

DOCUMENT 2/2
DOCUMENT NUMBER
@: unavailable

DETAIL **JAPANESE** **LEGAL STATUS**

- 1. [JP.2008-026625.A](#)
- 2. [JP.2007-248975.A](#)

PATENT ABSTRACTS OF JAPAN

(11)Publication number : **2007-248975**
(43)Date of publication of application : **27. 09. 2007**

(51) Int. Cl. **G10L 21/02 (2006. 01)**
H04R 1/40 (2006. 01)

(21)Application number : **2006-074423** (71)Applicant : **DOSHISHA
UNIV NACIONAL DE ENTRE RIOS
UNIV NACIONAL DEL LITORAL**
(22)Date of filing : **17. 03. 2006** (72)Inventor : **DI PERSIA LEANDRO EZEQUIEL
MILONE DIEGO
YANAGIDA MASUZO**

(54) PERMUTATION-FREE BLIND SOUND SOURCE SEPARATING METHOD AND DEVICE

(57)Abstract:
PROBLEM TO BE SOLVED: To provide a frequency range BSS (Blind Source Separation) technique which causes no permutation problem.
SOLUTION: When sounds output from a plurality of sound sources are received by a plurality of sound receiving devices and a target sound is extracted, a signal from a sound source is extracted first, and the extracted signal in a time range is converted into a signal in a frequency range by analytic frame. Then time variations of Fourier transformations of frequencies by frequency bin are read out of frequency spectra of all analytic frames and time variations in the value of Fourier transformation at respective frequencies are connected as to all frequency sampling points. Then the connected frequency signal is separated into signals of the frequency ranges of the respective sound sources. Then the separated signals in the frequency ranges are put back into signals in time ranges, which are output.



sinc(i) Research Center for Signals, Systems and Computational Intelligence (fich.unl.edu.ar/sinc)
L. Di Persia, D. H. Milone & Masuzo Yanagida; "Permutation-free Blind Sound Source Separating Method and Device"
Japan Patent Office, No. 2007-24897, 2007.

BACK NEXT
MENU SEARCH
NUMBER LIST HELP

(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2007-248975

(P2007-248975A)

(43) 公開日 平成19年9月27日 (2007.9.27)

(51) Int. Cl.		F I	テーマコード (参考)
G10L 21/02 (2006.01)		G10L 21/02 204	
H04R 1/40 (2006.01)		G10L 21/02 201D	
		H04R 1/40 320Z	

審査請求 未請求 請求項の数 2 O L (全 14 頁)

(21) 出願番号 特願2006-74423 (P2006-74423)
 (22) 出願日 平成18年3月17日 (2006.3.17)

(71) 出願人 503027931
 学校法人同志社
 京都府京都市上京区今出川通烏丸東入玄武町601

(71) 出願人 506092444
 ウニベルシダー・ナシオナル・デ・エントレ・リオス
 アルゼンチン国エントレリオス州コンセプション・デル・ウルグアイ市エバ・ペロン通り24

(71) 出願人 506092466
 ウニベルシダー・ナシオナル・デル・リトラール
 アルゼンチン国サンタフェ州サンタフェ市ペレグリーニ大通り2750番地
 最終頁に続く

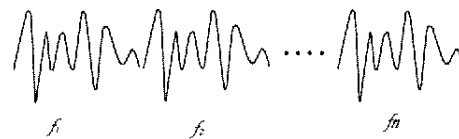
(54) 【発明の名称】 パーミュテーションフリー・ブライント音源分離方法および装置

(57) 【要約】

【課題】 パーミュテーション問題が起きないように周波数領域BSS手法を提案する。

【解決手段】 複数の音源から出力された音声を複数の受音装置で受音して目的の音声を抽出する場合、まず、音源から信号を取り出し、この取り出された時間領域の信号を分析フレーム毎に周波数領域の信号に変換する。次に、全分析フレームにおける周波数スペクトルから、各周波数ビン毎の周波数のフーリエ変換の時間変化を読み取り、各周波数でのフーリエ変換の値の時間変化を全周波数標本点について連結する。その後、この連結された周波数信号を各音源における周波数領域の信号に分離する。そして、分離された周波数領域での信号を時間領域の信号に戻し、この時間領域の信号を出力する

