

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5123595号
(P5123595)

(45) 発行日 平成25年1月23日(2013.1.23)

(24) 登録日 平成24年11月2日(2012.11.2)

(51) Int.Cl.

F I

G 1 O L 21/0208 (2013.01)	G 1 O L 21/02 1 O 2 B
H O 4 R 3/00 (2006.01)	H O 4 R 3/00 3 2 O
G 1 O L 21/0308 (2013.01)	G 1 O L 21/02 2 O 3 Z
G 1 O L 21/028 (2013.01)	G 1 O L 21/02 2 O 1 C
H O 4 R 1/40 (2006.01)	H O 4 R 1/40 3 2 O A

請求項の数 4 (全 24 頁)

(21) 出願番号 特願2007-198664 (P2007-198664)
 (22) 出願日 平成19年7月31日(2007.7.31)
 (65) 公開番号 特開2009-36810 (P2009-36810A)
 (43) 公開日 平成21年2月19日(2009.2.19)
 審査請求日 平成22年7月29日(2010.7.29)

(73) 特許権者 301022471
 独立行政法人情報通信研究機構
 東京都小金井市貫井北町4-2-1
 (74) 代理人 100099933
 弁理士 清水 敏
 (72) 発明者 西村 電一
 東京都小金井市貫井北町4-2-1 独立
 行政法人情報通信研究機構内
 (72) 発明者 井ノ上 直己
 東京都小金井市貫井北町4-2-1 独立
 行政法人情報通信研究機構内
 審査官 毛利 太郎

最終頁に続く

(54) 【発明の名称】 近傍場音源分離プログラム、及びこのプログラムを記録したコンピュータ読取可能な記録媒体、並びに近傍場音源分離方法

(57) 【特許請求の範囲】

【請求項1】

2次元平面上において一の音源に対して所定の間隔を隔てて配置された第1及び第2のマイクロホンからの音源観測信号間の振幅比に基づいて近傍場音源を分離するために、コンピュータを、

強調したい音源からの距離が入力されると、これに応じた重み関数を複数の重み関数の中から選択するための重み関数選択手段として機能させる近傍場音源分離プログラムであって、当該複数の重み関数は、各音源距離に対するSN比と振幅比との関係を距離減衰に基づき予め計算することによりデータベース化されており、

前記近傍場音源分離プログラムはさらに、前記コンピュータを、

前記第2のマイクロホンからの音源観測信号の振幅に対する前記第1のマイクロホンからの音源観測信号の振幅比を周波数成分ごとに求め、当該振幅比から、周波数成分ごとの重み係数を前記重み関数選択手段により選択された重み関数により求め、前記第1のマイクロホンの出力を周波数成分ごとに前記重み係数により重み付けし、得られた結果を時間領域に変換することにより、近傍場音の強調を行なうための近傍場音強調手段、として機能させる近傍場音源分離プログラム。

【請求項2】

前記近傍場音強調手段は、

入力された各マイクロホンの音源観測信号を周波数領域に変換する第1の線形変換を行なうための第1の手段、

10

20

前記第 1 の手段による第 1 の線形変換で得られた結果を基にして各マイクロホンからの音源観測信号の振幅を周波数成分ごとに抽出するための第 2 の手段、

前記第 2 の手段により抽出された、前記第 2 のマイクロホンの音源観測信号の振幅に対する前記第 1 のマイクロホンの音源観測信号の振幅比を周波数成分ごとに求めるための第 3 の手段、

前記重み関数選択手段により選択された重み関数に基づいて、前記第 3 の手段により求められたマイクロホン間の周波数成分ごとの振幅比に応じた重み係数を求めることにより、前記第 1 のマイクロホンの出力の周波数成分ごとの重みを求めるための第 4 の手段、

前記第 1 の手段による第 1 の線形変換で得られた、前記第 1 のマイクロホン側の結果と前記第 4 の手段で求められた重みとを周波数成分ごとに掛け合わせるための第 5 の手段、及び

前記第 5 の手段により周波数成分ごとに得られた演算結果に対して前記第 1 の線形変換の逆変換に対応する第 2 の線形変換を行なって時間領域に変換するための第 6 の手段を含む、請求項 1 に記載の近傍場音源分離プログラム。

【請求項 3】

2次元平面上において一の音源に対して所定の間隔を隔てて配置された第 1 及び第 2 のマイクロホンからの音源観測信号間の振幅比に基づいて近傍場音源を分離するために、コンピュータを、

強調したい音源からの距離が入力されると、これに応じた重み関数を複数の重み関数の中から選択するための重み関数選択手段として機能させる近傍場音源分離プログラムを記録したコンピュータ読取可能な記録媒体であって、当該複数の重み関数は、各音源に対する SN 比と振幅比との関係を距離減衰に基づき予め計算することによりデータベース化されており、

前記近傍場音源分離プログラムはさらに、前記コンピュータを、

前記第 2 のマイクロホンからの音源観測信号の振幅に対する前記第 1 のマイクロホンからの音源観測信号の振幅比を周波数成分ごとに求め、当該振幅比から、周波数成分ごとの重み係数を前記重み関数選択手段により選択された重み関数により求め、前記第 1 のマイクロホンの出力を周波数成分ごとに前記重み係数により重み付けし、得られた結果を時間領域に変換することにより、近傍場音の強調を行なうための近傍場音強調手段、として機能させる近傍場音源分離プログラムを記録したコンピュータ読取可能な記録媒体。

【請求項 4】

2次元平面上において一の音源に対して所定の間隔を隔てて配置された第 1 及び第 2 のマイクロホンからの音源観測信号間の振幅比に基づいて近傍場音源を分離するための方法であって、

強調したい音源からの距離が入力されると、これに応じた重み関数を複数の重み関数の中から選択するための第 1 のステップを含み、当該複数の重み関数は、各音源距離に対する SN 比と振幅比との関係を距離減衰に基づき予め計算することによりデータベース化されており、

前記第 2 のマイクロホンからの音源観測信号の振幅に対する前記第 1 のマイクロホンからの音源観測信号の振幅比を周波数成分ごとに求め、当該振幅比から、周波数成分ごとの重み係数を前記第 1 のステップにおいて選択された重み関数により求め、前記第 1 のマイクロホンの出力を周波数成分ごとに前記重み係数により重み付けし、得られた結果を時間領域に変換することにより、近傍場音の強調を行なうための第 2 のステップと、を含む近傍場音源分離方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、マイクロホン間の振幅比に基づいて近傍場音源を分離するための、プログラム、及び同プログラムを記録したコンピュータ読取可能な記録媒体、並びに方法に関する。

10

20

30

40

50

【背景技術】

【0002】

情報化社会においては、より多くの情報を収集し、それを分析・解析することが新しい価値及び情報を生み出すと言われている。例えば、相手の現在の状況及び表情、仕草等から相手の心情を理解することができれば、より円滑なコミュニケーションの実現が可能となる。しかし、現在の遠隔通信では、通信帯域の制約により、音響情報だけを取ってみても、「ことば」を伝達することに主眼が置かれ、声色又は話者の置かれている環境の音等、他の要素の伝達を犠牲にしている面が少なくない。

【0003】

これに対し、「超臨場感通信」では、上記の補助情報も伝達することで、より高次のコミュニケーションの実現を目指している。その一例を挙げると、動作には音響的イベントを伴うことが多いため、相手の仕草及び状況は視覚ばかりではなく、音響的にも伝達することが可能である。これらの音は一般的にエネルギーが小さい。そのため、音源の近傍でしか聴取できないが、逆にそれらが聞こえることで、臨場感を高められると期待できる。

【0004】

ところが、実際の環境下では、エネルギーの大きな数々の外来音が重畳するため、上記の音のみを收音することは非常に困難なタスクとなる。

【0005】

また、頭部伝達関数(HRTF: Head Related Transfer Function)に基づく立体音合成技術(非特許文献1及び非特許文献2参照)では、音源が頭部近傍に存在するときと遠方に存在するときとで、用いるべきHRTFを違える必要があると予想される。これは、音源が近傍に存在する場合には、音源が遠方にある場合に比べて、頭部又は肩における反射及び回折による影響が大きくなることに依存している。したがって、臨場感通信における受信側において、上記の頭部伝達関数に基づく立体音響技術を用いて臨場感音場再生を行なうには、送信側においても近傍音と遠方音とを分離して收音する必要がある。

【0006】

この目的を実現するには、マイクロホンアレイを用いたアレイ信号処理による音源分離手法が有力である。最も基本的な構造の1つは、delay-and-sumに基づく手法である。ただし、この手法は、マイクロホンアレイへの入射角度に応じてゲイン特性が変化するフィルタを形成するものであって、音源の距離に応じて変化するものでは必ずしもない。

【0007】

一方で、非常に多くのマイクロホンを用いて、ある程度広い空間の音場全体を收音する試みも行われている(非特許文献3及び非特許文献4参照)。しかし、大掛かりなシステムでは汎用性に欠け、利用できる場面が限られる、という実際的な応用面での問題が考えられる。

