

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2009-36810

(P2009-36810A)

(43) 公開日 平成21年2月19日(2009.2.19)

(51) Int.Cl.		F I		テーマコード (参考)
G10L 21/02	(2006.01)	G10L 21/02	102B	5D018
H04R 3/00	(2006.01)	H04R 3/00	320	5D020
H04R 1/40	(2006.01)	G10L 21/02	203Z	
		G10L 21/02	201C	
		H04R 1/40	320A	

審査請求 未請求 請求項の数 4 O L (全 22 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)

(71) 出願人 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
 Fターム(参考) 5D018 BB24
 5D020 BB07

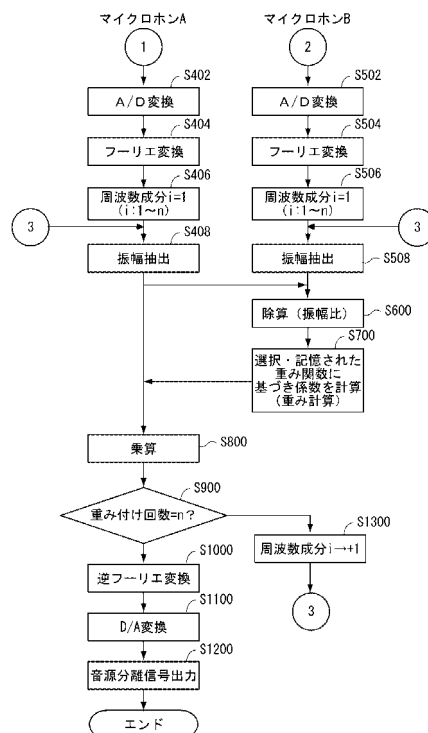
(54) 【発明の名称】 近傍場音源分離プログラム、及びこのプログラムを記録したコンピュータ読取可能な記録媒体、並びに近傍場音源分離方法

(57) 【要約】

【課題】 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、遠方音を抑制して近傍音のみを強調し得る、近傍場音源分離を提供すること。

【解決手段】 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に、各マイクロホンの振幅を抽出しマイクロホン間の振幅比を求めて当該振幅比からSN比を推定し、これに、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入力された強調したい音源からの距離に応じて選択された重み関数を用いて重み付けを行ない、それによって近傍場音の強調を行なう。

【選択図】 図4



【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するための重み関数選択手段、

各マイクロホンの振幅を抽出しマイクロホン間の振幅比を求めて当該振幅比からSN比を推定し、これに前記重み関数選択手段により選択された重み関数を用いて重み付けを行ない、それによって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。

10

【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に対して第1の線形変換を行なうための第1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結果を基にして各マイクロホンの振幅を抽出するための第2の手段、

前記第2の手段により抽出された各マイクロホンの振幅を除算し、それによって前記SN比を推定するためにマイクロホン間の振幅比を求めるための第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振幅比に応じた重みを求めるための第4の手段、

20

前記第1の手段による第1の線形変換で得られた、対をなすマイクロホン間の一方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の手段による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の線形変換を行なうための第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に、コンピュータを、

30

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するための重み関数選択手段、

各マイクロホンの振幅を抽出しマイクロホン間の振幅比を求めて当該振幅比からSN比を推定し、これに前記重み関数選択手段により選択された重み関数を用いて重み付けを行ない、それによって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方法であって、

40

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するための第1のステップと、

各マイクロホンの振幅を抽出しマイクロホン間の振幅比を求めて当該振幅比からSN比を推定し、これに前記第1のステップにおいて選択された重み関数を用いて重み付けを行ない、それによって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】**

50

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。

【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いても、「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。

10

【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。

20

【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。

30

【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。

40

【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】

50

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。

【 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 nen: "Creating interactive virtual acoustic environments", J. Audio Eng. Soc., 47, 9, pp. 675-705 (1999).）

20

【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).）

30

【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 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. アブハヤバラ、「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).）

40

【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 localization 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).）

50

【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」IEE 予稿集、140、6、362-370 ページ(1993 年)。(J. F. Cardoso and A. Souloumiac: "Blind beam forming for non-gaussian signals", IEE Proceedings, 140, 6, pp. 362-370 (1993).)

