

(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2010-72164  
(P2010-72164A)

(43) 公開日 平成22年4月2日(2010.4.2)

(51) Int.Cl.	F I	テーマコード (参考)
<b>G 1 0 L 11/02 (2006.01)</b>	G 1 0 L 11/02	5 D 0 1 5
<b>G 1 0 L 21/02 (2006.01)</b>	G 1 0 L 21/02 1 0 1 B	
<b>G 1 0 L 15/04 (2006.01)</b>	G 1 0 L 21/02 1 0 2 A	
<b>G 1 0 L 15/20 (2006.01)</b>	G 1 0 L 15/04 3 0 0 A	
	G 1 0 L 15/20 3 7 0 D	
審査請求 未請求 請求項の数 12 O L (全 24 頁) 最終頁に続く		

(21) 出願番号	特願2008-237642 (P2008-237642)	(71) 出願人	000004226 日本電信電話株式会社 東京都千代田区大手町二丁目3番1号
(22) 出願日	平成20年9月17日 (2008.9.17)	(71) 出願人	504132272 国立大学法人京都大学 京都府京都市左京区吉田本町36番地1
		(74) 代理人	100121706 弁理士 中尾 直樹
		(74) 代理人	100066153 弁理士 草野 卓
		(74) 代理人	100128705 弁理士 中村 幸雄
		(72) 発明者	石塚 健太郎 東京都千代田区大手町二丁目3番1号 日 本電信電話株式会社内
		最終頁に続く	

(54) 【発明の名称】 目的信号区間推定装置、目的信号区間推定方法、目的信号区間推定プログラム及び記録媒体

(57) 【要約】

【課題】 方向性の雑音や拡散性の雑音が含まれる環境であっても、信号の音源数や到来方向を事前を知ることなく、精度よく目的信号区間を推定する。

【解決手段】 複数のセンサで観測された各信号を所定の時間区間であるフレーム毎に切り出し、切り出された各センサについての各フレームの信号を周波数領域に変換し、周波数領域信号を各センサについて生成する。また、基準センサに対応する周波数領域信号を基準として、当該基準センサ以外のセンサに対応する各周波数領域信号を正規化し、信号の到来方向に対応する正規化信号値を生成し、それを用いて量子化された到来方向ごとの空間パワー分布値を求め、その統計的性質の時間変化に基づいて各フレームが目的信号区間に対応するかどうかを判定する。

【選択図】 図1

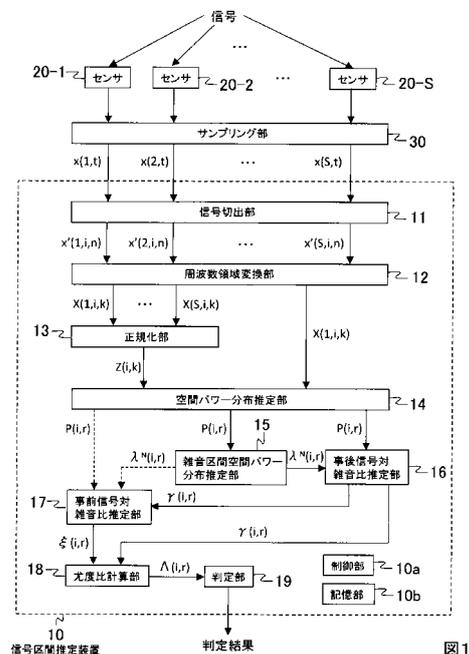


図1

**【特許請求の範囲】****【請求項 1】**

目的信号が存在する目的信号区間を推定する目的信号区間推定装置であって、

複数のセンサで観測された各信号をそれぞれ所定の時間区間であるフレーム毎に切り出す信号抽出部と、

上記信号抽出部で切り出された各フレームの信号を周波数領域に変換した周波数領域信号を各センサについて生成する周波数領域変換部と、

上記センサに含まれる特定の基準センサに対応する上記周波数領域信号を基準として、少なくとも当該基準センサ以外の何れかの上記センサに対応する周波数領域信号を正規化し、上記センサで観測された信号の到来方向に対応する正規化信号値を各フレームについて生成する正規化部と、

10

上記正規化信号値がとり得る全範囲を複数に分割した各区分を各到来方向区分とし、上記正規化信号値に対応するフレームの上記周波数領域信号の振幅の絶対値に対して単調増加する値を、当該正規化信号値が属する上記到来方向区分の当該フレームでの空間パワー分布値とする空間パワー分布推定部と、

目的信号が存在しないと推定される雑音区間のフレームに対応する各到来方向区分の空間パワー分布値を用い、雑音区間の空間パワー分布値の推定値である雑音区間空間パワー分布値を各到来方向区分について生成する雑音区間空間パワー分布推定部と、

各到来方向区分の各フレームにおける上記空間パワー分布値と、各到来方向区分の上記雑音区間空間パワー分布値とを用い、当該空間パワー分布値と当該雑音区間空間パワー分布値との比である事後信号対雑音比を、到来方向区分ごとに各フレームについて生成する事後信号対雑音比推定部と、

20

上記事後信号対雑音比を少なくとも用い、到来方向区分ごとに、各フレームが雑音区間である尤度と各フレームが目的信号区間である尤度との比である尤度比又はその関数値を生成する尤度比計算部と、

を有する目的信号区間推定装置。

**【請求項 2】**

請求項 1 の目的信号区間推定装置であって、

上記事後信号対雑音比を少なくとも用い、上記雑音区間空間パワー分布値と、空間パワー分布値の目的信号成分である目的信号空間パワー分布値との比である事前信号対雑音比を、到来方向区分ごとに各フレームについて生成する事前信号対雑音比推定部をさらに有し、

30

上記尤度比計算部は、

上記事後信号対雑音比と上記事前信号対雑音比とを用いて上記尤度比又はその関数値を生成する、

ことを特徴とする目的信号区間推定装置。

**【請求項 3】**

請求項 1 又は 2 の目的信号区間推定装置であって、

上記尤度比又はその関数値を指標とし、各フレームが上記目的信号区間であるか否かを判定する判定部をさらに有する、

40

ことを特徴とする目的信号区間推定装置。

**【請求項 4】**

請求項 3 の目的信号区間推定装置であって、

上記判定部は、

上記尤度比又はその関数値をフレームごとに上記到来方向区分について平均し、その演算結果を指標として、各フレームが上記目的信号区間であるか否かを判定する、

ことを特徴とする目的信号区間推定装置。

**【請求項 5】**

請求項 1 から 4 の何れかの目的信号区間推定装置であって、

上記正規化部は、

50

周波数ビンとフレームとの組ごとに上記正規化信号値を生成し、  
上記空間パワー分布推定部は、

特定の到来方向区分に属する上記正規化信号値の周波数ビン及びフレームに対してハイレベル値をとり、当該特定の到来方向区分に属しない上記正規化信号値の周波数ビン及びフレームに対してローレベル値をとる、当該特定の到来方向区分と周波数ビンとフレームとを変数とする関数である時間周波数マスクを生成する時間周波数マスク生成部と、

周波数ビンとフレームとの組が同一である上記周波数領域信号と上記時間周波数マスクとの積に対して単調増加する値を用い、当該フレームと当該到来方向区分とに対応する空間パワー分布値を生成するマスク適用部と、

を有することを特徴とする目的信号区間推定装置。

10

【請求項 6】

請求項 1 から 5 の何れかの目的信号区間推定装置であって、

上記雑音区間空間パワー分布推定部は、

複数の上記雑音区間のフレームでの各到来方向区分の空間パワー分布値を、到来方向区分ごとに当該雑音区間のフレームについて平均した値を、各到来方向区分についての雑音区間空間パワー分布値として生成する、

ことを特徴とする目的信号区間推定装置。

【請求項 7】

請求項 1 から 6 の何れかの目的信号区間推定装置であって、

上記事後信号対雑音比推定は、

20

同一の到来方向区分に対応する上記空間パワー分布値と上記雑音区間空間パワー分布値との組ごとに、当該空間パワー分布値を当該雑音区間空間パワー分布値で除算し、その演算結果を、当該空間パワー分布値の到来方向区分及びフレームに対応する上記事後信号対雑音比として生成し、

上記事前信号対雑音比推定部は、

上記事後信号対雑音比から 1 を減じた値を、当該事後信号対雑音比の到来方向区分及びフレームに対応する事前信号対雑音比として生成する、

ことを特徴とする目的信号区間推定装置。

【請求項 8】

請求項 1 から 6 の何れかの目的信号区間推定装置であって、

30

上記事後信号対雑音比推定は、

同一の到来方向区分に対応する上記空間パワー分布値と上記雑音区間空間パワー分布値との組ごとに、当該空間パワー分布値を当該雑音区間空間パワー分布値で除算し、その演算結果を、当該空間パワー分布値の到来方向区分及びフレームに対応する上記事後信号対雑音比として生成し、

上記事前信号対雑音比推定部は、

到来方向区分  $r$  及びフレーム  $i$  に対応する上記事後信号対雑音比を  $(i, r)$  とし、到来方向区分  $r$  及びフレーム  $i$  に対応する上記事前信号対雑音比を  $(i, r)$  とし、到来方向区分  $r$  及びフレーム  $i$  に対応する空間パワー分布値を  $P(i, r)$  とし、到来方向区分  $r$  及びフレーム  $i$  に対応する上記雑音区間空間パワー分布値を  $N(i, r)$  とし、 $\max(\cdot)$  を  $\cdot$  の最大値を出力する関数とし、 $\Gamma(\cdot)$  をガンマ関数とし、 $I_0(\cdot)$  と  $I_1(\cdot)$  をそれぞれ 0 次と 1 次の第 1 種修正ベッセル関数とし、 $(i-1, r) = (i-1, r) \cdot (i-1, r) / (1 + (i-1, r))$  とし、 $S^\wedge(i-1, r) = \{ (\Gamma(i-1, r))^{1/2} / (i-1, r) \} \cdot (1.5) \cdot \exp(- (i-1, r) / 2) \cdot [(1 + (i-1, r)) \cdot I_0((i-1, r) / 2) + (i-1, r) \cdot I_1((i-1, r) / 2)] \cdot (P(i-1, r))^{1/2}$  とし、 $(i-1, r)$  を 0 以上 1 以下の定数とした場合における、 $(i-1, r) = |S^\wedge(i-1, r)|^2 / N(i-1, r) + (1 - (i-1, r)) \cdot \max((i, r) - 1, 0)$  を到来方向区分  $r$  及びフレーム  $i$  に対応する上記事前信号対雑音比として生成する、

40

ことを特徴とする目的信号区間推定装置。

【請求項 9】

請求項 1 から 8 の何れかの目的信号区間推定装置であって、

50

上記尤度比計算部は、

到来方向区分  $r$  及びフレーム  $i$  に対応する上記事後信号対雑音比を  $(i, r)$  とし、到来方向区分  $r$  及びフレーム  $i$  に対応する上記事前信号対雑音比を  $(i, r)$  とした場合における、 $(i, r) = \{1 / (1 + (i, r))\} \cdot \exp\{(i, r) \cdot (i, r) / (1 + (i, r))\}$  を上記尤度比として生成する、

