

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6392450号
(P6392450)

(45) 発行日 平成30年9月19日(2018.9.19)

(24) 登録日 平成30年8月31日(2018.8.31)

(51) Int. Cl. F I
G 1 O L 25/54 (2013.01) G 1 O L 25/54
G 1 O L 25/18 (2013.01) G 1 O L 25/18

請求項の数 10 (全 23 頁)

(21) 出願番号	特願2017-512524 (P2017-512524)	(73) 特許権者	000004226
(86) (22) 出願日	平成28年4月11日 (2016.4.11)		日本電信電話株式会社
(86) 国際出願番号	PCT/JP2016/061683		東京都千代田区大手町一丁目5番1号
(87) 国際公開番号	W02016/167216	(73) 特許権者	504137912
(87) 国際公開日	平成28年10月20日 (2016.10.20)		国立大学法人 東京大学
審査請求日	平成29年10月6日 (2017.10.6)		東京都文京区本郷七丁目3番1号
(31) 優先権主張番号	特願2015-81769 (P2015-81769)	(74) 代理人	100121706
(32) 優先日	平成27年4月13日 (2015.4.13)		弁理士 中尾 直樹
(33) 優先権主張国	日本国(JP)	(74) 代理人	100128705
			弁理士 中村 幸雄
		(74) 代理人	100147773
			弁理士 義村 宗洋
		(72) 発明者	守谷 健弘
			東京都千代田区大手町一丁目5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】 マッチング装置、判定装置、これらの方法、プログラム及び記録媒体

(57) 【特許請求の範囲】

【請求項1】

パラメータ を正の数として、所定の時間長の時系列信号に対応するパラメータ を、その時系列信号に対応する周波数領域サンプル列の絶対値の 乗をパワースペクトルと見做すことにより推定されたスペクトル包絡で上記周波数領域サンプル列を除算した系列である白色化スペクトル系列のヒストグラムを近似する一般化ガウス分布の形状パラメータとして、

第一信号を構成する少なくとも1つの所定の時間長の時系列信号にそれぞれ対応するパラメータ の第一系列と、第二信号を構成する少なくとも1つの所定の時間長の時系列信号にそれぞれ対応するパラメータ の第二系列とに基づいて、上記第一信号と上記第二信号との一致度合い、及び/又は、上記第一信号と上記第二信号とが一致しているかどうかを判定するマッチング部、

を含むマッチング装置。

【請求項2】

請求項1のマッチング装置において、

パラメータ 。及びパラメータ を正の数として、所定の方法で定められるパラメータ 。を用いて、入力された所定の時間長の時系列信号に対応する周波数領域サンプル列の絶対値の 。乗をパワースペクトルと見做してスペクトル包絡の推定を行うスペクトル包絡推定部と、上記スペクトル包絡で上記周波数領域サンプル列を除算した系列である白色化スペクトル系列を得る白色化スペクトル系列生成部と、パラメータ を形状パラメータ

とする一般化ガウス分布が上記白色化スペクトル系列のヒストグラムを近似するパラメータを求め、求めたパラメータを上記入力された所定の時間長の時系列信号に対応するパラメータとするパラメータ取得部と、を含むパラメータ決定部を含み、

上記パラメータ決定部は、上記第一信号を構成する少なくとも1つの所定の時間長の時系列信号のそれぞれを入力とする処理を行うことにより、上記第一系列を得る、
マッチング装置。

【請求項3】

請求項1又は2のマッチング装置において、

上記第二系列が記憶されている第二系列記憶部を含み、

上記マッチング部は、上記第二系列記憶部から読み込んだ上記第二系列を用いて、上記判定を行う、
マッチング装置。

10

【請求項4】

請求項1から3の何れかのマッチング装置において、

上記第一信号を構成する少なくとも1つの所定の時間長の時系列信号は、上記第一信号を構成する所定の時間長の時系列信号の全部又は一部であり、

上記第二信号を構成する少なくとも1つの所定の時間長の時系列信号は、上記第二信号を構成する所定の時間長の時系列信号の全部又は一部であり、
マッチング装置。

【請求項5】

20

請求項1から3のマッチング装置において、

複数の信号のそれぞれを上記第二信号として、上記判定を行う、

マッチング装置。

【請求項6】

パラメータを正の数として、所定の時間長の時系列信号に対応するパラメータを、その時系列信号に対応する周波数領域サンプル列の絶対値の乗をパワースペクトルと見做すことにより推定されたスペクトル包絡で上記周波数領域サンプル列を除算した系列である白色化スペクトル系列のヒストグラムを近似する一般化ガウス分布の形状パラメータとし、第一信号を構成する少なくとも1つの所定の時間長の時系列信号にそれぞれ対応するパラメータの系列を第一系列として、

30

上記第一系列に基づいて、上記第一信号の中の所定の種類の信号の区間、及び/又は、上記第一信号の種類を判定する判定部、

を含む判定装置。

【請求項7】

パラメータを正の数として、所定の時間長の時系列信号に対応するパラメータを、その時系列信号に対応する周波数領域サンプル列の絶対値の乗をパワースペクトルと見做すことにより推定されたスペクトル包絡で上記周波数領域サンプル列を除算した系列である白色化スペクトル系列のヒストグラムを近似する一般化ガウス分布の形状パラメータとして、

マッチング部51が、第一信号を構成する少なくとも1つの所定の時間長の時系列信号にそれぞれ対応するパラメータの第一系列と、第二信号を構成する少なくとも1つの所定の時間長の時系列信号にそれぞれ対応するパラメータの第二系列とに基づいて、上記第一信号と上記第二信号との一致度合い、及び/又は、上記第一信号と上記第二信号とが一致しているかどうかを判定するマッチングステップ、

40

を含むマッチング方法。

【請求項8】

パラメータを正の数として、所定の時間長の時系列信号に対応するパラメータを、その時系列信号に対応する周波数領域サンプル列の絶対値の乗をパワースペクトルと見做すことにより推定されたスペクトル包絡で上記周波数領域サンプル列を除算した系列である白色化スペクトル系列のヒストグラムを近似する一般化ガウス分布の形状パラメータ

50

とし、第一信号を構成する少なくとも1つの所定の時間長の時系列信号にそれぞれ対応するパラメータの系列を第一系列として、

判定部が、上記第一系列に基づいて、上記第一信号の中の所定の種類の信号の区間、及び/又は、上記第一信号の種類を判定する判定ステップ、
を含む判定方法。

【請求項9】

請求項1から5の何れかのマッチング装置又は請求項6の判定装置の各部としてコンピュータを機能させるためのプログラム。

【請求項10】

請求項1から5の何れかのマッチング装置又は請求項6の判定装置の各部としてコンピュータを機能させるためのプログラムが記録されたコンピュータ読み取り可能な記録媒体

10

【発明の詳細な説明】

【技術分野】

【0001】

この発明は、音信号に基づいてマッチング又は信号の区間若しくは種類の判定を行う技術に関する。

【背景技術】

【0002】

音信号等の時系列信号の特徴を表すパラメータとして、LSP等のパラメータが知られている（例えば、非特許文献1参照）。

20

【0003】

LSPは、複数次なので直接的に音の分類や区間推定に使うのは扱いが難しい場合がある。例えば、LSPは複数次であるため、LSPを用いた閾値に基づく処理は容易とは言えない。

【0004】

ところで、公知とはなっていないが、発明者によりパラメータが提案されている。このパラメータは、例えば3GPP EVS (Enhanced Voice Services) 規格で使われているような線形予測包絡線を利用する周波数領域の係数の量子化値を算術符号化する符号化方式において、算術符号の符号化対象の属する確率分布を定める形状パラメータである。パラメータは、符号化対象の分布と関連性を有しており、パラメータを適宜定めると効率の良い符号化及び復号を行うことが可能である。

30

【0005】

また、パラメータは、時系列信号の特徴を表す指標と成り得る。このため、パラメータは、上記符号化処理以外の技術、例えばマッチング技術や信号の区間又は種類の判定技術等の音声音響関連技術に用いることができる。

【0006】

さらに、パラメータは1次の値であるため、パラメータを用いた閾値に基づく処理はLSPを用いた閾値に基づく処理と比較すると容易である。このため、パラメータは、マッチング技術や信号の区間又は種類の判定技術等の音声音響関連技術に容易に用いることができる。

40

【先行技術文献】

【非特許文献】

【0007】

【非特許文献1】守谷健弘，「高圧縮音声符号化の必須技術：線スペクトル対（LSP）」，NTT技術ジャーナル，2014年9月，P. 58 - 60

【発明の概要】

【発明が解決しようとする課題】

【0008】

しかしながら、パラメータを用いたマッチング技術及び信号の区間又は種類の判定技

50

術は知られていなかった。

【0009】

本発明は、パラメータを用いてマッチングを行うマッチング装置、パラメータを用いて信号の区間又は種類の判定を行う判定装置、これらの方法、プログラム及び記録媒体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】

