する。 z 変換領域では、 d サンプルによる遅延は z^{-d} の係数で表される（遅延は小数サンプルであり得ることを観察されたい）。

$$A = \begin{bmatrix} 1 & \alpha_0 z^{-d_0} \\ \alpha_1 z^{-d_1} & 1 \end{bmatrix}$$

したがって、非混合行列はその逆である。

【 0 1 3 0 】

$$A^{-1} = \frac{1}{1 - \alpha_0 \alpha_1 z^{-d_0 - d_1}} \begin{bmatrix} 1 & \alpha_0 z^{-d_0} \\ \alpha_1 z^{-d_1} & 1 \end{bmatrix}$$

実際に性能を犠牲にすることなく行列の前に行列式を用いて画分を落とすことによって、混合解除行列を単純化できることが分かる。

【 0 1 3 1 】

係数計算

係数 a 、および d は、最適化によって得られる必要がある。係数を見つけるために、結果として生じる信号の相互情報の最小化を再び使用することができる。相互情報のためには、結合エントロピーが必要である。Pythonでは、`numpy.histogram2d` を使用して2次元確率密度関数からそれを計算することができ、これを `hist2d`、`xedges`、`yedges = np.histogram2d(x[:, 0], x[:, 1], bins = 100)` と呼ぶ。

【 0 1 3 2 】

目的関数の最適化

我々の目的関数は、いくつかの最小値を有し得ることに留意されたい！これは、その開始点が大域的な最小値に十分に近くない場合、凸最適化を使用できないことを意味する。

したがって、非凸最適化を使用する必要があり、Pythonは非常に強力な非凸最適化方法を有する。

【 0 1 3 3 】

`scipy.optimize.differential_evolution`。発明者らは、`coeffs_minimized = opt.differential_evolution(mutualinfo_coeffs, bounds, args = (X,), tol = 1e-4, disp = True)` と呼ぶ。

【 0 1 3 4 】

代替目的関数

しかし、これは計算が複雑である。したがって、同じ最小値を有する別の目的関数を探す。カルバックライブラーおよびItakura-Saito発散を考慮する

2つの確率分布 P および Q のカルバックライブラー発散は、以下のように定義される。

$$D_{KL}(P, Q) = \sum_{n=i} P(i) \cdot \log \frac{P(i)}{Q(i)}$$

【 0 1 3 5 】

i は、（離散的な）分布上を走る。ヒストグラムの計算を回避するために、本発明者らは、単に時間領域サンプルの正規化された大きさを確率分布として錯覚として扱う。これらは非類似性尺度であるため、最大化する必要があり、したがって負の値を最小化する必要がある。

カルバックライブラー Python 関数

Pythonでは、この目的関数は、

`def minabsklcoeffs(coeffs, X):`

10

20

30

40

50

```

# チャンネルの正規化された大きさを計算し、次いで、
# the Kullback - Leibler divergence
X_prime = unmixing ( coeffs, X )
X_abs = np . abs ( X_prime )
# sum ( ) = 1 に正規化して、それを確率のように見せる。
X_abs [ : , 0 ] = X_abs [ : , 0 ] / np . sum ( X_abs [ : , 0 ]
X_abs [ : , 1 ] = X_abs [ : , 1 ] / np . sum ( X_abs [ : , 1 ]
)
# print ( " Kullback - Leibler Divergence calculation " )
abs_kl = np . sum ( X_abs [ : , 0 ] * np . log ( ( X_abs [ : , 0 ] + 1 e - 6 ) / ( X_abs [ : , 1 ] + 1 e - 6 ) )
)
戻り - abs_kl
( ここで、 min_abs_kl_coeffs は、アルゴリズム 1 の min_abs_kl_i に
対応する )

```

目的機能の比較

遅延係数のうちの 1 つを除く係数を固定して、プロット内の目的関数を比較する。

【 0 1 3 6 】

図 6 a は、例示的な信号および例示的な係数の目的関数を示す。関数が実際に同じ最小値を有することに留意されたい！「abs_kl」は、信号の絶対値におけるカルバックライブラーであり、最も滑らかである。

【 0 1 3 7 】

最適化例

相互情報を用いた最適化：python

ICAmutualinfo_puredelay.py

(これには約 1 2 1 より遅い反復が必要である)

- カルバックライブラーを使用した最適化：python ICAabskl_puredelay.py

(これには約 3 9 より速い反復が必要である)

得られた信号を聞くと、以前はそうではなかった明瞭さのために十分に分離されていることが確認される。

【 0 1 3 8 】

さらなる高速化最適化

最適化をさらにスピードアップするために、凸最適化を可能にする必要がある。そのためには、(最良の場合には) 1 つの最小値のみを有する目的関数を有する必要がある。より小さい極小値の多くは、信号の高周波成分が急速に変化することに起因する。アプローチ：本発明者らは、本発明者らの信号の低域通過およびダウンサンプリングバージョンに基づいて目的関数を計算する(これはまた、目的関数の計算をさらに高速化する)。ローパスは、極小値を除去するのに十分に狭いが、依然として十分に正確な係数を得るのに十分に広い必要がある。

【 0 1 3 9 】

さらなる高速化最適化、低域通過

ダウンサンプリング係数 8、したがって全帯域の 1 / 8 のローパスを選択する。全帯域の約 1 / 8 (ナイキスト周波数の 1 / 8) の帯域幅を有する以下のローパスフィルタを使用する。

図 6 b : 使用されたローパスフィルタ、振幅周波数応答。x 軸は正規化周波数であり、2 はサンプリング周波数である。

【 0 1 4 0 】

低域通過からの目的関数

ここでも、目的関数をプロットして比較することができる。

図 6 c : ローパスフィルタリングされた例示的な信号および例示的な係数の目的関数を示す図である。ここで、関数は実際には主に 1 つの最小値を有することに留意されたい！「 $a b s k l$ 」は、やはり最も滑らかである（小さな極小値はない）。

さらなる高速化の最適化、実施例

ここで、凸最適化、例えば共役勾配の方法を試すことができる。

- Python ICA abskl__pure delay__low pass . py

- 観測：最適化がほぼ瞬時に成功した！

- 結果として生じる分離は、以前と同じ品質を有する。

【0141】

その他の説明

上記では、異なる本発明の実施例および態様が記載されている。また、さらなる例は、添付の特許請求の範囲によって定義される（例は特許請求の範囲にもある）。特許請求の範囲によって定義される任意の例は、前のページで説明した詳細（特徴および機能）のいずれかによって補足することができることに留意されたい。また、上述した例は、個別に使用することができ、別の章の特徴のいずれか、または特許請求の範囲に含まれる任意の特徴によって補足することもできる。

【0142】

丸括弧および角括弧内のテキストは任意選択であり、さらなる実施形態（さらに特許請求の範囲によって定義されるもの）を定義する。また、本明細書に記載の個々の態様は、個別にまたは組み合わせて使用することができることに留意されたい。したがって、詳細は、前記態様の別の 1 つに詳細を追加することなく、前記個々の態様の各々に追加することができる。

本開示は、移動通信デバイスおよび移動通信システムの受信機の特徴を明示的または暗黙的に記述することにも留意されたい。

【0143】

特定の実装要件に応じて、例はハードウェアで実装されてもよい。実施態様は、それぞれの方法が実行されるようにプログラマブルコンピュータシステムと協働する（または協働することができる）、電子的に読み取り可能な制御信号が記憶されたデジタル記憶媒体、例えばフロッピーディスク、デジタル多用途ディスク（DVD）、ブルーレイディスク、コンパクトディスク（CD）、読み出し専用メモリ（ROM）、プログラマブル読み出し専用メモリ（PROM）、消去可能およびプログラマブル読み出し専用メモリ（EEPROM）またはフラッシュメモリを使用して実行されてもよい。したがって、デジタル記憶媒体はコンピュータ可読であってもよい。

【0144】

一般に、例は、プログラム命令を有するコンピュータプログラム製品として実装されてもよく、プログラム命令は、コンピュータプログラム製品がコンピュータ上で実行されるときに方法のうちの 1 つを実行するように動作する。プログラム命令は、例えば、機械可読媒体に記憶されてもよい。

【0145】

他の例は、機械可読キャリアに格納された、本明細書に記載の方法の 1 つを実行するためのコンピュータプログラムを含む。言い換えれば、したがって、方法の一例は、コンピュータプログラムがコンピュータ上で実行されるときに、本明細書に記載の方法のうちの 1 つを実行するためのプログラム命令を有するコンピュータプログラムである。