【非特許文献 9】F. エーラーズ及び 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するための重み関数選択手段、各マイクロホンの振幅を抽出しマイクロホン間の振幅比を求めて当該振幅比から SN 比を推定し、これに重み関数選択手段により選択された重み関数を用いて重み付けを行ない、それによって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である。

【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に対して第 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の振幅を抽出するための第 2 の手段、第 2 の手段により抽出された各マイクロホンの振幅を除算し、それによって SN 比を推定するためにマイクロホン間の振幅比を求めるための第 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振幅比に応じた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、対をなすマイクロホン間の一方の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掛け合わせるための第 5 の手段、及び第 5 の手段による演算結果に対して第 1 の

10

20

30

40

50

線形変換の逆変換に対応する第2の線形変換を行なうための第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するための重み関数選択手段、各マイクロホンの振幅を抽出しマイクロホン間の振幅比を求めて当該振幅比からSN比を推定し、これに重み関数選択手段により選択された重み関数を用いて重み付けを行ない、それによって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である。

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するための第1のステップと、各マイクロホンの振幅を抽出しマイクロホン間の振幅比を求めて当該振幅比からSN比を推定し、これに第1のステップにおいて選択された重み関数を用いて重み付けを行ない、それによって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である。

20

【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>

30

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、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。

40

【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

50

る。

【 0 0 2 8 】

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遍在しており、その遍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。

【 0 0 2 9 】

これまでも、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。

【 0 0 3 0 】

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 と他方のマイクロホン B 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 R は、下記式 (1) となる。

【 0 0 3 1 】

【 数 1 】

$$\begin{aligned}
 R &= \frac{|S_A + N_A|}{|S_B + N_B|} \\
 &= \frac{\sqrt{(S_A + N_A)(S_A + N_A)^*}}{\sqrt{(S_B + N_B)(S_B + N_B)^*}} \\
 &= \frac{\sqrt{|S_A|^2 + |N_A|^2 + 2|S_A||N_A|\cos(\arg S_A - \arg N_A)}}{\sqrt{|S_B|^2 + |N_B|^2 + 2|S_B||N_B|\cos(\arg S_B - \arg N_B)}}
 \end{aligned}
 \tag{1}$$

ここで、* は複素共役を表す。

【 0 0 3 2 】

目的音は、マイクロホン A, B に近い場合、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「0」に近似する。

【 数 3 】

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

【 0 0 3 4 】

そのため、次の式 (2) で書き表せる。

【 数 4 】

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

【 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、下記の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 】

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

【 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 】

そのため、ここでは上記の両振幅の比 C をパラメータとして検討することにする。

【 数 7 】

$$C = |S_A|/|S_B|$$

【 0 0 4 1 】

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。

10

20

30

40

50

【 0 0 4 2 】

【 数 9 】

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

【 0 0 4 3 】

同様に、S N比が低い場合の振幅比 R の近似式は、式 (5) で表される。

【 数 1 0 】

$$R \simeq \frac{1 + C \cdot D \cos(\arg S_A - \arg N_A)}{1 + D \cos(\arg S_B - \arg N_B)} \quad (5) \quad 10$$

【 0 0 4 4 】

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 1 1 】

$$R_{\max} = \begin{cases} \frac{C \cdot D + 1}{D - 1} & D \geq 1 \\ -\frac{C \cdot D + 1}{D - 1} & D < 1 \end{cases} \quad (6) \quad 20$$

$$R_{\min} = \begin{cases} \frac{C \cdot D - 1}{D + 1} & \text{for } C \geq 1 & \text{for } C < 1 \\ 0 & 1 > D \geq 1/C & 1/C > D \geq 1 \\ -\frac{C \cdot D - 1}{D + 1} & 1/C > D & 1 > D \end{cases} \quad (7) \quad 30$$

【 0 0 4 5 】

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 S N比 (D) の範囲が求まることを意味している。

【 0 0 4 6 】

そこで、上記の D の範囲の代表値を S N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。

【 0 0 4 7 】

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。

40

【 0 0 4 8 】

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。

【 0 0 4 9 】

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。

【 0 0 5 0 】

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

50

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】

【表1】

表1

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

【0056】

この推定されたSN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 】

 $f(\hat{D})$

10

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 】

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。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

20

【 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 1 3 】

30

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

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。

【 0 0 6 3 】

< 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。

40

【 0 0 6 4 】

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。

【 0 0 6 5 】

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

50

調距離入力部 30 及び重み関数選択・記憶部 40 が含まれる。

【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。

10

【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 HDD 12 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。