【0008】

また、何れの場合でも、位相差に着目した手法は、空間エイリアシングの問題が不可避であり、適用範囲を高い周波数まで広げるためには、マイクロホンを密に配置しなければならない。これにより、システムの規模及び隣接するマイクロホンの影響が新たな問題を生む可能性もある。

【0009】

近傍場音に対するビームフォーミングに関する研究としては、遠方場に対する指向性合成を基にして近傍場用のフィルタを設計する手法が提案されている(非特許文献5参照)。しかし、この手法も、指向性の制御であり、距離に関する分離は基本的に困難である。

【0010】

また、部分空間法に基づいて近傍場の音源分離を行なう手法も提案されている(非特許文献6参照)。この手法では、近傍場に存在する音源の位置を部分空間法により推定し、その推定した位置に基づいて空間逆フィルタを設計することで分離を行なうものである。

10

20

30

40

50

【 0 0 1 1 】

ここで、近傍場であるという条件は、部分空間法により音源方向に基づいて処理する（非特許文献7参照）ばかりでなく、距離に関する情報も得るために利用されている。

【 0 0 1 2 】

一方、音源位置に関する先験的知識を一切使用せずに音源信号の性質のみに着目した分離手法として、独立成分分析（ICA）に基づく手法が高い性能を示している（非特許文献8～非特許文献10）。この分析手法は、音源信号の独立性が規範となるため、複数話者の混合音の分離等では有力であるが、同一話者を音源とする反射音のある空間での分離、及び衣擦れ音又はキーボードのタイプ音のように音源自体の性質が白色性の信号の分離では、必ずしも高い性能が期待できない。

10

【 0 0 1 3 】

【非特許文献1】E. M. ウェンゼル、J. D. ミラー及びJ. S. アベル、「対話型空間音声合成のためのソフトウェアベースシステム」ICAD 2000 予稿集（2000年）。（E. M. Wenzel, J. D. Miller and J. S. Abel: "A software-based system for interactive spatial sound synthesis", In proc. of ICAD 2000 (2000).）

【非特許文献2】L. サビオラ、J. フオパニエミ、T. ロッキ及びR. ヴェナーネン、「対話型仮想音響環境の生成」音声工学学会誌、47、9、675-705ページ（1999年）。（L. Savioja, J. Huopaniemi, T. Lokki and R. V a a n a n e n: "Creating interactive virtual acoustic environments", J. Audio En 20 g. Soc., 47, 9, pp. 675-705 (1999).）

【非特許文献3】M. S. ブランドシュタイン及びD. B. ワード、「マイクロホンアレイでの音声獲得のためのセルベースのビーム形成（ce-base）」IEEE 音声及び音響処理トランザクション、8、6、738-743ページ（2000年）。（M. S. Brandstein and D. B. Ward: "Cell-based beam forming (ce-base) for speech acquisition with microphone arrays", IEEE Trans. on Speech and Audio Processing, 8, 6, pp.738-743 (2000).）

【非特許文献4】T. オカモト、R. ニシムラ及びY. イワヤ、「周囲のマイクロホンアレイを用いた音源位置の推定」、音響科学及び技術、28、3、181-189ページ（2007年）。（T. Okamoto, R. Nishimura and Y. Iwaya: "Estimation of 30 sound source positions using a surrounding microphone array", Acoust. Sci. & Tech., 28, 3, pp. 181-189 (2007).）

【非特許文献5】R. A. ケネディ、D. B. ワード及びT. D. アブハヤパラ、「ラジアル相互関係を用いた近傍場ビーム形成」IEEE 信号処理トランザクション、47、1、33-40ページ（1999年）。（R. A. Kennedy, D. B. Ward and T. D. Abhayapala: "Near field beam forming using radial reciprocity", IEEE Trans. on Signal Processing, 47, 1, pp. 33-40 (1999).）

【非特許文献6】F. アサノ、H. アソー及びT. マツイ、「近傍場における音源の局所化及び分離」IEICE 基礎工学トランザクション、83、11、2286-2294ページ（2000年）。（F. Asano, H. Asoh and T. Matsui: "Sound source l 40 ocalization and separation in near field", IEICE Trans. on Fundamentals, 83, 11, pp. 2286-2294 (2000).）

【非特許文献7】F. アサノ、S. ハヤミズ、T. ヤマダ及びS. ナカムラ、「サブ空間法に基づく音声の向上」IEEE 音声及び音響処理トランザクション、8、5、497-507ページ（2000年）。（F. Asano, S. Hayamizu, T. Yamada and S. Nakamura: "Speech enhancement based on the subspace method", IEEE Trans. on Speech and Audio Processing, 8, 5, pp.497-507 (2000).）

【非特許文献8】J. F. カルドーソ及びA. ソウルオウミアク、「非ガウス信号のためのブラインドビーム形成」IEEE 予稿集、140、6、362-370ページ（1993年）。（J. F. Cardoso and A. Souloumiac: "Blind beam forming for non- 50

gaussian signals”, IEE Proceedings, 140, 6, pp. 362 - 370 (1993).)

【非特許文献9】F. エーラース及びH. G. シュスター、「畳込み混合のブラインド分離と雑音環境での自動音声認識における応用」IEEE 信号処理トランザクション、45、10、2608 - 2612 ページ(1997年)。(F. Ehlers and H. G. Schuster: “Blind separation of convolutive mixtures and an application in automatic speech recognition in a noisy environment”, IEEE Trans. on Signal Processing, 45, 10, pp. 2608-2612 (1997).)

【非特許文献10】H. サワダ、R. ムカイ、S. アラキ及びS. マキノ、「周波数ドメインのブラインド音源分離における置換問題の頑健かつ精密な解決方法」、IEEE 音声及び音響処理トランザクション、12、5、530 - 538 ページ(2004年)。(H. Sawada, R. Mukai, S. Araki and S. Makino: “A robust and precise method for solving the permutation problem of frequency-domain blind source separation”, IEEE Trans. on Speech and Audio Processing, 12, 5, pp. 530-538 (2004).)

【発明の開示】

【発明が解決しようとする課題】

【0014】

本発明は、上記技術的課題に鑑みなされたもので、室伝達関数の推定とその逆フィルタを用いることもなく、遠方音を抑圧して近傍音のみを強調し得る、近傍場音源プログラム、及びこのプログラムを記録したコンピュータ読取可能な記録媒体、並びに近傍場音源分離方法の提供を目的とする。

【課題を解決するための手段】

【0015】

上記目的を達成するため、本発明者等は、物理的な性質として、音源からの距離が遠くなるに従って単位距離当りの振幅(エネルギー)の減衰率が低下することを音源分離に利用すればよいのではないかと着想した。

【0016】

かかる着想に基づく具体的な発明は、以下の通りである。

【0017】

本発明を第1の局面から観れば、2次元平面上において一の音源に対して所定の間隔を隔てて配置された第1及び第2のマイクロホンからの音源観測信号間の振幅比に基づいて近傍場音源を分離するために、コンピュータを、強調したい音源からの距離が入力されると、これに応じた重み関数を複数の重み関数の中から選択するための重み関数選択手段として機能させる近傍場音源分離プログラムである。複数の重み関数は、各音源距離に対するSN比と振幅比との関係を、距離減衰に基づき予め計算することによりデータベース化されている。この近傍場音源分離プログラムは、コンピュータをさらに、第2のマイクロホンからの音源信号の振幅に対する第1のマイクロホンからの音源観測信号の振幅比を周波数成分ごとに求め、当該振幅比から、周波数成分ごとの重み係数を重み関数選択手段により選択された重み関数により求め、第1のマイクロホンの出力を周波数成分ごとにこの重み係数により重み付けし、得られた結果を時間領域に変換することにより、近傍場音の強調を行なうための近傍場音強調手段、として機能させる。

【0018】

上記近傍場音源分離プログラムにおいて、近傍場音強調手段は、入力された各マイクロホンの音源観測信号を周波数領域に変換する第1の線形変換を行なうための第1の手段、第1の手段による第1の線形変換で得られた結果を基にして各マイクロホンからの音源観測信号の振幅を周波数成分ごとに抽出するための第2の手段、第2の手段により抽出された、第2のマイクロホンの音源観測信号の振幅に対する第1のマイクロホンの音源観測信号の振幅比を周波数成分ごとに求めるための第3の手段、重み関数選択手段により選択された重み関数に基づいて、第3の手段により求められたマイクロホン間の周波数成分ごとの重みを求めるための第4の手段、第1の手段による第1の線形変換で得られた、第1の

10

20

30

40

50

マイクロホン側の結果と第4の手段で求められた重みとを周波数成分ごとに掛け合わせるための第5の手段、及び第5の手段により周波数成分ごとに得られた演算結果に対して第1の線形変換の逆変換に対応する第2の線形変換を行なって時間領域に変換するための第6の手段を含む。