【0146】

したがって、方法のさらなる例は、本明細書に記載の方法の 1 つを実行するためのコンピュータプログラムを記録して含むデータキャリア媒体（またはデジタル記憶媒体、またはコンピュータ可読媒体）である。データキャリア媒体、デジタル記憶媒体、または記録媒体は、無形で一時的な信号ではなく、有形および/または非一時的である。

10

20

30

40

50

さらなる例は、処理ユニット、例えばコンピュータ、または本明細書に記載の方法のうちの1つを実行するプログラブル論理デバイスを含む。

さらなる例は、本明細書に記載の方法のうちの1つを実行するためのコンピュータプログラムをインストールしたコンピュータを含む。

【0147】

さらなる例は、本明細書に記載された方法のうちの1つを実行するためのコンピュータプログラムを受信機へ転送する（例えば、電子的または光学的に）装置またはシステムを備える。受信機は、例えば、コンピュータ、モバイルデバイス、メモリデバイスなどであってもよい。装置またはシステムは、例えば、コンピュータプログラムを受信機に転送するためのファイルサーバを備えることができる。

10

【0148】

いくつかの例では、プログラブル・ロジック・デバイス（例えば、フィールド・プログラマブル・ゲート・アレイ）を使用して、本明細書に記載の方法の機能の一部またはすべてを実行することができる。いくつかの例では、フィールド・プログラマブル・ゲート・アレイは、本明細書に記載の方法のうちの1つを実行するためにマイクロプロセッサと協働することができる。一般に、方法は、任意の適切なハードウェア装置によって実行されてもよい。

【0149】

上述の例は、上述の原理の例示である。本明細書に記載の構成および詳細の修正および変形は明らかであることが理解される。したがって、本明細書の実施例の説明および説明として提示される特定の詳細によってではなく、添付の特許請求の範囲によって限定されることが意図される。

20

【0150】

参考文献

[1] A. Hyvarinen, J. Karhunen, and E. Oja, "Independent component analysis." John Wiley & Sons, 2001.

【0151】

[2] G. Evangelista, S. Marchand, M. D. Plumbley, and E. Vincent, "Sound source separation," in DAFX: Digital Audio Effects, second edition ed. John Wiley and Sons, 2011.

30

【0152】

[3] J. Tariquallah, W. Wang, and D. Wang, "A multistage approach to blind separation of convolutive speech mixtures," in Speech Communication, 2011, vol. 53, pp. 524 - 539.

【0153】

[4] J. Benesty, J. Chen, and E. A. Habets, "Speech enhancement in the STFT domain," in Springer, 2012.

40

【0154】

[5] J. Jansky, Z. Koldovsky, and N. Ono, "A computationally cheaper method for blind speech separation based on auxiva and incomplete demixing transform," in IEEE International Workshop on Acoustic Signal Enhancement (IWAENC), Xi'an, China, 2016.

【0155】

[6] D. Kitamura, N. Ono, H. Sawada, H. Kameoka

50

, and H. Saruwatari, "Determined blind source separation unifying independent vector analysis and nonnegative matrix factorization," in *IEEE/ACM Trans. ASLP*, vol. 24, no. 9, 2016, pp. 1626 - 1641.

【0156】

[7] "Determined blind source separation with independent low-rank matrix analysis," in Springer, 2018, p. 31.

【0157】

[8] H.-C. Wu and J.C. Principe, "Simultaneous diagonalization in the frequency domain (sdif) for source separation," in *Proc. ICA*, 1999, pp. 245 - 250.

【0158】

[9] H. Sawada, N. Ono, H. Kameoka, and D. Kitamura, "Blind audio source separation on tensor representation," in *ICASSP*, Apr. 2018.

【0159】

[10] J. Harris, S.M. Naqvi, J.A. Chambers, and C. Jutten, "Realtime independent vector analysis with student's t source prior for convolutive speech mixtures," in *IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP*, Apr. 2015.

【0160】

[11] H. Buchner, R. Aichner, and W. Kellermann, "Trinicon: A versatile framework for multichannel blind signal processing," in *IEEE International Conference on Acoustics, Speech, and Signal Processing, Montreal, Quebec, Canada, 2004*.

【0161】

[12] J. Chua, G. Wang, and W.B. Kleijn, "Convolutional blind source separation with low latency," in *Acoustic Signal Enhancement (IWAENC), IEEE International Workshop, 2016*, pp. 1 - 5.

【0162】

[13] W. Kleijn and K. Chua, "Non-iterative impulse response shortening method for system latency reduction," in *Acoustics, Speech and Signal Processing (ICASSP), 2017*, pp. 581 - 585.

【0163】

[14] I. Senesnick, "Low-pass filters realizable as all-pass sums: design via a new flat delay filter," in *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 46, 1999.

10

20

30

40

50

【0164】

[15] T. I. Laakso, V. Valimäki, M. Karjalainen, and U. K. Laine, "Splitting the unit delay," in *IEEE Signal Processing Magazine*, Jan. 1996.

【0165】

[16] M. Brandstein and D. Ward, "Microphone arrays, signal processing techniques and applications," in Springer, 2001.

【0166】

[17] "Beamforming," <http://www.labbookpages.co.uk/audio/beamforming/delaySum.html>, accessed: 2019-04-21.

10

【0167】

[18] Benesty, Sondhi, and Huang, "Handbook of speech processing," in Springer, 2008.

【0168】

[19] J. P. Thiran, "Recursive digital filters with maximally flat group delay," in *IEEE Trans. on Circuit Theory*, vol. 18, no. 6, Nov. 1971, pp. 659-664.

20

【0169】

[20] S. Das and P. N. Suganthan, "Differential evolution: A survey of the state-of-the-art," in *IEEE Trans. on Evolutionary Computation*, Feb. 2011, vol. 15, no. 1, pp. 4-31.

【0170】

[21] R. Storn and K. Price, "Differential evolution - a simple and efficient heuristic for global optimization over continuous spaces," in *Journal of Global Optimization*. 11(4), 1997, pp. 341-359.

30

【0171】

[22] "Differential evolution," <http://www1.icsi.berkeley.edu/~storn/code.html>, accessed: 2019-04-21.

【0172】

[23] J. Garofolo et al., "Timit acoustic-phonetic continuous speech corpus," 1993.

[24] "Microphone array speech processing," <https://github.com/ZitengWang/MASP>, accessed: 2019-07-29.

40

[25] "Ilrma," <https://github.com/d-kitamura/ILRMA>, accessed: 2019-07-29.

【0173】

[26] R. B. Stephens and A. E. Bate, *Acoustics and Vibrational Physics*, London, U.K., 1966.

【0174】

[27] J. B. Allen and D. A. Berkley, *Image me*

50

thod for efficiently simulating small room acoustics. J. Acoust. Soc. Amer., 1979, vol. 65.

【0175】

[28] C. Fevotte, R. Gribonval, and E. Vincent, "Bss eval toolbox user guide," in Tech. Rep. 1706, IRISA Technical Report 1706, Rennes, France, 2005.

【0176】

[29] E. G. Learned-Miller, Entropy and Mutual Information. Department of Computer Science University of Massachusetts, Amherst Amherst, MA 01003, 2013. 10

【0177】

[30] "Comparison of blind source separation techniques," <https://github.com/TUIMenauAMS/>