【0010】

本発明の一態様によるマッチング装置によれば、パラメータを正の数として、所定の時間長の時系列信号に対応するパラメータを、その時系列信号に対応する周波数領域サンプル列の絶対値の乗をパワースペクトルと見做すことにより推定されたスペクトル包絡で周波数領域サンプル列を除算した系列である白色化スペクトル系列のヒストグラムを近似する一般化ガウス分布の形状パラメータとして、第一信号を構成する少なくとも1つの所定の時間長の時系列信号にそれぞれ対応するパラメータの第一系列と、第二信号を構成する少なくとも1つの所定の時間長の時系列信号にそれぞれ対応するパラメータの第二系列とに基づいて、第一信号と第二信号との一致度合い、及び/又は、第一信号と第二信号とが一致しているかどうかを判定するマッチング部、を備えている。

10

【0011】

本発明の一態様による判定装置によれば、パラメータを正の数として、所定の時間長の時系列信号に対応するパラメータを、その時系列信号に対応する周波数領域サンプル列の絶対値の乗をパワースペクトルと見做すことにより推定されたスペクトル包絡で周波数領域サンプル列を除算した系列である白色化スペクトル系列のヒストグラムを近似する一般化ガウス分布の形状パラメータとし、第一信号を構成する少なくとも1つの所定の時間長の時系列信号にそれぞれ対応するパラメータの系列を第一系列として、第一系列に基づいて、第一信号の中の所定の種類の信号の区間、及び/又は、第一信号の種類を判定する判定部を備えている。

20

【発明の効果】

【0012】

パラメータを用いてマッチング又は信号の区間若しくは種類の判定を行うことができる。

【図面の簡単な説明】

30

【0013】

【図1】マッチング装置の例を説明するためのブロック図。

【図2】マッチング方法の例を説明するためのフローチャート。

【図3】判定装置の例を説明するためのブロック図。

【図4】判定方法の例を説明するためのフローチャート。

【図5】パラメータ決定部の例を説明するためのブロック図。

【図6】パラメータ決定部の例を説明するためのフローチャート。

【図7】一般化ガウス分布を説明するための図。

【発明を実施するための形態】

【0014】

40

[マッチング装置及び方法]

マッチング装置及び方法の一例について説明する。

【0015】

マッチング装置は、図1に示すように、パラメータ決定部27'、マッチング部51及び第二系列記憶部52を例えば備えている。マッチング装置の各々が、図2に例示する各処理を行うことによりマッチング方法が実現される。

【0016】

以下、マッチング装置の各部について説明する。

【0017】

< パラメータ決定部27' >

50

パラメータ決定部 27' には、時系列信号である第一信号が所定の時間長のごとに入力される。第一信号の例は、音声デジタル信号又は音響デジタル信号等の音信号である。

【0018】

パラメータ決定部 27' は、入力された所定の時間長の時系列信号に基づいて、後述する処理により、その入力された所定の時間長の時系列信号パラメータ を決定する（ステップ F1）。これにより、第一信号を構成する少なくとも 1 つの所定の時間長の時系列信号にそれぞれ対応するパラメータ の系列を得る。この第一信号を構成する少なくとも 1 つの所定の時間長の時系列信号にそれぞれ対応するパラメータ の系列のことを「第一系列」と呼ぶことにする。このように、パラメータ決定部 27' は、所定の時間長のフレームごとに処理を行う。

10

【0019】

なお、第一信号を構成する少なくとも 1 つの所定の時間長の時系列信号は、第一信号を構成する所定の時間長の時系列信号の全部であってもよいし、一部であってもよい。

【0020】

パラメータ決定部 27' により決定されたパラメータ の第一系列は、マッチング部 51 に出力される。

【0021】

パラメータ決定部 27' の構成例を図 5 に示す。パラメータ決定部 27' は、図 5 に示すように、周波数領域変換部 41 と、スペクトル包絡推定部 42 と、白色化スペクトル系列生成部 43 と、パラメータ取得部 44 とを例えば備えている。スペクトル包絡推定部 42 は、線形予測分析部 421 及び非平滑化振幅スペクトル包絡系列生成部 422 を例えば備えている。例えばこのパラメータ決定部 27' により実現されるパラメータ決定方法の各処理の例を図 6 に示す。

20

【0022】

以下、図 5 の各部について説明する。

【0023】

<周波数領域変換部 41>

周波数領域変換部 41 には、所定の時間長の時系列信号が入力される。

【0024】

周波数領域変換部 41 は、所定の時間長のフレーム単位で、入力された所定の時間長の時系列信号である時間領域の音信号を周波数領域の N 点の MDCT 係数列 $X(0), X(1), \dots, X(N-1)$ に変換する。N は正の整数である。

30

【0025】

得られた MDCT 係数列 $X(0), X(1), \dots, X(N-1)$ は、スペクトル包絡推定部 42 及び白色化スペクトル系列生成部 43 に出力される。

【0026】

特に断りがない限り、以降の処理はフレーム単位で行われるものとする。

【0027】

このようにして、周波数領域変換部 41 は、所定の時間長の時系列信号に対応する、例えば MDCT 係数列である周波数領域サンプル列を求める（ステップ C41）。

40

【0028】

<スペクトル包絡推定部 42>

スペクトル包絡推定部 42 には、周波数領域変換部 21 が得た MDCT 係数列 $X(0), X(1), \dots, X(N-1)$ が入力される。

【0029】

スペクトル包絡推定部 42 は、所定の方法で定められるパラメータ θ_0 に基づいて、時系列信号に対応する周波数領域サンプル列の絶対値の θ_0 乗をパワースペクトルとして用いたスペクトル包絡の推定を行う（ステップ C42）。

【0030】

50

推定されたスペクトル包絡は、白色化スペクトル系列生成部 4 3 に出力される。

【 0 0 3 1 】

スペクトル包絡推定部 4 2 は、例えば以下に説明する線形予測分析部 4 2 1 及び非平滑化振幅スペクトル包絡系列生成部 4 2 2 の処理により、非平滑化振幅スペクトル包絡系列を生成することによりスペクトル包絡の推定を行う。

【 0 0 3 2 】

パラメータ α_0 は所定の方法で定められるとする。例えば、 α_0 を 0 より大きい所定の数とする。例えば、 $\alpha_0 = 1$ とする。また、現在パラメータ α_0 を求めようとしているフレームよりも前のフレームで求めた α_0 を用いてもよい。現在パラメータ α_0 を求めようとしているフレーム（以下、現フレームとする。）よりも前のフレームとは、例えば現フレームのよりも前のフレームであって現フレームの近傍のフレームである。現フレームの近傍のフレームは、例えば現フレームの直前のフレームである。

10

【 0 0 3 3 】

< 線形予測分析部 4 2 1 >

線形予測分析部 4 2 1 には、周波数領域変換部 4 1 が得た MDCT 係数列 $X(0), X(1), \dots, X(N-1)$ が入力される。

【 0 0 3 4 】

線形予測分析部 4 2 1 は、MDCT 係数列 $X(0), X(1), \dots, X(N-1)$ を用いて、以下の式 (C1) により定義される $\tilde{R}(0), \tilde{R}(1), \dots, \tilde{R}(N-1)$ を線形予測分析して線形予測係数 $\alpha_1, \alpha_2, \dots, \alpha_p$ を生成し、生成された線形予測係数 $\alpha_1, \alpha_2, \dots, \alpha_p$ を符号化して線形予測係数符号と線形予測係数符号に対応する量子化された線形予測係数である量子化線形予測係数 $\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_p$ とを生成する。

20

【 0 0 3 5 】

【数 1】

$$\tilde{R}(k) = \sum_{n=0}^{N-1} |X(n)|^{\alpha_0} \exp\left(-j \frac{2\pi kn}{N}\right), \quad k = 0, 1, \dots, N-1 \quad \dots(C1)$$

【 0 0 3 6 】

生成された量子化線形予測係数 $\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_p$ は、非平滑化スペクトル包絡系列生成部 4 2 2 に出力される。

30

【 0 0 3 7 】

具体的には、線形予測分析部 4 2 1 は、まず MDCT 係数列 $X(0), X(1), \dots, X(N-1)$ の絶対値の α_0 乗をパワースペクトルと見做した逆フーリエ変換に相当する演算、すなわち式 (C1) の演算を行うことにより、MDCT 係数列 $X(0), X(1), \dots, X(N-1)$ の絶対値の α_0 乗に対応する時間領域の信号列である擬似相関関数信号列 $\tilde{R}(0), \tilde{R}(1), \dots, \tilde{R}(N-1)$ を求める。そして、線形予測分析部 4 2 1 は、求めた擬似相関関数信号列 $\tilde{R}(0), \tilde{R}(1), \dots, \tilde{R}(N-1)$ を用いて線形予測分析を行って、線形予測係数 $\alpha_1, \alpha_2, \dots, \alpha_p$ を生成する。そして、線形予測分析部 4 2 1 は、生成された線形予測係数 $\alpha_1, \alpha_2, \dots, \alpha_p$ を符号化することにより、線形予測係数符号と、線形予測係数符号に対応する量子化線形予測係数 $\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_p$ とを得る。