【0069】

HDD 12 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。

20

【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)

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。

30

【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。

【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。

40

【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。

【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

50

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。

【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。

10

【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。

【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。

20

【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。

【0081】

D/A変換器216は、逆フーリエ変換器214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218を介して、図1に示すスピーカ14に出力する。

30

【0082】

<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10が、このプログラムを実行することによって実現される。

【0083】

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。

40

【0084】

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。

【0085】

図3を参照して、まず、近傍場音源分離装置10は、ステップS100において、同装置10の強調距離入力部30により強調したい音源16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16からの距離が入力された場合には、制御はステップS200に進む。

【0086】

50

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、H D D 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った場合には、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った場合には、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。

10

【 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って制御がステップ 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。

【 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。

20

【 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 」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 S 4 0 4 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 2 0 の除算器 2 0 8 に与えられる。その後、制御はステップ S 6 0 0 に進む。

30

【 0 0 9 1 】

他方、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って制御がステップ S 5 0 2 に進むと、音源分離部 2 0 の A / D 変換器 2 0 2 B 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 2 0 のフーリエ変換器 2 0 4 B に与えられる。その後、制御はステップ S 5 0 4 に進む。

【 0 0 9 2 】

制御がステップ S 5 0 4 に進むと、フーリエ変換器 2 0 4 B は、上記ステップ S 5 0 2 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 2 0 の振幅抽出器 2 0 6 B に与えられる。その後、制御はステップ S 5 0 6 及びステップ S 5 0 8 に進む。

40

【 0 0 9 3 】

制御がステップ S 5 0 6 及び S 5 0 8 に進むと、振幅抽出器 2 0 6 B は、音源 1 6 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 S 5 0 4 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 2 0 の除算器 2 0 8 に与えられる。その後、制御はステップ S 6 0 0 に進む。

【 0 0 9 4 】

50

制御がステップ S 6 0 0 に進むと、除算器 2 0 8 は、振幅抽出器 2 0 6 A から与えられた上記音源 1 6 に近い方のマイクロホン A の振幅と振幅抽出器 2 0 6 B から与えられた上記音源 1 6 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 S N 比を推定して、音源分離部 2 0 の係数計算器 2 1 0 に与えられる。その後、制御はステップ S 7 0 0 に進む。

【 0 0 9 5 】

制御がステップ S 7 0 0 に進むと、係数計算器 2 1 0 は、重み関数選択・記憶部 4 0 内の重み関数格納領域 4 0 2 に格納されている重み関数（上記ステップ S 2 0 0（図 3 参照）で強調したい音源 1 6 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 S 6 0 0 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 2 0 の乗算器 2 1 2 に与えられる。その後、制御はステップ S 8 0 0 に進む。

10

【 0 0 9 6 】

制御がステップ S 8 0 0 に進むと、乗算器 2 1 2 は、音源 1 6 に近い方のマイクロホン A 用のフーリエ変換器 2 0 4 A から与えられた変換結果と係数計算器 2 1 0 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 S 9 0 0 に進む。

【 0 0 9 7 】

制御がステップ S 9 0 0 に進むと、音源分離部 2 0 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 2 0 は、制御をステップ S 1 0 0 0 に進める。このとき、乗算器 2 1 2 による掛算の結果が音源分離部 2 0 の逆フーリエ変換器 2 1 4 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 2 0 は、周波数成分 i に「1」を加算し（ステップ S 1 3 0 0）、その後、制御をステップ S 4 0 8 及び S 5 0 8 に戻す。これは、ステップ S 4 0 8 及び S 5 0 8 並びに S 6 0 0 ~ S 8 0 0 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。

20

【 0 0 9 8 】

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。

30

【 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。

40

【 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

50

。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、SN比を推定するためにマイクロホンA、B間の振幅比が求められる。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16に近い方のマイクロホンA側の結果と上記重みとが掛け合わされ、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。

【0104】

<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。

10

【0105】

<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Allen等による手法、J. B. Allen及びD. A. Berkley、「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、65、4、943-950ページ(1979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、20kHzサンプリングで8912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。

20

【0106】

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20m、高さが3m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、高さ方向については、マイクロホンは床面から1.5m、音源についてはどちらも床面から1.0mとした。音源は、NTT-AT社のATR音声データベース(現在は、ATR-Promotionsから販売)の中から表2に示す音源を使用した。