【0019】

本発明を第2の局面から観れば、2次元平面上において一の音源に対して所定の間隔を隔てて配置された第1及び第2のマイクロホンからの音源観測信号間の振幅比に基づいて近傍場音源を分離するために、コンピュータを、強調したい音源からの距離が入力されると、これに応じた重み関数を複数の重み関数の中から選択するための重み関数選択手段として機能させる近傍場音源分離プログラムを記録したコンピュータ読取可能な記録媒体である。当該複数の重み関数は、各音源距離に対するSN比と振幅比との関係を距離減衰に基づき予め計算することによりデータベース化されている。近傍場音源分離プログラムはさらに、コンピュータを、第2のマイクロホンからの音源観測信号の振幅に対する第1のマイクロホンからの音源観測信号の振幅比を周波数成分ごとに求め、当該振幅比から、周波数成分ごとの重み係数を重み関数選択手段により選択された重み関数により求め、第1のマイクロホンの出力をこれら重み係数により周波数成分ごとに重み付けし、得られた結果を時間領域に変換することにより、近傍場音の強調を行なうための近傍場音強調手段、として機能させる。

10

【0020】

本発明を第3の局面から観れば、2次元平面上において一の音源に対して所定の間隔を隔てて配置された第1及び第2のマイクロホンからの音源観測信号間の振幅比に基づいて近傍場音源を分離するための方法であって、強調したい音源からの距離が入力されると、これに応じた重み関数を複数の重み関数の中から選択するための第1のステップを含む近傍場音源分離方法である。複数の重み関数は、各音源距離に対するSN比と振幅比との関係を距離減衰に基づき予め計算することによりデータベース化されている。この方法はさらに、第2のマイクロホンからの音源観測信号の振幅に対する第1のマイクロホンからの音源観測信号の振幅比を周波数ごとに求め、当該振幅比から周波数成分ごとの重み係数を第1のステップにおいて選択された重み関数により求め、第1のマイクロホンの出力をこれら重み係数により周波数成分ごとに重み付けし、得られた結果を時間領域に変換することにより、近傍場音の強調を行なうための第2のステップを含む。

20

30

【発明の効果】

【0021】

本発明によると、近傍場音源分離は、マイクロホン間の振幅比に基づいて行なわれるので、室伝達関数の推定とその逆フィルタを用いることなく、遠方音を抑圧して近傍音のみを強調することが可能となる。

【発明を実施するための最良の形態】

【0022】

以下、本発明の実施の形態について、添付図面に基づき詳細に説明する。

【0023】

<概要>

40

頭部伝達関数に基づく仮想立体音響再生技術は、実測又は合成した聴取者の頭部伝達関数を利用している。この関数の形状は、音源が聴取者の近傍にある場合には、頭部又は上半身の反射及び回折により、音源が遠方にある場合のそれとは異なる。したがって、臨場感遠隔通信を行なう際の送信側では、仮想的な聴取者の位置に対して、近傍に存在する音源と遠方に存在する音源を分離する必要がある。

【0024】

そこで、本実施の形態では、少数のマイクロホンを用いて、聴取位置から近傍の音のみを強調する手法の一例を説明する。

【0025】

本実施の形態では、音源から放射された音の振幅が、距離に反比例して減少するため、

50

音源が近傍に存在する場合は2点間の振幅比が大きい、遠方の音源に対しては、同じ距離離れていても振幅比が小さくなることを利用している。

【0026】

ここで、本実施の形態の具体的な構成を説明する前に、まず、本近傍場音源分離に関するアルゴリズム(定式化)について説明する。

【0027】

<アルゴリズム>

球面進行波の音圧の振幅は、音源からの距離に反比例して減少する。そのため、単位距離当りの振幅の減衰量は、音源の近傍では大きい、音源から離れるにつれて急激に減少する。この旨は、「城戸(編): “基礎音響工学”, コロナ社(1990).」にて開示されている。

10

【0028】

また、音声信号は、時間周波数平面上で一様に分布しているわけではなく偏在しており、その偏在の様相も話者又は発話内容によって異なる。したがって、音源からの距離が異なる複数のマイクロホンで信号を観測して、それぞれの信号を時間周波数表現に変換した後、それらと比較すると、振幅差が大きくなる領域は近傍場の音が優勢な領域であると考えられる。そのため、その領域の信号を増幅することで近傍場音の音声強調が実現できる。逆に、振幅差の小さな領域に対しては、信号を小さくすることで、遠方場から到来する音を抑制することが実現できる。

【0029】

20

これまでにも、各領域ごとに、どの方向から到来した音が優勢なのかを推定して音源分離を行なう手法(M. アオキ、M. オカモト、S. アオキ、H. マツイ、T. サクライ及びY. カネダ、「多数のマイクロホンを用いて獲得した入力信号の各周波数成分の入射角推定に基づく音源分離」、音響科学及び工学、22, 2, 149 - 157ページ、2001年(M. Aoki, M. Okamoto, S. Aoki, H. Matsui, T. Sakurai and Y. Kaneda: “Sound source segregation based on estimating incident angle of each frequency component of input signals acquired by multiple microphones”, Acoust. Sci. & Tech., 22, 2, pp. 149-157 (2001).)が提案されている。しかし、本実施の形態は、音の到来方向ではなく、音源までの距離を規範としている点で、この手法とは大きく異なる。

30

【0030】

ここで、2つのマイクロホンを用いる場合を考える。一方のマイクロホンA(第1のマイクロホン)と他方のマイクロホンB(第2のマイクロホン)とで受信する、近傍場から到来する目的信号の複素振幅を S_A , S_B とし、遠方場から到来する妨害音の複素振幅を N_A , N_B とする。このとき、上記2つのマイクロホンA及びBで受信する信号の振幅の比(音源観測信号間の振幅比)Rは、下記式(1)

【0031】

【数1】

$$R = \frac{|S_A + N_A|}{|S_B + N_B|}$$

40

$$= \frac{\sqrt{(S_A + N_A)(S_A + N_A)^*}}{\sqrt{(S_B + N_B)(S_B + N_B)^*}}$$

$$= \frac{\sqrt{|S_A|^2 + |N_A|^2 + 2|S_A||N_A|\cos(\arg S_A - \arg N_A)}}{\sqrt{|S_B|^2 + |N_B|^2 + 2|S_B||N_B|\cos(\arg S_B - \arg N_B)}}$$

(1)

ここで、*は複素共役を表す。

50

【 0 0 3 2 】

目的音は、マイクロホン A , B に近いため、以下の関係を有することが期待できる。

【数 2】

$$|S_A| \gg |N_A|, |S_B| \gg |N_B|$$

【 0 0 3 3 】

上記の関係を式 (1) に適用し、さらにテイラー展開を行なうと、下記のように 2 つの値を「 0 」と見なすことができるので、

【数 3】

$$(|N_A|/|S_A|)^2 \simeq 0, (|N_B|/|S_B|)^2 \simeq 0$$

10

【 0 0 3 4 】

式 (1) は、次の式 (2) のように書き直せる。

【数 4】

$$\begin{aligned} & \frac{\sqrt{|S_A|^2 + |N_A|^2 + 2|S_A||N_A| \cos(\arg S_A - \arg N_A)}}{\sqrt{|S_B|^2 + |N_B|^2 + 2|S_B||N_B| \cos(\arg S_B - \arg N_B)}} \\ & \simeq \frac{|S_A| \sqrt{1 + 2|N_A|/|S_A| \cos(\arg S_A - \arg N_A)}}{|S_B| \sqrt{1 + 2|N_B|/|S_B| \cos(\arg S_B - \arg N_B)}} \\ & \simeq \frac{|S_A| + |N_A| \cos(\arg S_A - \arg N_A)}{|S_B| + |N_B| \cos(\arg S_B - \arg N_B)} \end{aligned} \quad (2)$$

20

【 0 0 3 5 】

妨害音は、遠方場から到来する信号であることから、2 つのマイクロホン A , B での受信信号の振幅差はあまり生じず、下記関係を有することが期待できる。

【数 5】

$$|N_A| \simeq |N_B|$$

【 0 0 3 6 】

そこで、これらを改めて下記の形式で表記すると、

【外 1】

【 0 0 3 7 】

$$|N|$$

式 (2) は下記の式 (3) で書き表すことができる。

【数 6】

$$R \simeq \frac{|S_A| + |N| \cos(\arg S_A - \arg N_A)}{|S_B| + |N| \cos(\arg S_B - \arg N_B)} \quad (3)$$

40

【 0 0 3 8 】

次に、この振幅比 R の取り得る値の最小値 R_{\min} と最大値 R_{\max} とが、SN 比 (Signal - to - Noise Ratio) によってどのように変化するのかに着目すると、これらの値は、下記の振幅の両方の値に依存する。

【外 2】

【 0 0 3 9 】

$$|S_A| \text{ と } |S_B|$$

【 0 0 4 0 】

50

そのため、ここでは上記の両振幅の比（音源距離に対する振幅比） C をパラメータとして検討することにする。

【0041】

他方のマイクロホンBでの振幅と雑音の比を下記のように表記すると、

【数8】

$$D = |S_B|/|N|$$

式(3)は、式(4)と書き表せる。

【0042】

【数9】

$$R \simeq \frac{C \cdot D + \cos(\arg S_A - \arg N_A)}{D + \cos(\arg S_B - \arg N_B)} \quad (4)$$