Comparison of Blind Source Separation techniques, accessed: 2019-07-29. 20

【図面】

【図1】

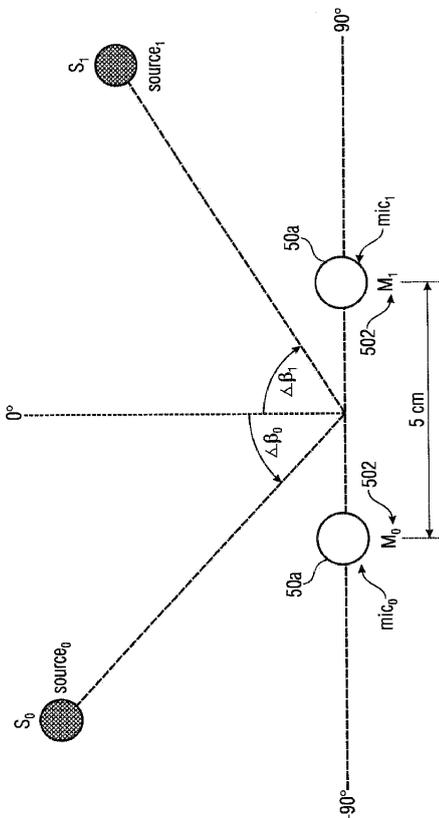


Fig. 1

【図2a】

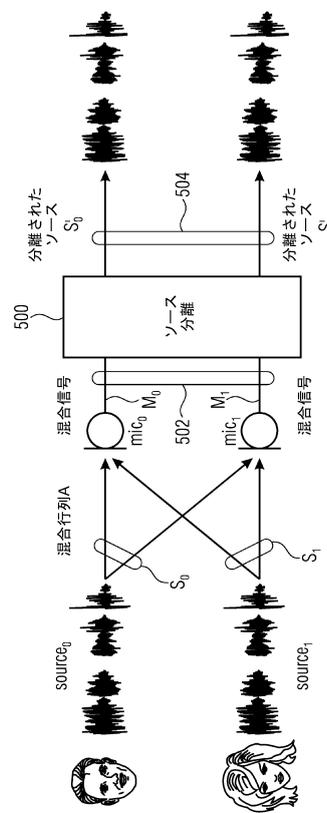


Fig. 2a

20

30

40

50

【 図 2 b 】

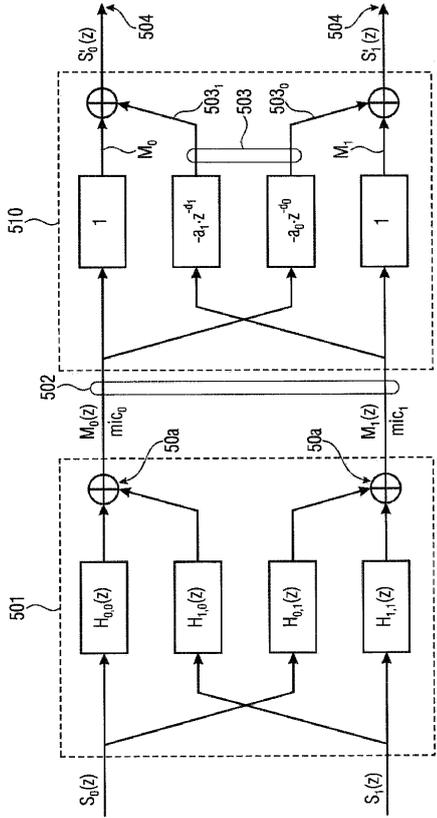
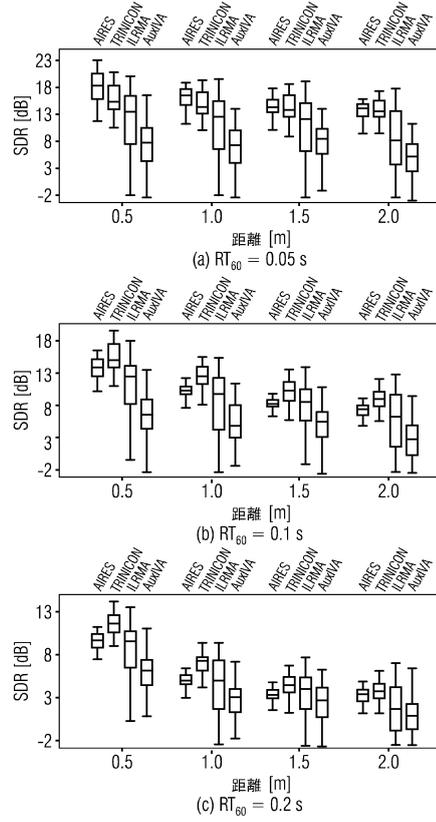


