

(19) 日本国特許庁(JP)

再公表特許(A1)

(11) 国際公開番号

WO2017/204226

発行日 平成31年3月22日 (2019. 3. 22)

(43) 国際公開日 平成29年11月30日 (2017. 11. 30)

(51) Int. Cl.	F I	テーマコード (参考)
G 1 O L 21/0264 (2013. 01)	G 1 O L 21/0264 Z	
G 1 O L 21/0208 (2013. 01)	G 1 O L 21/0208 1 0 0 A	
G 1 O L 21/0232 (2013. 01)	G 1 O L 21/0208 1 0 0 Z	
	G 1 O L 21/0232	

審査請求 未請求 予備審査請求 未請求 (全 25 頁)

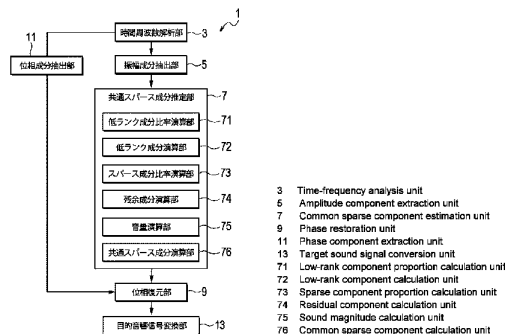
出願番号 特願2018-519566 (P2018-519566)	(71) 出願人 504132272 国立大学法人京都大学 京都府京都市左京区吉田本町 3 6 番地 1
(21) 国際出願番号 PCT/JP2017/019259	
(22) 国際出願日 平成29年5月23日 (2017. 5. 23)	
(31) 優先権主張番号 特願2016-102063 (P2016-102063)	(74) 代理人 100091443 弁理士 西浦 ▲嗣▼晴
(32) 優先日 平成28年5月23日 (2016. 5. 23)	
(33) 優先権主張国 日本国 (JP)	(72) 発明者 坂東 宜昭 京都府京都市左京区吉田本町 3 6 番地 1 国立大学法人京都大学内
	(72) 発明者 吉井 和佳 京都府京都市左京区吉田本町 3 6 番地 1 国立大学法人京都大学内
	(72) 発明者 糸山 克寿 京都府京都市左京区吉田本町 3 6 番地 1 国立大学法人京都大学内

最終頁に続く

(54) 【発明の名称】 目的音響信号復元システム及び方法

(57) 【要約】

事前情報を用いずに雑音が含まれる音響信号から目的音響信号を高い精度で復元することができる音声信号復元システム及び方法を提供する。Mチャンネルの振幅スペクトログラムを入力として、Mチャンネルの振幅スペクトログラムのうち最も多くのチャンネルの振幅スペクトログラムに共通に含まれている可能性が高いスパース時間周波数成分を含む共通スパース成分を推定する共通スパース成分推定部7を設ける。位相復元部9は、共通スパース成分の位相を復元して目的音響複素スペクトログラムとする。目的音響信号変換部13は、目的音響複素スペクトログラムを時間信号である目的音響信号に変換する。



【特許請求の範囲】**【請求項 1】**

M本（Mは2以上の整数）のマイクロホンによって採取したMチャンネルの音響信号に含まれる目的音響信号を復元する目的音響信号復元システムであって、

前記M本（Mは2以上の整数）のマイクロホンによって採取した前記Mチャンネルの音響信号を時間周波数解析してMチャンネルの複素スペクトログラムを得る時間周波数解析部と

、
前記Mチャンネルの複素スペクトログラムからMチャンネルの振幅スペクトログラムを抽出する振幅成分抽出部と、

前記Mチャンネルの振幅スペクトログラムを入力として、前記Mチャンネルの振幅スペクトログラムのうち最も多くのチャンネルの振幅スペクトログラムに共通に含まれている可能性が高いスパース時間周波数成分を含む共通スパース成分を推定する共通スパース成分推定部と、

前記共通スパース成分の位相を復元して目的音響複素スペクトログラムとする位相復元部と、

前記目的音響複素スペクトログラムを時間信号である前記目的音響信号に変換する目的音響信号変換部とからなる目的音響信号復元システム。

【請求項 2】

前記共通スパース成分推定部は、前記Mチャンネルの振幅スペクトログラムからそれぞれ反復推定 $i-1$ 回目（ i は2以上の正数）の低ランク成分を除いて得たM個のスパース成分を含む残余の総和を、前記共通スパース成分に前記M個のスパース成分が含まれている含有比率の総和で除算して得た結果を反復推定 i 回目の前記共通スパース成分として推定する請求項1に記載の目的音響信号復元システム。

【請求項 3】

前記共通スパース成分推定部は、

前記Mチャンネルの振幅スペクトログラムに含まれる低ランク成分の比率を演算する低ランク成分比率演算部と、

前記低ランク成分の比率に基づいて前記Mチャンネルの振幅スペクトログラムに含まれるM個の低ランク成分を演算する低ランク成分演算部と、

前記Mチャンネルの振幅スペクトログラムに含まれるスパース成分の比率を演算するスパース成分比率演算部と、

前記スパース成分の比率に基づいて前記Mチャンネルの振幅スペクトログラムに含まれるM個のスパース成分を含む残余成分を演算する残余成分演算部と、

前記M個のスパース成分を含む残余成分と前記共通スパース成分とに基づいて、前記共通スパース成分に前記M個のスパース成分が含まれている含有比率を前記M個のスパース成分の音量として演算する音量演算部と、

前記M個のスパース成分を含む残余成分の総和を前記M個のスパース成分の音量の総和で除算して前記共通スパース成分を演算する共通スパース成分演算部とを備えて、

前記低ランク成分比率演算部、低ランク成分演算部、スパース成分比率演算部、残余成分演算部、音量演算部及び前記共通スパース成分演算部において反復演算を行うことにより前記共通スパース成分を推定する請求項2に記載の音声信号復元システム。

【請求項 4】

前記共通スパース成分推定部は、変分ベイズEM法または逐次変分ベイズEM法により前記共通スパース成分をベイズ推定するベイズ推定器によって構成されている請求項1に記載の音声信号復元システム。

【請求項 5】

コンピュータを用いて、M本（Mは2以上の整数）のマイクロホンによって採取したMチャンネルの音響信号に含まれる目的音響信号を復元する目的音響信号復元方法であって、

前記M本（Mは2以上の整数）のマイクロホンによって採取した前記Mチャンネルの音響信号を時間周波数解析してMチャンネルの複素スペクトログラムを得る時間周波数解析ステ

10

20

30

40

50

ップと、

前記 M チャンネルの複素スペクトログラムから M チャンネルの振幅スペクトログラムを抽出する振幅成分抽出ステップと、

前記 M チャンネルの振幅スペクトログラムを入力として、前記 M チャンネルの振幅スペクトログラムのうち最も多くのチャンネルの振幅スペクトログラムに共通に含まれている可能性が高いスパース時間周波数成分を含む共通スパース成分を推定する共通スパース成分推定ステップと、

前記共通スパース成分の位相を復元して目的音響複素スペクトログラムとする位相復元ステップと、

前記目的音響複素スペクトログラムを時間信号である前記目的音響信号に変換する目的音響信号変換ステップとからなる目的音響信号復元方法。

10

【請求項 6】

前記共通スパース成分推定ステップでは、前記 M チャンネルの振幅スペクトログラムからそれぞれ反復推定 $i - 1$ 回目の低ランク成分を除いて得た M 個のスパース成分を含む残余の総和を、前記共通スパース成分に前記 M 個のスパース成分が含まれている含有比率の総和で除算して得た結果を反復推定 i 回目の前記共通スパース成分として推定する請求項 5 に記載の目的音響信号復元方法。

【請求項 7】

前記共通スパース成分推定ステップは、

前記 M チャンネルの振幅スペクトログラムに含まれる低ランク成分の比率を演算する低ランク成分比率演算ステップと、

20

前記低ランク成分の比率に基づいて前記 M チャンネルの振幅スペクトログラムに含まれる M 個の低ランク成分を演算する低ランク成分演算ステップと、

前記 M チャンネルの振幅スペクトログラムに含まれるスパース成分の比率を演算するスパース成分比率演算ステップと、

前記スパース成分の比率に基づいて前記 M チャンネルの振幅スペクトログラムに含まれる M 個のスパース成分を含む残余成分を演算する残余成分演算ステップと、

前記 M 個のスパース成分を含む残余成分と前記共通スパース成分とに基づいて、前記共通スパース成分に前記 M 個のスパース成分が含まれている含有比率を前記 M 個のスパース成分の音量として演算する音量演算ステップと、