【0043】

同様に、SN比が低い場合の振幅比Rの近似式は、式(5)で表される。

【数10】

$$R \simeq \frac{1 + C \cdot D \cos(\arg S_A - \arg N_A)}{1 + D \cos(\arg S_B - \arg N_B)} \quad (5)$$

【0044】

したがって、振幅比Rの最大値 R_{\max} と最小値 R_{\min} は、 $D \geq 1$ のときには式(4)を用い、 $D < 1$ のときには式(5)を用いて条件分けされ、以下のようにまとめられる。

【数11】

$$R_{\max} = \begin{cases} \frac{C \cdot D + 1}{D - 1} & D \geq 1 \\ -\frac{C \cdot D + 1}{D - 1} & D < 1 \end{cases} \quad (6)$$

$$R_{\min} = \begin{cases} \frac{C \cdot D - 1}{D + 1} & \text{for } C \geq 1 & \text{for } C < 1 \\ 0 & 1 > D \geq 1/C & 1/C > D \geq 1 \\ -\frac{C \cdot D - 1}{D + 1} & 1/C > D & 1 > D \end{cases} \quad (7)$$

【0045】

パラメータCは、目的音源までの距離(と角度)を与えると一意に定まるので、上記の関係式は、振幅比Rの値からその成分のSN比(D)の範囲が求まることを意味している。

【0046】

そこで、上記のDの範囲の代表値をSN比の推定値とし、それに応じて重み付けを行なうことで近傍場音の強調を行なうのが、本実施の形態の基本的な考え方である。

【0047】

以下に、この近傍場音源分離の基本的な考え方を、具体例を挙げて説明する。

【0048】

10

20

30

40

50

2つのマイクロホンA, B間における振幅の比Cが、2つのマイクロホンA, B間距離と各マイクロホンA及びBと音源との距離とに対して、どのような振る舞いを示すのかを調査し、その計算シミュレーションの結果を図5に示す。

【0049】

同図は、音源が2つのマイクロホンA, Bを結ぶ直線上に位置すると仮定し、マイクロホンA, B間距離をdとし、音源に近い方のマイクロホンまでの距離をrとしたときの振幅比、つまり、 $(r+d)/r$ を図示したものである。

【0050】

もし、基準とするマイクロホンを逆にとると、図5では対数表現にしているため、符号が逆転するだけで同じ形状のグラフが描かれる。また、rの最大値 R_{max} と最小値 R_{min} とSN比(D)の関係を、幾つかのCの値(妨害音が無いときの2つのマイクロホンA, Bの目的音の振幅比)に対して、上記の式(6)及び(7)を基にして図示したものが図6である。

【0051】

同図では、最小値 R_{min} が不連続になっているが、これは、 $D=1$ を境にして使用する式を違えていることに起因する。また、式(6)及び(7)は、近似式として得られたものであるので、SN比の絶対値が大きくなる以外では、必ずしも厳密には成立しない。

【0052】

図6を基にして、振幅比Rからその成分のSN比を推定するために、ここでは、SN比に関する事前情報は一切得られていないと仮定し、与えられた振幅比Rに対して、そのときに起こり得るSN比が図6で示された上限と下限との間で等確率に出現するものとする。このときの期待値

【外3】

【0053】

$$\hat{D}$$

は、振幅比Rの最大値 R_{max} と最小値 R_{min} の中間値となる。したがって、観測された振幅比Rに対するSN比の推定値

【外4】

【0054】

$$\hat{D}$$

は、上記のCの値によって異なり、表1に示す式で与えられる。また、この式を幾つかのCの値に対して図示すると図7を得る。

【0055】

10

20

30

【表 1】
表 1

\hat{D}	$C \geq 3$	$3 > C \geq 1$	$1 > C \geq 1/3$	$1/3 > C$
$\sqrt{\frac{R^2 - 1}{R^2 - C^2}}$	$R \geq C$	$R \geq C,$ $1 > R \geq \frac{C-1}{2}$	$R \geq 1,$ $C > R \geq -\frac{C-1}{2}$	$R \geq 1$
$\sqrt{\frac{1 - R^2}{R^2 - C^2}}$	$C > R \geq \frac{C-1}{2}$	$C > R \geq 1$	$1 > R \geq C$	$1 > R \geq -\frac{C-1}{2}$
$\sqrt{\frac{R-1}{R+C}}$	$\frac{C-1}{2} > R \geq 1$			
$\sqrt{\frac{-(R-1)}{R+C}}$	$1 > R$	$\frac{C-1}{2} > R$		
$\sqrt{\frac{R+1}{R-C}}$				$-\frac{C-1}{2} > R \geq C$
$\sqrt{\frac{-(R+1)}{R-C}}$			$-\frac{C-1}{2} > R$	$C > R$

10

【 0 0 5 6 】

20

この推定された S N 比

【外 5】

【 0 0 5 7 】

\hat{D}

に基づいて、対象としている成分に重み付けを行なう。

【 0 0 5 8 】

ここで定める重み関数

【外 6】

【 0 0 5 9 】

30

$f(\hat{D})$

は、処理後の音声の明瞭度にも影響を与える。そのため、十分に注意して決定する必要があるが、今回は S N 比の期待値として が得られる可能性があることから、この問題を回避するためにシグモイド関数を基にして、重み関数を式 (8) とする。

【数 1 2】

$$f(\hat{D}) = \frac{1}{1 + e^{-\alpha 10 \log \hat{D}}} \quad (8)$$

【 0 0 6 0 】

40

ここで、 は傾斜を調整するパラメータであり、この の値を「 1 」とした。幾つかのパラメータ C に対する重み関数を、図 8 に図示する。図 8 から分かるように、C を変化させると、重み関数のグラフの形が変化する。マイクロホンから特定の近距離に存在する音源からの信号の推定値

【外 7】

【 0 0 6 1 】

$\hat{s}(n)$

【 0 0 6 2 】

は、マイクロホンで実際に受信した信号 y (n) に対してフーリエ変換を行ない、全ての

50

周波数成分を対象として各周波数成分ごとに重みを計算して掛け合わせ、逆フーリエ変換することで、式(9)のようにして得られる。

【数13】

$$\hat{s}(n) = \text{IFFT}[f(\hat{D}) \cdot \text{FFT}[y(n)]], \quad (9)$$

ここで、 $\text{FFT}[\cdot]$ はフーリエ変換を、
 $\text{IFFT}[\cdot]$ は逆フーリエ変換を表す。

【0063】

<システムの構成>

図1は本発明の実施の形態に係るシステムの構成を示すブロックダイアグラムである。図1を参照して、本実施の形態のシステムは、2つのマイクロホンA、B、近傍場音源分離装置10、ハードディスク装置(以下、「HDD」という)12及びスピーカ14を含む。

10

【0064】

マイクロホンA、Bは、例えば、図1に示すように、1つの音源16を挟んで同一直線上で対向しており、所定の距離dを以ってして互いに離間して配置されている。具体的には、両マイクロホンA、Bの配置態様は、一方のマイクロホンAの方が他方のマイクロホンBよりも音源16に近くなるように構成されている。

【0065】

近傍場音源分離装置10は、実質的にはコンピュータであって、そのソフトウェア資源を利用して(換言すると、プログラムに従って)、マイクロホンA、B間の振幅比に基づいて近傍場の音源分離を行なう。この近傍場音源分離装置10には、音源分離部20、強調距離入力部30及び重み関数選択・記憶部40が含まれる。

20

【0066】

音源分離部20は、上述したフーリエ変換、マイクロホンA、B間の振幅比、重み付け計算に用いられる係数、及び逆フーリエ変換等の計算処理を行なう。この音源分離部20には、音源16の音響振動から各マイクロホンA及びBで各々得られた電気信号が入力される。なお、この音源分離部20の詳細な構成については後述する。

【0067】

強調距離入力部30は、強調したい音源16からの距離を入力するためのものであって、例えば、テンキー等の入力デバイスが備えられている。この強調距離入力部30では、強調した音源16の距離が入力されると、その入力された強調距離が重み関数選択・記憶部40に与えられる。

30

【0068】

重み関数選択・記憶部40は、強調距離入力部30から入力された強調したい音源16の距離に基づいてHDD12を参照し、それによって入力された強調距離に応じた重み関数を選択して記憶すると共に、記憶した選択重み関数を所定のタイミングで音源分離部20に出力する。この重み関数選択・記憶部40には、強調したい音源16からの距離に基づき選択された重み関数を格納しておくための重み関数格納領域402が形成されている。

40

【0069】

HDD12には、複数の重み関数が蓄積されている。この蓄積された複数の重み関数は、各音源距離に対するSN比と振幅比との関係を距離減衰に基づき事前に計算してデータベース化されている。