40

【 0 0 3 8 】

線形予測係数 $\alpha_1, \alpha_2, \dots, \alpha_p$ は、MDCT 係数列 $X(0), X(1), \dots, X(N-1)$ の絶対値の α_0 乗をパワースペクトルと見做したときの時間領域の信号に対応する線形予測係数である。

【 0 0 3 9 】

線形予測分析部 4 2 1 による線形予測係数符号の生成は、例えば従来の符号化技術によって行われる。従来の符号化技術とは、例えば、線形予測係数そのものに対応する符号を線形予測係数符号とする符号化技術、線形予測係数を LSP パラメータに変換して LSP パラメータに対応する符号を線形予測係数符号とする符号化技術、線形予測係数を PARCOR 係数に変換して PARCOR 係数に対応する符号を線形予測係数符号とする符号化技術などである。

50

【 0 0 4 0 】

このようにして、線形予測分析部 4 2 1 は、例えば MDCT 係数列である周波数領域サンプル列の絶対値の γ_0 乗をパワースペクトルと見做した逆フーリエ変換を行うことにより得られる疑似相関関数信号列を用いて線形予測分析を行い線形予測係数を生成する（ステップ C 4 2 1）。

【 0 0 4 1 】

< 非平滑化振幅スペクトル包絡系列生成部 4 2 2 >

非平滑化振幅スペクトル包絡系列生成部 4 2 2 には、線形予測分析部 4 2 1 が生成した量子化線形予測係数 $\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_p$ が入力される。

【 0 0 4 2 】

非平滑化振幅スペクトル包絡系列生成部 4 2 2 は、量子化線形予測係数 $\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_p$ に対応する振幅スペクトル包絡の系列である非平滑化振幅スペクトル包絡系列 $\hat{H}(0), \hat{H}(1), \dots, \hat{H}(N-1)$ を生成する。

【 0 0 4 3 】

生成された非平滑化振幅スペクトル包絡系列 $\hat{H}(0), \hat{H}(1), \dots, \hat{H}(N-1)$ は、白色化スペクトル系列生成部 4 3 に出力される。

【 0 0 4 4 】

非平滑化振幅スペクトル包絡系列生成部 4 2 2 は、量子化線形予測係数 $\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_p$ を用いて、非平滑化振幅スペクトル包絡系列 $\hat{H}(0), \hat{H}(1), \dots, \hat{H}(N-1)$ として、式 (C2) により定義される非平滑化振幅スペクトル包絡系列 $\hat{H}(0), \hat{H}(1), \dots, \hat{H}(N-1)$ を生成する。

【 0 0 4 5 】

【 数 2 】

$$\hat{H}(k) = \left(\frac{1}{2\pi} \frac{1}{\left| 1 + \sum_{n=1}^p \hat{\beta}_n \exp(-j2\pi kn/N) \right|^2} \right)^{1/\gamma_0} \quad \dots (C2)$$

【 0 0 4 6 】

このようにして、非平滑化振幅スペクトル包絡系列生成部 4 2 2 は、疑似相関関数信号列に対応する振幅スペクトル包絡の系列を $1/\gamma_0$ 乗した系列である非平滑化スペクトル包絡系列を線形予測分析部 4 2 1 により生成された線形予測係数に変換可能な係数に基づいて得ることによりスペクトル包絡の推定を行う（ステップ C 4 2 2）。

【 0 0 4 7 】

なお、非平滑化スペクトル包絡系列生成部 4 2 2 は、量子化線形予測係数 $\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_p$ に代えて線形予測分析部 4 2 1 が生成した線形予測係数 $\alpha_1, \alpha_2, \dots, \alpha_p$ を用いることにより、非平滑化振幅スペクトル包絡系列 $\hat{H}(0), \hat{H}(1), \dots, \hat{H}(N-1)$ を得てもよい。この場合は、線形予測分析部 4 2 1 は、量子化線形予測係数 $\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_p$ を得る処理をしなくてもよい。

【 0 0 4 8 】

< 白色化スペクトル系列生成部 4 3 >

白色化スペクトル系列生成部 4 3 には、周波数領域変換部 4 1 が得た MDCT 係数列 $X(0), X(1), \dots, X(N-1)$ 及び非平滑化振幅スペクトル包絡生成部 4 2 2 が生成した非平滑化振幅スペクトル包絡系列 $\hat{H}(0), \hat{H}(1), \dots, \hat{H}(N-1)$ が入力される。

【 0 0 4 9 】

白色化スペクトル系列生成部 4 3 は、MDCT 係数列 $X(0), X(1), \dots, X(N-1)$ の各係数を、対応する非平滑化振幅スペクトル包絡系列 $\hat{H}(0), \hat{H}(1), \dots, \hat{H}(N-1)$ の各値で除算することにより、白色化スペクトル系列 $X_w(0), X_w(1), \dots, X_w(N-1)$ を生成する。

【 0 0 5 0 】

生成された白色化スペクトル系列 $X_w(0), X_w(1), \dots, X_w(N-1)$ は、パラメータ取得部 4 4 に

10

20

30

40

50

出力される。

【 0 0 5 1 】

白色化スペクトル系列生成部 4 3 は、例えば、 $k=0, 1, \dots, N-1$ として、MDCT係数列 $X(0), X(1), \dots, X(N-1)$ の各係数 $X(k)$ を非平滑化振幅スペクトル包絡系列 $\hat{H}(0), \hat{H}(1), \dots, \hat{H}(N-1)$ の各値 $\hat{H}(k)$ で除算することにより、白色化スペクトル系列 $X_W(0), X_W(1), \dots, X_W(N-1)$ の各値 $X_W(k)$ を生成する。すなわち、 $k=0, 1, \dots, N-1$ として、 $X_W(k)=X(k)/\hat{H}(k)$ である。

【 0 0 5 2 】

このようにして、白色化スペクトル系列生成部 4 3 は、例えば非平滑化振幅スペクトル包絡系列であるスペクトル包絡で例えばMDCT係数列である周波数領域サンプル列を除算した系列である白色化スペクトル系列を得る（ステップ C 4 3）。

10

【 0 0 5 3 】

< パラメータ取得部 4 4 >

パラメータ取得部 4 4 には、白色化スペクトル系列生成部 4 3 が生成した白色化スペクトル系列 $X_W(0), X_W(1), \dots, X_W(N-1)$ が入力される。

【 0 0 5 4 】

パラメータ取得部 4 4 は、パラメータ を形状パラメータとする一般化ガウス分布が白色化スペクトル系列 $X_W(0), X_W(1), \dots, X_W(N-1)$ のヒストグラムを近似するパラメータ を求める（ステップ C 4 4）。言い換えれば、パラメータ取得部 4 4 は、パラメータ を形状パラメータとする一般化ガウス分布が白色化スペクトル系列 $X_W(0), X_W(1), \dots, X_W(N-1)$ のヒストグラムの分布に近くなるようなパラメータ を決定する。

20

【 0 0 5 5 】

パラメータ を形状パラメータとする一般化ガウス分布は、例えば以下のように定義される。 は、ガンマ関数である。

【 0 0 5 6 】

【 数 3 】

$$f_{GG}(X|\phi, \eta) = \frac{A(\eta)}{\phi} \exp\left(-\left|B(\eta) \frac{X}{\phi}\right|^\eta\right),$$

$$A(\eta) = \frac{\eta B(\eta)}{2\Gamma(1/\eta)}, \quad B(\eta) = \sqrt{\frac{\Gamma(3/\eta)}{\Gamma(1/\eta)}}, \quad \Gamma(x) = \int_0^\infty e^{-t} t^{x-1} dt$$

30

【 0 0 5 7 】

一般化ガウス分布は、形状パラメータである を変えることにより、図 3 のように $\eta = 1$ の時はラプラス分布、 $\eta = 2$ の時はガウス分布、といったように様々な分布を表現することができるものである。 は、0 より大きい所定の数である。 は、0 より大きい 2 以外の所定の数であってもよい。具体的には、 は、2 未満の所定の正の数であってもよい。 は分散に対応するパラメータである。

【 0 0 5 8 】

ここで、パラメータ取得部 4 4 が求める は、例えば以下の式(C3)により定義される。 F^{-1} は、関数Fの逆関数である。この式は、いわゆるモーメント法により導出されるものである。

40

【 0 0 5 9 】

【数4】

$$\eta = F^{-1}\left(\frac{m_1}{\sqrt{m_2}}\right) \dots (C3)$$

$$F(\eta) = \frac{\Gamma(2/\eta)}{\sqrt{\Gamma(1/\eta)\Gamma(3/\eta)}}$$

$$m_1 = \frac{1}{N} \sum_{k=0}^{N-1} |X_w(k)|, \quad m_2 = \frac{1}{N} \sum_{k=0}^{N-1} |X_w(k)|^2$$

10

【0060】

逆関数 F^{-1} が定式化されている場合には、パラメータ取得部44は、定式化された逆関数 F^{-1} に $m_1/((m_2)^{1/2})$ の値を入力したときの出力値を計算することによりパラメータを求めることができる。

