

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号
特許第5410603号
(P5410603)

(45) 発行日 平成26年2月5日 (2014. 2. 5)

(24) 登録日 平成25年11月15日 (2013. 11. 15)

(51) Int. Cl.	F I
HO 4 R 3/00 (2006. 01)	HO 4 R 3/00 3 2 O
G 1 O L 21/0332 (2013. 01)	G 1 O L 21/02 3 O 1 A
HO 4 R 1/10 (2006. 01)	HO 4 R 1/10 1 O 1 A
HO 4 R 1/40 (2006. 01)	HO 4 R 1/40 3 2 O A

請求項の数 31 (全 57 頁)

(21) 出願番号	特願2012-515105 (P2012-515105)	(73) 特許権者	595020643
(86) (22) 出願日	平成22年6月9日 (2010. 6. 9)		クォアルコム・インコーポレイテッド
(65) 公表番号	特表2012-529868 (P2012-529868A)		Q U A L C O M M I N C O R P O R A T E D
(43) 公表日	平成24年11月22日 (2012. 11. 22)		アメリカ合衆国、カリフォルニア州 9 2
(86) 国際出願番号	PCT/US2010/037973		1 2 1 - 1 7 1 4、サン・ディエゴ、モア
(87) 国際公開番号	W02010/144577		ハウス・ドライブ 5 7 7 5
(87) 国際公開日	平成22年12月16日 (2010. 12. 16)	(74) 代理人	100108855
審査請求日	平成24年2月9日 (2012. 2. 9)		弁理士 蔵田 昌俊
(31) 優先権主張番号	61/185, 518	(74) 代理人	100159651
(32) 優先日	平成21年6月9日 (2009. 6. 9)		弁理士 高倉 成男
(33) 優先権主張国	米国 (US)	(74) 代理人	100091351
(31) 優先権主張番号	61/227, 037		弁理士 河野 哲
(32) 優先日	平成21年7月20日 (2009. 7. 20)	(74) 代理人	100088683
(33) 優先権主張国	米国 (US)		弁理士 中村 誠

最終頁に続く

(54) 【発明の名称】 マルチチャネル信号の位相ベースの処理のためのシステム、方法、装置、およびコンピュータ可読媒体

(57) 【特許請求の範囲】

【請求項 1】

マルチチャネル信号を処理する方法であって、
前記マルチチャネル信号の複数の異なる周波数成分の各々に関して、複数の計算された位相差を取得するために、前記マルチチャネル信号の第 1 のチャンネルにおける前記周波数成分の位相と前記マルチチャネル信号の第 2 のチャンネルにおける前記周波数成分の位相との間の差を計算することと；
前記第 1 のチャンネルのレベルと前記第 2 のチャンネルの対応するレベルとを計算することと；
前記第 1 のチャンネルの前記計算されたレベルと前記第 2 のチャンネルの前記計算されたレベルと前記複数の計算された位相差の少なくとも 1 つとに基づいて、利得係数の更新された値を計算することと；及び
前記更新された値に従って、前記第 1 のチャンネルの対応する振幅に関して前記第 2 のチャンネルの振幅を修正することによって、処理されたマルチチャネル信号を生成することと；
を備える、マルチチャネル信号を処理する方法。

【請求項 2】

前記第 1 のチャンネルの前記計算されたレベルは第 1 の周波数サブバンドにおける前記第 1 のチャンネルの計算されたエネルギーであり、前記第 2 のチャンネルの前記計算されたレベルは前記第 1 の周波数サブバンドにおける前記第 2 のチャンネルの計算されたエネルギーで

あり、且つ

前記第1のチャンネルの前記振幅は前記第1の周波数サブバンドにおける前記第1のチャンネルの振幅であり、前記第2のチャンネルの前記対応する振幅は前記第1の周波数サブバンドにおける前記第2のチャンネルの振幅であり、且つ

前記方法は：

前記第1の周波数サブバンドとは異なる第2の周波数サブバンドにおける前記第1のチャンネルのエネルギーを計算することと；

前記第2の周波数サブバンドにおける前記第2のチャンネルのエネルギーを計算することと；及び

前記第2の周波数サブバンドにおける前記第1のチャンネルの前記計算されたエネルギーと、前記第2の周波数サブバンドにおける前記第2のチャンネルの前記計算されたエネルギーと、前記複数の計算された位相差の少なくとも1つとに基づいて、第2の利得係数の更新された値を計算することと；

を備えており、

処理されたマルチチャンネル信号を前記生成することは、前記第2の周波数サブバンドにおける前記第1のチャンネルの振幅に関して前記第2の周波数サブバンドにおける前記第2のチャンネルの振幅を、前記第2の利得係数の前記更新された値に従って、修正することによって前記処理されたマルチチャンネル信号を生成することを含む、請求項1に記載のマルチチャンネル信号を処理する方法。

【請求項3】

前記方法は、前記複数の計算された位相差からの情報に基づいて、少なくとも前記複数の異なる周波数成分の到着方向間のコヒーレンスの度合いを示すコヒーレンシー測定値を計算することを備えており、且つ

利得係数の更新された値を前記計算することは、前記コヒーレンシー測定の前記計算された値に基づく、請求項1および2のいずれか一項に記載のマルチチャンネル信号を処理する方法。

【請求項4】

前記第2のチャンネルの対応する振幅に関して前記第1のチャンネルの振幅を前記修正することは、前記コヒーレンシー測定の前記値をある閾値と比較した結果に応じて実行される、請求項3に記載のマルチチャンネル信号を処理する方法。

【請求項5】

前記方法は、前記マルチチャンネル信号の推定されたピッチ周波数に基づいて前記複数の異なる周波数成分を選択することを含む、請求項1乃至4のいずれか一項に記載のマルチチャンネル信号を処理する方法。

【請求項6】

利得係数の前記更新された値は、前記第1のチャンネルの前記計算されたレベルと前記第2のチャンネルの前記計算されたレベルとの比に基づく、請求項1乃至5のいずれか一項に記載のマルチチャンネル信号を処理する方法。

【請求項7】

前記更新された値に従って、前記第1のチャンネルの対応する振幅に関して前記第2のチャンネルの振幅を修正することによって、処理されたマルチチャンネル信号を前記生成することは、前記第1および第2のチャンネルの前記計算されたレベル間の不均衡を低減することを備える、請求項1乃至6のいずれか一項に記載のマルチチャンネル信号を処理する方法。

【請求項8】

処理されたマルチチャンネル信号を前記生成することは、前記更新された値に従って、前記マルチチャンネル信号の複数の連続するセグメントの各々において前記第1のチャンネルの対応する振幅に関して前記第2のチャンネルの振幅を修正することを含む、請求項1乃至7のいずれか一項に記載のマルチチャンネル信号を処理する方法。

【請求項9】

前記方法は、前記処理されたマルチチャンネル信号の第1のチャンネルのレベルと前記処理

10

20

30

40

50

されたマルチチャネル信号の第2のチャンネルのレベルとの間の関係に基づいて音声活動の存在を示すことを備える、請求項1乃至8のいずれか一項に記載のマルチチャネル信号を処理する方法。

【請求項10】

前記方法は、前記処理されたマルチチャネル信号の第1のチャンネルのレベルと前記処理されたマルチチャネル信号の第2のチャンネルのレベルとの間の関係に基づいて、及び、前記コヒーレンシー測定の前記値をある閾値と比較した結果に応じて、前記マルチチャネル信号の前記第1および第2のチャンネルの少なくとも1つからの音響情報に従って雑音推定値を更新することを備える、請求項3および4のいずれか一項に記載のマルチチャネル信号を処理する方法。

10

【請求項11】

プロセッサによって読み取られたときに請求項1乃至10のいずれか一項による方法を前記プロセッサに実行させるプログラムを記録したコンピュータ可読媒体。

【請求項12】

マルチチャネル信号を処理するための装置であって、

前記マルチチャネル信号の複数の異なる周波数成分の各々に関して、前記マルチチャネル信号の第1のチャンネルにおける前記周波数成分の位相と前記マルチチャネル信号の第2のチャンネルにおける前記周波数成分の位相との間の差を計算することによって複数の計算された位相差を取得するように構成された第1の計算器と；

前記第1のチャンネルのレベルと前記第2のチャンネルの対応するレベルとを計算するように構成された第2の計算器と；

20

前記第1のチャンネルの前記計算されたレベルと前記第2のチャンネルの前記計算されたレベルと前記複数の計算された位相差の少なくとも1つとに基づいて、利得係数の更新された値を計算するように構成された第3の計算器と；及び

前記更新された値に従って前記第1のチャンネルの対応する振幅に関して前記第2のチャンネルの振幅を修正することによって、処理されたマルチチャネル信号を生成するように構成された利得制御要素と；

を備える、マルチチャネル信号を処理するための装置。

【請求項13】

前記第1のチャンネルの前記計算されたレベルは第1の周波数サブバンドにおける前記第1のチャンネルの計算されたエネルギーであり、前記第2のチャンネルの前記計算されたレベルは前記第1の周波数サブバンドにおける前記第2のチャンネルの計算されたエネルギーであり、且つ

30

前記第1のチャンネルの前記振幅は前記第1の周波数サブバンドにおける前記第1のチャンネルの振幅であり、前記第2のチャンネルの前記対応する振幅は前記第1の周波数サブバンドにおける前記第2のチャンネルの振幅であり、且つ

前記第2の計算器は、前記第1の周波数サブバンドとは異なる第2の周波数サブバンドにおける前記第1のチャンネルのエネルギーを計算するように、及び前記第2の周波数サブバンドにおける前記第2のチャンネルのエネルギーを計算するように構成され、且つ

前記第3の計算器は、前記第2の周波数サブバンドにおける前記第1のチャンネルの前記計算されたエネルギーと、前記第2の周波数サブバンドにおける前記第2のチャンネルの前記計算されたエネルギーと、前記複数の計算された位相差と、のうち少なくとも1つに基づいて、第2の利得係数の更新された値を計算するように構成され、

40

前記利得制御要素は、前記第2の利得係数の前記更新された値に従って、前記第2の周波数サブバンドにおける前記第1のチャンネルの振幅に関して前記第2の周波数サブバンドにおける前記第2のチャンネルの振幅を修正することによって前記処理されたマルチチャネル信号を生成するように構成されている、請求項12に記載の装置。

【請求項14】

前記第3の計算器は、前記複数の計算された位相差からの情報に基づいて少なくとも前記複数の異なる周波数成分の到着方向間のコヒーレンスの度合いを示すコヒーレンシー測

50

定の値を計算するように構成され、且つ

前記第3の計算器は、前記コヒーレンシー測定の前記計算された値に基づいて利得係数の前記更新された値を計算するように構成されている、請求項12および13のいずれか一項に記載の装置。

【請求項15】

前記第3の計算器は、前記コヒーレンシー測定の前記値をある閾値と比較するように構成され、且つ

前記利得制御要素は、前記コヒーレンシー測定の前記値をある閾値と前記比較した結果に応じて前記第2のチャンネルの対応する振幅に関して前記第1のチャンネルの振幅を修正するように構成されている、請求項14に記載の装置。

10

【請求項16】

前記位相差計算器は、前記マルチチャンネル信号の推定されたピッチ周波数に基づいて前記複数の異なる周波数成分を選択するように構成されている、請求項12乃至15のいずれか一項に記載の装置。

【請求項17】

利得係数の前記更新された値は、前記第1のチャンネルの前記計算されたレベルと前記第2のチャンネルの前記計算されたレベルとの比に基づく、請求項12乃至16のいずれか一項に記載の装置。

【請求項18】

前記利得制御要素は、前記更新された値に従って、前記第1のチャンネルの対応する振幅に関して前記第2のチャンネルの振幅を修正することによって前記第1および第2のチャンネルの前記計算されたレベル間の不均衡を低減するように構成されている、請求項12乃至17のいずれか一項に記載の装置。

20

【請求項19】

前記利得制御要素は、前記更新された値に従って、前記マルチチャンネル信号の複数の連続するセグメントの各々において前記第1のチャンネルの対応する振幅に関して前記第2のチャンネルの振幅を修正することによって前記処理されたマルチチャンネル信号を生成するように構成されている、請求項12乃至18のいずれか一項に記載の装置。

【請求項20】

前記装置は、前記処理されたマルチチャンネル信号の第1のチャンネルのレベルと前記処理されたマルチチャンネル信号の第2のチャンネルのレベルとの間の関係に基づいてボイス活動の存在を示すように構成されたボイス活動検出器を含む、請求項12乃至19のいずれか一項に記載の装置。

30

【請求項21】

前記方法は、前記処理されたマルチチャンネル信号の第1のチャンネルのレベルと前記処理されたマルチチャンネル信号の第2のチャンネルのレベルとの間の関係に基づいて、及び前記コヒーレンシー測定の前記値をある閾値と比較した結果に応じて、前記マルチチャンネル信号の前記第1および第2のチャンネルの少なくとも1つからの音響情報に従って雑音推定値を更新することを備える、請求項14および15のいずれか一項に記載の装置。

【請求項22】

40

マルチチャンネル信号を処理するための装置であって、

複数の計算された位相差を取得するために、前記マルチチャンネル信号の複数の異なる周波数成分の各々に関して、前記マルチチャンネル信号の第1のチャンネルにおける前記周波数成分の位相と前記マルチチャンネル信号の第2のチャンネルにおける前記周波数成分の位相との間の差を計算するための手段と；

前記第1のチャンネルのレベルと前記第2のチャンネルの対応するレベルを計算するための手段と；

前記第1のチャンネルの前記計算されたレベルと、前記第2のチャンネルの前記計算されたレベルと、前記複数の計算された位相差の少なくとも1つとに基づいて、利得係数の更新された値を計算するための手段と；及び

50

前記更新された値に従って、前記第 1 のチャンネルの対応する振幅に関して前記第 2 のチャンネルの振幅を修正することによって、処理されたマルチチャンネル信号を生成するための手段と；

を備える、マルチチャンネル信号を処理するための装置。

【請求項 2 3】

前記第 1 のチャンネルの前記計算されたレベルは第 1 の周波数サブバンドにおける前記第 1 のチャンネルの計算されたエネルギーであり、前記第 2 のチャンネルの前記計算されたレベルは前記第 1 の周波数サブバンドにおける前記第 2 のチャンネルの計算されたエネルギーであり、且つ

前記第 1 のチャンネルの前記振幅は前記第 1 の周波数サブバンドにおける前記第 1 のチャンネルの振幅であり、前記第 2 のチャンネルの前記対応する振幅は前記第 1 の周波数サブバンドにおける前記第 2 のチャンネルの振幅であり、且つ

前記装置は：

前記第 1 の周波数サブバンドとは異なる第 2 の周波数サブバンドにおける前記第 1 のチャンネルのエネルギーを計算するための手段と；

前記第 2 の周波数サブバンドにおける前記第 2 のチャンネルのエネルギーを計算するための手段と；及び

前記第 2 の周波数サブバンドにおける前記第 1 のチャンネルの前記計算されたエネルギーと、前記第 2 の周波数サブバンドにおける前記第 2 のチャンネルの前記計算されたエネルギーと、前記複数の計算された位相差の少なくとも 1 つとに基づいて、第 2 の利得係数の更新された値を計算するための手段と；

を備えており、

処理されたマルチチャンネル信号を生成するための前記手段は、前記第 2 の利得係数の前記更新された値に従って、前記第 2 の周波数サブバンドにおける前記第 1 のチャンネルの振幅に関して前記第 2 の周波数サブバンドにおける前記第 2 のチャンネルの振幅を修正することによって、前記処理されたマルチチャンネル信号を生成するための手段を含む、請求項 2 2 に記載の装置。

【請求項 2 4】

前記装置は、前記複数の計算された位相差からの情報に基づいて、少なくとも前記複数の異なる周波数成分の到着方向間のコヒーレンスの度合いを示すコヒーレンシー測定の値を計算するための手段を備えており、且つ

利得係数の更新された値を計算するための前記手段は、前記コヒーレンシー測定の前記計算された値に基づいて前記利得係数の前記更新された値を計算するように構成されている、請求項 2 2 および 2 3 のいずれか一項に記載の装置。

【請求項 2 5】

前記第 2 のチャンネルの対応する振幅に関して前記第 1 のチャンネルの振幅を修正するための前記手段は、前記コヒーレンシー測定の前記値をある閾値と比較するための前記手段の出力に応じてこのような修正を実行するように構成されている、請求項 2 4 に記載の装置。

【請求項 2 6】

前記装置は、前記マルチチャンネル信号の推定されたピッチ周波数に基づいて前記複数の異なる周波数成分を選択するための手段を含む、請求項 2 2 乃至 2 5 のいずれか一項に記載の装置。

【請求項 2 7】

利得係数の前記更新された値は、前記第 1 のチャンネルの前記計算されたレベルと前記第 2 のチャンネルの前記計算されたレベルとの比に基づく、請求項 2 2 乃至 2 6 のいずれか一項に記載の装置。

【請求項 2 8】

前記更新された値に従って、前記第 1 のチャンネルの対応する振幅に関して前記第 2 のチャンネルの振幅を修正することによって、処理されたマルチチャンネル信号を生成するための

10

20

30

40

50

前記手段は、前記第1および第2のチャンネルの前記計算されたレベル間の不均衡を低減するように構成されている、請求項22乃至27のいずれか一項に記載の装置。

【請求項29】

処理されたマルチチャンネル信号を生成するための前記手段は、前記更新された値に従って、前記マルチチャンネル信号の複数の連続するセグメントの各々において前記第1のチャンネルの対応する振幅に関して前記第2のチャンネルの振幅を修正するための手段を含む、請求項22乃至28のいずれか一項に記載の装置。

【請求項30】

前記装置は、前記処理されたマルチチャンネル信号の第1のチャンネルのレベルと前記処理されたマルチチャンネル信号の第2のチャンネルのレベルとの間の関係に基づいて、ボイス活動の存在を示すための手段を備える、請求項22乃至29のいずれか一項に記載の装置。

10

【請求項31】

前記装置は、前記処理されたマルチチャンネル信号の第1のチャンネルのレベルと前記処理されたマルチチャンネル信号の第2のチャンネルのレベルとの間の関係に基づいて、及び前記コヒーレンシー測定の前記値をある閾値と比較した結果に応じて、前記マルチチャンネル信号の前記第1および第2のチャンネルの少なくとも1つからの音響情報に従って雑音推定値を更新するための手段を備える、請求項24および25のいずれか一項に記載の装置。

【発明の詳細な説明】

【技術分野】

【0001】

20

[米国特許法第119条下での優先権の主張]

本特許出願は、2009年6月9日出願され、また本出願の譲受人に譲渡された「Systems, methods, apparatus, and computer-readable media for coherence detection」（コヒーレンス検出のためのシステム、方法、装置、およびコンピュータ可読媒体）と題する米国仮特許出願第61/185,518号に対する優先権を主張する。本特許出願はまた、2009年9月8日出願され、また本出願の譲受人に譲渡された「Systems, methods, apparatus, and computer-readable media for coherence detection」（コヒーレンス検出のためのシステム、方法、装置、およびコンピュータ可読媒体）と題する、米国仮特許出願第61/240,318号に対する優先権を主張する。

【0002】

30

本特許出願はまた、2009年7月20日出願され、また本出願の譲受人に譲渡された「Systems, methods, apparatus, and computer-readable media for phase-based processing of multichannel signal」（マルチチャンネル信号の位相ベースの処理のためのシステム、方法、装置、およびコンピュータ可読媒体）と題する米国仮特許出願第61/227,037号、代理人明細書（attorney docket）第091561P1号に対する優先権を主張する。本特許出願はまた、2009年9月8日出願され、また本出願の譲受人に譲渡された「Systems, methods, apparatus, and computer-readable media for phase-based processing of multichannel signal」（マルチチャンネル信号の位相ベースの処理のためのシステム、方法、装置、およびコンピュータ可読媒体）と題する米国仮特許出願第61/240,320号に対する優先権を主張する。

40

【0003】

[技術分野]

本開示は、信号処理に関する。

【背景技術】

【0004】

以前は静かなオフィスまたは家庭環境において実行されていた多くの活動は、今日では自動車、街頭またはカフェのような音響的に変わりやすい状況で実行されている。例えば、ある人はボイス通信チャンネルを使用して他の人と通信することを望むことができる。このチャンネルは、例えば移動無線送受話器またはヘッドホン、ウォークトーカー、送受兼用無線機、カーキット（car-kit）、または他の通信デバイスによって用意され得る。そ

50

の結果、かなりの量のボイス通信は、典型的には人々が集まりやすい場所で出会う種類の雑音内容を有する、ユーザが他の人々に取り囲まれる環境で移動デバイス（例えば、スマートフォン、送受話器および／またはヘッドホン）を使用して行われている。このような雑音は、電話会話の遠端にいるユーザの気を紛らわしたり、ユーザを悩ましたりする傾向がある。更に、多くの標準的な自動化されたビジネス取引（例えば、勘定収支または株価チェック）はボイス認識ベースのデータ問合せを使用し、またこれらのシステムの精度は妨害雑音によって大きく妨げられ得る。

【0005】

雑音の多い環境で通信が行われる用途に関しては、背景雑音から所望のスピーチ信号を分離することが望ましいことがあり得る。雑音は所望の信号に干渉する、またはそうでなければ所望の信号を劣化させるすべての信号の組合せと定義され得る。背景雑音は、所望の信号から、および／または他の信号のいずれかから生成される反響および残響ばかりでなく他の人々の背景会話といった音響環境内で生成される非常に多くの雑音信号を含み得る。所望のスピーチ信号が背景雑音から分離されない場合には、スピーチ信号を信頼度高く効率的に使用することは困難である可能性がある。1つの特定の例では、スピーチ信号は雑音の多い環境で生成され、この環境雑音からスピーチ信号を分離するためにスピーチ処理方法が使用される。

【0006】

移動環境で遭遇する雑音は、競合する話者、ミュージック、片言、街頭雑音、および／または空港雑音といった種々の異なる成分を含み得る。このような雑音のシグネチャー（signature）は典型的には非定常的であってユーザ自身の周波数シグネチャーに近いので、雑音は従来型の単一マイクロホンまたは固定式ビーム形成タイプの方法を使用してモデル化するのが困難であり得る。単一マイクロホン雑音低減技法は、典型的には最適な性能を達成するためにかなりのパラメータ調整（parameter tuning）を必要とする。例えば、適当な雑音基準はこのような場合には直接利用可能でない可能性があり、間接的に雑音基準を導き出すことを必要とする可能性がある。従って、雑音の多い環境におけるボイス通信のために移動デバイスの使用をサポートするために、多数のマイクロホンに基づく進歩した信号処理が望ましいことができる。

【発明の概要】

【0007】

全体的構成によるマルチチャネル信号を処理する方法は、マルチチャネル信号の複数の異なる周波数成分の各々に関して、複数の計算された位相差を取得するために、マルチチャネル信号の第1のチャネルの周波数成分の位相とマルチチャネル信号の第2のチャネルの周波数成分の位相との間の差を計算することを含む。この方法は第1のチャネルのレベルと第2のチャネルの対応するレベルとを計算することを含む。この方法は、第1のチャネルの計算されたレベルと第2のチャネルの計算されたレベルと複数の計算された位相差の少なくとも1つとに基づいて利得係数の更新された値を計算することと、更新された値に従って第1のチャネルの対応する振幅に関して第2のチャネルの振幅を修正することによって、処理されたマルチチャネル信号を生成することと、を含む。これらの活動の各々を実行するための手段を含む装置も本明細書で開示される。このような方法を実行するための機械実行可能命令を記憶する具体的特徴機能を有するコンピュータ可読媒体も本明細書で開示される。

【0008】

全体的構成によるマルチチャネル信号を処理するための装置は、マルチチャネル信号の複数の異なる周波数成分の各々に関して、マルチチャネル信号の第1のチャネルの周波数成分の位相とマルチチャネル信号の第2のチャネルの周波数成分の位相との間の差を計算することによって複数の計算された位相差を取得するように構成された第1の計算器を含む。この装置は、第1のチャネルのレベルと第2のチャネルの対応するレベルとを計算するように構成された第2の計算器と、第1のチャネルの計算されたレベルと第2のチャネルの計算されたレベルと複数の計算された位相差の少なくとも1つとに基づいて利得係

10

20

30

40

50

の更新された値を計算するように構成された第3の計算器と、を含む。この装置は、この更新された値に従って第1のチャンネルの対応する振幅に関して第2のチャンネルの振幅を修正することによって、処理されたマルチチャンネル信号を生成するように構成された利得制御要素を含む。