30

【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)

40

【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。

【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】

50

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】

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

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

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。

【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7dB程度、女性話者妨害音のときには9dB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11に示す。

【0115】

マイクロホン及びスピーカは、床面から95cm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、PC(Personal Computer)にUSB(Universal Serial Bus)ケーブルで接続したRoland EDIROLUA-101でD/A変換した後、オーディオアンプ(Accuphase E-305)で増幅して、スピーカ(DIATONE DS-107V)で再生した。

【0116】

また、マイクロホン(B&K 4951)で観測した信号は、日本ナショナルインスツルメンツのPXI-4462 DAQモジュールを用いて20kHzで標本化を行った。

【0117】

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収録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SN比を算出した。求めた処理後のSN比の結果を図12に示す。

【0118】

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8dBの改善が見られ、男性話者に対しては約6dB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。

【0119】

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。

【0120】

10

20

30

40

50

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、SN比の改善量に少な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。この点を確認するため、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0.52、女性話者の妨害音では0.42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
10
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、SN比の期待値

【外8】

【0121】

\hat{D}

は、観測された振幅比 R に依存せず常に0dB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
20
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。

【0122】

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、SNRの尤度を基に、表1における当該成分のSNR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。

【0123】

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
30
が、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。

【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
40
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。

【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 S/N 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 S/N 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。

【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 S/N 比及び $S/D/R$ の入力及び出力を示す図である。

10

【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 S/N 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

