

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2017-151216

(P2017-151216A)

(43) 公開日 平成29年8月31日(2017.8.31)

(51) Int. Cl.	F I	テーマコード (参考)
G 1 O L 25/51 (2013.01)	G 1 O L 25/51 4 0 0	5 D 0 1 8
H O 4 R 3/00 (2006.01)	H O 4 R 3/00 3 2 0	5 D 2 2 0
H O 4 R 1/40 (2006.01)	H O 4 R 1/40 3 2 0 A	

審査請求 未請求 請求項の数 11 O L (全 14 頁)

(21) 出願番号 特願2016-32281 (P2016-32281)
 (22) 出願日 平成28年2月23日 (2016. 2. 23)

(71) 出願人 504133110
 国立大学法人電気通信大学
 東京都調布市調布ヶ丘一丁目5番地1
 (74) 代理人 100121131
 弁理士 西川 孝
 (74) 代理人 100082131
 弁理士 稲本 義雄
 (72) 発明者 田中 龍亮
 東京都調布市調布ヶ丘一丁目5番地1 国
 立大学法人電気通信大学内
 (72) 発明者 羽田 陽一
 東京都調布市調布ヶ丘一丁目5番地1 国
 立大学法人電気通信大学内
 Fターム(参考) 5D018 BB23
 5D220 BA06 BC05

(54) 【発明の名称】 音源方向推定装置、音源方向推定方法、およびプログラム

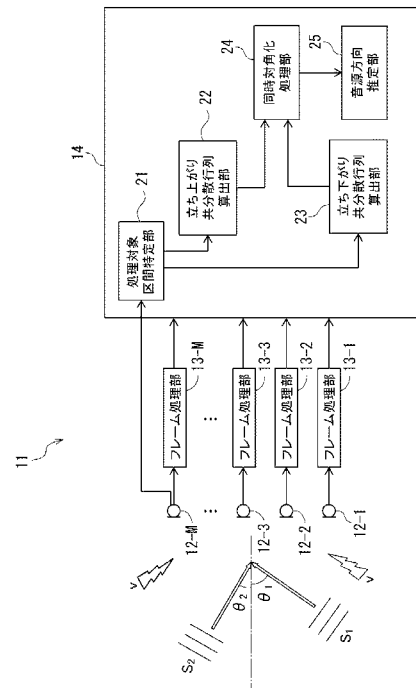
(57) 【要約】

【課題】より短時間で、反射音や残響音などの外乱が存在する環境下における音源方向を高精度に推定する。

【解決手段】複数のマイクロホンにより観測される音のレベルが所定の値よりも大きな観測信号のうち、所望の音源の方向を推定する処理に用いる対象とする観測信号を区分けする立ち上がり区間および立ち下がり区間が特定される。また、立ち上がり区間の観測信号を用いて立ち上がり共分散行列が算出されるとともに、立ち下がり区間の観測信号を用いて立ち下がり共分散行列が算出される。そして、立ち上がり共分散行列と立ち下がり共分散行列とを同時対角化する固有ベクトルが求められ、その固有ベクトルを用いて所望の音源の方向が推定される。本技術は、例えば、マイクロホンアレーにより收音される音の音源方向を推定する音源方向推定装置に適用できる。

【選択図】 図 1

Fig. 1



【特許請求の範囲】**【請求項 1】**

音を受音する複数の受音素子により観測される音のレベルが所定の値よりも大きな観測信号のうち、所望の音源の方向を推定する処理に用いる対象とする前記観測信号を区分けする時間区間を特定する処理対象区間特定部と、

前記処理対象区間特定部により特定された前記時間区間において観測された前記観測信号を用いて共分散行列を算出する共分散行列算出部と、

前記共分散行列算出部により算出された前記共分散行列に基づいて、前記所望の音源の方向を推定する音源方向推定部と

を備える音源方向推定装置。

10

【請求項 2】

前記処理対象区間特定部は、前記観測信号により観測される特定の強さの音のピークが検出されたタイミングから一定時間だけ後にある所定の前記時間区間を立ち下がり区間として特定し、

前記共分散行列算出部は、前記立ち下がり区間において観測された前記観測信号を用いて立ち下がり共分散行列を算出する

請求項 1 に記載の音源方向推定装置。

【請求項 3】

前記共分散行列算出部により算出される前記立ち下がり共分散行列と、前記観測信号の全体を用いて算出される共分散行列とを同時対角化する固有ベクトルを求める同時対角化処理部

をさらに備え、

前記音源方向推定部は、前記同時対角化処理部により求められた前記固有ベクトルを用いて前記所望の音源の方向を推定する

請求項 2 に記載の音源方向推定装置。

20

【請求項 4】

前記処理対象区間特定部は、前記観測信号により観測される特定の強さの音のピークが検出されたタイミングを含み、そのタイミングの前後にある所定の前記時間区間を立ち上がり区間として特定し、

前記共分散行列算出部は、前記立ち上がり区間において観測された前記観測信号を用いて立ち上がり共分散行列を算出する

請求項 1 に記載の音源方向推定装置。

30

【請求項 5】

前記音源方向推定部は、前記共分散行列算出部により算出される前記立ち上がり共分散行列から固有ベクトルを求める固有値分解処理を行い、その固有値分解により求められた前記固有ベクトルを用いて前記所望の音源の方向を推定する

請求項 4 に記載の音源方向推定装置。

【請求項 6】

前記処理対象区間特定部は、前記観測信号により観測される特定の強さの音のピークが検出されたタイミングを含み、そのタイミングの前後にある所定の前記時間区間を立ち上がり区間として特定するとともに、前記ピークが検出されたタイミングから一定時間だけ後にある所定の前記時間区間を立ち下がり区間として特定し、