【図面の簡単な説明】

【0009】

【図1】図1は、使用中のヘッドホンD100の側面図を示す。

【図2】図2は、ユーザの耳に装着されたヘッドホンD100の上面図を示す。

【図3A】図3Aは、使用中の送受話器D300の側面図を示す。

【図3B】図3Bは、マイクロホンアレイに関する幅広面領域およびエンドファイア (endfire) 領域の例を示す。 10

【図4A】図4Aは、全体的構成によるマルチチャンネル信号を処理する方法M100のための流れ図を示す。

【図4B】図4Bは、タスクT100の実現形態T102の流れ図を示す。

【図4C】図4Cは、タスクT110の実現形態T112の流れ図を示す。

【図5A】図5Aは、タスクT300の実現形態T302の流れ図を示す。

【図5B】図5Bは、タスクT300の代替実現形態T304の流れ図を示す。

【図5C】図5Cは、方法M100の実現形態M200の流れ図を示す。

【図6A】図6Aは、到着方向を推定するためのアプローチを示す幾何学的近似の一例を示す。 20

【図6B】図6Bは、第2象限値および第3象限値に関して図6Aの近似を使用することの一例を示す。

【図7】図7は、球形波面を想定するモデルの一例を示す。

【図8A】図8Aは、パスバンドとストップバンドとの間の比較的急激な遷移を有するマスキング関数の一例を示す。

【図8B】図8Bは、マスキング関数のための線形ロールオフの一例を示す。

【図8C】図8Cは、マスキング関数のための非線形ロールオフの一例を示す。

【図9A】図9Aは、異なるパラメータ値のための非線形関数の例を示す。

【図9B】図9Bは、異なるパラメータ値のための非線形関数の例を示す。

【図9C】図9Cは、異なるパラメータ値のための非線形関数の例を示す。 30

【図10】図10は、マスキング関数の指向性パターンの前方ローブおよび後方ローブを示す。

【図11A】図11Aは、方法M100の実現形態M110の流れ図を示す。

【図11B】図11Bは、タスクT360の実現形態T362の流れ図を示す。

【図11C】図11Cは、タスクT360の実現形態T364の流れ図を示す。

【図12A】図12Aは、方法M100の実現形態M120の流れ図を示す。

【図12B】図12Bは、方法M100の実現形態M130の流れ図を示す。

【図13A】図13Aは、方法M100の実現形態M140の流れ図を示す。

【図13B】図13Bは、方法M100の実現形態M150の流れ図を示す。

【図14A】図14Aは、3つの異なる閾値に対応する近接検出領域の境界の一例を示す。 40

【図14B】図14Bは、スピーカーカバレッジのコーンを取得するための近接バブルとある範囲の許容された方向との交差の一例を示す。

【図15】図15は、図14Bに示された音源選択領域境界の上面図を示す。

【図16】図16は、図14Bに示された音源選択領域境界の側面図を示す。

【図17A】図17Aは、方法M100の実現形態M160の流れ図を示す。

【図17B】図17Bは、方法M100の実現形態M170の流れ図を示す。

【図18】図18は、方法M170の実現形態M180の流れ図を示す。

【図19A】図19Aは、全体的構成による方法M300の流れ図を示す。

【図19B】図19Bは、方法M300の実現形態M310の流れ図を示す。 50

【図 2 0 A】図 2 0 A は、方法 M 3 1 0 の実現形態 M 3 2 0 の流れ図を示す。
 【図 2 0 B】図 2 0 B は、全体的構成による装置 G 1 0 0 のブロック図を示す。
 【図 2 1 A】図 2 1 A は、全体的構成による装置 A 1 0 0 のブロック図を示す。
 【図 2 1 B】図 2 1 B は、装置 A 1 1 0 のブロック図を示す。
 【図 2 2】図 2 2 は、装置 A 1 2 0 のブロック図を示す。
 【図 2 3 A】図 2 3 A は、アレイ R 1 0 0 の実現形態 R 2 0 0 のブロック図を示す。
 【図 2 3 B】図 2 3 B は、アレイ R 2 0 0 の実現形態 R 2 1 0 のブロック図を示す。
 【図 2 4 A】図 2 4 A は、全体的構成によるデバイス D 1 0 のブロック図を示す。
 【図 2 4 B】図 2 4 B は、デバイス D 1 0 の実現形態 D 2 0 のブロック図を示す。
 【図 2 5 A】図 2 5 A は、マルチマイクロホン無線ヘッドホン D 1 0 0 の種々の図を示す

10

。【図 2 5 B】図 2 5 B は、マルチマイクロホン無線ヘッドホン D 1 0 0 の種々の図を示す

。【図 2 5 C】図 2 5 C は、マルチマイクロホン無線ヘッドホン D 1 0 0 の種々の図を示す

。【図 2 5 D】図 2 5 D は、マルチマイクロホン無線ヘッドホン D 1 0 0 の種々の図を示す

。【図 2 6 A】図 2 6 A は、マルチマイクロホン無線ヘッドホン D 2 0 0 の種々の図を示す

。【図 2 6 B】図 2 6 B は、マルチマイクロホン無線ヘッドホン D 2 0 0 の種々の図を示す

20

。【図 2 6 C】図 2 6 C は、マルチマイクロホン無線ヘッドホン D 2 0 0 の種々の図を示す

。【図 2 6 D】図 2 6 D は、マルチマイクロホン無線ヘッドホン D 2 0 0 の種々の図を示す

。【図 2 7 A】図 2 7 A は、マルチマイクロホン通信送受話器 D 3 0 0 の（中心軸に沿った）断面図を示す。

【図 2 7 B】図 2 7 B は、デバイス D 3 0 0 の実現形態 D 3 1 0 の断面図を示す。

【図 2 8 A】図 2 8 A は、マルチマイクロホン・メディア・プレーヤ D 4 0 0 の図を示す

30

。【図 2 8 B】図 2 8 B は、マルチマイクロホン・メディア・プレーヤ D 4 1 0 の図を示す

。【図 2 8 C】図 2 8 C は、マルチマイクロホン・メディア・プレーヤ D 4 2 0 の図を示す

。【図 2 9】図 2 9 は、マルチマイクロホン・ハンズフリー・カーキット D 5 0 0 の図を示す。

【図 3 0】図 3 0 は、デバイス D 1 0 のマルチマイクロホン携帯型オーディオセンシング実現形態 D 6 0 0 の図を示す。

【発明を実施するための形態】

【0 0 1 0】

40

現実世界は、残響という結果を招く多数のサウンドにしばしば侵入する（transgress）シングルポイント雑音源を含む多数の雑音源で溢れている。背景音響雑音は、所望のサウンド信号から、および／または他の信号のいずれかから生成される反響および残響ばかりでなく、一般環境によって生成される多数の雑音信号および他の人々の背景会話によって生成される干渉信号を含み得る。

【0 0 1 1】

環境雑音は、近端スピーチ信号といった感知されたオーディオ信号の理解度に影響を及ぼすことができる。背景雑音から所望のオーディオ信号を区別するために信号処理を使用することは望ましいことであり得る。通信が雑音の多い環境で行われることができる用途に関しては、例えば、背景雑音からスピーチ信号を区別してスピーチ信号の理解度を高め

50

るためのスピーチ処理方法を使用することが望ましいことができる。現実世界の状況では雑音はほとんど常に存在するので、このような処理は毎日の通信の多くの領域において重要であり得る。

【0012】

音響信号を受信するように構成された2つ以上のマイクロホンのアレイR100を有する携帯型オーディオセンシング・デバイスを製造することが望ましいことができる。このようなアレイを含むように実現され得る、そしてオーディオレコーディングおよび／またはボイス通信用途のために使用され得る携帯型オーディオセンシング・デバイスの例は、電話送受話器（例えば、携帯電話送受話器またはスマートフォン）；有線または無線ヘッドホン（例えば、Bluetooth（登録商標）ヘッドホン）、ハンドヘルド・オーディオおよび／またはビデオレコーダー；オーディオおよび／またはビデオコンテンツを記録するように構成されたパーソナル・メディア・プレーヤー；パーソナル・デジタル・アシスタント（PDA）または他のハンドヘルド・コンピューティング・デバイス；およびノートブックコンピュータ、ラップトップコンピュータ、ネットブックコンピュータ、または他のポータブル・コンピューティング・デバイスを含む。

【0013】

通常の使用時に、携帯型オーディオセンシング・デバイスは所望の音源に関してある範囲の標準的な方位の間でも動作し得る。例えば、異なるユーザが異なる仕方でデバイスを装着または保持することがあり、また同じユーザが同じ使用期間内（例えば、1回の電話通話中）であっても異なる時に異なる仕方でデバイスを装着または保持することがあり得る。図1は、ユーザの口に関してデバイスのある範囲の標準的な方位にある2つの例を含む使用中のヘッドホンD100の側面図を示す。ヘッドホンD100は、デバイスの典型的な使用時に、より直接的にユーザのボイスを受け入れるように位置付けられた1次マイクロホンMC10と、デバイスの典型的な使用時に、より直接的でなくユーザのボイスを受け入れるように位置付けられた2次マイクロホンMC20を含むアレイR100の一事例を有する。図2は、ユーザの口に関して標準的な方位にあるユーザの耳に装着されたヘッドホンD100の上面図を示す。図3Aは、ユーザの口に関してデバイスのある範囲の標準的な方位にある2つの例を含む使用中の送受話器D300の側面図を示す。

【0014】

文脈によって明確に限定されていない場合には、用語「信号」は本明細書では、ワイヤ、バスまたは他の伝送媒体について表されるようなメモリ位置（または、1セットのメモリ位置）の状態を含むこの用語の通常の意味のいずれかを指すために使用される。文脈によって明確に限定されていない場合には、用語「生成すること（generating）」は本明細書では、コンピューティング（computing）またはそうでなければ作り出すこと（producing）といったこの用語の通常の意味のいずれかを指すために使用される。文脈によって明確に限定されていない場合には、用語「計算すること（calculating）」は本明細書では、コンピューティング、評価すること（evaluating）、平滑化すること（smoothing）および／または複数の値から選択することといったこの用語の通常の意味のいずれかを指すために使用される。文脈によって明確に限定されていない場合には、用語「取得すること（obtaining）」は本明細書では、計算すること、導き出すこと、受信すること（例えば、外部デバイスから）、および／または（例えば、1アレイの記憶要素から）検索することといったこの用語の通常の意味のいずれかを指すために使用される。文脈によって明確に限定されていない場合には、用語「選択すること（selecting）」は本明細書では、識別すること、指示すること、適用すること、および／または2つ以上のセットの内の少なくとも1つ、およびすべてより少なく、を使用すること、といったこの用語の通常の意味のいずれかを指すために使用される。用語「備えること（comprising）」が本説明および請求項において使用される場合には、この用語は他の要素または動作を除外しない。用語「に基づく（based on）」（「AはBに基づく」におけるような）は、ケース（i）「から導き出される」（例えば、「BはAの先行体である」、（ii）「少なくとも・・・に基づく」（例えば、「Aは少なくともBに基づく」）、および特定の文脈において適当で

10

20

30

40

50

あれば、(i i i) 「・・・に等しい」(例えば、「AはBに等しい」)を含むこの用語の通常の意味のいずれかを指すために使用される。同様に用語「・・・に応じて(in response to)」は「少なくとも・・・に応じて」を含むこの用語の通常の意味のいずれかを指すために使用される。

【0015】

マルチマイクロホン・オーディオセンシング・デバイスのマイクロホンの「位置(location)」への言及は、文脈によって別に指示されていない場合には、マイクロホンの音響的に敏感な面の中心の位置を指す。用語「チャンネル(channel)」は、特定の文脈に従って、時には信号経路を指すために、また他の時にはこのような経路によって伝達される信号を指すために使用される。別に指示されていない場合には、用語「シリーズ(series)」は2つ以上の項目の一続きを指すために使用される。用語「対数(logarithm)」は10を底とする対数を指すために使用されるが、他の底へのこのような演算の拡張も本開示の範囲内にある。用語「周波数成分(frequency component)」は、信号の周波数ドメイン表現(例えば、高速フーリエ変換によって作り出されるような)のサンプル(または、「ビン(bin)」)、または信号のサブバンド(例えば、バークスケール(Bark scale)サブバンド)といった信号の1セットの周波数または周波数帯域の間の1つを指すために使用される。

【0016】

別に指示されていない場合には、特定の特徴機能を有する装置の動作のいかなる開示もアナログ的特徴機能を有する方法を開示するように明確に意図されており(逆もまた同様である)、また特定の構成による装置の動作のいかなる開示もアナログ的構成による方法を開示するように明確に意図されている(逆もまた同様である)。用語「構成(configuration)」はこの用語の特定の文脈によって示されるような方法、装置および/またはシステムへの参照時に使用され得る。用語「方法」、「プロセス」、「手順」、および「技法」は、特定の文脈によって別に指示されていない場合には、一般的にまた相互交換可能に使用される。用語「装置(apparatus)」および「デバイス(device)」もまた、特定の文脈によって別に指示されていない場合には、一般的にまた相互交換可能に使用される。用語「要素」および「モジュール」は典型的には、より大きな構成の一部分を指すために使用される。用語「システム」は、この用語の文脈によって明確に限定されていない場合には、本明細書では「共通目的に役立つために相互作用する要素の1グループ」を含むこの用語の通常の意味のいずれかを指すために使用される。文書の一部の参照によるいかなる抱合も、抱合された部分において参照される何らかの数字(figure)と同様に用語または変数の定義がこの文書のほかの場所に現れる部分内で参照される用語または変数の定義を抱合すると理解されるべきである。

【0017】

近距離音場(near-field)は、サウンド受信機(例えば、マイクロホンアレイ)から1波長未満離れた空間の領域として定義され得る。この定義の下で、この領域の境界までの距離は周波数とは逆に変化する。例えば、200Hz、700Hzおよび2000Hzの周波数において1波長境界までの距離は、それぞれ約170、49および17センチメートルである。その代わりに近距離音場/遠距離音場境界がマイクロホンアレイから特定の距離(例えば、このアレイのマイクロホンから、またはこのアレイの重心から50センチメートル、またはこのアレイのマイクロホンから、またはこのアレイの重心から1メートルまたは1.5メートル)にあると考えることは有用であり得る。

【0018】

マイクロホンアレイは、各チャンネルが音響環境に対するマイクロホンの対応する1つの応答に基づいているマルチチャンネル信号を作り出す。異なる音源から受信された信号の成分間を弁別するためにマルチチャンネル信号に対して空間選択的処理(SSP)動作を実行することが望ましいことができる。例えば、指向性サウンドの所望の音源(例えば、ユーザの口)からのサウンド成分と、拡散背景雑音および/または指向性干渉雑音の1つ以上の音源(例えば、競合スピーカー)からのサウンド成分との間を弁別することは望ましい

可能性がある。SSP動作の例は、ビーム形成アプローチ（例えば、汎用サイドローブ・キャンセレーション（GSC）、最小分散無歪み応答（MVDR）、および／または線形拘束最小分散（LCMV）ビームフォーマー）、ブラインドソース分離（BSS）および他の適応型学習アプローチ、および利得ベース近接検出を含む。SSP動作の典型的な用途は、携帯型オーディオセンシング・デバイスのためのマルチマイクロホン雑音低減方式を含む。

【0019】

SSP動作といったアレイR100によって作り出されたマルチチャンネル信号に対する動作の成果は、アレイチャンネルの応答特性が互いにどれほどよく整合しているかに依存することができる。例えば、チャンネルのレベルはそれぞれのマイクロホンの応答特性の差、それぞれの前処理ステージの利得レベルの差、および／またはチャンネルの回路雑音レベルの差に起因して異なる可能性がある。このような場合、結果として得られたマルチチャンネル信号は、チャンネル応答特性間の不整合（「チャンネル応答不均衡」とも呼ばれる）が補正され得ない場合には、音響環境の正確な表現を与えないことができる。

【0020】

このような補正なしでは、このような信号に基づくSSP動作は誤った結果を与えることができる。チャンネル間の利得差が指向性音源の相対的近接度を示すために使用される動作に関しては、チャンネルの応答間の不均衡は近接度指示の精度を下げる傾向があるであろう。もう1つの例では、低周波数（すなわち、約100Hz～1kHz）における1または2デシベルほどの小さいチャンネル間の振幅応答偏差は低周波数指向性を著しく低減することができる。アレイR100のチャンネルの応答間の不均衡の影響は、2個より多いマイクロホンを有するアレイR100の実現形態からのマルチチャンネル信号を処理する用途のために特に有害であり得る。

【0021】

正確なチャンネル較正は、ヘッドホン用途のために特に重要であり得る。例えば、近距離音場音源から到着するサウンド成分と遠距離音場音源から到着するサウンド成分との間を弁別するように携帯型オーディオセンシング・デバイスを構成することが望ましいことができる。このような弁別は、マルチチャンネル信号の2つのチャンネルの利得レベル間の差（すなわち、「チャンネル間利得レベル差」）がアレイのエンドファイア方向（すなわち、対応するマイクロホンの中心を通る直線の近く）に配置された近距離音場音源からのサウンド成分のために、より高いことが予期され得るので、この差に基づいて実行され得る。

【0022】

マイクロホン間の距離が減少すると、近距離音場信号に関するチャンネル間利得レベル差も減少する。ハンドヘルド用途のために、近距離音場信号に関するチャンネル間利得レベル差は典型的には、遠距離音場信号に関するチャンネル間利得レベル差から約6デシベルである。しかしながらヘッドホン用途のためには、典型的な近距離音場サウンド成分に関するチャンネル間利得レベル差は、典型的な遠距離音場音響成分に関するチャンネル間利得レベル差の3デシベル（または、これより小さい）以内にあり得る。このような場合には、ほんの数デシベルのチャンネル応答不均衡がこのような成分間を弁別する能力を厳しく妨げることができるが、3デシベル以上の不均衡はこの能力を台無しにすることができる。

【0023】

アレイチャンネルの応答間の不均衡は、マイクロホン自体の応答間の差から発生し得る。アレイR100の製造時にばらつきが発生し得るので、1バッチの量産された一見同じマイクロホンの間でも感度はマイクロホンによってかなり変わることができる。携帯型の大量市場向けオーディオセンシング・デバイスでの使用のためのマイクロホンは、例えばプラスマイナス3デシベルの感度公差で製造され得るので、アレイR100の実現形態における2つのこのようなマイクロホンの感度は6デシベルほどの大きさだけ異なることができる。

【0024】

チャンネル応答不均衡の問題は、応答が既に整合させられている（例えば、選別または廃

10

20

30

40

50

棄プロセスを介して) マイクロホンを使用することによって携帯型オーディオセンシング・デバイスの製造時に取られ得る。代替としてまたは更に、チャンネル較正手順は、研究所においておよび／または工場といった製造施設においてアレイ R 1 0 0 のマイクロホン（または、アレイを含むデバイス上の）に実行され得る。このような手順は、均衡したマルチチャンネル信号を作り出すために 1 つ以上の利得係数を計算してこのような係数を対応するチャンネルに適用することによって不均衡を補正し得る。サービスの前に実行され得る較正手順の例は、2 0 0 9 年 5 月 2 8 日に出願された、「SYSTEMS, METHODS, AND APPARATUS FOR MULTICHANNEL SIGNAL BALANCING」（マルチチャンネル信号均衡化のためのシステム、方法、および装置）と題する米国特許出願第 1 2 / 4 7 3, 9 3 0 号と、2 0 0 8 年 1 2 月 1 2 日に出願された、「SYSTEMS, METHODS, AND APPARATUS FOR MULTI-MICROPHONE BASED SPEECH ENHANCEMENT」（マルチマイクロホンベースの音声改善のためのシステム、方法、および装置）と題する米国特許出願第 1 2 / 3 3 4, 2 4 6 号とに説明されている。このような整合または較正動作はデバイスを製造するコストを増加させることができるが、デバイスのサービス寿命（例えば、老化による）の期間中に発生するチャンネル応答不均衡に対しては無効でもあり得る。

【0 0 2 5】

代替としてまたは更に、チャンネル較正は稼働中に（例えば、米国特許出願第 1 2 / 4 7 3, 9 3 0 号に説明されているように）実行され得る。このような手順は、時間の経過と共に発生する応答不均衡を修正するために、および／または初期応答不均衡を修正するために使用され得る。初期応答不均衡は、例えばマイクロホン不整合および／または誤った較正手順（例えば、この手順中にマイクロホンが触られるか、またはカバーされる）に起因することができる。変動するチャンネルレベルによってユーザの気を紛らわすことを防止するために、このような手順が時間の経過と共に徐々に変化する補正を適用することが望ましいことができる。しかしながら初期応答不均衡が大きい場合にはこのような徐々の補正は、マルチチャンネル信号に対する S S P 動作がうまく実行できず、不満足なユーザ体験につながる恐れがある、長い収束期間（例えば、1 ～ 1 0 分以上）を引き起こすことができる。

【0 0 2 6】

マルチチャンネル信号の時間・周波数ポイントを分類するために位相分析が使用され得る。例えば、信号のチャンネルの推定された位相間の複数の異なる周波数の各々における差に基づいてマルチチャンネル信号の時間・周波数ポイントを分類するようにシステム、方法または装置を構成することが望ましいことができる。このような構成は本明細書では「位相ベース」と呼ばれる。

【0 0 2 7】

特定の位相差特性を示す時間・周波数ポイントを識別するために位相ベース方式を使用することは望ましいことができる。例えば、位相ベース方式は、感知されたマルチチャンネル信号の特定の周波数成分がアレイ軸に関する可能な角度の範囲内から、またはこの範囲外から発生したかどうかを決定するために、マイクロホン間距離とチャンネル間位相差とに関する情報を適用するように構成され得る。このような決定は、異なる方向から到着する音響成分間を弁別するために（例えば、可能な範囲内から発生するサウンドが選択されてこの範囲外から発生するサウンドが拒絶されるように）、および／または近距離音場音源および遠距離音場音源から到着するサウンド成分間を弁別するために使用され得る。

【0 0 2 8】

典型的な用途ではこのようなシステム、方法、または装置は、マルチチャンネル信号の少なくとも一部分に亘る（例えば、特定の範囲の周波数に亘る、および／または特定の時間間隔に亘る）各時間・周波数ポイントに関するマイクロホンペアに関して到着方向を計算するために使用される。指向性マスキング関数は、所望の範囲内の到着方向を有するポイントを他の到着方向を有するポイントから区別するためにこれらの結果に適用され得る。指向性マスキング演算からの結果は、このマスクの外側の到着方向を有する時間・周波数ポイントを廃棄する、または減衰させることによって望ましくない方向からのサウンド成

分を減衰させるために使用され得る。

【0029】

上記のように、多くのマルチマイクロホン空間処理動作は本来的にマイクロホンチャンネルの相対的利得応答に依存しているので、チャンネル利得応答の較正はこのような空間処理動作を可能にするために必要であり得る。製造時にこのような較正を実行することは典型的には多大の時間を必要とする、および／またはそうでなくとも高価である。しかしながら位相ベースの方式は、入力チャンネル間の利得不均衡によって比較的影響されないように実現され得るので、対応するチャンネルの利得応答が互いに整合させられる度合いは、計算された位相差の精度およびこれらに基づく引き続きの動作（例えば、指向性マスキング）に対する限定要因ではない。

10

【0030】

本明細書に説明されているようなチャンネル較正動作（「チャンネル均衡化」動作とも呼ばれる）をサポートするために、このような方式の分類結果を使用することによって位相ベース方式のチャンネル不均衡に対する強固さを活用することが望ましいことができる。例えば、チャンネル均衡化のために有用であり得る記録されたマルチチャンネル信号の周波数成分および／または時間間隔を識別するために位相ベース方式を使用することが望ましいことができる。このような方式は、到着方向が各チャンネルにおいて比較的等しい応答を作り出すように予期されることを示す時間・周波数ポイントを選択するように構成され得る。

【0031】

図3Bに示されているような2マイクロホンアレイに関するある範囲の音源方向に関して、チャンネル較正のための幅広面方向（すなわち、アレイ軸に直交する方向）から到着するサウンド成分だけを使用することが望ましいことができる。このような状況は、例えば近距離音場音源が動作しておらず音源が分散している（例えば、背景雑音）ときに見出され得る。較正のための遠距離音場エンドファイア音源から到着するサウンド成分は無視可能なチャンネル間利得レベル差（例えば、分散に起因する）を引き起こすと予期され得るので、このようなサウンド成分を使用することも受入れ可能であり得る。しかしながら、アレイのエンドファイア方向（すなわち、アレイ軸に近い方向）から到着する近距離音場サウンド成分は、チャンネル不均衡よりむしろ音源位置情報を表すチャンネル間利得差を有すると予期されるであろう。その結果、較正のためにこのような成分を使用することは誤った結果を作り出す可能性があり、幅広面方向から到着するサウンド成分からこのような成分を区別するために指向性マスキング演算を使用することが望ましいことができる。

20

30

【0032】

このような位相ベースの分類方式は、実行時間における（例えば、デバイスの使用時に連続的または間欠的いずれにおいても）較正動作をサポートするために使用され得る。このような仕方では、チャンネル利得応答不均衡にそれ自身は影響されない敏速で正確なチャンネル較正動作が達成され得る。代替として、選択された時間・周波数ポイントからの情報は、後にチャンネル較正動作をサポートするために、ある時間に亘って蓄積され得る。

