

(19) 日本国特許庁 (JP)

(12) 特許公報 (B2)

(11) 特許番号

特許第3584001号

(P3584001)

(45) 発行日 平成16年11月4日 (2004.11.4)

(24) 登録日 平成16年8月6日 (2004.8.6)

(51) Int. Cl.⁷

F I

G 1 0 L 15/02
G 1 0 L 11/00
G 1 0 L 15/00

G 1 0 L 7/08 A
G 1 0 L 7/08 Z
G 1 0 L 9/00 A

請求項の数 5

(全 9 頁)

(21) 出願番号 特願2001-70803(P2001-70803)
(22) 出願日 平成13年3月13日 (2001.3.13)
(65) 公開番号 特開2002-268669 (P2002-268669A)
(43) 公開日 平成14年9月20日 (2002.9.20)
審査請求日 平成14年8月13日 (2002.8.13)

(73) 特許権者 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
(74) 代理人 100066153
弁理士 草野 卓
(74) 代理人 100100642
弁理士 稲垣 稔
(72) 発明者 守谷 健弘
東京都千代田区大手町二丁目3番1号 日
日本電信電話株式会社内
(72) 発明者 千喜良 和明
東京都千代田区大手町二丁目3番1号 日
日本電信電話株式会社内

最終頁に続く

(54) 【発明の名称】 音響信号分析方法、音響信号分析装置、音響信号分析プログラム、音響信号分析プログラムを記憶したコンピュータが読取り可能な記憶媒体

(57) 【特許請求の範囲】

【請求項 1】

A、音響信号をフレーム毎に分析し、特徴パラメータを算出するパラメータ算出処理と、
B、このパラメータ算出処理で得られた各フレーム毎の特徴パラメータを少なくとも2以上のフレームに跨って回帰分析し、特徴パラメータの2次以内の回帰係数のエネルギーの総和を算出する変化率算出処理と、
C、時間的変化率算出処理により算出された特徴パラメータの2次以内の回帰係数のエネルギーの総和をあらかじめ設定したしきい値と比較し、その比較結果により入力信号の種類を分析する分析処理と、
を含むことを特徴とする音響信号分析方法。

【請求項 2】

請求項1記載の音響信号分析方法において、上記音響信号の分析は線形予測分析であり、上記特徴パラメータはフレーム毎のエネルギー、線形予測ケプストラム係数、線スペクトル対パラメータ、周波数領域の伸縮による変形を行った線形予測ケプストラム係数の何れか一つで実行することを特徴とする音響信号分析方法。

【請求項 3】

A、音響信号をフレーム毎に分析して特徴パラメータを算出するパラメータ算出手段と、
B、このパラメータ算出手段が算出した各フレーム毎の特徴パラメータを少なくとも2以上のフレームに跨って回帰分析し、特徴パラメータの2次以内の回帰係数のエネルギーの総和を算出する変化率算出手段と、

C、この変化率算出処理により算出された特徴パラメータの2次以内の回帰係数のエネルギーの総和に適応して入力信号の分析を実行する信号分析手段と、
 によって構成したことを特徴とする音響信号分析装置。

【請求項4】

A、音響信号をフレーム毎に分析し、特徴パラメータを算出するパラメータ算出プログラムと、

B、このパラメータ算出プログラムが算出した各フレーム毎の特徴パラメータを少なくとも2以上のフレームに跨って回帰分析し、フレーム間の特徴パラメータの2次以内の回帰係数のエネルギーの総和を算出する変化率算出プログラムと、

C、この変化率算出プログラムにより算出した特徴パラメータの2次以内の回帰係数のエネルギーの総和に適応させた信号分析プログラムと、

によって構成される音響信号分析プログラム。

【請求項5】

A、音響信号をフレーム毎に分析し、特徴パラメータを算出するパラメータ算出プログラムと、

B、このパラメータ算出プログラムが算出した各フレーム毎の特徴パラメータを少なくとも2以上のフレームに跨って回帰分析し、特徴パラメータの2次以内の回帰係数のエネルギーの総和を算出する変化率算出プログラムと、

C、この変化率算出プログラムにより算出した特徴パラメータの2次以内の回帰係数のエネルギーの総和に適応させた分析処理を実行する分析処理プログラムと、

によって構成された音響信号分析プログラムを記憶したコンピュータが読取り可能な記憶媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