前記共分散行列算出部は、前記立ち上がり区間において観測された前記観測信号を用いて立ち上がり共分散行列を算出するとともに、前記立ち下がり区間において観測された前記観測信号を用いて立ち下がり共分散行列を算出する

請求項 1 に記載の音源方向推定装置。

40

【請求項 7】

前記共分散行列算出部により算出される前記立ち上がり共分散行列と前記立ち下がり共分散行列とを同時対角化する固有ベクトルを求める同時対角化処理部

をさらに備え、

50

前記音源方向推定部は、前記同時対角化処理部により求められた前記固有ベクトルを用いて前記所望の音源の方向を推定する

請求項 6 に記載の音源方向推定装置。

【請求項 8】

前記処理対象区間特定部は、前記観測信号により観測される前記音のピーク値を時間方向に減衰させながら保持し、保持した値よりも大きな値が現れたときに、その値に前記ピーク値を更新するピークホールド処理を行うことにより、特定の強さの音のピークを検出する

請求項 1 乃至 7 のいずれかに記載の音源方向推定装置。

【請求項 9】

前記音源方向推定部は、前記観測信号を用いて算出される共分散行列から求められる所定の固有ベクトルのうち、前記複数の受音素子の数から、所定数の前記所望の音源の数を減じた数に対応する前記固有ベクトルに対して直交するベクトルを求めることにより、前記所望の音源の方向を推定する

請求項 1 乃至 8 のいずれかに記載の音源方向推定装置。

【請求項 10】

音を受音する複数の受音素子により観測される音のレベルが所定の値よりも大きな観測信号のうち、所望の音源の方向を推定する処理に用いる対象とする前記観測信号を区別する時間区間を特定し、

特定された前記時間区間において観測された前記観測信号を用いて共分散行列を算出し

算出された前記共分散行列に基づいて、前記所望の音源の方向を推定する

ステップを含む音源方向推定方法。

【請求項 11】

音を受音する複数の受音素子により観測される音のレベルが所定の値よりも大きな観測信号のうち、所望の音源の方向を推定する処理に用いる対象とする前記観測信号を区別する時間区間を特定し、

特定された前記時間区間において観測された前記観測信号を用いて共分散行列を算出し

算出された前記共分散行列に基づいて、前記所望の音源の方向を推定する

ステップを含む音源方向推定処理をコンピュータに実行させるプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本開示は、音源方向推定装置、音源方向推定方法、およびプログラムに関し、特に、より短時間で、反射音や残響音などの外乱が存在する環境下における音源方向を高精度に推定することができるようにした音源方向推定装置、音源方向推定方法、およびプログラムに関する。

【背景技術】

【0002】

従来、複数のマイクロホンがアレー状に配置されたマイクロホンアレーを用いて音源の方向を推定する音源方向推定方法について、様々な利用方法が提案されている。例えば、テレビ会議における話者の位置を検出することや、防犯カメラにおいて音源を追跡することなどに、音源方向推定方法を利用することができる。

【0003】

しかしながら、実際に、マイクロホンアレーを用いて音源方向を推定するとき、反射音や残響音などの外乱が存在する環境下では、それらの外乱による悪影響が発生することによって推定精度が低下してしまう。このような推定精度の低下を回避するためには、外乱の影響を排除する必要があり、例えば、ICA (Independent Component Analysis) などの音源分離手法が提案されている。

10

20

30

40

50

【 0 0 0 4 】

また、特許文献 1 には、例えば、複数のマイクロホンを搭載して構成されるマイクロホンアレーの出力信号を利用して、パワースペクトルおよび残響スペクトルを推定することにより、收音される所望信号の音質を向上させる收音装置が開示されている。

【 先行技術文献 】

【 特許文献 】

【 0 0 0 5 】

【 特許文献 1 】 特開 2 0 0 9 - 0 0 5 2 6 1 号 公 報

【 発明の概要 】

【 発明が解決しようとする課題 】

【 0 0 0 6 】

しかしながら、上述したような音源分離手法を用いる場合、音源方向を推定するまでに長時間を要することがあり、より短時間で、反射音や残響音などの外乱が存在する環境下における音源方向を高精度に推定することが求められている。

【 0 0 0 7 】

本開示は、このような状況に鑑みてなされたものであり、より短時間で、反射音や残響音などの外乱が存在する環境下における音源方向を高精度に推定することができるようにするものである。

【 課題を解決するための手段 】

【 0 0 0 8 】

本開示の一側面の音源方向推定装置は、音を受音する複数の受音素子により観測される音のレベルが所定の値よりも大きな観測信号のうち、所望の音源の方向を推定する処理に用いる対象とする前記観測信号を区別する時間区間を特定する処理対象区間特定部と、前記処理対象区間特定部により特定された前記時間区間において観測された前記観測信号を用いて共分散行列を算出する共分散行列算出部と、前記共分散行列算出部により算出された前記共分散行列に基づいて、前記所望の音源の方向を推定する音源方向推定部とを備える。

【 0 0 0 9 】

本開示の一側面の音源方向推定方法またはプログラムは、音を受音する複数の受音素子により観測される音のレベルが所定の値よりも大きな観測信号のうち、所望の音源の方向を推定する処理に用いる対象とする前記観測信号を区別する時間区間を特定し、特定された前記時間区間において観測された前記観測信号を用いて共分散行列を算出し、算出された前記共分散行列に基づいて、前記所望の音源の方向を推定するステップを含む。

【 0 0 1 0 】

本開示の一側面においては、音を受音する複数の受音素子により観測される音のレベルが所定の値よりも大きな観測信号のうち、所望の音源の方向を推定する処理に用いる対象とする観測信号を区別する時間区間が特定され、その特定された時間区間において観測された観測信号を用いて共分散行列が算出され、共分散行列に基づいて、所望の音源の方向が推定される。