30

前記 M 個のスパース成分を含む残余成分の総和を前記 M 個のスパース成分の音量の総和で除算して前記共通スパース成分を演算する共通スパース成分演算ステップを備えて、

前記低ランク成分比率演算ステップ、前記低ランク成分演算ステップ、前記スパース成分比率演算ステップ、前記残余成分演算ステップ、前記音量演算ステップ及び前記共通スパース成分演算ステップにおいて反復演算を行うことにより前記共通スパース成分を推定する請求項 6 に記載の音声信号復元方法。

【請求項 8】

M 本 (M は 2 以上の整数) のマイクロホンによって採取した M チャンネルの音響信号に含まれる目的音響信号を復元する目的音響信号復元方法を、コンピュータを用いて実現するためにコンピュータ読み取り可能な記憶手段に記憶されたコンピュータプログラムであって、

40

前記 M 本 (M は 2 以上の整数) のマイクロホンによって採取した前記 M チャンネルの音響信号を時間周波数解析して M チャンネルの複素スペクトログラムを得る時間周波数解析ステップと、

前記 M チャンネルの複素スペクトログラムから M チャンネルの振幅スペクトログラムを抽出する振幅成分抽出ステップと、

前記 M チャンネルの振幅スペクトログラムを入力として、前記 M チャンネルの振幅スペクトログラムのうち最も多くのチャンネルの振幅スペクトログラムに共通に含まれている可能性が高いスパース時間周波数成分を含む共通スパース成分を推定する共通スパース成分推定ステップと、

50

前記共通スパー成分の位相を復元して目的音響複素スペクトログラムとする位相復元ステップと、

前記目的音響複素スペクトログラムを時間信号である前記目的音響信号に変換する目的音響信号変換ステップを前記コンピュータで実現するための目的音響信号復元用コンピュータプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、低ランク性雑音により妨害された目的音響信号を、観測した複数チャンネルの音響信号から復元する目的音響信号復元システム及び方法に関するものである。

10

【背景技術】

【0002】

非特許文献1には、雑音と目的音響信号の混合音を観測した多チャンネル音響信号から、目的音響信号のみを分離できる技術が開示されている。

【0003】

非特許文献2には、複数のマイクロホンによって採取した複数チャンネルの音響信号に含まれる音声信号を復元する技術が開示されている。具体的には、マイクロホンアレイを搭載した柔軟索状ロボットで収録した音響信号の音声強調のために、RPCA（ロバスト主成分分析）をマイクロホンアレイに適用する雑音抑圧法が開示されている。この技術では、まず各チャンネルにそれぞれRPCAを適用し、各チャンネルの共通成分を抽出するために中央値を取って統合している。

20

【0004】

また非特許文献3に記載の技術では、雑音の事前情報無しで雑音の低ランク性と音声のスパー成分性から音声強調を行う。

【0005】

さらに特許文献1（特許第5752324号公報）に記載の技術では、単チャンネル音響信号からインパルス性（突発的）雑音を除去する。この技術は、インパルス性雑音を除去する性能は高いが、一方で想定しない持続的な非定常雑音（低ランク雑音）では性能が劣化する。

【0006】

30

また特許文献2（特開2009-116275公報）に記載の技術では、単チャンネル音響信号から平均二乗誤差最小法（MMSE）に基づいて雑音を抑圧する。MMSEは、雑音の定常性を仮定するため非定常雑音の抑圧では性能が劣化する。

【0007】

特許文献3（特開2014-503849公報）に記載の技術では、雑音源の近くに子機マイクを配置し、本マイクの情報を積極的に利用して音声強調を行う技術が開示されている。本技術では、雑音源の位置が特定されており、また子機マイクをその雑音源近くに配置する必要がある。

【0008】

特許文献4（特開2015-095897公報）には、ビデオ信号に対し低ランク成分とスパー成分を抽出することで、背景映像と移動する物体の映像を分離する技術が開示されている。本手法を音声強調へ応用することは可能であるが、一部のマイクが障害物等で音声を十分録音出来なかったときに性能が大きく劣化する。

40

【0009】

特許文献5（特開2014-058399公報）に記載の技術では、任意の数の音源信号の混合音を観測した多チャンネル音響信号から各音源信号を分離抽出する。本技術では、各マイクロホンの位置および音源位置が固定であると仮定されており、これらが動く場合性能が劣化する。

【先行技術文献】

【非特許文献】

50

【 0 0 1 0 】

【非特許文献1】高田一真，北村大地，中嶋広明，小山翔一，猿渡洋，小野順貴，牧野昭二，"教師あり多チャンネル NMF と統計的音声強調を用いた柔軟索状ロボットにおける音源分離，"日本音響学会 2016年春季研究発表会（2016年3月9日～11日），3-3-2

【非特許文献2】坂東宜昭，糸山克寿，昆陽雅司，田所諭，中臺一博，吉井和佳 及び奥乃博「ロバスト主成分分析を用いた動作雑音抑圧に基づく柔軟索状ロボットのための音声強調」第33回日本ロボット学会学術講演会（2015年9月3日～5日）

【非特許文献3】C. Sun, Q. Zhang, G. Wang, & J. Xie, "Noise reduction based on robust principal component analysis," Journal of Computational Information Systems, vol. 10, no. 10, pp. 4403-4410, 2014

10

【特許文献】

【 0 0 1 1 】

【特許文献1】特許第5752324号公報

【特許文献2】特開2009-116275公報

【特許文献3】特開2014-503849公報

【特許文献4】特開2015-095897公報

【特許文献5】特開2014-058399公報

【発明の概要】

【発明が解決しようとする課題】

【 0 0 1 2 】

20

しかしながら、非特許文献1に示された技術では、雑音を抑圧し音声を強調するには事前に雑音の音色情報の収録が不可欠であり、雑音がシステムの使用環境に依存して変化する場合などでは使用が困難だった。

【 0 0 1 3 】

また非特許文献2に示された技術では、多チャンネル音響信号から雑音の低ランク性と音声のスパース性を用いて目的音声を強調する。本技術では、各チャンネルの振幅スペクトログラムに対し個別にロバスト主成分分析を用いて低ランク成分とスパース成分の分離を行い、その後マイクロホンごとのスパース成分について各時間周波数点で中央値を選択して音声を強調していた。本技術では、複数のマイクロホンが中央値で全マイクの信号を統合するため、一部のマイクが障害物等で音声を十分録音出来なかったときに性能が大きく劣化する問題があった。

30

【 0 0 1 4 】

また非特許文献3に示された従来技術では、実数値行列の解析に特化しており、音響信号の振幅スペクトログラムである非負値行列の解析には不向きで、音響信号処理のための多チャンネル拡張や信頼度推定機能の実現が困難だった。

【 0 0 1 5 】

従来、多チャンネルのマイクロホンにより音声を集音した場合に、一部のマイクロホンが障害物等で音声を十分大きな音量で収録できない場合でも頑健に音声を強調できる技術は提案されていない。例えば、瓦礫の狭い隙間に侵入し被災者を捜索する柔軟索状レスキューロボットでは、自身の動作雑音により被災者の声を聞き取りづらくなる問題があった。

40

【 0 0 1 6 】

本発明の目的は、事前情報を用いずに雑音が含まれる音響信号から目的音響信号を高い精度で復元することができる音声信号復元システム及び方法を提供することにある。

【課題を解決するための手段】

【 0 0 1 7 】

本発明は、M本（Mは2以上の整数）のマイクロホンによって採取したMチャンネルの音響信号に含まれる目的音響信号を復元する目的音響信号復元システムを対象として、時間周波数解析部と、振幅成分抽出部と、共通スパース成分推定部と、位相復元部と、目的音響信号復元部とを備えている。時間周波数解析部は、M本のマイクロホンによって採取したMチャンネルの音響信号を時間周波数解析してMチャンネルの複素スペクトログラムを得る

