

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第6460676号
(P6460676)

(45) 発行日 平成31年1月30日 (2019. 1. 30)

(24) 登録日 平成31年1月11日 (2019. 1. 11)

(51) Int. Cl.

F I

H O 4 R 3/00 (2006. 01)

H O 4 R 3/00 3 2 O

H O 4 R 1/40 (2006. 01)

H O 4 R 1/40 3 2 O A

H O 4 N 5/232 (2006. 01)

H O 4 N 5/232 O 3 O

請求項の数 23 (全 23 頁)

(21) 出願番号 特願2014-159761 (P2014-159761)
 (22) 出願日 平成26年8月5日 (2014. 8. 5)
 (65) 公開番号 特開2016-39410 (P2016-39410A)
 (43) 公開日 平成28年3月22日 (2016. 3. 22)
 審査請求日 平成29年7月20日 (2017. 7. 20)

(73) 特許権者 000001007
 キヤノン株式会社
 東京都大田区下丸子3丁目30番2号
 (74) 代理人 100076428
 弁理士 大塚 康德
 (74) 代理人 100112508
 弁理士 高柳 司郎
 (74) 代理人 100115071
 弁理士 大塚 康弘
 (74) 代理人 100116894
 弁理士 木村 秀二
 (74) 代理人 100130409
 弁理士 下山 治
 (74) 代理人 100134175
 弁理士 永川 行光

最終頁に続く

(54) 【発明の名称】 信号処理装置および信号処理方法

(57) 【特許請求の範囲】

【請求項 1】

形状が可変である所定の機器を装着可能な信号処理装置であって、前記信号処理装置が有するマイクから複数チャンネルの音響信号を取得する音取得手段と

、

前記信号処理装置に装着された前記機器の形状の変化に伴って変化するパラメータを取得するパラメータ取得手段と、方向に応じた音の伝搬特性を表す情報であって前記機器の形状に応じた伝搬特性を表す情報を、前記パラメータ取得手段により取得されたパラメータに基づいて取得する情報取得手段と、

前記音取得手段により取得された音響信号を、前記情報取得手段により取得された情報を用いて処理する処理手段と、を有することを特徴とする信号処理装置。

【請求項 2】

前記パラメータ取得手段は、前記信号処理装置としてのカメラに前記所定の機器としてのズームレンズが装着された場合に、前記ズームレンズの形状の変化に伴って変化するパラメータとして、ズーム倍率を表すパラメータを取得することを特徴とする請求項 1 に記載の信号処理装置。

【請求項 3】

前記パラメータ取得手段は、前記ズームレンズと通信することにより前記ズーム倍率を表すパラメータを取得することを特徴とする請求項 2 に記載の信号処理装置。

10

20

【請求項 4】

前記音取得手段により取得された音響信号に対する前記処理手段による処理の結果に基づいて、当該音響信号に係る音源の方向を推定する推定手段を有することを特徴とする請求項 1 乃至 3 の何れか 1 項に記載の信号処理装置。

【請求項 5】

前記音取得手段により取得された音響信号から、前記推定手段により推定された方向に対応する音響信号を分離する分離手段を有することを特徴とする請求項 4 に記載の信号処理装置。

【請求項 6】

前記伝搬特性を表す情報は、アレイ・マニフォールド・ベクトルを示す情報であることを特徴とする請求項 1 乃至 5 の何れか 1 項に記載の信号処理装置。

10

【請求項 7】

前記情報取得手段は、複数の所定のアレイ・マニフォールド・ベクトルのうち、各アレイ・マニフォールド・ベクトルに対応する感度曲線の極値に基づいて選択されたアレイ・マニフォールド・ベクトルを、前記伝搬特性を表す情報として取得することを特徴とする請求項 1 乃至 5 の何れか 1 項に記載の信号処理装置。

【請求項 8】

前記情報取得手段は、保持部により保持される複数のアレイ・マニフォールド・ベクトルのうち、前記パラメータ取得手段により取得されたパラメータに対応するアレイ・マニフォールド・ベクトルを、前記伝搬特性を表す情報として取得することを特徴とする請求項 1 乃至 5 の何れか 1 項に記載の信号処理装置。

20

【請求項 9】

前記情報取得手段は、前記パラメータ取得手段により取得されたパラメータに対応するアレイ・マニフォールド・ベクトルが前記保持部により保持されていない場合には、前記保持部により保持されるアレイ・マニフォールド・ベクトルに対応するパラメータと前記パラメータ取得手段により取得されたパラメータとの差に基づいて選択されたアレイ・マニフォールド・ベクトルを、前記伝搬特性を表す情報として取得することを特徴とする請求項 8 に記載の信号処理装置。

【請求項 10】

前記情報取得手段は、前記パラメータ取得手段により取得されたパラメータに対応するアレイ・マニフォールド・ベクトルが前記保持部により保持されていない場合には、前記保持部により保持される 2 以上のアレイ・マニフォールド・ベクトルを用いた補間処理により生成されたアレイ・マニフォールド・ベクトルを、前記伝搬特性を表す情報として取得することを特徴とする請求項 8 に記載の信号処理装置。

30

【請求項 11】

前記音取得手段は、少なくとも一部が前記信号処理装置の筐体内に存在する第 1 マイクから複数チャネルの音響信号を取得する第 1 取得手段と、

前記筐体の外部に備えられる第 2 マイクから複数チャネルの音響信号を取得する第 2 取得手段と、を有し、

40

前記処理手段は、前記第 1 取得手段により取得された音響信号を処理する場合には、方向に応じた音の伝搬特性を表す第 1 情報を用いて音響信号を処理し、前記第 2 取得手段により取得された音響信号を処理する場合には、方向に応じた音の伝搬特性を表す第 2 情報であって前記第 1 情報が表す伝搬特性とは異なる伝搬特性を表す第 2 情報を用いて音響信号を処理することを特徴とする請求項 1 乃至 10 の何れか 1 項に記載の信号処理装置。

【請求項 12】

前記第 1 情報は前記筐体の影響を含む伝搬特性を表し、前記第 2 情報は前記筐体の影響を含まない伝搬特性を表すことを特徴とする請求項 11 に記載の信号処理装置。

【請求項 13】

前記第 1 情報を用いて音響信号を処理するか前記第 2 情報を用いて音響信号を処理する

50

かを選択する選択手段を有することを特徴とする請求項 1 1 又は 1 2 に記載の信号処理装置。

【請求項 1 4】

前記第 2 情報は、前記筐体の影響を含まない音の伝搬特性であって前記第 2 マイクが有する複数のマイク素子の間隔に基づいて決まる伝搬特性を表すことを特徴とする請求項 1 1 乃至 1 3 の何れか 1 項に記載の信号処理装置。

【請求項 1 5】

前記第 2 マイクが有する複数のマイク素子の間隔を示す情報を取得する第 3 取得手段をさらに備え、

前記処理手段は、前記第 2 取得手段により取得された音響信号を処理する場合には、前記第 3 取得手段により取得される情報に基づいて取得される前記第 2 情報を用いて音響信号を処理することを特徴とする請求項 1 1 乃至 1 4 の何れか 1 項に記載の信号処理装置。

【請求項 1 6】

複数パターンのマイク素子の間隔に対応する複数の感度曲線の極値に基づいて、前記第 2 マイクが有する複数のマイク素子の間隔を推定する第 2 推定手段をさらに備えることを特徴とする請求項 1 4 に記載の信号処理装置。

【請求項 1 7】

前記第 2 情報を前記第 2 マイクから取得する第 4 取得手段を有することを特徴とする請求項 1 1 乃至 1 6 の何れか 1 項に記載の信号処理装置。

【請求項 1 8】

前記筐体の形状を判定する判定手段をさらに備え、

前記第 1 情報は、前記判定手段による判定結果に対応する伝搬特性を表すことを特徴とする請求項 1 1 乃至 1 7 のいずれか 1 項に記載の信号処理装置。

【請求項 1 9】

形状が可変である所定の機器を装着可能な信号処理装置により実行される信号処理方法であって、

前記信号処理装置が有するマイクから複数チャンネルの音響信号を取得する音取得工程と、

前記信号処理装置に装着された前記機器の形状の変化に伴って変化するパラメータを取得するパラメータ取得工程と、

方向に応じた音の伝搬特性を表す情報であって前記機器の形状に応じた伝搬特性を表す情報を、前記パラメータ取得工程において取得されたパラメータに基づいて取得する情報取得工程と、

前記音取得工程において取得された音響信号を、前記情報取得工程において取得された情報を用いて処理する処理工程と、を有することを特徴とする信号処理方法。

【請求項 2 0】

前記パラメータ取得工程は、前記信号処理装置としてのカメラに前記所定の機器としてのズームレンズが装着された場合に、前記ズームレンズの形状の変化に伴って変化するパラメータとして、ズーム倍率を表すパラメータを取得することを特徴とする請求項 1 9 に記載の信号処理方法。

【請求項 2 1】

前記情報取得工程は、保持部に保持される複数のアレイ・マニフォールド・ベクトルのうち、前記パラメータ取得工程において取得されたパラメータに対応するアレイ・マニフォールド・ベクトルを、前記伝搬特性を表す情報として取得することを特徴とする請求項 1 9 又は 2 0 に記載の信号処理方法。

【請求項 2 2】

前記情報取得工程は、前記パラメータ取得工程において取得されたパラメータに対応するアレイ・マニフォールド・ベクトルが前記保持部により保持されていない場合には、前記保持部により保持される 2 以上のアレイ・マニフォールド・ベクトルを用いた補間処理により生成されたアレイ・マニフォールド・ベクトルを、前記伝搬特性を表す情報として

10

20

30

40

50

取得することを特徴とする請求項 2 1 に記載の信号処理方法。

【請求項 2 3】

コンピュータを、請求項 1 乃至 1 8 のいずれか 1 項に記載の信号処理装置の各手段として機能させるためのプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、音響処理を行う信号処理装置および信号処理方法に関する。

【背景技術】

【0002】

音響信号から不要な雑音を除去する技術は、音響信号に含まれる目的音に対する聴感を改善し、また、音声認識においては認識率を高めるために重要な技術である。音響信号の雑音を除去する技術として代表的なものにビームフォーマがある。これは、複数のマイク素子で収音した複数チャンネルのマイク信号にそれぞれフィルタリングを施してから加算し、単一の出力信号を得るものである。上記のフィルタリングと加算の処理が、複数のマイク素子で指向性、すなわち方向選別性を持つ空間的なビームパターンを形成することに対応するため、ビームフォーマと呼ばれる。