【 発明の効果 】

【 0 0 1 1 】

本開示の一側面によれば、より短時間で、反射音や残響音などの外乱が存在する環境下における音源方向を高精度に推定することができる。

【 図面の簡単な説明 】

【 0 0 1 2 】

【 図 1 】 本技術を適用した音源方向推定システムの一実施の形態の構成例を示すブロック図である。

【 図 2 】 ピークホールド処理と、立ち上がり区間および立ち下がり区間について説明する図である。

【 図 3 】 シミュレーション条件を示す図である。

10

20

30

40

50

【図4】シミュレーション結果を示す図である。

【図5】音源方向推定処理を説明するフローチャートである。

【図6】本技術を適用したコンピュータの一実施の形態の構成例を示すブロック図である。

【発明を実施するための形態】

【0013】

以下、本技術を適用した具体的な実施の形態について、図面を参照しながら詳細に説明する。

【0014】

図1は、本技術を適用した音源方向推定システム11の一実施の形態の構成例を示すブロック図である。

【0015】

図1に示すように、音源方向推定システム11は、M個のマイクロホン12-1乃至12-Mからなるマイクロホンアレーにより、1カ所以上の音源S（図1の例では2カ所の音源 S_1 および S_2 ）から到達する音を観測して、音源Sの方向を推定する。音源方向推定システム11では、従来から提案されている様々な音源推定方法を利用することができるが、以下では、音源推定方法の一つであるMUSIC (Multiple Signal Classification) 法を利用する例について説明する。

【0016】

例えば、Lカ所の音源 S_1 乃至 S_L があり、M個のマイクロホン12-1乃至12-Mにより観測を行うとき、その観測により取得される音のレベルが所定の値よりも大きな観測信号zは、混合行列A、音源ベクトルs、および雑音ベクトルvを用いて、次の式(1)で表される。

【0017】

【数1】

$$z = [z_1 z_2 \cdots z_M]^T = As + v$$

$$s = [s_1 s_2 \cdots s_L]^T \quad \cdots (1)$$

$$v = [v_1 v_2 \cdots v_M]^T$$

【0018】

ここで、式(1)において、混合行列Aは、i番目($i = 1 \sim L$)の音源 S_i の音源方向 θ_i （即ち、音源 S_i からの平面波がマイクロホン12-1乃至12-Mに到来する角度）に対するステアリングベクトル $a(\theta_i)$ を列にもつ行列である。また、音源ベクトルsは、Lカ所の音源からの音に従った音源信号 s_1 乃至 s_L （音源 S_i の複素振幅 s_i ）を表すベクトルであり、雑音ベクトルvは、M個のマイクロホン12-1乃至12-Mに混入する雑音に従った雑音信号 v_1 乃至 v_M を表すベクトルである。

【0019】

また、観測信号zの全体を用いて算出される共分散行列である観測共分散行列Rは、観測信号zに含まれている音源からの音の成分からなる音源信号共分散行列、および、観測信号zに含まれている雑音の成分からなる雑音共分散行列Kを用いて、次の式(2)で表される。なお、式(2)において、 $E[\cdot]$ は期待値演算を示し、Hは共役転置を示す。

【0020】

【数2】

$$R = E[zz^H] = A\Gamma A^H + K$$

$$\Gamma = E[ss^H] \quad \cdots (2)$$

$$K = E[vv^H]$$

10

20

30

40

50

【 0 0 2 1 】

そして、MUSIC法では、混合行列 A と直交する雑音部分空間 V を求め、次の式 (3) で表される関数 $P_{\text{MUSIC}}(\theta)$ のピークを探索することによって、音源方向の推定が行われる。

【 0 0 2 2 】

【 数 3 】

$$P_{\text{MUSIC}}(\theta) = \frac{a^H(\theta)a(\theta)}{\|V^H a(\theta)\|^2} \quad \dots (3)$$

【 0 0 2 3 】

即ち、式 (3) で示される関数 $P_{\text{MUSIC}}(\theta)$ は、雑音部分空間 V とステアリングベクトル $a(\theta)$ とが直交するときにピークを持ち、雑音部分空間 V とステアリングベクトル $a(\theta)$ との直行性から、ピークに対応する方向が音源方向となる。

【 0 0 2 4 】

また、雑音部分空間 V は、マイクロホン $12 - 1$ 乃至 $12 - M$ ごとに雑音が互いに無相関であり、雑音共分散行列 K が対角行列であれば、観測共分散行列 R を固有値分解して音源部分空間と雑音部分空間 V に分離することにより求められる。ところが、雑音共分散行列 K が非対角行列のとき、真の音源方向のステアリングベクトル $a(\theta)$ と直交する雑音部分空間 V を求めることは困難となる。

【 0 0 2 5 】

一方、雑音共分散行列 K が雑音部分空間 V のみから推定可能であると仮定すると、観測共分散行列 R および雑音共分散行列 K の関係について、固有ベクトル e および音源の自己パワースペクトルを用いて、次の式 (4) が成り立つ。

【 0 0 2 6 】

【 数 4 】

$$Re = \lambda Ke \quad \dots (4)$$

【 0 0 2 7 】

そして、式 (4) に示す観測共分散行列 R および雑音共分散行列 K の一般化固有値分解を解き、観測共分散行列 R および雑音共分散行列 K を同時対角化することで、雑音共分散行列 K の性質に依存しない雑音部分空間 V を推定することができる。

【 0 0 2 8 】

しかしながら、上述したように、観測信号 z に反射音や残響音などの外乱が存在する場合、雑音共分散行列 K は非対角行列となり、音源方向の推定精度が低下することになる。

