

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第5105336号
(P5105336)

(45) 発行日 平成24年12月26日 (2012.12.26)

(24) 登録日 平成24年10月12日 (2012.10.12)

(51) Int. Cl. F I
G 1 0 L 21/02 (2006.01)
 G 1 0 L 21/02 1 0 1 B
 G 1 0 L 21/02 1 0 2 A
 G 1 0 L 21/02 2 0 1 C

請求項の数 4 (全 14 頁)

<p>(21) 出願番号 特願2009-282026 (P2009-282026) (22) 出願日 平成21年12月11日 (2009.12.11) (65) 公開番号 特開2011-123370 (P2011-123370A) (43) 公開日 平成23年6月23日 (2011.6.23) 審査請求日 平成21年12月11日 (2009.12.11)</p> <p>(出願人による申告) 国等の委託研究の成果に係る特許出願 (平成19年、20年度経済産業省戦略的技術開発委託費 (情報家電センサー・ヒューマンインターフェイスデバイス活用技術開発<音声認識基盤技術の開発>に係るもの) に関する再委託研究、産業技術力強化法第19条の適用を受けるもの)</p>	<p>(73) 特許権者 000000295 沖電気工業株式会社 東京都港区虎ノ門一丁目7番12号 (73) 特許権者 899000068 学校法人早稲田大学 東京都新宿区戸塚町1丁目104番地 (74) 代理人 100090620 弁理士 工藤 宣幸 (74) 代理人 100161861 弁理士 若林 裕介 (72) 発明者 森戸 誠 東京都港区西新橋三丁目16番11号 沖電気工業株式会社内</p>
---	---

最終頁に続く

(54) 【発明の名称】 音源分離装置、プログラム及び方法

(57) 【特許請求の範囲】

【請求項1】

間隔を置いて配置された複数個のマイクロフォンのうち、2個のマイクロフォンの受信信号のスペクトルについて、目的音が到来すると想定される想定到来方向を含む所定の範囲内で、それぞれ異なる方向に、成分抑圧の指向性に向けて処理する複数の目的音抑圧部を用いて、上記受信信号のスペクトルから、上記目的音の成分を抑圧した目的音抑圧スペクトルを生成する目的音抑圧スペクトル生成手段と、

上記受信信号のスペクトルについて、上記所定の範囲以外の任意の方向から到来する妨害音を抑圧した目的音優勢スペクトルを生成する目的音優勢スペクトル生成手段と、

目的音抑圧スペクトルと、目的音優勢スペクトルとを用いて、上記受信信号について、上記妨害音の成分と上記目的音の成分とを分離する分離手段とを有し、

上記目的音抑圧スペクトル生成手段は、目的音抑圧スペクトルの各成分について、上記目的音抑圧部の処理結果のうち最も絶対値の小さい値を適用する

ことを特徴とする音源分離装置。

【請求項2】

上記目的音抑圧スペクトル生成手段、及び上記目的音優勢スペクトル生成手段を有するスペクトル生成処理部を複数備え、

上記分離手段は、それぞれの上記スペクトル生成処理部が生成する、目的音抑圧スペクトル及び目的音優勢スペクトルを利用して、上記受信信号について、上記妨害音の成分と上記目的音の成分とを分離する

ことを特徴とする請求項 1 に記載の音源分離装置。

【請求項 3】

音源分離装置に搭載されたコンピュータを、

間隔を置いて配置された複数個のマイクロフォンのうち、2 個のマイクロフォンの受信信号のスペクトルについて、目的音が到来すると想定される想定到来方向を含む所定の範囲内で、それぞれ異なる方向に、成分抑圧の指向性を向けて処理する複数の目的音抑圧部を用いて、上記受信信号のスペクトルから、上記目的音の成分を抑圧した目的音抑圧スペクトルを生成する目的音抑圧スペクトル生成手段と、

上記受信信号のスペクトルについて、上記所定の範囲以外の任意の方向から到来する妨害音を抑圧した目的音優勢スペクトルを生成する目的音優勢スペクトル生成手段と、

目的音抑圧スペクトルと、目的音優勢スペクトルとを用いて、上記受信信号について、上記妨害音の成分と上記目的音の成分とを分離する分離手段として機能させ、

上記目的音抑圧スペクトル生成手段は、目的音抑圧スペクトルの各成分について、上記目的音抑圧部の処理結果のうち最も絶対値の小さい値を適用する

ことを特徴とする音源分離プログラム。

【請求項 4】

音源分離装置により行われる音源分離方法において、

目的音抑圧スペクトル生成手段、目的音優勢スペクトル生成手段、分離手段を有し、

上記目的音抑圧スペクトル生成手段は、間隔を置いて配置された複数個のマイクロフォンのうち、2 個のマイクロフォンの受信信号のスペクトルについて、目的音が到来すると想定される想定到来方向を含む所定の範囲内で、それぞれ異なる方向に、成分抑圧の指向性を向けて処理する複数の目的音抑圧部を用いて、上記受信信号のスペクトルから、上記目的音の成分を抑圧した目的音抑圧スペクトルを生成し、

上記目的音優勢スペクトル生成手段は、上記受信信号のスペクトルについて、上記所定の範囲以外の任意の方向から到来する妨害音を抑圧した目的音優勢スペクトルを生成し、

上記分離手段は、目的音抑圧スペクトルと、目的音優勢スペクトルとを用いて、上記受信信号について、上記妨害音の成分と上記目的音の成分とを分離し、

上記目的音抑圧スペクトル生成手段は、目的音抑圧スペクトルの各成分について、上記目的音抑圧部の処理結果のうち最も絶対値の小さい値を適用する

ことを特徴とする音源分離方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、音源分離装置、プログラム及び方法に関し、例えば、電話装置や音声認識装置等の音声捕捉における雑音除去に適用し得る。