ことを特徴とする目的信号区間推定装置。

【請求項 10】

目的信号が存在する目的信号区間を推定する目的信号区間推定装置の目的信号区間推定方法であって、

信号抽出部が、複数のセンサで観測された各信号をそれぞれ所定の時間区間であるフレーム毎に切り出すステップと、

周波数領域変換部が、上記信号切出部で切り出された各フレームの信号を周波数領域に変換した周波数領域信号を各センサについて生成するステップと、

正規化部が、上記センサに含まれる特定の基準センサに対応する上記周波数領域信号を基準として、少なくとも当該基準センサ以外の何れかの上記センサに対応する周波数領域信号を正規化し、上記センサで観測された信号の到来方向に対応する正規化信号値を各フレームについて生成するステップと、

空間パワー分布推定部が、上記正規化信号値がとり得る全範囲を複数に分割した各区分を各到来方向区分とし、上記正規化信号値に対応するフレームの上記周波数領域信号の振幅の絶対値に対して単調増加する値を、当該正規化信号値が属する上記到来方向区分の当該フレームでの空間パワー分布値とするステップと、

雑音区間空間パワー分布推定部が、目的信号が存在しないと推定される雑音区間のフレームに対応する各到来方向区分の空間パワー分布値を用い、雑音区間の空間パワー分布値の推定値である雑音区間空間パワー分布値を各到来方向区分について生成するステップと、

事後信号対雑音比推定部が、各到来方向区分の各フレームにおける上記空間パワー分布値と、各到来方向区分の上記雑音区間空間パワー分布値とを用い、当該空間パワー分布値と当該雑音区間空間パワー分布値との比である事後信号対雑音比を、到来方向区分ごとに各フレームについて生成するステップと、

尤度比計算部が、上記事後信号対雑音比を少なくとも用い、到来方向区分ごとに、各フレームが雑音区間である尤度と各フレームが目的信号区間である尤度との比である尤度比又はその関数値を生成するステップと、

を有する目的信号区間推定方法。

【請求項 11】

請求項 1 から 9 の何れかに記載の目的信号区間推定装置としてコンピュータを機能させるための目的信号区間推定プログラム。

【請求項 12】

請求項 11 に記載の目的信号区間推定プログラムを格納したコンピュータ読み取り可能な記録媒体。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、信号処理技術に関し、特に、雑音が含まれる観測信号から目的信号が存在する区間を推定する技術に関する。

【背景技術】

【0002】

音声信号や音楽信号などの目的信号を処理対象とする符号化、雑音信号の抑圧、残響除去、自動音声認識などの音響信号処理技術では、複数種類の信号が含まれる入力音響信号から目的信号が存在する区間を推定する必要がある。この目的信号区間推定の精度はその後の信号処理性能に大きく影響する。

10

20

30

40

50

## 【 0 0 0 3 】

従来、目的信号の到来方向を既知として到来方向の信号を強調することで目的信号区間推定を行う手法（非特許文献1）、目的信号の到来方向を推定した後に到来方向からの信号の周波数スペクトルの情報を用いて目的信号の有無を推定する手法（非特許文献2）、音声強調により得られた目的信号を観測信号から差し引くことにより目的信号の存在区間を推定する手法（特許文献1）、パラメトリックな方法で求められた空間スペクトルのピークの有無で音声の有無を推定する手法（非特許文献3）、目的信号を抑圧して得られた雑音信号を用いて周波数スペクトル上で目的信号を強調する手法（非特許文献4）などがある。しかし、これらの方法で十分な精度を得るためには、目的信号の到来方向や音源数が既知であるか、事前に目的信号の到来方向の正確な推定が必要であったり、拡散性・方向性の雑音がない環境での收音が必要であったりする。また、これらの前提を必要としない手法として、雑音のコヒーレンス関数を観測信号のコヒーレンス関数から引くことで目的信号の有無を推定する手法（非特許文献5）、雑音のCSP（Cross-power Spectrum Phase）係数を観測信号のCSP係数から差し引くことで目的信号の有無を推定する手法（非特許文献6）がある。しかし、これらの周波数スペクトルの相互相関を用いる手法は、特に低周波数帯域において拡散性雑音の影響を十分に取り除くことができないという問題がある。

10

## 【 0 0 0 4 】

その他、あらゆる方向からあらゆる周波数帯域の音響信号が同時に到来するような環境（例えば、街頭、駅、空港のような日常環境）で信号の到達時間差を用いて十分な目的信号区間推定精度を得るための手法として、一定範囲の時間周波数領域で推定した信号の到達時間差がある一定の値に偏る度合いを利用する手法（非特許文献7）がある。しかし、この手法では、方向性の雑音が存在する場合に、それも目的信号として検出してしまおうという問題点がある。

20

【非特許文献1】 Alvarez, A., Gomez, P., Nieto, V., Martinez, R., and Rodellar, V., "Application of a first-order differential microphone for efficient voice activity detection in a car platform", Proceedings of Interspeech, 2669-2672, 2005.

【非特許文献2】 Potamitis, I. and Fishler, E., "Speech activity detection and enhancement of a moving speaker based on the wideband generalized likelihood ratio and microphone arrays," Journal of the Acoustical Society of America, vol. 116, pp. 2406-2415, 2004.

30

【非特許文献3】 Asano, F., Yamamoto, K., Ogata, J., Yamada, M., and Nakamura, M., "Detection and separation of speech events in meeting recordings using a microphone array," EURASIP Journal Audio, Speech, and Music Processing, vol. 2007, Article ID 27616, 2007.

【非特許文献4】 Takahashi, Y., Takatani, T., Saruwatari, H., and Shikano, K., "Blind spatial subtraction array with independent component analysis for hands-free speech recognition," Proceedings of IWAENC, 2006.

【非特許文献5】 Le Bouquin-Jeannes R., Azirani, A. A., and Faucon, G., "Enhancement of speech degraded by coherent and incoherent noise using a cross-spectral estimator," IEEE Transactions on Speech and Audio Processing, vol. 5, pp. 484-487, 1997.

40

【非特許文献6】 Denda, Y., Nishiura, T., and Yamashita, Y., "Robust talker direction estimation based on weighted CSP analysis and maximum likelihood estimation," IEICE Transactions on Information and Systems, vol. E89-D, pp. 1050-1057, 2006.

【非特許文献7】 Juan E. Rubio, Kentaro Ishizuka, Hiroshi Sawada, Shoko Araki, Tomohiro Nakatani, and Masakiyo Fujimoto, "Two-Microphone Voice Activity Detection Based on the Homogeneity of the Direction of Arrival Estimates," Proceedings of the 32nd International Conference on Acoustics, Speech, and Signal Processing,

50

Vol. 4, pp. 385-388, 2007.

【特許文献1】特許第2913105号公報

【発明の開示】

【発明が解決しようとする課題】

【0005】

以上のように従来手法では、方向性の雑音や拡散性の雑音が含まれる環境であっても、信号の音源数や到来方向を事前に知ることなく、精度よく目的信号区間を推定することができなかった。

【0006】

本発明はこのように鑑みてなされたものであり、方向性の雑音や拡散性の雑音が含まれる環境であっても、信号の音源数や到来方向を事前に知ることなく、精度よく目的信号区間を推定することが可能な技術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】

【0007】

本発明では上記課題を解決するために、まず、信号抽出部が、複数のセンサで観測された各信号をそれぞれ所定の時間区間であるフレーム毎に切り出し、周波数領域変換部が、信号切出部で切り出された各フレームの信号を周波数領域に変換した周波数領域信号を各センサについて生成する。また、正規化部が、センサに含まれる特定の基準センサに対応する周波数領域信号を基準として、少なくとも当該基準センサ以外の何れかのセンサに対応する周波数領域信号を正規化し、センサで観測された信号の到来方向に対応する正規化信号値を各フレームについて生成する。そして、空間パワー分布推定部が、正規化信号値がとり得る全範囲を複数に分割した各区分を各到来方向区分とし、正規化信号値に対応するフレームの周波数領域信号の振幅の絶対値に対して単調増加する値を、当該正規化信号値が属する到来方向区分の当該フレームでの空間パワー分布値とし、雑音区間空間パワー分布推定部が、目的信号が存在しないと推定される雑音区間のフレームに対応する各到来方向区分の空間パワー分布値を用い、雑音区間の空間パワー分布値の推定値である雑音区間空間パワー分布値を各到来方向区分について生成する。その後、事後信号対雑音比推定部が、各到来方向区分の各フレームにおける空間パワー分布値と、各到来方向区分の雑音区間空間パワー分布値とを用い、当該空間パワー分布値と当該雑音区間空間パワー分布値との比である事後信号対雑音比を、到来方向区分ごとに各フレームについて生成する。そして、尤度比計算部が、事後信号対雑音比を少なくとも用い、到来方向区分ごとに、各フレームが雑音区間である尤度と各フレームが目的信号区間である尤度との比である尤度比又はその関数値を生成する。

【0008】

ここで、事後信号対雑音比推定部によって生成される到来方向区分ごとの事後信号対雑音比は、各フレームにおける空間パワー分布値と、目的信号が存在しないと推定される雑音区間のフレームに対応する空間パワー分布値を用いて生成された雑音区間空間パワー分布値との比である。また、当該事後信号対雑音比は、到来方向区分ごとに生成される。つまり、各事後信号対雑音比は、各フレームにおける空間パワー分布値が、目的信号が存在しないと推定される雑音区間のフレームに対応する空間パワー分布値に対してどの程度変化したかを到来方向区分ごとに示すものとなる。その結果、当該事後信号対雑音比を少なくとも用いて生成される尤度比又はその関数値にも、この変化が反映される。すなわち、本発明では、雑音区間を基準とした空間パワー分布値の統計的性質の変化が反映される尤度比又はその関数値を、到来方向区分ごとに各フレームについて求める。方向性の雑音及び拡散性の雑音は、通常ほぼ一定の統計的性質を持ち、或る到来方向から目的信号が到来した場合、その統計的性質は変化する。本発明では、統計的性質の変化が反映される上記尤度比又はその関数値を各フレームについて算出するため、それを用いることで、方向性の雑音や拡散性の雑音が含まれる環境であっても精度よく目的信号区間を推定することができる。また、このような尤度比又はその関数値は、各到来方向区分について算出されるため、信号の音源数や到来方向を事前に知る必要はない。さらに、このような尤度比又は

10

20

30

40

50

その関数値の算出に用いられる事後信号対雑音比を到来方向区分ごとに算出するため、到来方向区分ごとの統計量の変化を個別に捉えることができ、目的信号区間を制度よく推定できる。

【0009】

また、本発明において、事後信号対雑音比を少なくとも用い、雑音区間空間パワー分布値と、空間パワー分布値の目的信号成分である目的信号空間パワー分布値との比である事前信号対雑音比を、到来方向区分ごとに各フレームについて生成する事前信号対雑音比推定部をさらに設け、尤度比計算部が、上記事後信号対雑音比と上記事前信号対雑音比とを用いて上記尤度比又はその関数値を生成することとしてもよい。これにより、より高い精度で尤度比又はその関数値を推定することが可能となる。