【 0 0 2 9 】

そこで、音源が音声であるとして、音声の振幅の変化に着目すると、音声の振幅は一定ではなく、振幅の小さな区間の後の振幅の大きな区間は残響による影響が小さい一方、振幅の大きな区間の後の振幅の小さな区間は残響による影響が大きいと考えられる。

【 0 0 3 0 】

このことより、音源方向推定システム 11 は、音声が発せられる発話の立ち上がり区間、および、その立ち下がり区間により、音源方向を推定する処理に用いる対象とする観測信号を区分けする時間区間を特定する手法を採用する。このような手法によって、発話の立ち上がり区間における観測信号から算出される共分散行列と、発話の立ち下がり区間における観測信号から算出される共分散行列とを同時対角化するMUSIC法により、音源方向の推定精度を向上させることができる。即ち、音源方向推定システム 11 では、観測信号のみから雑音共分散行列 K を求めることにより、より短時間で、反射音や残響音などの外乱が存在する環境下における音源方向を高精度に推定することができる。

【 0 0 3 1 】

図 1 に示すように、音源方向推定システム 11 は、 M 個のマイクロホン $12 - 1$ 乃至 $12 - M$ 、 M 個のフレーム処理部 $13 - 1$ 乃至 $13 - M$ 、および音源方向推定装置 14 を備

10

20

30

40

50

えて構成される。

【0032】

マイクロホン12 - 1乃至12 - Mは、様々な方向から到達する音を受音する受音素子であり、受音した音を、その音の大きさに従った振幅の電気信号に変換することで観測信号を生成して、それぞれ対応するフレーム処理部13 - 1乃至13 - Mに供給する。

【0033】

フレーム処理部13 - 1乃至13 - Mは、それぞれ対応するマイクロホン12 - 1乃至12 - Mから供給される観測信号を、所定の短時間ごとのフレームに分割する。例えば、フレーム処理部13 - 1乃至13 - Mは、分割するフレームの長さよりも短い間隔のシフト幅で、次のフレームの先頭に移動するようにフレームを分割する。そして、フレーム処理部13 - 1乃至13 - Mは、フレームに窓関数を掛けた後、短時間フーリエ変換を行うことで、観測信号をフレームごとの周波数信号に分解して、音源方向推定装置14に供給する。

10

【0034】

音源方向推定装置14は、処理対象区間特定部21、立ち上がり共分散行列算出部22、立ち下がり共分散行列算出部23、同時対角化処理部24、および音源方向推定部25を備えて構成される。

【0035】

処理対象区間特定部21には、マイクロホン12 - 1乃至12 - Mのいずれか1つ(図1の例では、マイクロホン12 - M)から観測信号が供給される。そして、処理対象区間特定部21は、マイクロホン12 - 1乃至12 - Mから出力される観測信号のうち、音源方向を推定する処理に用いる対象とする観測信号を区別する時間区間として、発話の立ち上がり区間および立ち下がり区間を特定する。

20

【0036】

例えば、処理対象区間特定部21は、観測信号の振幅二乗値に対してピークホールド処理を行うことにより、発話の立ち上がり区間および立ち下がり区間を処理対象区間として特定する。ピークホールド処理は、観測信号により観測される音のピーク値を時間方向に減衰させながら保持し、保持した値よりも大きな値が現れたときに、その値にピーク値を更新し、保持した値を超える値が現れるまではピークを認識しない処理である。このようなピークホールド処理により、処理対象区間特定部21は、特定の強さの音のピークを検出することができる。

30

【0037】

例えば、処理対象区間特定部21は、1以下の所定の値に設定された減衰率 α を用いて、時刻 t のピーク値 $P(t)$ に減衰率 α を乗算したホールド値 $(P(t) \times \alpha)$ と、次の時刻 $t + 1$ の観測信号の振幅二乗値 $z^2(t + 1)$ とを比較してピークホールド処理を行う。即ち、処理対象区間特定部21は、ホールド値 $(P(t) \times \alpha)$ が観測信号の振幅二乗値 $z^2(t + 1)$ 以上である場合、次の時刻 $t + 1$ のピーク値 $P(t + 1)$ としてホールド値 $(P(t) \times \alpha)$ を用いる。一方、処理対象区間特定部21は、ホールド値 $(P(t) \times \alpha)$ が、観測信号の振幅二乗値 $z^2(t + 1)$ 未満である場合、次の時刻 $t + 1$ のピーク値 $P(t + 1)$ として観測信号の振幅二乗値 $z^2(t + 1)$ を用いる。

40

【0038】

これにより、図2の左側に示すような観測信号の振幅二乗値に対してピークホールド処理を施すことによって、図2の右側に示すように、観測信号の振幅二乗値の波形から特定の強さの音のピークが検出される波形を得ることができる。ここで、ピークホールド処理を利用するのは、振幅の大きさだけでは残響にマスクされているかどうかは分からず、DRR(Direct to Reverberation Ratio)が高いとは限らないためである。従って、残響により直接的に音のマスクを考慮することができるピークホールド処理は、立ち上がりを検出するのに適切である。

【0039】

そして、処理対象区間特定部21は、特定の強さの音のピークが検出されたタイミング

50

を含み、その前後にある所定の時間区間を発話の立ち上がり区間として特定し、その発話の立ち上がり区間を、立ち上がり共分散行列算出部 2 2 に通知する。また、処理対象区間特定部 2 1 は、特定の強さの音のピークが検出されたタイミングから一定時間だけ後に離れ、立ち上がり区間よりも後（例えば、2 フレーム後）にある所定の時間区間を発話の立ち下がり区間として特定し、その発話の立ち下がり区間を、立ち下がり共分散行列算出部 2 3 に通知する。