【0003】

ビームパターンの感度（ゲイン）がピークとなる部分をメインローブといい、メインローブが目的音の方向を向くようにビームフォーマを構成すれば、目的音を強調し、同時に目的音と異なる方向に存在する雑音を抑制することができる。しかしながら、ビームパターンのメインローブは、特にマイク素子数が少ない場合に広い幅を持つなだらかな曲線となる。このため、ビームパターンのそのようなメインローブを目的音の方向に向けても、目的音に近接するような雑音を十分に除去することはできない。

【0004】

そこでメインローブではなく、ビームパターンの感度がディップとなる部分である、ヌル（死角）を利用した雑音除去の方法が提案されている。すなわち、鋭いヌルを雑音の方向に向けることで、近接する目的音を削ることなく、雑音のみを十分に除去することができる。このように、特定の方向に固定的にヌルを形成するビームフォーマを固定ビームフォーマと呼ぶ。ここで、ヌルを向ける方向が正確でないと、雑音除去の性能が大きく劣化するため、音源方向の推定が重要となってくる。

【0005】

固定ビームフォーマに対し、ビームパターンのヌルが自動的に形成されるビームフォーマを適応ビームフォーマと呼び、適応ビームフォーマを音源方向の推定に用いることができる。目的音や雑音をパワーが空間的に一点に集中した方向性音源と考えれば、出力パワーを最小化する規範に基づく適応ビームフォーマを用いることで、音源方向にヌルを自動的に形成するようなフィルタ係数が得られる。よって、音源方向を知るためには、適応ビームフォーマのフィルタ係数で形成されるビームパターンを算出して、そのヌル方向を求めればよい。ビームパターンは、アレイ・マニフォールド・ベクトルと呼ばれる各方向の音源と各マイク素子間の伝達関数を、フィルタ係数に掛け合わせることで算出することができる。例えば、方向 - 180° から 180° まで 1° 刻みのアレイ・マニフォールド・ベクトルを用いて、フィルタ係数が何度の方向に、感度のディップであるヌルを持っているかを調べる。

【0006】

ここで、一般にビームフォーマのような音源分離においては、マイクが自由空間に配置されたものと仮定して、自由空間の理論式によるアレイ・マニフォールド・ベクトルを用いることが多い。自由空間では遮るものなく理想的に音が伝搬するため、例えばマイク素子間の伝搬遅延時間差、すなわちアレイ・マニフォールド・ベクトル要素間の周波数ごとの位相差は、幾何的にマイク間隔をパラメータとする理論式で求められる。一方、マイクが自由空間ではなく、筐体の近傍や内部に配置された場合は、筐体による音の回折・遮断

10

20

30

40

50

・散乱などが生じるため、上記位相差は自由空間の理論値から乖離する。また、各音源方向に対するマイク素子間の信号の振幅差も、マイク素子が配置される筐体の影響を受ける。

【0007】

以上のように、マイク素子間の振幅差や位相差は、マイクが配置される筐体の影響によって大きく変化するため、各方向の音源と各マイク素子間の伝達関数であるアレイ・マニフォールド・ベクトルも、筐体の影響によって変化する。ビームパターンの算出に用いられるアレイ・マニフォールド・ベクトルがそのような変化に追従していないと、音源方向の正確な推定ができなくなる。特許文献1では、独立成分分析を利用して筐体の影響を含むアレイ・マニフォールド・ベクトルを推定することが示されている。また、特許文献2では、折り畳み式携帯電話などで筐体可動部にマイクが付けられた場合に、筐体可動部の開閉状態に応じて変化するマイク位置座標を逐次取得し、音源分離処理のパラメータとして用いることが記載されている。

10

【先行技術文献】

【特許文献】

【0008】

【特許文献1】特開2011-199474号公報

【特許文献2】特開2010-278918号公報

【発明の概要】

【発明が解決しようとする課題】

20

【0009】

しかしながら、特許文献1、2に記載された方法では、依然として音源推定の精度を維持できない場合がある。特許文献1の方法は、例えばカムコードの内蔵マイクを使う場合は、カムコードの筐体の影響を含むアレイ・マニフォールド・ベクトルを推定して用いることができると考えられる。しかし、音響信号の取得に用いるマイクを内蔵マイクから外部マイクに切り替えるような場合、外部マイクはカムコードから離れることで、カムコードの筐体の影響を受けにくくなる。すなわち、内蔵マイクと外部マイクとではアレイ・マニフォールド・ベクトルは大きく変化する。特許文献1では、このようなマイクが切り替わる場合を想定したアレイ・マニフォールド・ベクトルの選択について何等考慮されていない。

30

【0010】

また、特許文献2の方法は、マイク位置座標を音源分離処理のパラメータとしているため、自由空間を想定していると考えられる。しかし、実際のカムコードなどの音響処理においては、音響処理に用いるアレイ・マニフォールド・ベクトルが筐体による回折等の影響を受ける。さらに、たとえマイク位置座標が変化しなかったとしても、例えばカムコードのレンズ交換やズームによって筐体の形状が変化すると、それに伴ってアレイ・マニフォールド・ベクトルも変化すると考えられる。しかしながら、特許文献2では、このような筐体の形状変化による回折等への影響を考慮した、アレイ・マニフォールド・ベクトルの選択については考えられていない。

【0011】

40

本発明は上述した問題を解決するためになされたものであり、高精度な音響処理を実現する信号処理装置および信号処理方法を提供することを目的とする。

【課題を解決するための手段】

【0012】

上記目的を達成するための本発明の一態様による信号処理装置は以下の構成を有する。すなわち、

形状が可変である所定の機器を装着可能な信号処理装置であって、

前記信号処理装置が有するマイクから複数チャネルの音響信号を取得する音取得手段と

、

前記信号処理装置に装着された前記機器の形状の変化に伴って変化するパラメータを取

50

得するパラメータ取得手段と、

方向に応じた音の伝搬特性を表す情報であって前記機器の形状に応じた伝搬特性を表す情報を、前記パラメータ取得手段により取得されたパラメータに基づいて取得する情報取得手段と、

前記音取得手段により取得された音響信号を、前記情報取得手段により取得された情報を用いて処理する処理手段と、を有する。

【発明の効果】

【0013】

本発明によれば、機器の状態の変化に応じてアレイ・マニフォールド・ベクトルが選択されるので、高精度な音響処理を実現することができる。

10

【図面の簡単な説明】

【0014】

【図1】実施形態に係る信号処理装置の構成例を示すブロック図。

【図2】筐体のアレイ・マニフォールド・ベクトルに及ぼす影響を説明する図。

【図3】アレイ・マニフォールド・ベクトルの選択がビームパターンに及ぼす影響を説明する図。

【図4】音源方向の推定精度が雑音除去の性能に及ぼす影響を説明する図。

【図5】実施形態に係る音響処理を説明するフローチャート。

【図6】実施形態に係る平均ビームパターン算出処理を示すフローチャート。

【図7A】実施形態に係る外部マイク間隔推定処理を説明する図。

20

【図7B】実施形態に係る外部マイク間隔推定処理を説明する図。

【図8】実施形態に係る外部マイク間隔推定処理のフローチャート。

【図9】実施形態に係る代用アレイ・マニフォールド・ベクトル選択処理のフローチャート。

【発明を実施するための形態】

【0015】

以下、添付の図面を参照して、本発明の好適な実施形態の一例を詳細に説明する。なお、以下の実施形態において示される構成等は一例に過ぎず、本発明は実施形態において記載され、図示された構成に限定されるものではない。なお、図面において各方向の音源と各マイク素子間の伝達関数を示すアレイ・マニフォールド・ベクトル (Array Manifold Vector) はAMVと略記している。

30

【0016】

まず、筐体の影響によるマイク素子間の位相差の理論値からの乖離について図2を参照して説明する。図2(a)の細線は、2つのマイク素子を内蔵マイクとして持つカムコーダについて、無響室のトラバース装置で実測した、各音源方向に対するマイク素子間の位相差を周波数ごとに示したものである。ここで、2つの内蔵マイク素子を結ぶ線分の垂直二等分線の方に、カムコーダの撮影方向である正面0°がある。また、周波数については187.5Hzから187.5Hzおきに1875Hzまで表示しており、周波数が高いほど位相差も大きくなる傾向がある。一方、図2(a)の滑らかな太線は、上記内蔵マイクの間隔をパラメータとして用いた、周波数ごとの自由空間の理論値を表している。各周波数において、2つのマイク素子を結ぶ線分の方角である±90°方向で、幾何的に位相差が最大となっている。ここで、同じ周波数における位相差の理論値と実測値を比較すると、カムコーダの筐体による回折等の影響によって、実測値は自由空間の理論値より大きくなる傾向があることが分かる。

40

【0017】

同様に図2(b)の細線は、上記カムコーダについて、各音源方向に対するマイク素子間の振幅差の実測値を、周波数ごとに示したものである。ここで、振幅差は振幅和によって正規化されており、-1から1の範囲を取るものとする。位相差と同様に、周波数が高いほど、また横方向である±90°付近で、振幅差が大きくなる傾向がある。一方、図2(b)の太線は、逆二乗則による距離減衰を考慮した自由空間の理論値であるが、数cm

50

程度のマイク間隔ではほとんど振幅差を生じないことが分かる。以上のように、マイク素子間の振幅差や位相差は、マイクが配置される筐体の影響を受けて大きく変化する。

【 0 0 1 8 】

次に、アレイ・マニフォールド・ベクトルの選択がビームパターンに及ぼす影響について具体的に説明する。図3は、適応ビームフォーマのビームパターンの算出に用いるアレイ・マニフォールド・ベクトルの選択が、ビームパターンおよび音源方向推定に及ぼす影響を示したものである。ここで、ビームパターンは周波数ごとに得られるが、図3の細線はその一部として、750Hzから750Hzおきに7500Hzまでのビームパターンを表示している。また、図3の太線は、各周波数のビームパターンを平均化した、平均ビームパターンを表示したものである。