【0033】

図4Aは、タスクT100、T200、T300、およびT400を含む全体的構成によるマルチチャンネル信号を処理する方法M100のための流れ図を示す。タスクT100は、この信号の複数の異なる周波数成分の各々に関してマルチチャンネル信号のチャンネル（例えば、マイクロホンチャンネル）間の位相差を計算する。タスクT200は、マルチチャンネル信号の第1のチャンネルのレベルとマルチチャンネル信号の第2のチャンネルの対応するレベルとを計算する。計算されたレベルと計算された位相差の少なくとも1つとに基づいて、タスクT300は利得係数値を更新する。更新された利得係数値に基づいて、タスクT400は処理された（例えば、均衡化された）マルチチャンネル信号を作り出すために、第1のチャンネルの対応する振幅に関して第2のチャンネルの振幅を修正する。方法M100はまた、SSP動作といった（例えば、本明細書で更に詳細に説明されるような）マルチチャンネル信号に対する更なる動作をサポートするためにも使用され得る。

40

【0034】

50

方法M100は、マルチチャネル信号を一連のセグメントとして処理するように構成され得る。典型的なセグメント長は、約5または10ミリセカンドから約40または50ミリセカンドの範囲にあり、これらのセグメントはオーバーラップしていることも（例えば、隣接セグメントと25%または50%だけオーバーラップしている）、オーバーラップしていないこともあり得る。1つの特定の例ではマルチチャネル信号は、各々が10ミリセカンドの長さを有する一連のオーバーラップしていないセグメントまたは「フレーム」に分割される。タスクT100は、セグメントの各々に関する1セットの（例えば、1ベクトルの）位相差を計算するように構成され得る。方法M100のある幾つかの実現形態では、タスクT200は各チャネルのセグメントの各々に関するレベルを計算するように構成され、タスクT300はセグメントの少なくとも一部に関する利得係数値を更新するように構成されている。方法M100の他の実現形態では、タスクT200は各チャネルのセグメントの各々に関する1セットのサブバンドレベルを計算するように構成され、タスクT300は1セットのサブバンド利得係数値の1つ以上を更新するように構成されている。方法M100によって処理されたセグメントは、異なる動作によって処理された更に大きなセグメントのセグメント（すなわち、「サブフレーム」）でもあり得るが、逆もまた同様である。

【0035】

図4BはタスクT100の実現形態T102の流れ図を示す。各マイクロホンチャンネルに関しては、タスクT102は異なる周波数成分の各々に関してこのチャンネルに関する位相を推定するサブタスクT110のそれぞれの事例を含む。図4Cは、サブタスクT112およびT1122を含むタスクT110の実現形態T112の流れ図を示す。タスクT1121は、高速フーリエ変換（FFT）または離散コサイン変換（DCT）といったチャンネルの周波数変換を計算する。タスクT1121は典型的には、各セグメントに関するチャンネルの周波数変換を計算するように構成されている。例えば、各セグメントの128ポイントまたは256ポイントFFTを実行するようにタスクT1121を構成することは望ましい可能性がある。タスクT1121の代替実現形態は1バンクのサブバンドフィルタを使用してチャンネルの種々の周波数成分を分離するように構成されている。

【0036】

タスクT1122は異なる周波数成分（「ビン」とも呼ばれる）の各々に関するマイクロホンチャンネルの位相を計算する（例えば、推定する）。推定されるべき各周波数成分に関して、例えばタスクT1122は対応するFFT係数の虚数項対FFT係数の実数項の比の逆タンジェント（アークタンジェントとも呼ばれる）として位相を推定するように構成され得る。

【0037】

タスクT102はまた、各チャンネルに関して推定された位相に基づいて異なる周波数成分の各々に関する位相差 $\Delta\phi$ を計算するサブタスクT120を含む。タスクT120は、1つのチャンネルにおける周波数成分に関して推定された位相を、他のチャンネルにおける周波数成分に関して推定された位相から差し引くことによって、位相差を計算するように構成され得る。例えば、タスクT120は1次チャンネルにおける周波数成分に関して推定された位相を、もう1つの（例えば、2次）チャンネルにおける周波数成分に関して推定された位相から差し引くことによって、位相差を計算するように構成され得る。このような場合、1次チャンネルはデバイスの典型的な使用時に最も直接的にユーザのボイスを受け入れると予期されるマイクロホンに対応するチャンネルといった最も高い信号対雑音比を有すると予期されるチャンネルであり得る。

【0038】

周波数の広帯域範囲に亘るマルチチャネル信号のチャンネル間の位相差を推定するように方法M100（または、このような方法を実行するように構成されたシステムまたは装置）を構成することが望ましいことができる。このような広帯域範囲は、例えばゼロ、50、100または200Hzという低い周波数範囲から3、3.5または4kHzという高い周波数範囲（または、より高い最大7または8kHz以上といった）まで広がり得る。

しかしながら、タスクT100が信号の全帯域幅に亘って位相差を計算することは必要でない可能性がある。例えば、このような広帯域幅における多くの帯域に関して、位相推定は実用的でない、または不必要であることができる。極めて低い周波数において受信された波形の位相関係の実際的評価は典型的には、対応するようにトランスデューサ間の大きな間隔を必要とする。この結果、マイクロホン間の最大利用可能間隔は低周波数範囲を確立し得る。他方では、マイクロホン間の距離は空間エイリアシングを防止するために、最小波長の半分を超えるべきでない。例えば8kHzのサンプリング速度はゼロから4kHzの帯域幅を与える。4kHz信号の波長は約8.5センチメートルであるから、この場合、隣接マイクロホン間の間隔は約4センチメートルを超えるべきでない。マイクロホンチャンネルは空間エイリアシングを引き起こす可能性がある周波数を除去するために、ロー

10

【0039】

従って、タスクT1121によって作り出された周波数成分のすべてより少ないものに関して（例えば、タスクT1121によって実行されたFFTの周波数サンプルのすべてより少ないものに関して）位相推定値を計算するようにタスクT1122を構成することが望ましいことができる。例えば、タスクT1122は、約50、100、200、または300Hzから約500または1000Hzまでの周波数範囲に関して位相推定値を計算するように構成され得る（これら8つの組合せの各々は明確に考えられ、開示されている）。このような範囲が較正のために特に有用である成分を含み、較正のためにあまり有用でない成分を除外するであろうことは予期され得る。

20

【0040】

チャンネル較正以外の目的のために使用される位相推定値を計算するようにもタスクT100を構成することが望ましいことができる。例えば、タスクT100はまた、（例えば、下記により詳細に説明されるように）ユーザのボイスを追跡および／または増強するために使用される位相推定値を計算するようにも構成され得る。1つのこのような例では、タスクT1122はまたユーザのボイスのエネルギーの大部分を含むと予期され得る700Hzから2000Hzの周波数範囲に関する位相推定値を計算するようにも構成され得る。4kHz帯域幅信号の128ポイントFFTに関して、700～2000Hzの範囲は概して、第10番目サンプルから第32番目サンプルまでの23個の周波数サンプルに対応する。更なる例では、タスクT1122は約50、100、200、300、または

30

500Hzの低周波数範囲から約700、1000、1200、1500、または2000Hzの高周波数範囲に広がる周波数範囲に亘って位相推定値を計算するように構成されている（これら低周波数範囲および高周波数範囲の25個の組合せの各々は明確に考えられて開示されている）。

【0041】

レベル計算タスクT200は、マルチチャンネル信号の対応するセグメントにおける第1および第2のチャンネルの各々に関するレベルを計算するように構成されている。代替としてタスクT200は、マルチチャンネル信号の対応するセグメントの1セットのサブバンドの各々における第1および第2のチャンネルの各々に関するレベルを計算するように構成され得る。このような場合、タスクT200は同じ幅（例えば、500、1000または1200Hzの均一な幅）を有する1セットのサブバンドの各々に関するレベルを計算するように構成され得る。代替としてタスクT200は、サブバンドの少なくとも2つ（場合によってはすべて）が異なる幅を有する1セットのサブバンド（例えば、信号スペクトルのBarkまたはMelスケール分割による幅といった不均一幅を有する1セットのサブバンド）の各々に関するレベルを計算するように構成され得る。

40

【0042】

タスクT200は、対応する時間に亘る（例えば、対応するセグメントに亘る）チャンネルにおけるサブバンドの振幅またはマグニチュード（「絶対振幅」または「修正振幅」とも呼ばれる）の測定値としてタイムドメインにおける選択されたサブバンドの各チャンネルに関するレベルLを計算するように構成され得る。振幅またはマグニチュードの測定値の

50

例は、全マグニチュード、平均マグニチュード、二乗平均平方根（RMS）振幅、中央値マグニチュードおよびピークマグニチュードを含む。デジタルドメインにおいて、このような測定値は下記の1つといった式に従って、 n 個のサンプル値 x_t ， $t = 1, 2, \dots, n$ の1ブロック（または、「フレーム」）に亘って計算され得る：

【数1】

$$\sum_{i=1}^n |x_i| \quad (\text{全マグニチュード}); \quad (1)$$

$$\frac{1}{n} \sum_{i=1}^n |x_i| \quad (\text{平均マグニチュード}); \quad (2)$$

$$\sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2} \quad (\text{RMS 振幅}); \quad (3)$$

$$\text{median}_{i=1,2,\dots,n} |x_i| \quad (\text{中間値マグニチュード}); \quad (4)$$

$$\max_{i=1,2,\dots,n} |x_i| \quad (\text{ピークマグニチュード}). \quad (5)$$

【0043】

タスクT200はまた、このような式に従って周波数ドメイン（例えば、フーリエ変換ドメイン）またはもう1つの変換ドメイン（例えば、離散型コサイン変換（DCT）ドメイン）における選択されたサブバンドの各チャンネルに関するレベル L を計算するようにも構成され得る。タスクT200はまた、同様の式に従って（例えば、合計の代わりに積分を使用して）、アナログドメインにおけるレベルを計算するようにも構成され得る。

【0044】

代替として、タスクT200は、対応する時間に亘る（例えば、対応するセグメントに亘る）サブバンドのエネルギーの測定値としてタイムドメインにおける選択されたサブバンドの各チャンネルに関するレベル L を計算するように構成され得る。エネルギーの測定値の例は全エネルギーと平均エネルギーとを含む。デジタルドメインでは、これらの測定値は下記のような式に従って n 個のサンプル値 x_t ， $t = 1, 2, \dots, n$ の1ブロックに亘って計算され得る：

【数2】

$$\sum_{i=1}^n x_i^2 \quad (\text{全エネルギー}); \quad (6)$$

$$\frac{1}{n} \sum_{i=1}^n x_i^2 \quad (\text{平均エネルギー}). \quad (7)$$

【0045】

タスクT200はまた、このような式に従って、周波数ドメイン（例えば、フーリエ変換ドメイン）またはもう1つの変換ドメイン（例えば、離散型コサイン変換（DCT）ドメイン）における選択されたサブバンドの各チャンネルに関するレベル L を計算するようにも構成され得る。タスクT200はまた、同様の式に従って（例えば、合計の代わりに積分を使用して）アナログドメインにおけるレベルを計算するようにも構成され得る。更なる代替として、タスクT200は、対応する時間に亘る（例えば、対応するセグメントに亘る）サブバンドのパワースペクトル密度（PSD）として、選択されたサブバンドの各チャンネルに関するレベルを計算するように構成される。

【0046】

代替としてタスクT200は、タイムドメインにおける、または周波数ドメインにおける、またはもう1つの変換ドメインにおけるマルチチャネル信号の選択されたセグメントの各チャネル*i*に関するレベル*L_i*を、このチャネルにおけるセグメントの振幅、マグニチュードまたはエネルギーの測定値として計算するようにアナログ的仕方で構成され得る。例えば、タスクT200はセグメントのチャネルに関するレベル*L*を、このチャネルにおけるセグメントのタイムドメイン・サンプル値の2乗の合計として、またはこのチャネルにおけるセグメントの周波数ドメインサンプル値の2乗の合計として、またはこのチャネルにおけるセグメントのPSDとして、計算するように構成され得る。タスクT300によって処理されたセグメントはまた、異なる動作によって処理された、より大きなセグメントのセグメント（すなわち、「サブフレーム」）でもあり得るが、この逆もまた同様である。

10

【0047】

レベル値を計算する前にオーディオ信号チャネルに1つ以上のスペクトル形成動作を実行するようにタスクT200を構成することが望ましいことができる。このような動作はアナログおよび／またはデジタルドメインにおいて実行され得る。例えば、対応するレベル値（単数またな複数）を計算する前にそれぞれのチャネルからの信号に、（例えば、200、500または1000Hzの遮断周波数を有する）ローパスフィルタまたは（例えば、200Hzから1kHzのパスバンドを有する）バンドパスフィルタを適用するようにタスクT200を構成することが望ましいことができる。

20

【0048】

利得係数更新タスクT300は、計算されたレベルに基づいて少なくとも1つの利得係数の各々に関する値を更新するように構成されている。例えば、タスクT200によって計算された対応する選択された周波数成分における各チャネルのレベル間の観測された不均衡に基づいて利得係数値の各々を更新するようにタスクT300を構成することは望ましい可能性がある。

【0049】

タスクT300のこのような実現形態は、観測された不均衡を線形レベル値の関数として（例えば、 L_1 および L_2 がそれぞれ第1および第2のチャネルのレベルを表す場合に L_1/L_2 といった式にしたがう比として）計算するように構成され得る。代替として、タスクT300のこのような実現形態は、観測された不均衡を対数ドメインにおけるレベル値の関数として（例えば、 $L_1 - L_2$ といった式にしたがう差として）計算するように構成され得る。

30

【0050】

タスクT300は、観測された不均衡に対応する周波数成分に関する更新された利得係数値として使用するように構成され得る。代替として、タスクT300は利得係数の対応する前の値を更新するために観測された不均衡を使用するように構成され得る。このような場合、タスクT300は下記のような式に従って、更新された値を計算するように構成され得る：

【数3】

40

$$G_{in} = (\mu_i)G_{i(n-1)} + (1 - \mu_i)R_{in}, \quad (8)$$

【0051】

ここで、 G_{in} は周波数成分*i*に関するセグメント*n*に対応する利得係数値を表し、 $G_{i(n-1)}$ は周波数成分*i*に関する前のセグメント(*n*−1)に対応する利得係数値を表し、 R_{in} はセグメント*n*における周波数成分*i*に関して計算された観測された不均衡を表し、そして μ_i は0.3、0.5または0.7といった0.1（最大平滑化）から1（無平滑化）までの範囲内の値を有する時間的平滑化係数を表す。タスクT300のこのような実現形態が各周波数成分に関して平滑化係数 μ_i の同じ値を使用することは典型的で

50

はあるが必要ではない。観測された不均衡の計算に先立って観測されたレベルの値を時間的に平滑化するように、および／または更新された利得係数値の計算に先立って観測されたチャンネル不均衡の値を時間的に平滑化するように、タスクT300を構成することも可能である。

【0052】

下記に更に詳細に説明されるように、利得係数更新タスクT300はまた、タスクT100において計算された複数の位相差からの情報（例えば、マルチチャンネル信号の音響的に均衡した部分の識別情報）に基づいて少なくとも1つの利得係数の各々に関する値を更新するようにも構成されている。マルチチャンネル信号の何らかの特定のセグメントにおいてタスクT300は、1セットの利得係数値のすべてより少ないものを更新し得る。例えば、較正動作時に周波数成分を音響的に不均衡状態のままに留まらせる音源の存在は、タスクT300が観測された不均衡とこの周波数成分に関する新しい利得係数値とを計算するのを妨げることができる。その結果、周波数に亘って観測されたレベル、観測された不均衡および／または利得係数の値を平滑化するようにタスクT300を構成することが望ましいことができる。例えば、タスクT300は選択された周波数成分の観測されたレベルの（または、観測された不均衡または利得係数の）平均値を計算して、この計算された平均値を、無選択周波数成分に割り当てるように構成され得る。もう1つの例では、タスクT300は下記のような式に従って無選択周波数成分*i*に対応する利得係数値を更新するように構成されている：

【数4】

$$G_{in} = (\beta)G_{i(n-1)} + (1-\beta)G_{(i-1)n}, \quad (9)$$

【0053】

ここで、 G_{in} は周波数成分*i*に関するセグメント*n*に対応する利得係数値を表し、 $G_{i(n-1)}$ は周波数成分*i*に関する前のセグメント(*n*−1)に対応する利得係数値を表し、 $G_{(i-1)n}$ は近隣周波数成分(*i*−1)に関するセグメント*n*に対応する利得係数値を表し、そしてβはゼロ（無更新）から1（無平滑化）までの範囲内の値を有する周波数平滑化係数を表す。更なる例では、式(9)は $G_{(i-1)n}$ の代わりに最も近い選択された周波数成分に関する利得係数値を使用するように変更される。タスクT300は時間的平滑化として同じ時間の前または後における、または同じ時間における周波数に亘って平滑化を実行するように構成され得る。

【0054】

タスクT400は、タスクT300において更新された少なくとも1つの利得係数値に基づいて、マルチチャンネル信号のもう1つのチャンネルの対応する応答特性に関してマルチチャンネル信号の1つのチャンネルの応答特性（例えば、利得応答）を修正することによって、処理されたマルチチャンネル信号（「均衡化された」または「較正された」信号とも呼ばれる）を作り出す。タスクT400は、第1のチャンネルにおける周波数成分の振幅に関して第2のチャンネルにおける対応する周波数成分の振幅を変えるために1セットのサブバンド利得係数値の各々を使用することによって、処理されたマルチチャンネル信号を作り出すように構成され得る。タスクT400は、例えば応答性のより小さいチャンネルからの信号を増幅するように構成され得る。代替として、タスクT400は、2次マイクロホンに対応するチャンネルにおける周波数成分の振幅を制御する（例えば、増幅する、または減衰させる）ように構成され得る。上記のように、マルチチャンネル信号の任意の特定のセグメントにおいて、1セットの利得係数値のすべてより少ないものが更新されることは可能である。

【0055】

タスクT400は、単一の利得係数値を信号の各セグメントに適用することによって、

またはそうではなく利得係数値を1つより多い周波数成分に適用することによって、処理されたマルチチャンネル信号を作り出すように構成され得る。例えば、タスクT400は、1次マイクロホンチャンネルの対応する振幅に関して2次マイクロホンチャンネルの振幅を修正するために（例えば、1次マイクロホンチャンネルに関して2次マイクロホンチャンネルを増幅する、または減衰させるために）更新された利得係数値を適用するように構成され得る。

【0056】

タスクT400は、線形ドメインにおいてチャンネル応答均衡化を実行するように構成され得る。例えば、タスクT400は第2のチャンネルにおけるセグメントのタイムドメイン・サンプルの値の各々にこのセグメントに対応する利得係数の値を掛けることによって、セグメントの第2のチャンネルの振幅を制御するように構成され得る。サブバンド利得係数に関して、タスクT400は、第2のチャンネルにおける対応する周波数成分の振幅に利得係数の値を掛けることによって、またはタイムドメインにおける対応するサブバンドに利得係数を適用するためにサブバンドフィルタを使用することによって、第2のチャンネルにおける対応する周波数成分の振幅を制御するように構成され得る。

10

【0057】

代替として、タスクT400は対数ドメインにおいてチャンネル応答均衡化を実行するように構成され得る。例えば、タスクT400は、利得係数の対応する値をセグメントの持続時間に亘ってセグメントの第2のチャンネルに適用された対数利得制御値に加えることによって、セグメントの第2のチャンネルの振幅を制御するように構成され得る。サブバンド利得係数に関して、タスクT400は、対応する利得係数の値を第2のチャンネルにおける周波数成分の振幅に加えることによって、第2のチャンネルにおける周波数成分の振幅を制御するように構成され得る。このような場合、タスクT400は、対数値（例えば、デシベル単位の）として振幅および利得係数値を受信するように、および／または（例えば、 x_{lin} は線形値であり、 x_{log} は対応する対数値であるとして $x_{log} = 20 \log x_{lin}$ のような式に従って）線形振幅または利得係数値を対数値に変換するように構成され得る。

20

【0058】

タスクT400は、チャンネル（単数または複数）の他の振幅制御（例えば、自動利得制御（AGC）または自動ボリューム制御（AVC）モジュール、ユーザ操作ボリューム制御など）と組み合され得る、または他の振幅制御の上流または下流で実行され得る。

30

【0059】

2つより多いマイクロホンのアレイに関して、各チャンネルの応答が少なくとも1つの他のチャンネルの応答に均衡するように、2対以上のチャンネルの各々に方法M100のそれぞれの事例を実行することが望ましいことができる。例えば、方法M100の1つの事例（例えば、方法M110）は1対のチャンネル（例えば、第1および第2のチャンネル）に基づいてコヒーレンシー測定値を計算するように実行され得るが、方法M100のもう1つの事例はもう1対のチャンネル（例えば、第1のチャンネルおよび第3のチャンネル、または第3および第4のチャンネル）に基づいてコヒーレンシー測定値を計算するように実行される。しかしながら、1対のチャンネルに対して共通の動作が実行されない場合には、この1対のチャンネルの均衡化は省略され得る。

40

【0060】

利得係数更新タスクT300は、各チャンネルにおいて同じレベルを有すると予期されるマルチチャンネル信号の周波数成分および／またはセグメント（例えば、本明細書で「音響的に均衡した部分」とも呼ばれるそれぞれのマイクロホンチャンネルによって等しい応答をもたらすと予期される周波数成分および／またはセグメント）を示すために、またこれらの部分からの情報に基づいて1つ以上の利得係数値を計算するために、計算された位相差からの情報を使用することを含み得る。アレイR100の幅広面方向における音源から受信されたサウンド成分はマイクロホンMC10およびMC20によって等しい応答をもたらすことが予期され得る。これとは逆に、アレイR100のエンドファイア方向のいずれ

50

かにおける近距離音場音源から受信されたサウンド成分は、一方のマイクロホンに他方のマイクロホンより高い出力レベルを持たせる（すなわち、「音響的に不均衡化される」）ことが予期され得る。従って、マルチチャネル信号の対応する周波数成分が音響的に均衡化されるか、音響的に不均衡化されるかどうかを決定するために、タスクT100において計算された位相差を使用するようにタスクT300を構成することが望ましいことができる。

【0061】

タスクT300は、対応する周波数成分の各々に関するマスクスコア（mask score）を取得するために、タスクT100によって計算された位相差に指向性マスク演算を実行するように構成され得る。限定された周波数範囲に亘るタスクT100による位相推定に関する上記の論議によれば、タスクT300は信号の周波数成分のすべてより少ないものに関する（例えば、タスクT1121によって実行されたFFTの周波数サンプルのすべてより少ないものに関する）マスクスコアを取得するように構成され得る。

10

【0062】

図5AはサブタスクT310、T320およびT340を含むタスクT300の実現形態T302の流れ図を示す。タスクT100からの複数の計算された位相差の各々に関して、タスクT310は対応する方向インジケータを計算する。タスクT320は方向インジケータを評価するために（例えば、方向インジケータの値を振幅またはマグニチュードスケールにおける値に変換またはマッピングするために）、指向性マスク関数を使用する。タスクT320によって作り出された評価に基づいて、タスクT340は（例えば、上記の式（8）または（9）に従って）更新された利得係数値を計算する。例えば、タスクT340は、信号の周波数成分が音響的に均衡化されていることを評価が示す信号の周波数成分を選択するように、そしてこの成分に関するチャネル間の観測された不均衡に基づくこれらの成分の各々に関する更新された利得係数値を計算するように、構成され得る。

20

【0063】

タスクT310は、マルチチャネル信号の対応する周波数成分 f_i の到着方向 θ_i として方向インジケータの各々を計算するように構成され得る。例えば、タスクT310は、 c がサウンドの速度（約340m/sec）を表し、 d がマイクロホン間の距離を表し、 $\Delta\phi_i$ が2つのマイクロホンに関する対応する位相推定値間の差をラジアン単位で表し、 f_i が位相推定値が対応する周波数成分（例えば、対応するFFTサンプルの周波数、または対応するサブバンドの中心周波数またはエッジ周波数）である場合に、量 $c\Delta\phi_i/d2\pi f_i$ の逆コサイン（アークコサインとも呼ばれる）として到着方向 θ_i を推定するように構成され得る。代替として、タスクT310は、 λ_i が周波数成分 f_i の波長を表す場合に、量 $\lambda_i\Delta\phi_i/d2\pi$ の逆コサインとして到着方向 θ_i を推定するように構成され得る。

30

【0064】

図6Aは2マイクロホンアレイMC10、MC20のマイクロホンMC20に関する到着方向 θ を推定する、このアプローチを示す幾何学的近似の一例を示す。この例では、 $\theta_i=0$ という値は基準エンドファイア方向（すなわち、マイクロホンMC10の方向）からマイクロホンMC20に到着する信号を表し、 $\theta_i=\pi$ という値は他のエンドファイア方向から到着する信号を表し、 $\theta_i=\pi/2$ という値は幅広面方向から到着する信号を表す。別の例では、タスクT310は、異なる基準位置（例えば、マイクロホンMC10、またはマイクロホン間の中間点といった他のポイント）および／または異なる基準方向（例えば、他のエンドファイア方向、幅広面方向など）に関して θ_i を評価するように構成され得る。