【 0 0 4 0 】

立ち上がり共分散行列算出部 2 2 は、フレーム処理部 1 3 - 1 乃至 1 3 - M から供給される周波数信号に分解されたフレームのうち、処理対象区間特定部 2 1 により特定された発話の立ち上がり区間におけるフレームを処理の対象とする。そして、立ち上がり共分散行列算出部 2 2 は、処理の対象としたフレームごとに順次、マイクロホン 1 2 - 1 乃至 1 2 - M の順番に従って周波数ごとに信号を並べてベクトル化し、それらのベクトルどうしの掛け算を行って共分散行列を求める。そして、立ち上がり共分散行列算出部 2 2 は、そのようにして求められたフレームごと共分散行列を、発話の立ち上がり区間のフレームで平均化処理したものを立ち上がり共分散行列 R_A として、同時対角化処理部 2 4 に供給する。

10

【 0 0 4 1 】

立ち下がり共分散行列算出部 2 3 は、フレーム処理部 1 3 - 1 乃至 1 3 - M から供給される周波数信号に分解されたフレームのうち、処理対象区間特定部 2 1 により特定された発話の立ち下がり区間におけるフレームを処理の対象とする。そして、立ち下がり共分散行列算出部 2 3 は、立ち上がり共分散行列算出部 2 2 と同様に求められたフレームごと共分散行列を、発話の立ち下がり区間のフレームで平均化処理したものを立ち下がり共分散行列 R_D として、同時対角化処理部 2 4 に供給する。

20

【 0 0 4 2 】

同時対角化処理部 2 4 は、立ち上がり共分散行列算出部 2 2 から供給される立ち上がり共分散行列 R_A (= 観測共分散行列 R) と、立ち下がり共分散行列算出部 2 3 から供給される立ち下がり共分散行列 R_D (= 雑音共分散行列 K) とを用いて、上述した式 (4) を構築する。そして、同時対角化処理部 2 4 は、立ち上がり共分散行列 R_A と立ち下がり共分散行列 R_D とを同時対角化する固有ベクトル e を求める同時対角化処理を行う。

【 0 0 4 3 】

音源方向推定部 2 5 は、同時対角化処理部 2 4 により求められる固有ベクトル e を用いて MUSIC スペクトルを求めることで、立ち上がり共分散行列 R_A に含まれる雑音の影響を除去し、その雑音の影響が軽減された音源方向を推定することができる。例えば、音源方向推定部 2 5 は、固有ベクトル e のうち、マイクロホン 1 2 の個数 M から、所定数の音源の数を減じた数に対応する固有ベクトル e に対して直交するベクトルを求めることにより、音源の方向を推定する。

30

【 0 0 4 4 】

即ち、音源方向推定部 2 5 は、音場に存在すると仮定した音源の数を、マイクロホン 1 2 の数から減じた数を雑音ベクトルの数として、固有値を大きい順に並べた後に、固有値の小さい方から雑音ベクトルの数分だけ雑音ベクトルとする。そして、音源方向推定部 2 5 は、音源があると推定される方向に向けたアレマニューフォールドベクトルを想定し、これと雑音ベクトルとの内積を算出する。ここで、アレマニューフォールドベクトルは、音源方向を決めれば、その音源方向とマイクロホン 1 2 との位置関係から一意に決まる音の到来時間差をマイクロホン 1 2 ごとの要素として持つベクトルである。

40

【 0 0 4 5 】

そして、音源方向推定部 2 5 は、内積の値が 0 に近ければ雑音と直交していることより、つまり、音源の方向であると考えられることより、これを判断するために、内積を分母に持ってきた MUSIC スペクトルを算出する。このとき分母が 0 に近づくということは、MUSIC スペクトルとしては大きな値となることより、MUSIC スペクトルがピークとなる方向が、音源方向を表していることになる。

50

【0046】

以上のように、音源方向推定システム11では、立ち上がり共分散行列 R_A と立ち下がり共分散行列 R_D とを同時対角化する固有ベクトル e を求める同時対角化処理を行うことにより、残響などの雑音の影響を抑制した固有ベクトル e を推定することができる。これにより、MUSIC法による音源方向を推定する精度を向上させることができる。

【0047】

即ち、発話の立ち上がり区間(上述の図2参照)では、観測信号の振幅が大きく、発話に対する反射音や残響音などの外乱の影響が小さくなることより、DRRが高くなる。従って、音源方向推定システム11は、発話の立ち上がり区間のフレームから求められる立ち上がり共分散行列 R_A を用いることで、音源方向の推定精度を向上させることができる。

10

【0048】

さらに、発話の立ち下がり区間では、発話に対する反射音や残響音などの外乱の影響が大きくなっており、DRRが低くなる。従って、音源方向推定システム11は、発話の立ち下がり区間のフレームから求められる立ち下がり共分散行列 R_D を立ち上がり共分散行列 R_A と同時対角化することにより、より高精度に音源方向を推定することができる。このように、立ち上がり共分散行列 R_A に含まれる外乱(主に残響)の寄与を白色化し、固有ベクトルを求めることで、MUSIC法による音源方向の推定精度の向上を図ることができる。

【0049】

図3および図4を参照して、音源方向推定システム11における音源方向推定のシミュレーション結果について説明する。

20

【0050】

図3には、シミュレーション条件が示されている。

【0051】

即ち、図3Aに示すように、マイクロホン数や、マイクロホン間隔、音源方向、音源数などがシミュレーション条件として設定される。また、このシミュレーション条件で用いられる7個のマイクロホン12-1乃至12-7からなるマイクロホンアレーは、図3Bに示すように、X方向、Y方向、およびZ方向に設置される。そして、シミュレーションは、スピーカとマイクロホンアレーとの距離、および、マイクロホンアレーから見た音源方向を、常に一定に保ちながらスピーカおよびマイクロホンアレーの位置をランダムに変更して複数回の試行を繰り返して行った。