10

【 0 0 1 9 】

図3(a)は、-30°方向に音源を配置し、自由空間に配置したマイクで音響信号を取得して適応ビームフォーマのフィルタ係数を算出し、そのビームパターンを算出・表示したものである。ここで、アレイ・マニフォールド・ベクトルは、マイク間隔をパラメータとする自由空間の理論式により生成したものをを用いている。これにより、マイクの自由空間配置という音響信号取得時の状態に対応するアレイ・マニフォールド・ベクトルを選択して用いていることと等価になる。この結果、図3(a)の太線のように、音源方向である-30°方向にヌルが形成された平均ビームパターンが得られ、図3(a)の縦点線で示される平均ビームパターンのヌル方向から音源方向を正確に知ることができる。なお、-90°から0°を通過して90°までのビームパターンと、-90°から±180°を

20

【 0 0 2 0 】

一方、図3(b)および図3(c)は、-40°方向に音源を配置し、カムコーダの内蔵マイクで音響信号を取得して適応ビームフォーマのフィルタ係数を算出し、そのビームパターンを算出・表示したものである。図3(b)では、この内蔵マイクの間隔をパラメータとして用いて自由空間の理論式で生成したアレイ・マニフォールド・ベクトルを用いている。この状況は、カムコーダの筐体の影響を受けている音響信号取得時の状態とは異なるアレイ・マニフォールド・ベクトルを選択して用いていることになる。この結果、たとえば図3(b)の太線のように、平均ビームパターンは-90°を中心として広く浅くへこんでいるだけとなり、適切にヌルが形成されているとは言い難い。このため、平均ビームパターンのヌル方向から正確に音源方向を推定することはできない。

30

【 0 0 2 1 】

図3(c)では、各方向の音源とカムコーダの内蔵マイク間の伝達関数として、無響室で実測したアレイ・マニフォールド・ベクトルを用いている。これは、カムコーダの筐体の影響を受けている音響信号取得時の状態に対応するアレイ・マニフォールド・ベクトルを選択して用いていることを意味する。この結果、図3(c)の太線のように、音源方向である-40°方向にヌルが形成された平均ビームパターンが得られ、図3(c)の縦点線で示される平均ビームパターンのヌル方向から、音源方向を正確に知ることができる。なお、カムコーダのように筐体の形状が撮影方向に対して概ね対称形である場合は、-90°から0°を通過して90°までのビームパターンと、-90°から±180°を通過して90°までのビームパターンも概ね対称形となる。

40

【 0 0 2 2 】

以上から、ビームフォーマのビームパターン算出において、音響信号取得時の状態に対応するアレイ・マニフォールド・ベクトルを選択して用いることが、ビームパターンのヌルから音源方向を推定する上で重要となることが分かる。ここで、音響信号取得時の状態は、筐体の形状などの影響を受ける。

【 0 0 2 3 】

図4はさらに、アレイ・マニフォールド・ベクトルの選択と音源方向の推定精度が、雑音除去の性能に及ぼす影響を示した図である。たとえば、ピアノの演奏会をカムコーダで撮影をしていたときに、図4(a)で示されるような正面方向のピアノの音に加えて、図

50

4 (b) で示されるような観客の咳の音が、 - 40 ° 方向から飛び込んできたような場合を考える。この場合、カムコーダの内蔵マイクで取得される音響信号の各チャンネルは、図 4 (c) のようにピアノの音と咳の音が混合したものとなる。そこで、このような音響信号から雑音である咳の音を除去することを考える。

【 0 0 2 4 】

図 4 の太線 4 0 1 で囲まれる部分では咳の音が卓越しているため、このときの音響信号から適応ビームフォーマを構成すれば、咳の方向にヌルを自動的に形成するようなフィルタ係数が得られる。よって、このフィルタ係数で形成されるビームパターンを算出することで、そのヌル方向から咳の方向を推定できる。しかしながら、上述したように、カムコーダの内蔵マイクで音響信号を取得しているにもかかわらず、自由空間の理論式で生成したアレイ・マニフォールド・ベクトルを用いてしまうと、たとえば図 3 (b) に示したように適切にヌルが形成されない。一方、カムコーダの筐体の影響を含むアレイ・マニフォールド・ベクトルを用いれば、たとえば図 3 (c) のように平均ビームパターンのヌル方向から、咳の方向を - 40 ° と正確に推定することができる。

【 0 0 2 5 】

図 4 (d) は、図 3 (b) の縦点線で示される - 90 ° をとりあえずのヌル方向として、その方向に固定ビームフォーマでヌルを向けた結果である。しかしながら、ヌルを向けた方向 (- 90 °) が咳の方向 (- 40 °) とずれているため、咳の音を効果的に除去することができていない。一方、図 4 (e) は、図 3 (c) の縦点線で示される - 40 ° に固定ビームフォーマでヌルを向けた結果である。ヌルを向けた方向が咳の方向と一致しているため、咳の音を効果的に除去できている。

【 0 0 2 6 】

以上のように、音源方向の推定精度は、雑音除去の性能に大きく影響する。また、音源方向推定に加え、上記固定ビームフォーマのフィルタ係数の算出には、ヌルを向ける方向のアレイ・マニフォールド・ベクトルが必要となる。そのため、固定ビームフォーマのフィルタ係数の算出にもアレイ・マニフォールド・ベクトルの選択の適否が影響してくる。したがって、雑音除去のような音響処理においては、筐体の形状など、マイク素子による收音時の環境に適したアレイ・マニフォールド・ベクトルを選択することが肝要となる。以上を踏まえ、本実施形態では、雑音除去のような音響処理において、筐体の影響で大きく変化する音響信号取得時の状態に対応するアレイ・マニフォールド・ベクトルを選択して用いることが可能な信号処理装置を開示する。

【 0 0 2 7 】

図 1 は、実施形態によるビデオカメラ (カムコーダ) の構成例を示すブロック図である。信号処理装置 1 0 0 は、全構成要素の統御を行うシステム制御部 1 0 1、各種データを記憶しておく記憶部 1 0 2、信号の解析処理を行う信号解析部 1 0 3 を備える。

【 0 0 2 8 】

ビデオカメラにおいて、收音系の機能を実現する要素としては、内蔵マイク 1 1 1、音響信号入力部 1 1 2 を備える。また、任意の外部マイク 1 1 9 が、信号処理装置 1 0 0 に接続可能である。本実施形態において内蔵マイク 1 1 1 および外部マイク 1 1 9 は、2 つのマイク素子が間隔を持って配置された 2 c h ステレオマイクでそれぞれ構成されるものとする。なお、マイク素子の数はそれぞれ複数であればよく、3 個以上でも構わない。すなわち、本発明はマイク素子の数が 2 個の場合に限定されるものではない。

【 0 0 2 9 】

音響信号入力部 1 1 2 は、外部マイク 1 1 9 の接続を検出し、外部マイク 1 1 9 が接続されている場合は、内蔵マイク 1 1 1 ではなく外部マイク 1 1 9 から音響信号を入力する。また、音響信号入力部 1 1 2 は、内蔵マイク 1 1 1 または外部マイク 1 1 9 の、各マイク素子からのアナログ音響信号に増幅および A D 変換を施して、所定の音響サンプリングレートに対応する周期でデジタル音響信号である 2 c h マイク信号を生成する。

【 0 0 3 0 】

ビデオカメラにおいて、撮像系の機能を実現する要素としては、レンズ部 1 2 0、映像

10

20

30

40

50

信号入力部 124 を備える。レンズ部 120 はさらに、光学レンズ 121、レンズ制御部 122、レンズ内記憶部 123 を備える。レンズ部 120 は、光学レンズ 121 に入射する光に光電変換を施して、アナログ映像信号を生成する。映像信号入力部 124 は、レンズ部 120 からのアナログ映像信号に A/D 変換とゲイン調整を施して、所定の映像フレームレートに対応する周期でデジタル映像信号を生成する。レンズ制御部 122 は、システム制御部 101 と通信を行い、光学レンズ 121 を駆動制御したり、レンズ部 120 に係る情報をやり取りしたりする。レンズ内記憶部 123 は、レンズ部 120 に係る情報を保持している。本実施形態においてレンズ部 120 は、交換可能かつ、ズーム倍率に応じてレンズ筐体が伸縮する交換レンズで構成されるものとする。また、本実施形態のビデオカメラは、ユーザの操作を受け付けたり、ユーザに操作メニューや映像信号などを提示したりするための要素として、入出力 UI 部 131 を備える。入出力 UI 部 131 は、たとえばタッチパネルで構成される。

10

【0031】

以上のような構成を備えた本実施形態のビデオカメラ（信号処理装置 100）による音響信号の処理について、以下、詳細に説明する。はじめに、信号処理装置 100 での撮影に先立って、撮影時の音響処理に用いる種々のアレイ・マニフォールド・ベクトルの取得が行われる。

【0032】

非撮影時に外部マイク 119 が接続されると、音響信号入力部 112 により検出される。この検出は、音響信号入力部 112 からシステム制御部 101 に伝えられる。次に、システム制御部 101 からの指示によって、入出力 UI 部 131 は、外部マイク 119 のマイク素子間の間隔である外部マイク間隔の入力をユーザに促す。そして、ユーザによって例えば mm 単位で入力された数値は、外部マイク 119 の外部マイク間隔として設定され、記憶部 102 に保持される。マイク間隔が既知であれば、自由空間の理論式によりアレイ・マニフォールド・ベクトルを生成することが可能である。なお、外部マイク間隔がどのくらいかユーザが把握できていない場合は、外部マイク間隔を未設定のままとすることもできる。

20

【0033】

また、非撮影時にレンズ交換によってレンズ部 120 が装着されると、該装着がシステム制御部 101 により検出される。次に、システム制御部 101 は、レンズ部 120 のレンズ制御部 122 と通信を行い、装着されているレンズ部 120 のタイプを識別する。さらに、システム制御部 101 は、レンズ制御部 122 を介して、レンズ内記憶部 123 が保持している複数のアレイ・マニフォールド・ベクトルのうち、信号処理装置 100 用のアレイ・マニフォールド・ベクトルを取得して、記憶部 102 に保存する。信号処理装置 100 用のアレイ・マニフォールド・ベクトルとは、レンズ部 120 が信号処理装置 100 に装着された状態で、信号処理装置 100 の内蔵マイク 111 により音響信号を取得する場合のアレイ・マニフォールド・ベクトルである。なお、レンズ内記憶部 123 に複数のアレイ・マニフォールド・ベクトルが記憶されているのは、筐体形状が異なる複数種類のビデオカメラに対応するためである。