Fig. 2b

【 図 3 】



10

20

【 図 4 】

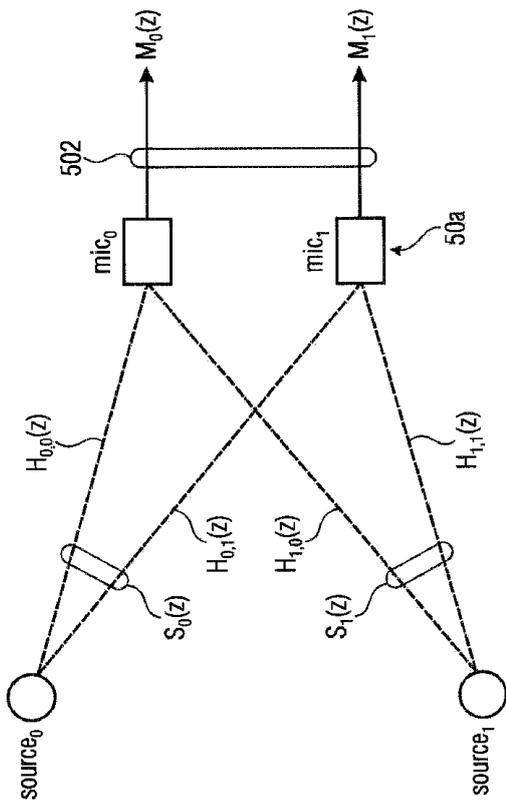


Fig. 4

【 図 5 】

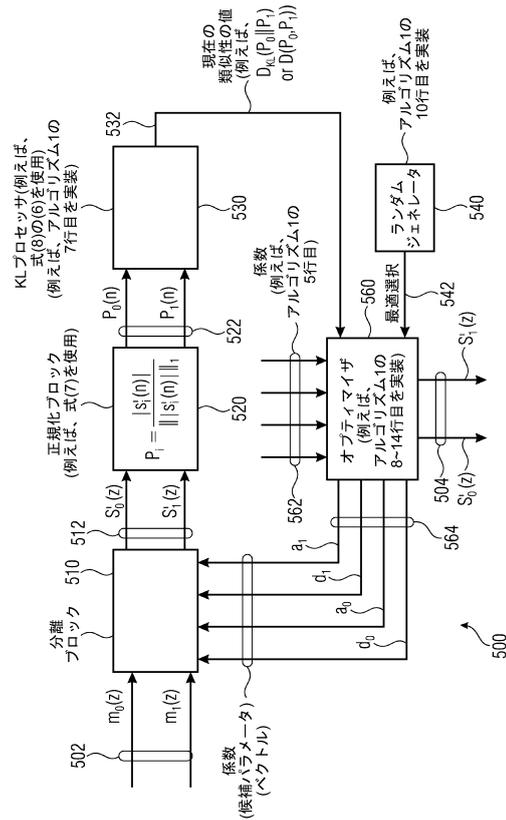


Fig. 5

30

40

50

【 図 6 a 】

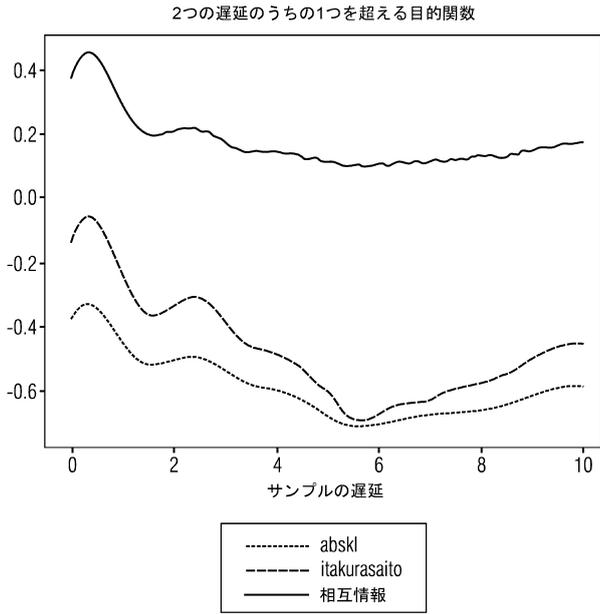


Fig. 6a

【 図 6 b 】

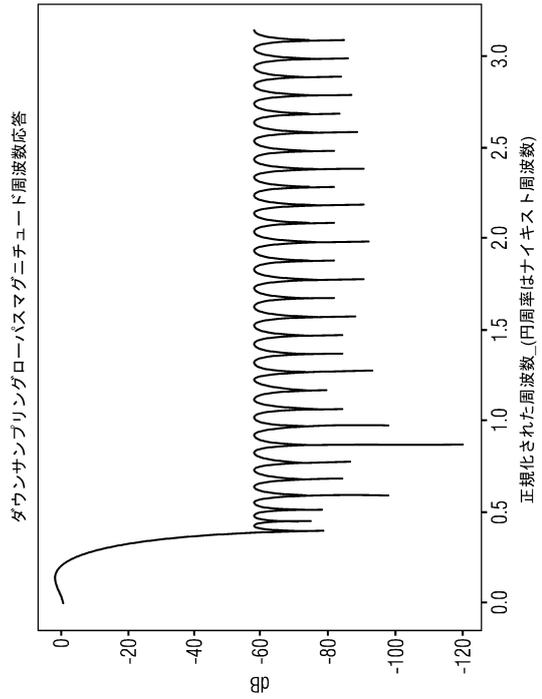


Fig. 6a

10

20

【 図 6 c 】

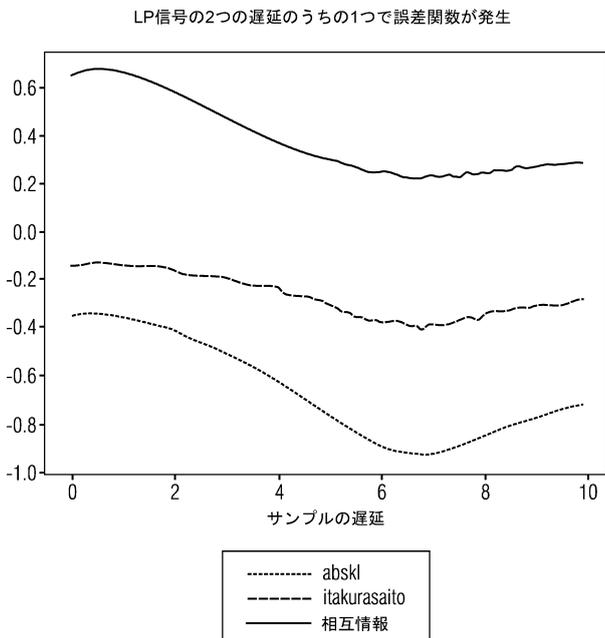


Fig. 6c

【 図 7 】

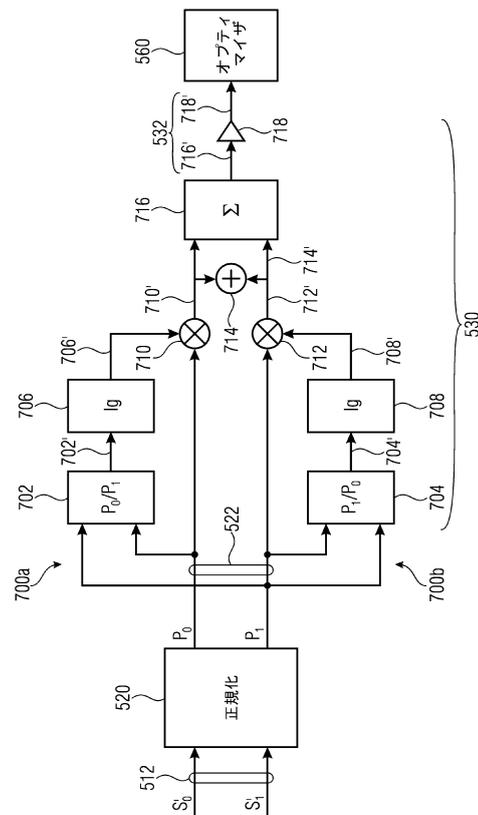


Fig. 7

30

40

50

【手続補正書】

【提出日】令和3年7月12日(2021.7.12)

【手続補正1】

【補正対象書類名】特許請求の範囲

【補正対象項目名】全文

【補正方法】変更

【補正の内容】

【特許請求の範囲】

【請求項1】

複数の入力信号(502、 M_0 、 M_1)に基づいて、異なる音源($source_0$ 、 $source_1$)に関連付けられた複数の出力信号(504、 S'_0 、 S'_1)を取得するための装置(500)であって、前記音源($source_0$ 、 $source_1$)からの信号(S_0 、 S_1)が結合され(501)、

前記装置は、第1の入力信号(502、 M_0)またはその処理されたバージョンを第2の入力信号(M_1)の遅延およびスケールされたバージョン(503₁)と組み合わせて(510)、第1の出力信号(504、 S'_0)を取得するように構成され、

前記装置は、第2の入力信号(502、 M_1)、またはその処理されたバージョンを前記第1の入力信号(M_0)の遅延およびスケールされたバージョン(503₀)と組み合わせて(510)、第2の出力信号(504、 S'_1)を取得するように構成され、

前記装置は、ランダム方向最適化(560)を使用して、

前記第1の入力信号(502、 M_0)の前記遅延されスケールされたバージョン(503₀)を取得するために使用される第1のスケール値(564、 a_0)と、

前記第1の入力信号(502、 M_0)の前記遅延されスケールされたバージョン(503₀)を取得するために使用される第1の遅延値(564、 d_0)と、

前記第2の入力信号(502、 M_1)の前記遅延されスケールされたバージョン(503₁)を取得するために使用される第2のスケール値(564、 a_1)と、

前記第2の入力信号の前記遅延およびスケールされたバージョン(503₁)を取得するために使用される第2の遅延値(564、 d_1)と、

を決定するように構成され、

前記ランダム方向最適化(560)は、候補パラメータが候補のベクトル(564、 a_0 、 a_1 、 d_0 、 d_1)を形成するようなものであり、前記候補のベクトルは、前記候補のベクトルをランダムな方向に修正することによって反復的に洗練される、

装置(500)。

【請求項2】

前記第1の入力信号(502、 M_0)と結合されるべき前記第2の入力信号(502、 M_1)の前記遅延されスケールされたバージョン(503₁)は、前記第2の入力信号(502、 M_1)に部分遅延を適用することによって得られる、請求項1に記載の装置。

【請求項3】

前記第2の入力信号(502、 M_1)と結合されるべき前記第1の入力信号(502、 M_0)の前記遅延されスケールされたバージョン(503₀)は、前記第1の入力信号(502、 M_0)に部分遅延を適用することによって得られる、請求項1または2に記載の装置。

【請求項4】

正規化された大きさの値の第1のセットのそれぞれの要素(P_0)と、

正規化された大きさの値の第1のセット(522)のそれぞれの要素(P_0)および

正規化された大きさの値の第2のセット(522)のそれぞれの要素(P_1)

に基づいて形成された商(702')の対数(706')と、

の間の複数の積(712'、710')を合計する(714、716)ように構成されており、

その結果、前記第 1 の正規化された大きさの値のセット (P_0) によって記述される信号部分 (s'_0) と前記第 2 の正規化された大きさの値のセット (P_1) によって記述される信号部分 (s'_1) との間の類似性または非類似性を記述する値 (D_{KL} 、 D 、 532) を取得する (530)、請求項 1 から 3 のいずれか一項に記載の装置。

【請求項 5】

前記ランダム方向最適化 (560) が、前記第 1 の出力信号と前記第 2 の出力信号との間の前記類似性または非類似性を示すメトリックおよび / または値 (D_{KL} 、 D 、 532) が測定され、前記第 1 の出力測定値および前記第 2 の出力信号 (S'_0 、 S'_1) が、前記値に関連付けられた前記候補パラメータに関連付けられた測定値、または最も低い類似性または最も高い非類似性を示すメトリックであるように選択される、請求項 1 から 4 のいずれか一項に記載の装置。

10

【請求項 6】

前記第 1 のスケール値および前記第 2 のスケール値ならびに前記第 1 の遅延値および前記第 2 の遅延値は、目的関数を最小化することによって取得される、請求項 1 から 5 のいずれか一項に記載の装置。

【請求項 7】

前記第 1 の入力信号 (502 、 M_0)、またはその処理されたバージョンを、前記時間領域および / または前記 z 変換もしくは周波数領域において、前記第 2 の入力信号 (502 、 M_1) の前記遅延されスケールされたバージョン (503_1) と結合するようにさらに構成され、

20

前記第 2 の入力信号 (502 、 M_1)、またはその処理されたバージョンを、前記時間領域および / または前記 z 変換もしくは周波数領域において前記第 1 の入力信号 (502 、 M_0) の前記遅延およびスケールされたバージョン (503_0) と組み合わせるようにさらに構成されている、