30

【0052】

図4には、図3に示したシミュレーション条件に従って、鏡像法を用いて音源方向推定するシミュレーションを行ったシミュレーション結果として、マイクロホンアレーから音源までの距離ごとの正答率が示されている。このシミュレーション結果は、例えば、100回の試行における $\pm 5^\circ$ を誤差としたときに音源方向を推定した正答率である。

【0053】

また、図4には、それぞれ同一のシミュレーション条件で、音源方向推定に用いられる共分散行列の算出対象区間が異なる4つのシミュレーション結果が示されている。

【0054】

40

即ち、第1のシミュレーション結果は、観測信号の全体を用いて算出される共分散行列と、固有値とを利用した通常MUSIC法により音源方向を推定するシミュレーションを行って求められた正答率である。また、第2のシミュレーション結果は、観測信号の立ち上がり区間を用いて算出される立ち上がり共分散行列 R_A から固有ベクトルを求める固有値分解処理を行い、その固有値分解処理で求められた固有ベクトルを利用したMUSIC法により音源方向を推定するシミュレーションを行って求められた正答率である。

【0055】

また、第3のシミュレーション結果は、観測信号の全体を用いて算出される共分散行列と、観測信号の立ち下がり区間を用いて算出される立ち下がり共分散行列 R_D とを同時対角化してMUSIC法により音源方向を推定するシミュレーションを行って求められた正答率

50

である。そして、第4のシミュレーション結果は、観測信号の立ち上がり区間を用いて算出される立ち上がり共分散行列 R_A と、観測信号の立ち下がり区間を用いて算出される立ち下がり共分散行列 R_D とを同時対角化してMUSIC法により音源方向を推定するシミュレーションを行って求められた正答率である。

【0056】

図4に示すように、マイクロホンアレーから音源までの距離を1m、2m、および3mとしてシミュレーションを行った結果、第1乃至第4のシミュレーション結果のいずれも、この距離に依らず同様の傾向を示していることが分かる。

【0057】

例えば、観測信号の全区間の共分散行列と固有値とを利用した第1のシミュレーション結果と比較して、立ち上がり共分散行列 R_A から求められる固有ベクトルを利用した第2のシミュレーション結果は、正答率が上昇している。そして、第2のシミュレーション結果と比較して、観測信号の全区間の共分散行列と立ち下がり共分散行列 R_D とを同時対角化する第3のシミュレーション結果、および、立ち上がり共分散行列 R_A と立ち下がり共分散行列 R_D とを同時対角化する第4のシミュレーション結果は、いずれも正答率が飛躍的に上昇しており、音源方向を推定する推定精度の改善が図られている。

【0058】

このように、音源方向推定システム11では、例えば、立ち上がり共分散行列 R_A と立ち下がり共分散行列 R_D とを同時対角化してMUSIC法により音源方向を推定する手法により、通常のMUSIC法よりも非常に高精度に音源方向を推定することができる。この手法の他、音源方向推定システム11では、観測信号の全区間の共分散行列と立ち下がり共分散行列 R_D とを同時対角化してMUSIC法により音源方向を推定する手法、または、立ち上がり共分散行列 R_A から求められる固有ベクトルを利用したMUSIC法により音源方向を推定する手法を採用しても、通常のMUSIC法よりも高精度に音源方向を推定することができる。

【0059】

次に、図5は、図1の音源方向推定システム11において行われる音源方向推定処理を説明するフローチャートである。

【0060】

例えば、マイクロホン12-1乃至12-Mからフレーム処理部13-1乃至13-Mに観測信号が供給されると処理が開始される。ステップS11において、フレーム処理部13-1乃至13-Mは、それぞれ対応するマイクロホン12-1乃至12-Mから供給される観測信号を、順次、所定の短時間ごとのフレームに分割する。

【0061】

ステップS12において、フレーム処理部13-1乃至13-Mは、ステップS11で分割したフレームに窓関数を掛けた後、短時間フーリエ変換を行うことで、フレームごとに観測信号を周波数信号に分解して、順次、音源方向推定装置14に供給する。

【0062】

ステップS13において、処理対象区間特定部21は、所定のマイクロホン12から供給される観測信号の振幅二乗値に対して、図2を参照して上述したようなピークホールド処理を施し、特定の強さの音のピークが検出されるまで処理を待機する。そして、処理対象区間特定部21は、特定の強さの音のピークが検出されると、そのピークが検出されたタイミングを含む発話の立ち上がり区間を特定して、処理はステップS14に進む。

【0063】

ステップS14において、立ち上がり共分散行列算出部22は、フレーム処理部13-1乃至13-Mにより周波数信号に分解されたフレームのうち、ステップS13において処理対象区間特定部21により特定された発話の立ち上がり区間のフレームを用いて、立ち上がり共分散行列 R_A を算出して同時対角化処理部24に供給する。

【0064】

ステップS15において、処理対象区間特定部21は、ピークホールド処理によりピークが検出されたタイミングから一定時間だけ後にある発話の立ち下がり区間を特定し、処

10

20

30

40

50

理はステップ S 1 6 に進む。

【 0 0 6 5 】

ステップ S 1 6 において、立ち下がり共分散行列算出部 2 3 は、フレーム処理部 1 3 - 1 乃至 1 3 - M により周波数信号に分解されたフレームのうち、ステップ S 1 5 において処理対象区間特定部 2 1 により特定された発話の立ち下がり区間のフレームを用いて、立ち下がり共分散行列 R_D を算出して同時対角化処理部 2 4 に供給する。