40

【0065】

図6Aに示されている幾何学的近似は距離 s が距離 L に等しいことを想定しており、ここで s はマイクロホンMC20の位置と、音源とマイクロホンMC20との間の直線へのマイクロホンMC10の位置の直交投影と、の間の距離であり、 L は各マイクロホンの音

50

源までの距離間の実際の差である。誤差 ($s - L$) は、マイクロホン M C 2 0 に関する到着方向 θ がゼロに近づくにつれて小さくなる。この誤差はまた、音源とマイクロホンアレイとの間の相対的距離が増加するにつれて小さくなる。

【0066】

図 6 A に示されている方式は、 $\Delta \phi_i$ の第 1 象限および第 4 象限の値 (すなわち、ゼロから $+\pi/2$ およびゼロから $-\pi/2$) のために使用され得る。図 6 B は、 $\Delta \phi_i$ の第 2 象限および第 3 象限の値 (すなわち、 $+\pi/2$ から $-\pi/2$) のために同じ近似を使用する一例を示す。この場合、到着方向 θ_i を生み出すために π ラジアンから差し引かれる角度 ζ を評価するために、上記のように逆コサインが計算され得る。現役のエンジニアは、到着方向 θ_i が度で表され得ること、またはラジアンに代わりに特定用途のために適当な他の任意の単位で表され得ることを理解するであろう。

10

【0067】

$\pi/2$ ラジアンに近い到着方向 (アレイの幅広面方向) を有する周波数成分を選択するようにタスク T 3 0 0 を構成することが望ましいことができる。その結果、一方における $\Delta \phi_i$ の第 3 および第 4 象限の値と他方における $\Delta \phi_i$ の第 2 および第 3 象限の値との間の差異は較正目的のためには重要でなくなる。

【0068】

代替実現形態では、タスク T 3 1 0 はマルチチャネル信号の対応する周波数成分 f_i の到着時間遅延 τ_i (例えば、秒単位の) として方向インジケータの各々を計算するように構成されている。タスク T 3 1 0 は $\tau_i = \lambda_i \Delta \phi_i / c 2 \pi$ または $\tau_i = \Delta \phi_i / 2 \pi f_i$ といった数式を使用してマイクロホン M 1 0 に関連してマイクロホン M 2 0 における到着時間遅延 τ_i を推定するように構成され得る。これらの例では、 $\tau_i = 0$ という値は幅広面方向から到着する信号を表し、 τ_i の大きな正の値は基準エンドファイア方向から到着する信号を表し、 τ_i の大きな負の値は他の基準エンドファイア方向から到着する信号を表す。値 τ_i を計算する際に、サンプリング周期といった特定の用途のために適当であると考えられる時間の単位 (例えば、8 kHz のサンプリング速度のための 125 マイクロ秒単位) または秒の何分の一 (例えば、 10^{-3} 、 10^{-4} 、 10^{-5} または 10^{-6} 秒) を使用することが望ましいことができる。タスク T 3 1 0 がタイムドメインにおける各チャネルの周波数成分 f_i を相互相関させることによって到着時間遅延 τ_i を計算するようにも構成され得ることに留意されたい。

20

30

【0069】

同じポイントの音源から直接到着するサウンド成分に関して、 $\Delta \phi / f$ の値は理想的にはすべての周波数に関して定数 k に等しく、ここで、 k の値は到着方向 θ と到着時間遅延 τ とに関連する。もう 1 つの代替実施形態ではタスク T 3 1 0 は、推定された位相差 $\Delta \phi_i$ と周波数 f_i との比 r_i (例えば、 $r_i = \Delta \phi_i / f_i$ または $r_i = f_i / \Delta \phi_i$) として方向インジケータの各々を計算するように構成されている。

【0070】

式 $\theta_i = \cos^{-1} (c \Delta \phi_i / d 2 \pi f_i)$ または $\theta_i = \cos^{-1} (\lambda_i \Delta \phi_i / d 2 \pi)$ は、遠距離音場モデル (すなわち、平面波面を想定したモデル) に従って方向インジケータ θ_i を計算するが、式 $\tau_i = \lambda_i \Delta \phi_i / c 2 \pi$ 、 $\tau_i = \Delta \phi_i / 2 \pi f_i$ 、 $r_i = \Delta \phi_i / f_i$ および $r_i = f_i / \Delta \phi_i$ は近距離音場モデル (すなわち、図 7 に示されているような球形波面を想定したモデル) に従って方向インジケータ τ_i および r_i を計算する。近距離音場モデルに基づく方向インジケータは計算することがより正確および/または容易である結果を与え得るが、遠距離音場モデルに基づく方向インジケータは方法 M 1 0 0 のある幾つかの構成のために望ましいことができる位相差と方向インジケータとの間の非線形マッピングを与える。

40

【0071】

タスク T 3 0 2 はまた、タスク T 3 1 0 によって作り出された方向インジケータを評価するサブタスク T 3 2 0 を含む。タスク T 3 2 0 は、吟味されるべき周波数成分に関して、方向インジケータの値を振幅、マグニチュード、または合格/不合格 (pass/fail) ス

50

ケール（「マスキスコア」とも呼ばれる）についての対応する値に変換またはマッピングすることによって、方向インジケータを評価するように構成され得る。例えば、タスク T 3 2 0 は、指示された方向がマスキング関数のパスバンド内に入るかどうか（および／またはどれほどうまく入るか）を示すマスキスコアに各方向インジケータの値をマッピングするために、指向性マスキング関数を使用するように構成され得る。（この文脈では、用語「パスバンド」はマスキング関数によって通された到着方向の範囲を指す。）種々の周波数成分に関するこの 1 セットのマスキスコアはベクトルと考えられ得る。タスク T 3 2 0 は種々の方向インジケータを連続しておよび／または並行して評価するように構成され得る。

【0072】

10

マスキング関数のパスバンドは所望の信号方向を含むように選択され得る。マスキング関数の空間選択性はパスバンドの幅を変えることによって制御され得る。例えば、収束速度と較正精度との間のトレードオフに従ってパスバンド幅を選択することが望ましいことができる。より幅広いパスバンドは周波数成分のより多くが較正動作に寄与することを可能にすることによって、より速い収束を可能にし得るが、アレイの幅広面軸からより遠い方向から到着する成分を受け入れることによって、より不正確になることも预期されるであろう（従って、マイクロホンに異なる影響を与えることが预期され得る）。1 つの例では、タスク T 3 0 0（例えば、下記のようなタスク T 3 2 0 またはタスク T 3 3 0）はアレイの幅広面軸の 50 度以内の方向から到着する成分（すなわち、75～105 度または同等に $5\pi/12 \sim 7\pi/12$ ラジアン）の範囲内の到着方向を有する成分）を選択するよう

20

【0073】

図 8 A は、パスバンドとストップバンド（「ブリックウォール (brickwall)」プロファイルとも呼ばれる）との間の比較的急な遷移と、到着方向 $\theta = \pi/2$ に中心を持つパスバンドと、を有するマスキング関数の一例を示す。1 つのこのような場合には、タスク T 3 2 0 は、方向インジケータがマスキング関数のパスバンド内の方向を示すときに第 1 の値（例えば、1）を有するバイナリ値マスキスコアを割り当て、方向インジケータがこの関数のパスバンド外の方向を示すときには第 2 の値（例えば、ゼロ）を有するマスキスコアを割り当てるように構成されている。信号対雑音比（SNR）、雑音レベルなどといった 1 つ以上の因子に依存してストップバンドとパスバンドとの間の遷移の位置を変えること（例えば、SNR が高いときに、より狭いパスバンドを使用して較正精度に悪影響を与え得る所望の指向性信号の存在を示すこと）が望ましいことができる。

30

【0074】

代替として、パスバンドとストップバンドとの間にあまり急激でない遷移（例えば、非 2 成分値マスキスコアを生み出す、より緩やかなロールオフ）を有するマスキング関数を使用するようにタスク T 3 2 0 を構成することが望ましいことができる。図 8 B は到着方向 $\theta = \pi/2$ に中心を持つパスバンドを有するマスキング関数に関する線形ロールオフの一例を示し、図 8 C は到着方向 $\theta = \pi/2$ に中心を持つパスバンドを有するマスキング関数に関する非線形ロールオフの一例を示す。SNR、雑音レベルなどといった 1 つ以上の因子に依存して、ストップバンドとパスバンドとの間の遷移の位置および／または急激さ

40

を変えること（例えば、SNR が高いときに、より急激なロールオフを使用して較正精度に悪影響を与え得る所望の指向性信号の存在を示すこと）が望ましいことができる。マスキング関数（例えば、図 8 A～図 8 C に示されているような）が、方向 θ よりむしろ時間遅延 τ または比 $r = \Delta\phi/f$ に対応する。

【0075】

非線形マスキング関数の一例は、

【数 5】

$$m = \frac{1}{1 + \exp(\gamma[|\theta - \theta_T| - (\frac{w}{2})])}$$

【0076】

のように表され得、ここで、 ζ_T は目標到着方向を表し、 w はラジアン単位で所望のマスク幅を表し、 γ は急激さパラメータを表す。図9A～図9Cは、それぞれ $(8, \pi/2, \pi/2)$ 、 $(20, \pi/4, \pi/2)$ 、および $(50, \pi/8, \pi/5)$ に等しい (γ, w, θ_T) に関するこのような関数の例を示す。このような関数が方向 θ よりむしろ時間遅延 τ または比 r の観点からも表され得ることは無論である。SNR、雑音レベルなどといった1つ以上の因子に依存してマスクの幅および/または急激さを変えること（例えば、SNRが高いときに、より狭いマスクおよび/またはより急激なロールオフを使用すること）が望ましいことができる。

10

【0077】

図5Bは、タスクT300の代替実現形態の流れ図を示す。複数の方向インジケータの各々を評価するために同じマスク関数を使用する代わりに、タスクT304は、対応する指向性マスク関数 m_i を使用して各位相差 $\Delta\phi_i$ を評価し、計算された位相差を方向インジケータとして使用するサブタスクT330を含む。例えば、 θ_L から θ_H までの範囲内の方向から到着する音響成分を選択することが望まれる場合には、各マスク関数 m_i は、 $\Delta\phi_L = (d/2\pi f_i/c) \cos\theta_H$ （同等に、 $\Delta\phi_L = (d/2\pi/\lambda_i) \cos\theta_H$ ）および $\Delta\phi_H = (d/2\pi f_i/c) \cos\theta_L$ （同等に、 $\Delta\phi_H = (d/2\pi/\lambda_i) \cos\theta_L$ ）として、 $\Delta\phi_L$ から $\Delta\phi_H$ の範囲にあるパスバンドを有するように構成され得る。 τ_L から τ_H の到着時間遅延の範囲に対応する方向から到着するサウンド成分を選択することが望まれる場合には、各マスク関数 m_i は、 $\Delta\phi_{Li} = 2\pi f_i \tau_L$ （同等に、 $\Delta\phi_{Li} = c/2\pi \tau_L/\lambda_i$ ）および $\Delta\phi_{Hi} = 2\pi f_i \tau_H$ （同等に、 $\Delta\phi_{Hi} = c/2\pi \tau_H/\lambda_i$ ）として、 $\Delta\phi_{Li}$ から $\Delta\phi_{Hi}$ の範囲にあるパスバンドを有するように構成され得る。 r_L から r_H の位相差対周波数の比の範囲に対応する方向から到着するサウンド成分を選択することが望まれる場合には、各マスク関数 m_i は、 $\Delta\phi_{Li} = f_i r_L$ および $\Delta\phi_{Hi} = f_i r_H$ として、 $\Delta\phi_{Li}$ から $\Delta\phi_{Hi}$ の範囲にあるパスバンドを有するように構成され得る。タスクT320に関して前に論じられたように、各マスク関数のプロファイルはSNR、雑音レベルなどといった1つ以上の因子に従って選択され得る。

20

30

【0078】

周波数成分の1つ以上（場合によってはすべて）の各々に関するマスクスコアを時間的に平滑化された値として作り出すようにタスクT300を構成することが望ましいことができる。タスクT300のこのような実現形態は、 m の可能な値が5、10、20、および50を含むとして、このような値をごく最近の m 個のフレームに亘る周波数成分に関するマスクスコアの平均値として計算するように構成され得る。より一般的には、タスクT300のこのような実現形態は、有限または無限インパルス応答（FIRまたはIIR）フィルタといった時間的平滑化関数を使用して平滑化された値を計算するように構成され得る。1つのこのような例では、 $v_i(n-1)$ は前のフレームのための周波数成分 i に関するマスクスコアの平滑化された値を表し、 $c_i(n)$ は周波数成分 i に関するマスクスコアの現在値を表し、 α_i はゼロ（無平滑化）から1（無更新）までの範囲から選択され得る平滑化係数であるとした場合に、タスクT300は、 $v_i(n) = \alpha_i v_i(n-1) + (1 - \alpha_i) c_i(n)$ といった式に従ってフレーム n の周波数成分 i に関するマスクスコアの平滑化された値 $v_i(n)$ を計算するように構成されている。この1次IIRフィルタはまた「漏洩積分器（leaky integrator）」とも呼ばれ得る。

40

【0079】

平滑化係数 α_i の典型的な値は0.99、0.09、0.95、0.9および0.8を

50

含む。タスクT300が1フレームの各周波数成分に関して α_i の同じ値を使用することは、典型的ではあるが必要ではない。初期収束期間の間（例えば、オーディオセンシング回路の電源投入または他の活性化動作の直後）に、タスクT300がより短い間隔に亘って、平滑化された値を計算すること、または引き続き定常状態動作中より小さい値を平滑化係数 α_i の1つ以上（場合によってはすべて）に関して使用すること、が望ましいことができる。

【0080】

タスクT340は信号の音響的に均衡化された部分を選択するために複数のマスクスコアからの情報を使用するように構成され得る。タスクT340は音響的均衡の方向インジケータとして、2成分値マスクスコアを採用するように構成され得る。例えば、パスバンドがアレイR100の幅広面方向にあるマスクに関して、タスクT340は1というマスクスコアを有する周波数成分を選択するように構成され得るが、パスバンドがアレイR100のエンドファイア方向（例えば、図3Bに示されているような）にあるマスクに関しては、タスクT340はゼロというマスクスコアを有する周波数成分を選択するように構成され得る。

10

【0081】

非2成分値マスクスコアの場合には、タスクT340はマスクスコアをある閾値と比較するように構成され得る。例えば、パスバンドがアレイR100の幅広面方向にあるマスクに関しては、マスクスコアが閾値より大きい（代替として、小さくない）場合に、タスクT340が周波数成分を音響的に均衡化された部分として識別することが望ましいことができる。同様に、パスバンドがアレイR100のエンドファイア方向にあるマスクに関しては、マスクスコアが閾値より小さい（代替として、小さくない）場合に、タスクT340が周波数成分を音響的に均衡化された部分として識別することが望ましいことができる。

20

【0082】

タスクT340のこのような実現形態は、周波数成分のすべてに関して同じ閾値を使用するように構成され得る。代替として、タスクT340は周波数成分の2つ以上（場合によってはすべて）の各々に関して異なる閾値を使用するように構成され得る。タスクT340は、一定の閾値（単数または複数）を使用するように構成され得るが、代替として信号の特性（例えば、フレームエネルギー）および／またはマスクの特性（例えば、パスバンド幅）に基づいて経過時間に亘って1つのセグメントからもう1つのセグメントに閾値（単数または複数）を適応させるように構成され得る。

30

【0083】

図5Cは、タスクT200の実現形態T205と；タスクT300（例えば、タスクT302またはT304）の実現形態T305と；タスクT400の実現形態T405と；を含む方法M100の実現形態M200の流れ図を示す。タスクT205は（少なくとも）2つのサブバンドの各々における各チャンネルに関するレベルを計算するように構成されている。タスクT305はこれらのサブバンドの各々に関する利得係数値を更新するように構成され、またタスクT405はサブバンドにおける第1のチャンネルの振幅に関して対応するサブバンドにおける第2のチャンネルの振幅を修正するために各更新された利得係数を適用するように構成されている。

40

【0084】

信号が理想的なポイント音源から残響なしに受信されるときには、すべての周波数成分は同じ到着方向を持つはずである（例えば、比 $\Delta\phi/f$ の値はすべての周波数に亘って一定であるはずである）。信号の異なる周波数成分が同じ到着方向を有する度合いは「方向的コヒーレンス」とも呼ばれる。マイクロホンアレイが遠距離音場（例えば、背景雑音源）から発生したサウンドを受けるとき、結果として得られたマルチチャンネル信号は、典型的には近距離音場音源から発生する受信サウンド（例えば、ユーザのボイス）に関するより方向的コヒーレンスに乏しいであろう。例えば、異なる周波数成分の各々におけるマイクロホンチャンネル間の位相差は典型的には、近距離音場音源から発生する受信サウンドに

50

関する周波数より遠距離音場音源から発生する受信サウンドに関する周波数に、より少ない相関性を持つであろう。

【0085】

マルチチャネル信号の一部分（例えば、セグメントまたはサブバンド）が音響的に均衡化されているか、音響的に不均衡化されているかどうかを示すために、到着方向と同様に方向的コヒーレンスを使用するようにタスクT300を構成することが望ましいことができる。例えば、これらの部分における周波数成分が方向的にコヒーレントである度合いに基づいて、マルチチャネル信号の音響的に均衡化された部分を選択するように、タスクT300を構成することが望ましいことができる。方向的コヒーレンスの使用は、例えばアレイのエンドファイア方向に位置する方向的にコヒーレントな音源（例えば、近距離音場音源）による活動を含むセグメントまたはサブバンドの拒絶を可能にすることによって、チャンネル較正動作の高められた精度および／または信頼度をサポートし得る。

10

【0086】

図10は、タスクT300の一実現形態によって、2マイクロホンアレイR100からマルチチャネル信号に適用され得るようなマスキング関数の指向性パターンの前方および後方ローブを示す。アレイR100の幅広面方向における近距離音場音源または任意方向における遠距離音場音源といったこのパターンの外側に位置する音源から受信されたサウンド成分は、音響的に均衡化されるであろう（すなわち、マイクロホンMC10およびMC20による等しい応答をもたらすであろう）ことが予期され得る。同様に、このようなパターンの前方または後方ローブ内の音源（すなわち、アレイR100のエンドファイア方向のいずれかにおける近距離音場音源）から受信されたサウンド成分は、音響的に不均衡化されるであろう（すなわち、一方のマイクロホンが他方のマイクロホンより高い出力レベルを持たせるであろう）ことが予期され得る。従って、このようなマスキング関数パターンのいずれのローブ内にも音源を持たないセグメントまたはサブバンド（例えば、方向的にコヒーレントでない、または幅広面方向にだけコヒーレントであるセグメントまたはサブバンド）を選択するように、タスクT300の対応する実現形態を構成することが望ましいことができる。

20

【0087】

上記のように、タスクT300はマルチチャネル信号の音響的に均衡化された部分を識別するために、タスクT100によって計算された位相差からの情報を使用するように構成され得る。識別されたサブバンドまたはセグメントに関してだけ、対応する利得係数値の更新が実行されるように、タスクT300はサブバンドまたはセグメントがアレイの幅広面方向において方向的にコヒーレントである（または、代替として、エンドファイア方向には方向的にコヒーレントでない）ことをマスクスコアが示す信号のサブバンドまたはセグメントとして、音響的に均衡化された部分を識別するように実現され得る。

30

【0088】

図11Aは、タスクT300の実現形態T306を含む方法M100の実現形態M110の流れ図を示す。タスクT306は、タスクT100によって計算された位相差からの情報に基づいて、コヒーレンシー測定値を計算するサブタスクT360を含む。図11Bは、上記のサブタスクT312およびT322の事例とサブタスクT350とを含むタスクT360の実現形態T362の流れ図を示す。図11Cは、上記のサブタスクT332の事例とサブタスクT350とを含むタスクT360の実現形態T364の流れ図を示す。

40

【0089】

タスクT350は、サブバンドに関するコヒーレンシー測定値を取得するために各サブバンドにおける周波数成分のマスクスコアを組み合わせるように構成され得る。1つのこのような例では、タスクT350は特定の状態を有するマスクスコアの数に基づいてコヒーレンシー測定値を計算するように構成されている。もう1つの例では、タスクT350はマスクスコアの合計としてコヒーレンシー測定値を計算するように構成されている。更なる例では、タスクT350はマスクスコアの平均値としてコヒーレンシー測定値を計算

50

するように構成されている。これらのケースのいずれにおいても、タスクT350はマスクスコアの各々を等しく重み付けするように、（例えば、各マスクスコアを1で重み付けするように）または1つ以上のマスクスコアを互いに異なるように重み付けするように（例えば、低周波数または高周波数成分に対応するマスクスコアを中音域周波数成分に対応するマスクスコアより重くなく重み付けするように）構成され得る。

【0090】

パスバンドがアレイR100の幅広面方向にあるマスク（例えば、図8A～図8Cおよび図9A～図9Cに示されているような）に関して、タスクT350は、例えばマスクスコアの合計または平均がある閾値より小さくない（代替として、より大きい）場合に、またはサブバンドにおける少なくとも最小数の（代替として、最小数より多い）周波数成分が1というマスクスコアを有する場合に第1の状態（例えば、高い、または「1」）を有し、そうでない場合には第2の状態（例えば、低い、または「0」）を有するコヒーレンシー指示を作り出すように構成され得る。パスバンドがアレイR100のエンドファイア方向にあるマスクに関して、タスクT350は、例えばマスクスコアの合計または平均がある閾値より大きくない（代替として、より小さい）場合に、またはサブバンドにおける最大数より大きくない（代替として、より小さい）数の周波数成分が1というマスクスコアを有する場合に第1の状態を有し、そうでない場合には第2の状態を有するコヒーレンシー測定値を作り出すように構成され得る。

【0091】

タスクT350は、各サブバンドに関して同じ閾値を使用するように、またはサブバンドの2つ以上（おそらくはすべて）の各々に関して異なる閾値を使用するように、構成され得る。各閾値は発見的に決定される可能性があり、またパスバンド幅、信号の1つ以上の特性（例えば、SNR、雑音レベル）などといった1つ以上の因子に依存して経過時間に亘って閾値を変えることが望ましいことができる。（同じ原理は前のパラグラフで述べられた最大および最小数に当てはまる。）

代替として、タスクT350はマルチチャネル信号の一連のセグメントの各々に関して、対応する方向的コヒーレンシー測定値を作り出すように構成され得る。この場合、タスクT350は（例えば、上記のように、特定の状態を有するマスクスコアの数に基づいて、またはマスクスコアの合計または平均に基づいて）セグメントに関するコヒーレンシー測定値を取得するために、各セグメントにおける周波数成分の2つ以上（場合によってはすべて）のマスクスコアを組み合わせるように構成され得る。タスクT350のこのような実現形態は各セグメントに関して同じ閾値を使用するように、または上記のように1つ以上の因子に依存して経過時間に亘って閾値を変えるように構成され得る（例えば、最大数または最小数のマスクスコアに同じ原理が当てはまる）。

【0092】

セグメントのすべての周波数成分のマスクスコアに基づいて各セグメントに関するコヒーレンシー測定値を計算するようにタスクT350を構成することが望ましいことができる。代替として、限定された周波数範囲に亘る周波数成分のマスクスコアに基づいて各セグメントに関するコヒーレンシー測定値を計算するようにタスクT350を構成することが望ましいことができる。例えば、タスクT350は、約50、100、200、または300Hzから約500または1000Hzの周波数範囲（これら8つの組合せの各々は明確に考えられて開示されている）に亘る周波数成分のマスクスコアに基づいてコヒーレンシー測定値を計算するように構成され得る。例えば、チャンネルの応答特性間の差はこのような周波数範囲に亘るチャンネルの利得応答における差によって実質的に特徴付けられることが決定され得る。

【0093】

タスクT340は、タスクT360によって識別された音響的に均衡化された部分からの情報に基づいて少なくとも1つの利得係数の各々に関する更新された値を計算するように構成され得る。例えば、対応するセグメントまたはサブバンドにおいてマルチチャネル信号が方向的にコヒーレントであるという指示に応じて（例えば、対応するコヒーレンス

指示の状態によって示されるようにタスクT360におけるサブバンドまたはセグメントの選択に応じて)、更新された利得係数を計算するようにタスクT340を構成することが望ましいことができる。