【0061】

逆関数 F^{-1} が定式化されていない場合には、パラメータ取得部44は、式(C3)で定義されるの値を計算するために、例えば以下に説明する第一方法又は第二方法によりパラメータを求めてもよい。

【0062】

パラメータを求めるための第一方法について説明する。第一の方法では、パラメータ取得部44は、白色化スペクトル系列に基づいて $m_1/((m_2)^{1/2})$ を計算し、予め用意しておいた異なる複数の、と対応する $F(\quad)$ のペアを参照して、計算された $m_1/((m_2)^{1/2})$ に最も近い $F(\quad)$ に対応するを取得する。

20

【0063】

予め用意しておいた異なる複数の、と対応する $F(\quad)$ のペアは、パラメータ取得部44の記憶部441に予め記憶しておく。パラメータ取得部44は、記憶部441参照して、計算された $m_1/((m_2)^{1/2})$ に最も近い $F(\quad)$ を見つけ、見つかった $F(\quad)$ に対応するを記憶部441から読み込み出力する。

【0064】

計算された $m_1/((m_2)^{1/2})$ に最も近い $F(\quad)$ とは、計算された $m_1/((m_2)^{1/2})$ との差の絶対値が最も小さくなる $F(\quad)$ のことである。

30

【0065】

パラメータを求めるための第二方法について説明する。第二の方法では、逆関数 F^{-1} の近似曲線関数を例えば以下の式(C3')で表される \tilde{F}^{-1} として、パラメータ取得部44は、白色化スペクトル系列に基づいて $m_1/((m_2)^{1/2})$ を計算し、近似曲線関数 \tilde{F}^{-1} に計算された $m_1/((m_2)^{1/2})$ を入力したときの出力値を計算することによりを求める。この近似曲線関数 \tilde{F}^{-1} は使用する定義域において出力が正値となる単調増加関数であればよい。

【0066】

【数5】

$$\eta = \tilde{F}^{-1}\left(\frac{m_1}{\sqrt{m_2}}\right)$$

40

$$\tilde{F}^{-1}(x) = \frac{0.2718}{0.7697 - x} - 0.1247 \dots (C3')$$

【0067】

なお、パラメータ取得部44が求めるは、式(C3)ではなく、式(C3')のように予め定めた正の整数 q_1 及び q_2 を用いて(ただし $q_1 < q_2$)式(C3)を一般化した式により定義されてもよい。

【0068】

50

【数6】

$$\eta = F'^{-1} \left(\frac{m_{q_1}}{(m_{q_2})^{q_1/q_2}} \right) \dots (C3'')$$

$$F'(\eta) = \frac{\Gamma((q_1+1)/\eta)}{(\Gamma(1/\eta))^{q_1/q_2} (\Gamma((q_2+1)/\eta))^{q_1/q_2}}$$

$$m_{q_1} = \frac{1}{N} \sum_{k=0}^{N-1} |X_w(k)|^{q_1}, \quad m_{q_2} = \frac{1}{N} \sum_{k=0}^{N-1} |X_w(k)|^{q_2}$$

10

【0069】

なお、 η が式(C3'')により定義される場合も、 η が式(C3)により定義されている場合と同様の方法により、 η を求めることができる。すなわち、パラメータ取得部44が、白色化スペクトル系列に基づいてその q_1 次モーメントである m_{q_1} とその q_2 次モーメントである m_{q_2} とに基づく値 $m_{q_1}/((m_{q_2})^{q_1/q_2})$ を計算した後、例えば上記の第一及び第二の方法と同様、予め用意しておいた異なる複数の η と対応する $F'(\eta)$ のペアを参照して、計算された $m_{q_1}/((m_{q_2})^{q_1/q_2})$ に最も近い $F'(\eta)$ に対応する η を取得するか、逆関数 F'^{-1} の近似曲線関数を $\sim F'^{-1}$ として、近似曲線関数 $\sim F'^{-1}$ に計算された $m_{q_1}/((m_{q_2})^{q_1/q_2})$ を入力したときの出力値を計算して η を求めることができる。

【0070】

20

このように、 η は次数が異なる2つの異なるモーメント m_{q_1}, m_{q_2} に基づく値であるとも言える。例えば、次数が異なる2つの異なるモーメント m_{q_1}, m_{q_2} のうち、次数が低い方のモーメントの値又はこれに基づく値(以下、前者とする。)と次数が高い方のモーメントの値又はこれに基づく値(以下、後者とする。)との比の値、この比の値に基づく値、又は、前者を後者で割って得られる値に基づき、 η を求めてもよい。モーメントに基づく値とは、例えば、そのモーメントを m とし Q を所定の実数として m^Q のことである。また、これらの値を近似曲線関数 $\sim F'^{-1}$ に入力して η を求めてもよい。この近似曲線関数 $\sim F'^{-1}$ は上記同様、使用する定義域において出力が正值となる単調増加関数であればよい。

【0071】

パラメータ決定部27'は、ループ処理によりパラメータ η を求めてもよい。すなわち、パラメータ決定部27'は、パラメータ取得部44で求まるパラメータ η を所定の方法で定められるパラメータ η_0 とする、スペクトル包絡推定部42、白色化スペクトル系列生成部43及びパラメータ取得部44の処理を更に1回以上行ってもよい。

30

【0072】

この場合、例えば、図5で破線で示すように、パラメータ取得部44で求めたパラメータ η は、スペクトル包絡推定部42に出力される。スペクトル包絡推定部42は、パラメータ取得部44で求めた η をパラメータ η_0 として用いて、上記説明した処理と同様の処理を行いスペクトル包絡の推定を行う。白色化スペクトル系列生成部43は、新たに推定されたスペクトル包絡に基づいて、上記説明した処理と同様の処理を行い白色化スペクトル系列を生成する。パラメータ取得部44は、新たに生成された白色化スペクトル系列に基づいて、上記説明した処理と同様の処理を行いパラメータ η を求める。

40

【0073】

例えば、スペクトル包絡推定部42、白色化スペクトル系列生成部43及びパラメータ取得部44の処理は、所定の回数である N 回だけ更に行われてもよい。 N は所定の正の整数であり、例えば $N=1$ 又は $N=2$ である。

【0074】

また、スペクトル包絡推定部42は、今回求めたパラメータ η と前回求めたパラメータ η_{prev} との差の絶対値が所定の閾値以下となるまで、スペクトル包絡推定部42、白色化スペクトル系列生成部43及びパラメータ取得部44の処理を繰り返してもよい。

【0075】

50

< 第二系列記憶部 5 2 >

第二系列記憶部 5 2 には、第二信号を構成する少なくとも 1 つの所定の時間長の時系列信号にそれぞれ対応するパラメータ の系列である第二系列が記憶されている。

【 0 0 7 6 】

第二信号は、第一信号に対するマッチングの対象となる音声デジタル信号又は音響デジタル信号等の音信号である。

【 0 0 7 7 】

第二系列は、例えば、パラメータ決定部 2 7 ' により求められて、第二系列記憶部 5 2 に記憶される。すなわち、第二信号を構成する少なくとも 1 つの所定の時間長の時系列信号のそれぞれがパラメータ決定部 2 7 ' に入力され、パラメータ決定部 2 7 ' は、第一系列を求めた処理と同様の処理により、第二系列を求めて第二系列記憶部 5 2 に記憶させてもよい。

【 0 0 7 8 】

なお、第二信号を構成する少なくとも 1 つの所定の時間長の時系列信号は、第二信号を構成する所定の時間長の時系列信号の全部であってもよいし、一部であってもよい。

【 0 0 7 9 】

マッチング部 5 1 が、複数の信号のそれぞれを第二信号として、後述する判定を行う場合には、複数の信号のそれぞれに対応する第二系列が第二系列記憶部 5 2 に記憶されているとする。

【 0 0 8 0 】

なお、パラメータ決定部 2 7 ' により求められた第二系列が、第二系列記憶部 5 2 を介さずに直接マッチング部 5 1 に入力されてもよい。この場合、第二系列記憶部 5 2 はマッチング装置に備えられていなくてもよい。また、この場合、パラメータ決定部 2 7 ' が、例えば複数の信号（複数の楽曲）が記憶されている図示していないデータベースから各信号を読み込み、読み込んだ信号から第二系列を求めて、マッチング部 5 1 へ出力する。

【 0 0 8 1 】

< マッチング部 5 1 >

マッチング部 5 1 には、パラメータ決定部 2 7 ' が得た第一系列及び例えば第二系列記憶部 5 2 から読み込んだ第二系列が入力される。

【 0 0 8 2 】

マッチング部 5 1 は、第一系列と第二系列とに基づいて、第一信号と第二信号との一致度合い、及びノ又は、第一信号と第二信号とが一致しているかどうかを判定して、その判定結果を出力する（ステップ F 2 ）。

【 0 0 8 3 】

第一系列を $(1, 1, 1, 2, \dots, 1, N_1)$ と表記し、第二系列を $(2, 1, 2, 2, \dots, 2, N_2)$ と表記する。N1 は、第一系列を構成するパラメータ の数である。N2 は、第二系列を構成するパラメータ の数である。N1 N2 であるとする。