【背景技術】

【0002】

電話装置や音声認識装置では、マイクロフォンによりユーザ音声を捕捉するが、周囲雑音によって、音声認識の精度は極度に劣化したり、録音した音声が雑音のために聞き取りにくい場合がある。

【0003】

このため、従来は、マイクロフォンアレーにより指向特性を制御する等して、所望の目的音だけを選択的に捕捉する試みがなされているが、このような指向特性の制御だけでは、所望の音声を背景雑音から分離して取り出すことは困難であった。

【0004】

従来のマイクロフォンアレーの技術としては、例えば、遅延和アレー (DSA: Delayed Sum Array) や、BF (Beam-Forming) と呼ばれる指向特性制御に関する技術、あるいは DCMP (Directionally Constrained Minimization of Power) アダプティブアレーによる指向特性制御に関する技術等がある。

【 0 0 0 5 】

一方、遠隔発話による音声を分離する技術として、複数の固定マイクロフォンの出力信号を狭帯域スペクトル分析し、周波数帯域毎に最も大きな振幅を与えたマイクロフォンにその周波数帯域の音を割り当てる技術（S A F I Aと称されている）としては、特許文献1の記載技術がある。特許文献1に記載されている帯域選択（B S : B a n d S e l e c t i o n）による音声を分離技術では、所望の音声を得るために、所望の音声を発する音源に最も近いマイクロフォンを選び、そのマイクロフォンに割り当てられた周波数帯域の音を使って音声を合成する。

【 0 0 0 6 】

また、更なる技術として、帯域選択の方法に改良を加えた技術が特許文献2に記載されている。

10

【 0 0 0 7 】

特許文献2の記載技術では、目的音到来方向と直角または略直角をなす方向に並べて配置された2個のマイクロフォンに入力された信号を用いて、妨害音を抑圧して捕捉対象である目的音を強調した目的音優勢信号と、目的音を抑圧して妨害音を強調した目的音劣勢信号を作成し、その2種類の信号を目的音と妨害音の分離に利用している。

【 0 0 0 8 】

特許文献2では、目的音優勢信号及び目的音劣勢信号の生成について「空間フィルタ」と呼ばれるフィルタを用いて実現している。

【 0 0 0 9 】

図3は、空間フィルタの特性について示した説明図である。

20

【 0 0 1 0 】

以下では、2つのマイクロフォンM1、M2を結ぶ線に対する垂直平面を0度の方向と呼び、時計回りの方向を正の角度、反時計回りの方向を負の角度として方向を表すものとする。すなわち、上述の方向は-180度~180度（-180度と180度は同じ方向）の範囲で表される。

【 0 0 1 1 】

図3では、間隔dで配置された2つのマイクロフォンM1、M2に対して角度θの方向から入力される音源があった場合について説明している。この場合、角度θの方向から入力される音源から2つのマイクロフォンM1、M2への距離で、 $d \times \sin \theta$ の距離差が生じ、結果として音の到達時間については、マイクロフォンM1、M2の間で、以下の(1)式で表される時間差τが生じる。

30

【 数 1 】

$$\tau = \frac{d \times \sin \theta}{(\text{音の伝播速度})} \dots (1)$$

【 0 0 1 2 】

そこで、マイクロフォンM2の出力からマイクロフォンM1の出力を時間差τ遅延させた出力を減じると、お互いが相殺されθ方向の音は抑圧される。以下では、空間フィルタにおいて音を抑圧する方向の角度（上述の例ではθ）を「抑圧角度」と呼ぶ。

40

【 0 0 1 3 】

図4は、空間フィルタにおける指向特性について示した説明図である。

【 0 0 1 4 】

図4において、曲線Lは、空間フィルタの抑圧角度をθに設定した場合の指向特性を表しており、マイクロフォンM1、M2を結ぶ線の midpoint からの距離が長くなっている方向ほど、利得が大きく（抑圧の強度が弱い）、距離が短いほど利得が小さい（抑圧の強度が強い）ことを示している。

【 0 0 1 5 】

50

図4では、空間フィルタの抑圧角度を の方向に設定しているため、その方向の抑圧の強度が最も大きくなるように設定されている様子を示している。

【先行技術文献】

【特許文献】

【0016】

【特許文献1】特開平10-313497号公報

【特許文献2】特開2006-197552号公報

【発明の概要】

【発明が解決しようとする課題】

【0017】

しかしながら、特許文献1の記載技術では、2つの音が重なった状況において、よく両者を分離することができるが、音源が3つ以上となると、理論的には分離可能とされているものの、分離性能は極端に劣化する。従って、複数の雑音源が存在する状況下で、これらの複数の雑音から目的音を精度よく分離することは困難である。

【0018】

また、特許文献2の記載技術では、空間フィルタを用いて、目的音を分離する処理を行っているが、分離する処理をしている途中で目的音の到来方向がずれた場合に、空間フィルタの特性が、分離後の目的音の品質に影響を及ぼす恐れがある。以下、特許文献2に記載の方法において、空間フィルタの特性が、分離後の目的音に及ぼす恐れがある影響について説明する。

【0019】

図5は、空間フィルタにおける抑圧角度に近い方向の利得の変化特性について示した説明図である。

【0020】

図5では、空間フィルタの抑圧角度を とし、目的音が0度の方向（正面）から到来した場合の利得をG1、0度から反時計回りに僅かにずれた方向から目的音が到来した場合の利得をG2として説明している。

【0021】

空間フィルタにおいて、抑圧角度の近くで、角度の変化に応じた利得の変化率が大きい場合には、図5に示すように、利得がG1となる方向と、G2となる方向との角度のずれが僅かであっても、G1とG2の差は大きくなってしまふ恐れがある。