50

。振幅成分抽出部は、Mチャンネルの複素スペクトログラムからMチャンネルの振幅スペクトログラムを抽出する。共通スパース成分推定部は、Mチャンネルの振幅スペクトログラムを入力として、Mチャンネルの振幅スペクトログラムのうち最も多くのチャンネルの振幅スペクトログラムに共通に含まれている可能性が高いスパース周波数成分を含む共通スパース成分を推定する。ここで「Mチャンネルの振幅スペクトログラムのうち最も多くのチャンネルの振幅スペクトログラム」の代表的な例としては、M本のマイクロホンのうち例えば2本のマイクロホンが音響信号を殆ど採取していなかったとすると、この2本のマイクロホンからの音響信号から得た2チャンネルの振幅スペクトログラムは、「最も多くのチャンネルの振幅スペクトログラム」には含まれない。そして残りのM-2本のマイクロホンが採取した音響信号から得たM-2チャンネルの音響信号から得たM-2チャンネルの振幅スペクトログラムが「最も多くのチャンネルの振幅スペクトログラム」になる。位相復元部では、共通スパース成分の位相を復元して目的音響複素スペクトログラムとする。位相は、Mチャンネルの振幅スペクトログラム、共通スパース成分等から推定すればよく、位相を求める方法は任意である。そして目的音響信号変換部は、目的音響複素スペクトログラムを時間信号である目的音響信号に変換する。

10

20

30

40

50

【0018】

本発明において、共通スパース成分推定部では、「各チャンネルの音響信号が低ランク成分とチャンネル間共通の共通スパース成分に分解される」と仮定して共通スパース成分を推定し、この共通スパース成分の位相を復元し、且つ復元した目的音響複素スペクトログラムを目的音響信号に変換することにより、雑音抑圧を行う。そして共通スパース成分は、共通スパース成分に含まれている各チャンネルのスパース成分の含有比率（本願明細書では、この含有比率を「音量」ということがある。）のみ違うものとモデル化することにより、この含有比率（音量）の推定により目的音響信号を十分に収録できないマイクロホンが存在しても頑強な目的音響信号強調を実現する。

【0019】

具体的に、本発明では、雑音の事前情報無しで雑音の低ランク性と目的音響信号のスパース性から音声強調を行う。本発明においては、Mチャンネルの振幅スペクトログラムのうち最も多くのチャンネルの振幅スペクトログラムに共通に含まれている可能性が高いスパース周波数成分を含む共通スパース成分を推定するため、低ランク成分の影響を極力受けることなく目的音響信号を復元することができる。その結果、復元の精度を従来よりも高くすることができる。

【0020】

本発明では、非負実数値行列である振幅スペクトログラムの解析に特化しており、マイクロホンの配置などのシステムの使用環境に依存せず、マイクロホンアレイの周囲に障害物が多い環境や、事前に雑音に関する情報を得られない場合でも頑健に動作する目的音響信号強調を実現できる。さらに一部のマイクロホンが障害物等で音声を十分大きな音量で収録できない場合でも頑健に目的音響信号を強調できる。例えば、瓦礫の狭い隙間に侵入し被災者を捜索する柔軟索状レスキューロボットでは、自身の動作雑音により被災者の声を聞き取りづらくなる問題があったが、本発明により瓦礫内でも頑健に音声を強調することができる。

【0021】

共通スパース成分推定部は、Mチャンネルの振幅スペクトログラムからそれぞれ反復推定 $i-1$ 回目の低ランク成分を除いて得たM個のスパース成分を含む残余の総和を、共通スパース成分にM個のスパース成分が含まれている含有比率（音量）の総和で除算して得た結果を反復推定 i 回目の共通スパース成分として推定する。この含有比率は、変分ベイズEM法や逐次変分ベイズEM法等の反復推定法を用いた反復推定の過程で徐々に収束することになる。

【0022】

反復推定法を用いる場合の共通スパース成分推定部は、Mチャンネルの振幅スペクトログラムに含まれる低ランク成分の比率を演算する低ランク成分比率演算部と、低ランク成分

の比率に基づいてMチャンネルの振幅スペクトログラムに含まれるM本の低ランク成分を演算する低ランク成分演算部と、Mチャンネルの振幅スペクトログラムに含まれるスパース成分の比率を演算するスパース成分比率演算部と、スパース成分の比率に基づいてMチャンネルの振幅スペクトログラムに含まれるM個のスパース成分を含む残余成分を演算する残余成分演算部と、M個のスパース成分を含む残余成分と共通スパース成分とに基づいて、共通スパース成分にM個のスパース成分が含まれている含有比率をM個のスパース成分の音量として演算する音量演算部と、M個のスパース成分を含む残余成分の総和をM個のスパース成分の音量の総和で除算して共通スパース成分を演算する共通スパース成分演算部とを備える。そして低ランク成分比率演算部、低ランク成分演算部、スパース成分比率演算部、残余成分演算部、音量演算部及び共通スパース成分演算部において反復演算を行うことにより共通スパース成分を推定する。このような構成で、反復推定法を用いると、適宜の反復演算を所定回数繰り返すことにより、高い精度で共通スパース成分を推定することができる。

10

【0023】

共通スパース成分推定部は、例えば、変分ベイズEM法や逐次変分ベイズEM法により共通スパース成分をベイズ推定するベイズ推定器によって構成することができる。ベイズ推定器を用いると、簡単に共通スパース成分を推定することができる。

【0024】

本発明は、コンピュータを用いて、M本(Mは2以上の整数)のマイクロホンによって採取したMチャンネルの音響信号に含まれる目的音響信号を復元する目的音響信号復元方法としても特定することができる。この方法では、M本のマイクロホンによって採取したMチャンネルの音響信号を時間周波数解析してMチャンネルの複素スペクトログラムを得る時間周波数解析ステップと、Mチャンネルの複素スペクトログラムからMチャンネルの振幅スペクトログラムを抽出する振幅成分抽出ステップと、Mチャンネルの振幅スペクトログラムを入力として、前記Mチャンネルの振幅スペクトログラムのうち最も多くのチャンネルの振幅スペクトログラムに共通に含まれている可能性が高いスパース時間周波数成分を含む共通スパース成分を推定する共通スパース成分推定ステップと、共通スパース成分の位相を復元して目的音響複素スペクトログラムとする位相復元ステップと、目的音響複素スペクトログラムを時間信号である目的音響信号に変換する目的音響信号変換ステップとがコンピュータで実施される。

20

30

【0025】

共通スパース成分推定ステップでは、Mチャンネルの振幅スペクトログラムからそれぞれ反復推定*i*-1回目の低ランク成分を除いて得たM個のスパース成分を含む残余の総和を、共通スパース成分にM個のスパース成分が含まれている含有比率の総和で除算して得た結果を反復推定*i*回目の共通スパース成分として推定する。

【0026】

共通スパース成分推定ステップは、Mチャンネルの振幅スペクトログラムに含まれる低ランク成分の比率を演算する低ランク成分比率演算ステップと、低ランク成分の比率に基づいてMチャンネルの振幅スペクトログラムに含まれるM個の低ランク成分を演算する低ランク成分演算ステップと、Mチャンネルの振幅スペクトログラムに含まれるスパース成分の比率を演算するスパース成分比率演算ステップと、スパース成分の比率に基づいてMチャンネルの振幅スペクトログラムに含まれるM個のスパース成分を含む残余成分を演算する残余成分演算ステップと、M個のスパース成分を含む残余成分と共通スパース成分とに基づいて、共通スパース成分にM個のスパース成分が含まれている含有比率をM個のスパース成分の音量として演算する音量演算ステップと、M個のスパース成分を含む残余成分の総和を前記M個のスパース成分の音量の総和で除算して共通スパース成分を演算する共通スパース成分演算ステップとからなる。そして低ランク成分比率演算ステップ、低ランク成分演算ステップ、スパース成分比率演算ステップ、残余成分演算ステップ、音量演算ステップ及び共通スパース成分演算ステップにおいて反復演算を行うことにより共通スパース成分を推定する。