【0070】

スピーカ14は、音源分離部20で音源分離された音響パワーを周囲の媒体に出力する。

【0071】

(音源分離部20の構成)

50

図2は音源分離部20の構成を示すブロックダイアグラムである。図2を参照して、音源分離部20は、音源16から近い方のマイクロホンA専用の入力端子200A、アナログ-デジタル変換器(A/D変換器)202A、フーリエ変換器204A及び振幅抽出器206Aと、音源16に遠い方のマイクロホンB専用の入力端子200B、アナログ-デジタル変換器(A/D変換器)202B、フーリエ変換器204B及び振幅抽出器206Bと、除算器208と、係数計算器210と、乗算器212と、逆フーリエ変換器214と、デジタル-アナログ変換器(D/A変換器)216と、スピーカ出力端子218とを含む。

【0072】

A/D変換器202Aは、入力端子200Aを介して入力された、音源16に近い方のマイクロホンAで得た電気信号にA/D変換処理を行ない、この変換処理によりデジタル化された当該マイクロホンAの電気信号をフーリエ変換器204Aに与える。

10

【0073】

フーリエ変換器204Aは、A/D変換器202Aより与えられた上記のマイクロホンAのデジタル信号に対してフーリエ変換処理を行ない、その結果を振幅抽出器206A及び乗算器212の両者に振り分けて与える。

【0074】

他方、A/D変換器202Bは、入力端子200Bを介して入力された、音源16に遠い方のマイクロホンBで得た電気信号にA/D変換処理を行ない、この変換処理によりデジタル化された当該マイクロホンBの電気信号をフーリエ変換器204Bに与える。

20

【0075】

フーリエ変換器204Bは、A/D変換器202Bより与えられた上記のマイクロホンBのデジタル信号に対してフーリエ変換処理を行ない、その結果を振幅抽出器206Bに与える。

【0076】

振幅抽出器206Aは、フーリエ変換器204Aから与えられた結果を基に音源16から近い方のマイクロホンAの振幅を抽出して演算し、当該マイクロホンAの振幅を除算器208に与える。一方、振幅抽出器206Bは、フーリエ変換器204Bから与えられた結果を基に音源16に遠い方のマイクロホンBの振幅を抽出して演算し、当該マイクロホンBの振幅を除算器208に与える。

30

【0077】

除算器208は、振幅抽出器206Aから与えられた上記音源16に近い方のマイクロホンAの振幅と振幅抽出器206Bから与えられた上記音源16に遠い方のマイクロホンBの振幅との割算を行ない、それによってマイクロホンA、B間の振幅比を得る。この振幅比は、SN比の推定に用いられる。

【0078】

係数計算器210は、図1に示す重み関数選択・記憶部40で強調したい音源16からの距離に応じて選択された重み関数に基づいて、除算器208で求められたマイクロホンA、Bの振幅比に応じた重み付け係数を計算する。この係数計算器210による重み計算は、全ての周波数成分を対象としてマイクロホンA及びBの各周波数成分ごとに行なわれる。この各周波数成分ごとに計算された重み付け係数は、係数計算器210から乗算器212に与えられる。

40

【0079】

乗算器212は、音源16に近い方のマイクロホンA用のフーリエ変換器204Aから与えられた変換結果と係数計算器210から与えられた重み付け係数とを掛け合わせる。この乗算器212による掛算は、マイクロホンA、Bの全ての周波数成分を対象として実行され、その結果が逆フーリエ変換器214に与えられる。

【0080】

逆フーリエ変換器214は、乗算器212からの掛算の結果に対して逆フーリエ変換処理を行なう。

50

【 0 0 8 1 】

D / A 変換器 2 1 6 は、逆フーリエ変換器 2 1 4 から与えられた変換結果に対して D / A 変換処理を行ない、この変換処理によりアナログ化された電気信号（データ）を、スピーカ出力端子 2 1 8 を介して、図 1 に示すスピーカ 1 4 に出力する。

【 0 0 8 2 】

< ソフトウェアの構成 >

本実施の形態においては、マイクロホン A , B 間の振幅比に基づいて近傍場の音源分離を行なうようにプログラミングされている。このプログラムは、上述したように近傍場音源分離装置のソフトウェア資源として構成されており、次に説明する近傍場音源分離機能を実現する。この機能は、上記の実質的にコンピュータからなる近傍場音源分離装置 1 0 が、このプログラムを実行することによって実現される。

10

【 0 0 8 3 】

図 3 及び図 4 は本実施の形態の近傍場音源分離機能を実現するプログラムの構造をフローチャート形式で示す図である。

【 0 0 8 4 】

なお、以下の説明においては、マイクロホン A 及び B の周波数成分 i は、それぞれ、「1」～「n」に分割されているものとして説明を行なう。

【 0 0 8 5 】

図 3 を参照して、まず、近傍場音源分離装置 1 0 は、ステップ S 1 0 0 において、同装置 1 0 の強調距離入力部 3 0 により強調したい音源 1 6 からの距離が入力されるのを待つ。ここで、強調したい音源 1 6 からの距離が入力された場合には、制御はステップ S 2 0 0 に進む。

20

【 0 0 8 6 】

制御がステップ S 2 0 0 に進むと、近傍場音源分離装置 1 0 の重み関数選択・記憶部 4 0 は、HDD 1 2 内のデータベースから当該入力された強調音源距離に応じた重み関数を選択すると共に、この選択した重み関数を同選択・記憶部 4 0 内の重み関数格納領域 4 0 2 に記憶する。その後、制御はステップ S 3 0 0 に進む。

【 0 0 8 7 】

制御がステップ S 3 0 0 に進むと、近傍場音源分離装置 1 0 の音源分離部 2 0 は、マイクロホン入力端子 2 0 0 A , 2 0 0 B を介してマイクロホン A 及び B の各音源観測信号（音源 1 6 の音響振動からマイクロホン A 及び B で各々得られた電気信号）が入力されるのを待つ。ここで、音源 1 6 に近い方のマイクロホン A からの音源観測信号の入力があつた場合には、音源分離部 2 0 は、制御を図 4 に示すステップ S 4 0 2 に進める。これに対し、音源 1 6 に遠い方のマイクロホン B からの音源観測信号の入力があつた場合には、音源分離部 2 0 は、制御を図 4 に示すステップ S 5 0 2 に進める。

30

【 0 0 8 8 】

図 4 を参照して、音源 1 6 に近い方のマイクロホン A からの音源観測信号の入力があつて制御がステップ S 4 0 2 に進むと、音源分離部 2 0 の A / D 変換器 2 0 2 A は、当該入力されたマイクロホン A からの音源観測信号に対して A / D 変換処理を行なう。この変換処理によりデジタル化されたマイクロホン A の音源観測信号は、音源分離部 2 0 のフーリエ変換器 2 0 4 A に与えられる。その後、制御はステップ S 4 0 4 に進む。

40

【 0 0 8 9 】

制御がステップ S 4 0 4 に進むと、フーリエ変換器 2 0 4 A は、上記ステップ S 4 0 2 において A / D 変換処理された上記のマイクロホン A のデジタル信号に対してフーリエ変換処理を行なう。この結果は、音源分離部 2 0 の振幅抽出器 2 0 6 A 及び乗算器 2 1 2 の両者に振り分けて与えられる。その後、制御はステップ S 4 0 6 及びステップ S 4 0 8 に進む。

【 0 0 9 0 】

制御がステップ S 4 0 6 及び S 4 0 8 に進むと、振幅抽出器 2 0 6 A は、音源 1 6 に近い方のマイクロホン A の周波数成分 i を「1」にセットすると共に、この周波数成分「1

50

」のマイクロホンAの振幅の抽出処理を行なう。すなわち、上記ステップS404においてフーリエ変換処理された結果を基に上記マイクロホンAの周波数成分「1」の振幅が抽出・演算される。このようにして得られたマイクロホンAの振幅は、音源分離部20の除算器208に与えられる。その後、制御はステップS600に進む。

【0091】

他方、音源16に遠い方のマイクロホンBからの音源観測信号の入力があって制御がステップS502に進むと、音源分離部20のA/D変換器202Bは、当該入力されたマイクロホンBからの音源観測信号に対してA/D変換処理を行なう。この変換処理によりデジタル化されたマイクロホンBの音源観測信号は、音源分離部20のフーリエ変換器204Bに与えられる。その後、制御はステップS504に進む。

10

【0092】

制御がステップS504に進むと、フーリエ変換器204Bは、上記ステップS502においてA/D変換処理された上記のマイクロホンBのデジタル信号に対してフーリエ変換処理を行なう。この結果は、音源分離部20の振幅抽出器206Bに与えられる。その後、制御はステップS506及びステップS508に進む。