この発明はデジタル音響信号を処理する方法で、例えば音声信号かそれ以外の信号かを分析することに利用することができる音響信号分析方法及びこの方法を用いて動作する装置、この装置を構成するためのコンピュータプログラム及びこれを記憶した記憶媒体に関する。

【0002】

【従来の技術】

従来から、音声が存在する区間と無音の区間を判別する技術や、音声認識の技術ではスペクトル包絡パラメータ、特にLPCケプストラムの時間変化率を使った特徴パラメータの抽出や音素の境界を判別する技術がよく使われている。

【0003】

【発明が解決しようとする課題】

音声或いは音楽などを含む一般の音響信号のうち、例えば音声信号を自動的に判別したり、或いは音声信号らしさを自動的に推定することができるとその後の処理が適応して制御できることになるため、各信号に適した処理を行うことができる。

この発明の目的は、音声或いは音楽などを含む一般の音響信号の予め設定した特定の信号か、それ以外の信号かを自動的に分析することができる音響信号分析方法及びこの分析方法を用いた装置を提案しようとするものである。

【0004】

【課題を解決するための手段】

この発明の請求項1では分析すべき信号の特徴パラメータを求めると共に、この特徴パラメータの時間的変化率を算出し、この算出された特徴パラメータの時間的変化率に応じて信号の分析処理を行う音響信号分析方法を提案する。

音響信号をフレーム毎に分析し、特徴パラメータを算出するパラメータ算出処理と、

このパラメータ算出処理で得られた各フレーム毎の特徴パラメータを少なくとも2以上のフレームに跨って回帰分析し、特徴パラメータの時間的変化率を算出する変化率算出処理

と、
この変化率算出処理により算出された特徴パラメータの時間的変化率に適応した分析処理
と、
を含む音響信号分析方法を提案する。

【0005】

分析処理は時間的変化率算出処理により算出された特徴パラメータの時間的変化率をあらかじめ設定したしきい値と比較し、その比較結果により入力信号の種類を分析する分析処理とする音響信号分析方法を提案する。

この発明の請求項2では請求項1記載の音響信号分析方法において、音響信号の分析は線形予測分析であり、特徴パラメータはフレーム毎のエネルギー、線形予測ケプストラム係数、線スペクトル対パラメータ、周波数領域の伸縮による変形を行った線形予測ケプストラム係数の何れか一つで実行する音響信号分析方法を提案する。

【0006】

この発明の請求項3では音響信号をフレーム毎に分析して特徴パラメータを算出するパラメータ算出手段と、

このパラメータ算出手段が算出した各フレーム毎の特徴パラメータを少なくとも2以上のフレームに跨って回帰分析し、特徴パラメータの時間的変化率を算出する変化率算出手段と、

この変化率算出処理により算出された特徴パラメータの時間的変化率に適応して入力信号の分析を実行する信号分析手段と、

によって構成した音響信号分析装置を提案する。

【0007】

この発明の請求項4では音響信号をフレーム毎に分析し、特徴パラメータを算出するパラメータ算出プログラムと、

このパラメータ算出プログラムが算出した各フレーム毎の特徴パラメータを少なくとも2以上のフレームに跨って回帰分析し、フレーム間の変化率を算出する変化率算出プログラムと、

この変化率算出プログラムにより算出した特徴パラメータの時間的変化率に適応させた信号分析プログラムと、

によって構成される音響信号分析プログラムを提案する。

【0008】

この発明の請求項5では音響信号をフレーム毎に分析し、特徴パラメータを算出するパラメータ算出プログラムと、

このパラメータ算出プログラムが算出した各フレーム毎の特徴パラメータを少なくとも2以上のフレームに跨って回帰分析し、特徴パラメータの時間的変化率を算出する変化率算出プログラムと、

この変化率算出プログラムにより算出した特徴パラメータの時間的変化率に適応させた分析処理を実行する分析処理プログラムと、

によって構成された音響信号分析プログラムを記憶したコンピュータが読取り可能な記憶媒体を提案する。

【0009】

作用

この発明による音響信号分析方法及び音響信号分析装置によれば、分析すべき入力信号の特徴パラメータを求めると共に、この特徴パラメータの時間的変化率を算出して、この特徴パラメータの時間的変化率の値から、信号の種類が例えば音声信号であるか、音楽信号であるか、或いはその中間的な音の信号であるかを分析する方法を採ったから、信頼性よく分析処理することができる。