40

50

【 0 0 2 7 】

本発明は、M本（Mは2以上の整数）のマイクロホンによって採取したMチャンネルの音響信号に含まれる目的音響信号を復元する目的音響信号復元方法を、コンピュータを用いて実現するためにコンピュータ読み取り可能な記憶手段に記憶されたコンピュータプログラムとしても特定することができる。このコンピュータプログラムは、M本のマイクロホンによって採取したMチャンネルの音響信号を時間周波数解析してMチャンネルの複素スペクトログラムを得る時間周波数解析ステップと、Mチャンネルの複素スペクトログラムからMチャンネルの振幅スペクトログラムを抽出する振幅成分抽出ステップとMチャンネルの振幅スペクトログラムを入力として、Mチャンネルの振幅スペクトログラムのうち最も多くのチャンネルの振幅スペクトログラムに共通に含まれている可能性が高いスパース時間周波数成分を含む共通スパース成分を推定する共通スパース成分推定ステップと、共通スパース成分の位相を復元して目的音響複素スペクトログラムとする位相復元ステップと、目的音響複素スペクトログラムを時間信号である目的音響信号に変換する目的音響信号変換ステップとをコンピュータで実現する。

10

【 図面の簡単な説明 】

【 0 0 2 8 】

【 図 1 】 本発明の目的音響信号復元システムの実施の形態の一例の構成を示すブロック図である。

【 図 2 】 図 1 の共通スパース成分推定部を、コンピュータを用いて反復推定法により実現する場合に用いるコンピュータプログラムのアルゴリズムを示すフローチャートである。

20

【 図 3 】 マイクロホンアレイを搭載した柔軟索状ロボットの写真である。

【 図 4 】 ロボットと音声を再生するスピーカ（音源）の配置条件 1 と条件 2 を説明するために用いる図である。

【 図 5 】 (A) 及び (B) は、条件 1 と条件 2 の配置条件及びSNR 条件での音声強調性能を信号対歪比(SDR)で示す図である。

【 図 6 】 多チャンネルの振幅スペクトログラムの 8 チャンネル中 4 チャンネルを抜粋した図である。

【 図 7 】 8 チャンネル中 4 チャンネルの低ランク成分 L_m を抜粋した図である。

【 図 8 】 8 チャンネル中 4 チャンネルの共通スパース成分での音量 g_m を示した図である。

【 図 9 】 共通スパース成分を示す図である。

30

【 図 1 0 】 MNMF での強調結果を示す図である。

【 図 1 1 】 Med-RPCA での強調結果を示す図である。

【 図 1 2 】 RPCA での強調結果を示す図である。

【 図 1 3 】 共通スパース成分の位相を復元して、目的音響複素スペクトログラムとし、この目的音響複素スペクトログラムを時間信号である目的音響信号に変換した結果を示す図である。

【 発明を実施するための形態 】

【 0 0 2 9 】

以下図面を参照して本発明の実施の形態を詳細に説明する。

【 0 0 3 0 】

40

（実施の形態の構成）

図 1 は、コンピュータまたは複数のプロセッサと複数のメモリを用いて実現される本発明の目的音響信号復元システムの実施の形態の一例の構成を示すブロック図である。本実施の形態では、具体的に、低ランク性雑音により妨害された音声を観測した多チャンネル音響信号から、目的音響信号として音声信号を抽出する。そして各マイクロホンが移動し、一部のマイクロホンが障害物により音声を十分大きな音量で観測出来なくても、頑健に音声信号を抽出することを可能にする。本実施の形態の目的音響信号復元システム 1 は、M本（Mは2以上の整数）のマイクロホンによって採取したMチャンネルの音響信号に含まれる目的音響信号を復元する。

【 0 0 3 1 】

50

本実施の形態の目的音響信号復元システム 1 は、コンピュータによってそれぞれ実現されるか、1 以上のプロセッサと 1 以上のメモリを用いてそれぞれ実現される、時間周波数解析部 3 と、振幅成分抽出部 5 と、共通スパース成分推定部 7 と、位相復元部 9 と、位相成分抽出部 11 と、目的音響信号変換部 13 とを備えている。時間周波数解析部 3 は、例えば、非特許文献 2「ロバスト主成分分析を用いた動作雑音抑圧に基づく柔軟索状ロボットのための音声強調」に示された柔軟索状ロボットに設けられた M 本（M は 2 以上の整数）のマイクロホンによって採取した M チャンネルの音響信号を時間周波数解析して M チャンネルの複素スペクトログラムを得る。振幅成分抽出部 5 は、M チャンネルの複素スペクトログラムから M チャンネルの振幅スペクトログラムを抽出する。共通スパース成分推定部 7 は、M チャンネルの振幅スペクトログラムを入力として、M チャンネルの振幅スペクトログラムのうち最も多くのチャンネルの振幅スペクトログラムに共通に含まれている可能性が高いスパース周波数成分を含む共通スパース成分を推定する。位相復元部 9 では、共通スパース成分の位相を復元して目的音響複素スペクトログラムとする。本実施の形態では、位相情報を、時間周波数解析部 3 から位相成分抽出部 11 により抽出している。なお位相は、M チャンネルの振幅スペクトログラム、共通スパース成分等から推定すればよく、位相を求める方法は任意である。したがって本発明は、位相成分抽出部 11 を設けることに限定されるものではない。そして目的音響信号変換部 13 は、目的音響複素スペクトログラムを時間信号である目的音響信号に変換する。

10

20

30

40

50

【0032】

本実施の形態では、共通スパース成分推定部 7 で、反復推定法を用いる。そのため共通スパース成分推定部 7 は、低ランク成分比率演算部 71 と、低ランク成分演算部 72 と、スパース成分比率演算部 73 と、残余成分演算部 74 と、音量演算部 75 と、共通スパース成分演算部 76 とを備える。低ランク成分比率演算部 71 は、M チャンネルの振幅スペクトログラムに含まれる低ランク成分の比率を演算する。低ランク成分演算部 72 は、低ランク成分の比率に基づいて M チャンネルの振幅スペクトログラムに含まれる M 個の低ランク成分を演算する。スパース成分比率演算部 73 は、M チャンネルの振幅スペクトログラムに含まれるスパース成分の比率を演算する。そして残余成分演算部 74 は、スパース成分の比率に基づいて M チャンネルの振幅スペクトログラムに含まれる M 個のスパース成分を含む残余成分を演算する。音量演算部 75 は、M 個のスパース成分を含む残余成分と共通スパース成分とに基づいて、共通スパース成分に M 個のスパース成分が含まれている含有比率を M 個のスパース成分の音量として演算する。この含有比率は、例えば、変分ベイズ EM 法や逐次変分ベイズ EM 法等の反復推定法を用いる際に、反復推定の過程で求める。共通スパース成分演算部 76 は、M チャンネルの振幅スペクトログラムからそれぞれ反復推定 $i - 1$ 回目の低ランク成分を除いて得た M 個のスパース成分を含む残余の総和を、共通スパース成分に M 個のスパース成分が含まれている含有比率の総和で除算して得た結果を反復推定 i 回目の共通スパース成分として演算する。

【0033】

そして低ランク成分比率演算部 71、低ランク成分演算部 72、スパース成分比率演算部 73、残余成分演算部 74、音量演算部 75 及び共通スパース成分演算部 76 において反復演算を行うことにより共通スパース成分を推定する。

【0034】

本実施の形態では、共通スパース成分推定部 7 では、「各チャンネルの音響信号が低ランク成分とチャンネル間共通の共通スパース成分に分解される」と仮定して共通スパース成分を推定し、この共通スパース成分の位相を位相復元部 9 で復元し、復元した目的音響複素スペクトログラムを目的音響信号変換部 13 で変換することにより、雑音抑圧を行う。本実施の形態では、共通スパース成分は、共通スパース成分に含まれている各チャンネルのスパース成分の含有比率（音量）のみ違うものとモデル化することにより、この含有比率（音量）の推定により目的音響信号（音声等）を十分に収録できないマイクロホンが存在しても頑強な音声強調を実現する。

【0035】

具体的に、本発明では、雑音の事前情報無しで雑音の低ランク性と目的音響信号（音声等）のスパース性から目的音響信号（音声等）強調を行う。本発明においては、Mチャンネルの振幅スペクトログラムのうち最も多くのチャンネルの振幅スペクトログラムに共通に含まれている可能性が高いスパース周波数成分を含む共通スパース成分を推定するため、低ランク成分の影響を極力受けることなく目的音響信号を復元することができる。その結果、復元の精度を従来よりも高くすることができる。