【0022】

上述の特許文献2に記載されている目的音劣勢信号生成手段では、目的音が到来すると想定される方向に、空間フィルタの抑圧角度を向けて、目的音成分を抑圧するとともに、妨害音成分を抽出しているが、上述のように、目的音と妨害音とを分離する処理をしている途中で目的音の到来方向がずれると、僅かなずれであっても、出力音に大きなゆれを生じる結果となる恐れがある。

【0023】

そのため、目的音と、目的音の到来方向以外の任意の方向から到来する妨害音とを分離する処理において、目的音の到来方向がずれた場合でも、分離処理後の音の品質を保つことができる音源分離装置、プログラム及び方法が望まれている。

【課題を解決するための手段】

【0024】

第1の本発明の音源分離装置は、(1)間隔を置いて配置された複数個のマイクロフォンのうち、2個のマイクロフォンの受信信号のスペクトルについて、目的音が到来すると想定される想定到来方向を含む所定の範囲内で、それぞれ異なる方向に、成分抑圧の指向性を向けて処理する複数の目的音抑圧部を用いて、上記受信信号のスペクトルから、上記目的音の成分を抑圧した目的音抑圧スペクトルを生成する目的音抑圧スペクトル生成手段と、(2)上記受信信号のスペクトルについて、上記所定の範囲以外の任意の方向から到来する妨害音を抑圧した目的音優勢スペクトルを生成する目的音優勢スペクトル生成手段

10

20

30

40

50

と、(3)目的音抑圧スペクトルと、目的音優勢スペクトルとを用いて、上記受信信号について、上記妨害音の成分と上記目的音の成分とを分離する分離手段とを有し、(4)上記目的音抑圧スペクトル生成手段は、目的音抑圧スペクトルの各成分について、上記目的音抑圧部の処理結果のうち最も絶対値の小さい値を適用することを特徴とする。

【0025】

第2の本発明の音源分離プログラムは、音源分離装置に搭載されたコンピュータを、(1)間隔を置いて配置された複数個のマイクロフォンのうち、2個のマイクロフォンの受信信号のスペクトルについて、目的音が到来すると想定される想定到来方向を含む所定の範囲内で、それぞれ異なる方向に、成分抑圧の指向性を向けて処理する複数の目的音抑圧部を用いて、上記受信信号のスペクトルから、上記目的音の成分を抑圧した目的音抑圧スペクトルを生成する目的音抑圧スペクトル生成手段と、(2)上記受信信号のスペクトルについて、上記所定の範囲以外の任意の方向から到来する妨害音を抑圧した目的音優勢スペクトルを生成する目的音優勢スペクトル生成手段と、(3)目的音抑圧スペクトルと、目的音優勢スペクトルとを用いて、上記受信信号について、上記妨害音の成分と上記目的音の成分とを分離する分離手段として機能させ、(4)上記目的音抑圧スペクトル生成手段は、目的音抑圧スペクトルの各成分について、上記目的音抑圧部の処理結果のうち最も絶対値の小さい値を適用することを特徴とする。

【0026】

第3の本発明は、音源分離装置により行われる音源分離方法において、(1)目的音抑圧スペクトル生成手段、目的音優勢スペクトル生成手段、分離手段を有し、(2)上記目的音抑圧スペクトル生成手段は、間隔を置いて配置された複数個のマイクロフォンのうち、2個のマイクロフォンの受信信号のスペクトルについて、目的音が到来すると想定される想定到来方向を含む所定の範囲内で、それぞれ異なる方向に、成分抑圧の指向性を向けて処理する複数の目的音抑圧部を用いて、上記受信信号のスペクトルから、上記目的音の成分を抑圧した目的音抑圧スペクトルを生成し、(3)上記目的音優勢スペクトル生成手段は、上記受信信号のスペクトルについて、上記所定の範囲以外の任意の方向から到来する妨害音を抑圧した目的音優勢スペクトルを生成し、(4)上記分離手段は、目的音抑圧スペクトルと、目的音優勢スペクトルとを用いて、上記受信信号について、上記妨害音の成分と上記目的音の成分とを分離し、(5)上記目的音抑圧スペクトル生成手段は、目的音抑圧スペクトルの各成分について、上記目的音抑圧部の処理結果のうち最も絶対値の小さい値を適用することを特徴とする。

【発明の効果】

【0027】

本発明によれば、目的音と、目的音の到来方向以外の任意の方向から到来する妨害音とを分離する処理において、目的音の到来方向がずれた場合でも、分離処理後の音の品質を保つことができる。

【図面の簡単な説明】

【0028】

【図1】第1の実施形態に係る音源分離装置の機能的構成について示したブロック図である。

【図2】第2の実施形態に係る音源分離装置の機能的構成について示したブロック図である。

【図3】従来の空間フィルタの特性について示した説明図である。

【図4】従来の空間フィルタにおける指向特性について示した説明図である。

【図5】従来の空間フィルタにおける抑圧角度に近い方向の利得の変化特性について示した説明図である。

【発明を実施するための形態】

【0029】

(A)第1の実施形態

以下、本発明による音源分離装置、プログラム及び方法の第1の実施形態を、図面を参

10

20

30

40

50

照しながら詳述する。

【0030】

(A-1)第1の実施形態の構成及び動作

図1は、第1の実施形態の音源分離装置10の機能的構成について示したブロック図である。

【0031】

音源分離装置10は、目的音と、目的音の到来方向以外の任意の方向から到来する妨害音とを分離するものである。音源分離装置10の用途は限定されるものではないが、例えば、音声認識装置や、携帯電話などの電話装置に搭載して、音声捕捉に用いるようにしても良い。具体的には、例えば、音源分離装置10を電話会議装置に搭載して、遠隔発話を行う複数の話者による混合音声から任意の話者の音声を目的音として分離したり、遠隔発話を行う話者の音声とその他の音との混合音から話者の音声を目的音として分離したりすることに用いるようにしても良い。また、例えば、音声対話を行うロボット、カーナビゲーションシステム等の車載機器についての音声操作、会議の議事録作成等の音声認識において、目的音となるユーザの音声の分離に用いるようにしても良い。