【 0 0 8 4 】

第一信号と第二信号との一致度合いは、第一系列と第二系列との類似度である。第一系列と第二系列との類似度は、例えば、第二系列 $(2, 1, 2, 2, \dots, 2, N_2)$ に含まれる、第一系列 $(1, 1, 1, 2, \dots, 1, N_1)$ に最も近い系列と、第一系列 $(1, 1, 1, 2, \dots, 1, N_1)$ との距離である。第二系列 $(2, 1, 2, 2, \dots, 2, N_2)$ に含まれる、第一系列 $(1, 1, 1, 2, \dots, 1, N_1)$ に最も近い系列の要素数と、第一系列 $(1, 1, 1, 2, \dots, 1, N_1)$ の要素数とは同じであるとする。

【 0 0 8 5 】

この第一系列と第二系列との類似度は、例えば以下の式により定義される。min は、最小値を出力する関数である。この例では、距離として、ユークリッド距離を用いているが、マンハッタン距離等の他の既存の距離や誤差の標準偏差を用いてもよい。

【 0 0 8 6 】

10

20

30

40

【数7】

$$\min_{m \in \{0,1,\dots,N2-N1\}} \left(\sum_{k=1}^{N1} (\eta_{1,k} - \eta_{2,m+k})^2 \right)^{\frac{1}{2}}$$

【0087】

第一系列 ($\eta_{1,1}, \eta_{1,2}, \dots, \eta_{1,N1}$) から得られる、パラメータ η の代表値の系列を代表第一系列 ($\eta_{1,1}^r, \eta_{1,2}^r, \dots, \eta_{1,N1}^r$) とする。同様に、第二系列 ($\eta_{2,1}, \eta_{2,2}, \dots, \eta_{2,N2}$) から得られる、パラメータ η の代表値の系列を代表第二系列 ($\eta_{2,1}^r, \eta_{2,2}^r, \dots, \eta_{2,N2}^r$) とする。

10

【0088】

例えば、 c を $N1$ 及び $N2$ の約数である所定の正の整数として、 c 個のパラメータ η ごとに代表値を得るとすると、 $N1' = N1/c$ とし、 $k=1, 2, \dots, N1'$ として、代表値 $\eta_{1,k}^r$ は第一系列の中の系列 ($\eta_{1,(k-1)c+1}, \eta_{1,(k-1)c+2}, \dots, \eta_{1,kc}$) の代表値である。同様に、代表値 $\eta_{2,k}^r$ は第二系列の中の系列 ($\eta_{2,(k-1)c+1}, \eta_{2,(k-1)c+2}, \dots, \eta_{2,kc}$) の代表値である。

【0089】

$k=1, 2, \dots, N1'$ として、代表値 $\eta_{1,k}^r$ は、第一系列の中の系列 ($\eta_{1,(k-1)c+1}, \eta_{1,(k-1)c+2}, \dots, \eta_{1,kc}$) を代表する値であり、例えば、系列 ($\eta_{1,(k-1)c+1}, \eta_{1,(k-1)c+2}, \dots, \eta_{1,kc}$) の平均値、中央値、最大値又は最小値である。 $k=1, 2, \dots, N2'$ として、代表値 $\eta_{2,k}^r$ は、第二系列の中の系列 ($\eta_{2,(k-1)c+1}, \eta_{2,(k-1)c+2}, \dots, \eta_{2,kc}$) を代表する値であり、例えば、系列 ($\eta_{2,(k-1)c+1}, \eta_{2,(k-1)c+2}, \dots, \eta_{2,kc}$) の平均値、中央値、最大値又は最小値である。

20

【0090】

第一系列と第二系列との類似度は、代表第二系列 ($\eta_{2,1}^r, \eta_{2,2}^r, \dots, \eta_{2,N2}^r$) に含まれる、代表第一系列 ($\eta_{1,1}^r, \eta_{1,2}^r, \dots, \eta_{1,N1}^r$) に最も近い系列と、代表第一系列 ($\eta_{1,1}^r, \eta_{1,2}^r, \dots, \eta_{1,N1}^r$) との距離としてもよい。代表第二系列 ($\eta_{2,1}^r, \eta_{2,2}^r, \dots, \eta_{2,N2}^r$) に含まれる、代表第一系列 ($\eta_{1,1}^r, \eta_{1,2}^r, \dots, \eta_{1,N1}^r$) に最も近い系列の要素数と、代表第一系列 ($\eta_{1,1}^r, \eta_{1,2}^r, \dots, \eta_{1,N1}^r$) の要素数とは同じであるとする。

【0091】

この代表値を用いた第一系列と第二系列との類似度は、例えば以下の式により定義される。 \min は、最小値を出力する関数である。この例では、距離として、ユークリッド距離を用いているが、マンハッタン距離等の他の既存の距離や誤差の標準偏差を用いてもよい。

30

【0092】

【数8】

$$\min_{m \in \{0,1,\dots,N2'-N1'\}} \left(\sum_{k=1}^{N1'} (\eta_{1,k}^r - \eta_{2,m+k}^r)^2 \right)^{\frac{1}{2}}$$

40

【0093】

第一信号と第二信号とが一致しているかどうかは、例えば、第一信号と第二信号との一致度合いと所定の閾値とを比較することにより判定することができる。例えば、マッチング部51は、第一信号と第二信号との一致度合いが、所定の閾値より小さい又は所定の閾値以下である場合に、第一信号と第二信号とが一致していると判定し、そうでない場合には第一信号と第二信号とは一致していないと判定する。

【0094】

マッチング部51は、複数の信号のそれぞれを第二信号として、上記判定を行ってもよい。この場合、マッチング部51は、複数の信号のそれぞれと第一信号との一致度合いを計算し、複数の信号の中の、計算された一致度合いが最も小さい信号を選択して、その一

50

致度合いが最も小さい信号についての情報を出力してもよい。

【0095】

例えば、複数の楽曲のそれぞれに対応する第二系列及び情報が第二系列記憶部52に記憶されているとする。ユーザが、ある曲がどの楽曲であるのかを知りたいとする。この場合、ユーザは、その曲に対応する音信号を第一信号としてマッチング装置に入力して、マッチング部51は、その曲に対応する音信号と最も一致度合いが小さい楽曲の情報を第二系列記憶部52から得ることにより、その曲に対応する楽曲の情報を知ることができる。

【0096】

なお、マッチング部51は、第一系列($1,1, 1,2, \dots, 1,N_1$)の時間変化の系列である時間変化第一系列($1,1, 1,2, \dots, 1,N_1-1$)と、第二系列($2,1, 2,2, \dots, 2,N_2$)の時間変化の系列である時間変化第二系列($2,1, 2,2, \dots, 2,N_2-1$)とに基づいてマッチングを行ってもよい。ここで、例えば、 $1,k = 1,k+1 - 1,k$ ($k=1,2,\dots,N_1-1$)とし、 $2,k = 2,k+1 - 2,k$ ($k=1,2,\dots,N_2-1$)とする。

【0097】

例えば、上記の第一系列及び第二系列を用いたマッチング処理において、第一系列($1,1, 1,2, \dots, 1,N_1$)の代わりに時間変化第一系列($1,1, 1,2, \dots, 1,N_1-1$)を用い、第二系列($2,1, 2,2, \dots, 2,N_2$)の代わりに時間変化第二系列($2,1, 2,2, \dots, 2,N_2-1$)を用いることにより、時間変化第一系列と時間変化第二系列とに基づいてマッチングを行うことができる。

【0098】

また、マッチング部51は、第一系列及び第二系列のみならず、音の大きさを表す指標(例えば、振幅又はエネルギー)、音の大きさを表す指標の時間的変動、スペクトル形状、スペクトル形状の時間的変動、ピッチの周期性、基本周波数等の音響特徴量を更に用いてマッチングを行ってもよい。例えば、(1)第一系列及び第二系列と音の大きさを表す指標とに基づいてマッチングを行ってもよい。また、(2)第一系列及び第二系列と時系列信号の音の大きさを表す指標の時間的変動とに基づいてマッチングを行ってもよい。また、(3)第一系列及び第二系列と時系列信号のスペクトル形状とに基づいてマッチングを行ってもよい。また、(4)第一系列及び第二系列と時系列信号のスペクトル形状の時間的変動とに基づいてマッチングを行ってもよい。また、(5)第一系列及び第二系列と時系列信号のピッチの周期性とに基づいてマッチングを行ってもよい。

【0099】

さらに、マッチング部51は、SVM(Support Vector Machine)、Boosting等の識別技術を用いてマッチングを行ってもよい。

【0100】

なお、マッチング部51は、後述する判定部53の処理と同様の処理により第一信号を構成する各所定の時間長の時系列信号の種類を判定し、また、後述する判定部53の処理と同様の処理により第二信号を構成する各所定の時間長の時系列信号の種類を判定し、それらの判定結果が同一であるかを判定することによりマッチングを行ってもよい。例えば、マッチング部51は、第一信号についての判定結果が「音声 音楽 音声 音楽」であり、第二信号についての判定結果が「音声 音楽 音声 音楽」である場合には、第一信号と第二信号とが一致していると判定する。