【選択図】 図5



【特許請求の範囲】

【請求項1】

複数の音源から出力された音声を複数の受信装置で受信し、目的の音声を抽出するパーミュテーションフリー・ブラインド音源分離方法において、
 受信された時間領域の信号を分析フレーム毎に周波数領域の信号に変換するステップと、
 全分析フレームにおける周波数スペクトルから、各周波数ビン毎の周波数のフーリエ変換の時間変化を読み取るステップと、
 各周波数でのフーリエ変換の値の時間変化を全周波数標本点について連結するステップと、
 連結された周波数信号を、各音源における周波数領域の信号に分離するステップと、
 分離された周波数領域の信号を時間領域の信号に戻すステップと、
 時間領域の信号に戻された信号を出力するステップとを備えたことを特徴とするパーミュテーションフリー・ブラインド音源分離方法。

【請求項2】

複数の音源から出力された音声を複数の受信装置で受信し、目的の音声を抽出するパーミュテーションフリー・ブラインド音源分離装置において、
 受信装置で受信した時間領域の信号を分析フレーム毎に周波数領域の信号に変換する周波数領域変換部と、
 全分析フレームにおける周波数スペクトルから、各周波数ビン毎の周波数のフーリエ変換の時間変化を読み取る読取部と、
 各周波数でのフーリエ変換の値の時間変化を全周波数標本点について連結する連結部と、
 連結された周波数信号を、各音源における周波数領域の信号に分離する分離部と、
 分離された周波数領域での信号を時間領域の信号に戻す時間領域変換部とを備え、当該時間領域変換部によって変換された信号を出力するようにしたことを特徴とするパーミュテーションフリー・ブラインド音源分離装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、ブラインド分離手法を用いた音声分離方法においてパーミュテーション問題を考慮する必要のない手法に関するものである。

【背景技術】

【0002】

現在の音声認識システムは、マイクに接近した位置からの音声に対してはそれなりに高い認識精度を有する一方で、マイクから離れた位置からの音声に対しては、周囲の雑音や部屋の残響の影響を受けて認識率は著しく低下してしまう（非特許文献1）。実環境で音声認識を行うためには音声認識システムそのものの認識能力の向上とともに、それに入力される音声に対して何らかの処理を行って認識しやすい形にしておくことが有効であると考えられる。その方法の一つがブラインド音源分離（BSS: Blind Source Separation）である。BSSとは、複数のマイクロホン線形に混合された信号が入力された時に、音源信号や混合過程を知ることなく、観測信号のみから音源信号を推定し分離する技術である。マイクロホン数を M 、音源数を N とすると、マイクロホン M_m での観測信号は、

【0003】

【数1】

$$x_m(t) = \sum_{n=1}^N h_{mn}(t) * s_n(t)$$

h_{mn} : 音源からマイクロホン M_n
へのインパルス応答

$s_n(t)$: 音源 s_n の信号

* : 畳み込み積分
(コンボリューション)

と表される。数1において、観測可能な要素は x_m だけで、 h_{mn} と s_n は未知である。BSSはこの状況において、 h_{mn} と s_n を推定する問題である。しかし、時間領域のBSSでは、計算量が膨大になり、現実的な処理時間で分離フィルタを計算するのが困難であることが指摘されている。

【0004】

そこで、一般に、このフィルタを直接求めずにその周波数特性を求める周波数領域BSSを用いることが多い。マイクロホン数 M 、音源数 N の混合モデルを求めるために、数1の両辺をフーリエ変換すると、 $X_m(f) = \sum_n H_{mn}(f) S_n(f)$ と表される。 $X_m(f)$ 、 $H_{mn}(f)$ 、 $S_n(f)$ は、それぞれ $x_m(t)$ 、 $h_{mn}(t)$ 、 $s_n(t)$ のフーリエ変換である。そして、 $m=1 \sim M$ についてまとめると、周波数領域BSSの混合モデルは、

【0005】

【数2】

$$X(f) = H(f)S(f)$$

【0006】

と表される。ただし、 $X(f) = (X_1(f), \dots, X_M(f))^T$ は観測ベクトル、 $S(f) = (S_1(f), \dots, S_N(f))^T$ は音源信号ベクトルそして、 $H(f)$ は混合行列である。 f は周波数である。

【0007】

BSSを解く手法として、各音源信号の独立性仮定したICAが一般的に用いられる。ICAには時間領域の方法と周波数領域の方法がある。周波数領域ICAは、畳み込みを瞬時混合の問題に置き換えて解くことができるため、現在、実環境で使える手法として盛んに用いられている。