【0094】

タスクT400は、第1のチャンネルの振幅に関して第2のチャンネルの振幅を制御するためにタスクT300によって作り出された更新済み利得係数値を使用するように構成され得る。本明細書で説明されているように、音響的に均衡化されたセグメントの観測されたレベル不均衡に基づいて利得係数値を更新するようにタスクT300を構成することが望ましいことができる。音響的に均衡化されていない次のセグメントに関して、タスクT300が利得係数値を更新することを差し控えること、およびタスクT400がごく最近更新された利得係数値を適用し続けることが望ましいことができる。図12Aは、タスクT400のこのような実現形態T420を含む方法M100の実現形態M120の流れ図を示す。タスクT420は、マルチチャンネル信号の1シリーズの連続するセグメントの各々（例えば、1シリーズの音響的に不均衡化されたセグメントの各々）において第1のチャンネルの振幅に関して第2のチャンネルの振幅を修正するために更新済み利得係数値を使用するように構成されている。このような1シリーズは、タスクT300が利得係数値を再び更新するようにもう1つの音響的に均衡化されたセグメントが識別されるまで続き得る。

（このパラグラフで説明された原理は本明細書で説明されたようにサブバンド利得係数値の更新および利用にも適用され得る。）

方法M100の実現形態は、較正依存性であり得る空間選択性処理動作といったマルチチャンネル信号および／または処理済みマルチチャンネル信号に対する種々の更なる動作（例えば、オーディオセンシング・デバイスと特定の音源との間の距離を決定し、雑音を減らし、特定の方向から到着する信号成分を増強し、および／または1つ以上のサウンド成分を他の環境サウンドから分離する1つ以上の動作）をサポートするようにも構成され得る。例えば、均衡化されたマルチチャンネル信号（例えば、処理済みマルチチャンネル信号）の用途の範囲は、非定常拡散および／または指向性雑音の低減；近距離音場の所望スピーカーによって作り出されるサウンドの残響除去；マイクロホンチャンネル間で無相関である雑音（例えば、風および／またはセンサー雑音）の除去；望ましくない方向からのサウンドの抑制；任意の方向からの遠距離音場信号の抑制；直接経路対残響（direct-path-to-reverberation）信号強度の推定（例えば、遠距離音場音源からの干渉の大幅な低減）；近距離および遠距離音場音源間の識別を介した非定常雑音の低減；および典型的には利得ベースのアプローチでは達成できない休止中だけでなく近距離音場所望音源活動中の正面干渉体からのサウンドの低減；を含む。

【0095】

図12Bは、処理されたマルチチャンネル信号にボイス活動検出（VAD）動作を実行するタスクT500を含む方法M100の実現形態M130の流れ図を示す。図13Aは、処理されたマルチチャンネル信号からの情報に基づいて雑音推定値を更新し、そしてボイス活動検出動作を含み得るタスクT600を含む方法M100の実現形態M140の流れ図を示す。

【0096】

近距離音場および遠距離音場音源からのサウンド間を弁別する信号処理方式を実現する（例えば、より良好な雑音低減のために）ことが望ましいことができる。このような方式の1つの振幅ベースまたは利得ベースの例は、音源が近距離音場または遠距離音場であるかどうかを決定するために2つのマイクロホン間の圧力傾斜音場（pressure gradient field）を使用する。このような技法は近距離音場無音時に遠距離音場音源からの雑音を減らすために有用であり得るが、両音源が活動しているときには近距離音場信号と遠距離音場信号との間の弁別をサポートしない可能性がある。

【0097】

特定の角度範囲内に一貫性のあるピックアップを用意することが望ましいことができる。例えば、特定の範囲（例えば、マイクロホンアレイの軸に関して60度の範囲）内のす

10

20

30

40

50

すべての近距離音場信号を受け入れて、他のすべて（例えば、70度以上の角度における音源からの信号）を減衰させることが望ましいことができる。ビーム形成およびBSSによれば、角度的減衰は典型的にはこのような範囲に亘って一貫性のあるピックアップを防止する。このような方法はまた、後処理動作が再収束する前で、デバイスの方位変化（例えば、回転）後にボイス拒絶という結果を招く可能性もある。本明細書で説明されたような方法M100の実現形態は、所望スピーカーに対する方向がなお許容可能な方向の範囲内にあり、それによって収束遅延に起因するボイス変動および／または期限切れ雑音基準に起因するボイス減衰を防止する限り、デバイスの急激な回転に対して強固である雑音低減方法を取得するために使用され得る。

【0098】

10

均衡化されたマルチチャネル信号からの利得差と位相ベースの方向情報とを組み合わせることによって、信号の存在が監視され得るマイクロホンアレイの周りにおいて調整可能な空間領域が選択され得る。異なるサブタスクのための狭いまたは広いピックアップ領域を定義するために利得ベース範囲および／または方向範囲が設定され得る。例えば、所望のボイス活動を検出するためにはより狭い範囲が設定され得るが、雑音低減といった目的のためには選択された領域上のより広い範囲が使用され得る。位相相関および利得差評価の精度はSNRの低下と共に低下する傾向があり、誤警報率を制御するためにそれに応じて閾値および／または決定を調整することは望ましい可能性がある。

【0099】

20

処理されたマルチチャネル信号がボイス活動検出（VAD）動作をサポートするためだけに使用されている用途に関して、効果的で正確な雑音低減動作が短縮された雑音低減収束時間で、より敏速に実行され得るように、利得較正が低下した精度レベルで動作することは受入れ可能であり得る。

【0100】

音源とマイクロホンペアとの間の相対的距離が増加するにつれて、（例えば、残響の増加に起因して）異なる周波数成分の到着方向間のコヒーレンスは減少すると予期され得る。従って、タスクT360において計算されたコヒーレンシー測定値はまた、ある程度、近接度測定値としても役立つ。到着方向だけに基づく処理動作とは異なり、例えば本明細書で説明されたようなコヒーレンシー測定値に基づく時間依存性および／または周波数依存性振幅制御は、ユーザのスピーチまたは他の所望近距離音場音源を同じ方向の遠距離音場音源からの、競合するスピーカーのスピーチといった干渉音から区別するために有効であり得る。方向的コヒーレンシーが距離と共に減少する速度は環境によって変化し得る。例えば、自動車の内部は典型的には極めて残響が大きいので、広い範囲の周波数に亘る方向的コヒーレンシーは、音源からほんの約50センチメートルの範囲内だけで経過時間に亘って信頼できる安定なレベルに維持され得る。このような場合、スピーカーが指向性マスキング関数のパスバンド内に位置する場合でも、後部座席の乗客からのサウンドはコヒーレントでないとして拒絶される可能性がある。検出可能なコヒーレンスの範囲は（例えば、すぐ近くの天井からの反響によって）丈の高いスピーカーに関するこのような状況においても低減されることもあり得る。

30

【0101】

40

処理されたマルチチャネル信号は、BSS、到着遅延または他の指向性SSPといった他の空間選択性処理（SSP）、または近接検出といった距離SSPをサポートするために使用され得る。近接検出はチャネル間の利得差に基づき得る。（例えば、限定された周波数範囲に亘る、および／または多数のピッチ周波数における、コヒーレンスの測定値として）タイムドメインまたは周波数ドメインにおける利得差を計算することが望ましいことができる。

【0102】

携帯型オーディオセンシング・デバイスのためのマルチマイクロホン雑音低減方式は、ビーム形成アプローチとブラインド音源分離（BSS）アプローチとを含む。このようなアプローチは典型的には、所望の音源（例えば、近距離音場スピーカーのボイス）と同じ

50

方向から到着する雑音を抑制する能力のないことに悩まされる。特に、ヘッドホンおよび中音域音場または遠距離音場のハンドヘルドアプリケーション（例えば、送受話器またはスマートフォンのブラウジングトーク（browse-talk）およびスピーカーフォン・モード）において、マイクロホンアレイによって記録されたマルチチャネル信号は、干渉雑音源からのサウンドおよび／または所望近距離音場の話し手のスピーチの大きな残響を含み得る。特にヘッドホンに関して、ユーザの口までの大きな距離は、方向情報だけを使用して大きく抑制することが困難であり得る正面方向からの多量の雑音をマイクロホンアレイがピックアップすることを可能にする恐れがある。

【0103】

典型的なBSSまたは汎用サイドローブ・キャンセル（GSC）タイプの技法は、最初に所望のボイスを1つのマイクロホンチャンネルに分離し、それからこの分離されたボイスに後処理動作を実行することによって雑音低減を実行する。この手順は音響シナリオ変更の場合に長い収束時間を引き起こす可能性がある。例えば、ブラインド音源分離、GSC、または類似の適応型学習ルールに基づく雑音低減方式は、デバイスユーザ保持パターン（例えば、デバイスとユーザの口との間の方位）の変化時および／または音量の急激な変化時の長い収束時間を示す、および／または環境雑音（例えば、通過車両、公衆アドレス告知（public address announcement））のスペクトルシグネチャー（spectral signature）を示すことができる。残響の大きい環境（例えば、車両内部）では適応型学習方式はトラブル収束を有することができる。収束するためのこのような方式の不成功は、この方式に所望の信号成分を拒絶させることができる。ボイス通信の用途において、このような拒絶はボイス歪みを増大させることができる。

【0104】

デバイスユーザ保持パターンの変化に対するこのような方式の強固さを向上させるためには、および／または収束時間を高速化するためには、より迅速な初期雑音低減応答を供与するためにデバイスの周りの空間ピックアップ領域を限定することが望ましいことができる。このような方法は、ある角度方向に対する（例えば、マイクロホンアレイの軸といったデバイスの基準方向に関する）弁別によって限定された空間ピックアップ領域を定義するためにマイクロホン間の、および／または近距離および遠距離音場音源からの信号成分間の、位相および利得関係を活用するように構成され得る。常に基準線初期雑音低減を示す所望スピーカー方向におけるオーディオデバイスの周りの選択領域を有することによって、環境雑音の急激な変化ばかりでなくオーディオデバイスに関する所望ユーザの空間的变化に対する高度の強固さが達成され得る。

【0105】

均衡化されたチャンネル間の利得差は、より良好な正面雑音抑制（例えば、ユーザの前の干渉スピーカーの抑制）といった、より積極的な近距離音場／遠距離音場弁別をサポートし得る近接検出のために使用され得る。マイクロホン間の距離に依存して、均衡化されたマイクロホンチャンネル間の利得差は、典型的には音源が50センチメートルまたは1メートル以内にある場合にだけ発生するであろう。

【0106】

図13Bは、方法M100の実現形態M150の流れ図を示す。方法M150は、処理されたマルチチャネル信号に近接検出動作を実行するタスクT700を含む。例えば、タスクT700は、処理されたマルチチャネル信号のチャンネルのレベル間の差がある閾値より大きいとき（代替として、（A）較正されていないチャンネルのレベル差と（B）タスクT300の利得係数値との合計がこの閾値より大きいとき）セグメントが所望の音源からであることを検出するように（例えば、ボイス活動の検出を示すように）構成され得る。この閾値は発見的に決定され得るが、また信号対雑音比（SNR）、雑音レベルなどといった1つ以上の因子に依存して異なる閾値を使用することは（例えば、SNRが低いときに高い閾値を使用することは）望ましい可能性がある。図14Aは閾値が高くなるにつれて小さくなる領域を有する、3つの異なる閾値に対応する近接検出領域の境界の例を示す。

【0107】

スピーカーカバレッジ (speaker coverage) のコーン (cone) を取得するために、またこのゾーンの外側の音源からの非定常雑音を減衰させるために、許容された方向の範囲 (例えば、プラスマイナス45度) を近距離音場／遠距離音場近接バブルと組み合わせることが望ましいことができる。このような方法は、音源が許容可能な方向の範囲内にあるときでも遠距離音場音源からのサウンドを減衰させるために使用され得る。例えば、近距離音場／遠距離音場弁別器の積極的調整をサポートするために良好なマイクロホン較正を用意することが望ましいことができる。図14Bは、スピーカーカバレッジのこのようなコーンを取得するために許容可能な方向の範囲 (図10に示されているような) と近接バブル (図14Aに示されているような) との交差 (太線で示されている) の一例を示す。このような場合、タスクT100において計算された複数の位相差は、所望の範囲内の音源から発生したセグメントを識別するために (例えば、タスクT312、T322、およびT332を参照しながら上記に論じられたような) マスキング関数および／または (例えば、タスクT360を参照しながら上記に論じられたような) コヒーレンシー測定値を使用して許容可能な方向の範囲を強化するために使用され得る。このようなマスキング関数の方向およびプロファイルは、所望の用途に従って選択され得る (例えば、ボイス活動検出のためにはより急峻なプロファイル、または雑音成分の減衰のためにはより平滑なプロファイル)。

10

【0108】

上記のように、図2は、ユーザの口に関して標準的な方位にあるユーザの耳に装着されたヘッドホンの上面図を示す。図15および図16は、この用途に適用される図14Bに示されているような音源選択領域境界の上面図および側面図を示す。

20

【0109】

ボイス活動検出 (VAD) のために近接検出動作 (例えば、タスクT700) の結果を使用することは望ましい可能性がある。1つのこのような例では、(例えば、雑音周波数成分および／またはセグメントを減衰させるために) チャンネルの1つ以上における利得制御として非2成分改善VAD測定値が適用される。図17Aは、均衡化されたマルチチャンネル信号にこのような利得制御動作を実行するタスクT800を含む方法M100の実現形態M160の流れ図を示す。もう1つのこのような例では、(例えば、VAD動作によって雑音として分類された周波数成分またはセグメントを使用して) 雑音低減動作のために雑音推定値を計算する (例えば、更新する) ために2成分改善VADが適用される。図17Bは、近接検出動作の結果に基づいて雑音推定値を計算する (例えば、更新する) タスクT810を含む方法M100の実現形態M170の流れ図を示す。図18は、方法M170の実現形態M180の流れ図を示す。方法M180は、更新された雑音推定値に基づくマルチチャンネル信号の少なくとも1つのチャンネルに雑音低減動作 (例えば、スペクトル減算またはWienerフィルタリング動作) を実行するタスクT820を含む。

30

【0110】

(例えば、図14Bおよび／または図15および図16に示されているようなバブルを定義する) 近接検出動作および方向的コヒーレンス検出動作からの結果は、改善されたマルチチャンネルボイス活動検出 (VAD) 動作を取得するために組み合わされ得る。この組み合わせられたVAD動作は、非ボイスフレームの敏速な拒絶のために、および／または1次マイクロホンチャンネル上で動作する雑音低減方式を構築するために、使用され得る。このような方法は、較正とVADのための方向および近接情報を組み合わせることとVAD動作の結果に基づいて雑音低減動作を実行することとを含み得る。例えば、近接検出タスクT700の代わりに、方法M160、M170、またはM180においてこのような組み合わせられたVAD動作を使用することが望ましいことができる。

40

【0111】

典型的な環境における音響雑音は、片言雑音、空港雑音、街頭雑音、競合する話し手のボイス、および／または干渉音源 (例えば、TVセットまたはラジオ) からのサウンドを含み得る。その結果、このような雑音は典型的には非定常的であって、ユーザ自身のボイ

50

スの平均スペクトルに近い平均スペクトルを持つことができる。単一マイクロホン信号から計算されるような雑音パワー（エネルギー）基準信号は通常、単なる近似的定常雑音推定値である。更に、このような計算は一般に、雑音パワー推定遅延を伴うので、サブバンド利得の対応する調整はかなりの遅延後にだけ実行されることができる。環境雑音の信頼できる同時的な推定値を取得することが望ましいことができる。

【0112】

雑音推定値の例は、単一チャンネルVADに基づく単一チャンネル長期推定値とマルチチャンネルBSSフィルタによって作り出されるような雑音基準とを含む。タスクT810は、1次マイクロホンチャンネルの成分および／またはセグメントを分類するために近接検出動作からの（2重チャンネル）情報を使用することによって単一チャンネル雑音基準を計算するように構成され得る。このような雑音推定は、長期推定を必要としないので他のアプローチより遥かに敏速に利用可能になり得る。この単一チャンネル雑音基準は、典型的には非定常雑音の除去をサポートできない長期推定ベースのアプローチとは異なり、非定常雑音も捕捉できる。このような方法は速くて正確な非定常雑音基準を提供し得る。例えば、このような方法は、図14Bに示されているような前方コーン内に存在しないいかなるフレームに関する雑音基準も更新するように構成され得る。雑音基準は（例えば、場合によっては各周波数成分上にある1次スムーザー（first-degree smoother）を使用して）平滑化され得る。近接検出の使用は、このような方法を使用するデバイスが指向性マスキング関数の前方ローブ内に進入する車両の雑音の音響といったすぐ近くの遷移を拒絶することを可能にし得る。

【0113】

収束するマルチチャンネルBSS方式を待つよりむしろ直接1次チャンネルから雑音基準を取るようにタスクT810を構成することが望ましいことができる。このような雑音基準は、組み合わせられた位相・利得VADを使用するか、単に位相VADを使用して構成され得る。このようなアプローチはまた、スピーカーと電話との間の新しい空間構成に収束しながら、または送受話器が次善の空間構成で使用されているときに、ボイスを減衰させるBSS方式の問題を回避する助けにもなり得る。

【0114】

上記のようなVAD指示は雑音基準信号の計算をサポートするために使用され得る。例えば、あるフレームが雑音であることをVAD指示が示しているとき、このフレームは雑音基準信号（例えば、1次マイクロホンチャンネルの雑音成分のスペクトルプロファイル）を更新するために使用され得る。このような更新は、例えば周波数成分値を時間的に平滑化することによって（例えば、各成分の前の値を現在雑音推定値の対応する成分の値で更新することによって）周波数ドメインにおいて実行され得る。1つの例では、Wienerフィルタは1次マイクロホンチャンネルに雑音低減動作を実行するために雑音基準信号を使用する。もう1つの例では、スペクトル減算動作は（例えば、1次マイクロホンチャンネルから雑音スペクトルを減算することによって）1次マイクロホンチャンネルに雑音低減動作を実行するために雑音基準信号を使用する。あるフレームが雑音でないことをVAD指示が示すときには、このフレームは1次マイクロホンチャンネルの信号成分のスペクトルプロファイルを更新するために使用され得るが、このプロファイルも雑音低減動作を実行するためにWienerフィルタによって使用され得る。結果的に行われた動作は、2重チャンネルVAD動作を使用する擬似単一チャンネル（quasi-single-channel）雑音低減アルゴリズムであると考えられ得る。

【0115】

チャンネル較正が必要とされない（例えば、マイクロホンチャンネルが既に均衡化されている）状況においても本明細書で説明されたような近接検出動作が適用され得ることは明らかに注目される。図19Aは、本明細書で説明されたようなタスクT100の事例およびT360と、本明細書で説明されたようなコヒーレンシー測定および近接決定（例えば、図14Bに示されているようなバブル）に基づくVAD動作T900と、を含む全体の構成による方法M300の流れ図を示す。図19Bは、（例えば、タスクT810を参照し

ながら説明されたような) 雑音推定値計算タスクT 9 1 0を含む方法M 3 0 0の実現形態M 3 1 0の流れ図を示し、また図2 0 Aは、(例えば、タスクT 8 2 0を参照しながら説明されたような) 雑音低減タスクT 9 2 0を含む方法M 3 1 0の実現形態M 3 2 0の流れ図を示す。

【0 1 1 6】

図2 0 Bは全体的構成による装置G 1 0 0のブロック図を示す。装置G 1 0 0は(例えば、タスクT 1 0 0を参照しながら本明細書で説明されたような) 複数の位相差を取得するための手段F 1 0 0を含む。装置G 1 0 0はまた、(例えば、タスクT 2 0 0を参照しながら本明細書で説明されたような) マルチチャネル信号の第1および第2のチャンネルのレベルを計算するための手段F 2 0 0を含む。装置G 1 0 0はまた、(例えば、タスクT 3 0 0を参照しながら本明細書で説明されたような) 利得係数値を更新するための手段F 3 0 0を含む。装置G 1 0 0はまた、(例えば、タスクT 4 0 0を参照しながら本明細書で説明されたような) 更新された利得係数値に基づいて第1のチャンネルに関して第2のチャンネルの振幅を修正するための手段F 4 0 0を含む。

10

【0 1 1 7】

図2 1 Aは、全体的構成による装置A 1 0 0のブロック図を示す。装置A 1 0 0は(例えば、タスクT 1 0 0を参照しながら本明細書で説明されたような) マルチチャネル信号のチャンネルS 1 0 - 1およびS 1 0 - 2から複数の位相差を取得するように構成された位相差計算器1 0 0を含む。装置A 1 0 0はまた、(例えば、タスクT 2 0 0を参照しながら本明細書で説明されたような) マルチチャネル信号の第1および第2のチャンネルのレベルを計算するように構成されたレベル計算器2 0 0を含む。装置A 1 0 0はまた、例えば、タスクT 3 0 0を参照しながら本明細書で説明されたような) 利得係数値を更新するように構成された利得係数計算器3 0 0を含む。(装置A 1 0 0はまた、(例えば、タスクT 4 0 0を参照しながら本明細書で説明されたような) 更新された利得係数値に基づいて第1のチャンネルに関して第2のチャンネルの振幅を修正することによって処理済みマルチチャネル信号を作り出すように構成された利得制御要素4 0 0を含む。

20

【0 1 1 8】

図2 1 Bは、装置A 1 0 0と；周波数ドメインにおける信号S 1 0 - 1およびS 1 0 - 2をそれぞれ作り出すように構成されたFFTモジュールTM 1 0 aおよびTM 1 0 bと；処理済みマルチチャネル信号に(例えば、本明細書で説明されたような) 空間選択性処理動作を実行するように構成された空間選択性処理モジュールSS 1 0 0と；を含む装置A 1 1 0のブロック図を示す。図2 2は装置A 1 0 0とFFTモジュールTM 1 0 aおよびTM 1 0 bとを含む装置A 1 2 0のブロック図を示す。装置A 1 2 0はまた、(例えば、タスクT 7 0 0を参照しながら本明細書で説明されたような) 処理済みマルチチャネル信号に近接検出動作(例えば、ボイス活動検出動作)を実行するように構成された近接検出モジュール7 0 0(例えば、ボイス活動検出器)と；(例えば、タスクT 8 1 0を参照しながら本明細書で説明されたような) 雑音推定値を更新するように構成された雑音基準計算器8 1 0と；(例えば、タスクT 8 2 0を参照しながら本明細書で説明されたような) 処理済みマルチチャネル信号の少なくとも1つのチャンネルに雑音低減動作を実行するように構成された雑音低減モジュール8 2 0と；雑音低減された信号をタイムドメインに変換するように構成された逆FFTモジュールIM 1 0と；を含む。近接検出モジュール7 0 0に加えて、または代替として装置A 1 1 0は、処理済みマルチチャネル信号の方向性処理(例えば、図1 4 Bに示されているような前方ローブに基づくボイス活動検出)のためのモジュールを含み得る。

30

40

【0 1 1 9】

ある幾つかのマルチチャネル信号処理動作はマルチチャネル出力の各チャンネルを作り出すためにマルチチャネルの1つより多いチャンネルからの情報を使用する。このような動作の例は、ビーム形成動作とブラインド音源分離(BSS)動作とを含み得る。エコーキャンセル動作は各出力チャンネルにおける残留エコーを変える傾向があるので、エコーキャンセルをこのような技法に統合することは困難であり得る。本明細書で説明されているよう

50

に、方法M100は、マルチチャネル信号の1つ以上のチャネルの各々に（例えば、1次チャネルに）単一チャネル時間および／または周波数依存振幅制御（例えば、雑音低減動作）を実行するために計算された位相差からの情報を使用するように実現され得る。このような単一チャネル動作は、残留エコーが実質的に変わらないままに留まるように実現され得る。その結果、このような雑音低減動作を含む方法M100の一実現形態とのエコーキャンセル動作の統合は、2つ以上のマイクロホンチャネル上で動作する雑音低減動作とのエコーキャンセル動作の統合より容易であり得る。

【0120】

残留背景雑音を白化する(whiten)ことは望ましい可能性がある。例えば、雑音だけの間隔を識別して、このような間隔中の信号スペクトルを雑音スペクトルプロファイル（例えば、擬似ホワイトまたはピンク・スペクトルプロファイル）に圧縮・伸張または低減するためにVAD動作（例えば、本明細書で説明されたような方向および／または近接度ベースのVAD動作）を使用することが望ましいことができる。このような雑音白化は残留定常雑音レベルの感知を創造することができる、および／または背景内に入れられる、または引っ込む雑音の認知を引き起こすことができる。白化が適用されない間隔（例えば、スピーチ間隔）と白化が適用される間隔（例えば、雑音間隔）との間の遷移を取り扱うために時間的平滑化方式といった平滑化方式を含むことが望ましいことができる。このような平滑化は間隔間の平滑な遷移をサポートする助けとなり得る。

【0121】

マイクロホン（例えば、MC10およびMC20）がサウンド以外の放射線または放出体(emission)に敏感なトランスデューサとして、より一般的に実現され得ることは明らかに注目される。1つのこのような例では、マイクロホンペアは1対の超音波トランスデューサ（例えば、15、20、25、30、40、または50kHz以上より高い音響周波数に敏感なトランスデューサ）として実現される。