12 HDD

14 スピーカ

16 音源

20 音源分離部

204A, 204B フーリエ変換器

206A, 206B 振幅抽出器

208 除算器

210 係数計算器

212 乗算器

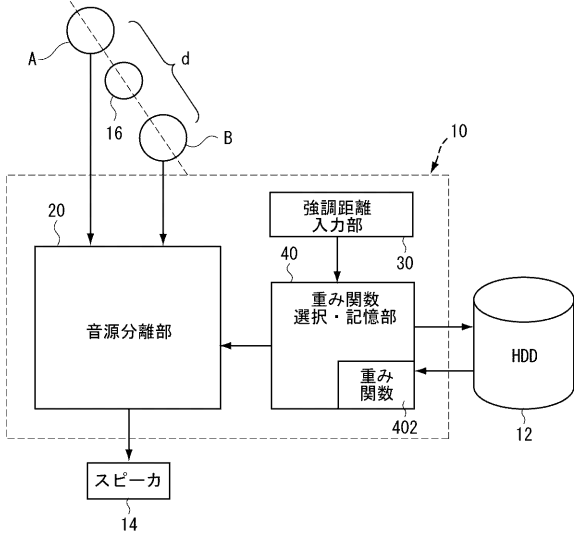
214 逆フーリエ変換器

30 強調距離入力部

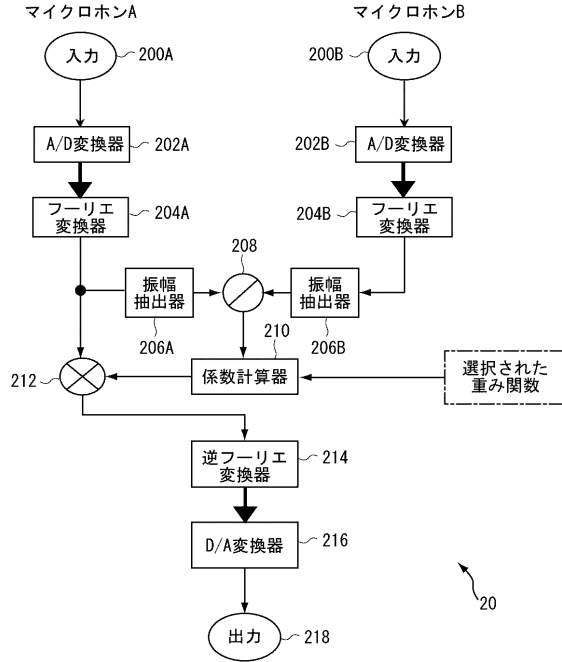
40 重み関数選択・記憶部

20

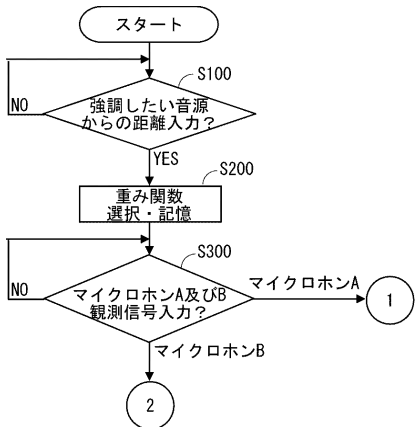
【図1】



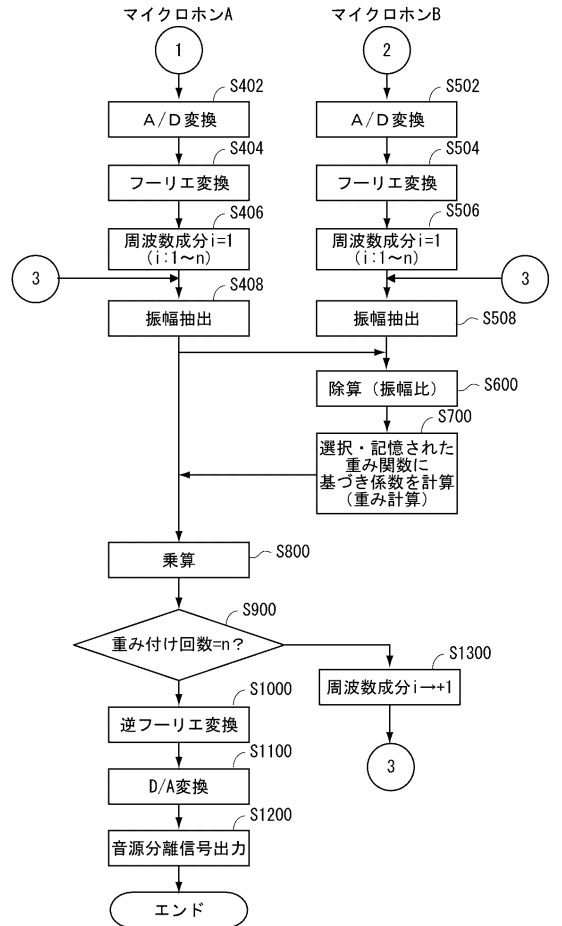
【図2】



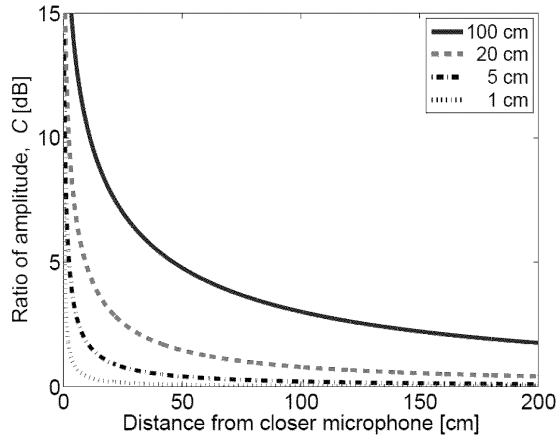
【図3】



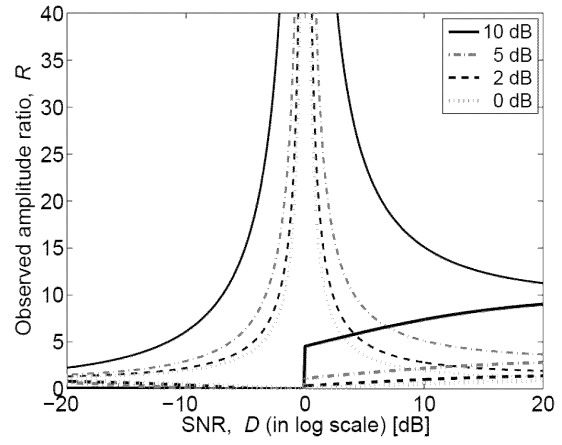