【0036】

本実施の形態では、非負実数値行列の解析に特化しており、雑音信号の音色情報を事前に収録しなくても、雑音の低ランク性と音声等の目的音響信号のスパース性から頑健に目的音響信号強調を実現できる。本特徴により、マイクロホンアレイの周囲に障害物が多い環境や、事前に雑音に関する情報を得られない場合でも頑健に動作する目的音響信号（音声等）強調を実現できる。さらに一部のマイクロホンが障害物等で音声を十分大きな音量で収録できない場合でも頑健に目的音響信号（音声等）を強調できる。

10

【0037】

（生成モデルの説明）

本実施の形態で採用する生成モデルについて説明する。なお以下の説明では便宜上目的音響信号を音声信号ということがある。この生成モデルで扱う音声強調の問題は以下で定義される。

【0038】

入力：Mチャンネルの振幅スペクトログラム

20

【0039】

【数1】

$$\mathbf{Y}_m \in \mathbb{R}_+^{F \times T}$$

出力：目的音声の振幅スペクトログラム

【0040】

【数2】

$$\mathbf{S} \in \mathbb{R}_+^{F \times T}$$

30

ここで、

【0041】

【数3】

$$\mathbb{R}_+$$

は非負実数値を表す。また、FおよびTはそれぞれ周波数ビン数、時間フレーム数を表す。以下の説明では、チャンネルの振幅スペクトログラムのことを音響信号と呼び、fおよびtを周波数ビン、時間フレームのインデックスとする。

【0042】

40

本実施の形態では、音声信号 $\mathbf{S} = [s_1, \dots, s_T]$ とそのマイクロホンでの観測 $\mathbf{Y}'_m = [y'_{m1}, \dots, y'_{mT}]$ の関係は、周波数非依存時変線形変換であると仮定する。

【0043】

【数4】

$$y'_{mt} \approx g_{mt} s_t \quad (1)$$

ここで、

【0044】

50

【数5】

$$g_{mt} \in \mathbb{R}_+$$

は本線形変換の係数（各マイクロホンでの音声の音量：各マイクロホンにおけるスパース成分の含有比率）を表す。本音声伝達モデルを用いて、マイクロホンの観測 y_{mt} は以下のように分解されると仮定する。

【0045】

【数6】

$$y_{mt} \approx l_{mt} + g_{mt}s_t. \quad (2)$$

10

ここで、 $L_m = [l_{m1}, \dots, l_{mT}]$ および $S = [s_1, \dots, s_T]$ はそれぞれ各マイクロホンに混入した低ランク雑音と、音声を表すスパース成分を表す。本低ランク成分は更に K 個の基底 $W_m = [w_{m1}, \dots, w_{mK}]$ （基底行列）および、各基底の各時刻での音量 $H_m = [h_{m1}, \dots, h_{mT}]$ （アクティベーション行列）の積で表現する。

【0046】

【数7】

$$y_{mt} \approx W_m h_{mt} + g_{mt}s_t. \quad (3)$$

以降では、各チャンネルの低ランク成分の低ランクらしさ及び、共通スパース成分のスパースらしさをモデル化するためのベイズ生成モデルについて説明する。

20

【0047】

（尤度モデル）

本モデルでは入力振幅スペクトログラムの近似誤差をKullback-Leibler (KL) 擬距離を用いて評価する。ベイズ生成モデルではKL 擬距離の最小化は、Poisson 分布尤度の最大化に対応するので、本モデルでは尤度モデルを以下のように定義する。

【0048】

【数8】

$$p(\mathbf{Y}_m | \mathbf{W}_m, \mathbf{H}_m, \mathbf{G}_m, \mathbf{S}) =$$

$$\prod_{f,t} \mathcal{P} \left(y_{mft} \mid \sum_k w_{mfk} h_{mkt} + g_{mt} s_{ft} \right). \quad (4)$$

30

ここで、 $P(x | k)$ はパラメータ

【0049】

【数9】

$$k \in \mathbb{R}_+$$

40

を持つPoisson 分布を表す。（4）式において、 \mathbf{Y}_m は M チャンネルの振幅スペクトログラムの m 番目の振幅スペクトログラム、 \mathbf{H}_m は各基底の各時刻での音量、 \mathbf{W}_m は K 個の基底、 \mathbf{S} は共通スパース成分、 \mathbf{G}_m はスパース成分の各マイクロホンにおける含有比率である。 s_{ft} 、 g_{mt} 、 w_{mfk} 、 h_{mkt} はそれぞれ、 \mathbf{S} 、 \mathbf{G}_m 、 \mathbf{W}_m 、 \mathbf{H}_m の要素を表し、 y_{mft} は、観測の複素スペクトログラムを表す。

【0050】

（低ランク成分の事前分布）

低ランク成分の基底行列とアクティベーション行列の事前分布は、Poisson 分布の共役事前分布であるGamma 分布を用いて定式化する。

【0051】

50

【数 1 0】

$$p(\mathbf{W}_m | \alpha^{wh}, \beta^{wh}) = \prod_{f,k} \mathcal{G}(w_{mfk} | \alpha^{wh}, \beta^{wh}) \quad (5)$$

$$p(\mathbf{H}_m | \alpha^{wh}, \beta^{wh}) = \prod_{k,t} \mathcal{G}(h_{mkt} | \alpha^{wh}, \beta^{wh}) \quad (6)$$

ここで、 $\mathcal{G}(x | \alpha, \beta)$ は shape パラメータ α および rate パラメータ β を持つガンマ分布を表す。また、

【0 0 5 2】

【数 1 1】

$$\alpha^{wh} \in \mathbb{R}_+$$

および

【0 0 5 3】

【数 1 2】

$$\beta^{wh} \in \mathbb{R}_+$$

は基底とアクティベーションの超パラメータを表す。本モデルでは、shape パラメータを 1 以下に設定することで基底とアクティベーション行列をスパースに制限でき、これによって低ランク成分 \mathbf{L} は低ランク行列に制限される。

【0 0 5 4】

(スパース成分の事前分布)

従来法の一つである Bayesian RPCA の生成モデルでは、スパース成分の表現に Gaussian 事前分布を置き、その精度パラメータに Jeffreys 超事前分布を置くことでスパース成分を表現していた。本実施の形態では、非負値行列である振幅スペクトログラムを表現するためにスパース成分の事前分布には Gamma 分布を置き、Gauss 分布における精度パラメータに対応する Gamma 分布の rate パラメータに Jeffreys 超事前分布を置くことでスパース成分をモデル化する。

【0 0 5 5】

【数 1 3】

$$p(\mathbf{S} | \alpha^s, \beta^s) = \prod_{f,t} \mathcal{G}(s_{ft} | \alpha^s, \beta_{ft}^s), \quad (7)$$

$$p(\beta_{ft}^s) \propto (\beta_{ft}^s)^{-1}. \quad (8)$$

ここで、

【0 0 5 6】

【数 1 4】

$$\alpha^s \in \mathbb{R}_+$$

はガンマ分布の超パラメータを表す。提案モデルでは、スパース成分のスパース度をこの超パラメータの値で調節する。

【0 0 5 7】

(音量変数の事前分布)

各マイクロホンのスパース成分の音量 (含有比率) g_{mt} には、Poisson 分布の共役事前

10

20

30

40

50

分布であるGamma 分布を置く。

【 0 0 5 8 】

【 数 1 5 】

$$p(g_{mt}|\alpha^g) = \mathcal{G}(g_{mt}|\alpha^g, \alpha^g) \quad (9)$$

ここで、 α^g は各マイクロホンでのスパース成分の音量のばらつきを調整する超パラメータである。

【 0 0 5 9 】

(変分ベイズEM 法によるベイズ推論)

入力多チャンネル振幅スペクトログラムが得られたときの本モデルの事後分布を解析的に導出することは困難なので、変分ベイズEM 法により近似推論を行う。以下では

【 0 0 6 0 】

【 数 1 6 】

⊖

を全てのパラメータの集合を表し、 $q(\cdot)$ を変分事後分布とする。変分近似では事後分布を以下のように分解近似し、真の事後分布とのKL 擬距離を最小化することで推論を行う。