【0008】

周波数領域ICAは、 x_m の短時間フーリエ変換により得られた観測信号ベクトル $X(f)$ に対して、各周波数ビンにおいて学習された $W(f)$ を用いて分離を行う。ここで、 $Y(f) = (Y_1(f), \dots, Y_N(f))^T$ を分離信号ベクトルとすると分離過程は、

【0009】

【数3】

$$Y(f) = W(f)X(f)$$

と表される。

【0010】

ICAは音源信号の独立性を最大にするという基準に基づいて分離を行っており、 $W(f)$ の行が入れ替わったとしても $Y(f)$ の独立性は保たれるので、任意の周波数 f_1 と f_2 において、 $Y_n(f_1)$ と $Y_n(f_2)$ が必ずしも同じ音源に対応しているとは限らない。従って、 $W(f)$ の行の並びを $Y_n(f_1)$ と $Y_n(f_2)$ が同じ音源に属するように並べ替える「パーミュテーション問題」を

解決する必要がある。

【0011】

パーミュテーション問題に対する従来の解決手法として、各周波数ビンでの指向特性を調べ、雑音源の方向を推定し、方向推定結果を基に並び替えを行う方法がある（非特許文献5）。しかし、すべての周波数において理想的なビームあるいは死角が形成できるわけではなく、特に低周波数における方向推定性能が低いと指摘されている。

【0012】

そこで、パーミュテーション問題への対処法として各周波数ビンでの相関を用いた方法（非特許文献2）が提案されている。各周波数ビンのエンベロープは、同じ音源については、特に近傍の周波数で高い相関を持つと考えるのが妥当である。そのため、周波数差が比較的小さい範囲内で相関の和が最大になるように決定していくことが考えられるが、ある周波数で間違えると以降の周波数でも間違えてしまうため、安定性に欠けるといわれている。また、ICAによる分離性能が比較的良好と思われる周波数ビンからパーミュテーションを決定していく方法が提案されている（非特許文献2）が、この方法では、離れた周波数に対してもエンベロープの相関が高いことが仮定されているため適切ではない。

【0013】

上述の欠点を補うために、方向推定と相関を統合した手法が提案されている（非特許文献6）。この方法では、まず指向特性を各周波数ビンに対して求め、確実に推定できる周波数に対してパーミュテーション問題を解決する。次に、未定の周波数に対して周波数差が小さく、かつパーミュテーションが決定されている周波数との相関の和が最大になるように決定していく。しかし、この方法でも残響が強い場合に不安定になることが報告されている（非特許文献7）。

【0014】

残響が強い場合に不安定になる原因として、低域での方向推定の困難さが挙げられている（非特許文献7）。そして、この問題を解決するために、調波構造を用いた方法が提案されている（非特許文献7）。音声は基本周波数の整数倍のところで相関が高くなる性質を持っており、調波性を用いることにより、より確実にパーミュテーション問題を解決することが可能となる。

【非特許文献1】中村哲：“実音響環境に頑健な音声認識を目指して”，電子情報通信学会技術報告，SP2002-12，pp.31-36，2002.

【非特許文献2】T.W.Lee：“Independent Component Analysis”，Kluwer，1998.

【非特許文献3】S.Ikeda, and N.Murata：“A method of ICA in time-frequency domain”，Proc WS on Independent Component Analysis and Blind Signal Separation (ICA' 99)，pp.365-371, Aussiaos, France, Jan., 1999.

【非特許文献4】A.Hyvarinen, J.Karhunen, and E.Oja：“Independent Component Analysis”，John Wiley, New York, 2001.

【非特許文献5】S. Kurita, H. Saruwatari, S. Kajita, K. Takeda, and F. Itakura：“Evaluation of blind signal separation method using directivity pattern under reverberant conditions”，Proc. ICASSP2000, pp3140-3143, Istanbul, Turkey, June, 2000.

【非特許文献6】澤田宏, 向井良, 荒木章子, 牧野昭二，“周波数領域ブラインド音源分離におけるpermutation問題の解法”，音響学会講演論文集, pp. 541-542, Sep, 2002.