【0010】

【発明の実施の形態】

図1にこの発明による音響信号分析装置10にその利用形態の一例を示す。この例では入

10

20

30

40

50

力信号 S を音響信号分析装置 10 に入力することにより、この入力信号の種類を例えば音声信号であるか、音楽信号であるか、或いはその中間の音の信号であるかを分析し、その分析結果に適応して、信号処理手段 20 は音響信号の分析結果に適応した信号処理を行う構成とした例を示す。

音響信号分析装置 10 では入力信号 S の特徴パラメータを求め、この特徴パラメータの時間的変化率を算出し、この時間的変化率の値に応じて入力信号の種類を判別する。つまり、特徴パラメータの時間的変化率が大きいほど音声信号らしさが強く、特徴パラメータの時間的変化率が小さいほど音楽信号らしさが強いと判定することができる。

【0011】

このように、入力信号 S の種類を特定することにより、その判定結果を利用することにより、次の処理ステップではそれぞれの信号に適した処理を行うことができる。

入力信号 S の種類を判別する具体的な方法を以下に説明する。この発明では、音声信号をフレームごとに分析し、特徴パラメータを算出する。この特徴パラメータを少なくとも 2 以上のフレームに跨って回帰分析し、フレーム間の変化率を算出する。

【0012】

特徴パラメータとしては、例えば LPC (線形予測) ケプストラム係数を適用することができる。図 2 に線形予測ケプストラム係数を適応した場合の音響信号分析装置 10 の一例を示す。線形予測ケプストラム係数は先ず線形予測係数算出手段 11 が音響信号からフレーム毎に P 次の線形予測係数を求める。この線形予測係数から特徴パラメータ算出手段 12 は一般的には 2P 次の線形予測ケプストラム係数 $y_{j, i}$ を求める。尚、音響信号から線形予測係数を求める方法及び線形予測係数から線形予測ケプストラム係数を求める方法は既によく知られており、例えば「守谷健弘著、音声符号化、電子情報通信学会編、2章」を参照。

【0013】

時間軸の複数のフレームにわたってケプストラム係数を求めた後で、変化率算出手段 13 A、13 B... 13 n はある次数 j の係数についてその時間軸 (フレーム番号の増加) に沿った変化の傾向 (時間変化率) を例えば回帰直線の傾きから求める。この時間変化率を求める処理をここでは変化率算出処理と称することにする。0 から m - 1 フレームの回帰係数は直線と実際の係数の 2 乗誤差の最小化から求める。尚、分析による遅延が許されない場合には現在のフレームを m とする。つまり、過去のフレーム 0 ~ m - 1 のフレームだけから分析する。遅延が許される場合には現在のフレーム (符号化の対象とするフレーム) を m / 2 とすることができる

実測値 y を近似するため、a, b をパラメータとする回帰直線を

$$y^{\wedge} = a x + b \tag{1}$$

とすると

【0014】

【数 1】

$$E = \begin{Bmatrix} x_0 & 1 \\ x_1 & 1 \\ \cdot & \cdot \\ x_{m-1} & 1 \end{Bmatrix} \begin{Bmatrix} a \\ b \end{Bmatrix} = \begin{Bmatrix} y_{j,0} \\ y_{j,1} \\ \cdot \\ y_{j,m-1} \end{Bmatrix} \tag{2}$$

【0015】

から $|E|^2$ を最小とする a, b を連立方程式から求める。実際は次の連立方程式の解と

なる。

【0016】

【数2】

$$\begin{Bmatrix} \sum_{i=0}^{m-1} x_i^2 & \sum_{i=0}^{m-1} x_i \\ \sum_{i=0}^{m-1} x_i & \sum_{i=0}^{m-1} 1 \end{Bmatrix} \begin{Bmatrix} a \\ b \end{Bmatrix} = \begin{Bmatrix} \sum_{i=0}^{m-1} y_{ji} x_i \\ \sum_{i=0}^{m-1} y_{ji} \end{Bmatrix} \quad (3)$$

10

【0017】

この処理を複数の次数 j で計算し、判定手段14は a_j のエネルギーの総和 A ($A = \sum_j a_j^2$) または A を時間方向又は振幅方向で正規化した値を求め、この値を変化率とする。尚、 a_j のエネルギーの総和 A は上記以外に重み w_j を付した $A = \sum_j w_j a_j^2$ 或いは $A = \sum_j |a_j|$ 等を用いることができる。また、ここでは x_i をフレームの番号として説明したが、サンプリング周波数やフレーム長によって x_i の実際の時間は異なる例えば1秒当たりの変化率を統一的に使う場合には、 a_j をフレーム更新周期で正規化した後で総エネルギー A を求めても良い。