請求項 1 から 6 のいずれか一項に記載の装置。

【請求項 8】

前記最適化は、前記 z 変換領域で実行される、
請求項 1 から 7 のいずれか一項に記載の装置。

【請求項 9】

前記最適化は、前記時間領域で実行される、
請求項 1 から 7 のいずれか一項に記載の装置。

30

【請求項 10】

前記最適化は、前記周波数領域で実行される、
請求項 1 から 7 のいずれか一項に記載の装置。

【請求項 11】

前記第 2 の入力信号 (502 、 M_1) に適用される前記遅延または部分遅延 (d_0) は

、
前記第 1 のマイクロフォン (mic_0) によって受信された前記第 1 の音源 ($source_0$) からの前記信号、および

前記第 2 のマイクロフォン (mic_1) によって受信された前記第 1 の音源 ($source_0$) からの前記信号、

40

の関係、および / または差もしくは到着を示す、
請求項 1 から 10 のいずれか一項に記載の装置。

【請求項 12】

前記第 1 の入力信号 (502 、 M_0) に適用される前記遅延または部分遅延 (d_1) は

、
前記第 2 のマイクロフォン (mic_1) によって受信された前記第 2 の音源 ($source_1$) からの前記信号、および

前記第 1 のマイクロフォン (mic_0) によって受信される前記第 2 の音源 ($source_1$) からの前記信号、

50

の間の関係、および/または差もしくは到着を示す、
請求項 1 から 1 1 のいずれか一項に記載の装置。

【請求項 1 3】

前記メトリック (5 3 2) は、カルバックライブラー発散として処理される、請求項 5
に記載の装置。

【請求項 1 4】

前記メトリックは、

$$D_{KL}(P||Q) = \sum_n P(n) \log \left(\frac{P(n)}{Q(n)} \right) \quad (6)$$

10

の形態で取得され、

式中、 $P(n)$ は前記第 1 の入力信号に関連付けられた要素であり、 $Q(n)$ は前記第
2 の入力信号に関連付けられた要素である、請求項 5 に記載の装置。

【請求項 1 5】

前記メトリックは、

$$D(P_0, P_1) = - \sum_n \left[P_0(n) \log \left(\frac{P_0(n)}{P_1(n)} \right) + P_1(n) \log \left(\frac{P_1(n)}{P_0(n)} \right) \right] \quad (8)$$

20

の形態で取得され、

式中、 $P_1(n)$ は、前記第 1 の入力信号に関連付けられた要素であり、 $P_2(n)$ は
、前記第 2 の入力信号に関連付けられた要素である、
請求項 5 に記載の装置。

【請求項 1 6】

前記取得された第 1 および第 2 の出力信号 (S'_0 、 S'_1) に関連する情報を周波数領
域に変換するようにさらに構成される、請求項 1 から 1 5 のいずれか一項に記載の装置。

30

【請求項 1 7】

前記取得された第 1 および第 2 の出力信号 (S'_0 、 S'_1) に関連する情報を符号化す
るようにさらに構成される、請求項 1 から 1 6 のいずれか一項に記載の装置。

【請求項 1 8】

前記取得された第 1 および第 2 の出力信号 (S'_0 、 S'_1) に関連する情報を格納する
ようにさらに構成された、請求項 1 から 1 7 のいずれか一項に記載の装置。

【請求項 1 9】

前記取得された第 1 および第 2 の出力信号 (S'_0 、 S'_1) に関連する情報を送信する
ようにさらに構成される、請求項 1 から 1 8 のいずれか一項に記載の装置。

40

【請求項 2 0】

前記第 1 の入力信号 (5 0 2、 M_0) を取得するための第 1 のマイクロフォン (mic_0)、および前記第 2 の入力信号 (5 0 2、 M_1) を取得するための第 2 のマイクロフォ
ン (mic_1) のうちの少なくとも 1 つと、を含む、請求項 1 から 1 9 のいずれか一項に
記載の装置。

【請求項 2 1】

前記取得された第 1 および第 2 の出力信号 (S'_0 、 S'_1) に関連する情報を送信する
ための機器と、請求項 1 から 2 0 のいずれか一項に記載の装置とを含む、テレビ会議のた
めの装置。

【請求項 2 2】

50

請求項 1 から 2 1 のいずれか一項に記載の装置を含むバイノーラルシステム。

【請求項 2 3】

複数の入力信号 (502、 M_0 、 M_1) に基づいて、異なる音源 (source₀、source₁) に関連付けられた複数の出力信号 (504、 S'_0 、 S'_1) を取得するための方法であって、前記音源 (source₀、source₁) からの信号が結合され、

前記方法は、

第 1 の出力信号 (504、 S'_0) を得るために、第 1 の入力信号 (502、 M_0) またはその処理されたバージョンを第 2 の入力信号 (502、 M_1) の遅延されスケールされたバージョン (503₁) と結合するステップと、

第 2 の入力信号 (502、 M_1) またはその処理されたバージョンを、前記第 1 の入力信号 (502、 M_0) の遅延されスケールされたバージョン (503₀) と結合して、第 2 の出力信号 (504、 S'_1) を取得するステップと、

ランダム方向最適化 (560) を使用して、

前記第 1 の入力信号 (502、 M_0) の前記遅延されスケールされたバージョン (503₀) を取得するために使用される第 1 のスケール値 (564、 a_0) と、

前記第 1 の入力信号 (502、 M_0) の前記遅延されスケールされたバージョン (503₀) を取得するために使用される第 1 の遅延値 (564、 d_0) と、

前記第 2 の入力信号 (502、 M_1) の前記遅延されスケールされたバージョン (503₁) を取得するために使用される第 2 のスケール値 (564、 a_1) と、

前記第 2 の入力信号の前記遅延およびスケールされたバージョン (503₁) を取得するために使用される第 2 の遅延値 (564、 d_1) と、

のうちの少なくとも 1 つを決定するステップと、を含み、

前記ランダム方向最適化 (560) は、候補パラメータが候補のベクトル (a_0 、 a_1 、 d_0 、 d_1) を形成するようなものであり、前記候補のベクトルは、前記候補のベクトルをランダムな方向に修正することによって反復的に洗練される、

方法。

【請求項 2 4】

請求項 1 から 2 1 のいずれか一項に記載の装置を使用するように構成された、請求項 2 3 に記載の方法。

【請求項 2 5】

前記部分遅延 (d_1) は、前記第 2 のソース (source₁) から前記第 1 のマイクロフォン (mic₀) に到達する前記信号 (M_0) の前記遅延と、前記第 2 のソース (source₁) から前記第 2 のマイクロフォン (mic₁) に到達する前記信号 (M_1) の前記遅延 (H_1 、 1) との間の関係および/または差を示す、請求項 2 3 に記載の方法。

【請求項 2 6】

プロセッサによって実行されると、前記プロセッサに請求項 2 3 から 2 5 のいずれか一項に記載の方法を実行させる命令を格納する非一時的記憶ユニット。

10

20

30

40

50

【手続補正書】

【提出日】令和3年10月19日(2021.10.19)

【手続補正1】

【補正対象書類名】特許請求の範囲

【補正対象項目名】全文

【補正方法】変更

【補正の内容】

【特許請求の範囲】

【請求項1】

複数の入力信号(502、 M_0 、 M_1)に基づいて、異なる音源($source_0$ 、 $source_1$)に関連付けられた複数の出力信号(504、 S'_0 、 S'_1)を取得するための装置(500)であって、前記音源($source_0$ 、 $source_1$)からの信号(S_0 、 S_1)が結合され(501)、

前記装置は、第1の入力信号(502、 M_0)またはその処理されたバージョンを第2の入力信号(M_1)の遅延およびスケールされたバージョン(503₁)と組み合わせて(510)、第1の出力信号(504、 S'_0)を取得するように構成され、

前記装置は、第2の入力信号(502、 M_1)、またはその処理されたバージョンを前記第1の入力信号(M_0)の遅延およびスケールされたバージョン(503₀)と組み合わせて(510)、第2の出力信号(504、 S'_1)を取得するように構成され、

前記装置は、ランダム方向最適化(560)を使用して、

前記第1の入力信号(502、 M_0)の前記遅延されスケールされたバージョン(503₀)を取得するために使用される第1のスケール値(564、 a_0)と、

前記第1の入力信号(502、 M_0)の前記遅延されスケールされたバージョン(503₀)を取得するために使用される第1の遅延値(564、 d_0)と、

前記第2の入力信号(502、 M_1)の前記遅延されスケールされたバージョン(503₁)を取得するために使用される第2のスケール値(564、 a_1)と、

前記第2の入力信号の前記遅延およびスケールされたバージョン(503₁)を取得するために使用される第2の遅延値(564、 d_1)と、

を決定するように構成され、

前記ランダム方向最適化(560)は、候補パラメータが候補のベクトル(564、 a_0 、 a_1 、 d_0 、 d_1)を形成するようなものであり、前記候補のベクトルは、前記候補のベクトルをランダムな方向に修正することによって反復的に洗練され、

前記ランダム方向最適化(560)は、前記第1の出力信号と前記第2の出力信号との間の類似性または非類似性を示すメトリック(D_{KL} 、 D 、532)が測定され、前記第1の出力信号および前記第2の出力信号(S'_0 、 S'_1)が、最も低い類似性または最も高い非類似性を示すメトリックに関連付けられた前記候補パラメータに関連付けられた測定値であるように選択されるようなものであり、