【図4】



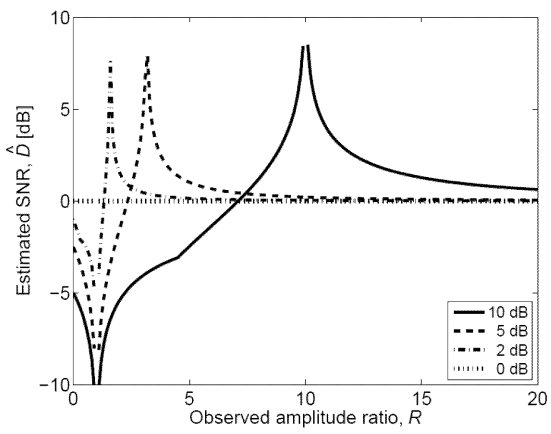
【 図 5 】



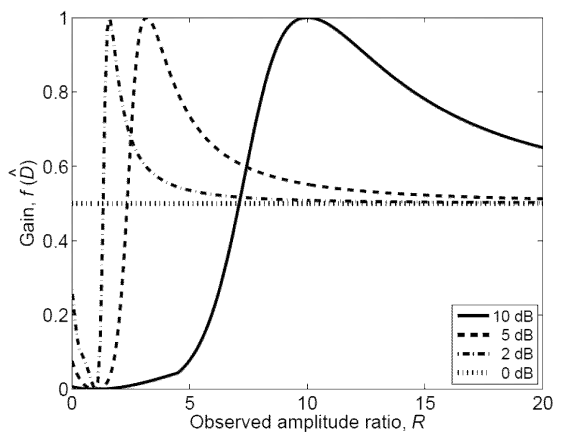
【 図 6 】



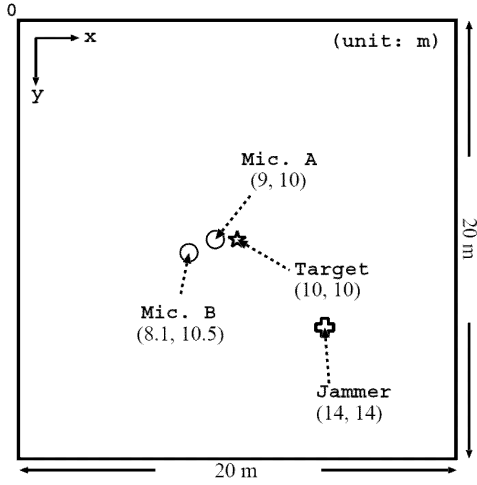
【 図 7 】



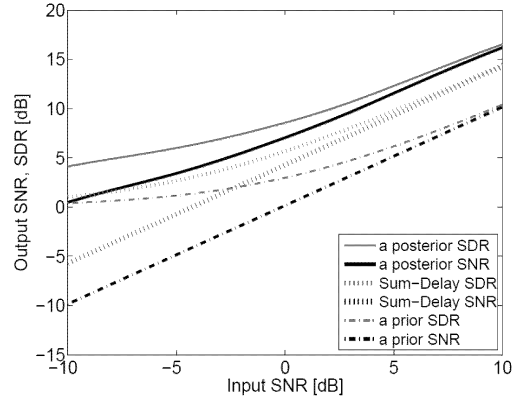
【 図 8 】



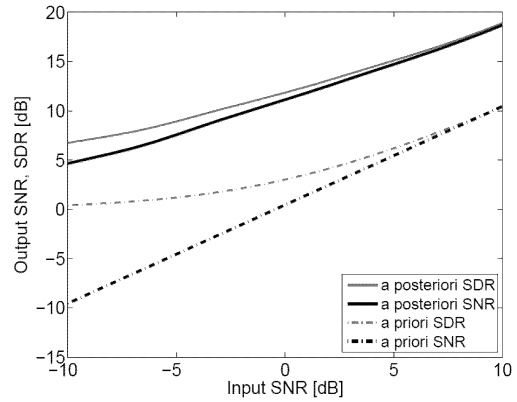
【 図 9 】



【 図 1 0 】

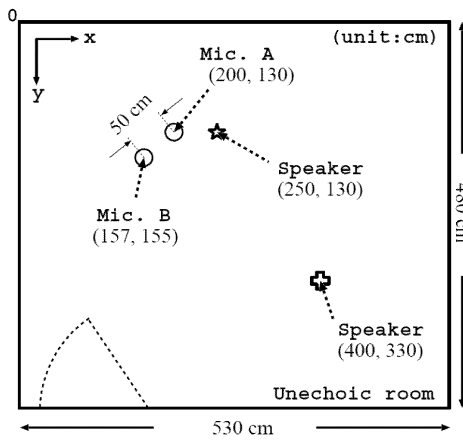


(a) Male voice interference



(b) Female voice interference

【 図 1 1 】



【 図 1 2 】