【非特許文献7】澤田宏, 向井良, 荒木章子, 牧野昭二，“周波数領域ブラインド音源分離におけるpermutation問題の頑健な解法”，音響学会講演論文集, pp. 777-778, Mar, 2003.

【発明の開示】

【発明が解決しようとする課題】

【0015】

しかし、これらの方法によっても本質的に並べ替えに失敗する可能性があり、パーミュテーション問題は解決されない。

【0016】

そこで、本発明は上記課題に着目してなされたもので、パーミュテーション問題を考慮する必要のないパーミュテーションフリー・ブラインド音源分離手法を提案することを目的とする。

【課題を解決するための手段】

【0017】

すなわち、本発明は上記課題を解決するために、複数の音源から出力された音声を複数のマイクで受音して目的の音声を抽出する場合、まず、音源から信号を取り出し、この取り出された時間領域の信号を分析フレーム毎に周波数領域の信号に変換する。次に、全分析フレームにおける周波数スペクトルから、各周波数ビン毎の周波数のフーリエ変換の時間変化を読み取り、各周波数でのフーリエ変換の値の時間変化を全周波数標本点について連結した後、この連結された周波数信号を各音源における周波数領域の信号に分離する。そして、分離された周波数領域での信号を時間領域の信号に戻し、この時間領域の信号を出力する。

【0018】

すなわち、従来の周波数領域BBSでは、受音した時間領域の信号を分析フレーム毎にフーリエ変換して並べ(図3)、これを各周波数ビン毎の周波数のフーリエ変換の時間変化ごとに読み出して(図4)、それぞれの分離行列を用いて分離していたが、この手法では、各周波数ビン毎に異なる値の分離行列を用いていたため、各分離された信号が音源1の信号なのか音源2の信号なのかを判別することができない。このため、組み合わせに自由度ができてしまい、いわゆるパーミュテーション問題を生じてしまう。さらに、学習された分離行列を用いて分離を行う場合、標本点の数が多ければ多いほど学習能力が高くなって分離精度がよくなるが、従来の方法であれば分析フレームの長さ分だけしか標本点を取ることができないため、分離の精度がよくなる。加えて、従来の手法では、周波数ビン毎に異なる値の分離行列を用いていたため、分離された信号の振幅に不確実性を生じてしまい、これを修正しなければならなくなる。これに対して、本発明によれば、図4や図5に示すように、各周波数ビン毎の周波数のフーリエ変換の時間変化を全周波数標本点について連結した後、一括して分離行列を用いて分離するようにしたので、組み合わせの自由度におけるパーミュテーションを考える必要がなくなる。また、連結された多くの標本点に基づいて一括した分離処理を行うので、分離精度もよくなる。加えて、一括した分離処理により、従来のように振幅における不確実性もなくなる。

【発明の効果】

【0019】

本発明によれば、各周波数でのフーリエ変換の値の時間変化を全周波数標本点について連結した後、一括して分離処理を行うようにしたので、パーミュテーション問題を考える必要がなくなり、また、多くの標本点に基づく処理を行うので分離精度もよくなる。加えて、一括した分離処理を行うので、従来のように振幅における不確実性もなくなる。

【発明を実施するための最良の形態】

【0020】

以下、本発明の一実施の形態について図面を参照して説明する。図1は、周波数領域BBS手法を用いた装置の機能ブロック図を示したものである。以下では、音源として2つの音源S1と音源S2を用い、集音装置として2つのマイクM1、M2を設けた場合を例とする。ここで、音源の種類としては、人間の声などのような可聴周波数領域の音波を出力する音源を用いる。また、マイクの本数は音源の本数以上であればよく、3個以上のマイクを設けるようにしてもよい。

【0021】

図1の機能ブロックについて説明すると、まず、受音装置1はマイクやマイクロホンアレイなどの装置によって構成され、好ましくは、全周波数に対して最適な指向特性を持つビームフォーマーを用いる。そして、図1に示す2つのビームフォーマーを求め、音源と雑音方向の推定を行う。このビームフォーマーを求めるために、図4や図5の関係を用いる。図4および図5については後述する。