【0122】

（例えば、図14Bに示されているような前方ローブを識別する）方向的信号処理アプリケーションに関して、スピーチ信号（または、他の所望信号）が方向的にコヒーレントであることが予期され得る特定の周波数成分または周波数範囲を目標にすることが望ましいことができる。指向性雑音（例えば、自動車といった音源からの）および／または拡散雑音といった背景雑音が同じ範囲に亘って方向的にコヒーレントでないことは予期され得る。音声は4から8kHzの範囲内で低いパワーを持つ傾向があるので、4kHzより高い周波数に関連して方向的コヒーレンスを決定することが望ましいことができる。例えば、約700Hzから約2kHzの範囲に亘って方向的コヒーレンスを決定することが望ましいことができる。

【0123】

上記のように、限定された周波数範囲に亘る周波数成分の位相差に基づいてコヒーレンシー測定値を計算するようにタスクT360を構成することが望ましいことができる。更に、または代替として、多数のピッチ周波数における周波数成分に基づいてコヒーレンシー測定値を計算するために、タスクT360および／または（特に、図14Bに示されているような前方ローブを定義するといったスピーチアプリケーションのための）別の方向的処理タスクを構成することが望ましいことができる。

【0124】

発声されたスピーチ（例えば、母音）のエネルギースペクトルは、ピッチ周波数の高調波においてローカルピークを有する傾向がある。他方、背景雑音のエネルギースペクトルは相対的に構造化されない傾向がある。その結果、ピッチ周波数の高調波における入力チャネルの成分は、他の成分より高い信号対雑音比(SNR)を有すると予期され得る。方法M100のスピーチ処理アプリケーション（例えば、ボイス活動検出アプリケーション）のための方向的処理タスクに関して、推定されたピッチ周波数の多数に対応する位相差だけを考慮するようにタスクを構成すること（例えば、前方ローブ識別タスクを構成すること）が望ましいことができる。

10

20

30

40

50

【0125】

典型的ピッチ周波数は、男性話者に関しては約70から100Hz、女性話者に関しては約150から200Hzの範囲にある。現在ピッチ周波数は、ピッチ周期を隣接ピッチピーク（例えば、1次マイクロホンチャンネルにおける）間の距離として計算することによって推定され得る。入力チャンネルのサンプルは、このエネルギーの測定値に基づいて（例えば、サンプルエネルギーとフレーム平均エネルギーとの間の比に基づいて）および／またはこのサンプルの近隣が既知のピッチピークの同様な近隣とどれほどよい相関性があるかの測定値に基づいてピッチピークとして識別され得る。ピッチ推定手順は、例えばwww-dot-3gpp-dot-orgにおいてオンラインで利用可能なE V R C（Enhanced Variable Rate Code（改善可変速度コード））文書C S 0 0 1 4－Cのセクション4. 6. 3（4－4 4から4－4 9ページ）に説明されている。（例えば、ピッチ周期または「ピッチラグ」の推定値の形をした）ピッチ周波数の現在推定値は典型的には既に、スピーチ符号化および／または復号を含むアプリケーション（例えば、符号励起線形予測（C E L P）および原型波形補間（P W I）といったピッチ推定を含むコーデックを使用するボイス通信）において利用可能であろう。

10

【0126】

ピッチ周波数の多数に対応するこれらの位相差だけを考慮することによって、考慮されるべき位相差の数は大幅に削減され得る。更に、これらの選択された位相差が計算される周波数係数は考慮される周波数範囲内の他の周波数係数に関して高いS N Rを有することが予期され得る。より一般的な場合では、他の信号特性も考慮され得る。例えば、計算された位相差の少なくとも25、50、または75パーセントが推定ピッチ周波数の多数に対応するように、方向的処理タスクを構成することが望ましいことができる。同じ原理は他の所望高調波信号にも適用され得る。

20

【0127】

上記のように、音響信号を受信するように構成された2つ以上のマイクロホンのアレイR 1 0 0を有する携帯型オーディオセンシング・デバイスを製造することが望ましいことができる。このようなアレイを含むように実現され得る、そしてオーディオレコーディングおよび／またはボイス通信アプリケーションのために使用され得る携帯型オーディオセンシング・デバイスの例は、電話送受話器（例えば、携帯電話送受話器）；有線または無線ヘッドホン（例えば、B l u e t o o t hヘッドホン）；ハンドヘルド・オーディオおよび／またはビデオレコーダー；オーディオおよび／またはビデオコンテンツを記録するように構成されたパーソナル・メディア・プレーヤー；パーソナル・ディジタル・アシスタント（P D A）または他のハンドヘルド・コンピューティングデバイス；およびノートブックコンピュータ、ラップトップコンピュータ、ネットブックコンピュータ、または他の携帯型コンピューティングデバイス；を含む。

30

【0128】

アレイR 1 0 0の各マイクロホンは全方向性、双方向性または一方向性（例えば、心臓形）である応答を持ち得る。アレイR 1 0 0において使用され得る種々のタイプのマイクロホンは圧電性マイクロホン、動的マイクロホンおよびエレクトレットマイクロホンを含む（限定なしに）。送受話器またはヘッドホンといった携帯型音声通信のためのデバイスでは、アレイR 1 0 0の隣接マイクロホン間の中心間の間隔は典型的には、約1. 5 cmから約4. 5 cmの範囲にあるが、送受話器といったデバイスでは、より大きな間隔（例えば、最大10または15 cm）も可能である。補聴器では、アレイR 1 0 0のマイクロホン間の中心間の間隔は約4または5 mmほどに小さい可能性がある。アレイR 1 0 0のマイクロホンは、1直線に沿って、または代替としてこれらの中心が2次元形状（例えば、3角形）または3次元形状の頂点にあるように、配置され得る。

40

【0129】

マルチマイクロホン・オーディオセンシング・デバイス（例えば、本明細書で説明されるようなデバイスD 1 0 0、D 2 0 0、D 3 0 0、D 4 0 0、D 5 0 0、またはD 6 0 0）の動作時に、アレイR 1 0 0は各チャンネルがマイクロホンのうちの音響環境に対応する

50

1つのマイクロホンの応答に基づくマルチチャネル信号を作り出す。1つのマイクロホンはもう1つのマイクロホンより直接的に特定の音を受信し得るので、単一のマイクロホンを使用して捕捉され得るより完全な音響環境の表現をまとめて与えるために、対応するチャネルは互いに異なる。

【0130】

マルチチャネル信号S10を作り出すために、これらのマイクロホンによって作り出された信号に1つ以上の処理動作をアレイR100が実行することが望ましいことができる。図23Aは、インピーダンス整合、アナログ・デジタル変換、利得制御、および／またはアナログおよび／またはデジタルドメインにおけるフィルタリングを含み得る（限定なしに）1つ以上のこのような動作を実行するように構成されたオーディオ前処理ステージAP10を含むアレイR100の実現形態R200のブロック図を示す。

10

【0131】

図23Bは、アレイR200の実現形態R210のブロック図を示す。アレイR210はアナログ前処理ステージP10aおよびP10bを含むオーディオ前処理ステージAP10の実現形態AP20を含む。1つの例では、ステージP10aおよび10bは各々、対応するマイクロホン信号にハイパスフィルタリング動作を（例えば、50、100または200Hzの遮断周波数で）実行するように構成されている。

【0132】

アレイR100がマルチチャネル信号をデジタル信号として、すなわち一連のサンプルとして作り出すことが望ましいことができる。アレイR210は例えば、各々が対応するアナログチャネルをサンプリングするように整えられたアナログ・デジタル変換器（ADC）C10aおよびC10bを含む。音響アプリケーションのための典型的なサンプリングレートは8kHz、12kHz、16kHzの周波数、および約8から約16kHzの範囲内の他の周波数を含むが、約44kHzほどの高いサンプリングレートも使用され得る。この特定の例ではアレイR210はまた各々が、対応するデジタル化されたチャネルに1つ以上の前処理動作（例えば、エコーキャンセル、雑音低減および／またはスペクトル形成）を実行するように構成されたデジタル前処理ステージP20aおよびP20bを含む。

20

【0133】

アレイR100のマイクロホンがサウンド以外の放射線または放出体に敏感なトランスデューサとして、より一般的に実現され得ることは明らかに注目される。1つのこのような例では、アレイR100のマイクロホンは超音波トランスデューサ（例えば、15、20、25、30、40、または50kHz以上より高い音響周波数に敏感なトランスデューサ）として実現される。

30

【0134】

図24Aは、全体的構成によるデバイスD10のブロック図を示す。デバイスD10は本明細書で開示されたマイクロホンアレイR100のいずれかの実現形態の事例を含み、また本明細書で開示されたオーディオセンシング・デバイスのいずれもデバイスD10の一事例として実現され得る。デバイスD10はまた、コヒーレンシー測定 of 値を計算するためにアレイR100によって作り出されたマルチチャネル信号を処理するように構成された装置A10の実現形態の一事例を含む。例えば、装置A10は本明細書で開示された方法M100の実現形態のうちのいずれかの実現形態の一事例に従ってマルチチャネルオーディオ信号を処理するように構成され得る。装置A10はハードウェアにおいて、および／またはソフトウェア（例えば、ファームウェア）において実現され得る。例えば、装置A10は、処理されたマルチチャネル信号に上記の空間処理動作（例えば、オーディオセンシング・デバイスと特定の音源との間の距離を決定する、雑音を減らす、特定の方向から到着する信号成分を増強する、および／または他の環境サウンドから1つ以上のサウンド成分を分離する1つ以上の動作）を実行するようににも構成されたデバイスD10のプロセッサ上に実現され得る。上記のような装置A10は装置A10の一事例として実現され得る。

40

50

【0135】

図24Bは、デバイスD10の一実現形態である通信デバイスD20のブロック図を示す。デバイスD20は装置A10を含むチップまたはチップセットCS10（例えば、移動局モデム（MSM）チップセット）を含む。チップ／チップセットCS10は装置A10の全部または一部を実行する（例えば、命令として）ように構成され得る1つ以上のプロセッサを含み得る。チップ／チップセットCS10はまたアレイR100の処理要素（例えば、オーディオ前処理ステージAP10の要素）も含み得る。チップ／チップセットCS10は、無線周波数（RF）通信信号を受信してRF信号内の符号化されたオーディオ信号を復号して再生するように構成された受信機と、装置A10によって作り出された処理済み信号に基づくオーディオ信号を符号化してこの符号化されたオーディオ信号を記述するRF通信信号を送信するように構成された送信機と、を含む。例えば、チップ／チップセットCS10の1つ以上のプロセッサは、符号化されたオーディオ信号が雑音低減された信号に基づくようにマルチチャネル信号の1つ以上のチャンネルに上記のような雑音低減動作を実行するように構成され得る。

10

【0136】

デバイスD20は、アンテナC30を介してRF通信信号を受信および送信するように構成されている。デバイスD20はまた、アンテナC30までの経路にダイプレクサー（diplexer）と1つ以上のパワーアンプとを含み得る。チップ／チップセットCS10はまた、キーパッドC10を介してユーザ入力を受信し、ディスプレイC20を介して情報を表示するようにも構成されている。この例では、デバイスD20はまた、全世界測位システム（GPS）位置探索サービスを、および／または無線（例えば、Bluetooth（商標））ヘッドホンといった外部デバイスとの短距離通信を、サポートするための1つ以上のアンテナC40も含む。もう1つの例では、このような通信デバイスはそれ自身がBluetoothヘッドホンであって、キーパッドC10、ディスプレイC20およびアンテナC30を欠いている。

20

【0137】

本明細書で説明されているような装置A10の実現形態はヘッドホンおよび送受話器を含む種々のオーディオセンシング・デバイスに具体化され得る。送受話器実現形態の一例は、マイクロホン間に6.5センチメートル間隔を有するアレイR100の前向き2重マイクロホン実現形態を含む。2重マイクロホン・マスキングアプローチの実現形態は、直接的にスペクトログラムにおけるマイクロホンペアの位相関係を分析することと、望ましくない方向からの時間・周波数ポイントをマスキングすることと、を含み得る。

30

【0138】

図25A～図25Dは、デバイスD10のマルチマイクロホン携帯型オーディオセンシング実現形態D100の種々の図を示す。デバイスD100は、アレイR100の2マイクロホン実現形態を保持するハウジングZ10とこのハウジングから延びるイヤホンZ20とを含む無線ヘッドホンである。このようなデバイスは、（例えば、Bluetooth特別関心グループ社、ベレビュー、ワシントン州（Bluetooth Special Interest Group, Inc., Bellevue, WA）によって公表されているBluetooth（商標）プロトコルの1バージョンを使用して）携帯電話送受話器といった電話機デバイスとの通信を介して半二重または全二重電話方式をサポートするように構成され得る。一般に、ヘッドホンのハウジングは、矩形であるか、またはそうでなければ図25A、図25B、および図25Dに示されているように細長くされる（例えば、ミニブームのように形作られる）可能性があり、あるいは、より丸く、円形にさえされ得る。ハウジングはまた、バッテリーとプロセッサおよび／または他の処理回路（例えば、プリント回路基板とこれに搭載された部品）とを収容することもあり、また電気ポート（例えば、ミニ・ユニバーサル・シリアル・バス（USB）またはバッテリー充電のための他のポート）と1つ以上のボタンスイッチといったユーザインタフェース機構および／またはLEDとを含み得る。典型的には、ハウジングの長軸に沿った長さは1インチから3インチの範囲内にある。

40

【0139】

50

典型的には、アレイ R 1 0 0 の各マイクロホンはデバイス内、音響ポートとして役立つハウジングの1つ以上の小さな孔の後方に装着される。図 2 5 B～図 2 5 Dは、デバイス D 1 0 0 のアレイの1次マイクロホンのための音響ポート Z 4 0 とデバイス D 1 0 0 のアレイの2次マイクロホンのための音響ポート Z 5 0 との位置を示す。

【0 1 4 0】

ヘッドホンはまた、典型的にはヘッドホンから取り外し可能であるイヤフック Z 3 0 といった固定デバイスも含み得る。外部のイヤフックは、例えばユーザがどちらの耳でも使用できるようにヘッドホンを構成することを可能にするために両側使用可能 (reversible) であり得る。代替としてヘッドホンのイヤホンは、異なるユーザが特定のユーザの外耳道の外側部分へのより良好な嵌合のために異なるサイズ (例えば、直径) の受話口を使用

10

【0 1 4 1】

図 2 6 A～図 2 6 Dは、無線ヘッドホンのもう1つの例であるデバイス D 1 0 のマルチマイクロホン携帯型オーディオセンシング実現形態 D 2 0 0 の種々の図を示す。デバイス D 2 0 0 は、丸い楕円形のハウジング Z 1 2 とイヤプラグとして構成され得るイヤホン Z 2 2 とを含む。図 2 6 A～図 2 6 Dはまた、デバイス D 2 0 0 のアレイの1次マイクロホンのための音響ポート Z 4 2 と2次マイクロホンのための音響ポート Z 5 2 との位置を示す。2次マイクロホンポートが少なくとも部分的に閉塞され得る (例えば、ユーザインタフェース・ボタンによって) ことはあり得る。

20

【0 1 4 2】

図 2 7 Aは、通信送受話器であるデバイス D 1 0 のマルチマイクロホン携帯型オーディオセンシング実現形態 D 3 0 0 の (中心軸に沿った) 断面図を示す。デバイス D 3 0 0 は、1次マイクロホン M C 1 0 と2次マイクロホン M C 2 0 とを有するアレイ R 1 0 0 の一実現形態を含む。この例では、デバイス D 3 0 0 はまた1次ラウドスピーカー S P 1 0 および2次ラウドスピーカー S P 2 0 も含む。このようなデバイスは、1つ以上の符号化および復号方式 (「コーデック」とも呼ばれる) を介して無線でボイス通信を送信および受信するように構成され得る。このようなコーデックの例は、2 0 0 7 年 2 月の「Enhanced Variable Rate Codec, Speech Service Options 3, 68 and 70 for Wideband Spread Spectrum Digital Systems」 (広帯域拡散スペクトル・デジタルシステムのための改良型可変速度コーデック、スピーチサービスオプション 3、6 8、および 7 0) と題する第 3 世代パートナーシッププロジェクト 2 (3 G P P 2) 文書 C. S 0 0 1 4 - C, v 1. 0 (www-dot-3gpp-dot-orgにおいてオンラインで利用可能) ; 2 0 0 4 年 1 月の「Selectable Mode Vocoder (SMV) Service Option for Wideband Spread Spectrum Communication Systems」 (広帯域拡散スペクトル通信システムのための選択可能モードボコーダ (S M V) サービスオプション) と題する 3 G P P 2 文書 C. S 0 0 3 0 - 0, v 3. 0 (www-dot-3gpp-dot-orgにおいてオンラインで利用可能) ; 文書 E T S I T S 1 2 6 0 9 2 V 6. 0. 0 ユーロッパ電気通信規格協会 (European Telecommunications Standards Institute (ETSI)、ソフィア・アンチポリス・セデックス、フランス、2 0 0 4 年 1 2 月 (Sophia Antipolis Cedex, FR, December 2004) に記載の適応型マルチレート (Adaptive Multi Rate (AMR)) スピーチコーデック ; および文書 E T S I T S 1 2 6 1 9 2 V 6. 0. 0 (E T S I、2 0 0 4 年 1 2 月) に記載の A M R 広帯域スピーチコーデック ; に記載されたような改良型可変速度コーデックを含む。図 3 A の例では、送受話器 D 3 0 0 はクラムシェルタイプの携帯電話送受話器 (「フリップ」ハンドセットとも呼ばれる) である。このようなマルチマイクロホン通信送受話器の他の構成はバータイプおよびスライダタイプの電話送受話器を含む。図 2 7 Bは、第 3 のマイクロホン M C 3 0 を含むアレイ R 1 0 0 の 3 マイクロホン実現形態 D 3 1 0 の断面図を示す。

30

40

【0 1 4 3】

図 2 8 Aは、メディアプレーヤーであるデバイス D 1 0 のマルチマイクロホン携帯型オーディオセンシング実現形態 D 4 0 0 の図を示す。このようなデバイスは、標準的圧縮フ

50

フォーマット（例えば、Moving Pictures Experts Group (MPEG)-1 Audio Layer 3 (MP3)、MPEG-4 Part 14 (MP4)、Windows（登録商標）Media Audio/Video (WMA/WMV)（マイクロソフト社、レッドモンド、ワシントン州 (Microsoft Corp., Redmond, WA)）の1バージョン、国際電気通信連合 (International Telecommunication Union)（ITU）-T H. 264など）に従って符号化されたファイルまたはストリームといった圧縮されたオーディオまたはオーディオビジュアル情報のプレイバック（再生）のために構成され得る。デバイスD400は、デバイスの前面に配置されたディスプレイスクリーンSC10とラウドスピーカーSP10とを含み、アレイR100のマイクロホンMC10およびMC20はデバイスの同じ面に（例えば、この例のように上面の両反対側に、または正面の両反対側に）配置される。図28BはマイクロホンMC10およびMC20がデバイスの両反対側に配置されたデバイスD400のもう1つの実現形態D410を示し、図28CはマイクロホンMC10およびMC20がデバイスの隣接面に配置されたデバイスD400の更なる実現形態D420を示す。メディアプレーヤーは意図した使用の間中、長軸が水平になっているようにもデザインされ得る。

【0144】

図29は、ハンズフリーカーキットであるデバイスD10のマルチマイクロホン携帯型オーディオセンシング実現形態D500の図を示す。このようなデバイスは、ダッシュボード、フロントガラス、バックミラー、サンバイザー、または車両のもう1つの内面にまたは内面上に設置されるように、または取り外し可能に固定されるように構成され得る。デバイスD500はラウドスピーカー85とアレイR100の一実現形態とを含む。この特定の例では、デバイスD500は直線状アレイに配置された4個のマイクロホンとしてのアレイR100の実現形態R102を含む。このようなデバイスは上記にリストアップされた例といった1つ以上のコーデックを介して無線でボイス通信データを送信および受信するように構成され得る。代替として、または更に、このようなデバイスは、（例えば、上記のようなBluetooth（商標）プロトコルの1バージョンを使用して）携帯電話送受話器といった電話デバイスとの通信を介して半二重または全二重電話方式をサポートするように構成され得る。

【0145】

図30は、ハンドヘルドアプリケーションのためのデバイスD10のマルチマイクロホン携帯型オーディオセンシング実現形態D600の図を示す。デバイスD600は、タッチスクリーン・ディスプレイTS10、3個の前部マイクロホンMC10～MC30、1個の後部マイクロホン、2個のラウドスピーカーSP10およびSP20、左側ユーザインタフェース・コントロール（例えば、選択用）UI10、および右側ユーザインタフェース・コントロール（例えば、ナビゲーション用）UI20を含む。ユーザインタフェース・コントロールの各々は、押しボタン、トラックボール、クリックホイール、タッチパッド、ジョイスティックおよび／または他のポインティングデバイスなどの1つ以上を使用して実現され得る。ブラウジングトーク (browse talk) モードまたはゲームプレイ・モードで使用され得るデバイスD800の典型的なサイズは約15センチメートル×20センチメートルである。本明細書で開示されたシステム、方法、および装置の適用可能性が図25A～図30に示されている特定の例に限定されないことは明らかに開示されている。このようなシステム、方法、および装置が適用され得る携帯型オーディオセンシング・デバイスの他の例は補聴器を含む。

【0146】

本明細書で開示された方法および装置は一般に、いかなる送受通信アプリケーションおよび／またはオーディオセンシング・アプリケーションにおいても、特に移動通信またはそうでなければこのようなアプリケーションの他の携帯型事例においても適用され得る。例えば、本明細書で開示された構成の範囲は、符号分割多元接続 (CDMA) 無線インタフェースを使用するように構成された無線電話通信システムに常駐する通信デバイスを含む。それにもかかわらず、本明細書で説明されたような特徴を有する方法および装置が有線および／または無線（例えば、CDMA、TDMA、FDMA、および／またはTD-

SCDMA) 伝送チャネルに亘ってボイスオーバーIP (VoIP) を使用するシステムといった当業者に知られた広範囲の技術を使用する種々の通信システムのいずれにも常駐し得ることは当業者によって理解されるであろう。

【0147】

本明細書で開示された通信デバイスがパケット交換されるネットワーク (例えば、VoIP といったプロトコルに従ってオーディオ伝送体を搬送するように整えられた有線および/または無線ネットワーク) および/または回線交換されるネットワークにおける使用のために適応し得ることは明確に考えられ、本明細書で開示されている。本明細書で開示された通信デバイスが狭帯域符号化システム (例えば、約 4 または 5 kHz のオーディオ周波数範囲を符号化するシステム) での使用のために、および/または全帯域広帯域符号化システムと分割帯域広帯域符号化システムとを含む広帯域符号化システム (例えば、5 kHz より高いオーディオ周波数を符号化するシステム) での使用のために、適応し得ることも明確に考えられ、本明細書で開示されている。

10

【0148】

本明細書で説明された構成の表現は、いかなる当業者も本明細書で開示された方法および他の構成を行う、または使用することを可能にするために提供されている。本明細書で図示され説明された流れ図、ブロック図および他の構成図は単なる例であってこれらの構成の他の変形版も本開示の範囲内にある。これらの構成に対する種々の修正は可能であり、本明細書で提示された一般的原理は他の構成にも同様に適用され得る。このように、本開示は上記の構成に限定されるようには意図されておらず、むしろ本明細書で任意の仕方で開示されて、出願された添付の請求項に含まれる、オリジナルな開示の一部を形成する原理と新規な特徴とに一致する最も広い範囲に合致させられるべきである。

20

【0149】

当業者は、情報および信号が種々の異なる技術および技法のいずれかを使用して表現され得ることを理解しているであろう。例えば、上記の説明全体を通して参照され得るデータ、命令、コマンド、情報、信号、ビット、および記号は、電圧、電流、電磁波、磁界または磁気粒子、光場または光粒子、またはこれらの任意の組合せによって表現され得る。

【0150】

本明細書で開示されたような構成の実現のための重要なデザイン要件は特に、圧縮されたオーディオまたはオーディオビジュアル情報 (例えば、本明細書で識別された例の 1 つといった圧縮フォーマットに従って符号化されたファイルまたはストリーム) のプレイバックといった計算集成的なアプリケーションまたは広帯域通信 (例えば、12、16、または 44 kHz といった 8 kHz より高いサンプリングレートでのボイス通信) のためのアプリケーションのために、処理遅延および/または計算の複雑さ (典型的には、1 秒当たり百万命令単位または MIPS 単位で測定される) を最小にすることを含み得る。

30

【0151】

マルチマイクロホン処理システムの目標は、全体的雑音低減において 10～12 デシベルを達成すること、所望スピーカーの移動中にも音声のレベルおよびカラーを維持すること、積極的な雑音除去の代わりに雑音が背景に移されたという認識を取得すること、スピーチの残響除去、および/またはより積極的な雑音低減のための後処理のオプションを取得すること、を含み得る。