【0018】

音声信号では、多くの音楽の信号と比較して、 A の値が大きくなる傾向があるので、この A の値そのものを利用して音声らしさ、音楽らしさ、中間的な音らしさを決定することができる。また、 A の値を長期的に平均化し、その平均化した A の値を判定基準に使うことにより音声信号か音楽か、或いは中間的な音かを判別することができる。図2に示す信号分析手段14は出力端子10A、10B、10Cを有し、その分析結果をこれらの出力端子10A、10B、10Cに出力し、その分析結果により信号処理手段20は音声信号に適応した信号処理、中間的な音の信号に適した信号処理、音楽に適した信号処理を実行する構成とした場合を示す。

20

【0019】

判定結果の利用方法としては例えば A の値に適応して音声らしさが強いときほど、信号の伝送路の帯域を狭くする方向に制限して入力信号 S を音声符号化処理ステップに送り込むことにより歪みの少ない音声符号化を実現することができる。また音楽らしさが強いときは帯域を広げて符号化処理ステップに送り込むことにより、音楽らしさを損うことなく符号化を実行することができる。入力信号の性格に適応した処理を実行することができる。上述では図3のように、線形(1次式)の回帰係数である傾きをパラメータとしているが、高次の回帰係数を利用することも可能である。例えば図4のように2次の場合は y を実測値、 y^{\wedge} を回帰推定値とすると、

30

$$y^{\wedge} = a x^2 + b x + c \quad (4)$$

で

$$E = \sum (y - y^{\wedge})^2 \quad (5)$$

) を最小にする a , b を連立方程式で求め、 a , b の絶対値で変化の程度が分かる。

40

【0020】

【数3】

$$E = \begin{pmatrix} x_0^2 & x_0 & 1 \\ x_1^2 & x_1 & 1 \\ \cdot & \cdot & \cdot \\ x_{m-1}^2 & x_{m-1} & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} y_{j,0} \\ y_{j,1} \\ \cdot \\ y_{j,m-1} \end{pmatrix} \quad (6)$$

10

$$\begin{pmatrix} \sum_{i=0}^{m-1} x_i^4 & \sum_{i=0}^{m-1} x_i^3 & \sum_{i=0}^{m-1} x_i^2 \\ \sum_{i=0}^{m-1} x_i^3 & \sum_{i=0}^{m-1} x_i^2 & \sum_{i=0}^{m-1} x_i \\ \sum_{i=0}^{m-2} x_i^2 & \sum_{i=0}^{m-1} x_i & \sum_{i=0}^{m-1} 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} \sum_{i=0}^{m-1} y_{j,i} x_i^2 \\ \sum_{i=0}^{m-1} y_{j,i} x_i \\ \sum_{i=0}^{m-1} y_{j,i} \end{pmatrix} \quad (7)$$

20

【0021】

換言すれば、

$$y = a x^2 + b x + c \\ = a (x - b/2a)^2 + c - b^2/4a$$

となるから、

aを一次回帰式の場合の係数 a_j と同じように扱うことができる。つまり、 a^2 が大きい程、音声らしさが強くなる。また、a、bを2次元に配置して音声らしい領域を実験的に抽出することも可能である。更に a^2 、 b^2 或いは $A = a^2$ 、 $B = b^2$ を2次元に配置してもよい。

30

【0022】

上述と同様の効果は信号波形のエネルギー、p次の線スペクトル対パラメータ、周波数軸を伸縮して変形した線形予測ケプストラム係数などにも同様に適用可能である。また上記係数 a_j を変形したり重みをつけて総和を求めたりすることでより判別精度を改善することも可能である。

上述した、音響信号からフレーム毎の特徴パラメータを求める特徴パラメータ算出手段12及び特徴パラメータの時間的変化率を求める変化率算出手段13A~13n、信号分析手段14はそれぞれコンピュータ上で動作するプログラムによって構成することができる。これらの各プログラムは記憶媒体に記憶されて販売されるか或いは通信回線を通じて要求先のコンピュータに伝送されてインストールされ、そのインストールにより売買が成立するなどの形態で販売される。