【0022】

この受信装置1によって受信された信号は、まず、時間領域の信号として記憶部に格納される。そして、この信号を周波数領域変換部2にて周波数領域の信号に変換し、同様に記憶部に格納していく。時間領域の信号を周波数領域の信号に変換する場合、分析フレームをずらしながらその分析フレーム内において短時間フーリエ変換(STFT)を行う。ここでは分析フレームの個数をm個としている。これにより、図3や図4に示すように、周波数-時間-スペクトルからなる3次元空間内における離散的な信号が得られる。このうち時間軸方向については、m個の分析フレームに基づく離散的な信号が存在し、また、周波数軸方向については、n個の離散的な信号が存在している。なお、図3と図4は同じ空間を示しており、図3は時間軸に切った信号を示しており、また、図4はある周波数で切った信号を示している。

【0023】

読取部3は、この記憶部に格納されている周波数-時間-スペクトル空間の信号から、各周波数ビン毎の周波数のフーリエ変換の時間変化を読み取る。すなわち、周波数領域変換部2によって分析フレーム毎に図3のようにマッピングされた信号を、今度は、図4に示すような周波数軸の標本点毎に信号として抽出する。

【0024】

そして、本提案手法においては特徴的に、図5に示すように、読み取られた各周波数ビン毎の周波数のフーリエ変換の時間変化を強制的に連結し、一つの連続した信号とする。連結に際しては、低周波数ビンから順に高周波数ビンに連続して連結するが、この順序に限定されるものではない。このとき、連結された信号は、周波数 f_1 におけるスペクトルの時間変化 $t=t_1 \sim t_m$ の信号の後に、周波数 f_2 におけるスペクトルの時間変化 $t=t_1 \sim t_m$ の信号が連結される。以下、同様にすべての周波数標本点におけるスペクトルの時間変化 $t=t_1 \sim t_m$ の信号を連結する。

【0025】

一括分離部4は、連結された周波数信号を、各音源の周波数領域の信号に分離する。この分離手法としては、JADEやFAST ICAなどの手法が存在するが、JADEを用いた場合について説明する。

【0026】

JADEとは、Cardosoらにより提案された手法で、4次クロススキュラントを対角化する手法である。4次クロススキュラントは次のように定義される。

【0027】

【数4】

$$\begin{aligned} \text{cum}(x_i, x_j, x_k, x_l) \\ = E(x_i, x_j, x_k, x_l) - E(x_i, x_j)E(x_k, x_l) \\ - E(x_i, x_k)E(x_j, x_l) - E(x_i, x_l)E(x_j, x_k) \end{aligned}$$

【0028】

観測信号 x が平均0で無相関化したものを z とし、音源信号 s と無相関化された観測信号 z はある直交行列 $U = (u_1, \dots, u_n)$ により、

【0029】

【数5】

$$z = Us$$

という関係で結ばれている。独立性の仮定から、

【0030】

【数6】

$$\text{cum}(s_i, s_j, s_k, s_l) = \begin{cases} \kappa_i, & i = j = k = l \\ 0, & \text{otherwise} \end{cases}$$

【0031】

となる。ここで、 $N \times N$ 行列 $M=(m_{ij})$ により縮約されたkurtosisの行列の第 i, j 要素は

【0032】

【数7】

$$C_{i,j}(M) = \left(\sum_{k,l=1}^N \text{cum}(z_i, z_j, z_k, z_l) m_{kl} \right)$$

【0033】

を考える。 s のインデックスが全て同じ時だけが問題であるから、 $\kappa_i = \text{cum}(s_i, s_i, s_i, s_i)$ として数7を変形すると

【0034】

【数8】

$$C(M) = \sum_{i=1}^N (\kappa_i w_i^T M w_i) w_i w_i^T$$

となる。ここで

【0035】

【数9】

$$D(M) = \text{diag}(\kappa_1 w_1^T M w_1, \dots, \kappa_N w_N^T M w_N)$$

とすると、

【0036】

【数10】

$$C(M) = W D(M) W^T$$

直交行列である W で対角行列を挟んでいるので、 $C(M)$ は対称行列になっている。この $C(M)$ を対角化する W を探すのがJADEである。行列の対角化にはJacobi法を用いることができる。Jacobi法は理想的には2次収束することが知られているので、JADEは収束の速いアルゴリズムといえる。ただし、JADEは4次元配列を用いるので、信号数があまり多くない時には問題がないが、信号数が多い時には計算上の問題がある。

【0037】

周波数上でICAを行う場合は、各周波数ビンに対して分離行列 $w(f)$ を求める必要があり、通常はここで、パーミュテーションが問題となる。しかし、提案法では、分離行列を求めるのは1度でよいので、パーミュテーションが問題にならない。