10

【0010】

また、上記の尤度比を指標として各フレームが目的信号区間であるか否かを判定する場合、例えば、判定部が、尤度比又はその関数値をフレームごとに到来方向区分について平均し、その演算結果を指標として、各フレームが目的信号区間であるか否かを判定してもよい。この場合、突発的に生じたパワーの小さい方向性の雑音の影響を低減させ、目的信号区間の推定精度を向上させることができる。

【0011】

また、本発明において好ましくは、正規化部は、周波数ビンとフレームとの組ごとに正規化信号値を生成し、空間パワー分布推定部は、特定の到来方向区分に属する正規化信号値の周波数ビン及びフレームに対してハイレベル値をとり、当該特定の到来方向区分に属しない正規化信号値の周波数ビン及びフレームに対してローレベル値をとる、当該特定の到来方向区分と周波数ビンとフレームとを変数とする関数である時間周波数マスクを生成し、周波数ビンとフレームとの組が同一である周波数領域信号と時間周波数マスクとの積に対して単調増加する値を用い、当該フレームと当該到来方向区分とに対応する空間パワー分布値を生成する。これにより、信号源数とセンサ数との大小関係にかかわらず、空間パワー分布値を正確に生成できる。

20

【0012】

また、本発明において好ましくは、雑音区間空間パワー分布推定部は、複数の雑音区間のフレームでの各到来方向区分の空間パワー分布値を、到来方向区分ごとに当該雑音区間のフレームについて平均した値を、各到来方向区分についての雑音区間空間パワー分布値として生成する。これにより、妥当な雑音区間空間パワー分布値が得られるため、目的信号区間の推定精度が向上する。

30

【0013】

また、事後信号対雑音比推定は、例えば、同一の到来方向区分に対応する空間パワー分布値と雑音区間空間パワー分布値との組ごとに、当該空間パワー分布値を当該雑音区間空間パワー分布値で除算し、その演算結果を、当該空間パワー分布値の到来方向区分及びフレームに対応する事後信号対雑音比として生成し、事前信号対雑音比推定部は、事後信号対雑音比から1を減じた値を、当該事後信号対雑音比の到来方向区分及びフレームに対応する事前信号対雑音比として生成する。これにより、少ない演算量で事前信号対雑音比を算出できる。

40

【0014】

また、事後信号対雑音比推定が、同一の到来方向区分に対応する空間パワー分布値と雑音区間空間パワー分布値との組ごとに、当該空間パワー分布値を当該雑音区間空間パワー分布値で除算し、その演算結果を、当該空間パワー分布値の到来方向区分及びフレームに対応する事後信号対雑音比として生成し、事前信号対雑音比推定部が、到来方向区分  $r$  及びフレーム  $i$  に対応する事後信号対雑音比を  $(i, r)$  とし、到来方向区分  $r$  及びフレーム  $i$  に対応する事前信号対雑音比を  $(i, r)$  とし、到来方向区分  $r$  及びフレーム  $i$  に対応する空間パワー分布値を  $P(i, r)$  とし、到来方向区分  $r$  及びフレーム  $i$  に対応する雑音区間空間パワー分布値を  $N(i, r)$  とし、 $\max(\cdot)$  を  $\cdot$  の最大値を出力する関数とし、 $(\cdot)$  をガンマ関数とし、 $I_0(\cdot)$  と  $I_1(\cdot)$  をそれぞれ0次と1次の第1種修正ベッセル関数とし、

50

$(i-1, r) = (i-1, r) \cdot (i-1, r) / (1 + (i-1, r))$  とし、 $S^{(i-1, r)} = \{((i-1, r))^{1/2} / (i-1, r)\} \cdot (1.5) \cdot \exp(-(i-1, r)/2) \cdot [(1 + (i-1, r)) \cdot I_0((i-1, r)/2) + (i-1, r) \cdot I_1((i-1, r)/2)] \cdot (P(i-1, r))^{1/2}$  とし、 $\alpha$  を 0 以上 1 以下の定数とした場合における、 $(i-1, r) = \alpha \cdot |S^{(i-1, r)}|^2 / N(i-1, r) + (1 - \alpha) \cdot \max((i, r) - 1, 0)$  を到来方向区分  $r$  及びフレーム  $i$  に対応する事前信号対雑音比として生成してもよい。これにより、定数  $\alpha$  の値を適宜選択することで、環境に適した事前信号対雑音比を生成することができる。

【0015】

また、尤度比計算部は、例えば、到来方向区分  $r$  及びフレーム  $i$  に対応する事後信号対雑音比を  $(i, r)$  とし、到来方向区分  $r$  及びフレーム  $i$  に対応する事前信号対雑音比を  $(i, r)$  とした場合における、 $(i, r) = \{1 / (1 + (i, r))\} \cdot \exp\{(i, r) \cdot (i, r) / (1 + (i, r))\}$  を尤度比として生成する。

10

【発明の効果】

【0016】

以上のように本発明では、方向性の雑音や拡散性の雑音が含まれる環境であっても、信号の音源数や到来方向を事前に知ることなく、精度よく目的信号区間を推定することが可能となる。

【発明を実施するための最良の形態】

【0017】

以下、本発明を実施するための最良の形態を図面を参照して説明する。

図1は、本形態の目的信号区間推定装置10の全体構成を例示したブロック図である。また、図2(a)は、本形態の空間パワー分布推定部14の詳細を例示するブロック図であり、図2(b)は、本形態の判定部19の詳細を例示するブロック図である。

20

【0018】

<構成>

図1に例示するように、本形態の目的信号区間推定装置10は、信号切出部11と周波数領域変換部12と正規化部13と空間パワー分布推定部14と雑音区間空間パワー分布推定部15と事後信号対雑音比推定部16と事前信号対雑音比推定部17と尤度比計算部18と判定部19と制御部10aと記憶部10bとを具備し、 $S$  ( $S \geq 2$ ) 個のセンサ20-1 ~  $S$  で観測され、サンプリング部30でサンプリングされた信号が入力され、目的信号区間の分析結果を出力する装置である。また、図2(a)に例示する空間パワー分布推定部14は、時間周波数マスク生成部14aとマスク適用部14bとを有し、図2(b)に例示する判定部19は、平均算出部19aと閾値比較部19bとを有する。

30

【0019】

なお、本形態の目的信号区間推定装置10は、例えば、CPU (central processing unit), RAM (random access memory), ROM (read only memory) 等から構成される公知のコンピュータに所定のプログラムを実行させることによって構成されるものである。すなわち、信号切出部11と周波数領域変換部12と正規化部13と空間パワー分布推定部14と雑音区間空間パワー分布推定部15と事後信号対雑音比推定部16と事前信号対雑音比推定部17と尤度比計算部18と判定部19と制御部10aは、それぞれ、例えば、CPUが所定のプログラムを実行して構築される処理部であり、記憶部10bは、例えば、ハードディスクなどの補助記憶装置、RAM、キャッシュメモリ若しくはレジスタ、又は、それらの少なくとも一部が結合して構成される記憶領域である。

40

【0020】

<処理>

次に、本形態の目的信号区間推定方法について説明する。

本形態の目的信号区間推定方法では、複数のセンサ20-1 ~  $S$  ( $S \geq 2$ ) で観測された各信号を時間周波数分析し、特定の基準センサを基準とした正規化信号値を求め、正規化信号値がとり得る全範囲を複数に分割した各区分である各到来方向区分におけるパワーの統計的性質の時間的变化に基づいて、目的信号の有無を検出して出力する。なお、本形態では、複数のセンサ20-1 ~  $S$  としてマイクロホンを利用し、それらで観測された各

50

音響信号を用い、音声信号や音楽信号などの目的信号の有無を検出して出力する場合を例示する。また、以下では明記しないが、目的信号区間推定装置 10 は、制御部 10 a の制御に基づいて各演算処理を実行する。また、本形態の場合、各演算処理の過程で得られたデータは記憶部 10 b に逐次格納され、それ以降の各演算処理において読み出されて利用される。すなわち、本形態において、処理部 A からデータが出力され、このデータが処理部 B に入力されるとは、処理部 A から出力されたデータが記憶部 10 b に格納され、処理部 B がこのデータを記憶部 10 b から読み込むことを意味する。

【0021】

図3は、本形態の目的信号区間推定方法を説明するためのフローチャートである。また、図4(a)は、ステップS5の詳細を例示するためのフローチャートであり、図4(b)は、ステップS10の詳細を例示するためのフローチャートである。以下、これらのフローチャートに沿って本形態の目的信号区間推定方法を説明していく。

10

【0022】

まず、 $S(S-2)$ 個のセンサ20-1~Sでそれぞれ観測された各信号がサンプリング部30に入力される。これらの信号には音声信号や音楽信号等目的信号の他に環境雑音信号も含まれる。サンプリング部30は、各信号を所定(例えば16,000Hz)のサンプリング周波数 $f_s$ でサンプリングし、これにより各センサ20-1~Sに対応する時間領域の信号 $x(1,t), \dots, x(S,t)$ を抽出する(ステップS1)。なお、 $t$ は $t$ 番目のサンプリング点を示す。

20

【0023】

サンプリング部30で抽出された各時間領域の信号 $x(1,t), \dots, x(S,t)$ は、目的信号区間推定装置10の信号切出部11に入力される。信号切出部11は、入力された各信号 $x(1,t), \dots, x(S,t)$ を所定の時間区間であるフレーム毎に切り出し、各センサ20-1~Sについての各フレーム $i$ ( $i$ はフレームインデックスを示す)の信号 $x'(1,i,n), \dots, x'(S,i,n)$ を抽出する(ステップS2)。なお、 $n$ はフレーム $i$ における $n$ 番目のサンプル点を表す。具体的には、信号切出部11は、例えば、入力された各信号 $x(1,t), \dots, x(S,t)$ に対し、それぞれ所定の窓関数を例えば時間軸方向に16msずつ移動(シフト)させながら乗じ、これにより、例えば32msの時間長の信号 $x'(1,i,n), \dots, x'(S,i,n)$ を切り出す。より具体的には、例えば、サンプリング周波数が16,000Hzの場合、信号切出部11は、入力された各信号 $x(1,t), \dots, x(S,t)$ に対し、それぞれ例えば式(1)のハニング窓を256サンプル点(16,000Hz×16ms)ずつ移動(シフト)させながら乗じ、センサ20-1~S毎に、512サンプル点(16,000Hz×32ms)の離散信号を1フレーム分の信号として切り出す。ここで、 $L$ は切り出される1フレーム分の信号のサンプル点数(フレーム長:上述の例では $L=512$ )を表す。

30

【0024】

【数1】

$$w(n) = 0.5 - 0.5 \cdot \cos \frac{2\pi n}{L-1} \quad \dots (1)$$

【0025】

信号切出部11は、以上のように切り出した各センサ20-1~Sについての各フレーム $i$ の信号 $x'(1,i,n), \dots, x'(S,i,n)$ を出力し、これらは周波数領域変換部12に入力される。

40

【0026】

周波数領域変換部12では、各センサ20-1~Sについての各フレーム $i$ の信号 $x'(1,i,n), \dots, x'(S,i,n)$ を周波数領域に変換し、時間周波数ビン $(i,k)$ 毎の周波数領域信号(周波数領域スペクトル) $X(1,i,k), \dots, X(S,i,k)$ を各センサ20-1~Sについて生成する(ステップS3)。離散フーリエ変換によってこの変換を行う場合、周波数領域変換部12は、以下の式(2)のように周波数領域信号 $X(1,i,k), \dots, X(S,i,k)$ を算出する。