【 0 0 6 1 】

【 数 1 7 】

$$p(\Theta|Y_{1:M}) \approx \left\{ \prod_m q(\mathbf{W}_m)q(\mathbf{H}_m)q(\mathbf{g}_m) \right\} q(\mathbf{S})q(\beta^s). \quad (10)$$

本実施の形態のモデルでは、共役指数分布族上でモデル化されているため、各更新則は Jensen の不等式とLagrange 未定乗数法を用いることで容易に導出できる。 \cdot を確率変数の平均とすると、各辺分事後分布は以下の更新則を他のパラメータを固定して反復更新することで得られる。

【 0 0 6 2 】

10

20

30

【数 18】

$$q(w_{mfk}) = \mathcal{G}(\alpha^{wh} + \sum_t y_{mft} \lambda_{mftk}^{wh}, \beta^{wh} + \sum_t \langle h_{mtk} \rangle) \quad (11)$$

$$q(h_{mtk}) = \mathcal{G}(\alpha^{wh} + \sum_f y_{mft} \lambda_{mftk}^{wh}, \beta^{wh} + \sum_f \langle w_{mfk} \rangle) \quad (12)$$

$$q(s_{ft}) = \mathcal{G}(\alpha^s + \sum_m s'_{mft}, \langle \beta_{ft}^s \rangle + \sum_m \langle g_{mt} \rangle) \quad (13)$$

$$q(g_{mt}) = \mathcal{G}(\alpha^g + \sum_f s'_{mft}, \alpha^g + \sum_f \langle s_{ft} \rangle) \quad (14)$$

$$q(\beta_{ft}^s) = \mathcal{G}(\alpha^s, \langle s_{ft} \rangle). \quad (15)$$

$$s'_{mft} = y_{mft} \lambda_{mft}^{gs} \quad (16)$$

$$\lambda_{mftk}^{wh} = \frac{\mathbb{G}[w_{mfk}] \mathbb{G}[h_{mtk}]}{\sum_k \mathbb{G}[w_{mfk}] \mathbb{G}[h_{mtk}] + \mathbb{G}[g_{mt}] \mathbb{G}[s_{ft}]} \quad (17)$$

$$\lambda_{mft}^{gs} = \frac{\mathbb{G}[g_{mt}] \mathbb{G}[s_{ft}]}{\sum_k \mathbb{G}[w_{mfk}] \mathbb{G}[h_{mtk}] + \mathbb{G}[g_{mt}] \mathbb{G}[s_{ft}]} \quad (18)$$

ここで、 s'_{mft} は低ランク成分の残余を表し、

【0063】

【数 19】

$$\lambda_{mftk}^{wh}$$

および

【0064】

【数 20】

$$\lambda_{mft}^{gs}$$

はそれぞれ低ランク成分が含まれている比率及びスパース成分が含まれている比率を表す。

【0065】

図2には、図1の共通スパース成分推定部7をコンピュータを用いて反復推定法により実現する場合に用いるコンピュータプログラムのアルゴリズムを示すフローチャートを示してある。図2には、上記(11)式~(18)式までが使用されるステップに、各式の表示を付してある。反復推定の終了条件は、200回繰り返し、または各推定値のデータ Y_m 、 H_m 、 W_m 、 S 、 s 、 g_m について、前回処理時との比較を行い、比較結果が近似になったら終了するものとした。なお図2においても、 Y_m はMチャンネルの振幅スペクトログラムのm番目の振幅スペクトログラム、 H_m は各基底の各時刻での音量、 W_m はK個の基底、 S は共通スパース成分、 s はベイズ推定の係数、 g_m はスパース成分の含有比率である。

【0066】

共通スパース成分演算部76で行う、Mチャンネルの振幅スペクトログラムからそれぞれ反復推定 $i-1$ 回目の低ランク成分を除いて得たM個のスパース成分を含む残余の総和の演算は、式(13)で更新される変分事後分布の平均を計算する上で、式(13)中の「 $\langle g_{mt} \rangle$ 」の前の式を用いて実施される。そして残余の総和を、共通スパース成分にM個のスパース成分が含まれている含有比率(音量)の総和で除算する演算は、式(13)で得られる変分事後分布の平均を計算するとき実施される。この結果を、反復推定 i 回目の共通スパース成分として推定する。含有比率の総和 [式(13)中の $\langle g_{mt} \rangle$] は、変分ベ

10

20

30

40

50

イズEM法等の反復推定法を用いる際に、反復推定の過程で徐々に収束することになる。式(13)中の s_{ft} はベイズ推定の係数である。

【0067】

図2の場合、終了条件が成立したら位相復元器9に、共通スパース成分Sが与えられ、位相復元器9は式(19)で目的音響スペクトログラム s'_{ft} (出力)を計算する。

【0068】

【数21】

$$s'_{ft} = s_{ft} \frac{y'_{mft}}{\|y'_{mft}\|} \quad (19)$$

10

なお上記式において、 s_{ft} はそれぞれ、Sの要素を表し、 y_{mft} は、観測の複素スペクトログラムの要素を表す。

【0069】

また位相復元部9は、共通スパース成分Sの他に、各基底の各時刻での音量 H_m 、K個の基底 W_m 、スパース成分の含有比率 g_m を含めて目的音響スペクトログラム s'_{ft} (出力)を復元してもよい。この場合、位相復元部9は式(20)で目的音響スペクトログラム s'_{ft} を計算する。

20

【0070】

【数22】

$$s'_{ft} = \frac{(g_{mt}s_{ft})^2}{(\sum_k w_{mfk}h_{mkt})^2 + (g_{mt}s_{ft})^2} y'_{mft} \quad (20)$$

なお上記式において、 s_{ft} 、 g_{mt} 、 w_{mfk} 、 h_{mkt} はそれぞれ、S、 g_m 、 W_m 、 H_m の要素を表し、 y_{mft} は、観測の複素スペクトログラムの要素を表す。

【0071】

[評価実験]

30

駆動機構とマイクロホンアレイを有する柔軟索状ロボットの動作雑音を用いて本実施の形態の音声強調性能を評価した。

【0072】

(使用した柔軟索状ロボット)

図3に、マイクロホンアレイを搭載した柔軟索状ロボットの写真を示す。本体は、直径38mmのコルゲートチューブからなり、全長3mである。8本のマイクロホンアレイ(M=8)をロボット表面に40cm間隔で90度ずつ回転して装着した。両端のマイクロホン間の距離は2.8mである。マイクロホンは手元から順番にインデックスmで区別する(m=1, ..., M)。本ロボットは、NamariらのTube-type Active Scope Camera [J. Fukuda, et al. Remote vertical exploration by active scope camera into collapsed buildings. In IEEE/RSJ IROS, pp. 1882-1888, 2014.]と同様、繊毛と振動モータを用いた駆動で前進する。振動モータはロボット内に40cm間隔で7個直列に装着されている。

40

【0073】

(実験設定)

(録音条件)

柔軟索状ロボットを用いて音声と動作雑音を個別に録音し、SNRを-20dBから+5dBまで5dB分ずつ変化させて混合し、音声の強調性能を評価した。図4に示すように、ロボットと音声を再生するスピーカ(音源)の配置条件を条件1と条件2のように定めた。

50

【 0 0 7 4 】

条件 1 : ロボットは自由空間に配置され、音源はロボットの正面に配置されている。部屋の残響時間(RT60) は750 ms だった。

【 0 0 7 5 】

条件 2 : ロボットはドアの隙間に配置され、音源はロボットの正面に配置されている。4つのマイクロホンがドアにより音源から遮られている。残響時間(RT60)は990 ms だった。

【 0 0 7 6 】

ロボットを駆動させ、手と振動モータを使って左右にロボット振りながら、60 秒の動作雑音を録音した。目的音である音声は、ノイズを軽減するために、ロボットが静止時のインパルス応答に60 秒の録音音声を畳み込んで作成した。録音音声は男声 2 種、女声 2 種の計 4 種(240 秒) を用いた。これらの録音は 8 c h 同期、2 4 bit 量子化、16 kHz サンプルングで行った。

【 0 0 7 7 】

(比較手法)