10

【0032】

音源分離装置10は、大きくは、入力手段20、分析手段30、分離手段40、除去手段50、生成手段60を有する。

【0033】

音源分離装置10は、マイクロフォン等のハードウェア以外の構成要素に関しては、プロセッサ(CPU等)を有する装置に、実施形態の音源分離プログラムをインストールすることにより実現するようにしても良いし、一部又は全部の構成要素について、専用のハードウェア(例えば、半導体チップ)を用いて実現するようにしても良い。

20

【0034】

入力手段20は、間隔を置いて配置された2個のマイクロフォン21、22と、これらの2個のマイクロフォン21、22の受信信号をアナログ/デジタル信号変換器(図示せず)を用いてデジタル信号に変換し、そのデジタル信号を分析手段30に与える。

【0035】

以下では、上述の図3~図5と同様に、2つのマイクロフォン21、22を結ぶ線に対する垂直平面を0度の方向と呼び、時計回りの方向を正の角度、反時計回りの方向を負の角度として方向を表すものとする。すなわち、上述の方向は-180度~180度(-180度と180度とは同じ方向)の範囲で表される。

30

【0036】

また、以下では、例として、音源分離装置10は、目的音が概ね0度の方向から到来することを想定した構成として説明する。

【0037】

以下の説明においては、マイク21から出力されるデジタル音声信号を $x_1(n)$ とする。また、同様にマイク22から出力されるデジタル音声信号を $x_2(n)$ とする。但し、 n は、 n 番目のデータ(サンプル)を表すものとする。

【0038】

デジタル音声信号 $x_1(n)$ 、 $x_2(n)$ は、例えば、マイクロフォンなどの音声入力装置から入力されたアナログ音声信号を、アナログ/デジタル変換し、標本化周期 T 毎に標本化することにより得られるものである。標本化周期 T は、例えば、31.25マイクロ秒~125マイクロ秒程度とすることが望ましい。

40

【0039】

同一時間区間における、 N 個の連続する $x_1(n)$ 、 $x_2(n)$ を1つの分析単位(フレーム)として、後述する分析手段30、分離手段40、除去手段50、生成手段60の処理は行われるものとする。

【0040】

以下の説明において、音源分離装置10では、例として $N=1024$ とする。そして、

50

音源分離装置 10 では、処理対象分析単位に対する当該音源分離の一連の処理が終了すると、 $x_1(n)$ 、 $x_2(n)$ のうち後半の $3N/4$ 個のデータを前半にシフトし、新たに入力された連続する $N/4$ 個のデータを後半に接続するものとする。これにより、音源分離装置 10 では、新たな N 個の連続する $x_1(n)$ 、 $x_2(n)$ を生成し、1つの分析単位として新たな処理を行うものとする。音源分離装置 10 では、このような処理対象分析単位の処理を繰り返すようになされているものとする。

【0041】

なお、分析手段 30 に入力されるデジタル音声信号はマイクロフォンが捕捉してアナログ/デジタル変換されたものに限定されない。例えば、記録媒体などから読み出されたものであっても良いし、他の装置から通信によって与えられたものであっても良い。すなわち、音源分離装置 10 において、 $x_1(n)$ 、 $x_2(n)$ が保持できれば、入力手段 20 を省略した構成としても良い。

10

【0042】

分析手段 30 は、雑音の混在したデジタル音声信号 $x_1(n)$ 、 $x_2(n)$ が、入力手段 20 から与えられると、 $x_1(n)$ を周波数分析部 31 で、 $x_2(n)$ を周波数分析部 32 で、それぞれ FFT (高速フーリエ変換) 処理等を行い、その結果を分離手段 40 に与える。分析手段 30 では、FFT 処理にあたっては、 N 個の連続する $x_1(n)$ 、 $x_2(n)$ に対し、窓関数をかけるものとする。なお、窓関数 $w(n)$ としては、各種の窓関数を適用可能であるが、例えば、以下の (2) 式に示すようなハニング窓を適用するようにしても良い。

20

【数 2】

$$w(n) = 0.5 - 0.5 \cdot \cos\left(\frac{2n\pi}{N}\right) \quad \dots (2)$$

【0043】

分析手段 30 による上述の窓処理は、後述する生成手段 60 における分析単位の接続処理を考慮してなされる処理である。ただし、上述の窓関数を適用することは好ましいが、必須ではない。

30

【0044】

以下では、周波数分析部 31、32 の出力を、それぞれ $D_1(m)$ 、 $D_2(m)$ と表すものとする。なお、 $D_1(m)$ 、 $D_2(m)$ は複素数である。

【0045】

なお、分析手段 30 における分析方法は、FFT に限定されず、DFT (離散フーリエ変換) などの他の周波数分析方法を適用するようにしても良い。

【0046】

また、音源分離装置 10 が搭載される装置によっては、他の目的の処理装置における分析に関する構成を、この音源分離装置 10 の構成として流用するようにしても良い。例えば、当該音源分離装置 10 が搭載される装置が IP 電話装置の場合には、このような流用が可能である。IP 電話装置の場合、IP パケットのペイロードには FFT 出力を符号化したものを挿入するが、その FFT 出力を、上述した分析手段 30 の出力として流用することができる。

40

【0047】

また、後述する分離手段 40 の処理では、スペクトル $D(m)$ の性質 $D(m) = D^*(N-m)$ (ただし、 $1 \leq m \leq N/2 - 1$ 、 $D^*(N-m)$ は $D(N-m)$ の共役複素数を表す) から、 $0 \leq m \leq N/2$ の範囲で行えば良い。