【0027】

50

【数2】

$$X(s, i, k) = \sum_{n=0}^{M-1} x'(s, i, n) \cdot \exp\left(-j \frac{2\pi}{M} \cdot k \cdot n\right) \quad \dots (2)$$

【0028】

ここで、 $j$  は虚数単位を示し、 $s$  ( $s = \{1, \dots, S\}$ ) は各センサ 20 - 1 ~  $S$  の番号を示す。また、 $k$  ( $k=0, \dots, M-1$ ) は周波数インデックスであり、サンプリング周波数  $f_s$  を  $M$  等分した離散点を表す。 $M$  はフレーム長  $L$  以上の自然数であり、例えば、 $M=512$  とする。周波数領域変換部 12 は、以上のような変換によって得られた周波数領域信号 (周波数スペクトル)  $X(1, i, k), \dots, X(S, i, k)$  を出力する。

10

【0029】

周波数領域変換部 12 から出力された周波数領域信号  $X(1, i, k), \dots, X(S, i, k)$  は、正規化部 13 に入力される。正規化部 13 は、特定の基準センサ  $s_B$  ( $\{1, \dots, S\}$ ) に対応する周波数領域信号  $X(s_B, i, k)$  を基準として、少なくとも当該基準センサ  $s_B$  以外のセンサ  $s$  ( $s_B$ ) に対応する各周波数領域信号  $X(s, i, k)$  を正規化し、センサで観測された信号の到来方向に対応する正規化信号値  $Z(i, k)$  を各フレーム  $i$  及び各周波数ビン  $k$  について生成する (ステップ S4)。以下に正規化部 13 が生成する正規化信号値  $Z(i, k)$  の例を示す。

【0030】

[正規化信号値  $Z(i, k)$  の例]

本形態では正規化信号値  $Z(i, k)$  の一例として、 $S=2$  とし、基準センサ 20 - 1 に対応する周波数領域信号  $X(1, i, k)$  と、他方のセンサ 20 - 2 に対応する周波数領域信号  $X(2, i, k)$  とから信号到来方向を推定し、その信号到来方向推定値を正規化信号値  $Z(i, k)$  とする (正規化信号値  $Z(i, k)$  の例 1)。この例では、正規化部 13 は、以下の式 (3) (4) によって算出された信号到来方向推定値  $\tau(i, k)$  を正規化信号値  $Z(i, k)$  を算出する。なお、 $v$  は音速 (約 340 m/秒) を示し、 $d$  はセンサ間距離 (m) を示し、 $f$  は周波数インデックス  $k$  に対応する離散周波数  $f = f_s \cdot k/M$  を示し、 $\arg(\cdot)$  は  $\cdot$  の位相 (偏角) を示す。また、 $\tau(i, k)$  は信号源から各センサ 20 - 1, 2 までの信号到達時間差を示し、 $\theta(i, k)$  は信号到来方向推定値を示す。また、式 (4) によって算出される信号到来方向  $\theta(i, k)$  は、センサ 20 - 1, 2 を結ぶ線分の中点を通り、その線分と直交する方向を 0 radian とした角度 (radian) である。なお、このように算出された正規化信号値  $Z(i, k)$  は、周波数成分  $f$  が正規化され、周波数依存性が排除された値となる。

20

30

【0031】

【数3】

$$\tau(i, k) = \frac{\arg(X(2, i, k)/X(1, i, k))}{2 \cdot \pi \cdot f} \quad \dots (3)$$

$$\theta(i, k) = \arcsin\left(\frac{v \cdot \tau(i, k)}{d}\right) \quad \dots (4)$$

40

【0032】

また、前述の式 (3) で算出された信号到達時間差  $\tau(i, k)$  を正規化信号値  $Z(i, k)$  としてもよい (正規化信号値  $Z(i, k)$  の例 2)。なお、このように算出された正規化信号値  $Z(i, k)$  も周波数成分  $f$  が正規化され、周波数依存性が排除された値となる。

【0033】

また、周波数領域信号  $X(1, i, k)$  の位相に対する周波数領域信号  $X(1, i, k)$  の位相差  $\arg(X(2, i, k)/X(1, i, k))$  を正規化信号値  $Z(i, k)$  としてもよい (正規化信号値  $Z(i, k)$  の例 3)、周波数領域信号  $X(1, i, k)$  の位相と周波数領域信号  $X(1, i, k)$  の位相との差  $\arg(X(2, i, k)) - \arg(X(1, i, k))$  を正規化信号値  $Z(i, k)$  としてもよい (正規化信号値  $Z(i, k)$  の例 4)。さらに、周波数領域信号  $X(1, i, k)$  の振幅に対する周波数領域信号  $X(1, i, k)$  の振幅の比  $|X(2,$

50

$|X(1, i, k)|$  を正規化信号値  $Z(i, k)$  としてもよいし (正規化信号値  $Z(i, k)$  の例 5)、周波数領域信号  $X(1, i, k)$  のパワーに対する周波数領域信号  $X(1, i, k)$  のパワーの比  $|X(2, i, k)|^2 / |X(1, i, k)|^2$  を正規化信号値  $Z(i, k)$  としてもよい (正規化信号値  $Z(i, k)$  の例 6)。

【0034】

また、上記ではセンサの数が 2 つの場合を例示したが、センサの数が 3 以上の場合は、例えば以下のように、目的信号の到来方位角推定値  $(i, k)$  と仰角推定値  $(i, k)$  とを求め、それら 2 つの値を時間周波数ビン  $(i, k)$  に対する正規化信号値  $Z(i, k)$  としてもよい (正規化信号値  $Z(i, k)$  の例 7)。

【0035】

まず、各センサ  $20 - s$  ( $s=1, \dots, S$ ) の空間中の座標ベクトルを  $d_s = [x \text{ 座標}, y \text{ 座標}, z \text{ 座標}]$  とする。また、 $J$  ( $J(1, \dots, S)$ ) 番目のセンサ  $20 - J$  を基準センサとし、基準センサ  $20 - J$  と各センサ  $20 - s$  との距離ベクトル  $D$  を以下の式 (5) のように設定する。なお、 $[\cdot]^T$  はベクトル  $\cdot$  の転置を示す。

【0036】

$$D = [d_1 - d_J, d_2 - d_J, \dots, d_S - d_J]^T \quad \dots (5)$$

また、基準センサ  $20 - J$  と各センサ  $20 - s$  との信号到達時間差  $(s, i, k)$  を以下の式 (6) で求め、それらを要素とした信号到達時間差ベクトル  $\tau(i, k)$  を以下の式 (7) のように求める。

【0037】

【数 4】

$$\tau(s, i, k) = \frac{\arg(X(s, i, k) / X(J, i, k))}{2 \cdot \pi \cdot f} \quad \dots (6)$$

$$\tau(i, k) = [\tau(1, i, k), \tau(2, i, k), \dots, \tau(S, i, k)]^T \quad \dots (7)$$

上述の式 (5) ~ (7) には以下の式 (8) の関係が成り立ち、以下の式 (8) から目的信号の到来方位角推定値  $(i, k)$  と仰角推定値  $(i, k)$  とを求める。なお、式 (8) における  $D^{-1}$  はムーア・ペンローズ型一般化逆行列などの一般化逆行列である。また、目的信号の到来方位角とは  $x - y$  平面上の目的信号の到来方向を意味し、目的信号の仰角とは  $x - z$  平面上の目的信号の到来方向を意味する。また、 $y$  軸線方向が 0 radian である。

$$D^{-1} \cdot \tau(i, k) = [\cos \theta(i, k) \cos \phi(i, k), \sin \theta(i, k) \sin \phi(i, k), \sin \theta(i, k)]^T \quad \dots (8)$$

【0038】

また、上述の正規化信号値  $Z(i, k)$  の例 1 ~ 7 で例示した正規化信号値  $Z(i, k)$  を組み合わせ、時間周波数ビン  $(i, k)$  毎に 2 以上の正規化信号値  $Z(i, k)$  を算出する構成であってもよい (正規化信号値  $Z(i, k)$  の例 8)。例えば、位相差  $\arg(X(2, i, k) / X(1, i, k))$  と振幅の比  $|X(2, i, k)| / |X(1, i, k)|$  との組を時間周波数ビン  $(i, k)$  の正規化信号値  $Z(i, k)$  としてもよい。また、例えば  $S=3$  とし、位相差  $\arg(X(2, i, k) / X(1, i, k))$  と位相差  $\arg(X(3, i, k) / X(1, i, k))$  との組を時間周波数ビン  $(i, k)$  の正規化信号値  $Z(i, k)$  としてもよい。また、上述のように生成した値の写像を正規化信号値  $Z(i, k)$  としてもよい ([正規化信号値  $Z(i, k)$  の例] の説明終わり)。

【0039】

以上のようにステップ S4 では、正規化部 13 が上述のような正規化信号値  $Z(i, k)$  を生成し、出力する。

【0040】

正規化部 13 から出力された正規化信号値  $Z(i, k)$  と、周波数領域変換部 12 から出力された周波数領域信号とは、空間パワー分布推定部 14 に入力される。なお、空間パワー分布推定部 14 に入力される周波数領域信号は何れかのセンサ 20 に対応するものでよく、以下では周波数領域信号  $X(1, i, k)$  が空間パワー分布推定部 14 に入力される例を示す。しかし、その他のセンサ 20 に対応する周波数領域信号  $X(s, i, k)$  を用いてもよく、また、複

10

20

30

40

50

数のセンサに対応する周波数領域信号を平均したものをを用いてもよい。

【 0 0 4 1 】

空間パワー分布推定部 1 4 は、これらを用い、正規化信号値 $Z(i,k)$ に対応するフレーム $i$ の周波数領域信号 $X(1,i,k)$ の振幅の絶対値に対して単調増加する値を、当該正規化信号値 $Z(i,k)$ が属する到来方向区分 $r$ の当該フレーム $i$ での空間パワー分布値 $P(i,r)$ とし、出力する(ステップ S 5)。なお、到来方向区分 $r$ とは、正規化信号値 $Z(i,k)$ がとり得る全範囲を複数に分割した各区分を意味する。例えば、正規化信号値 $Z(i,k)$ がとり得る全範囲を均等に $R$ 個の区分に分割し、各区分を到来方向区分 $r$ ( $r \in \{1, \dots, R\}$ )としてもよいし、正規化信号値 $Z(i,k)$ がとり得る全範囲を不均等に $R$ 個の区分に分割し、各区分を到来方向区分 $r$ ( $r \in \{1, \dots, R\}$ )としてもよい。なお、 $r$ は各到来方向区分に対応するインデックスである。また、周波数領域信号 $X(1,i,k)$ の振幅の絶対値 $|X(1,i,k)|$ に対して単調増加する値としては、例えば、振幅の絶対値 $|X(1,i,k)|$ やパワー $|X(1,i,k)|^2$ を例示できる。以下にステップ S 5 の詳細を例示する。

10

【 0 0 4 2 】

[ ステップ S 5 の例 ]