30

【0034】

また、一般に、交換レンズには、焦点距離や絞り値等が異なる様々なタイプがあり、タイプごとにレンズ筐体の形状が異なる。このため、信号処理装置 100 に装着されると、タイプごとに信号処理装置 100 の筐体形状が変化することと同義であるため、レンズ部 120 のタイプごとにアレイ・マニフォールド・ベクトルも変化すると考えられる。さらに、レンズがズームレンズである場合には、ズーム倍率によってレンズ筐体の形状が伸縮する。これは、ズーム倍率によってビデオカメラ（信号処理装置 100）の筐体形状が変化することと同義であるため、レンズ部 120 のズーム倍率によってアレイ・マニフォールド・ベクトルも変化すると考えられる。よって、レンズ部 120 がズームレンズである場合は、システム制御部 101 は、ズーム倍率ごとのアレイ・マニフォールド・ベクトルを取得し、記憶部 102 に保存する。

40

50

【0035】

こうして、レンズ部120から取得された種々のアレイ・マニフォールド・ベクトルは、交換レンズのタイプ（レンズ部120のタイプ）やズーム倍率と対応付けられて、記憶部102に保存される。なお、信号処理装置100にデフォルトで装着されているレンズ、信号処理装置100に装着され得る代表的な交換レンズ、レンズが装着されていない状態などに対応するアレイ・マニフォールド・ベクトルは、記憶部102が予め保持するようにしてもよい。

【0036】

なお、信号処理装置100の筐体の影響を含むアレイ・マニフォールド・ベクトルは、レンズ部120のタイプやズーム倍率ごとに、無響室のトラバース装置などで、内蔵マイク111を用いて実測できる。またはCADデータをもとに、有限要素法や境界要素法といった、波動性を考慮したシミュレーションにより、アレイ・マニフォールド・ベクトルを生成してもよい。

10

【0037】

なお、方向別伝達関数であるアレイ・マニフォールド・ベクトルは、周波数領域のデータであるが、レンズ部120のレンズ内記憶部123では、もとなる方向別インパルス応答の形で保持しておいてもよい。そして、方向別インパルス応答を記憶部102に取り込んだ際に、信号処理装置100で行う音響処理の周波数解像度に合わせて、信号解析部103でフーリエ変換し、得られたアレイ・マニフォールド・ベクトルを記憶部102に保存するようにしてもよい。

20

【0038】

次に、信号処理装置100による撮影動作について説明する。入出力UI部131の画面上には、撮像系が捉えている映像信号がリアルタイムに映し出されている。このとき、ズーム倍率を示す画面上のスライダバーのつまみを動かすことで、その指定値がシステム制御部101に伝えられる。そして、システム制御部101からの指示によって、レンズ制御部122は光学レンズ121を駆動制御し、指定ズーム倍率に応じた光学ズーム処理を行う。

【0039】

ユーザは、撮影を開始したいシチュエーションになったら、入出力UI部131上に表示されているメニューから、「REC」をタッチして選択する。この選択に応じて、信号処理装置100は、撮像系で捉えている映像信号および、收音系で捉えている音響信号の記憶部102への記録を開始する。收音系で取得した音響信号である2chマイク信号は、記憶部102へ逐次記録され、図5のフローチャートに沿って、本実施形態の音響処理である音源方向推定処理および雑音除去処理が行われる。なお、音響サンプリングレートは48kHzとして説明を行う。

30

【0040】

ビームフォーマにおいてマイク信号のフィルタリングを行う信号サンプル単位を時間ブロックと呼び、本実施形態では時間ブロックの長さを1024サンプル（約21ms）とする。また、この時間ブロック長の半分である512サンプル（約11ms）ずつ信号サンプル範囲をシフトしながら、時間ブロックループの中でマイク信号のフィルタリングを行っていく。すなわち、第1時間ブロックではマイク信号の第1サンプルから第1024サンプルを、第2時間ブロックでは第513サンプルから第1536サンプルをフィルタリングする。図5のフローチャートは、時間ブロックループ内のひとつの時間ブロックにおける処理を表すものとする。

40

【0041】

S501では、システム制御部101が音響信号入力部112と通信を行い（結線不図示）、外部マイク119が接続されているかを調べる。外部マイク119が接続されている、すなわち外部マイク119で音響信号を取得している場合は、S502へ進む。S502では、システム制御部101が外部マイク119の外部マイク間隔が設定されているかを調べ、外部マイク間隔が設定されている場合はS503へ進む。

50

【 0 0 4 2 】

S 5 0 3では、信号解析部 1 0 3が、設定されている外部マイク間隔をパラメータとしてアレイ・マニフォールド・ベクトルを生成する。生成されたアレイ・マニフォールド・ベクトルは、外部マイク 1 1 9で現在取得している時間ブロックの音響信号の処理に用いるためのアレイ・マニフォールド・ベクトルとして選択される。外部マイク 1 1 9は信号処理装置 1 0 0から離れているため、信号処理装置 1 0 0の筐体の影響を受けにくいと考えられる。よって、以下の式(1)で表される自由空間の理論式と外部マイク間隔によりアレイ・マニフォールド・ベクトル $a(f, \theta)$ を生成し、これを後の音響処理のために選択する。

【数 1】

10

$$a(f, \theta) = \exp(-j2f\tau(\theta, d)) \quad (1)$$

【 0 0 4 3 】

ここで、 j は虚数単位を、 f は周波数を表す。また、外部マイク 1 1 9の、2つのマイク素子の真ん中を中心とする単位球面を考える。すると、単位球面上の方位角 θ の点から、各マイク素子までの伝搬遅延時間は、 $\tau_i(\theta, d)$ ($i=1,2$)のように方位角 θ および外部マイク間隔 d の関数となり、まとめてベクトル $\tau(\theta)=[\tau_1(\theta) \quad \tau_2(\theta)]^T$ と置いている。ここで上付きの T は転置を表す。なお、2つの外部マイク素子を結ぶ線分の垂直二等分線の方に信号処理装置 1 0 0の撮影方向である正面($\theta = 0^\circ$)があるものとする。

20

【 0 0 4 4 】

一方、S 5 0 1で外部マイク 1 1 9が接続されていない、すなわち内蔵マイク 1 1 1で音響信号を取得している場合は、S 5 0 4へ進む。S 5 0 4では、システム制御部 1 0 1がレンズ部 1 2 0のレンズ制御部 1 2 2と通信を行い、レンズ部 1 2 0のタイプおよび、現在のズーム倍率である現ズーム倍率を取得する。S 5 0 5では、システム制御部 1 0 1が、S 5 0 4で取得したレンズ部 1 2 0のタイプに対応するアレイ・マニフォールド・ベクトルを、記憶部 1 0 2が保持しているか調べ、保持している場合、処理はS 5 0 6へ進む。

【 0 0 4 5 】

S 5 0 6では、信号解析部 1 0 3が、内蔵マイク 1 1 1で取得している現時間ブロックの音響信号の処理に用いるための、アレイ・マニフォールド・ベクトルを選択する。すなわち、S 5 0 4で取得したレンズ部 1 2 0のタイプおよび現ズーム倍率に対応する、アレイ・マニフォールド・ベクトルを $a(f, \theta)$ 選択する。ここでも、2つの内蔵マイク素子を結ぶ線分の垂直二等分線の方に、信号処理装置 1 0 0の撮影方向である正面($\theta = 0^\circ$)があるものとする。

30

【 0 0 4 6 】

なお、ズーム倍率については、現ズーム倍率と完全に一致するアレイ・マニフォールド・ベクトルがあるとは限らない。したがって、本実施形態では、現ズーム倍率と最も近いズーム倍率のアレイ・マニフォールド・ベクトルを選択することとする。もしくは、複数のズーム倍率(例えば2倍と3倍)のアレイ・マニフォールド・ベクトルを振幅および位相上で補間することで、現ズーム倍率(例えば2.5倍)に対応するアレイ・マニフォールド・ベクトルを生成して選択するようにしてもよい。なお、レンズが交換中で、レンズ部 1 2 0が信号処理装置 1 0 0に装着されていないような場合は、レンズが装着されていない状態のアレイ・マニフォールド・ベクトルを選択するようにしてもよい。

40

【 0 0 4 7 】

以上のようにしてS 5 0 3またはS 5 0 6の処理を終えると、処理はS 5 0 7へ進む。S 5 0 7以降の処理は、信号解析部 1 0 3が中心となって行う。S 5 0 7において、信号解析部 1 0 3は、平均ビームパターン算出処理を行う。平均ビームパターン算出処理について、図6のフローチャートを参照して詳細に説明する。

【 0 0 4 8 】

50

S 6 0 1において、信号解析部 1 0 3 は、現在の時間ブロックの 2 c h マイク信号をフーリエ変換して、複素数であるフーリエ係数を取得する。このとき時間ブロック長によって、フーリエ変換における時間解像度および周波数解像度が決まる。次のステップである S 6 0 2 において空間相関行列が算出されるが、統計量である空間相関行列の算出には平均化処理が必要なため、現在の時間ブロックを基準として時間フレームという単位が導入される。

【 0 0 4 9 】

時間フレーム長は時間ブロック長と同じ 1 0 2 4 サンプルであり、現在の時間ブロックの信号サンプル範囲を基準として、所定の時間フレームシフト長ずつシフトした信号サンプル範囲を時間フレームとする。本実施形態では時間フレームシフト長を 3 2 サンプルとし、上記平均化の回数に相当する時間フレーム数を 1 2 8 とする。すなわち第 1 時間ブロックにおいて、第 1 時間フレームは第 1 時間ブロックと同じくマイク信号の第 1 サンプルから第 1 0 2 4 サンプルを対象とし、第 2 時間フレームは第 3 3 サンプルから第 1 0 5 6 サンプルを対象とする。こうして、第 1 2 8 時間フレームは第 4 0 6 5 サンプルから第 5 0 8 8 サンプルを対象とするため、第 1 時間ブロックの空間相関行列は、第 1 サンプルから第 5 0 8 8 サンプルの 1 0 6 m s のマイク信号から算出されることになる。なお、時間フレームは現在の時間ブロックより前の信号サンプル範囲としてもよい。

【 0 0 5 0 】

以上を踏まえて S 6 0 1 では、第 i チャンネルのマイク信号の現時間ブロックに関する、周波数 f 、時間フレーム k におけるフーリエ係数を、 $Z_i(f, k)$ ($i=1, 2, k=1 \sim 128$) のように得る。なお、フーリエ変換の前にマイク信号に対して窓掛けを行うのが好適であり、窓掛けは逆フーリエ変換によって再び時間信号に戻した後にも行う。このため、50% ずつオーバーラップする時間ブロックに対し、2 回の窓掛けにおける再構成条件を考慮して、窓関数にはサイン窓などを用いる。