【0048】

分離手段 40 は、妨害音抑圧部 41 及び目的音抑圧部 42 を有している。

50

【 0 0 4 9 】

妨害音抑圧部 4 1 は、D 1 (m)、D 2 (m) を利用して、妨害音の成分を抑圧し、目的音の成分が強調されたスペクトルを生成する。そして、目的音抑圧部 4 2 は、D 1 (m)、D 2 (m) を利用して、目的音の成分を抑圧し、妨害音の成分が強調されたスペクトルを生成する。

【 0 0 5 0 】

次に、妨害音抑圧部 4 1 の構成について説明する。

【 0 0 5 1 】

妨害音抑圧部 4 1 は、2 つの空間フィルタ 4 1 1、4 1 2 及び最小選択部 4 1 3 を有している。

10

【 0 0 5 2 】

空間フィルタ 4 1 1、4 1 2 の抑圧角度は、それぞれ、9 0 度、- 9 0 度に設定されているものとする。

【 0 0 5 3 】

上述の通り、音源分離装置 1 0 では、目的音は、概ね 0 度の方向から到来することが想定されているため、妨害音抑圧部 4 1 では、目的音が到来する方向とは異なる方向に、空間フィルタの抑圧角度を向けているが、目的音が到来すると想定される方向に応じて、空間フィルタの数や抑圧角度の組み合わせを変更するようにしても良い。

【 0 0 5 4 】

空間フィルタ 4 1 1 の具体的な処理としては、以下の (3) 式を用いて、E 1 (m) を求める。また、空間フィルタ 4 1 2 は、以下の (4) 式を用いて、E 2 (m) を求める。以下の (3) 式、(4) 式において、f はサンプリング周波数であり、例えば、1 6 0 0 H z を適用するようにしても良い。

20

【 0 0 5 5 】

そして、最小選択部 4 1 3 は、以下の (5) 式に示すように、空間フィルタ 4 1 1 の出力 E 1 (m) と空間フィルタ 4 1 2 の出力 E 2 (m) の絶対値の最小値 M (m) を、算出して M (m) を求める。この出力 M (m) は、目的音の成分を抽出したのものとして、最小選択部 4 1 3 から除去手段 5 0 に与えられる。

【 数 3 】

$$E1(m) = D1(m) - D2(m) \cdot e^{\frac{-j2\pi mf}{N}\tau} \quad \dots (3)$$

30

$$E2(m) = D2(m) - D2(m) \cdot e^{\frac{-j2\pi mf}{N}\tau} \quad \dots (4)$$

$$M(m) = \min(\|E1(m)\|, \|E2(m)\|) \quad (\text{ただし } \|\cdot\| \text{ は絶対値を表す}) \quad \dots (5)$$

【 0 0 5 6 】

次に、目的音抑圧部 4 2 の構成について説明する。

40

【 0 0 5 7 】

目的音抑圧部 4 2 は、3 つの空間フィルタ 4 2 1、4 2 2、4 2 3 及び最小選択部 4 2 4 を有している。

【 0 0 5 8 】

空間フィルタ 4 2 1、4 2 2、4 2 3 の抑圧角度は、それぞれ、0 度、5 度、- 5 度の方向に設定されているものとする。

【 0 0 5 9 】

上述の通り、音源分離装置 1 0 では、目的音は、概ね 0 度の方向から到来することが想定されているため、目的音抑圧部 4 2 では、空間フィルタ 4 2 1 の抑圧角度を 0 度に設定し、0 度の方向から、わずか (± 5 度程度) にずらした方向に、空間フィルタ 4 2 2 及び

50

空間フィルタ 4 2 3 の抑圧角度を設定している。音源分離装置 1 0 では、上述の例のように、目的音が到来すると想定される方向を中心として、左右対称の対になるように空間フィルタの抑圧角度を設定することが望ましい。

【 0 0 6 0 】

目的音抑圧部 4 2 では、3つの空間フィルタを用いているが、目的音が到来すると想定される方向（音源分離装置 1 0 では 0 度）を含む所定の範囲内（音源分離装置 1 0 では - 5 度 ~ + 5 度の範囲内）で、複数の空間フィルタにより、それぞれ異なる抑圧角度が向けられていれば、空間フィルタの数やその抑圧角度の組み合わせは限定されないものである。

【 0 0 6 1 】

空間フィルタ 4 2 1 具体的な処理としては、以下の (6) 式を用いて、 $F 0 (m)$ を求める。

【 0 0 6 2 】

空間フィルタ 4 2 2 は、以下の (7) 式を用いて、 $F 1 (m)$ を求める。なお、(7) 式において、 τ_5 は抑圧角度 = + 5 度に相当する遅延である。

【 0 0 6 3 】

空間フィルタ 4 2 3 は、以下の (8) 式を用いて、 $F 2 (m)$ を求める。なお、(8) 式において、 τ_{-5} は抑圧角度 = - 5 度に相当する遅延である。

【 0 0 6 4 】

そして、最小選択部 4 2 4 は、以下の (9) 式に示すように、 $F 0 (m)$ 、 $F 1 (m)$ 、 $F 2 (m)$ の絶対値の最小値 $N (m)$ を算出する。この出力 $N (m)$ は、妨害音の成分を抽出したものととして、最小選択部 4 2 4 から除去手段 5 0 に与えられる。

【数 4】

$$F0(m) = D1(m) - D2(m) \quad \dots (6)$$

$$F1(m) = D2(m) - D1(m) \times e^{\frac{-j2\pi mf}{N} \tau_5} \quad \dots (7)$$

$$F2(m) = D2(m) - D1(m) \times e^{\frac{-j2\pi mf}{N} \tau_{-5}} \quad \dots (8)$$

$$N(m) = \min(\|F0(m)\|, \|F1(m)\|, \|F2(m)\|) \quad (\text{ただし } \|\cdot\| \text{ は絶対値を表す}) \dots (9)$$