前記メトリック(532)は、カルバックライブラー発散として処理される、装置(500)。

【請求項2】

前記第1の入力信号(502、 M_0)と結合されるべき前記第2の入力信号(502、 M_1)の前記遅延されスケールされたバージョン(503₁)は、前記第2の入力信号(502、 M_1)に部分遅延を適用することによって得られる、請求項1に記載の装置。

【請求項3】

前記第2の入力信号(502、 M_1)と結合されるべき前記第1の入力信号(502、 M_0)の前記遅延されスケールされたバージョン(503₀)は、前記第1の入力信号(502、 M_0)に部分遅延を適用することによって得られる、請求項1または2に記載の装置。

【請求項4】

10

20

30

40

50

正規化された大きさの値の第 1 のセットのそれぞれの要素 (P_0) と、
 正規化された大きさの値の第 1 のセット (522) のそれぞれの要素 (P_0) および
 正規化された大きさの値の第 2 のセット (522) のそれぞれの要素 (P_1)
 に基づいて形成された商 ($702'$) の対数 ($706'$) と、
 の間の複数の積 ($712'$ 、 $710'$) を合計する (714 、 716) ように構成されて
 おり、

その結果、前記第 1 の正規化された大きさの値のセット (P_0) によって記述される信
 号部分 (s'_0) と前記第 2 の正規化された大きさの値のセット (P_1) によって記述さ
 れる信号部分 (s'_1) との間の類似性または非類似性を記述する値 (D_{KL} 、 D 、 53
 2) を取得する (530)、請求項 1 から 3 のいずれか一項に記載の装置。 10

【請求項 5】

前記第 1 のスケーリング値および前記第 2 のスケーリング値ならびに前記第 1 の遅延値
 および前記第 2 の遅延値は、目的関数を最小化することによって取得される、請求項 1 か
 ら 4 のいずれか一項に記載の装置。

【請求項 6】

前記第 1 の入力信号 (502 、 M_0)、またはその処理されたバージョンを、前記時間
 領域および / または前記 z 変換もしくは周波数領域において、前記第 2 の入力信号 (50
 2 、 M_1) の前記遅延されスケーリングされたバージョン (503_1) と結合するように
 さらに構成され、

前記第 2 の入力信号 (502 、 M_1)、またはその処理されたバージョンを、前記時間
 領域および / または前記 z 変換もしくは周波数領域において前記第 1 の入力信号 (502
 2 、 M_0) の前記遅延およびスケーリングされたバージョン (503_0) と組み合わせるよ
 うにさらに構成されている、 20

請求項 1 から 5 のいずれか一項に記載の装置。

【請求項 7】

前記最適化は、前記 z 変換領域で実行される、
 請求項 1 から 6 のいずれか一項に記載の装置。

【請求項 8】

前記最適化は、前記時間領域で実行される、
 請求項 1 から 6 のいずれか一項に記載の装置。 30

【請求項 9】

前記最適化は、前記周波数領域で実行される、
 請求項 1 から 6 のいずれか一項に記載の装置。

【請求項 10】

前記第 2 の入力信号 (502 、 M_1) に適用される前記遅延または部分遅延 (d_0) は
 、

前記第 1 のマイクロフォン (mic_0) によって受信された前記第 1 の音源 ($source_0$) から
 の前記信号、および

前記第 2 のマイクロフォン (mic_1) によって受信された前記第 1 の音源 ($source_0$) から
 の前記信号、 40

の間の関係、および / または差もしくは到着を示す、

請求項 1 から 9 のいずれか一項に記載の装置。

【請求項 11】

前記第 1 の入力信号 (502 、 M_0) に適用される前記遅延または部分遅延 (d_1) は
 、

前記第 2 のマイクロフォン (mic_1) によって受信された前記第 2 の音源 ($source_1$) から
 の前記信号、および

前記第 1 のマイクロフォン (mic_0) によって受信される前記第 2 の音源 ($source_1$) から
 の前記信号、

の間の関係、および / または差もしくは到着を示す、 50

請求項 1 から 1 0 のいずれか一項に記載の装置。

【請求項 1 2】

前記メトリックは、

$$D_{KL}(P || Q) = \sum_n P(n) \log \left(\frac{P(n)}{Q(n)} \right)$$

の形態で取得され、

式中、 $P(n)$ は前記第 1 の入力信号に関連付けられた要素であり、 $Q(n)$ は前記第 2 の入力信号に関連付けられた要素である、

請求項 1 から 1 1 のいずれか一項に記載の装置。

10

【請求項 1 3】

前記メトリックは、

$$D(P_0, P_1) = - \sum_n \left[P_0(n) \log \left(\frac{P_0(n)}{P_1(n)} \right) + P_1(n) \log \left(\frac{P_1(n)}{P_0(n)} \right) \right]$$

の形態で取得され、

式中、 $P_1(n)$ は、前記第 1 の入力信号に関連付けられた要素であり、 $P_2(n)$ は、前記第 2 の入力信号に関連付けられた要素である、

請求項 1 から 1 2 のいずれか一項に記載の装置。

【請求項 1 4】

20

前記取得された第 1 および第 2 の出力信号 (S'_0 、 S'_1) に関連する情報を周波数領域に変換するようにさらに構成される、請求項 1 から 1 3 のいずれか一項に記載の装置。

【請求項 1 5】

前記取得された第 1 および第 2 の出力信号 (S'_0 、 S'_1) に関連する情報を符号化するようにさらに構成される、請求項 1 から 1 4 のいずれか一項に記載の装置。

【請求項 1 6】

前記取得された第 1 および第 2 の出力信号 (S'_0 、 S'_1) に関連する情報を格納するようにさらに構成された、請求項 1 から 1 5 のいずれか一項に記載の装置。

【請求項 1 7】

前記取得された第 1 および第 2 の出力信号 (S'_0 、 S'_1) に関連する情報を送信するようにさらに構成される、請求項 1 から 1 6 のいずれか一項に記載の装置。

30

【請求項 1 8】

前記第 1 の入力信号 (S_0 、 M_0) を取得するための第 1 のマイクロフォン (mic_0)、および前記第 2 の入力信号 (S_1 、 M_1) を取得するための第 2 のマイクロフォン (mic_1) のうちの少なくとも 1 つと、を含む、請求項 1 から 1 7 のいずれか一項に記載の装置。

【請求項 1 9】

前記取得された第 1 および第 2 の出力信号 (S'_0 、 S'_1) に関連する情報を送信するための機器と、請求項 1 から 1 8 のいずれか一項に記載の装置とを含む、テレビ会議のための装置。

40

【請求項 2 0】

請求項 1 から 1 9 のいずれか一項に記載の装置を含むバイノーラルシステム。

【請求項 2 1】

複数の入力信号 (S_0 、 M_0 、 M_1) に基づいて、異なる音源 ($source_0$ 、 $source_1$) に関連付けられた複数の出力信号 (S_0 、 S'_0 、 S'_1) を取得するための方法であって、前記音源 ($source_0$ 、 $source_1$) からの信号が結合され、

前記方法は、

第 1 の出力信号 (S_0 、 S'_0) を得るために、第 1 の入力信号 (S_0 、 M_0) またはその処理されたバージョンを第 2 の入力信号 (S_1 、 M_1) の遅延されスケーリン

50

グされたバージョン (503₁) と結合するステップと、

第2の入力信号 (502、M₁) またはその処理されたバージョンを、前記第1の入力信号 (502、M₀) の遅延されスケーリングされたバージョン (503₀) と結合して、第2の出力信号 (504、S'₁) を取得するステップと、

ランダム方向最適化 (560) を使用して、

前記第1の入力信号 (502、M₀) の前記遅延されスケーリングされたバージョン (503₀) を取得するために使用される第1のスケーリング値 (564、a₀) と、

前記第1の入力信号 (502、M₀) の前記遅延されスケーリングされたバージョン (503₀) を取得するために使用される第1の遅延値 (564、d₀) と、

前記第2の入力信号 (502、M₁) の前記遅延されスケーリングされたバージョン (503₁) を取得するために使用される第2のスケーリング値 (564、a₁) と、

前記第2の入力信号の前記遅延およびスケーリングされたバージョン (503₁) を取得するために使用される第2の遅延値 (564、d₁) と、

のうちの少なくとも1つを決定するステップと、を含み、

前記ランダム方向最適化 (560) は、候補パラメータが候補のベクトル (564、a₀、a₁、d₀、d₁) を形成するようなものであり、前記候補のベクトルは、前記候補のベクトルをランダムな方向に修正することによって反復的に洗練され、