【 0 0 5 1 】

S 6 0 2 から S 6 0 4 は周波数ごとの処理であり、周波数ループの中で行われる。S 6 0 2 では、信号解析部 1 0 3 は、マイク信号の空間的性質を表す統計量である、空間相関行列を算出する。まず、S 6 0 1 で得た各チャンネルのフーリエ係数をまとめてベクトル化し、 $z(f, k) = [Z_1(f, k) \ Z_2(f, k)]^T$ のように置く。 $z(f, k)$ を用いて、周波数 f 、時間フレーム k における行列 $R_k(f)$ を式 (2) のように定める。ここで、上付きの H は複素共役転置を表す。

【 数 2 】

$$R_k(f) = z(f, k) z^H(f, k) \quad (2)$$

なお、空間相関行列 $R(f)$ は、 $R_k(f)$ を全ての時間フレームに関して平均化、すなわち $R_1(f)$ から $R_{128}(f)$ を足して 1 2 8 で割ることで得られる。

【 0 0 5 2 】

S 6 0 3 において、信号解析部 1 0 3 は、適応ビームフォーマのフィルタ係数を算出する。第 i チャンネルのマイク信号をフィルタリングするフィルタ係数を $W_i(f)$ ($i=1, 2$) とし、ビームフォーマのフィルタ係数ベクトルを $w(f) = [W_1(f) \ W_2(f)]^T$ のように置く。ここでは、信号解析部 1 0 3 は、適応ビームフォーマのフィルタ係数を最小ノルム法で算出する。これは、出力パワー最小化の規範に基づくものであり、 $w(f)$ を非零ベクトルとするための制約条件を、フィルタ係数ノルムの指定によって記述する。ビームフォーマの周波数 f における平均出力パワーは $w^H(f) R(f) w(f)$ で表されるため、最小ノルム法による適応ビームフォーマのフィルタ係数は、式 (3) の制約付き最適化問題の解として得られる。

【数 3】

$$\begin{aligned} & \min_{\mathbf{w}} \mathbf{w}^H(f) \mathbf{R}(f) \mathbf{w}(f) \\ & \text{subject to } \mathbf{w}^H(f) \mathbf{w}(f) = 1 \end{aligned} \quad (3)$$

これは、エルミート行列である $\mathbf{R}(f)$ を係数行列とする二次形式の最小化問題であるため、 $\mathbf{R}(f)$ の最小固有値に対応する固有ベクトルが、最小ノルム法で算出される適応ビームフォーマのフィルタ係数ベクトル $\mathbf{w}_{MN}(f)$ となる。

【0053】

S604では、信号解析部103は、S603で算出した適応ビームフォーマのフィルタ係数 $\mathbf{w}_{MN}(f)$ と、現時間ブロックで選択されているアレイ・マニフォールド・ベクトル $\mathbf{a}(f, \theta)$ を用いて、適応ビームフォーマのビームパターンを算出する。ビームパターンの方位角 θ 方向の値 $\Psi(f, \theta)$ は、式(4)で得られる。

【0054】

【数 4】

$$\Psi(f, \theta) = \mathbf{w}_{MN}^H(f) \mathbf{a}(f, \theta) \quad (4)$$

【0055】

$\mathbf{a}(f, \theta)$ の θ を、例えば -180° から 180° まで 1° 刻みで変えながら $\Psi(f, \theta)$ を計算することで、水平方向のビームパターンが得られる。なお、計算量を抑えるために、ビームパターンの対称性に着目して、 -90° から 0° を通して 90° までのビームパターンのみを算出するようにしてもよい。また、 θ が小さくなるヌル付近のみの刻みを密にして、音源方向を知る上で大切なヌル付近をより正確に把握するようにしてもよい。

【0056】

S605では、S604で算出した周波数ごとのビームパターンを平均化して、平均ビームパターンを算出する。なお、必ずしも全周波数について平均化する必要はなく、例えば目的音の主要周波数帯の周波数のみ平均化してもよい。以上でS507の平均ビームパターン算出処理を終える。

【0057】

他方、S502で、外部マイク119の外部マイク間隔が未設定であった場合は、処理はS520へ進む。S520では、外部マイク間隔推定処理を行う。はじめに、外部マイク間隔推定処理の考え方について説明する。

【0058】

外部マイク間隔 d を少しずつ大きくしながら、式(1)で表される自由空間の理論式によりアレイ・マニフォールド・ベクトルを生成し、平均ビームパターン算出処理を行う。図7Aに示すように、外部マイク間隔を(a)の5mm、(b)の10mm、(c)の15mmと大きくしていくに従って、横点線で示される平均ビームパターンの最小値が小さくなり、縦点線で示されるヌル方向も変わって行くことが分かる。これらの関係をグラフにしたのが、図7Bの(d)および(e)である。この場合、正解の外部マイク間隔は15mm、音源方向は -30° であるが、図7Bの(d)より d が概ね正解に達したときに、平均ビームパターンの最小値が底を打って収束する傾向があることが分かる。このような収束時の d を外部マイク間隔として設定することで、図7Bの(e)のようにそのときのヌル方向から、音源方向を正確に推定することができる。

【0059】

上記の考え方に基づく、外部マイク間隔推定処理(S520)について、図8のフローチャートを参照して説明する。S801では、信号解析部103は、外部マイク間隔を d_0 (例えば1mm)のように初期化する。S802において、信号解析部103は、現在の d をパラメータとして用い、式(1)で表される自由空間の理論式によりアレイ・マニフォールド・ベクトルを生成して、次のステップで用いるものとして選択する。

【 0 0 6 0 】

S 8 0 3では、S 8 0 2で選択したアレイ・マニフォールド・ベクトルを用いて、平均ビームパターン算出処理を行う。平均ビームパターン算出処理は図6のフローチャートにより説明したとおりである。なお、図6のフローチャートで表されるS 8 0 3の平均ビームパターン算出処理において、S 6 0 1からS 6 0 3までの処理は、最初の $d=d_0$ のときのみ行えばよい。また、外部マイク間隔推定処理（S 5 2 0）が実行された場合は、当該処理内で平均ビームパターン算出処理が行われるのでS 5 0 7の平均ビームパターン算出処理を省略可能である。

【 0 0 6 1 】

S 8 0 4では、S 8 0 3で算出した平均ビームパターンの最小値が収束したかを判定し、収束していない場合はS 8 0 5へ進む。たとえば、信号解析部103は、S 8 0 3において、 $d=t-n \sim d=t$ （ t は現在の d の値、 $n-1$ ）の、 $n+1$ 個の平均ビームパターンの最小値が所定範囲内に収まっている場合、平均ビームパターンの最小値が収束したと判定する。収束していないと判定された場合、処理はS 8 0 5へ進み、信号解析部103は、 $d=d+1$ のように外部マイク間隔をインクリメントし、処理をS 8 0 2へ戻す。S 8 0 4で、平均ビームパターンの最小値が収束したと判定された場合は、処理はS 8 0 6へ進む。S 8 0 6では、信号解析部103は、平均ビームパターンの最小値が収束したときの d を、外部マイク119の外部マイク間隔として設定する。

【 0 0 6 2 】

図5に戻り、S 5 0 5で、レンズ部120のタイプに対応するアレイ・マニフォールド・ベクトルを保持していない場合は、処理はS 5 3 0へ進む。S 5 3 0では、信号解析部103は、代用のアレイ・マニフォールド・ベクトルを選択する処理を行う。レンズ筐体の形状がレンズ部120と全く異なるようなレンズに対応するアレイ・マニフォールド・ベクトルを用いると、図3（b）のようになり、平均ビームパターンのヌルが適切に形成されず、ヌルは浅く、広がってしまう。一方、レンズ筐体の形状が、レンズ部120と比較的近いレンズに対応するアレイ・マニフォールド・ベクトルを用いれば、図3（c）のように平均ビームパターンのヌルが深くなると考えられる。そこで、レンズ部120のタイプに対応するアレイ・マニフォールド・ベクトルを保持していない場合は、その代わりとして用いるアレイ・マニフォールド・ベクトルを、平均ビームパターンのヌルの深さから決定する。

【 0 0 6 3 】

上記の考え方に基づく、代用アレイ・マニフォールド・ベクトルの選択処理（S 5 3 0）について、図9のフローチャートを参照して説明する。S 9 0 1からS 9 0 3は、記憶部102が保持しているアレイ・マニフォールド・ベクトルごとの処理であり、アレイ・マニフォールド・ベクトルループの中で行う。

【 0 0 6 4 】

S 9 0 1では、処理ループ（AMVループ）で対象とするアレイ・マニフォールド・ベクトルを選択する。S 9 0 2では、信号解析部103は、S 9 0 1で選択したアレイ・マニフォールド・ベクトルを用いて、平均ビームパターン算出処理を行う。平均ビームパターン算出処理は図6のフローチャートで説明したとおりである。但し、S 9 0 2で実行される平均ビームパターン算出処理において、S 6 0 1からS 6 0 3までは、AMVループの最初の一回のみ行えばよい。また、代用アレイ・マニフォールド・ベクトルの選択処理（S 5 3 0）が実行された場合は、当該処理内で平均ビームパターン算出処理が行われるので、S 5 0 7の平均ビームパターン算出処理を省略可能である。

【 0 0 6 5 】

S 9 0 3において、信号解析部103は、S 9 0 2で算出した平均ビームパターンのヌルの深さを算出する。ヌルの深さは、図3（c）の双方向矢印で示すように、平均ビームパターンの最大値と最小値の差とすればよいが、より簡単には最小値のみで考えてもよい。S 9 0 4では、信号解析部103は、S 9 0 3で算出したヌルの深さをもとに、ヌルが最も深かったときのアレイ・マニフォールド・ベクトルを、代用アレイ・マニフォールド

10

20

30

40

50

・ベクトルとして選択する。

【 0 0 6 6 】

なお、必ずしも記憶部 1 0 2 が保持する全てのアレイ・マニフォールド・ベクトルを図 9 に示される A M V ループの対象としなくてもよい。例えば、焦点距離がレンズ部 1 2 0 に近く、レンズ筐体の形状もレンズ部 1 2 0 に近いと考えられるレンズに対応するアレイ・マニフォールド・ベクトルのみを A M V ループの処理対象としてもよい。またズーム倍率についても、現ズーム倍率に近いズーム倍率に対応するアレイ・マニフォールド・ベクトルのみを A M V ループの対象としてもよい。さらに、ヌルの深さが所定値（例えば 1 0 d B ）以上となった時点で A M V ループを抜け、そのときのアレイ・マニフォールド・ベクトルを、代用アレイ・マニフォールド・ベクトルとして選択するようにしてもよい。