【 0 0 6 5 】

次に、除去手段 5 0 の構成について説明する。

【 0 0 6 6 】

除去手段 5 0 は、分離手段 4 0 から与えられる $M (m)$ と $N (m)$ とを用いて、 $D 1 (m)$ における、妨害音を除去するための妨害音除去スペクトル $H (m)$ を求めて、生成手段 6 0 に与える。

【 0 0 6 7 】

以下では、除去手段 5 0 が求める妨害音除去スペクトル $H (m)$ の一例について説明する。

【 0 0 6 8 】

除去手段 5 0 では最小選択部 4 1 3 の出力 $M (m)$ と最小選択部 4 2 4 の出力 $N (m)$ から、以下の (1 0) 式を用いて、 $S (m)$ を求める。さらに、除去手段 5 0 は、 $0 \leq m \leq N / 2$ の範囲で求められた $S (m)$ に対して、以下の (1 1) 式を用いて、除去手段 5 0 の出力である妨害音除去スペクトル $H (m)$ を求める。なお、(1 0) 式、(1 1) 式において、 $D 1$ を $D 2$ に置き換えるようにしても良い。

10

20

30

40

【数5】

$$S(m) = \begin{cases} 0 & M(m) < \|N(m)\| \text{ のとき} \\ \frac{\sqrt{M(m) - N(m)}}{\|D1(m)\|} & M(m) \geq \|N(m)\| \text{ のとき} \end{cases} \quad \dots (10)$$

(ただし $\|\cdot\|$ は絶対値を表す)

10

【0069】

$$H(m) = S(m) D1(m) \quad \dots (11)$$

また、除去手段50は、 $H(m) = H^*(N - m)$ (ただし、 $N/2 + 1 \leq m \leq N - 1$) の性質を利用して、 $0 \leq m \leq N - 1$ の範囲の妨害音除去スペクトル $H(m)$ を求め、生成手段60に与える。

【0070】

生成手段60は、妨害音除去スペクトル $H(m)$ を N 点逆FFT処理し、音源分離信号 $h(n)$ を得る。そして、生成手段60は、以下の(12)式に示すように、現在の音源分離信号 $h(n)$ と、直前の分析単位についての音源分離信号 $h'(n)$ の後半の $3N/4$ 個のデータを加算して、出力 $y(n)$ を得る。

20

【0071】

$$y(n) = h(n) + h'(n + N/4) \quad \dots (12)$$

音源分離装置10では、相前後する分析単位でデータ(サンプル)を重複させるように、 $N/4$ 個のデータをシフトしながら、上述した処理を行う例について説明したが、これは、波形接続を円滑に行うためであるため、必ずしも必要な処理ではなく、 N 個ずつ処理するようにしても良い。なお、 $N/4$ 個のデータをシフトしながら処理する場合、1つの分析単位に対し、分析手段30から当該生成手段60までの上述した一連の処理に要する時間は、 $NT/4$ を上限とすることが望ましい。

【0072】

(A-2) 第1の実施形態の効果

30

第1の実施形態によれば、以下のような効果を奏することができる。

【0073】

音源分離装置10では、目的音抑圧部42の、3つの空間フィルタに、それぞれ0度、5度、-5度の指向性を持たせており、最小選択部424では、3つの空間フィルタの出力値のうち最も小さい絶対値となる出力値を $N(m)$ に適用している。すなわち、目的音抑圧部42では、目的音が0度方向の近辺から到来した場合には、0度方向の近辺の成分については、空間フィルタ421の出力値の絶対値が最も小さいため、これが $N(m)$ に反映される。一方、5度方向の近辺のから目的音が到来した場合には、5度方向の近辺の成分については、空間フィルタ422の出力値が $N(m)$ に反映されることになる。このように、目的音抑圧部42では、目的音の到来方向に応じて、選択して適用される空間フィルタ群を設けることにより、目的音が到来する方向が少しずれた場合でも、 $N(m)$ に目的音成分が混入することを防ぎ、音源分離装置10が出力する音質劣化を防いでいる。

40

【0074】

したがって、上述のように、目的音の到来方向に応じて選択して適用される空間フィルタ群を用いて目的音抑圧部42を構成することにより、目的音の到来方向がずれた場合でも、分離後の目的音の音声の品質を向上させ、聞きやすくすることができる。

【0075】

(B) 第2の実施形態

以下、本発明による音源分離装置、プログラム及び方法の第2の実施形態を、図面を参照しながら詳述する。

50

【 0 0 7 6 】

(B - 1) 第 2 の実施形態の構成及び動作

図 2 は、第 2 の実施形態の音源分離装置 1 0 A の全体構成を示すブロック図である。

【 0 0 7 7 】

第 1 の実施形態の音源分離装置 1 0 では、入力手段 2 0 、分析手段 3 0 、分離手段 4 0 をそれぞれ一つずつ有する構成であったが、第 2 の実施形態の音源分離装置 1 0 A では、入力手段 2 0 、分析手段 3 0 、分離手段 4 0 の組を、複数組備える点で異なっている。また、第 2 の実施形態の音源分離装置 1 0 A では、除去手段 5 0 が除去手段 5 0 A に置き換わっている点で、第 1 の実施形態と異なっている。