前記ランダム方向最適化 (560) は、前記第1の出力信号と前記第2の出力信号との間の類似性または非類似性を示すメトリック (D_{KL}、D、532) が測定され、前記第1の出力信号および前記第2の出力信号 (S'₀、S'₁) が、最も低い類似性または最も高い非類似性を示すメトリックに関連付けられた前記候補パラメータに関連付けられた測定値であるように選択されるようなものであり、

前記メトリック (532) は、カルバックライブラー発散として処理される、方法。

【請求項22】

請求項1から20のいずれか一項に記載の装置を使用するように構成された、請求項21に記載の方法。

【請求項23】

前記部分遅延 (d₁) は、前記第2のソース (source₁) から前記第1のマイクロフォン (mic₀) に到達する前記信号 (M₀) の前記遅延と、前記第2のソース (source₁) から前記第2のマイクロフォン (mic₁) に到達する前記信号 (M₁) の前記遅延 (H₁、₁) との間の関係および/または差を示す、請求項21に記載の方法。

【請求項24】

プロセッサによって実行されると、前記プロセッサに請求項21から23のいずれか一項に記載の方法を実行させる命令を格納する非一時的記憶ユニット。

10

20

30

40

50

【手続補正書】

【提出日】令和4年5月31日(2022.5.31)

【手続補正1】

【補正対象書類名】特許請求の範囲

【補正対象項目名】全文

【補正方法】変更

【補正の内容】

【特許請求の範囲】

【請求項1】

複数の入力信号(502、 M_0 、 M_1)に基づいて、異なる音源($source_0$ 、 $source_1$)に関連付けられた複数の出力信号(504、 S'_0 、 S'_1)を取得するための装置(500)であって、前記音源($source_0$ 、 $source_1$)からの信号(S_0 、 S_1)が結合され(501)、

前記装置は、第1の入力信号(502、 M_0)またはその処理されたバージョンを第2の入力信号(M_1)の遅延およびスケールされたバージョン(503₁)と組み合わせて(510)、第1の出力信号(504、 S'_0)を取得するように構成され、

前記装置は、第2の入力信号(502、 M_1)、またはその処理されたバージョンを前記第1の入力信号(M_0)の遅延およびスケールされたバージョン(503₀)と組み合わせて(510)、第2の出力信号(504、 S'_1)を取得するように構成され、

前記装置は、ランダム方向最適化(560)を使用して、

前記第1の入力信号(502、 M_0)の前記遅延されスケールされたバージョン(503₀)を取得するために使用される第1のスケール値(564、 a_0)と、

前記第1の入力信号(502、 M_0)の前記遅延されスケールされたバージョン(503₀)を取得するために使用される第1の遅延値(564、 d_0)と、

前記第2の入力信号(502、 M_1)の前記遅延されスケールされたバージョン(503₁)を取得するために使用される第2のスケール値(564、 a_1)と、

前記第2の入力信号の前記遅延およびスケールされたバージョン(503₁)を取得するために使用される第2の遅延値(564、 d_1)と、

を決定するように構成され、

前記ランダム方向最適化(560)は、候補パラメータが候補のベクトル(564、 a_0 、 a_1 、 d_0 、 d_1)を形成するようなものであり、前記候補のベクトルは、前記候補のベクトルをランダムな方向に修正することによって反復的に洗練され、

前記ランダム方向最適化(560)は、前記第1の出力信号と前記第2の出力信号との間の類似性または非類似性を示すメトリック(D_{KL} 、 D 、532)が測定され、前記第1の出力信号および前記第2の出力信号(S'_0 、 S'_1)が、最も低い類似性または最も高い非類似性を示すメトリックに関連付けられた前記候補パラメータに関連付けられた測定値であるように選択されるようなものであり、

前記メトリック(532)は、カルバックライブラー発散として処理される、

装置(500)。

【請求項2】

前記第1の入力信号(502、 M_0)と結合されるべき前記第2の入力信号(502、 M_1)の前記遅延されスケールされたバージョン(503₁)は、前記第2の入力信号(502、 M_1)に部分遅延を適用することによって得られる、請求項1に記載の装置。

【請求項3】

前記第2の入力信号(502、 M_1)と結合されるべき前記第1の入力信号(502、 M_0)の前記遅延されスケールされたバージョン(503₀)は、前記第1の入力信号(502、 M_0)に部分遅延を適用することによって得られる、請求項1または2に記載の装置。

【請求項4】

10

20

30

40

50

正規化された大きさの値の第 1 のセットのそれぞれの要素 (P_0) と、
 正規化された大きさの値の第 1 のセット (5 2 2) のそれぞれの要素 (P_0) および
 正規化された大きさの値の第 2 のセット (5 2 2) のそれぞれの要素 (P_1)
 に基づいて形成された商 (7 0 2') の対数 (7 0 6') と、
 の間の複数の積 (7 1 2'、7 1 0') を合計する (7 1 4、7 1 6) ように構成されて
 おり、

その結果、前記第 1 の正規化された大きさの値のセット (P_0) によって記述される信
 号部分 (s'_0) と前記第 2 の正規化された大きさの値のセット (P_1) によって記述さ
 れる信号部分 (s'_1) との間の類似性または非類似性を記述する値 (D_{KL} 、 D 、5 3
 2) を取得する (5 3 0)、請求項 1 から 3 のいずれか一項に記載の装置。

10

【請求項 5】

前記第 1 および第 2 の入力信号 (M_0 、 M_1) の少なくとも一方について、前記それぞ
 れの要素 (P_0) は、前記候補パラメータから得られる前記候補第 1 または第 2 の出力信
 号 (S'_0 、 S'_1) に基づく、請求項 4 に記載の装置。

【請求項 6】

前記第 1 および第 2 の入力信号 (M_0 、 M_1) のうちの少なくとも 1 つについて、前記
 それぞれの要素 (P_0) は、

候補の第 1 または第 2 の出力信号 (S'_0 、 S'_1) に関連付けられた値、および
 前記第 1 の出力信号または前記第 2 の出力信号の前記以前に取得された値に関連付けら
 れたノルム、との分数として得られる、

20

請求項 5 に記載の装置。

【請求項 7】

前記第 1 および第 2 の入力信号 (M_0 、 M_1) のうちの少なくとも一方について、前記
 それぞれの要素は、

$$P_i(n) = \frac{|s'_i(n)|}{\|s'_i\|_1} \quad (7)$$

30

によって取得される、請求項 5 に記載の装置。

【請求項 8】

前記第 1 のスケール値および前記第 2 のスケール値ならびに前記第 1 の遅延値
 および前記第 2 の遅延値は、目的関数を最小化することによって取得される、請求項 1 か
 ら 7 のいずれか一項に記載の装置。

【請求項 9】

前記第 1 の入力信号 (5 0 2、 M_0)、またはその処理されたバージョンを、前記時間
 領域および / または前記 z 変換もしくは周波数領域において、前記第 2 の入力信号 (5 0
 2、 M_1) の前記遅延されスケールされたバージョン (5 0 3₁) と結合するように
 さらに構成され、

40

前記第 2 の入力信号 (5 0 2、 M_1)、またはその処理されたバージョンを、前記時間
 領域および / または前記 z 変換もしくは周波数領域において前記第 1 の入力信号 (5 0 2
 、 M_0) の前記遅延およびスケールされたバージョン (5 0 3₀) と組み合わせるよ
 うにさらに構成されている、

請求項 1 から 8 のいずれか一項に記載の装置。

【請求項 10】

前記最適化は、前記 z 変換領域で実行される、

請求項 1 から 9 のいずれか一項に記載の装置。

【請求項 11】

50

前記最適化は、前記時間領域で実行される、
請求項 1 から 9 のいずれか一項に記載の装置。

【請求項 1 2】

前記最適化は、前記周波数領域で実行される、
請求項 1 から 9 のいずれか一項に記載の装置。

【請求項 1 3】

前記第 2 の入力信号 (5 0 2 、 M_1) に適用される前記遅延または部分遅延 (d_0) は

、
前記第 1 のマイクロフォン (mic_0) によって受信された前記第 1 の音源 ($source_0$) からの前記信号、および

前記第 2 のマイクロフォン (mic_1) によって受信された前記第 1 の音源 ($source_0$) からの前記信号、

の間の関係、および / または差もしくはは到着を示す、

請求項 1 から 1 2 のいずれか一項に記載の装置。

【請求項 1 4】

前記第 1 の入力信号 (5 0 2 、 M_0) に適用される前記遅延または部分遅延 (d_1) は

、
前記第 2 のマイクロフォン (mic_1) によって受信された前記第 2 の音源 ($source_1$) からの前記信号、および

前記第 1 のマイクロフォン (mic_0) によって受信される前記第 2 の音源 ($source_1$) からの前記信号、

の間の関係、および / または差もしくはは到着を示す、

請求項 1 から 1 3 のいずれか一項に記載の装置。

【請求項 1 5】

前記メトリックは、

$$D_{KL}(P||Q) = \sum_n P(n) \log \left(\frac{P(n)}{Q(n)} \right)$$

の形態で取得され、

式中、 $P(n)$ は前記第 1 の入力信号に関連付けられた要素であり、 $Q(n)$ は前記第 2 の入力信号に関連付けられた要素である、

請求項 1 から 1 4 のいずれか一項に記載の装置。

【請求項 1 6】

前記メトリックは、

$$D(P_0, P_1) = - \sum_n \left[P_0(n) \log \left(\frac{P_0(n)}{P_1(n)} \right) + P_1(n) \log \left(\frac{P_1(n)}{P_0(n)} \right) \right]$$