10

【 0 0 6 7 】

図 5 のフローチャートに戻り、S 5 0 8 において、信号解析部 1 0 3 は、平均ビームパターン算出処理（S 5 0 7、S 8 0 3、S 9 0 2）で算出した平均ビームパターンのヌル方向から、音源方向を推定する。すなわち、平均ビームパターンが極小値を取る点、より簡単には最小値を取る点から、ヌル方向 θ_{null} を決定し、推定音源方向とする。

【 0 0 6 8 】

本実施形態においては、内蔵マイク 1 1 1 と外部マイク 1 1 9 の切り替えによる筐体の影響の変化や、レンズ部 1 2 0 のタイプやズーム倍率による筐体の形状変化に応じて、時間ブロックごとに適切なアレイ・マニフォールド・ベクトルが選択される。このため、図 3 (a) や図 3 (c) のように、筐体の形状変化等の影響の有無にかかわらず平均ビームパターンのヌルが適切に音源方向に形成され、高精度な音源方向推定ができるようになる。

20

【 0 0 6 9 】

なお、上記の音源方向推定においては、式 (3) を用いることにより適応ビームフォーマのフィルタ係数は最小ノルム法で算出したがこれに限られるものではなく、たとえば、最小分散法 (Capon 法) などを用いてもよい。最小分散法も、最小ノルム法と同じく出力パワー最小化の規範に基づくものであるが、フィルタ係数ベクトルを非零ベクトルとするための制約条件として、メインローブの方向 θ_{main} を適当に指定する。最小分散法による適応ビームフォーマのフィルタ係数 $w_{MV}(f)$ は、式 (5) のように得られる。

【 数 5 】

30

$$w_{MV}(f,t) = \frac{R^{-1}(f)a(f,\theta_{main})}{a^H(f,\theta_{main})R^{-1}(f)a(f,\theta_{main})} \quad (5)$$

【 0 0 7 0 】

また、上記においては、式 (4) に示される音源方向に感度のディップ (ヌル) を形成するビームパターン (f, θ) から音源方向を推定したが、代わりに、音源方向に感度のピークを形成する空間スペクトル $P(f, \theta)$ を用いてもよい。たとえば、最小ノルム法を用いた場合の空間スペクトル $P_{MN}(f, \theta)$ は、式 (6) で得られる。

【 数 6 】

40

$$P_{MN}(f,\theta) = \frac{a^H(f,\theta)a(f,\theta)}{|w_{MN}^H(f)a(f,\theta)|^2} \quad (6)$$

【 0 0 7 1 】

また、最小ノルム法では、空間相関行列の最小固有値に対応する固有ベクトルを用いた。さらに、雑音部分空間に属する全ての固有ベクトルを並べた行列を E_n と置き、信号部分空間に属するアレイ・マニフォールド・ベクトルとの直交性を考えれば、MUSIC法の空間スペクトル $P_{MU}(f, \theta)$ が式 (7) で得られる。

【数 7】

$$P_{MV}(f, \theta) = \frac{\mathbf{a}^H(f, \theta) \mathbf{a}(f, \theta)}{\mathbf{a}^H(f, \theta) \mathbf{E}_n \mathbf{E}_n^H \mathbf{a}(f, \theta)} \quad (7)$$

【0072】

また、最小分散法を用いた場合の空間スペクトル $P_{MV}(f, \theta)$ は、式(8)で得られる。

【数 8】

$$P_{MV}(f, \theta) = \frac{1}{\mathbf{a}^H(f, \theta) \mathbf{R}^{-1}(f) \mathbf{a}(f, \theta)} \quad (8)$$

10

【0073】

以上のように、本実施形態の音源方向推定は、アレイ・マニフォールド・ベクトルと音響信号の空間相関行列を用いて音源方向に感度の極値を持つビームパターンや空間スペクトルといった感度曲線を算出し、感度曲線の極値点から音源方向を推定するものである。

【0074】

図5に戻り、S509では、信号解析部103は、S508で推定した推定音源方向が、目的音の範囲外であるかを調べる。推定音源方向が目的音の範囲外である場合は、現時間ブロックにおいて推定音源方向にある雑音が卓越していると見なされ、処理はS510、S511の雑音除去処理へと進む。他方、S509で推定音源方向が目的音の範囲外でない、すなわち目的音の範囲内である場合は、現時間ブロックにおいて推定音源方向にある目的音が卓越していると見なされ、処理はS510、S511の雑音除去処理をスキップしてS512へ進む。なお、目的音の範囲は、例えば信号処理装置100の撮影方向である正面 $\pm 30^\circ$ のように定めておいてもよいし、現ズーム倍率に応じて変化する撮像系の画角範囲としてもよい。また、入出力UI部131を介してユーザが設定するようにしてもよい。以下、雑音除去処理について説明する。

20

【0075】

S510、S511の処理は周波数ごとの処理であり、周波数ループの中で行われる。S510において、信号解析部103は、S508で推定した推定音源方向 θ_{null} に鋭いヌルを形成するための固定ビームフォーマのフィルタ係数 $w_{fix}(f)$ を算出する。固定ビームフォーマのビームパターンにおいて、推定音源方向 θ_{null} にヌルを形成する条件は、アレイ・マニフォールド・ベクトル $\mathbf{a}(f, \theta_{null})$ を用いて式(9)のように表される。

30

【数 9】

$$\mathbf{w}_{fix}^H(f) \mathbf{a}(f, \theta_{null}) = 0 \quad (9)$$

【0076】

ただし、式(9)だけでは解が零ベクトルになってしまうため、メインローブ方向 θ_{main} にメインローブを形成する条件として式(10)を加える。ここで、メインローブ方向 θ_{main} は、目的音範囲の中心である正面 0° などに定める。

40

【数 10】

$$\mathbf{w}_{fix}^H(f) \mathbf{a}(f, \theta_{main}) = 1 \quad (10)$$

【0077】

式(9)および式(10)をまとめて行列 $\mathbf{A}(f) = [\mathbf{a}(f, \theta_{null}) \mathbf{a}(f, \theta_{main})]$ を用いて表現すれば、式(11)のようになる。

【数 1 1】

$$A^H(f)w_{fix}(f) = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad (11)$$

【0 0 7 8】

よって、式 (1 1) の両辺に左から $A^H(f)$ の逆行列を掛けることで、固定ビームフォーマのフィルタ係数 $w_{fix}(f)$ が式 (1 2) のように得られる。

【数 1 2】

$$w_{fix}(f) = (A^H(f))^{-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad (12)$$

10

【0 0 7 9】

ここで、 $w_{fix}(f)$ のノルムは周波数ごとに異なるため、最小ノルム法の $w_{MN}(f)$ と同様に、ノルムが 1 となるよう正規化するのが好適である。なお、フィルタ係数ベクトル $w_{fix}(f)$ の要素数、すなわち収音系のマイク素子の数と、式 (9)、式 (1 0) のようなビームパターン上の制御点の数が異なる場合は、 $A(f)$ が正方行列ではないため一般化逆行列を用いる。

【0 0 8 0】

S 5 1 1 では、S 5 1 0 で算出した固定ビームフォーマのフィルタ係数を用いてフィルタリングを行い、雑音が除去されたマイク信号のフーリエ係数を取得する。一般に、ビームフォーマによるフィルタリングは、マイク信号に対して式 (1 3) のように行う。ここで、 $z(f)=z(f,1)$ であり、 $Y(f)$ が雑音除去信号のフーリエ係数である。

20

【数 1 3】

$$Y(f) = w_{fix}^H(f)z(f) \quad (13)$$

【0 0 8 1】

ただし、これでは雑音除去信号がモノラル信号になってしまうため、再び 2 c h マイク信号に戻すための Projection Back を行う。具体的には、行ベクトルである $w_{fix}^H(f)$ を横長行列と見なし、その一般化逆行列を式 (1 3) の両辺に左から掛けることで、式 (1 4) のように雑音除去された 2 c h マイク信号のフーリエ係数 $z_{PJ}(f)$ を取得する。

30

【0 0 8 2】

【数 1 4】

$$\begin{aligned} z_{PJ}(f) &= (w_{fix}^H(f))^+ Y(f) \\ &= w_{fix}(f) (w_{fix}^H(f)w_{fix}(f))^{-1} w_{fix}^H(f)z(f) \end{aligned} \quad (14)$$

ここで、上付きの+は一般化逆行列を表すものとする。

【0 0 8 3】

40

以上のように、本実施形態によれば、適切なアレイ・マニフォールド・ベクトルの選択によって音源方向を正確に推定することができる。そして、正確に推定された雑音方向に鋭いヌルを形成する固定ビームフォーマを用いることで、雑音が目的音に近接する場合でも、雑音のみを高精度に除去することができる。なお、本実施形態では雑音除去を説明したが、正確に推定された音源方向を音源の分離に用いることも可能である。

【0 0 8 4】

S 5 1 2 では、2 c h マイク信号のフーリエ係数を逆フーリエ変換し、現時間ブロックにおけるマイク信号を取得する。これを窓掛けして前時間ブロックまでのマイク信号にオーバーラップ加算していき、得られるマイク信号を記憶部 1 0 2 へ逐次記録する。以上のようにして得られたマイク信号は、記憶部 1 0 2 と相互に結ばれた、不図示のデータ入出

50

力部を介して外部に出力したり、イヤホン、ヘッドホン、スピーカといった不図示の音響再生系によって再生したりすることができる。

【0085】

なお、上記説明においては簡単のため、方向として方位角のみを考えていたが、仰角も考えることができる。すなわち、方位角および仰角ごとの伝達関数としてアレイ・マニフォールド・ベクトル $a(f, \theta, \phi)$ を用意し、方位角だけでなく、仰角も 0° 以外に -90° から 90° まで変えながら、ビームパターン (f, θ, ϕ) を算出する。そして、平均ビームパターンの極値点から、水平方向だけでなく垂直方向含む全方位の音源方向を推定することができる。