【0038】

分割部5は、一括分離部4によって分離された信号を周波数ビン毎の周波数のフーリエ変換の時間変化に戻す。すなわち、この分割においては、連結部によって連結された順序と逆の順序で周波数ビン毎の信号に分割する。そして、時間領域変換部6にて各周波数ビン毎に短時間フーリエ逆変換を行い、時間領域の信号に戻す。この時間信号に戻された信

号は、各音源毎に分離された信号として音声出力部7を介して出力される。

【0039】

次に、このように構成された周波数領域BBS手法を用いた処理のフローチャートについて図6を用いて説明する。

【0040】

まず、複数の音源から出力された信号を複数の受音装置1で受音する(ステップS1)。この受音された信号は、時間領域の波形として記憶部に一時記憶される。そして、周波数領域変換部2により、分析フレーム毎に短時間フーリエ変換を行い(ステップS2)、すべての分析フレームについて短時間フーリエ変換を行うまでこの処理を続ける(ステップS3)。このとき、図3に示すように、周波数-時間-スペクトルの3次元空間内において分析フレーム分(m個)の離散的な信号がマッピングされる。次に、このマッピングされた信号を、図4に示すように、各周波数ビン毎の周波数のフーリエ変換の時間変化として読み出し(ステップS4)、各周波数でのフーリエ変換の値の時間変化を全周波数標本点について連結する(ステップS5)。その後、この連結された周波数信号を各音源毎にJADEを用いて一括して分離し(ステップS6)、その分離された周波数領域での信号を各周波数ビンへの信号へ分割する(ステップS7)。そして、時間領域変換部6にて短時間フーリエ逆変換して時間領域の信号に戻し(ステップS8)、この時間領域の信号を音声出力する(ステップS9)。

【0041】

上述のように本実施の形態によれば、各周波数でのフーリエ変換の値の時間変化を全周波数標本点について連結した後、分離行列を用いて一括して分離処理を行うようにしたので、パーミュテーション問題を考える必要がなくなり、また、多くの標本点に基づいて分離処理を行うようにしたので分離精度も向上させることができる。加えて、一括した分離処理を行うので、従来のように振幅における不確実性もなくなる。

【実施例】

【0042】

以下に、上記実施の形態を用いたシステムの評価結果について説明する。

【0043】

1. 評価データ

音声の収録は一般家庭のリビングルームを想定した部屋で行った。TSPにより測定した残響時間は310msである。音源(スピーカ)と観測点(マイクロホン)の配置を図7に示す。観測点は固定し、音源の配置は図中の $\theta = -30^\circ, 0^\circ$ (時計回りを正)となるようにし、それぞれの配置をL, Mとする。マイクロホンには素子間隔5cmの2素子アレイを用いた。音源信号は、男女各1名ずつで、女性(配置:M)による10フレーズ、男性(配置:L)による10フレーズの組み合わせで、合計100個の混合音声について考える。女性および男性の音声を個別に収録し、S/Nの調整を行うことなくそのまま加算した。

【0044】

2. 評価指標

S/Nを用いて分離信号を評価する。S/Nは以下のように定義した。

【0045】

【数11】

$$SNR = \frac{\int |S(\omega)|^2 d\omega}{\int |S(\omega) - \hat{S}(\omega)|^2 d\omega}$$

$S(\omega)$ と $\hat{S}(\omega)$ は音源信号 $s(t)$ と分離信号 $\hat{s}(t)$ をそれぞれ周波数領域に変換したもの

【0046】

3. 評価対象

比較対象のアルゴリズムは、JADE法によって周波数領域ICAを行い、パーミュテーションに関しては、周波数ビン間の相関を用いるIkeda et al. (非特許文献2)の手法である。

【0047】

4. 評価結果

評価は、比較対象のアルゴリズムでパーミュテーション問題の解決ができない例に対して、提案法では解決できていることを、スペクトログラムをスペクトル上での比較によって行う。音響分析条件を表1に示す。