【 0 0 7 8 】

音源分離装置 1 0 A では、図 2 に示すように、入力手段 2 0 、分析手段 3 0 、分離手段 4 0 の組を 2 組有する。すなわち、2 つの入力手段 2 0 (2 0 - 1 、 2 0 - 2) 、 2 つの分析手段 3 0 (3 0 - 1 、 3 0 - 2) 、及び 2 つの分離手段 4 0 (4 0 - 1 、 4 0 - 2) を有している。また、入力手段 2 0 - 1 は、2 つのマイクロフォン 2 1 - 1 、 2 2 - 1 を有し、入力手段 2 0 - 2 も 2 つのマイクロフォン 2 1 - 2 、 2 2 - 2 を有している。

【 0 0 7 9 】

入力手段 2 0 - 1 、 2 0 - 2 、分析手段 3 0 - 1 、 3 0 - 2 、分離手段 4 0 - 1 、 4 0 - 2 のそれぞれの処理については、第 1 の実施形態の入力手段 2 0 、分析手段 3 0 、分離手段 4 0 と同様であるので詳しい説明は省略する。

【 0 0 8 0 】

また、以下では、分離手段 4 0 - 1 における、妨害音抑圧部の出力を $MA(m)$ 、目的音抑圧部の出力を $NA(m)$ と表す。また、分離手段 4 0 - 2 における、妨害音抑圧部の出力を $MB(m)$ 、目的音抑圧部の出力を $NB(m)$ と表す。また、マイクロフォン 2 1 - 1 からの信号を、分析手段 3 0 - 1 で処理したものを $D1(m)$ と表すものとする。

【 0 0 8 1 】

次に、除去手段 5 0 A の構成について説明する。

【 0 0 8 2 】

除去手段 5 0 A は、分離手段 4 0 - 1 から与えられる $MA(m)$ 、 $NA(m)$ と、4 0 - 2 から与えられる、 $MB(m)$ 、 $NB(m)$ とを用いて、 $D1(m)$ における、妨害音を除去するための妨害音除去スペクトル $H(m)$ を求めて、生成手段 6 0 に与える。

【 0 0 8 3 】

以下では、除去手段 5 0 A が求める妨害音除去スペクトル $H(m)$ の一例について説明する。

【 0 0 8 4 】

除去手段 5 0 A は、分離手段 4 0 - 1 から与えられる $MA(m)$ 、 $NA(m)$ と、4 0 - 2 から与えられる、 $MB(m)$ 、 $NB(m)$ とを、以下の (1 3) 式に適用し、 $S(m)$ を求める。さらに、除去手段 5 0 A は、 $0 \leq m \leq N/2$ の範囲で求められた $S(m)$ に対して、以下の (1 4) 式を用いて、除去手段 5 0 A の出力である妨害音除去スペクトル $H(m)$ を求める。なお、(1 3) 式、(1 4) 式において、 $D1$ を、他のマイクロフォンからの信号に基づくスペクトルに置き換えるようにしても良い。

10

20

30

40

【数6】

$$S(m) = \begin{cases} \frac{\sqrt{MB(m) - NB(m)}}{\|D1(m)\|} & MA(m) \geq MB(m) \text{かつ} MB(m) \geq \|NB(m)\| \text{のとき} \\ \frac{\sqrt{MA(m) - NA(m)}}{\|D1(m)\|} & MB(m) \geq MA(m) \text{かつ} MA(m) \geq \|NA(m)\| \text{のとき} \quad \dots (13) \\ 0 & \text{そのほか} \end{cases}$$

10

(ただし $\|\cdot\|$ は絶対値を表す)

【0085】

$$H(m) = S(m) D1(m) \dots (14)$$

また、 $H(m) = H^*(N - m)$ (ただし、 $N/2 + 1 \leq m \leq N - 1$) の性質を利用して、 $0 \leq m \leq N - 1$ の範囲の妨害音除去スペクトル $H(m)$ を求め、生成手段60に与える。

【0086】

生成手段60の処理については、第1の実施形態と同様であるので説明を省略する。

【0087】

20

(B-2) 第2の実施形態の効果

第2の実施形態の音源分離装置10Aでは、入力手段において、2個よりも多い数のマイクロフォンを用いた場合でも、第1の実施形態と同様の効果を奏することができる。

【0088】

(C) 他の実施形態

本発明は、上記の各実施形態に限定されるものではなく、以下に例示するような変形実施形態も挙げることができる。

【0089】

(C-1) 第1の実施形態において、音源分離装置10の用途によっては、生成手段60を省略したり、他の装置が有する生成部を流用したりすることができる。例えば、音源分離装置が音声認識装置に利用される場合であれば、分離スペクトル $H(m)$ を認識用特徴量として用いるようにして生成手段60を省略することができる。また例えば、音源分離装置がIP電話機に利用される場合であれば、IP電話機が生成部に相当する手段を有するので、その生成部を流用するようにしても良い。

30

【0090】

(C-2) 第2の実施形態において、4つのマイクロフォン21-1、22-1、21-2、22-2を用いる例について説明したが、入力手段20-1と入力手段20-2との間で、1個を共通に使うことにより3個のマイクロフォンによって構成するようにしても良い。この場合、共通に用いるマイクロフォンが受音した信号の処理を共通にすることができるため演算量を低減させることができる。また、用いるマイクロフォンの数をさらに増加させた場合でも、同様に入力手段の間で共通のマイクロフォンを用いるようにしても良い。