図 4 ( a ) に示した例の場合、まず、空間パワー分布推定部 1 4 の時間周波数マスク生成部 1 4 a ( 図 2 ( a ) ) に各正規化信号値 $Z(i,k)$ が入力される。時間周波数マスク生成部 1 4 a は、各正規化信号値 $Z(i,k)$ を用い、特定の到来方向区分 $r$ に属する正規化信号値 $Z(i,k)$ の周波数ビン $k$ 及びフレーム $i$ に対してハイレベル値をとり、当該特定の到来方向区分 $r$ に属しない正規化信号値 $Z(i,k)$ の周波数ビン $k$ 及びフレーム $i$ に対してローレベル値( $<$ ハイレベル値)をとる、当該特定の到来方向区分 $r$ と周波数ビン $k$ とフレーム $i$ とを変数とする関数である時間周波数マスク $Mask(i,k,r)$ を生成して出力する(ステップ S 5 1)。時間周波数マスク $Mask(i,k,r)$ の一例は、以下のバイナリマスクである。

20

【 0 0 4 3 】

【 数 5 】

$$Mask(i,k,r) = \begin{cases} a & \text{if } Z(i,k) \notin \Theta_r \\ b & \text{if } Z(i,k) \in \Theta_r \end{cases} \dots (9)$$

【 0 0 4 4 】

なお、 $\Theta_r$ は、到来方向区分 $r$ に属する正規化信号値 $Z(i,k)$ の範囲を意味する。また、式(9)の例では、 $a$ がローレベル値に相当し、 $b$ がハイレベル値に相当する。到来方向区分 $r$ ごとに信号を抜き出す時間周波数マスクとする場合には、例えば、 $a=0, b=1$ とする。また、式(9)の時間周波数マスク $Mask(i,k,r)$ は、ローレベル値 $a$ からハイレベル値 $b$ への推移及びハイレベル値 $b$ からローレベル値 $a$ への推移が不連続なバイナリマスクであるが、周波数ビン $k$ やフレーム $i$ の変化に伴う出力値の推移が、ハイレベル値からローレベル値及びローレベル値からハイレベル値にかけてそれぞれ滑らかな時間周波数マスク $Mask(i,k,r)$ が生成されてもよい。

30

【 0 0 4 5 】

時間周波数マスク生成部 1 4 a で生成された時間周波数マスク $Mask(i,k,r)$ は、空間パワー分布推定部 1 4 のマスク適用部 1 4 b ( 図 2 ( a ) ) に入力され、マスク適用部 1 4 b にはさらに周波数領域変換部 1 2 から出力された周波数領域信号 $X(1,i,k)$ が入力される。マスク適用部 1 4 b は、周波数ビン $k$ とフレーム $i$ との組が同一である周波数領域信号 $X(1,i,k)$ と時間周波数マスク $Mask(i,k,r)$ との積 $Mask(i,k,r) \cdot X(1,i,k)$ に対して単調増加する値を用い、当該フレーム $i$ と当該到来方向区分 $r$ とに対応する空間パワー分布値 $P(i,r)$ を全ての到来方向区分 $r$ ( $r \in \{1, \dots, R\}$ )について生成して出力する(ステップ S 5 2)。例えば、マスク適用部 1 4 b は、周波数ビン $k$ とフレーム $i$ との組が同一である周波数領域信号 $X(1,i,k)$ と時間周波数マスク $Mask(i,k,r)$ との積 $Mask(i,k,r) \cdot X(1,i,k)$ に対して単調増加する値を、対応するフレーム $i$ と到来方向区分 $r$ との組が同一である集合ごとに周波数ビン $k$ について平均し、当該フレーム $i$ と当該到来方向区分 $r$ とに対応する空間パワー分布

40

50

値 $P(i, r)$ を生成する。より具体的には、マスク適用部 14b は、例えば、以下のようにして空間パワー分布値 $P(i, r)$ を生成する。

【0046】

【数6】

$$P(i, r) = \frac{1}{K} \sum_{k=1}^K |\text{Mask}(i, k, r) \cdot X(1, i, k)|^2 \quad \dots (10)$$

【0047】

また、マスク適用部 14b が、積 $\text{Mask}(i, k, r) \cdot X(1, i, k)$ に対して単調増加する値を全ての周波数ビン $k$ について平均するのではなく、一部の周波数ビンのみについて平均し、空間パワー分布値 $P(i, r)$ を求めてもよい。さらに、マスク適用部 14b が、何れかの周波数ビンに対応する積 $\text{Mask}(i, k, r) \cdot X(1, i, k)$ に対して単調増加する値を空間パワー分布値 $P(i, r)$ としてもよい。例えば、支配的な1つの周波数ビン $k'$ に対する

$$P(i, r) = |\text{Mask}(i, k', r) \cdot X(1, i, k')|^2 \quad \dots (11)$$

を空間パワー分布値 $P(i, r)$ としてもよい。

【0048】

また、時間周波数マスクを用いるのではなく、代わりに遅延和法（大賀寿朗，山崎芳男，金田豊，“音響システムとデジタル処理”，社団法人電子情報通信学会，1995，1995，pp.197-208）などによって得られた各フレーム $i$ の到来方向区分 $r$ 毎の空間スペクトルを空間パワー分布値 $P(i, r)$ としてもよい（[ステップS5の例]の説明終わり）。

【0049】

以上のように空間パワー分布推定部 14 から出力された空間パワー分布値 $P(i, r)$ は、雑音区間空間パワー分布推定部 15 に入力される。

【0050】

雑音区間空間パワー分布推定部 15 は、目的信号が存在しないと推定される雑音区間のフレームに対応する各到来方向区分 $r$ の空間パワー分布値 $P(i, r)$ とを用い、雑音区間の空間パワー分布値の推定値である雑音区間空間パワー分布値 $\lambda^N(i, r)$ を各到来方向区分 $r$ について生成し、出力する（ステップS6）。なお、この場合の雑音区間のフレームには、例えば、「観測信号の冒頭数フレームには目的信号が含まれない」と仮定し、観測信号の冒頭数フレームを用いる。また、音声信号や音楽信号の持つ周期性成分を検出する手段（例えば、特開第2008-64821号公報）などを用いて目的信号が含まれないことが推定されるフレームを推定し、それを雑音区間のフレームとして用いてもよい。以下にステップS6の詳細を例示する。

【0051】

[ステップS6の例]

雑音区間空間パワー分布値 $\lambda^N(i, r)$ の生成は、例えば、複数の雑音区間のフレームでの各到来方向区分の空間パワー分布値 $P(i, r)$ を、到来方向区分 $r$ ごとに当該雑音区間のフレームについて平均した値を、各到来方向区分 $r$ についての雑音区間空間パワー分布値 $\lambda^N(i, r)$ とすることによって行われる。すなわち、雑音区間空間パワー分布推定部 15 は、例えば、以下の式によって、全ての到来方向区分 $r$ について雑音区間空間パワー分布値 $\lambda^N(i, r)$ を生成する。

【0052】

【数7】

$$\lambda^N(i, r) = \frac{1}{\text{Num}(I^N)} \sum_{i' \in I^N} p(i', r) \quad \dots (12)$$

【0053】

ここで、 $I^N$ は雑音区間に属するフレームのフレームインデックスの集合を意味し、 $\text{Num}(\cdot)$ は、集合 $\cdot$ に含まれる要素数を出力する関数を意味する。なお、式(12)の場合、雑音区間空間パワー分布値 $\lambda^N(i, r)$ はフレーム $i$ に依存せず、到来方向区分 $r$ が同一であれ

10

20

30

40

50

ば各フレーム*i*について一定となる。

【0054】

また、観測信号の冒頭数フレームを雑音区間のフレームとして式(12)によって雑音区間空間パワー分布値 $\lambda^N(i,r)$ を求めた後、以下のように雑音区間空間パワー分布値 $\lambda^N(i,r)$ を各フレームで更新してもよい。

【0055】

【数8】

$$\lambda^N(i,r) = \begin{cases} \alpha \cdot \lambda^N(i-1,r) + (1-\alpha) \cdot P(i,r) & \text{if } i \in I^N \\ \lambda^N(i-1,r) & \text{if } i \in I^N \end{cases} \quad \dots (13)$$

10

【0056】

なお、 $\alpha$ は0以上1以下の定数であり、例えば、 $\alpha=0.8$ とする。また、各フレーム*i*が雑音区間に属するか否かの判断は、前述の音声信号や音楽信号の持つ周期性成分を検出する手段などを用いて行われる。なお、式(13)によって得られる雑音区間空間パワー分布値 $\lambda^N(i,r)$ も、複数の雑音区間のフレームでの各到来方向区分の空間パワー分布値 $P(i,r)$ を、到来方向区分*r*ごとに当該雑音区間のフレームについて平均(加重平均)した値に含まれる。

【0057】

また、何れかの雑音区間のフレームでの各到来方向区分の空間パワー分布値 $P(i,r)$ を、そのまま、各到来方向区分*r*についての雑音区間空間パワー分布値 $\lambda^N(i,r)$ として用いてもよい([ステップS6の例]の説明終わり)。

20

【0058】

以上のように雑音区間空間パワー分布推定部15から出力された雑音区間空間パワー分布値 $\lambda^N(i,r)$ は、事後信号対雑音比推定部16に入力され、事後信号対雑音比推定部16には、さらに、空間パワー分布推定部14から出力された空間パワー分布値 $P(i,r)$ が入力される。事後信号対雑音比推定部16は、各到来方向区分*r*の各フレーム*i*における空間パワー分布値 $P(i,r)$ と、各到来方向区分*r*の雑音区間空間パワー分布値 $\lambda^N(i,r)$ とを用い、当該空間パワー分布値 $P(i,r)$ と当該雑音区間空間パワー分布値 $\lambda^N(i,r)$ との比である事後信号対雑音比 $\beta(i,r)$ を、到来方向区分*r*ごとに各フレーム*i*について生成し、出力する(ステップS7)。例えば、事後信号対雑音比推定部16は、同一の到来方向区分*r*に対応する空間パワー分布値 $P(i,r)$ と雑音区間空間パワー分布値 $\lambda^N(i,r)$ との組ごとに、当該空間パワー分布値 $P(i,r)$ を当該雑音区間空間パワー分布値 $\lambda^N(i,r)$ で除算し、その演算結果 $P(i,r)/\lambda^N(i,r)$ を、当該空間パワー分布値 $P(i,r)$ の到来方向区分*r*及びフレーム*i*に対応する事後信号対雑音比 $\beta(i,r)$ として生成する。なお、前述のように雑音区間空間パワー分布値 $\lambda^N(i,r)$ は、全てのフレーム*i*について同一であってもよいし、フレーム*i*ごとに更新されるものであってもよい。

30

【0059】

$$\beta(i,r) = P(i,r) / \lambda^N(i,r) \quad \dots (14)$$

事後信号対雑音比推定部16から出力された事後信号対雑音比 $\beta(i,r)$ は、事前信号対雑音比推定部17に入力される。事前信号対雑音比推定部17は、事後信号対雑音比 $\beta(i,r)$ を少なくとも用い、雑音区間空間パワー分布値 $\lambda^N(i,r)$ と、空間パワー分布値 $P(i,r)$ の目的信号成分である目的信号空間パワー分布値 $S(i,r)$ との比である事前信号対雑音比 $\alpha(i,r)$ を、到来方向区分*r*ごとに各フレーム*i*について生成し、出力する(ステップS8)。以下にステップS8の詳細を例示する。