【0101】

[判定装置及び方法]

判定装置及び方法の一例について説明する。

【0102】

判定装置は、図3に示すように、パラメータ決定部27'及び判定部53を例えば備えている。判定装置の各部が、図4に例示する各処理を行うことにより判定方法が実現される。

【0103】

以下、判定装置の各部について説明する。

【 0 1 0 4 】

< パラメータ決定部 2 7 ' >

パラメータ決定部 2 7 ' には、時系列信号である第一信号が所定の時間長のごとに入力される。第一信号の例は、音声デジタル信号又は音響デジタル信号等の音信号である。

【 0 1 0 5 】

パラメータ決定部 2 7 ' は、入力された所定の時間長の時系列信号に基づいて、後述する処理により、その入力された所定の時間長の時系列信号パラメータ を決定する（ステップ F 1）。これにより、第一信号を構成する少なくとも 1 つの所定の時間長の時系列信号にそれぞれ対応するパラメータ の系列を得る。この第一信号を構成する少なくとも 1 つの所定の時間長の時系列信号にそれぞれ対応するパラメータ の系列のことを「第一系列」と呼ぶことにする。このように、パラメータ決定部 2 7 ' は、所定の時間長のフレームごとに処理を行う。

10

【 0 1 0 6 】

なお、第一信号を構成する少なくとも 1 つの所定の時間長の時系列信号は、第一信号を構成する所定の時間長の時系列信号の全部であってもよいし、一部であってもよい。

【 0 1 0 7 】

パラメータ決定部 2 7 ' により決定されたパラメータ の第一系列は、判定部 5 3 に出力される。

【 0 1 0 8 】

パラメータ決定部 2 7 ' の詳細は、[マッチング装置及び方法] の欄で説明したものと同様であるため、ここでは重複説明を省略する。

20

【 0 1 0 9 】

< 判定部 5 3 >

判定部 5 3 には、パラメータ決定部 2 7 ' で決定された第一系列が入力される。

【 0 1 1 0 】

判定部 5 3 は、第一系列に基づいて、第一信号の中の所定の種類の信号の区間、及び/又は、第一信号の種類を判定する（ステップ F 3）。所定の種類の信号区間とは、例えば、音声の区間、音楽の区間、非定常音の区間、定常音の区間等の区間のことである。

【 0 1 1 1 】

第一系列を $(p_{1,1}, p_{1,2}, \dots, p_{1,N1})$ と表記する。N1は、第一系列を構成するパラメータ の数である。

30

【 0 1 1 2 】

第一信号の中の所定の種類の信号の区間の判定は、例えば、第一系列を構成するパラメータ $p_{1,k}$ ($k=1,2,\dots,N1$) と、所定の閾値とを比較することにより行うことができる。

【 0 1 1 3 】

例えば、判定部 5 3 は、パラメータ $p_{1,k}$ 閾値である場合には、第一信号の中の、そのパラメータ $p_{1,k}$ に対応する所定の時間長の時系列信号の区間を非定常音（音声、無音等）の区間であると判定する。

【 0 1 1 4 】

また、判定部 5 3 は、閾値 $>$ パラメータ $p_{1,k}$ である場合には、第一信号の中の、そのパラメータ $p_{1,k}$ に対応する所定の時間長の時系列信号の区間を定常音（時間的変動の緩やかな音楽等）の区間であると判定する。

40

【 0 1 1 5 】

また、複数の所定の閾値と比較することにより、第一信号の中の所定の種類の信号の区間の判定を行ってもよい。以下では、2 個の閾値（第一閾値及び第二閾値）を用いた判定の例について説明する。第一閾値 $>$ 第二閾値とする。

【 0 1 1 6 】

例えば、判定部 5 3 は、パラメータ $p_{1,k}$ 第一閾値である場合には、第一信号の中の、そのパラメータ $p_{1,k}$ に対応する所定の時間長の時系列信号の区間を無音の区間である

50

と判定する。

【0117】

また、判定部53は、第一閾値>パラメータ $\theta_{1,k}$ 第二閾値である場合には、第一信号の中の、そのパラメータ $\theta_{1,k}$ に対応する所定の時間長の時系列信号の区間を非定常音の区間であると判定する。

【0118】

また、判定部53は、第二閾値>パラメータ $\theta_{1,k}$ である場合には、第一信号の中の、そのパラメータ $\theta_{1,k}$ に対応する所定の時間長の時系列信号の区間を定常音の区間であると判定する。

【0119】

10

第一信号の種類判定は、例えば、信号の区間の種類判定結果に基づいて行うことができる。例えば、判定部53は、判定された信号の区間の種類ごとに第一信号の中のその種類の信号の区間が占める割合を計算して、最も割合が大きい信号の区間の種類割合の値が処理の閾値以上又は閾値より大である場合には、第一信号はその最も割合が大きい信号の区間の種類であると判定する。

【0120】

第一系列 $(\theta_{1,1}, \theta_{1,2}, \dots, \theta_{1,N1})$ から得られる、パラメータ $\theta_{1,k}$ の代表値の系列を代表第一系列 $(\theta_{1,1}^r, \theta_{1,2}^r, \dots, \theta_{1,N1}^r)$ とする。例えば、 c を $N1$ の約数である所定の正の整数として、 c 個のパラメータ $\theta_{1,k}$ ごとに代表値を得るとすると、 $N1' = N1/c$ とし、 $k=1, 2, \dots, N1'$ として、代表値 $\theta_{1,k}^r$ は第一系列の中の系列 $(\theta_{1,(k-1)c+1}, \theta_{1,(k-1)c+2}, \dots, \theta_{1,kc})$ の代表値である。 $k=1, 2, \dots, N1'$ として、代表値 $\theta_{1,k}^r$ は、第一系列の中の系列 $(\theta_{1,(k-1)c+1}, \theta_{1,(k-1)c+2}, \dots, \theta_{1,kc})$ を代表する値であり、例えば、系列 $(\theta_{1,(k-1)c+1}, \theta_{1,(k-1)c+2}, \dots, \theta_{1,kc})$ の平均値、中央値、最大値又は最小値である。

20

【0121】

判定部53は、代表第一系列 $(\theta_{1,1}^r, \theta_{1,2}^r, \dots, \theta_{1,N1}^r)$ に基づいて、第一信号の中の所定の種類の信号の区間、及び/又は、第一信号の種類を判定する第一信号の種類を判定してもよい。

【0122】

例えば、例えば、判定部53は、代表値 $\theta_{1,k}^r$ 第一閾値である場合には、第一信号の中の、その代表値 $\theta_{1,k}^r$ に対応する所定の時間長の時系列信号の区間を音声の区間であると判定する。

30

【0123】

ここで、代表値 $\theta_{1,k}^r$ に対応する所定の時間長の時系列信号の区間とは、代表値 $\theta_{1,k}^r$ に対応する第一系列の中の系列 $(\theta_{1,(k-1)c+1}, \theta_{1,(k-1)c+2}, \dots, \theta_{1,kc})$ のそれぞれのパラメータ $\theta_{1,k}$ に対応する所定の時間長の時系列信号の区間のことである。

【0124】

また、判定部53は、第一閾値>代表値 $\theta_{1,k}^r$ 第二閾値である場合には、第一信号の中の、その代表値 $\theta_{1,k}^r$ に対応する所定の時間長の時系列信号の区間を音楽の区間であると判定する。

40

【0125】

また、判定部53は、第二閾値>代表値 $\theta_{1,k}^r$ 第三閾値である場合には、第一信号の中の、その代表値 $\theta_{1,k}^r$ に対応する所定の時間長の時系列信号の区間を非定常音の区間であると判定する。

【0126】

さらに、判定部53は、第三閾値>代表値 $\theta_{1,k}^r$ である場合には、第一信号の中の、その代表値 $\theta_{1,k}^r$ に対応する所定の時間長の時系列信号の区間を定常音の区間であると判定する。

【0127】

なお、判定部53は、第一系列 $(\theta_{1,1}, \theta_{1,2}, \dots, \theta_{1,N1})$ の時間変化の系列である時

50

間変化第一系列 ($a_{1,1}, a_{1,2}, \dots, a_{1,N1-1}$) に基づいて判定処理を行ってもよい。ここで、例えば、 $a_{1,k} = a_{1,k+1} - a_{1,k}$ ($k=1, 2, \dots, N1-1$) とする。

【0128】

例えば、上記の第一系列を用いた判定処理において、第一系列 ($a_{1,1}, a_{1,2}, \dots, a_{1,N1}$) の代わりに時間変化第一系列 ($a_{1,1}, a_{1,2}, \dots, a_{1,N1-1}$) を用いることにより、時間変化第一系列に基づいて判定を行うことができる。