40

【0023】

【発明の効果】

以上説明したように、この発明によれば音声の特徴パラメータの時間的変化率を算出して音声らしさ或いは音楽らしさを判定する方法を採ったから、その判定は信頼性に富み、その判定結果を利用することにより、例えば音声に適した符号化処理或いは音楽に適した符号化処理を実行することができ、この結果として、音声に対しても、また音楽に対しても

50

歪みの小さい符号化を実現できる利点を得られる。

【0024】

尚、上述では信号の種別の例として音声、音楽、その中間的な音を例示して説明したが、この発明は特にこの音に限定されるものでなく、他のあらゆる音の判別に利用することが考えられよう。

【図面の簡単な説明】

【図1】この発明による音声信号判定装置の概要を説明するためのブロック図。

【図2】この発明による音声信号判定装置の要部の構成を説明するためのブロック図。

【図3】この発明の動作を説明するためのグラフ。

【図4】図3と同様のグラフ。

【符号の説明】

- 10 音響信号分析装置
- 11 線形予測係数算出手段
- 12 特徴パラメータ算出手段
- 13 A、13 B... 13 n 変化率算出手段
- 14 信号分析手段
- 20 信号処理手段

【図1】

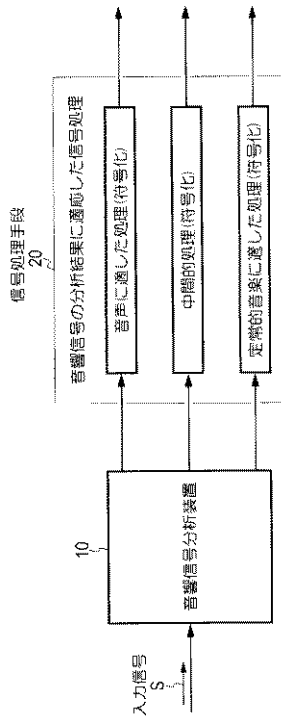


図 1

【図2】

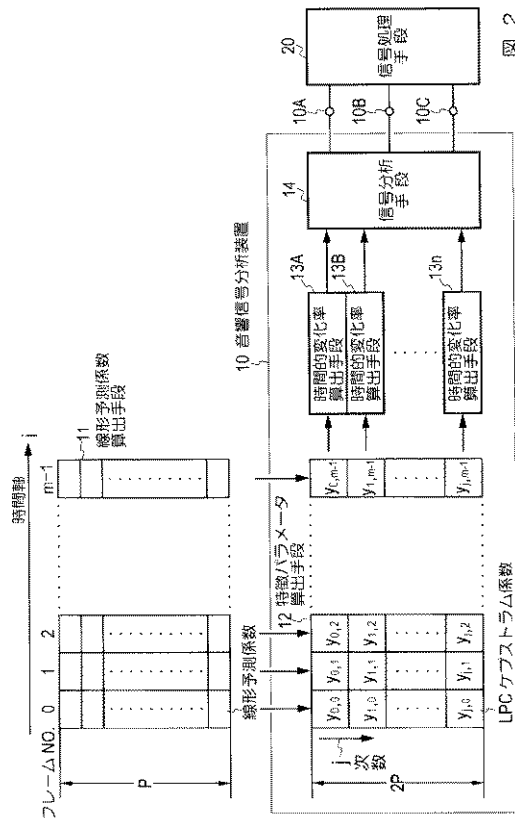


図 2

【図 3】

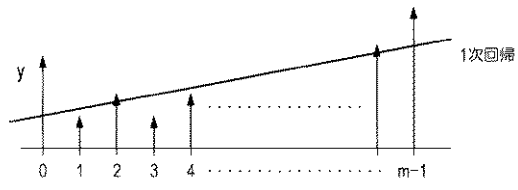


図 3

【図 4】

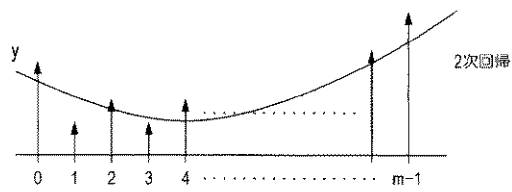


図 4

フロントページの続き

- (72)発明者 神 明夫
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 森 岳至
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

審査官 渡邊 聡

- (56)参考文献 特開2000-267699(JP,A)
特開平07-013586(JP,A)
特開平05-088695(JP,A)
特開平08-211897(JP,A)
特開平05-197397(JP,A)

- (58)調査した分野(Int.Cl.⁷, DB名)
G10L 11/00