の形態で取得され、

式中、 $P_1(n)$ は、前記第 1 の入力信号に関連付けられた要素であり、 $P_2(n)$ は、前記第 2 の入力信号に関連付けられた要素である、

請求項 1 から 1 5 のいずれか一項に記載の装置。

【請求項 1 7】

前記取得された第 1 および第 2 の出力信号 (S'_0 、 S'_1) に関連する情報を周波数領域に変換するようにさらに構成される、請求項 1 から 1 6 のいずれか一項に記載の装置。

【請求項 1 8】

前記取得された第 1 および第 2 の出力信号 (S'_0 、 S'_1) に関連する情報を符号化するようにさらに構成される、請求項 1 から 1 7 のいずれか一項に記載の装置。

【請求項 1 9】

前記取得された第 1 および第 2 の出力信号 (S'_0 、 S'_1) に関連する情報を格納するようにさらに構成された、請求項 1 から 1 8 のいずれか一項に記載の装置。

10

20

30

40

50

【請求項 20】

前記取得された第1および第2の出力信号 (S'_0 、 S'_1) に関連する情報を送信するようにさらに構成される、請求項1から19のいずれか一項に記載の装置。

【請求項 21】

前記第1の入力信号 (502 、 M_0) を取得するための第1のマイクロフォン (mic_0)、および前記第2の入力信号 (502 、 M_1) を取得するための第2のマイクロフォン (mic_1) のうちの少なくとも1つと、を含む、請求項1から20のいずれか一項に記載の装置。

【請求項 22】

前記取得された第1および第2の出力信号 (S'_0 、 S'_1) に関連する情報を送信するための機器と、請求項1から21のいずれか一項に記載の装置とを含む、テレビ会議のための装置。 10

【請求項 23】

請求項1から22のいずれか一項に記載の装置を含むバイノーラルシステム。

【請求項 24】

複数の入力信号 (502 、 M_0 、 M_1) に基づいて、異なる音源 ($source_0$ 、 $source_1$) に関連付けられた複数の出力信号 (504 、 S'_0 、 S'_1) を取得するための方法であって、前記音源 ($source_0$ 、 $source_1$) からの信号が結合され、

前記方法は、 20

第1の出力信号 (504 、 S'_0) を得るために、第1の入力信号 (502 、 M_0) またはその処理されたバージョンを第2の入力信号 (502 、 M_1) の遅延されスケールングされたバージョン (503_1) と結合するステップと、

第2の入力信号 (502 、 M_1) またはその処理されたバージョンを、前記第1の入力信号 (502 、 M_0) の遅延されスケールングされたバージョン (503_0) と結合して、第2の出力信号 (504 、 S'_1) を取得するステップと、

ランダム方向最適化 (560) を使用して、

前記第1の入力信号 (502 、 M_0) の前記遅延されスケールングされたバージョン (503_0) を取得するために使用される第1のスケールング値 (564 、 a_0) と、

前記第1の入力信号 (502 、 M_0) の前記遅延されスケールングされたバージョン (503_0) を取得するために使用される第1の遅延値 (564 、 d_0) と、 30

前記第2の入力信号 (502 、 M_1) の前記遅延されスケールングされたバージョン (503_1) を取得するために使用される第2のスケールング値 (564 、 a_1) と、

前記第2の入力信号の前記遅延およびスケールングされたバージョン (503_1) を取得するために使用される第2の遅延値 (564 、 d_1) と、

のうちの少なくとも1つを決定するステップと、を含み、

前記ランダム方向最適化 (560) は、候補パラメータが候補のベクトル (564 、 a_0 、 a_1 、 d_0 、 d_1) を形成するようなものであり、前記候補のベクトルは、前記候補のベクトルをランダムな方向に修正することによって反復的に洗練され、

前記ランダム方向最適化 (560) は、前記第1の出力信号と前記第2の出力信号との間の類似性または非類似性を示すメトリック (D_{KL} 、 D 、 532) が測定され、前記第1の出力信号および前記第2の出力信号 (S'_0 、 S'_1) が、最も低い類似性または最も高い非類似性を示すメトリックに関連付けられた前記候補パラメータに関連付けられた測定値であるように選択されるようなものであり、 40

前記メトリック (532) は、カルバックライブラー発散として処理される、方法。

【請求項 25】

請求項1から20のいずれか一項に記載の装置を使用するように構成された、請求項24に記載の方法。

【請求項 26】

前記部分遅延 (d_1) は、前記第 2 のソース ($source_1$) から前記第 1 のマイクロフォン (mic_0) に到達する前記信号 (M_0) の前記遅延と、前記第 2 のソース ($source_1$) から前記第 2 のマイクロフォン (mic_1) に到達する前記信号 (M_1) の前記遅延 ($H_{1,1}$) との間の関係および / または差を示す、請求項 24 に記載の方法。

【請求項 27】

プロセッサによって実行されると、前記プロセッサに請求項 24 から 26 のいずれか一項に記載の方法を実行させる命令を格納する非一時的記憶ユニット。

10

20

30

40

50

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT

International application No PCT/EP2020/077716

A. CLASSIFICATION OF SUBJECT MATTER INV. G10L21/0272 G06N3/12 ADD.		
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) G10L G06N		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) EPO-Internal, INSPEC		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X,P	GOLOKOLENKO OLEG ET AL: "A Fast Stereo Audio Source Separation for Moving Sources", 2019 53RD ASILOMAR CONFERENCE ON SIGNALS, SYSTEMS, AND COMPUTERS, IEEE, 3 November 2019 (2019-11-03), pages 1931-1935, XP033750517, DOI: 10.1109/IEEECONF44664.2019.9048652 the whole document ----- -/--	1-7, 16-53, 61,64, 65,67
<input checked="" 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 4 December 2020		Date of mailing of the international search report 05/02/2021
Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016		Authorized officer Dobler, Ervin

1

Form PCT/ISA/210 (second sheet) (April 2005)

page 1 of 2

10

20

30

40

50

INTERNATIONAL SEARCH REPORT

International application No.
PCT/EP2020/077716

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).

10

20

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 additional sheet

- 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 an 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, 61(completely); 16-53, 64, 65, 67(partially)

30

40

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.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

- 1. claims: 1-7, 61(completely); 16-53, 64, 65, 67(partially)

- Source separation with random direction optimization

- 2. claims: 8-10, 62(completely); 16-53, 64, 65, 67(partially)

- Source separation with optimization with fractional delays

- 3. claims: 11-15, 63(completely); 16-53, 64, 65, 67(partially)

- Source separation with a Kullback-Leibler fitting function

- 4. claims: 54-60, 66(completely); 67(partially)

- An optimization algorithm with randomly generated candidate vectors

10

20

30

40

50

フロントページの続き

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,N
E,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,DJ,DK,DM,DO,DZ,EC,EE,EG,ES,FI,GB,GD,GE,GH,GM,GT,HN,HR,HU,ID,IL,IN,IR,IS,IT,JO,JP,K
E,KG,KH,KN,KP,KR,KW,KZ,LA,LC,LK,LR,LS,LU,LY,MA,MD,ME,MG,MK,MN,MW,MY,MZ,NA,N
G,NI,NO,NZ,OM,PA,PE,PG,PH,PL,PT,QA,RO,RS,RU,RW,SA,SC,SD,SE,SG,SK,SL,ST,SV,SY,TH,TJ,TM,
TN,TR,TT,TZ,UA,UG,US,UZ,VC,VN,WS,ZA,ZM,ZW

(特許庁注：以下のものは登録商標)

1 . M A T L A B

ドイツ連邦共和国， 9 8 6 9 3 イルメナウ， エレンベルクシュトラッセ 3 1 フラウンホーファ
ー - インスティテュート・フル・ディギターレ・メーディエンテヒノロジー， イーデーエムター内