【0129】

また、判定部53は、時系列信号の音の大きさを表す指標（例えば、振幅又はエネルギー）、音の大きさを表す指標の時間的変動、スペクトル形状、スペクトル形状の時間的変動、ピッチの周期性、基本周波数等の音響特徴量を更に用いて判定を行ってもよい。例えば、判定部53は、(1)パラメータ $a_{1,k}$ と時系列信号の音の大きさを表す指標とに基づいて判定を行ってもよい。また、(2)パラメータ $a_{1,k}$ と時系列信号の音の大きさを表す指標の時間的変動とに基づいて判定を行ってもよい。また、(3)パラメータ $a_{1,k}$ と時系列信号のスペクトル形状とに基づいて判定を行ってもよい。また、(4)パラメータ $a_{1,k}$ と時系列信号のスペクトル形状の時間的変動とに基づいて判定を行ってもよい。また、(5)パラメータ $a_{1,k}$ と時系列信号のピッチの周期性とに基づいて判定を行ってもよい。

【0130】

以下、判定部53が、(1)パラメータ $a_{1,k}$ と時系列信号の音の大きさを表す指標とに基づいて判定を行う場合、(2)パラメータ $a_{1,k}$ と時系列信号の音の大きさを表す指標の時間的変動とに基づいて判定を行う場合、(3)パラメータ $a_{1,k}$ と時系列信号のスペクトル形状とに基づいて判定を行う場合、(4)パラメータ $a_{1,k}$ と時系列信号のスペクトル形状の時間的変動とに基づいて判定を行う場合、(5)パラメータ $a_{1,k}$ と時系列信号のピッチの周期性とに基づいて判定を行う場合のそれぞれについて説明する。

【0131】

(1)パラメータ $a_{1,k}$ と音の大きさを表す指標とに基づいて判定を行う場合、判定部53は、パラメータ $a_{1,k}$ に対応する時系列信号の音の大きさを表す指標が大きいかどうかを判定し、また、パラメータ $a_{1,k}$ が大きいかどうかを判定する。

【0132】

時系列信号の音の大きさを表す指標が小さく、かつ、パラメータ $a_{1,k}$ が大きい場合には、パラメータ $a_{1,k}$ に対応する時系列信号の区間を周囲雑音（騒音）の区間であると判定する。

【0133】

時系列信号の音の大きさを表す指標が大きいかどうかは、例えば所定の閾値 C_E に基づいて判定することができる。すなわち、時系列信号の音の大きさを表す指標の所定の閾値 C_E であれば時系列信号の音の大きさを表す指標が大きい、そうでない場合には時系列信号の音の大きさを表す指標は小さいと判定することができる。時系列信号の音の大きさを表す指標として、例えば、平均振幅（サンプルあたりの平均エネルギーの平方根）を用いた場合には、 $C_E = \text{最大振幅値} * (1/128)$ とする。例えば16ビット精度ならば最大振幅値は32768となるため、 $C_E = 256$ とする。

【0134】

パラメータ $a_{1,k}$ が大きいかどうかは、例えば所定の閾値 C に基づいて判定することができる。すなわち、パラメータ $a_{1,k}$ が所定の閾値 C であればパラメータ $a_{1,k}$ が大きい、そうでない場合にはパラメータ $a_{1,k}$ は小さいと判定することができる。例えば、 $C = 1$ とする。

【0135】

時系列信号の音の大きさを表す指標が小さく、かつ、パラメータ $a_{1,k}$ が小さい場合には、パラメータ $a_{1,k}$ に対応する時系列信号の区間をBGMのような特徴のある背景音の区間と判定する。

【0136】

10

20

30

40

50

時系列信号の音の大きさを表す指標が大きく、かつ、パラメータ $_{1,k}$ が大きい場合には、パラメータ $_{1,k}$ に対応する時系列信号の区間を音声やにぎやかな音楽の区間であると判定する。

【0137】

時系列信号の音の大きさを表す指標が大きく、かつ、パラメータ $_{1,k}$ が小さい場合には、パラメータ $_{1,k}$ に対応する時系列信号の区間を楽器の演奏等の音楽の区間と判定する。

【0138】

(2) パラメータ $_{1,k}$ と時系列信号の音の大きさを表す指標の時間的変動とに基づいて判定を行う場合には、判定部 53 は、パラメータ $_{1,k}$ に対応する時系列信号の音の大きさを表す指標の時間的変動が大きいかどうかを判定し、また、パラメータ $_{1,k}$ が大きいかどうかを判定する。

10

【0139】

時系列信号の音の大きさを表す指標の時間的変動が大きいかどうかは、例えば所定の閾値 C_E' に基づいて判定することができる。すなわち、時系列信号の音の大きさを表す指標の時間的変動が所定の閾値 C_E' であれば時系列信号の音の大きさを表す指標の時間的変動が大きい、そうでない場合には時系列信号の音の大きさを表す指標の時間的変動は小さいと判定することができる。時系列信号の音の大きさを表す指標として、例えば、時系列信号を構成する4個のサブフレームのエネルギーの相加平均を相乗平均で割った値 $F = ((1/4) \times 4 \text{個のサブフレームのエネルギー}) / ((\text{サブフレームのエネルギー})^{1/4})$ を用いた場合には、 $C_E' = 1.5$ とする。

20

【0140】

時系列信号の音の大きさを表す指標の時間的変動が小さく、かつ、パラメータ $_{1,k}$ が大きい場合には、パラメータ $_{1,k}$ に対応する時系列信号の区間を周囲雑音(騒音)の区間と判定する。

【0141】

時系列信号の音の大きさを表す指標の時間的変動が小さく、かつ、パラメータ $_{1,k}$ が小さい場合には、パラメータ $_{1,k}$ に対応する時系列信号の区間を継続音を主体とした管楽器や弦楽器の音楽の区間と判定する。

【0142】

時系列信号の音の大きさを表す指標の時間的変動が大きく、かつ、パラメータ $_{1,k}$ が大きい場合には、判定部 53 は、パラメータ $_{1,k}$ に対応する時系列信号の区間を音声の区間と判定する。

30

【0143】

時系列信号の音の大きさを表す指標の時間的変動が大きく、かつ、パラメータ $_{1,k}$ が小さい場合には、判定部 53 は、パラメータ $_{1,k}$ に対応する時系列信号の区間を時間変動の大きい音楽の区間と判定する。

【0144】

(3) パラメータ $_{1,k}$ と時系列信号のスペクトル形状とに基づいて判定を行う場合、判定部 53 は、パラメータ $_{1,k}$ に対応する時系列信号のスペクトル形状が平坦かどうかを判定し、また、パラメータ $_{1,k}$ が大きいかどうかを判定する。

40

【0145】

時系列信号のスペクトル形状が平坦であり、かつ、パラメータ $_{1,k}$ が大きい場合には、パラメータ $_{1,k}$ に対応する時系列信号の区間を定常な周囲雑音(騒音)の区間と判定する。パラメータ $_{1,k}$ に対応する時系列信号のスペクトル形状が平坦かどうかは、所定の閾値 E_V に基づいて判定することができる。例えば、パラメータ $_{1,k}$ に対応する第1次のPARCOR係数の絶対値が所定の閾値 E_V (例えば、 $E_V = 0.7$) 未満なら、パラメータ $_{1,k}$ に対応する時系列信号のスペクトル形状が平坦である、そうでない場合にはパラメータ $_{1,k}$ に対応する時系列信号のスペクトル形状は平坦でないと判定することができる。

【0146】

50

時系列信号のスペクトル形状が平坦であり、かつ、パラメータ $\alpha_{1,k}$ が小さい場合には、パラメータ $\alpha_{1,k}$ に対応する時系列信号の区間を時間変動の大きい音楽の区間と判定する。

【0147】

時系列信号のスペクトル形状が平坦ではなく、かつ、パラメータ $\alpha_{1,k}$ が大きい場合には、パラメータ $\alpha_{1,k}$ に対応する時系列信号の区間を音声の区間と判定する。

【0148】

時系列信号のスペクトル形状が平坦ではなく、かつ、パラメータ $\alpha_{1,k}$ が小さい場合には、パラメータ $\alpha_{1,k}$ に対応する時系列信号の区間を継続音を主体とした管楽器や弦楽器の音楽の区間と判定する。

10

【0149】

(4) パラメータ $\alpha_{1,k}$ と時系列信号のスペクトル形状の時間的変動とに基づいて判定を行う場合、判定部53は、パラメータ $\alpha_{1,k}$ に対応する時系列信号のスペクトル形状の時間的変動が大きいかどうかを判定し、また、また、パラメータ $\alpha_{1,k}$ が大きいかどうかを判定する。

【0150】

パラメータ $\alpha_{1,k}$ に対応する時系列信号のスペクトル形状の時間的変動が平坦かどうかは、所定の閾値 E_V' に基づいて判定することができる。例えば、時系列信号を構成する4個のサブフレームの第1次のPARCOR係数の絶対値の相加平均を相乗平均で割った値 $F_V = ((1/4) \times 4 \text{個のサブフレームの第1次のPARCOR係数の絶対値}) / ((\text{第1次のPARCOR係数の絶対値})^{1/4})$ が所定の閾値 E_V' (例えば、 $E_V' = 1.2$) 以上なら、パラメータ $\alpha_{1,k}$ に対応する時系列信号のスペクトル形状の時間的変動が大きい、そうでない場合にはパラメータ $\alpha_{1,k}$ に対応する時系列信号のスペクトル形状の時間的変動は小さいと判定することができる。