40

【0152】

本明細書で開示された ANC 装置の一実現形態の種々の要素は、意図された用途のために適していると見なされるハードウェア、ソフトウェアおよび/またはファームウェアのいかなる組合せにおいても具体化され得る。例えば、このような要素は、例えば同じチップ上に、または 1 チップセット内の 2 つ以上のチップの間に常駐する電子および/または光デバイスとして製造され得る。このようなデバイスの 1 つの例は、トランジスタまたは論理ゲートといった論理要素の固定された、またはプログラム可能なアレイであり、これらの要素のいずれも 1 つ以上のこのようなアレイとして実現され得る。これらの要素の任意の 2 つ以上またはすべては、同じアレイ (単数または複数) 内に実現され得る。このよ

50

うなアレイ（単数または複数）は1つ以上のチップ内に（例えば、2つ以上のチップを含む1つのチップセット内に）実現され得る。

【0153】

本明細書で開示されたA N C装置の種々の実現形態の1つ以上の要素はまた、マイクロプロセッサ、埋め込み型プロセッサ、I Pコア、デジタル信号プロセッサ、F P G A（フィールドプログラマブル・ゲートアレイ）、A S S P（特定アプリケーション向け標準製品）およびA S I C（特定アプリケーション向け集積回路）といった論理要素の1つ以上の固定された、またはプログラム可能なアレイを実行するように整えられた命令の1つ以上のセットとして全体的または部分的に実現され得る。本明細書で開示されたような装置の実現形態の種々の要素のいかなるものでも、1つ以上のコンピュータ（例えば、「プロセッサ」とも呼ばれる、命令の1つ以上のセットまたは列を実行するようにプログラムされた1つ以上のアレイを含む機械）として実現されることが可能であり、またこれらの要素のいかなる2つ以上またはすべても、同じこのようなコンピュータ（単数または複数）内に実現され得る。

10

【0154】

本明細書で開示されたプロセッサまたは他の処理のための手段は、例えば同じチップ上に、またはチップセット内の2つ以上のチップの間に常駐する電子および／または光デバイスとして製造され得る。このようなデバイスの1つの例は、トランジスタまたは論理ゲートといった論理要素の固定された、またはプログラム可能なアレイであり、これらの要素のいずれも1つ以上のこのようなアレイとして実現され得る。このようなアレイ（単数または複数）は1つ以上のチップ内に（例えば、2つ以上のチップを含む1つのチップセット内に）実現され得る。このようなアレイの例は、マイクロプロセッサ、埋め込み型プロセッサ、I Pコア、D S P、F P G A、A S S P、およびA S I Cといった論理要素の1つ以上の固定された、またはプログラム可能なアレイを含む。本明細書で開示されたようなプロセッサまたは他の処理するための手段はまた、1つ以上のコンピュータ（例えば、命令の1つ以上のセットまたは列を実行するようにプログラムされた1つ以上のアレイを含む機械）または他のプロセッサとして具体化されることもあり得る。プロセッサが埋め込まれたデバイスまたはシステム（例えば、オーディオセンシング・デバイス）の他の動作に関連するタスクといったコヒーレンシー検出手順に直接関連しないタスクを実行するために、または命令の他のセットを実行するために、本明細書で説明されたプロセッサが使用されることは可能である。本明細書で開示された方法の一部がオーディオセンシング・デバイスのプロセッサによって実行されることも、またこの方法の他の一部が1つ以上の他のプロセッサのコントロール下で実行されることも可能である。

20

30

【0155】

当業者は、本明細書で開示された構成に関連して説明された種々の例示的モジュール、論理ブロック、回路、およびテストおよび他の動作が電子ハードウェア、コンピュータソフトウェアまたは両者の組合せとして実現され得ることを認めるであろう。このようなモジュール、論理ブロック、回路、および動作は、汎用プロセッサ、デジタル信号プロセッサ（D S P）、A S I CまたはA S S P、F P G Aまたは他のプログラム可能な論理デバイス、個別ゲートまたはトランジスタ論理、個別ハードウェア部品、または本明細書で説明されたような構成を作り出すためにデザインされたこれらの任意の組合せ、によって実現または実行され得る。例えば、このような構成は少なくとも部分的には、配線接続された回路として、または特定用途向け集積回路に製造された回路構成として、または汎用プロセッサまたは他のデジタル信号処理ユニットといった論理要素のアレイによって実行可能な命令である機械可読コードとして不揮発性記憶装置にロードされたファームウェアプログラムまたはデータ記憶媒体からまたはデータ記憶媒体にロードされたソフトウェアプログラムとして、実現され得る。汎用プロセッサはマイクロプロセッサであり得るが、代替としてプロセッサはいかなる従来型プロセッサ、コントローラ、マイクロコントローラ、または状態機械でもあり得る。プロセッサはまた、コンピューティングデバイスの組合せとして、例えばD S Pとマイクロプロセッサとの組合せ、複数のプロセッサ、D S

40

50

Pコアと連動する1つ以上のマイクロプロセッサ、または他の任意のこのような構成、としても実現され得る。ソフトウェアモジュールは、RAM（ランダムアクセス・メモリ）、ROM（読取り専用メモリ）、フラッシュRAMといった不揮発性RAM（NVRAM）、消去可能プログラマブルROM（EPROM）、電氣的消去可能プログラマブルROM（EEPROM）、レジスタ、ハードディスク、リムーバブルディスク、CD-ROM、または当分野で周知の他の任意の形式の記憶媒体、に常駐し得る。例示的記憶媒体は、この記憶媒体から情報を読取ることができて、この記憶媒体に情報を書き込むことができるプロセッサに連結される。代替として、記憶媒体はプロセッサと一体化され得る。プロセッサおよび記憶媒体はASIC内に常駐し得る。ASICはユーザ端末内に常駐し得る。代替としてプロセッサおよび記憶媒体はユーザ端末内に個別部品として常駐し得る。

10

【0156】

本明細書で開示された種々の方法がプロセッサといった論理要素のアレイによって実行され得ること、および本明細書で説明された装置の種々の要素がこのようなアレイ上で実行するようにデザインされたモジュールとして実現され得ることに留意されたい。本明細書で使用されているように、用語「モジュール」または「サブモジュール」は、ソフトウェア、ハードウェア、またはファームウェア形式のコンピュータ命令（たとえば、論理式）を含む任意の方法、装置、デバイス、ユニットまたはコンピュータ可読データ記憶媒体を指すことができる。多数のモジュールまたはシステムが1つのモジュールまたはシステムに組み合されることが可能であり、また1つのモジュールまたはシステムが同じ機能を実行するために多数のモジュールまたはシステムに分離されることが可能であることは理解されるべきである。ソフトウェアまたは他のコンピュータ実行可能命令に実現されたとき、プロセスの要素は本質的に、ルーチン、プログラム、オブジェクト、コンポーネント、データ構造などといった関連タスクを実行するためのコードセグメントである。用語「ソフトウェア」は、ソースコード、アセンブリ言語コード、マシンコード、2進コード、ファームウェア、マクロコード、マイクロコード、論理要素のアレイによって実行可能な命令の任意の1つ以上のセットまたは列、およびこのような例の任意の組合せ、を含むと理解されるべきである。プログラムまたはコードセグメントは、プロセッサ可読媒体に記憶され得るか、または搬送波内で具体化されたコンピュータデータ信号によって伝送媒体または通信リンク上で伝送され得る。

20

【0157】

本明細書で開示された方法、方式、および技法の実現形態は、論理要素のアレイを含む機械（例えば、プロセッサ、マイクロプロセッサ、マイクロコントローラ、または他の有限状態機械）によって読取り可能および／または実行可能な命令の1つ以上のセットとして（例えば、本明細書でリストアップされたような1つ以上のコンピュータ可読媒体に）明確に具体化されることもあり得る。用語「コンピュータ可読媒体」は、揮発性、不揮発性、リムーバブル、およびノンリムーバブル媒体を含む、情報を記憶または移送できる任意の媒体を含み得る。コンピュータ可読媒体の例は、電子回路、半導体メモリデバイス、ROM、フラッシュメモリ、消去可能ROM（EROM）、フロッピー（登録商標）ディスクまたは他の磁気記憶装置、CD-ROM/DVDまたは他の光記憶装置、ハードディスク、光ファイバ媒体、無線周波数（RF）リンク、または所望の情報を記憶するために使用され得るそしてアクセスされ得る他の任意の媒体を含む。コンピュータデータ信号は、電子ネットワークチャネル、光ファイバ、空気、電磁、RFリンクなどといった伝送媒体上を伝播し得るいかなる信号も含み得る。コードセグメントは、インターネットまたはイントラネットといったコンピュータネットワークを介してダウンロードされ得る。いずれの場合にも、本開示の範囲はこのような実施形態によって限定されると解釈されるべきではない。

30

40

【0158】

本明細書で説明された方法のタスクの各々は直接的に、ハードウェアに、またはプロセッサによって実行されるソフトウェアモジュールに、またはこれら2つの組合せに、具体化され得る。本明細書で開示された方法の実現形態の典型的用途では、論理要素（例えば

50

、論理ゲート)のアレイは、この方法の種々のタスクのうちの1つのタスク、または1つより多いタスク、またはすべてのタスクを実行するように構成されている。これらのタスクの1つ以上(場合によってはすべて)はまた、論理要素のアレイ(例えば、プロセッサ、マイクロプロセッサ、マイクロコントローラ、または他の有限状態機械)を含む機械(例えば、コンピュータ)によって読取り可能および/または実行可能であるコンピュータプログラム製品(例えば、ディスク、フラッシュまたは他の不揮発性メモリカード、半導体メモリチップなどといった1つ以上のデータ記憶媒体)に具体化されたコード(例えば、命令の1つ以上のセット)としても実現され得る。本明細書で開示されたような方法の実現形態のタスクは1つより多いこのようなアレイまたは機械によって実行されることもあり得る。これらまたは他の実現形態ではこれらのタスクは、携帯電話またはこのような通信機能を有する他のデバイスといった無線通信デバイス内で実行され得る。このようなデバイスは、(例えば、V o I Pといった1つ以上のプロトコルを使用して)回線交換および/またはパケット交換ネットワークと通信するように構成され得る。例えば、このようなデバイスは符号化されたフレームを受信および/または送信するように構成されたR F回路を含み得る。

10

【0159】

本明細書で開示された種々の方法が送受話器、ヘッドホン、またはポータブル・デジタル・アシスタント(PDA)といった携帯型通信デバイスによって実行され得ること、および本明細書で説明された種々の装置がこのようなデバイス内に含まれ得ることは明確に開示されている。典型的なリアルタイム(例えば、オンライン)アプリケーションはこ

20

【0160】

1つ以上の例示的实施形態では本明細書で説明された動作は、ハードウェア、ソフトウェア、ファームウェア、またはこれらの任意の組合せにおいて実現され得る。ソフトウェアにおいて実現された場合には、このような動作は1つ以上の命令またはコードとしてコンピュータ可読媒体上に記憶され得るか、またはコンピュータ可読媒体上を伝送され得る。用語「コンピュータ可読媒体」は、1つの場所から他の場所へのコンピュータプログラムの移送を容易にするいかなる媒体も含むコンピュータ記憶媒体および通信媒体の両者を含む。記憶媒体は、コンピュータによってアクセスされ得るいかなる利用可能な媒体でもあり得る。限定ではなく例として、このようなコンピュータ可読媒体は、半導体メモリ(ダイナミックまたはスタティックRAM、ROM、EEPROM、および/またはフラッシュRAMを限定なしに含み得る)、または強誘電性、磁気抵抗性、オボニック(ovonic)、ポリマー、または位相変化メモリ;CD-ROMまたは他の光ディスク記憶装置、磁気ディスク記憶装置または他の磁気記憶デバイス、またはコンピュータによってアクセスされ得る具体的構造の命令またはデータ構造の形をした所望のプログラムコードを記憶するために使用され得る他の任意の媒体;といった記憶要素のアレイを備え得る。また、いかなる接続媒体も適切にコンピュータ可読媒体と名付けられる。例えば、ソフトウェアが同軸ケーブル、光ファイバケーブル、より対線、デジタル加入者線(DSL)、または赤外線、無線および/またはマイクロ波といった無線技術、を使用してウェブサイト、サーバまたは他の遠隔情報源から送信される場合には、これらの同軸ケーブル、光ファイバケーブル、より対線、DSL、または赤外線、無線および/またはマイクロ波といった無線技術、は媒体の定義に含まれる。本明細書で使用されるようなディスク(disk)およびディスク(disc)はコンパクトディスク(CD)、レーザーディスク(登録商標)、光ディスク、デジタル・バーサタイル・ディスク(DVD)、フロッピーディスクおよびブルーレイディスク(Blu-ray Disc(商標))(ブルーレイディスク・アソシエーション、ユニバーサルシティ、カリフォルニア州(Blu-Ray Disc Association, Universal City, CA))を含む。ここで、diskは通常、データを磁氣的に再生するが、discはデータをレーザによって光学的に再生する。上記のものの組合せもコンピュータ可読媒体の範囲内に含まれるべきである。

30

40

【0161】

50

本明細書で説明されたような音響信号処理装置は、ある幾つかの動作を制御するためにスピーチ入力を受け入れる、またはそうでなければ背景雑音からの所望の雑音の分離から利益を得ることがあり得る、通信デバイスといった電子デバイスに組み込まれ得る。多くの用途は、多数の方向から発生する背景サウンドから明瞭な所望のサウンドを増強または分離することから利益を得ることがあり得る。このような用途は、ボイス認識および検出、ボイス増強および分離、ボイス活性化制御などといった機能を組み入れている電子またはコンピューティングデバイスにおける人間・機械インタフェースを含み得る。単に限定された機能を提供するデバイスに適するようにこのような音響信号処理装置を実現することが望ましいことができる。

【0162】

10

本明細書で説明されたモジュール、要素、およびデバイスの種々の実現形態の要素は、例えば同じチップ上に、または1つのチップセット内の2つ以上のチップ間に常駐する電子および／または光デバイスとして製造され得る。このようなデバイスの1つの例は、トランジスタまたはゲートといった論理要素の固定された、またはプログラム可能なアレイである。本明細書で説明された装置の種々の実現形態の1つ以上の要素はまた、マイクロプロセッサ、埋め込み型プロセッサ、IPコア、デジタル信号プロセッサ、FPGA、ASSP、およびASICといった論理要素の1つ以上の固定された、またはプログラム可能なアレイ上で実行するように整えられた命令の1つ以上のセットとして全体的または部分的に実現され得る。

【0163】

20

本明細書で説明された装置の実現形態の1つ以上の要素が、この装置の動作に直接的には関連しない他の命令セットを実行するために、またはこの装置が埋め込まれたデバイスまたはシステムの他の動作に関連するタスクといったタスクを実行するために、使用されることは可能である。このような装置の実現形態の1つ以上の要素が、共通した構造（例えば、異なるときに異なる要素に対応するコードの部分を実行するために使用されるプロセッサ、異なるときに異なる要素に対応するタスクを実行するために実行される1セットの命令、または異なるときに異なる要素に関する動作を実行する電子および／または光デバイスの配置）を有することも可能である。

【図 1】

図 1

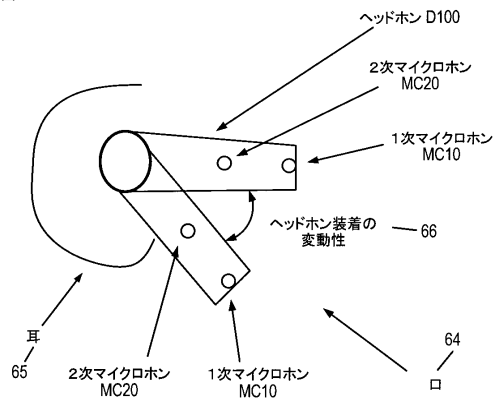


FIG. 1

【図 2】

図 2

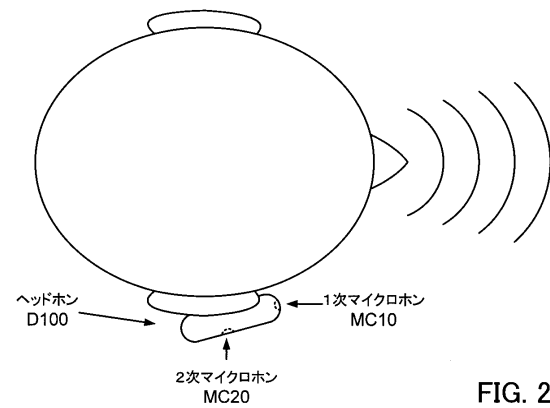


FIG. 2

【図 3 A】

図 3 A

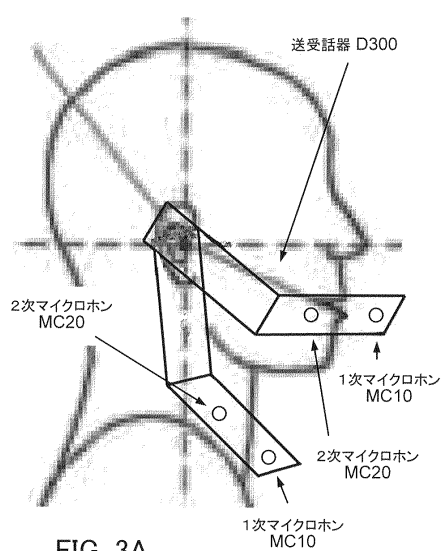


FIG. 3A

【図 3 B】

図 3 B

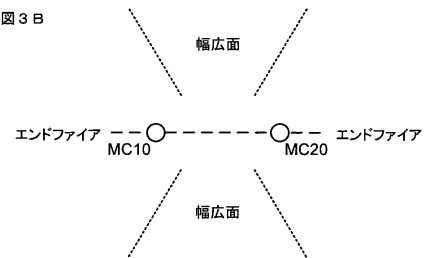


FIG. 3B

【図 4 A】

図 4 A

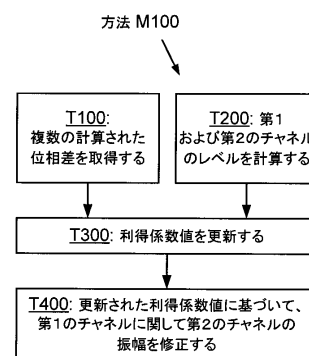
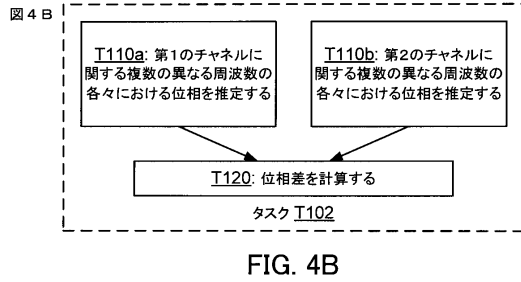
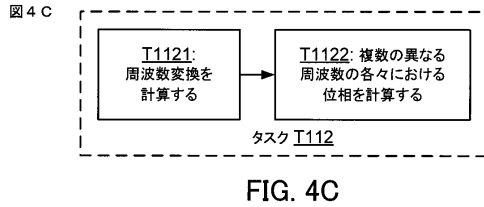