【0093】

制御がステップS506及びS508に進むと、振幅抽出器206Bは、音源16に遠い方のマイクロホンBの周波数成分*i*を「1」にセットすると共に、この周波数成分「1」のマイクロホンBの振幅の抽出処理を行なう。すなわち、上記ステップS504においてフーリエ変換処理された結果を基に上記マイクロホンBの周波数成分「1」の振幅が抽出・演算される。このようにして得られたマイクロホンBの振幅は、音源分離部20の除算器208に与えられる。その後、制御はステップS600に進む。

20

【0094】

制御がステップS600に進むと、除算器208は、振幅抽出器206Aから与えられた上記音源16に近い方のマイクロホンAの振幅と振幅抽出器206Bから与えられた上記音源16に遠い方のマイクロホンBの振幅との割算を行なうことによって、マイクロホンA、B間の振幅比を得る。この振幅比からSN比を推定して、音源分離部20の係数計算器210に与えられる。その後、制御はステップS700に進む。

【0095】

制御がステップS700に進むと、係数計算器210は、重み関数選択・記憶部40内の重み関数格納領域402に格納されている重み関数（上記ステップS200（図3参照）で強調したい音源16からの距離に応じて選択された重み関数）に基づいて、上記ステップS600において除算処理により得られたマイクロホンA、Bの振幅比に応じた重み係数を計算する。この重み付け係数は、音源分離部20の乗算器212に与えられる。その後、制御はステップS800に進む。

30

【0096】

制御がステップS800に進むと、乗算器212は、音源16に近い方のマイクロホンA用のフーリエ変換器204Aから与えられた変換結果と係数計算器210から与えられた重み付け係数とを掛け合わせる。その後、制御はステップS900に進む。

【0097】

制御がステップS900に進むと、音源分離部20は、重み付け回数が*n*回行なわれたか否か（即ち、マイクロホンA及びBの全周波数成分に対して重み付けがなされたか否か）を判断する。ここで、重み付けが全周波数成分に対してなされている場合には、音源分離部20は、制御をステップS1000に進める。このとき、乗算器212による掛算の結果が音源分離部20の逆フーリエ変換器214に与えられる。逆に、重み付けが全周波数成分に対してなされていない場合には、音源分離部20は、周波数成分*i*に「1」を加算し（ステップS1300）、その後、制御をステップS408及びS508に戻す。これは、ステップS408及びS508並びにS600～S800までの重み付けに関する一連の制御が全周波数成分に対してなされるまで繰返し実行されることを意味する。

40

【0098】

50

制御がステップ S 1 0 0 0 に進むと、逆フーリエ変換器 2 1 4 は、上記ステップ S 8 0 0 での乗算結果に対して逆フーリエ変換処理を行なう。この変換結果は、音源分離部 2 0 の D / A 変換器 2 1 6 に与えられる。その後、制御はステップ S 1 1 0 0 及び S 1 2 0 0 に進む。

【 0 0 9 9 】

制御がステップ S 1 1 0 0 及び S 1 2 0 0 に進むと、D / A 変換器 2 1 6 は、上記ステップ S 1 0 0 0 において逆フーリエ変換処理により得られた結果に対して D / A 変換処理を行ない、この変換処理によりアナログ化された電気信号を音源分離信号としてスピーカ 1 4 にスピーカ出力端子 2 1 8 を介して出力する。

【 0 1 0 0 】

< 動作 >

本実施の形態では、強調したい音源 1 6 からの距離が入力されると、これに応じた重み関数が各音源距離に対する S N 比と振幅比との関係を距離減衰に基づき計算してデータベース化された複数の重み関数の中から選択される。

【 0 1 0 1 】

そして、各マイクロホン A 及び B の振幅を抽出しマイクロホン A , B 間の振幅比を求めて当該振幅比から S N 比を推定し、これに上記の選択された重み関数を用いて重み付けが行なわれる。これによって、近傍場音の強調が行なわれることになる。

【 0 1 0 2 】

特に、この近傍場音の強調に関しては、マイクロホンから特定の近距離に存在する音源 1 6 からの信号の推定値を求めるために、以下の計算が行なわれる。

【 0 1 0 3 】

入力された各マイクロホン A 及び B の音源観測信号に対してフーリエ変換が行なわれる。このフーリエ変換で得られた結果を基にして、各マイクロホン A 及び B の振幅が周波数成分ごとに抽出される。抽出された各マイクロホン A 及び B の振幅を除算することによって、S N 比を推定するためにマイクロホン A , B 間の振幅比が周波数成分ごとに求められる。そうすると、上記の選択された重み関数に基づいて、上記マイクロホン A , B 間の振幅比に応じた重みが求められる。そして、フーリエ変換で得られた、音源 1 6 に近い方のマイクロホン A 側の結果と上記重みとが周波数成分ごとに掛け合わされ、その後、この演算結果に対して逆フーリエ変換が行われる。

【 0 1 0 4 】

< 作用・効果 >

以上の説明から明らかな通り、本実施の形態によると、近傍場音源分離はマイクロホン間の振幅比に基づいて行なわれる。そのため、近傍場音源分離に関しては、室伝達関数の推定とその逆フィルタを用いることなく、遠方音を抑圧して近傍音のみを強調することができる。

【 0 1 0 5 】

< 計算シミュレーション >

部屋における音源からマイクロホンまでの伝達関数を A l l e n 等による手法、J . B . アレン及び D . A . パークレー、「小部屋での音響の効率的シミュレーションのための画像方法」、米国音響学会誌、6 5、4、9 4 3 - 9 5 0 ページ (1 9 7 9 年) (J . B . Allen and D . A . Berkley: " Image method for efficiently simulating small-room acoustics ", J. Acoust. Soc. Am., 65, 4, pp. 943-950 (1979).) を用いて合成した。インパルス応答長は、2 0 k H z サンプリングで 8 9 1 2 ポイントである。ただし、自由音場を仮定しているため、床面も含めて壁の反射係数は全て「 0 」とした。

【 0 1 0 6 】

また、マイクロホン配置及び目的音と妨害音との位置関係については、図 9 に示す配置を用いた。部屋の大きさは、縦横 2 0 m、高さが 3 m の直方体である。括弧内の数値は、x 軸方向及び y 軸方向における座標をメートル単位で表記したものであり、高さ方向につ

10

20

30

40

50

いては、マイクロホンは床面から 1.5 m、音源についてはどちらも床面から 1.0 m とした。音源は、NTT - A T 社の A T R 音声データベース（現在は、A T R - P r o m o t i o n s から販売）の中から表 2 に示す音源を使用した。

【0107】

【表 2】

表 2

条件	音源	登録番号 (継続時間)
1	目的音	MAUSC405 (6.4 s)
	妨害音	MAUSC323 (7.5 s)
2	目的音	MAUSC405 (6.4 s)
	妨害音	FTKSC323 (8.3 s)

10

【0108】

表 2 を参照して、条件 1 では、目的音と妨害音はどちらも同一男性話者である。一方、条件 2 での妨害音は、女性話者の発話音を採用した。

【0109】

本計算シミュレーションで行った信号処理に関しては、図 1 及び図 2 に示すブロックダイアグラムに従って行なった。性能評価用のために使用する信号の取得方法も、これらの図のブロックダイアグラムには含まれている。

20

【0110】

本実施の形態による提案手法の性能評価は、下記の 2 つの指標に基づいて行なった。

【0111】

第 1 の指標である SN 比 (SNR) は、式 (10) で定義され、第 2 の指標である SDR は、式 (11) で定義される。

【数 14】

$$SNR = 10 \log \frac{\|\tilde{x}_A(n)\|^2}{\|\tilde{v}_A(n)\|^2} \quad (10)$$

30

$$SDR = 10 \log \frac{\|x_A(n)\|^2}{\|c \cdot \hat{x}_A(n) - x_A(n)\|^2} \quad (11)$$

ただし、

$$c = \frac{\|\hat{x}_A(n) \cdot x_A(n)\|^2}{\|\hat{x}_A(n) \cdot \hat{x}_A(n)\|^2}$$

【0112】

妨害音が男性話者の場合の SN 比及び SDR の結果を図 10 (a) に、女性話者の場合を図 10 (b) に示す。なお、図 10 (a) には、従来法との比較のために、delay - and - sum の結果も付記している。

40

【0113】

これらの図から、処理前の SN 比に関係なく、SN 比及び SDR が本実施の形態による提案手法により男性話者妨害音のときで 7 dB 程度、女性話者妨害音のときには 9 dB 程度改善していることが分かる。

【0114】

< 実験 >

上述した計算シミュレーションの結果を検証するため、無響室において本実施の形態に

50

よる提案手法の性能評価実験を行った。その測定環境を図 1 1 に示す。

【 0 1 1 5 】

マイクロホン及びスピーカは、床面から 9 5 c m の同一平面上に配置した。音源は、計算シミュレーションで用いたものと全く同じものを使用した。2 つのスピーカで再生する音声は、それぞれステレオの L チャンネルと R チャンネルとに記録し、P C (P e r s o n a l C o m p u t e r) に U S B (U n i v e r s a l S e r i a l B u s) ケーブルで接続した R o l a n d E D I R O L U A - 1 0 1 で D / A 変換した後、オーディオアンプ (A c c u p h a s e E - 3 0 5) で増幅して、スピーカ (D I A T O N E D S - 1 0 7 V) で再生した。