【 0 0 6 6 】

ステップ S 1 7 において、同時対角化処理部 2 4 は、ステップ S 1 4 で算出された立ち上がり共分散行列 R_A と、ステップ S 1 6 で算出された立ち下がり共分散行列 R_D とを同時対角化する固有ベクトル e を求める同時対角化処理を行う。

10

【 0 0 6 7 】

ステップ S 1 8 において、音源方向推定部 2 5 は、ステップ S 1 7 で求められた固有ベクトル e を用いて MUSIC スペクトルを求めることで、同時対角化処理部 2 4 により推定された外乱の影響が軽減された音源方向を推定する。

【 0 0 6 8 】

その後、処理はステップ S 1 1 に戻り、以下、上述と同様の処理を繰り返して行う。

【 0 0 6 9 】

以上のように、音源方向推定システム 1 1 は、立ち上がり共分散行列 R_A と立ち下がり共分散行列 R_D とを同時対角化する固有ベクトル e を求める同時対角化処理を行うことにより、MUSIC 法による音源方向の推定精度を向上させることができる。

20

【 0 0 7 0 】

これにより、例えば、通常の部屋などのように反射音や残響音などの外乱が存在する環境下であっても、より高精度に音源方向を推定することができる。例えば、音源方向を正確に推定することで、その方向に指向性ビームを向けたマイクロホンを構築することが可能となり、周囲の外乱を抑圧しながら音声をクリアに収録することができるようになることより、非常に有用である。

【 0 0 7 1 】

従って、音源方向推定システム 1 1 を音声認識装置に採用することにより、例えば、音声認識率の向上を図ることができる。さらに、音源方向推定システム 1 1 を介護や掃除などを行うロボットに搭載することで、例えば、利用者がロボットに呼びかける声に反応して、それらの利用者の位置に正確にロボットを近づけるような制御を行うことが可能となる。

30

【 0 0 7 2 】

なお、上述のフローチャートを参照して説明した各処理は、必ずしもフローチャートとして記載された順序に沿って時系列に処理する必要はなく、並列的あるいは個別に実行される処理（例えば、並列処理あるいはオブジェクトによる処理）も含むものである。また、プログラムは、1 つの CPU により処理されるものであっても良いし、複数の CPU によって分散処理されるものであってもよい。

【 0 0 7 3 】

また、上述した一連の処理（音源方向推定方法）は、ハードウェアにより実行することもできるし、ソフトウェアにより実行することもできる。一連の処理をソフトウェアにより実行する場合には、そのソフトウェアを構成するプログラムが、専用のハードウェアに組み込まれているコンピュータ、または、各種のプログラムをインストールすることで、各種の機能を実行することが可能な、例えば汎用のパーソナルコンピュータなどに、プログラムが記録されたプログラム記録媒体からインストールされる。

40

【 0 0 7 4 】

図 6 は、上述した一連の処理をプログラムにより実行するコンピュータのハードウェアの構成例を示すブロック図である。

【 0 0 7 5 】

コンピュータにおいて、CPU (Central Processing Unit) 1 0 1 , ROM (Read Only Mem

50

ory) 102, RAM (Random Access Memory) 103は、バス104により相互に接続されている。

【0076】

バス104には、さらに、入出力インタフェース105が接続されている。入出力インタフェース105には、キーボード、マウス、マイクロホンなどよりなる入力部106、ディスプレイ、スピーカなどよりなる出力部107、ハードディスクや不揮発性のメモリなどよりなる記憶部108、ネットワークインタフェースなどよりなる通信部109、磁気ディスク、光ディスク、光磁気ディスク、或いは半導体メモリなどのリムーバブルメディア111を駆動するドライブ110が接続されている。

【0077】

以上のように構成されるコンピュータでは、CPU101が、例えば、記憶部108に記憶されているプログラムを、入出力インタフェース105及びバス104を介して、RAM103にロードして実行することにより、上述した一連の処理が行われる。

【0078】

コンピュータ(CPU101)が実行するプログラムは、例えば、磁気ディスク(フレキシブルディスクを含む)、光ディスク(CD-ROM(Compact Disc-Read Only Memory), DVD(Digital Versatile Disc)等)、光磁気ディスク、もしくは半導体メモリなどよりなるパッケージメディアであるリムーバブルメディア111に記録して、あるいは、ローカルエリアネットワーク、インターネット、デジタル衛星放送といった、有線または無線の伝送媒体を介して提供される。

【0079】

そして、プログラムは、リムーバブルメディア111をドライブ110に装着することにより、入出力インタフェース105を介して、記憶部108にインストールすることができる。また、プログラムは、有線または無線の伝送媒体を介して、通信部109で受信し、記憶部108にインストールすることができる。その他、プログラムは、ROM102や記憶部108に、あらかじめインストールしておくことができる。

【0080】

また、本明細書において、システムとは、複数の装置により構成される装置全体を表すものである。

【0081】

なお、本実施の形態は、上述した実施の形態に限定されるものではなく、本開示の要旨を逸脱しない範囲において種々の変更が可能である。

【符号の説明】

【0082】

11 音源方向推定システム, 12 マイクロホン, 13 フレーム処理部, 14 音源方向推定装置, 21 処理対象区間特定部, 22 立ち上がり共分散行列算出部, 23 立ち下がり共分散行列算出部, 24 同時対角化処理部, 25 音源方向推定部