40

【0060】

[ステップS8の例]

事前信号対雑音比推定部17は、例えば、事後信号対雑音比 $\beta(i,r)$ から1を減じた値を、当該事後信号対雑音比 $\beta(i,r)$ の到来方向区分*r*及びフレーム*i*に対応する事前信号対雑音比 $\alpha(i,r) = S(i,r) / \lambda^N(i,r)$ として生成する。

50

【 0 0 6 1 】

(i, r) = (i, r) - 1 ... (15)

この場合には、少ない演算量で事前信号対雑音比 (i, r) が算出できる。また、各フレーム i の雑音区間空間パワー分布値 N(i, r) が実環境によく適合する場合、すなわち、

S(i, r) = P(i, r) - N(i, r) ... (16)

が実環境によく適合する場合にはよい推定となる。

【 0 0 6 2 】

また、事前信号対雑音比推定部 1 7 に、空間パワー分布推定部 1 4 から出力された空間パワー分布値 P(i, r) と、事後信号対雑音比推定部 1 6 から出力された事後信号対雑音比 (i, r) と、雑音区間空間パワー分布推定部 1 5 から出力された雑音区間空間パワー分布値 N(i, r) とが入力され、事前信号対雑音比推定部 1 7 が以下のように、到来方向区分 r 及びフレーム i に対応する事前信号対雑音比 (i, r) を生成してもよい。

10

【 0 0 6 3 】

(i-1, r) = (i-1, r) \* |S^(i-1, r)|^2 / (N(i-1, r) + (1 - (i-1, r)) \* max((i-1, r) - 1, 0)) ... (17)

S^(i-1, r) = {( (i-1, r) )^(1/2) / (i-1, r)} \* (1.5) \* exp(- (i-1, r) / 2) \* [(1 + (i-1, r)) \* I\_0((i-1, r) / 2) + (i-1, r) \* I\_1((i-1, r) / 2)] \* (P(i-1, r))^(1/2) ... (18)

(i-1, r) = (i-1, r) \* (i-1, r) / (1 + (i-1, r)) ... (19)

【 0 0 6 4 】

なお、max(・) は・の最大値を出力する関数であり、(・) はガンマ関数であり、I\_0(・) と I\_1(・) とは、それぞれ 0 次と 1 次の第 1 種修正ベッセル関数であり、は 0 以上 1 以下の定数である。また、前述のように雑音区間空間パワー分布値 N(i, r) は、全てのフレーム i について同一であってもよいし、フレーム i ごとに更新されるものであってもよい。また、式 ( 1 8 ) ( 1 9 ) は、周波数スペクトル上で雑音中の音声信号の振幅を最小二乗誤差推定する手法である「Ephraim, Y. and Malha, D., "Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator," IEEE Transactions on Acoustic, Speech, and Signal Processing, vol. ASSP-32, pp. 1109-1121, 1984」の式 ( 7 ) ~ ( 1 0 ) を、到来方向区分 r ごとの空間パワー分布推定に適用したものであり、|S^(i-1, r)|^2 が目的信号空間パワー分布値 S(i, r) の推定値となる ( [ ステップ S 8 の例 ] の説明終わり ) 。

20

【 0 0 6 5 】

以上のように事前信号対雑音比推定部 1 7 から出力された事前信号対雑音比 (i, r) は、尤度比計算部 1 8 に入力され、さらに、尤度比計算部 1 8 には、事後信号対雑音比推定部 1 6 から出力された事後信号対雑音比 (i, r) が入力される。

30

【 0 0 6 6 】

尤度比計算部 1 8 は、入力された事後信号対雑音比 (i, r) と事前信号対雑音比 (i, r) とを用い、到来方向区分 r に対応する観測信号 X(r) のフレーム i が雑音区間 (目的信号を含まない区間) である尤度 p\_i(X(r) | H\_0) と目的信号区間 (目的信号を含む区間) である尤度 p\_i(X(r) | H\_1) との比である尤度比 (i, r) を生成し、出力する (ステップ S 9) 。なお、H\_0 は、観測信号 X(r) のフレーム i が雑音区間であるとの仮定を意味し、H\_1 は、観測信号 X(r) のフレーム i が目的信号区間であるとの仮定を意味する。具体的には、本形態の尤度比計算部 1 8 は、例えば、以下のように尤度比 (i, r) を生成して出力する。

40

【 0 0 6 7 】

(i, r) = {1 / (1 + (i, r))} \* exp{( (i, r) \* (i, r) ) / (1 + (i, r))} ... (19)

以下にこの導出を行う。まず、尤度 p\_i(X(r) | H\_0) 及び p\_i(X(r) | H\_1) は、以下のように定義される (例えば、参考文献 1 「Sohn, J., Kim, N.-S., and Sung, W., "A Statistical Model-based Voice Activity Detection," IEEE Signal Processing Letters, Vol. 6, No. 1, pp. 1-3, 1999.」の式 ( 1 ) ( 2 ) 参照) 。

【 0 0 6 8 】

【数 9】

$$p_i(X(r) | H_0) = \frac{1}{\pi \cdot \lambda^N(i,r)} \exp\left(-\frac{P(i,r)}{\lambda^N(i,r)}\right) \quad \dots (20)$$

$$p_i(X(r) | H_1) = \frac{1}{\pi \cdot \{\lambda^N(i,r) + \lambda^S(i,r)\}} \exp\left(-\frac{P(i,r)}{\lambda^N(i,r) + \lambda^S(i,r)}\right) \quad \dots (21)$$

【0069】

10

これらの尤度の比をとり、 $(i,r) = S(i,r) / N(i,r)$ と  $(i,r) = P(i,r) / N(i,r)$ の関係をを用いることで、以下のように尤度比  $(i,r)$  が得られる。

【0070】

$$(i,r) = p_i(X(r) | H_1) / p_i(X(r) | H_0) \\ = \{1 / (1 + (i,r))\} \cdot \exp\{( (i,r) \cdot (i,r) ) / (1 + (i,r))\}$$

このように、本形態では、正確な  $S(i,r)$  を直接求めることができないため、以上のように事後信号対雑音比  $(i,r)$  と事前信号対雑音比  $(i,r)$  とを求め、それらを用いて尤度比  $(i,r)$  を計算する。

【0071】

20

尤度比計算部 18 から出力された尤度比  $(i,r)$  は判定部 19 に入力される。判定部 19 は、尤度比  $(i,r)$  を指標とし、各フレーム  $i$  が目的信号区間であるか否かを判定し、その判定結果を出力する (ステップ S 10)。以下にステップ S 10 の詳細を例示する。

【0072】

[ステップ S 10 の例]

図 4 (b) の例では、判定部 19 は、尤度比  $(i,r)$  又はその関数値をフレーム  $i$  ごとに到来方向区分  $r$  について平均し、その演算結果を指標として、各フレーム  $i$  が目的信号区間であるか否かを判定する。

【0073】

30

この例ではまず、判定部 19 の平均算出部 19 a (図 2 (a)) に尤度比  $(i,r)$  が入力され、平均算出部 19 a は、尤度比  $(i,r)$  又はその関数値をフレーム  $i$  ごとに到来方向区分  $r$  について平均した平均値  $Ave(i)$  を生成し、出力する (ステップ S 101)。平均値  $Ave(i)$  の一例は、尤度比  $(i,r)$  に対する対数尤度比  $\log (i,r)$  をフレーム  $i$  ごとに全到来方向区分  $r$  ( $r = \{1, \dots, R\}$ ) について平均した以下の平均対数尤度比  $\log (i)$  である。なお、式 (22) の対数は自然対数である。

【0074】

【数 10】

$$Ave(i) = \log \Lambda(i) = \frac{1}{R} \sum_{r=1}^R \log \Lambda(i,r) \quad \dots (22)$$

40

【0075】

また、平均算出部 19 a が、各フレーム  $i$  の前後  $W$  フレーム ( $i-W, \dots, i, \dots, i+W$ ) において尤度比  $(i,r)$  を乗算し、その乗算値又はその関数値を到来方向区分  $r$  について平均した平均値  $Ave(i)$  を生成し、出力してもよい。すなわち、 $(i)$  の代わりに以下の  $'(i)$  を用いてもよい。

【0076】

## 【数 1 1】

$$\Lambda'(i) = \prod_{p=i-W}^{i+W} \Lambda(p) \quad \dots (23)$$

## 【0077】

平均算出部 19 a から出力された平均値  $Ave(i)$  は、判定部 19 の閾値比較部 19 b に入力され、閾値比較部 19 b は、各フレーム  $i$  について、平均値  $Ave(i)$  が閾値  $th$  よりも大きいか否かを判断する (ステップ S 102)。ここで、平均値  $Ave(i) > th$  である場合 (「平均値  $Ave(i) > th$  である場合」としてもよい)、そのフレーム  $i$  は目的信号区間であるとして 1 を出力し (ステップ S 103)、平均値  $Ave(i) < th$  である場合 (「平均値  $Ave(i) < th$  である場合」としてもよい)、そのフレーム  $i$  は雑音区間であるとして 0 を出力する (ステップ S 104)。なお、閾値  $th$  は、平均値  $Ave(i)$  の時間長平均 (複数のフレーム  $i$  に対する平均) や分散などの統計量を用いて設定されてもよいし、 $th=0.2$  などの固定値を事前に設定しておいてもよい。分散などの統計量を用いて閾値  $th$  を設定する場合の一例としては、目的信号が存在しないと推定されるフレームを判定対象のフレームについて平均値  $Ave(i)$  を求め、それらの平均値から所定のマージンを設けた値を閾値  $th$  とする方法がある。

## 【0078】

[ステップ S 10 の変形例]

次に、ステップ S 10 の変形例を示す。

第 1 の変形例では、図 2 (a) の判定部 19 の代わりに、図 5 (a) の判定部 19 を用いる。この変形例の場合、判定部 19 の閾値比較部 19 c は、入力された尤度比  $(i, r)$  と、到来方向区分  $r$  ごとに設定された閾値  $th(r)$  とを到来方向区分  $r$  ごとに比較し、到来方向区分  $r$  ごとにそのフレーム  $i$  が目的信号区間であるか否かを判定し、その判定結果を出力する。例えば、尤度比  $(i, r) > th(r)$  である場合 (「尤度比  $(i, r) > th(r)$  である場合」としてもよい)、そのフレーム  $i$  はその到来方向区分  $r$  について目的信号区間であるとして 1 を出力し、尤度比  $(i, r) < th(r)$  である場合 (「尤度比  $(i, r) < th(r)$  である場合」としてもよい)、そのフレーム  $i$  はその到来方向区分  $r$  について雑音区間であるとして 0 を出力する。なお、閾値  $th(r)$  は、尤度比  $(i, r)$  の時間長平均 (複数のフレーム  $i$  に対する平均) や分散などの統計量を用いて設定されてもよいし、 $th=0.2$  などの固定値を事前に設定しておいてもよい。