【0048】

【表1】

音響分析条件

サンプリングレート	16ksamples/sec
量子化ビット数	16bits
フレームサイズ	1024 点
フレームシフト	128 点
分析窓	ハミング窓

【0049】

図8は、a)混合前の女性音声、b)混合前の男性音声、c)従来ICAによる分離信号1(女性)、d)従来ICAによる分離信号2(男性)、e)提案法による分離信号1(女性)そしてf)提案法による分離信号2(男性)のスペクトログラムを示している。図8のc)とd)において、特に4kHz以上の領域でパーミュテーションの解決に失敗していることがわかる。しかし、図8のe)とf)ではパーミュテーションの問題が起きていないことが確認できる。

【0050】

次に、パーミュテーション問題が解決できたことによりS/Nがどの程度向上するのかを、セグメンタルSNRによって評価する。

【0051】

図9と10を比較すると、女性音声側のセグメンタルSNRの平均値で約0.65dB改善し、男性音声側のセグメンタルSNRの平均値では提案法により約5.9dBの改善が確認できた。これにより、パーミュテーション問題を回避することによって解決することが分離性能の向上につながるということがわかる。

【0052】

さらに図11に分離結果100個についてのセグメンタルSNRの平均値の比較を示す。図1

1より、SNRはデータにより大きくばらついているが、女性については約1.7dBの改善、男性については平均値の改善についてはほぼ0であるが、標準偏差が小さくなっていることがわかる。図1 1からもパーミュテーション回避によるS/Nの改善が確認できた。

【図面の簡単な説明】

【0053】

【図1】本発明の一実施の形態におけるパーミュテーションフリー・ブラインド音源分離装置の機能ブロック図

【図2】同形態において推定すべき指向特性を示す図

【図3】同形態における分析フレーム区間毎にSTFTした値をマッピングした図

【図4】図3における信号を各周波数ビン毎にみた図

【図5】従来の周波数領域ICAと本実施の形態における対応関係を示す図

【図6】同形態における処理のフローチャートを示す図

【図7】本発明の実施例におけるデータ収録環境を示す図

【図8】同実施例におけるパーミュテーション問題解決の例

【図9】同実施例における女性音声のセグメンタルSNRの比較を示す図

【図10】同実施例における男性音声のセグメンタルSNRの比較を示す図

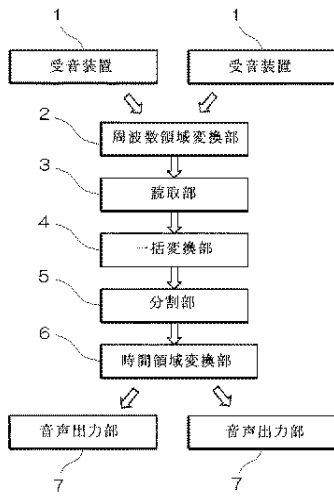
【図11】同実施例における全分離結果のセグメンタルSNRの平均値の比較を示す図

【符号の説明】

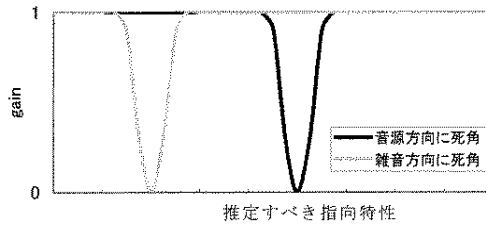
【0054】

- 1・・・受音装置
- 2・・・周波数領域変換部
- 3・・・読取部
- 4・・・一括分離部
- 5・・・分割部
- 6・・・時間領域変換部
- 7・・・音声出力部