20

【0151】

時系列信号のスペクトル形状の時間的変動が大きく、かつ、パラメータ $\alpha_{1,k}$ が大きい場合には、パラメータ $\alpha_{1,k}$ に対応する時系列信号の区間を音声の区間と判定する。

【0152】

時系列信号のスペクトル形状の時間的変動が大きく、かつ、パラメータ $\alpha_{1,k}$ が小さい場合には、パラメータ $\alpha_{1,k}$ に対応する時系列信号の区間を時間変動の大きい音楽の区間と判定する。

30

【0153】

時系列信号のスペクトル形状の時間的変動が小さく、かつ、パラメータ $\alpha_{1,k}$ が大きい場合には、パラメータ $\alpha_{1,k}$ に対応する時系列信号の区間を周囲雑音(騒音)の区間と判定する。

【0154】

時系列信号のスペクトル形状の時間的変動が小さく、かつ、パラメータ $\alpha_{1,k}$ が小さい場合には、パラメータ $\alpha_{1,k}$ に対応する時系列信号の区間を継続音を主体とした管楽器や弦楽器の音楽の区間と判定する。

【0155】

(5) パラメータ $\alpha_{1,k}$ と時系列信号のピッチの周期性とに基づいて判定を行う場合、判定部53は、パラメータ $\alpha_{1,k}$ に対応する時系列信号のピッチの周期性が大きいかどうかを判定し、また、また、パラメータ $\alpha_{1,k}$ が大きいかどうかを判定する。

40

【0156】

ピッチの周期性が大きいかどうかは、例えば所定の閾値 C_p に基づいて判定することができる。すなわち、ピッチの周期性が所定の閾値 C_p であればピッチの周期性が大きい、そうでない場合にはピッチの周期性は小さいと判定することができる。ピッチの周期性として、例えば、ピッチ周期 サンプル離れた系列との正規化相関関数

【0157】

【数9】

$$R(\tau) = \frac{\sum_{i=\tau}^N x(i)x(i-\tau)}{\sum_{i=\tau}^N x^2(i)}$$

【0158】

(ただし $x(i)$ は時系列のサンプル値、 N はフレームのサンプル数)を用いた場合には、 $C_p=0.8$ とする。

【0159】

ピッチの周期性が大きく、かつ、パラメータ $\alpha_{1,k}$ が大きい場合には、パラメータ $\alpha_{1,k}$ 10
に対応する時系列信号の区間を音声の区間と判定する。

【0160】

ピッチの周期性が大きく、かつ、パラメータ $\alpha_{1,k}$ が小さい場合には、パラメータ $\alpha_{1,k}$
に対応する時系列信号の区間を継続音を主体とした管楽器や弦楽器の音楽の区間と判定す
る。

【0161】

ピッチの周期性が小さく、かつ、パラメータ $\alpha_{1,k}$ が大きい場合には、パラメータ $\alpha_{1,k}$
に対応する時系列信号の区間を周囲雑音(騒音)の区間と判定する。

【0162】

ピッチの周期性が小さく、かつ、パラメータ $\alpha_{1,k}$ が小さい場合には、パラメータ $\alpha_{1,k}$ に 20
対応する時系列信号の区間を時間変動の大きい音楽の区間と判定する。さらに、判定部
53は、SVM(Support Vector Machine)、Boosting等の識別技術を用いて判定を行っても
よい。この場合、パラメータ $\alpha_{1,k}$ ごとに音声、音楽、無音等のラベルが対応付けされた学習
データを用意しておいて、この学習データを用いて予め学習をしておく。

【0163】

[プログラム及び記録媒体]

各装置又は各方法における各部をコンピュータによって実現してもよい。その場合、各
装置又は各方法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、各装置又は各方法における各部がコンピュータ上で
実現される。 30

【0164】

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、半導体メモリ等どのようなものでもよい。

【0165】

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-R
OM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
ラムを流通させてもよい。 40

【0166】

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
己の記憶部に格納する。そして、処理の実行時、このコンピュータは、自己の記憶部に格
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施形態として、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い。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
50

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、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。

【0167】

また、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、各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】

10

【0168】

マッチング装置、方法及びプログラムは、例えば、曲の出典の検索、違法コンテンツの検出、演奏楽器や曲の構成が似ている別の曲の検索に用いることができる。また、判定装置、方法及びプログラムは、例えば、著作権料の算出に用いることができる。

【図1】

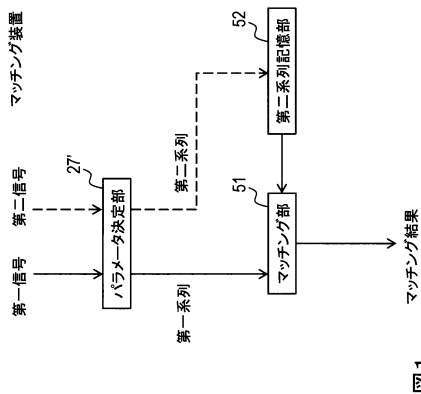


図1

【図2】

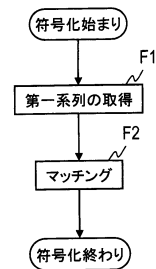


図2

【図3】

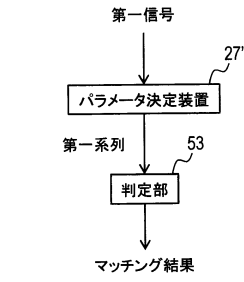


図3

【図4】

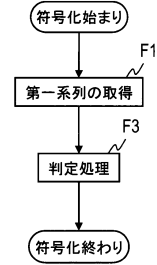


図4

【図5】

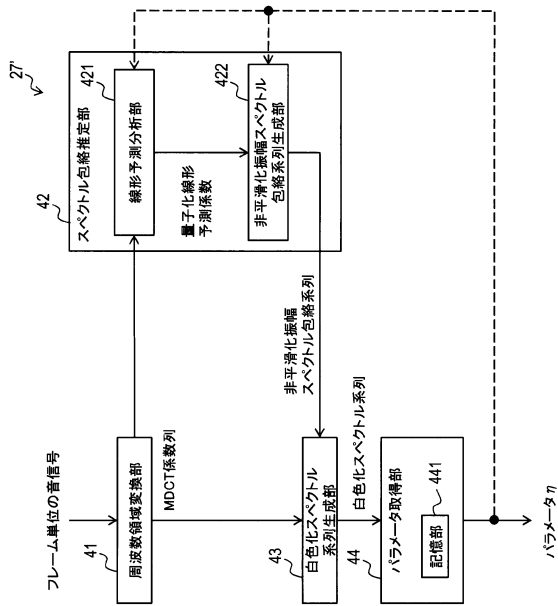


図5

【図6】

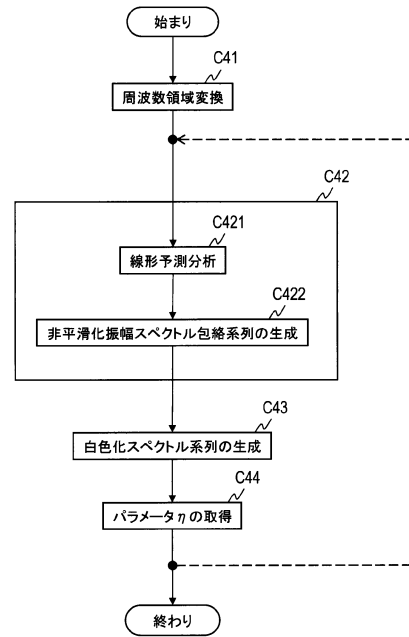


図6

【 図 7 】

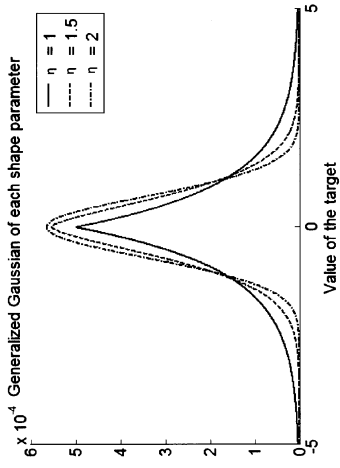


図7

フロントページの続き

- (72)発明者 川西 隆仁
東京都千代田区大手町一丁目5番1号 日本電信電話株式会社内
- (72)発明者 鎌本 優
東京都千代田区大手町一丁目5番1号 日本電信電話株式会社内
- (72)発明者 原田 登
東京都千代田区大手町一丁目5番1号 日本電信電話株式会社内
- (72)発明者 亀岡 弘和
東京都千代田区大手町一丁目5番1号 日本電信電話株式会社内
- (72)発明者 杉浦 亮介
東京都文京区本郷七丁目3番1号 国立大学法人東京大学内

審査官 安田 勇太

- (56)参考文献 特開平10-253444(JP,A)
国際公開第2013/035257(WO,A1)

- (58)調査した分野(Int.Cl.,DB名)
G10L 25/00 - 25/93