## 【0079】

第 2 の変形例では、図 2 (a) の判定部 19 の代わりに、図 5 (b) の判定部 19 を用いる。この変形例の場合、事前学習されたフレームの音響特徴量 ( $P(i, r)$  や  $(i, r)$ ) と当該フレームが目的信号区間であるか否かの判定結果との関係を用いたパターン認識により、偏在性指標値算出部で算出された偏在性指標値に対応するフレームが目的信号区間に対応するか否かを判定する。この例の場合、例えば、パラメータ学習部 19 d に、フレームの音響特徴量 ( $P(i, r)$  や  $(i, r)$ ) と当該フレームが目的信号区間であるか否かの判定結果との組からなる学習サンプルが入力され、パラメータ学習部 19 d がパターン認識学習を行い、モデルパラメータを求める。そして、パターン認識部 19 e にこのパラメータと判定対象のフレームの音響特徴量を入力し、パターン認識によってそのフレームが目的信号区間であるか否かを判定する。なお、パターン認識技術には、例えば、公知のサポートベクターマシン (津田宏治, “サポートベクターマシンとは何か”, 電子情報通信学会誌, 2000: 460 ~ 466 頁) や、隠れマルコフモデル (北研二, 中村哲, 永田昌明, “音声言語処理”, 森出版株式会社, 1996: 57 ~ 90 頁) 等を利用する。

## 【0080】

その他、判定部 19 が、目的信号区間であるか否かの判定結果を出力するのではなく、前述の尤度比  $(i, r)$  や平均対数尤度比  $\log (i)$  や  $(i)$  そのものや、それを 0 ~ 1 の収まる値に変換した  $(i) / (1 + (i))$  や  $\log (i) / (1 + \log (i))$  や  $(i) / (1 + (i))$  や、

尤度 $p_i(X(r)|H_0)$ や尤度 $p_i(X(r)|H_1)$ などを出力してもよい。

【0081】

<実験結果>

本形態の効果を示すために、複数のマイクロホンを用いて観測した、音声信号を雑音信号が混在する音響信号を本形態の目的信号区間推定方法によって分析し、得られた尤度比 $(i, r)$ (尤度比計算部18の出力)を用いて目的信号区間(この実験では音声信号区間)を検出する実験例を示す。使用した音響信号データは大学の研究室内で学生がポスターを用いて自身の研究の発表を行っている発話を収録した信号で、サンプリング周波数16kHz、量子化ビット数16ビットで離散サンプリングされたものを用いた。ここでの検出対象となる目的信号は発表者又は聞き手の発生した音声信号とする。この収録には、発表者1名と聞き手2名が参加した。収録された信号には、研究室内にある複数の計算機による雑音や、部屋外かの雑音が常に混入している。収録には8本の全指向性マイクロホン(センサ20-1~8)を用いており、これらは図6に示す間隔で配置されていた。このようにして収録された音響信号に対し、1フレームの時間量を32ms(512サンプル点)とし、16ms(256サンプル点)ごとにフレームの始点を移動させて、本形態の目的信号区間推定方法を適用し、各フレームで尤度比を推定した。また、正規化信号値 $Z(i, k)$ として信号到来方向 $(i, k)$ を用い、0~180度を1度毎に180分割した各区分を到来方向区分 $r$ ( $r \in \{1, \dots, 180\}$ )とした。

10

【0082】

図7(a)は、このような条件で尤度比計算部18から出力された尤度比 $(i, r)$ を例示するグラフであり、図7(b)は、平均算出部19aから出力された平均対数尤度比 $(i)$ を例示するグラフである。ここで、図7(a)(b)の横軸は各フレーム $i$ の中心時間[秒]を示す。また、図7(a)の縦軸は信号到来方向である方位角[度]を示し、図7(b)の縦軸は平均対数尤度比 $(i)$ を示す。また、図7(a)では、色が白に近いほど尤度比 $(i, r)$ が高い値であることを示し、色が黒に近いほど尤度比 $(i, r)$ が低い値であることを示している。図7(a)(b)に示すように、本形態の目的信号区間推定方法によって得られた尤度比 $(i, r)$ は、話者の位置と発話の有無に応じて変化する。

20

【0083】

また、本形態の効果を示すために、約15分程度のポスター発表の全データを用いてDER(Diarization Error Rate)による評価を行った。DERは、目的信号の無い箇所を誤って目的信号を検出した時間長FST(False-alarm Speech Time)及び目的信号の有る箇所を誤って目的信号無しとした時間長MST(Missed Speech Time)を統合して評価するもので、 $DER = (FST + MST) / \text{データの総時間長} \times 100(\%)$ として求められる。このDERが低いほど目的信号区間推定精度が高いことを意味する。なお、比較として、本形態の手法(本形態)によって目的信号区間を推定した場合の他、参考文献1の手法(従来手法A)によって目的信号区間を推定した場合と、非特許文献5の手法(従来手法B)によって目的信号区間を推定した場合とのDERも測定した。

30

【0084】

図8は、本形態と従来手法Aと従来手法Bでそれぞれ推定された目的信号区間のDERを示すグラフである。この図に示すように、本形態の手法が、従来手法Aと従来手法Bに比べ、方向性雑音や拡散性雑音のある環境下で、より高精度に目的信号区間を推定できる手法であることが分かる。

40

【0085】

〔変形例等〕

なお、本発明は上述の実施形態に限定されるものではない。例えば、上記の実施形態では、尤度比計算部18が、事後信号対雑音比推定部16で生成された事後信号対雑音比 $(i, r)$ と、事前信号対雑音比推定部17で生成された事前信号対雑音比 $(i, r)$ とを用い、尤度比 $(i, r)$ を計算して出力し、判定部19はそれを指標として目的信号区間であるか否かを判定していた。しかし、尤度比計算部18が、尤度比 $(i, r)$ ではなく尤度比 $(i, r)$ の関数値を出力し、判定部19がそれを指標として目的信号区間であるか否かを判定す

50

る構成でもよい。この場合には、判定部 19 での判定を容易にするため、尤度比  $(i, r)$  の関数値は、尤度比  $(i, r)$  に対して単調増加するものか、単調減少するものであることが望ましい。また、尤度比  $(i, r)$  の関数値の一例は、尤度比  $(i, r)$  の対数尤度比  $\log(i, r)$  である。尤度比計算部 18 が対数尤度比  $\log(i, r)$  を出力する例の場合、例えば、平均算出部 19 a (図 2 (a)) が、入力された対数尤度比  $\log(i, r)$  をフレーム  $i$  ごとに到来方向区分  $r$  について平均した平均値  $\text{Ave}(i)$  を生成して出力し、その後、前述したステップ S 102 ~ S 104 の処理を実行すればよい。

【0086】

また、事前信号対雑音比  $(i, r) = (i, r) - 1$  とする場合には、事前信号対雑音比推定部 17 で事前信号対雑音比  $(i, r)$  を計算することなく、尤度比計算部 18 が、事後信号対雑音比推定部 16 で生成された事後信号対雑音比  $(i, r)$  のみを用い、以下のように対数尤度比  $\log(i, r)$  を生成し、出力してもよい。この場合には、事前信号対雑音比推定部 17 が不要となる。

10

【0087】

$$\log(i, r) = (i, r) - \log(i, r) - 1 \quad \dots(24)$$

また、上述の各種の処理は、記載に従って時系列に実行されるのみならず、処理を実行する装置の処理能力あるいは必要に応じて並列的あるいは個別に実行されてもよい。また、信号区間推定装置 10 がサンプリング部 30 を包含する構成であってもよいし、信号区間推定装置 10 の機能を複数のコンピュータで分散処理する構成であってもよい。また、上述の実施形態では、マイクロホンで観測された音響信号を処理対象とし、目的信号区間を推定することとしたが、音響信号以外の信号、例えば、超音波、電磁波等を処理対象とし、目的信号区間を推定してもよい。その他、本発明の趣旨を逸脱しない範囲で適宜変更が可能であることはいうまでもない。

20

【0088】

また、上述の構成をコンピュータによって実現する場合、各装置が有すべき機能の処理内容はプログラムによって記述される。そして、このプログラムをコンピュータで実行することにより、上記処理機能がコンピュータ上で実現される。

【0089】

この処理内容を記述したプログラムは、コンピュータで読み取り可能な記録媒体に記録しておくことができる。コンピュータで読み取り可能な記録媒体としては、例えば、磁気記録装置、光ディスク、光磁気記録媒体、半導体メモリ等のようなものでもよい。

30

【0090】

また、このプログラムの流通は、例えば、そのプログラムを記録した DVD、CD-ROM 等の可搬型記録媒体を販売、譲渡、貸与等することによって行う。さらに、このプログラムをサーバコンピュータの記憶装置に格納しておき、ネットワークを介して、サーバコンピュータから他のコンピュータにそのプログラムを転送することにより、このプログラムを流通させる構成としてもよい。

【0091】

このようなプログラムを実行するコンピュータは、例えば、まず、可搬型記録媒体に記録されたプログラムもしくはサーバコンピュータから転送されたプログラムを、一旦、自己の記憶装置に格納する。そして、処理の実行時、このコンピュータは、自己の記録媒体に格納されたプログラムを読み取り、読み取ったプログラムに従った処理を実行する。また、このプログラムの別の実行形態として、コンピュータが可搬型記録媒体から直接プログラムを読み取り、そのプログラムに従った処理を実行することとしてもよく、さらに、このコンピュータにサーバコンピュータからプログラムが転送されるたびに、逐次、受け取ったプログラムに従った処理を実行することとしてもよい。また、サーバコンピュータから、このコンピュータへのプログラムの転送は行わず、その実行指示と結果取得のみによって処理機能を実現する、いわゆる ASP (Application Service Provider) 型のサービスによって、上述の処理を実行する構成としてもよい。なお、本形態におけるプログラムには、電子計算機による処理の用に供する情報であってプログラムに準ずるもの(コン

40

50

コンピュータに対する直接の指令ではないがコンピュータの処理を規定する性質を有するデータ等)を含むものとする。

【0092】

また、この形態では、コンピュータ上で所定のプログラムを実行させることにより、本装置を構成することとしたが、これらの処理内容の少なくとも一部をハードウェア的に実現することとしてもよい

【産業上の利用可能性】

【0093】

本発明の利用分野としては、例えば、音声信号や音楽信号などの目的信号が雑音信号とともに観測される環境においてなされる、目的信号の符号化、雑音信号の抑圧、残響除去、自動音声認識などの音響信号処理分野を例示できる。もちろん、音響信号以外の信号処理に本発明を適用してもかまわない。

【図面の簡単な説明】

【0094】

【図1】図1は、本形態の目的信号区間推定装置の全体構成を例示したブロック図である。

【図2】図2(a)は、本形態の空間パワー分布推定部の詳細を例示するブロック図であり、図2(b)は、本形態の判定部の詳細を例示するブロック図である。

【図3】図3は、本形態の目的信号区間推定方法を説明するためのフローチャートである。

【図4】図4(a)は、ステップS5の詳細を例示するためのフローチャートであり、図4(b)は、ステップS10の詳細を例示するためのフローチャートである。

【図5】図5(a)(b)は、判定部の変形例を示すブロック図である。

【図6】図6は、実験におけるマイクロホンの配置を説明するための図である。

【図7】図7(a)は、このような条件で尤度比計算部から出力された尤度比  $(i, r)$  を例示するグラフであり、図7(b)は、平均算出部から出力された平均対数尤度比  $(i)$  を例示するグラフである。