【 0 1 1 6 】

また、マイクロホン (B & K 4 9 5 1) で観測した信号は、日本ナショナルインスツルメンツの P X I - 4 4 6 2 D A Q モジュールを用いて 2 0 k H z で標本化を行った。

【 0 1 1 7 】

観測した信号を基に各時間周波数要素に対するゲインを決定し、目的音及び妨害音のみを收音した信号を用いて計算シミュレーションの時と同様の方法で S N 比を算出した。求めた処理後の S N 比の結果を図 1 2 に示す。

【 0 1 1 8 】

この図から明らかなように、女性話者の妨害音に対しては約 8 d B の改善が見られ、男性話者に対しては約 6 d B 程度の改善が確認された。これらの数値は、計算シミュレーションで得られた値よりも若干小さいがほぼ同じ値であり、実測データからも本実施の形態による提案手法の有効性が確認された。

【 0 1 1 9 】

なお、本発明は上記実施の形態に限定されるものではない。

【 0 1 2 0 】

例えば、上述したように、上記実施の形態で使用した男性話者と女性話者との妨害音は、同一の発話内容であるが、S N 比の改善量に少くない差異が確認された。これは、発話時間長が多少異なるために時間的な重なりもある程度影響していることが考えられるが、主な理由は時間周波数領域での重なりが女性話者の方が減少しているためであると考えられる。この点を確かめるため、実施の形態で用いた目的音声、男性話者の妨害音、及び女性話者の妨害音の短区間振幅スペクトルを求め、各フレーム毎に目的音声の短区間振幅スペクトルとの相関を計算して平均した。その結果、男性話者の妨害音では 0 . 5 2、女性話者の妨害音では 0 . 4 2 であった。このことから、女性話者妨害音の振幅スペクトルの分布は、男性話者妨害音の場合よりも目的音声のそれと異なっていたことが原因の 1 つであると推定される。図 6 において、C = 1 の場合には、R_{min} は R = 0 で連続に繋がり、R_{min} と R_{max} は共に正負方向に関して対称となる。これは、目的音が 2 つのマイクロホンから等距離に位置しており、目的音の両マイクロホンでの受信信号に関する振幅比が「1」であることに対応する。この場合、S N 比の期待値

【 外 8 】

【 0 1 2 1 】

\hat{D}

は、観測された振幅比 R に依存せずに常に 0 d B となるため、目的音の強調ができないことになる。このような場合には、マイクロホン数を増やして適切なマイクロホンの組み合わせを選択したり、又は複数の組み合わせで計算した結果を統合したりする等の工夫を施すようにすれば、上記の問題を回避することが可能となると考えられる。

【 0 1 2 2 】

また、もし音源に関する先験的な知識が利用可能であれば、S N R の尤度を基に、表 1 における当該成分の S N R の期待値を求める計算法を検討することで、より性能を高められると期待される。これには、尤度推定による重みの決定の最適化を検討すればよい。

【 0 1 2 3 】

さらに、上記実施の形態では、少数(2つ)のマイクロホンを用いて、マイクロホンアレイから近傍の音のみを強調する手法を提案した。この提案手法は、音源から放射された音の振幅が、距離に反比例して減少するため、音源が近傍に存在する場合は2点間の振幅比が大きい、遠方の音源に対しては、同じ距離離れていても振幅比が小さくなることを利用したものである。計算シミュレーション及び実験の結果、自由音場ではSN比で8dB程度の強調効果を実現できたが、重み付けの関数に関して、MAP(Maximum A Posteriori)推定のような確率的な考え方を導入すれば、本手法の性能を音質の面も含めてさらに向上させ得る可能性がある。

【0124】

その他、本明細書に添付の特許請求の範囲内での種々の設計変更及び修正を加え得ることは勿論である。

10

【0125】

すなわち、本明細書で開示した実施の形態は単に例示であって、本発明が上述した実施の形態のみに限定されるわけではない。本発明の範囲は、本明細書の記載内容を参酌した上で、別紙の特許請求の範囲における請求項によって示され、そこに記載された文言と均等の意味及び範囲内での全ての変更を含む。