FIG. 4A

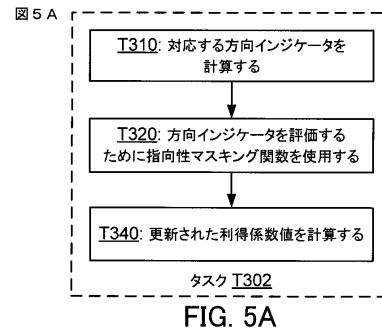
【図 4 B】



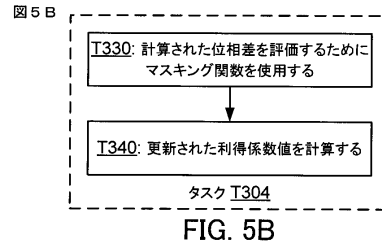
【図 4 C】



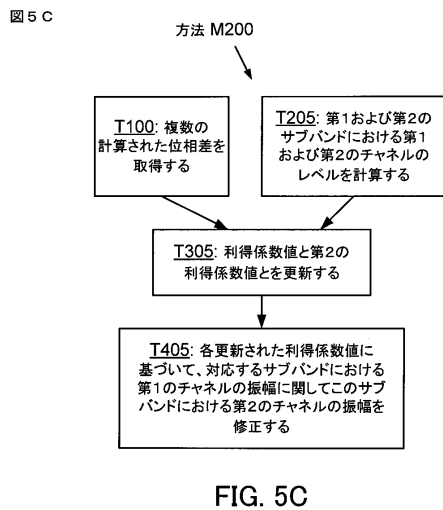
【図 5 A】



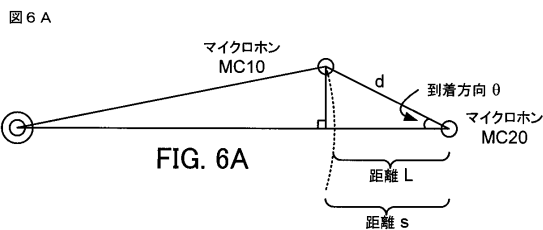
【図 5 B】



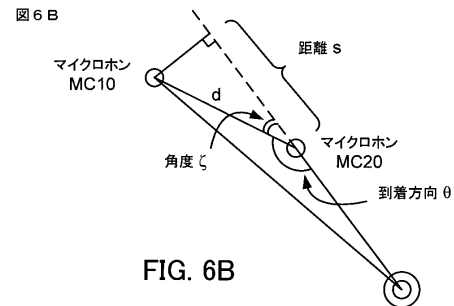
【図 5 C】



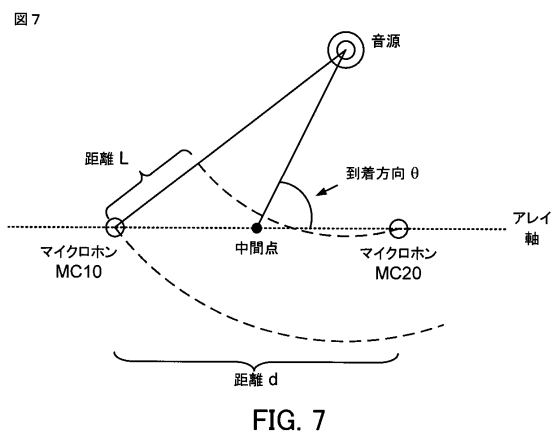
【図 6 A】



【図 6 B】



【図 7】



【図 8 A】

図 8 A

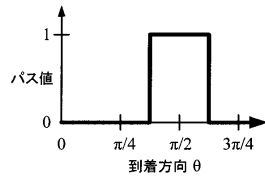


FIG. 8A

【図 9 A】

図 9 A

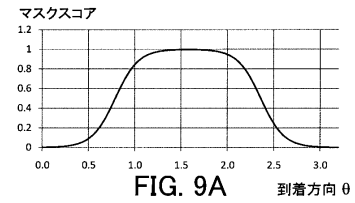


FIG. 9A

【図 8 B】

図 8 B

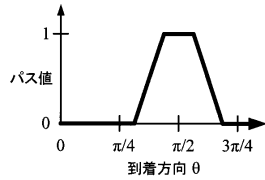


FIG. 8B

【図 9 B】

図 9 B

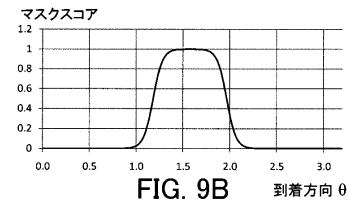


FIG. 9B

【図 8 C】

図 8 C

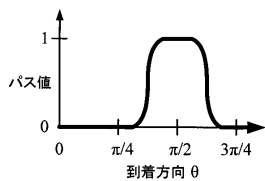


FIG. 8C

【図 9 C】

図 9 C

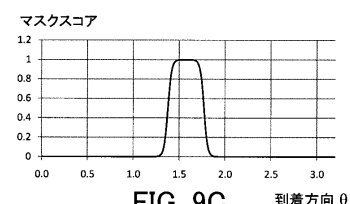


FIG. 9C

【図 10】

図 10

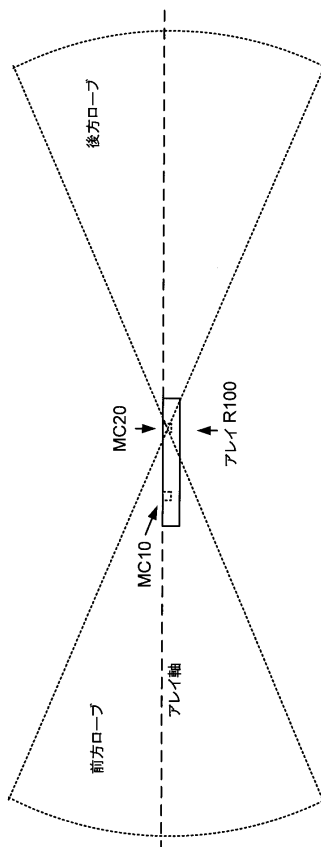


FIG. 10

【図 11 A】

図 11 A

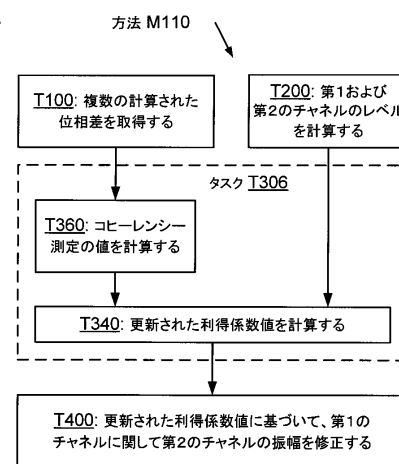


FIG. 11A

【図 1 1 B】

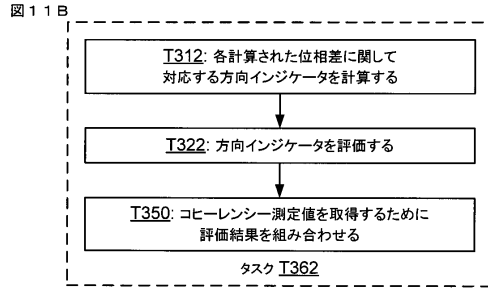


FIG. 11B

【図 1 1 C】

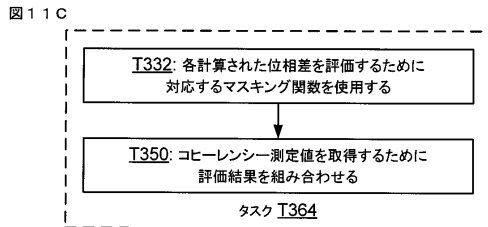


FIG. 11C

【図 1 2 A】

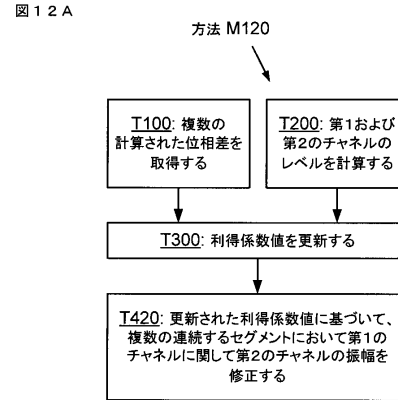


FIG. 12A

【図 1 2 B】

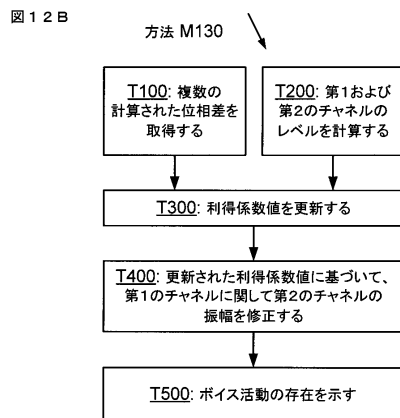


FIG. 12B

【図 1 3 A】

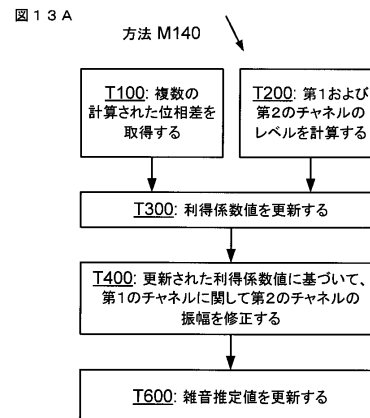


FIG. 13A

【図 1 3 B】

図 1 3 B

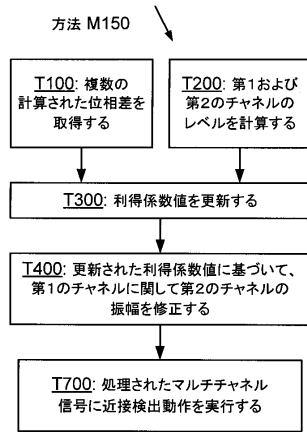


FIG. 13B

【図 1 4 A】

図 1 4 A

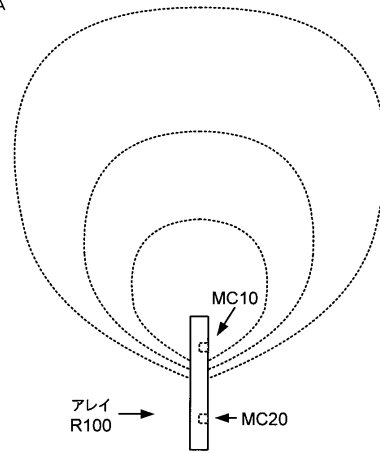


FIG. 14A

【図 1 4 B】

図 1 4 B

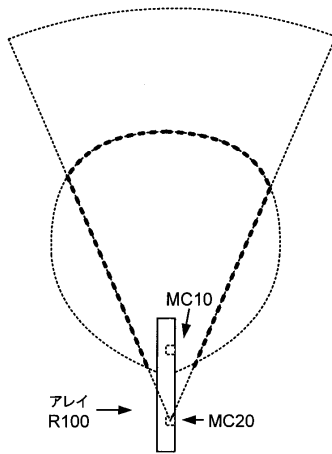


FIG. 14B

【図 1 5】

図 1 5

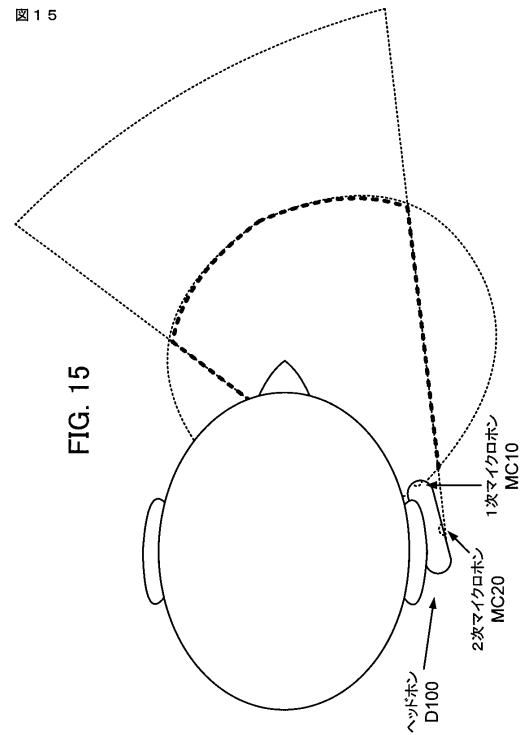
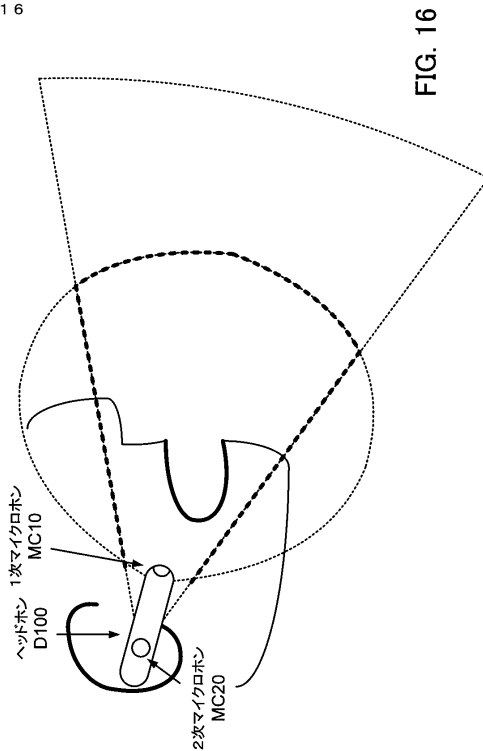


FIG. 15

【図 16】

図 16



【図 17 A】

図 17 A

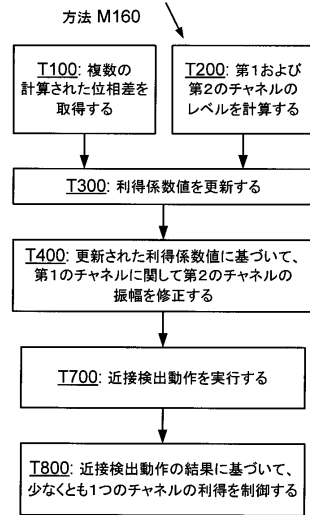


FIG. 17A

【図 17 B】

図 17 B

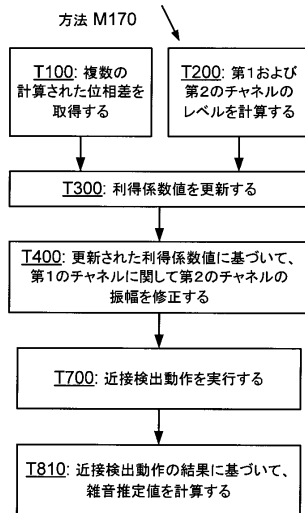


FIG. 17B

【図 18】

図 18

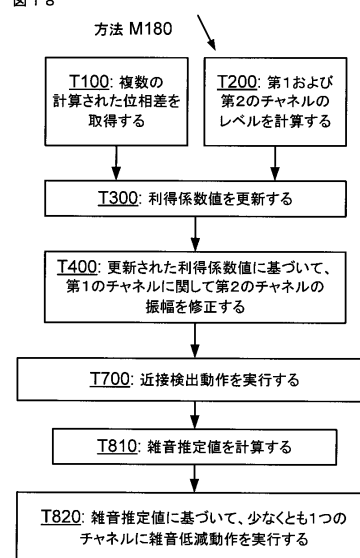


FIG. 18

【図 19 A】

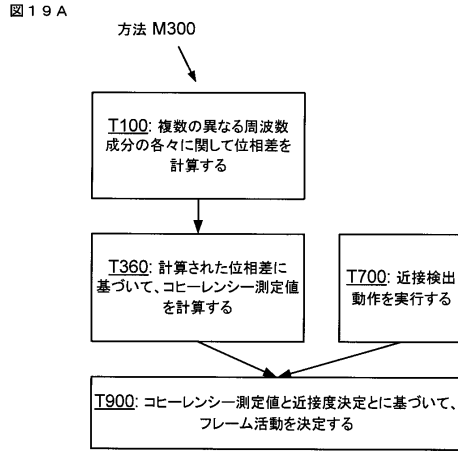


FIG. 19A

【図 19 B】

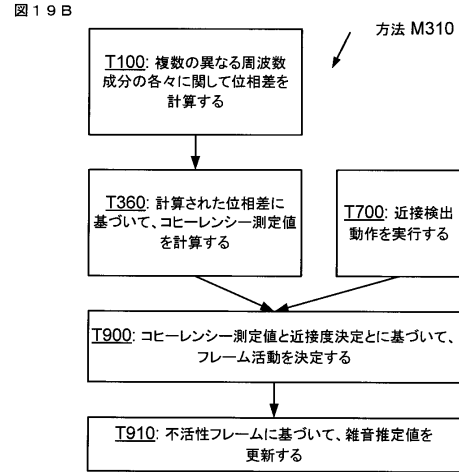


FIG. 19B

【図 20 A】

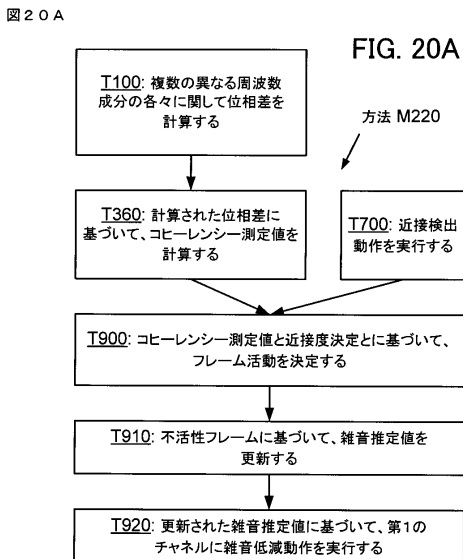


FIG. 20A

【図 20 B】

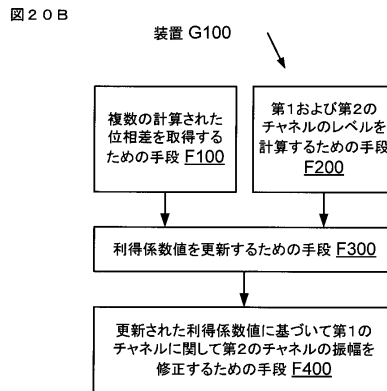
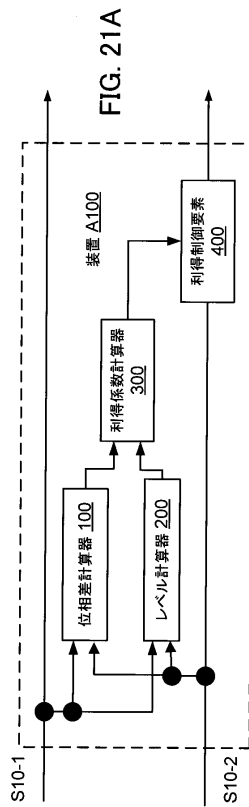


FIG. 20B

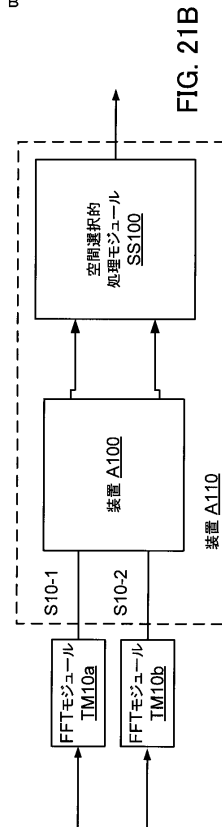
【図 2 1 A】

図 2 1 A



【図 2 1 B】

図 2 1 B



【図 2 2】

図 2 2

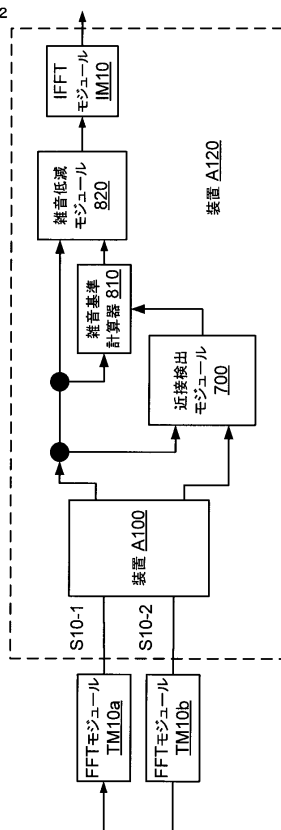
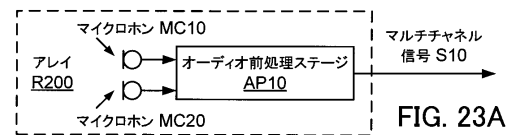


FIG. 22

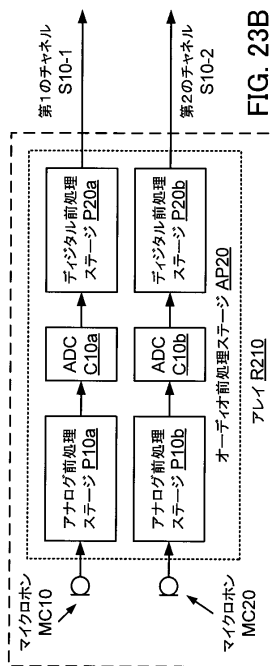
【図 2 3 A】

図 2 3 A



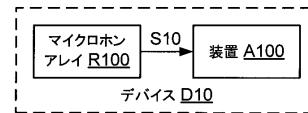
【図 2 3 B】

図 2 3 B



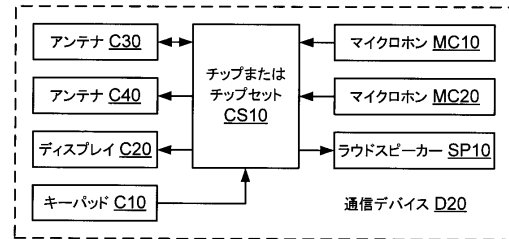
【図 2 4 A】

図 2 4 A



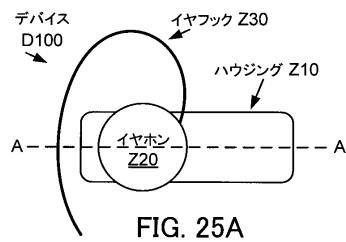
【図 2 4 B】

図 2 4 B



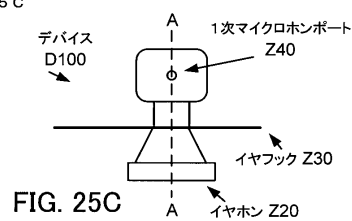
【図 2 5 A】

図 2 5 A



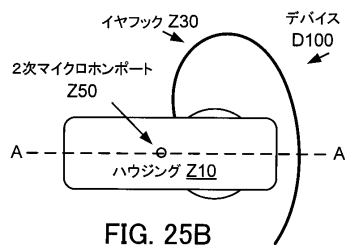
【図 2 5 C】

図 2 5 C



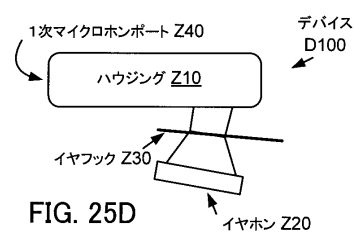
【図 2 5 B】

図 2 5 B

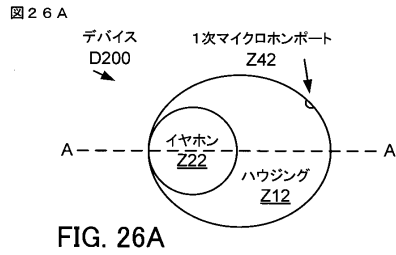


【図 2 5 D】

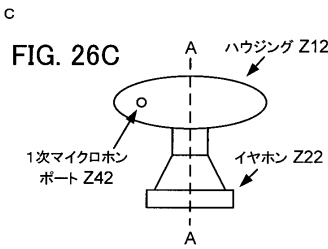
図 2 5 D



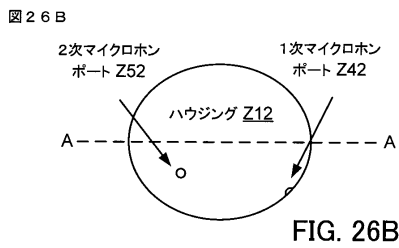
【図 26 A】



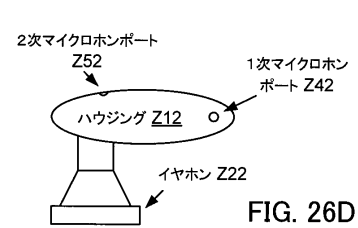
【図 26 C】



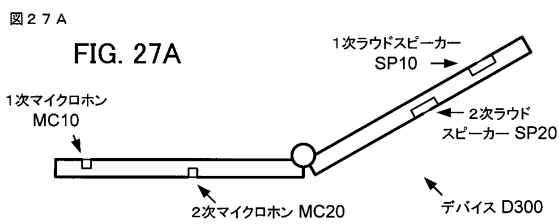
【図 26 B】



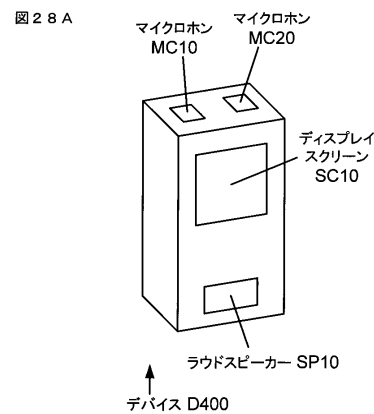
【図 26 D】



【図 27 A】



【図 28 A】



【図 27 B】

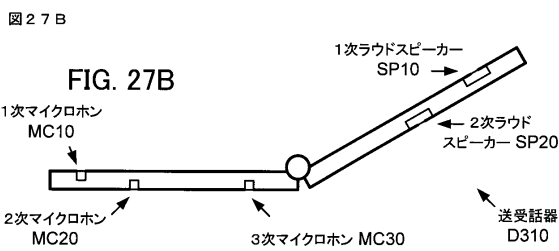


FIG. 28A

【図 28 B】

図 28 B

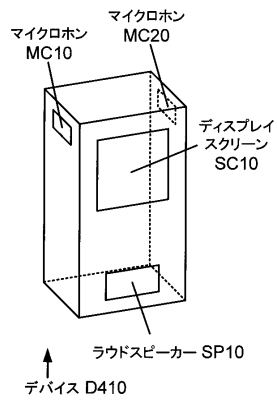


FIG. 28B

【図 28 C】

図 28 C

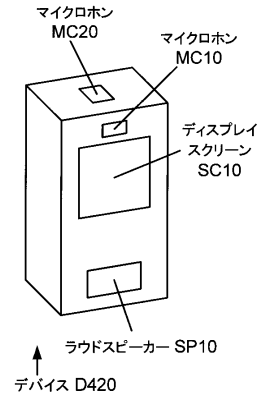


FIG. 28C

【図 29】

図 29

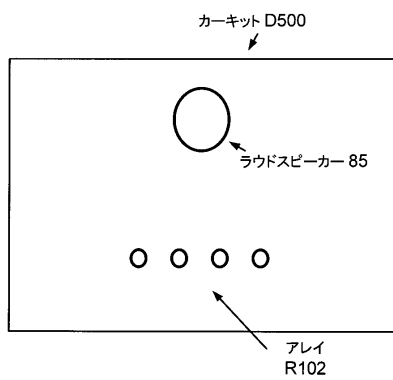


FIG. 29

【図 30】

図 30

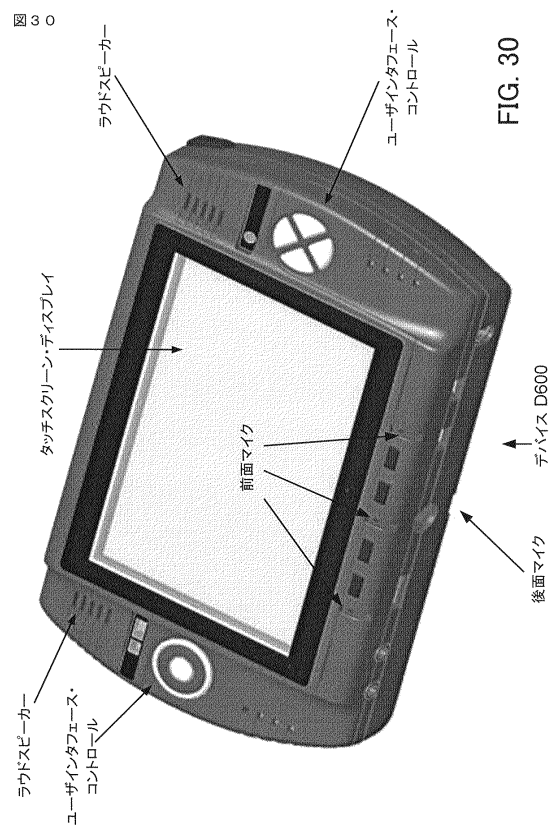


FIG. 30

フロントページの続き

- (31)優先権主張番号 61/240,318
(32)優先日 平成21年9月8日(2009.9.8)
(33)優先権主張国 米国(US)
(31)優先権主張番号 61/240,320
(32)優先日 平成21年9月8日(2009.9.8)
(33)優先権主張国 米国(US)
(31)優先権主張番号 12/796,566
(32)優先日 平成22年6月8日(2010.6.8)
(33)優先権主張国 米国(US)

- (74)代理人 100109830
弁理士 福原 淑弘
(74)代理人 100075672
弁理士 峰 隆司
(74)代理人 100095441
弁理士 白根 俊郎
(74)代理人 100084618
弁理士 村松 貞男
(74)代理人 100103034
弁理士 野河 信久
(74)代理人 100119976
弁理士 幸長 保次郎
(74)代理人 100153051
弁理士 河野 直樹
(74)代理人 100140176
弁理士 砂川 克
(74)代理人 100158805
弁理士 井関 守三
(74)代理人 100124394
弁理士 佐藤 立志
(74)代理人 100112807
弁理士 岡田 貴志
(74)代理人 100111073
弁理士 堀内 美保子
(74)代理人 100134290
弁理士 竹内 将訓
(72)発明者 ビッサー、エリック
アメリカ合衆国、カリフォルニア州 92121-1714、サン・ディエゴ、モアハウス・ドライブ 5775
(72)発明者 リウ、エルナン
アメリカ合衆国、カリフォルニア州 92121-1714、サン・ディエゴ、モアハウス・ドライブ 5775

審査官 大野 弘

(56)参考文献 特開2007-068125(JP, A)

(58)調査した分野(Int.Cl., DB名)

H04R	3/00
G10L	21/0332
H04R	1/10
H04R	1/40