【0086】

さらに、方向に加えて距離 r も考えることができる。すなわち、方位角、仰角、距離 r ごとの伝達関数としてアレイ・マニフォールド・ベクトル $a(f, \theta, \phi, r)$ を用意し、方位角、仰角に加えて、距離 r も例えば 0.5 m から 5 m まで変えながら、ビームパターン (f, θ, ϕ, r) を算出する。そして、平均ビームパターンの極値点から、音源方向に加えて音源距離も推定することができる。

【0087】

なお、音響処理において、雑音除去処理には固定ビームフォーマ以外の手法を用いても構わない。例えば、マイク信号のチャンネル間の位相差を周波数ごとに求め、推定音源方向に対応する位相差範囲の場合に、抑圧するようなマスク処理を用いてもよい。この場合も、推定音源方向に対応する位相差範囲の算出にはアレイ・マニフォールド・ベクトルが必要となるため、本実施形態のアレイ・マニフォールド・ベクトル選択を適用することができる。なお、適応ビームフォーマによる音源方向推定処理を行わず、固定ビームフォーマのみによって所定方向の雑音を除去するようにしてもよい。

【0088】

また、上記において、アレイ・マニフォールド・ベクトルの取得は非撮影時に行われるものとしていたが、撮影時の音響処理における割り込み処理として動的に行われるようにしてもよい。

【0089】

また、上記説明においては、音響信号取得時である撮影時に全ての音響処理を行っているが、これに限られるものではない。たとえば、各時間ブロックで選択するアレイ・マニフォールド・ベクトルを特定できるような付加情報とともに音響信号を記録しておくことで、非撮影時に後処理として音源方向推定処理や雑音除去処理を行うこともできる。そのような付加情報としては、たとえば、外部マイク119と内蔵マイク111の切り替えを示す外部マイク接続情報、外部マイク間隔、レンズ部120のタイプやズーム倍率、またはアレイ・マニフォールド・ベクトル識別IDなどが挙げられる。

【0090】

また、外部マイク119が、レンズ部120と同様に不図示の外部マイク制御部や外部マイク内記憶部を備える場合は、システム制御部101が外部マイク制御部と通信を行うことで、外部マイク119のタイプを識別することができる。さらに、システム制御部101は、外部マイク制御部を介して外部マイク間隔を取得し、外部マイクのタイプと対応付けて記憶部102に保存することができる。

【0091】

また、外部マイクの場合は、信号処理装置100の筐体の影響を受けにくいとして、自由空間の理論式によるアレイ・マニフォールド・ベクトルを選択していた。しかし、外部マイク自体の筐体の影響によって、アレイ・マニフォールド・ベクトルが自由空間の理論値から乖離するような場合も考えられる。そこで、外部マイク筐体の影響を含むアレイ・マニフォールド・ベクトルを外部マイク内記憶部が保持しておき、システム制御部101が外部マイク制御部を介して取得して、外部マイクのタイプと対応付けて記憶部102に保存するようにしてもよい。

【0092】

10

20

30

40

50

また、種々のアレイ・マニフォールド・ベクトルの取得方法は、上記した方法に限られるものではない。たとえば、アレイ・マニフォールド・ベクトルをデータ入出力部を介して任意の外部記憶部から取得したり、ネットワーク上のデータベースから取得したりするようにしてもよい。

【0093】

また、レンズ部120が交換レンズでなく、ズーム倍率によってもレンズ筐体の形状が伸縮しないような場合は、外部マイク用と、内蔵マイク用のデフォルトのアレイ・マニフォールド・ベクトルを切り替えるだけとなる。また、外部マイクを接続しないで内蔵マイクのみを用いる場合は、レンズ部120のタイプやズーム倍率によって筐体形状が変化する場合のみ、アレイ・マニフォールド・ベクトルを切り替えることになる。これらの場合も本発明に含まれることは言うまでもない。

10

【0094】

また、内蔵マイクと外部マイクの切り替えに限らず、使用マイク素子の変更等による内蔵マイク同士の切り替えや外部マイク同士の切り替えなど、任意のマイク切り替えにおけるアレイ・マニフォールド・ベクトル切り替えにも対応可能としてよい。

【0095】

また、上記説明においてシステム制御部101は、レンズ部120による信号処理装置100の筐体の形状変化を検出する検出手段として機能し、その検出結果に応じて信号解析部103がアレイ・マニフォールド・ベクトルを選択していた。これと同様に、次のような場合も本発明の範囲に含まれる。例えば入出力UI部131を構成するタッチパネルが開閉式である場合は、その開閉状態に応じて信号処理装置100の筐体形状が変化すると見なすことができる。そこで、タッチパネルの開閉状態を検出し、それに応じてアレイ・マニフォールド・ベクトルを選択するようにしてもよい。この考え方は、例えば折り畳み式携帯電話にも用いることができる。

20

【0096】

また、信号処理装置100には、レンズ部120以外にもストロボ等の様々なアクセサリが装着され得るため、その着脱状態に応じて信号処理装置100の筐体形状が変化すると見なすことができる。そこで、任意のアクセサリの着脱状態を検出し、それに応じてアレイ・マニフォールド・ベクトルを選択するようにしてもよい。

【0097】

以上説明した本発明によれば、マイクの切り替えや筐体の形状変化に応じてアレイ・マニフォールド・ベクトルを選択することで、高精度な音響処理を実現することができる。

30

【0098】

<その他の実施形態>

以上、実施形態例を詳述したが、本発明は例えば、システム、装置、方法、プログラム若しくは記録媒体(記憶媒体)等としての実施態様をとることが可能である。具体的には、複数の機器(例えば、ホストコンピュータ、インタフェース機器、撮像装置、webアプリケーション等)から構成されるシステムに適用しても良いし、また、一つの機器からなる装置に適用しても良い。

【0099】

また、本発明の目的は、以下のようにすることによって達成されることはいうまでもない。即ち、前述した実施形態の機能を実現するソフトウェアのプログラムコード(コンピュータプログラム)を記録した記録媒体(または記憶媒体)を、システムあるいは装置に供給する。係る記憶媒体は言うまでもなく、コンピュータ読み取り可能な記憶媒体である。そして、そのシステムあるいは装置のコンピュータ(またはCPUやMPU)が記録媒体に格納されたプログラムコードを読み出し実行する。この場合、記録媒体から読み出されたプログラムコード自体が前述した実施形態の機能を実現することになり、そのプログラムコードを記録した記録媒体は本発明を構成することになる。

40

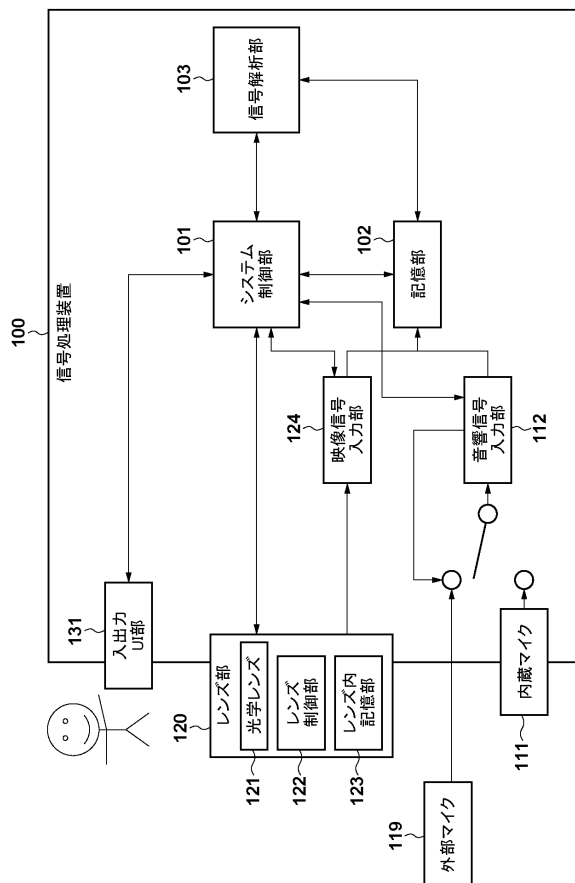
【符号の説明】

【0100】

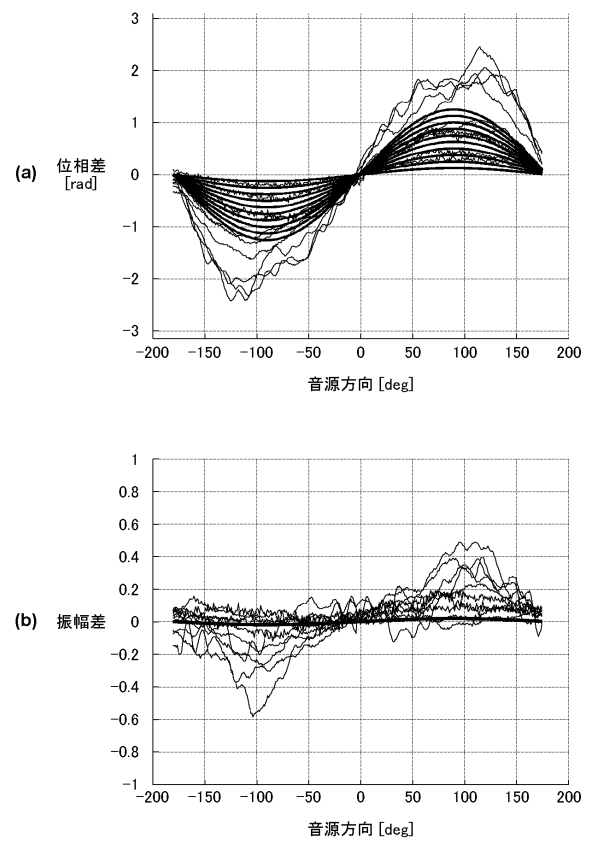
50

100：信号処理装置、101：システム制御部、102：記憶部、103：信号解析部、
111：内蔵マイク、112：音響信号入力部、119：外部マイク、120：レンズ部