【図面の簡単な説明】

【0126】

【図1】本発明の実施の形態に係るシステムの構成を示すブロックダイアグラムである。

【図2】音源分離部の構成を示すブロックダイアグラムである。

20

【図3】本実施の形態の近傍場音源分離機能を実現するプログラムの構造をフローチャート形式で示す図である。

【図4】本実施の形態の近傍場音源分離機能を実現するプログラムの構造をフローチャート形式で示す図であって、図3の続きを示す。

【図5】2つのマイクロホン間における振幅比Cが2つのマイクロホン間距離とマイクロホンと音源との距離に対してどのような振る舞いをなすのかを示す図である。

【図6】幾つかのパラメータCの値に対するSN比の関数としての振幅比Rの範囲を示す図である。

【図7】観測された振幅比Rの関数としての推定されたSN比を示す図である。

【図8】幾つかのパラメータCの値に対する重み関数を示す図である。

30

【図9】コンピュータシミュレーションにおけるマイクロホン及び音源の配置を示す図である。

【図10】提案手法及び従来手法に対するSN比及びSDRの入力及び出力を示す図である。

【図11】実験用の配置を示す図である。

【図12】実験において観測された提案手法のSN比の入力及び出力を示す図である。

【符号の説明】

【0127】

A, B マイクロホン

10 近傍場音源分離装置

40

12 HDD

14 スピーカ

16 音源

20 音源分離部

204 A, 204 B フーリエ変換器

206 A, 206 B 振幅抽出器

208 除算器

210 係数計算器

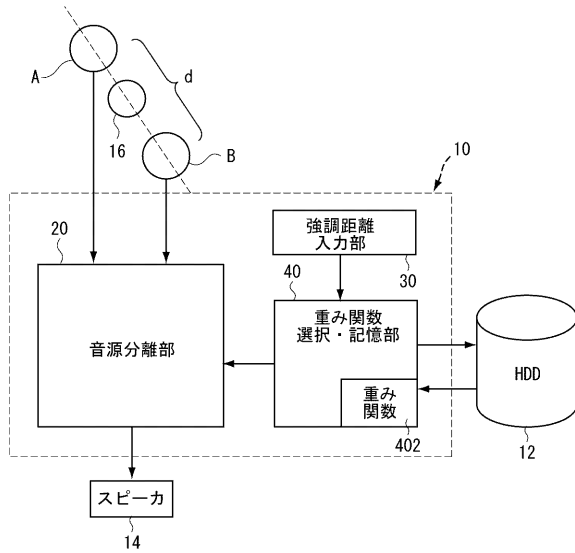
212 乗算器

214 逆フーリエ変換器

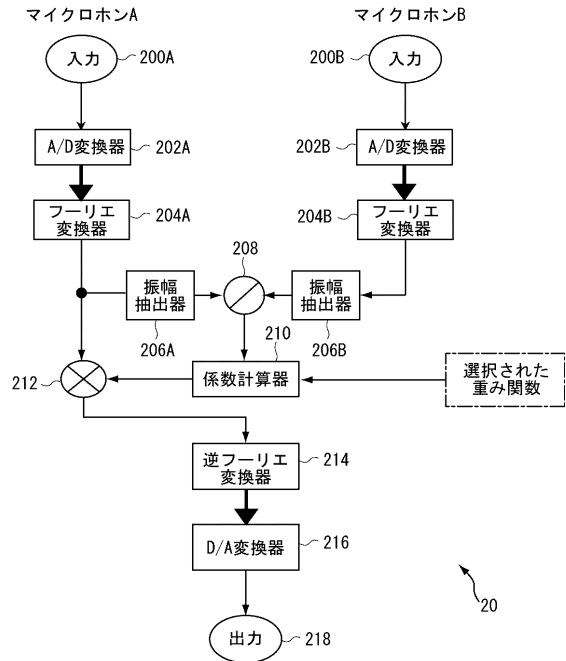
50

- 3 0 強調距離入力部
- 4 0 重み関数選択・記憶部

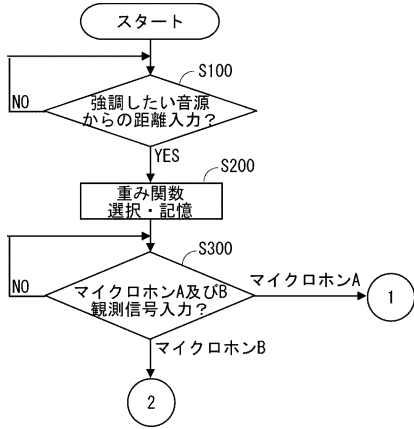
【図1】



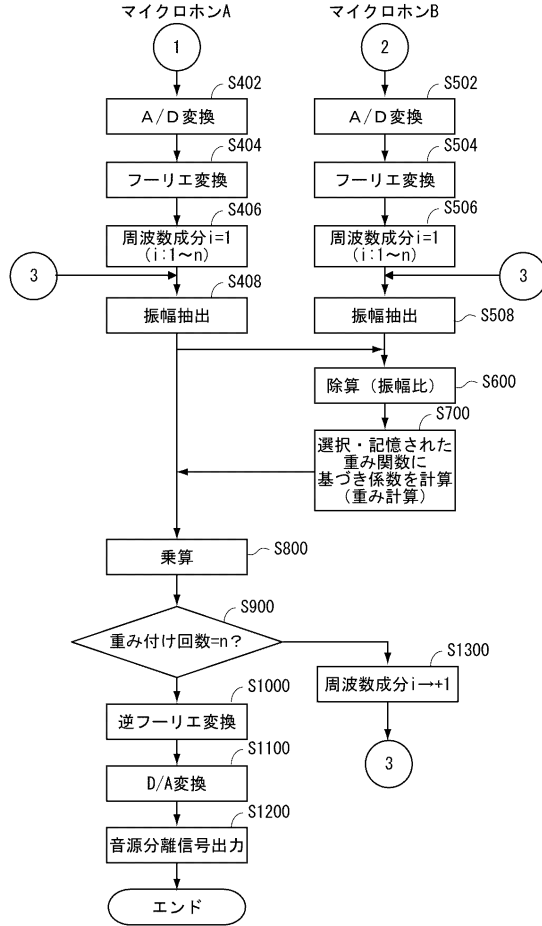
【図2】



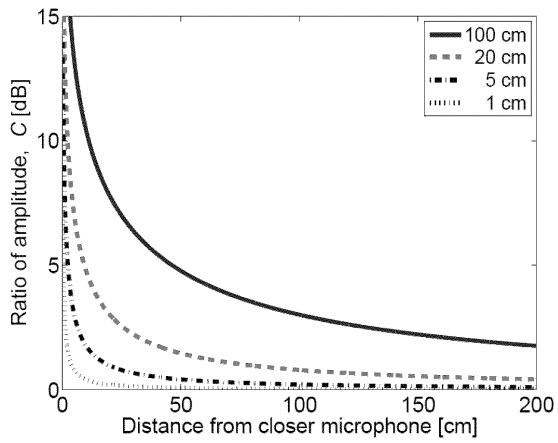
【図3】



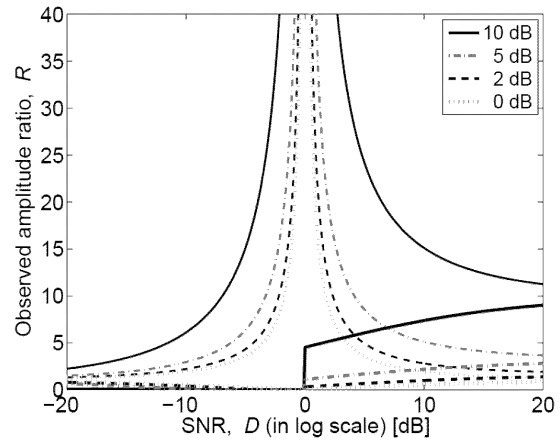
【図4】



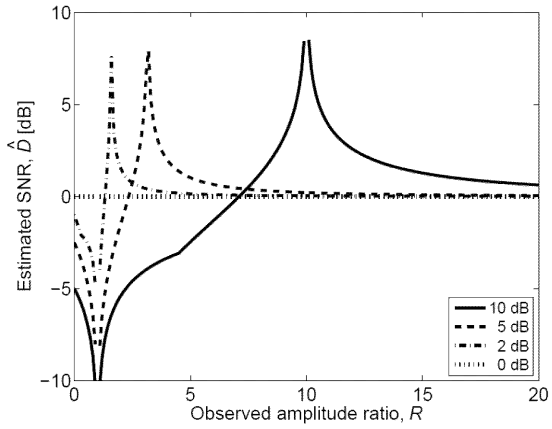
【図5】



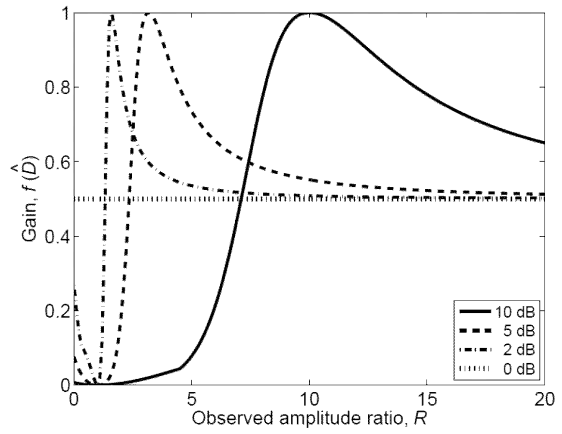
【図6】



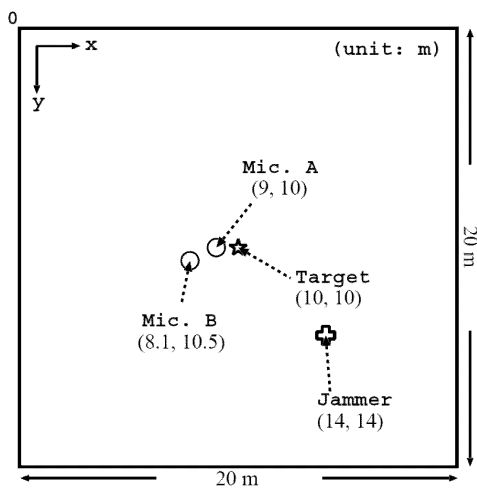
【 図 7 】



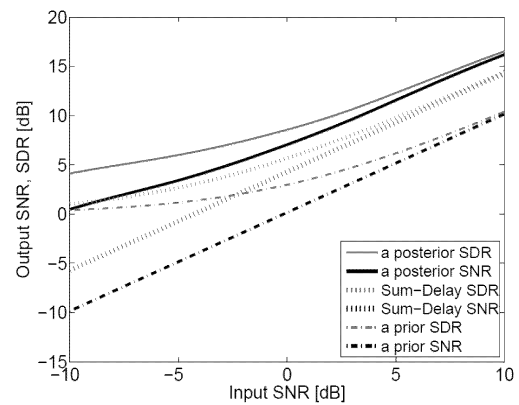
【 図 8 】



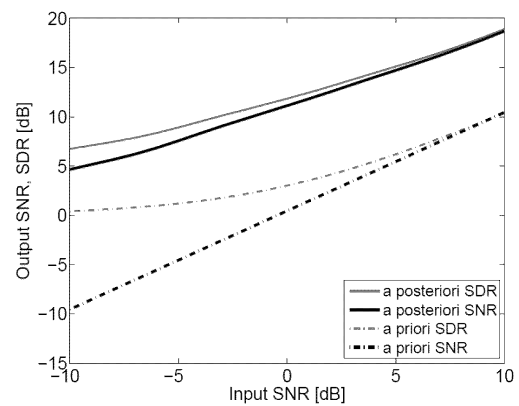
【 図 9 】



【 図 10 】

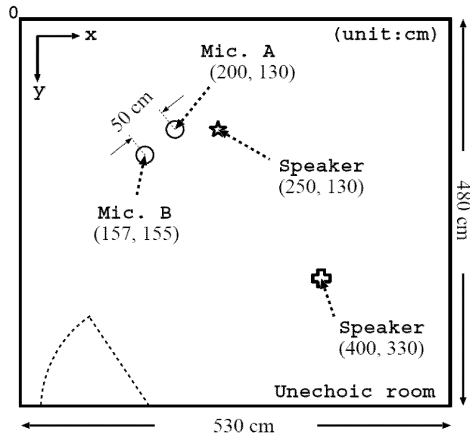


(a) Male voice interference

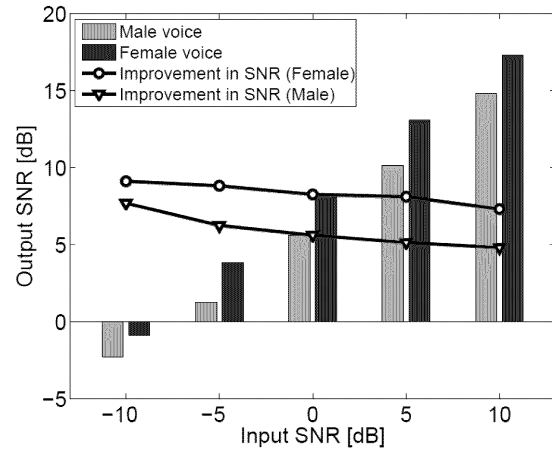


(b) Female voice interference

【 図 1 1 】



【 図 1 2 】



フロントページの続き

(56)参考文献 国際公開第2006/025106(WO, A1)
特開平09-251299(JP, A)

(58)調査した分野(Int.Cl., DB名)

G10L 21/00 - 21/06

G10L 11/00 - 11/06

H04R 1/40

H04R 3/00