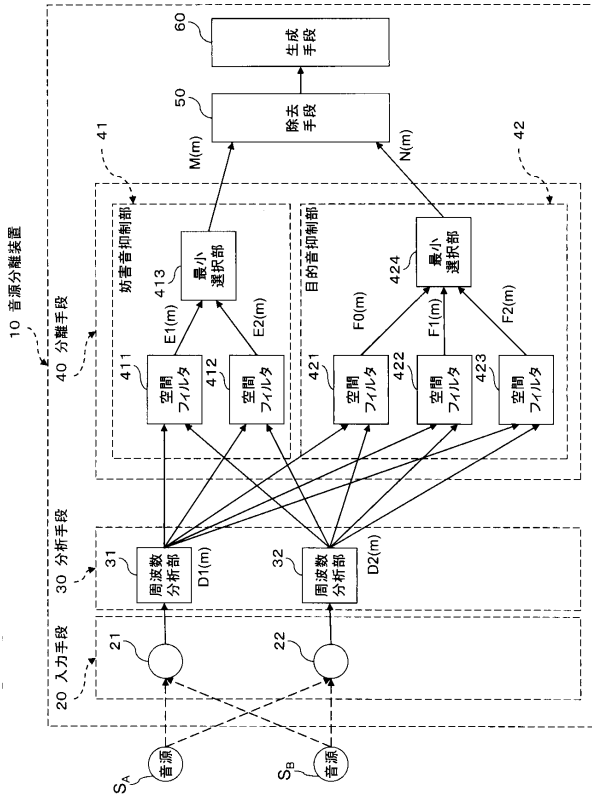
40

【符号の説明】

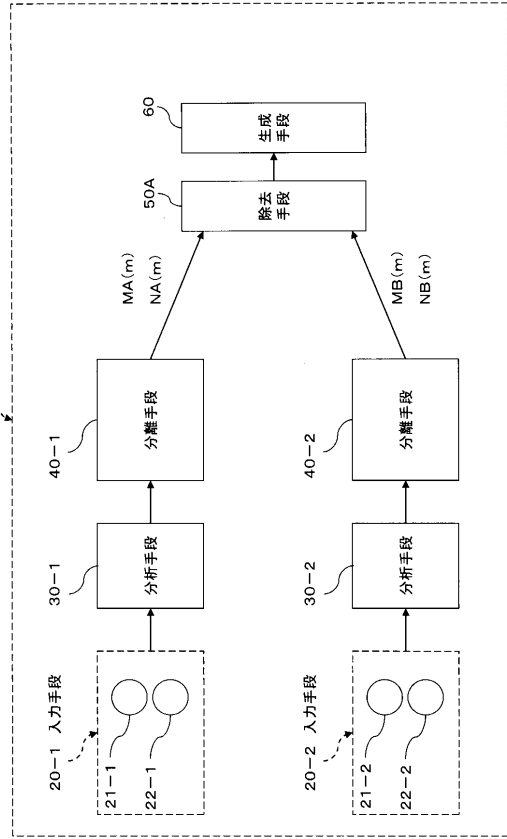
【0091】

10...音源分離装置、20...入力手段、21、22...マイクロフォン、30...分析手段、31、32...周波数分析部、40...分離手段、41...妨害音抑圧部、411、412...空間フィルタ、413...最小選択部、42...目的音抑圧部、421、422、423空間フィルタ、424...最小選択部、50...除去手段、60...生成手段。

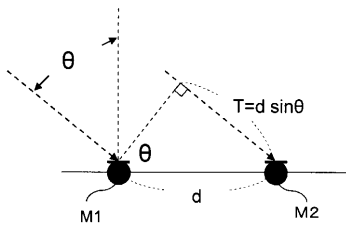
【図1】



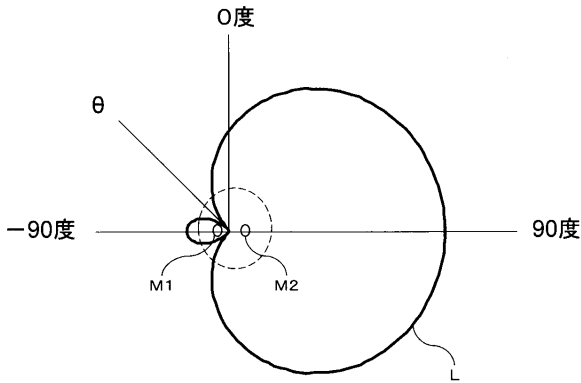
【図2】



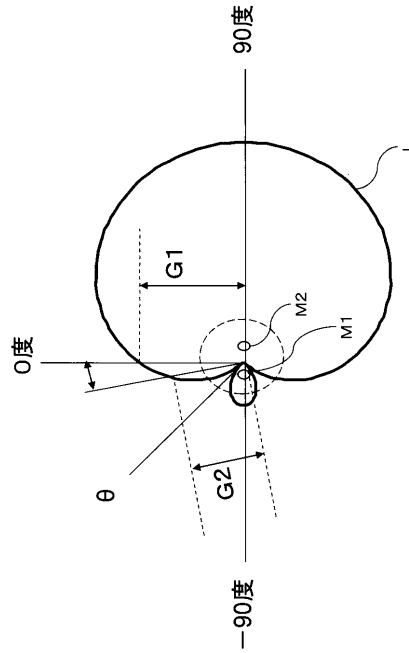
【図3】



【図4】



【図5】



フロントページの続き

- (72)発明者 矢頭 隆
東京都港区西新橋三丁目16番11号 沖電気工業株式会社内
- (72)発明者 山田 圭
東京都港区西新橋三丁目16番11号 沖電気工業株式会社内
- (72)発明者 小林 哲則
東京都新宿区大久保3丁目4番1号 学校法人 早稲田大学 理工学部内
- (72)発明者 赤桐 健三
東京都新宿区大久保3丁目4番1号 学校法人 早稲田大学 理工学部内
- (72)発明者 小川 哲司
東京都新宿区大久保3丁目4番1号 学校法人 早稲田大学 理工学部内

審査官 山下 剛史

- (56)参考文献 特開2006-197552(JP, A)
国際公開第2007/052604(WO, A1)
特開平11-41687(JP, A)
特開2008-164747(JP, A)
特開2007-13400(JP, A)
特開平4-212600(JP, A)

(58)調査した分野(Int.Cl., DB名)

G10L 15/20, 21/02
H04R 1/40, 3/00