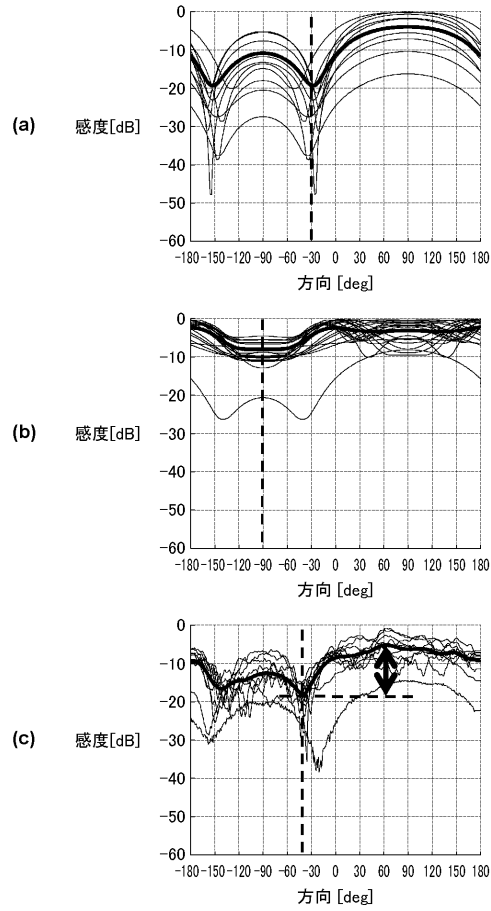
【図1】



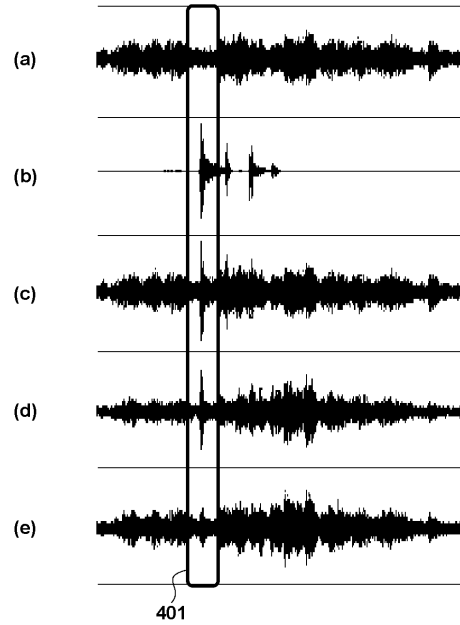
【図2】



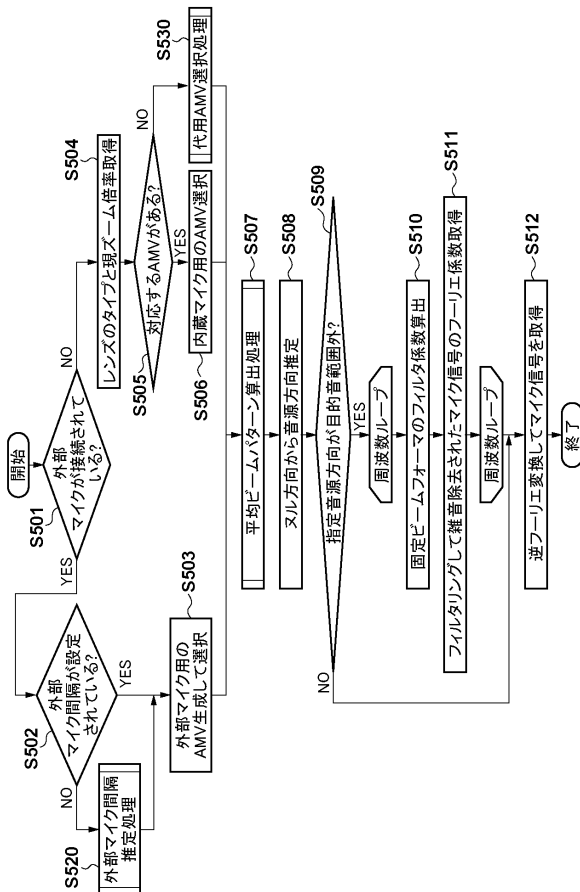
【図 3】



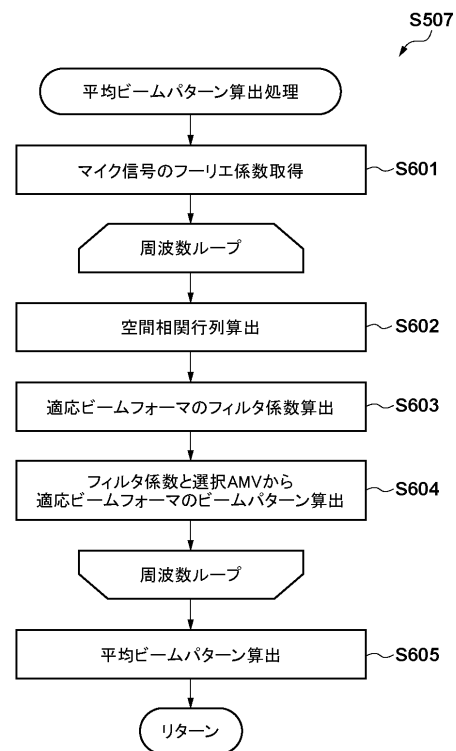
【図 4】



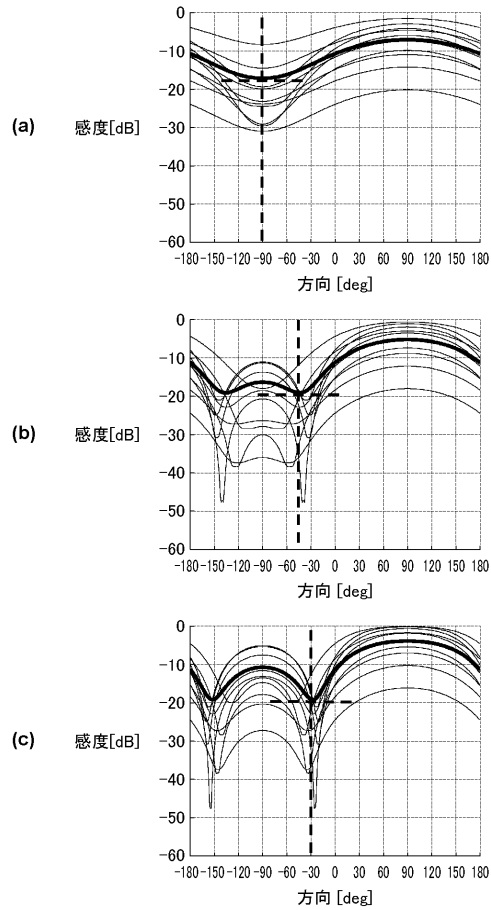
【図 5】



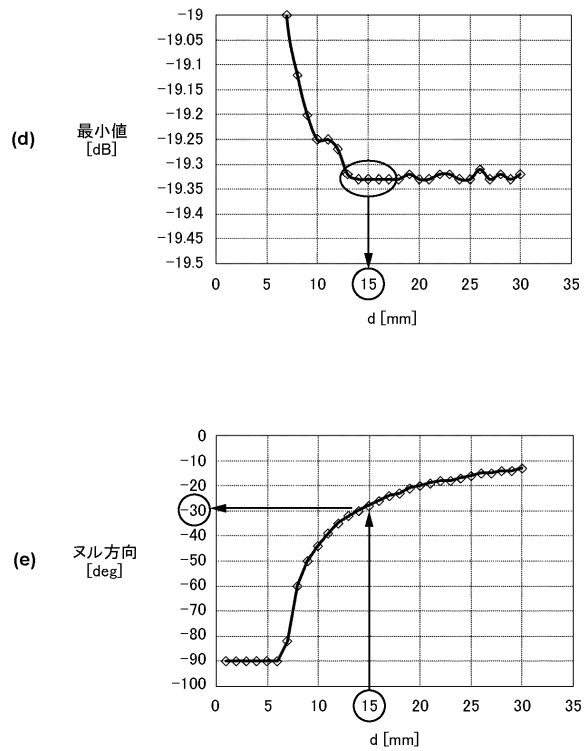
【図 6】



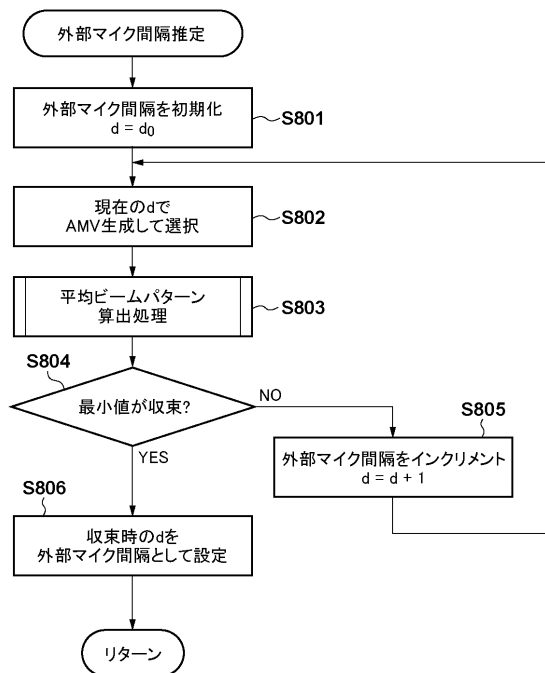
【図 7 A】



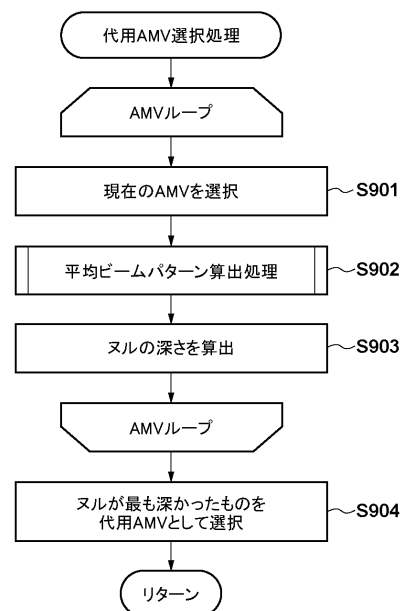
【図 7 B】



【図 8】



【図 9】



フロントページの続き

(72)発明者 多和田 典朗
東京都大田区下丸子3丁目30番2号 キヤノン株式会社内

審査官 須藤 竜也

(56)参考文献 特開2011-223491(JP,A)
特開2014-128013(JP,A)
特開2012-204900(JP,A)
特開2011-199474(JP,A)
特開2007-158731(JP,A)
特開2011-50009(JP,A)
特開2013-17160(JP,A)
特開2014-36395(JP,A)
米国特許出願公開第2013/0022217(US,A1)

(58)調査した分野(Int.Cl., DB名)

H04R	3/00
H04R	1/40
H04N	5/222 - 5/257