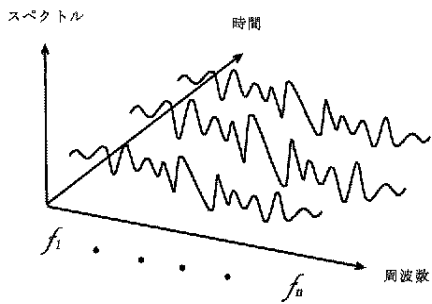
【図1】



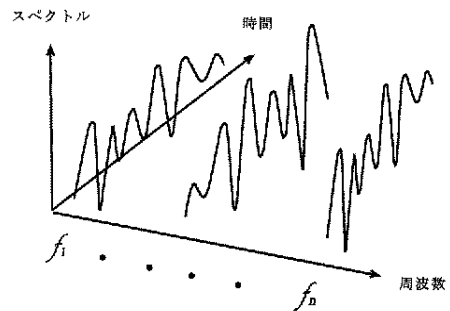
【図2】



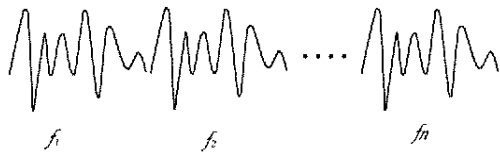
【図3】



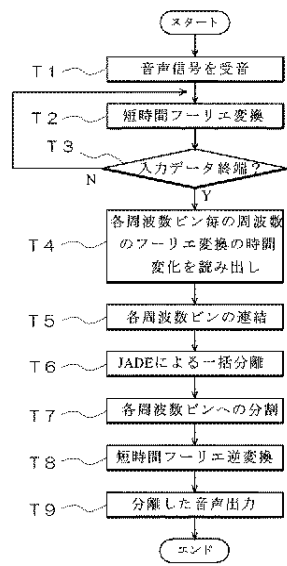
【図4】



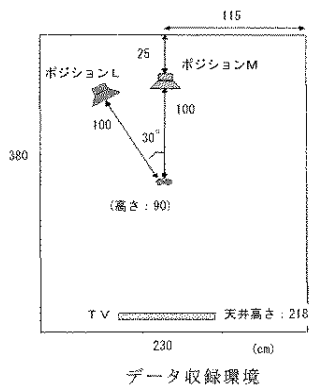
【図5】



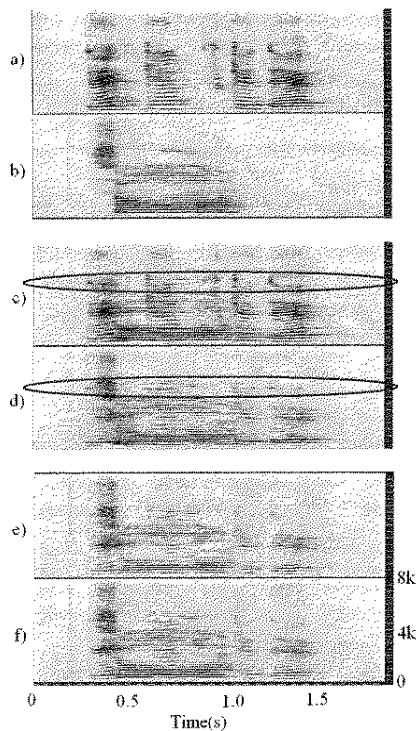
【図6】



【図7】

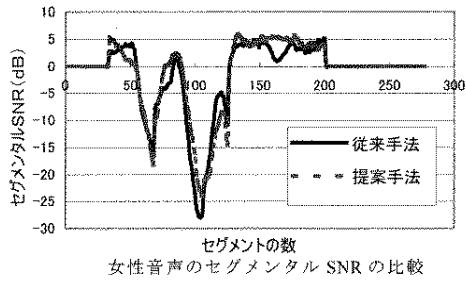


【図8】

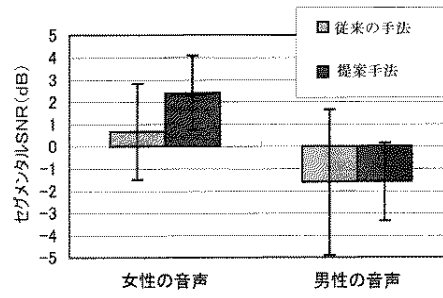


パーミュテーション問題解決の例

【図9】

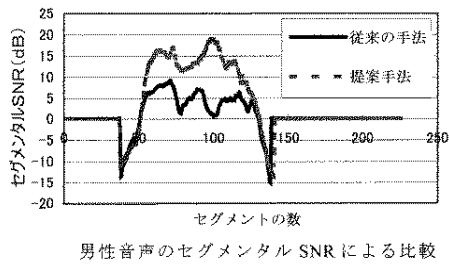


【図11】



全分離結果のセグメンタル SNR の平均値の比較

【図10】



- (74)代理人 100111349
弁理士 久留 徹
- (72)発明者 レアンドロ・エセキエル・ディ・ペルシア
アルゼンチン国エントレリオス州パラナー市ディアマンテ通り320番地9号室
- (72)発明者 ディエゴ・ミローネ
アルゼンチン国サンタフェ州サンタフェ市リバダビア通り2631番地2号室
- (72)発明者 柳田 益造
京都府京田辺市多々羅都谷1-3 同志社大学工学部内