実験では、本実施の形態の実施例と、Multi-channel non-negative matrix factorization (MNMF) [D. Kitamura, et al. Efficient multichannel nonnegative matrix factorization exploiting rank-1 spatial model. In IEEE ICASSP, pp. 276-280, 2015.]及びrobust principal component analysis (RPCA) [C. Sun, et al. Noise reduction based on robust principal component analysis. JCIS, Vol. 10, No. 10, pp.4403-4410, 2014.2] による比較例とを比較した。RPCAは先端のマイクの結果を使用した。更に、全マイクのRPCA の結果を中央値で統合した結果(Med-RPCA) とも比較した。本実施の形態の実施例では、従来法[坂東宜昭ほか. ロバスト主成分分析を用いた動作雑音抑圧に基づく柔軟索状ロボットのための音声強調. In RSJ2015] のオフライン実装となっている。

【 0 0 7 8 】

(評価尺度)

評価尺度には、信号対歪比(SDR) を用いた。信号対歪比(SDR)は総合的な分離精度を表す。

【 0 0 7 9 】

(実験結果)

図 5 には、条件 1 と条件 2 の配置条件及びSNR 条件での音声強調性能を信号対歪比(SDR)で示してある。各配置・SNR 条件での音声強調性能をSDR で示した場合、SDR が高いほど、音声強調性能が良いことを表す。すなわちSNRが高いほど音声が多く含まれていることを意味する。条件では、SNR が 0 dB 以下のとき、条件 2 ではSNR が - 15 dB 以上で 0 dB 以下のときに本実施の形態(提案法)の実施例では性能が最も高い。これに対して、条件 1 及び条件 2 のいずれにおいても、2 番目に性能が高いMed-RPCA の比較例では、一部のマイクロホンが隠れている条件 2 では性能が大きく劣化している。一方、条件 2 で 3 番目に性能が高いRPCA は条件 2 では、Med-RPCA や提案法より性能が低い。これらに比べて、本実施の形態の実施例(提案法)は両方の条件で高い性能を示しており、環境への依存性が低いことがわかる。

【 0 0 8 0 】

図 6 乃至図 1 3 には、本実施の形態の実施例(提案法)による音声強調結果および、従来法による強調結果を波形で示す。これらの波形をみれば、本実施の形態の実施例が、最も雑音を抑圧し、音声を強調できていることがわかる。図 6 は、多チャンネルの振幅スペクトログラム $Y_m(m = 1, \dots, 8)$ の 8 チャンネル中 4 チャンネルを抜粋したものである。図 7 は、8 チャンネル中 4 チャンネルの低ランク成分 L_m を抜粋したものである。図 8 は、8 チャンネル中 4 チャンネルの共通スパー成分での音量 g_m を示したものである。図 9 は、共通スパー成分を示したものである。そして図 1 0 は、MNMF での強調結果を示しており、図 1 1 はMed-RPCA での強調結果を示しており、図 1 2 はRPCA での強調結果を示している。そして図 1 3 は、共通スパー成分の位相を復元して、目的音響複素スペクトログ

10

20

30

40

50

ラムとし、この目的音響複素スペクトログラムを時間信号である目的音響信号に変換した結果を示している。

【産業上の利用可能性】

【0081】

本発明では、雑音の事前情報無しで雑音の低ランク性と目的音響信号のスパース性から音声強調を行うため、Mチャンネルの振幅スペクトログラムのうち最も多くのチャンネルの振幅スペクトログラムに共通に含まれている可能性が高いスパース周波数成分を含む共通スパース成分を推定する。そして共通スパース成分の位相を復元し、且つ復元した目的音響複素スペクトログラムを目的音響信号に変換することにより、雑音抑圧を行う、そのため低ランク成分の影響を極力受けることなく目的音響信号を復元することができ、復元の精度を従来よりも高くすることができる。

10

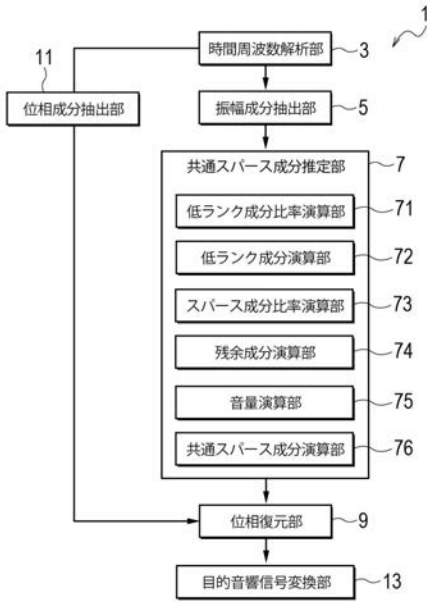
【符号の説明】

【0082】

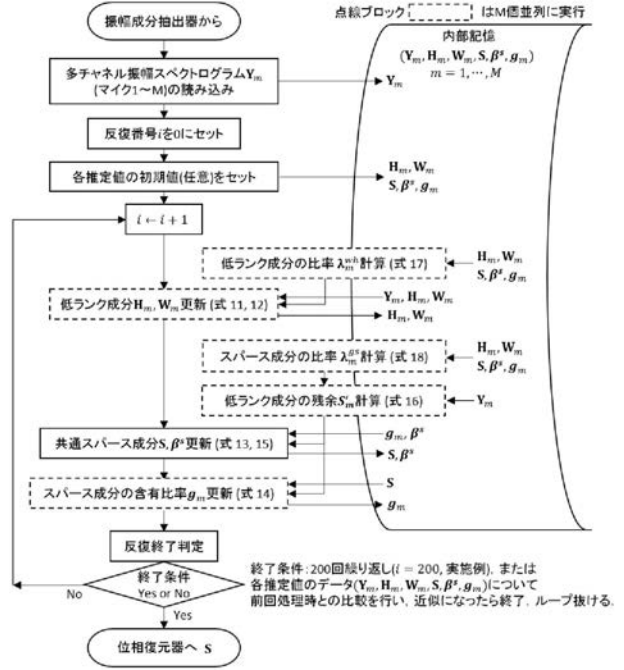
- 1 目的音響信号復元システム
- 3 時間周波数解析部
- 5 振幅成分抽出部
- 7 共通スパース成分推定部
- 9 位相復元部
- 11 位相成分抽出部
- 13 目的音響信号変換部
- 71 低ランク成分比率演算部
- 72 低ランク成分演算部
- 73 スパース成分比率演算部
- 74 残余成分演算部
- 75 音量演算部
- 76 共通スパース成分演算部

20

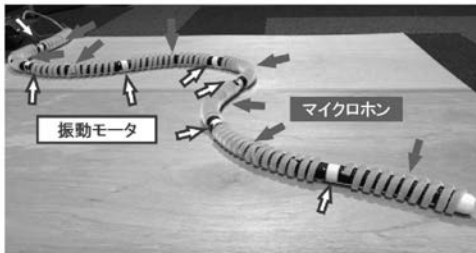
【 図 1 】



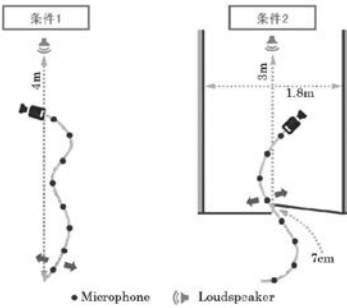
【 図 2 】



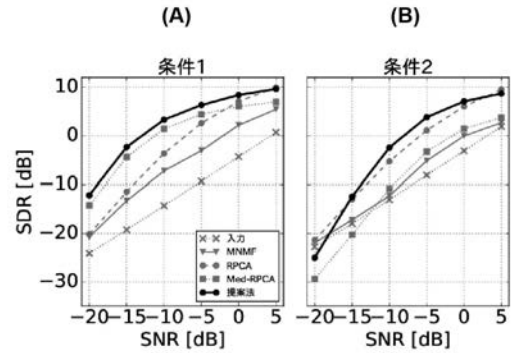
【 図 3 】



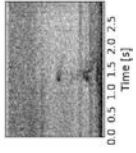
【 図 4 】



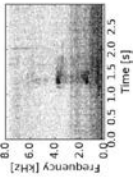
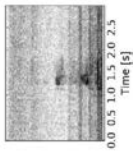
【 図 5 】



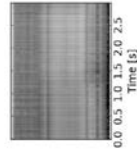
【 図 6 】



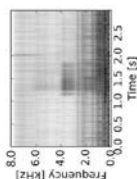
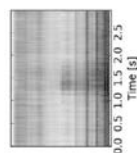
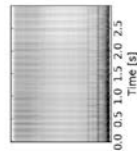
...