【図8】図8は、本形態と従来手法Aと従来手法Bでそれぞれ推定された目的信号区間のDERを示すグラフである。

【符号の説明】

【0095】

10 信号区間推定装置

10

20

30

【図1】

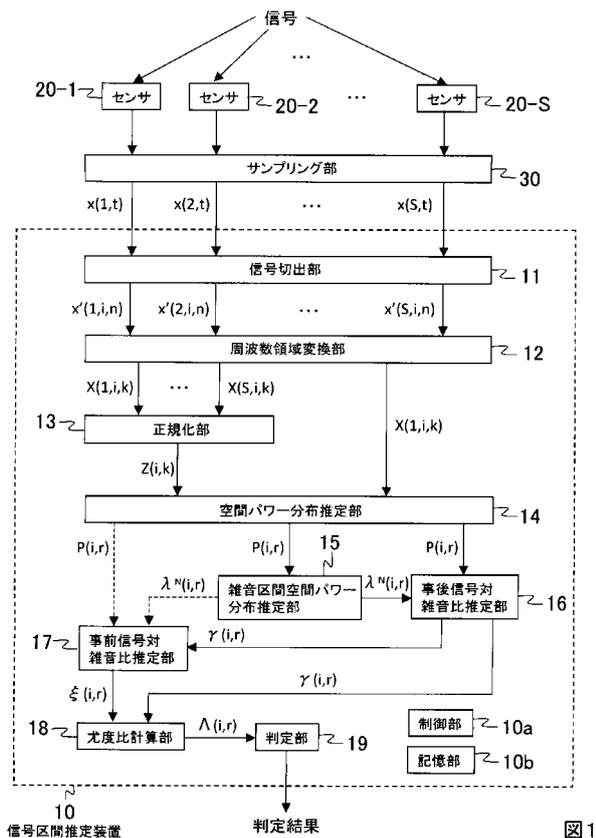


図1

【図2】

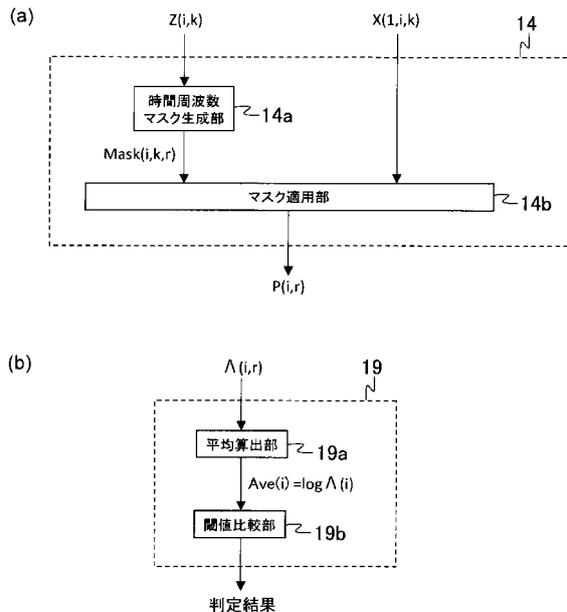


図2

【図3】

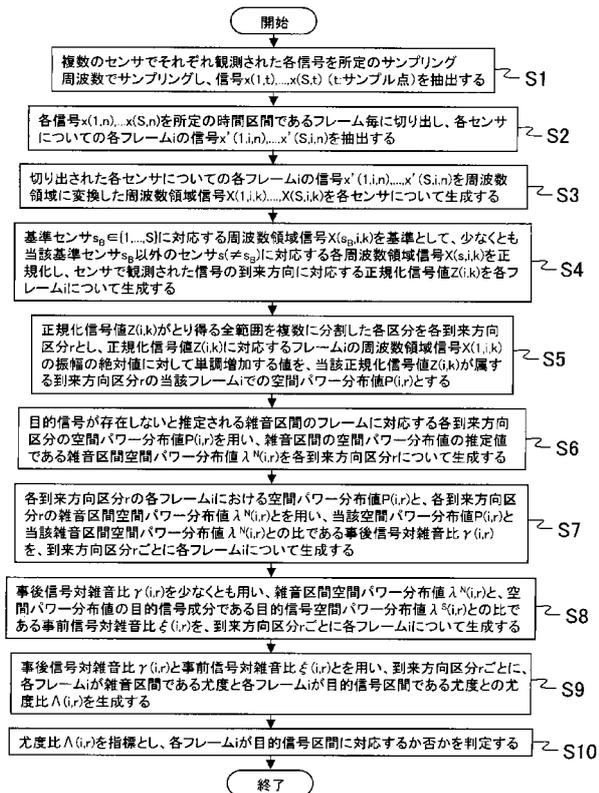


図3

【図4】

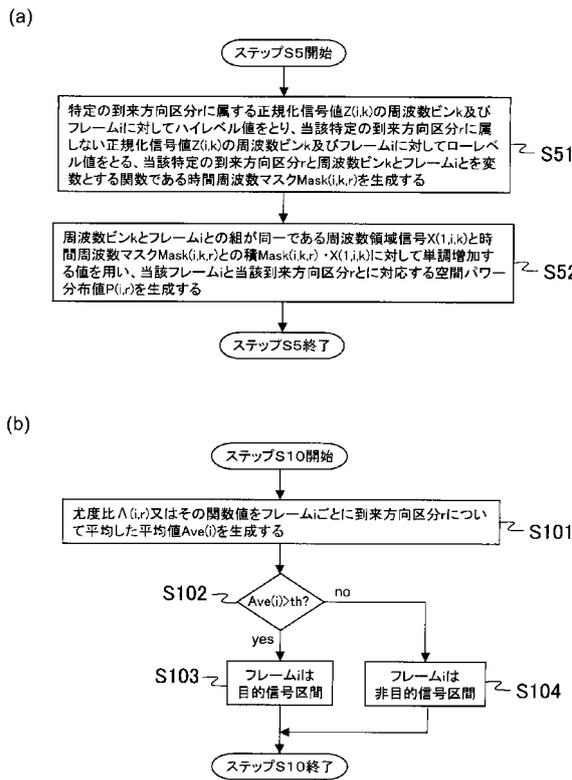
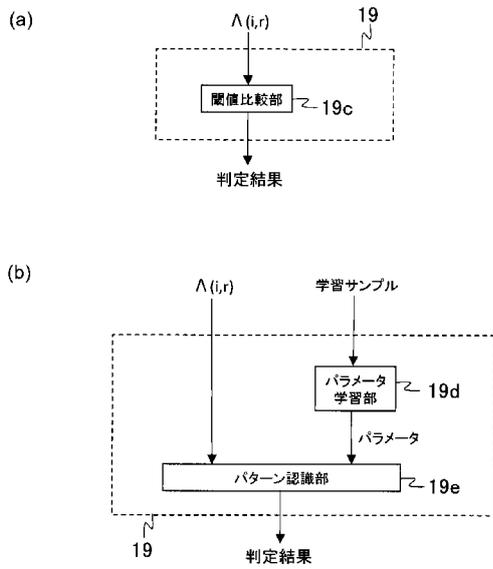


図4

【 図 5 】



【 図 6 】

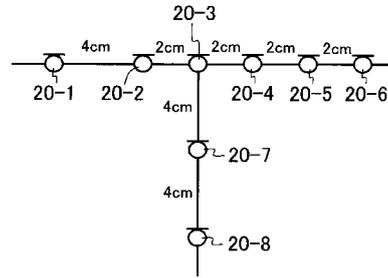


図6

図5

【 図 8 】

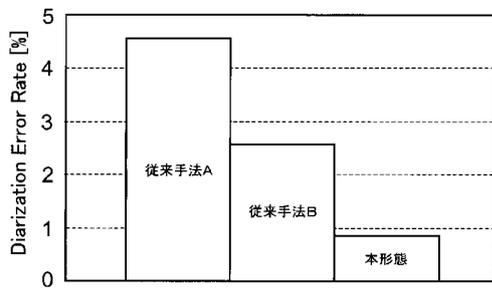
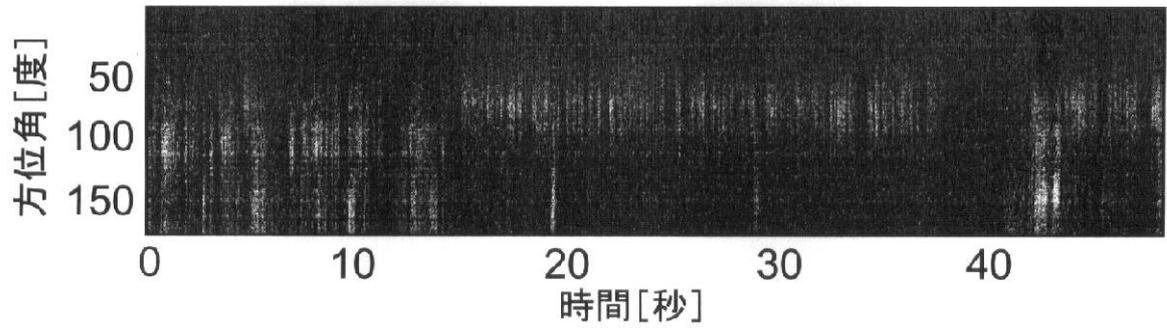


図8

【 図 7 】

(a)



(b)

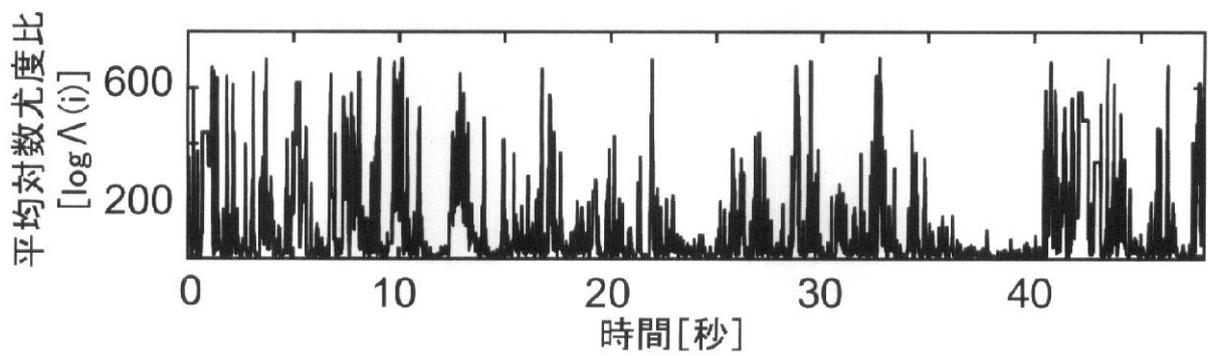


図7

---

フロントページの続き

(51)Int.Cl. F I テーマコード(参考)  
G 1 0 L 15/20 3 7 0 E

(72)発明者 荒木 章子  
東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内

(72)発明者 河原 達也  
京都府京都市左京区吉田本町 国立大学法人京都大学大学院情報学研究科内  
Fターム(参考) 5D015 EE05