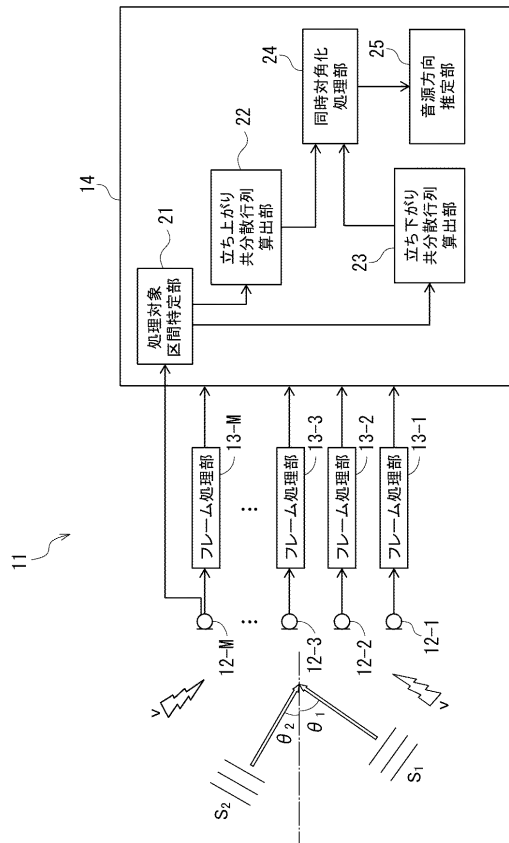
10

20

30

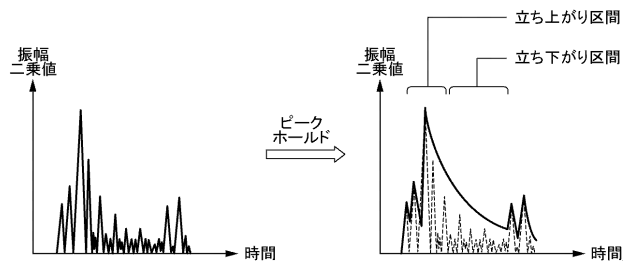
【図1】

Fig. 1



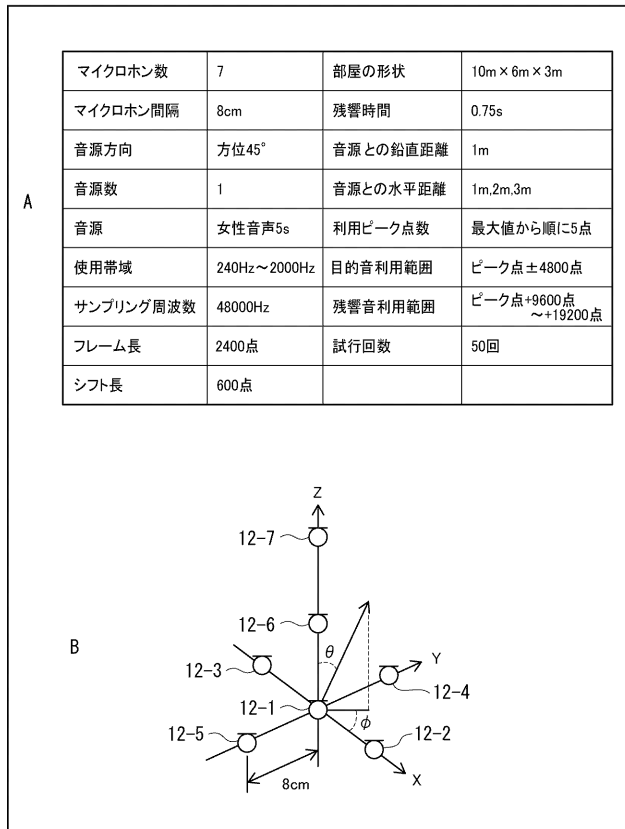
【図2】

Fig. 2



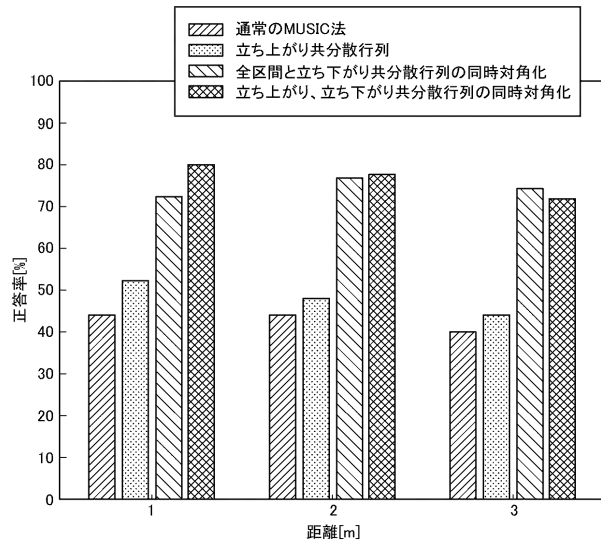
【図3】

Fig. 3



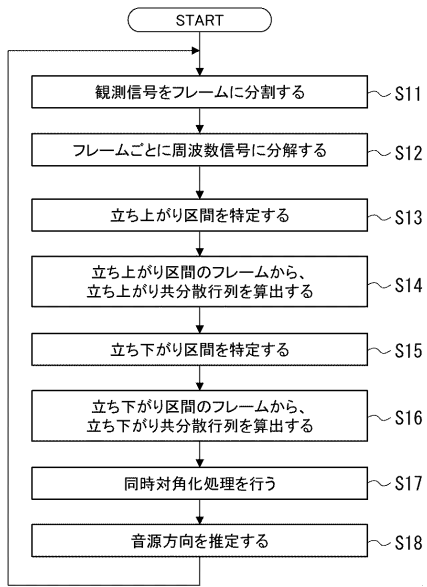
【図4】

Fig. 4



【 図 5 】

Fig. 5



【 図 6 】

Fig. 6