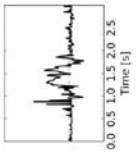
【 図 7 】



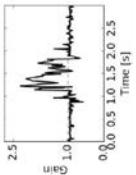
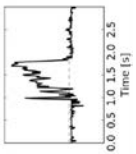
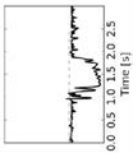
...



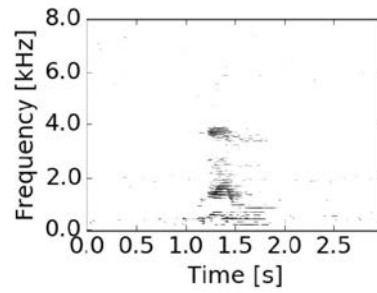
【 図 8 】



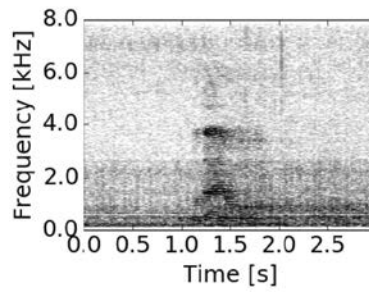
...



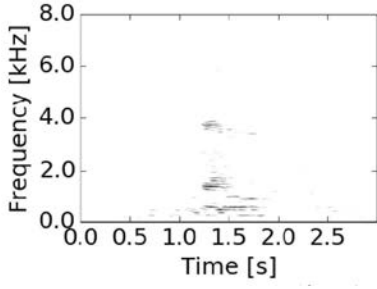
【 図 9 】



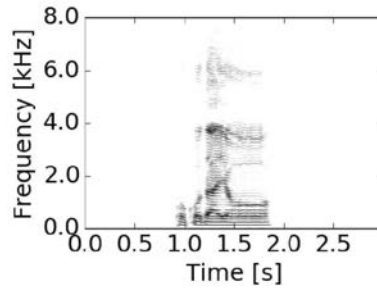
【 図 10 】



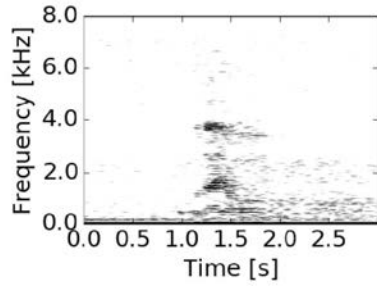
【 1 1 】



【 1 3 】



【 1 2 】



【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		International application No. PCT/JP2017/019259
A. CLASSIFICATION OF SUBJECT MATTER G10L21/0264(2013.01)i, G10L21/0208(2013.01)i 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) G10L21/0264, G10L21/0208 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Jitsuyo Shinan Koho 1922-1996 Jitsuyo Shinan Toroku Koho 1996-2017 Kokai Jitsuyo Shinan Koho 1971-2017 Toroku Jitsuyo Shinan Koho 1994-2017 Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P, X	Yoshiaki BANDO et al., "Henbun Bayes Ta-Channel RNMF ni Motozuku Junan Sakujo Rescue Robot no Tameno Onsei Kyocho", the 34th Annual Conference of the Robotics Society of Japan Yokoshu, The Robotics Society of Japan, 07 September 2016 (07.09.2016)	1-8
X A	Yoshiaki BANDO et al., "Junan Sakujo Rescue Robot no Tameno Robust Shuseibun Bunseki o Mochiita Soko Zatsuon Yokuatsu", Proceedings of the 77th National Convention of Information Processing Society of Japan (2), 17 March 2015 (17.03.2015), pages 2-505 to 2-506	1, 5, 8 2-4, 6-7
<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 08 August 2017 (08.08.17)		Date of mailing of the international search report 22 August 2017 (22.08.17)
Name and mailing address of the ISA/ Japan Patent Office 3-4-3, Kasumigaseki, Chiyoda-ku, Tokyo 100-8915, Japan		Authorized officer Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2017/019259

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2015/060375 A1 (Nagasaki University, National University Corp.), 30 April 2015 (30.04.2015), entire text; all drawings (Family: none)	1-8

国際調査報告		国際出願番号 PCT/J P 2 0 1 7 / 0 1 9 2 5 9									
A. 発明の属する分野の分類 (国際特許分類 (IPC)) Int.Cl. G10L21/0264(2013.01)i, G10L21/0208(2013.01)i											
B. 調査を行った分野 調査を行った最小限資料 (国際特許分類 (IPC)) Int.Cl. G10L21/0264, G10L21/0208											
最小限資料以外の資料で調査を行った分野に含まれるもの <table border="0"> <tr> <td>日本国実用新案公報</td> <td>1922-1996年</td> </tr> <tr> <td>日本国公開実用新案公報</td> <td>1971-2017年</td> </tr> <tr> <td>日本国実用新案登録公報</td> <td>1996-2017年</td> </tr> <tr> <td>日本国登録実用新案公報</td> <td>1994-2017年</td> </tr> </table>				日本国実用新案公報	1922-1996年	日本国公開実用新案公報	1971-2017年	日本国実用新案登録公報	1996-2017年	日本国登録実用新案公報	1994-2017年
日本国実用新案公報	1922-1996年										
日本国公開実用新案公報	1971-2017年										
日本国実用新案登録公報	1996-2017年										
日本国登録実用新案公報	1994-2017年										
国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)											
C. 関連すると認められる文献											
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号									
P, X	坂東 宜昭, 外7名, 変分ベイズ多チャネルRNMFに基づく柔軟 索状レスキューロボットのための音声強調, 第34回日本ロボット 学会学術講演会予稿集, 一般社団法人日本ロボット学会, 2016.09.07	1-8									
X A	坂東 宜昭, 外6名, 柔軟索状レスキューロボットのためのロバ スト主成分分析を用いた走行雑音抑圧, 情報処理学会第77回全国大 会講演論文集(2), 2015.03.17, pp. 2-505 - 2-506	1, 5, 8 2-4, 6-7									
<input checked="" type="checkbox"/> C欄の続きにも文献が列挙されている。 <input type="checkbox"/> パテントファミリーに関する別紙を参照。											
* 引用文献のカテゴリー		の日の後に公表された文献									
「A」特に関連のある文献ではなく、一般的技術水準を示すもの		「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの									
「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの		「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの									
「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)		「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの									
「O」口頭による開示、使用、展示等に言及する文献		「&」同一パテントファミリー文献									
「P」国際出願日前で、かつ優先権の主張の基礎となる出願											
国際調査を完了した日 08.08.2017		国際調査報告の発送日 22.08.2017									
国際調査機関の名称及びあて先 日本国特許庁 (ISA/J P) 郵便番号100-8915 東京都千代田区霞が関三丁目4番3号		特許庁審査官 (権限のある職員) 上田 雄	5Z 5095								
		電話番号 03-3581-1101 内線 3591									

国際調査報告		国際出願番号 PCT/JP2017/019259
C (続き) . 関連すると認められる文献		
引用文献の カテゴリ*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
A	WO 2015/060375 A1 (国立大学法人 長崎大学) 2015.04.30, 全文, 全図 (ファミリーなし)	1-8

フロントページの続き

(81)指定国 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, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, 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, NE, 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, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ

(出願人による申告)革新的研究開発推進プログラム(IMPACT)「極限環境での音環境理解に向けたロボット聴覚技術の実用化」委託研究、産業技術力強化法第19条の適用を受ける特許出願

(72)発明者 奥乃 博
東京都新宿区河田町3-51-3010

(注)この公表は、国際事務局(WIPO)により国際公開された公報を基に作成したものである。なおこの公表に係る日本語特許出願(日本語実用新案登録出願)の国際公開の効果は、特許法第184条の10第1項(実用新案法第48条の13第2項)により生ずるものであり、本掲載とは関係ありません。