

(19)日本国特許庁(JP)

(12)特許公報 (B 2)

(11)特許番号

特許第 3 2 8 5 1 7 8 号

(P 3 2 8 5 1 7 8)

(45)発行日 平成14年5月27日(2002.5.27)

(24)登録日 平成14年3月8日(2002.3.8)

(51)Int. Cl.⁷

識別記号

F I

G 1 0 L 21/02
11/02

G 1 0 L 9/14

F
D

請求項の数 9

(全 5 頁)

(21)出願番号 特願平6-188377
(22)出願日 平成6年8月10日(1994.8.10)
(65)公開番号 特開平8-54897
(43)公開日 平成8年2月27日(1996.2.27)
審査請求日 平成10年11月13日(1998.11.13)

(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
(72)発明者 岩上 直樹
東京都千代田区内幸町1丁目1番6号 日本
電信電話株式会社内
(72)発明者 守谷 健弘
東京都千代田区内幸町1丁目1番6号 日本
電信電話株式会社内
(72)発明者 三樹 聡
東京都千代田区内幸町1丁目1番6号 日本
電信電話株式会社内
(74)代理人 100066153
弁理士 草野 卓

審査官 山下 剛史

最終頁に続く

(54)【発明の名称】音信号立ち上がり検出方法

1

(57)【特許請求の範囲】

【請求項 1】 入力音響信号に対応する強度を求める第 1 の段階と、
その第 1 の段階で得られた強度に対して短時間の平滑化
を行い、短時間平滑化信号を得る第 2 の段階と、
その第 2 の段階で得られた短時間平滑化信号に対して長
時間の平滑化を行い、長時間平滑化信号を得る第 3 の段
階と、
その第 3 の段階で得られた長時間平滑化信号の不連続性
を計算し、得られた不連続点を立ち上がりとする第 4 の
段階とを有する音信号立ち上がり検出方法。

【請求項 2】 上記入力音響信号を線形予測分析し、そ
の結果の線形予測係数と対応したフィルタ係数をもつフ
ィルタに上記入力音響信号を通して残差信号を得、その
残差信号の強度を、上記第 1 の段階における上記入力音

2

響信号に対応する強度として用いることを特徴とする請
求項 1 記載の音信号立ち上がり検出方法。

【請求項 3】 上記第 3 の段階は、上記第 2 の段階で得
られた短時間平滑化信号にインパルス応答を畳み込んで
平滑化を行うことであることを特徴とする請求項 1 また
は 2 記載の音信号立ち上がり検出方法。

【請求項 4】 上記第 3 の段階は、上記第 2 の段階で得
られた短時間平滑化信号を、デジタルの検波回路に通
すことによって平滑化することであることを特徴とする
請求項 1 または 2 記載の音信号立ち上がり検出方法。

【請求項 5】 上記第 4 の段階は、上記第 3 の段階で得
られた過去の長時間平滑化信号から現在の長時間平滑化
信号を予測合成する第 5 の段階と、その第 5 の段階で得
られた予測合成された長時間平滑化信号と、実際の現在
における長時間平滑化信号とを比較し、後者が前者より

10

著しく大きい場合に立ち上がりと判断する第 6 の段階とからなることを特徴とする請求項 1 ないし 4 の何れかに記載の音信号立ち上がり検出方法。

【請求項 6】 上記第 5 の段階は、過去の長時間平滑化信号を線形和することにより予測合成を行うことを特徴とする請求項 5 記載の音信号立ち上がり検出方法。

【請求項 7】 上記第 5 の段階は、過去の長時間平滑化信号を近似表現する関数によって補外することによって予測合成を行うことを特徴とする請求項 5 記載の音信号立ち上がり検出方法。

【請求項 8】 上記第 6 の段階は、上記第 3 の段階で得られた現在の長時間平滑化信号と、上記第 5 の段階で得られた予測合成された長時間平滑化信号との比を計算し、この比があるしきい値を越えた場合立ち上がりと判断することを特徴とする請求項 5 ないし 7 記載の音信号立ち上がり検出方法。

【請求項 9】 上記第 6 の段階は、上記第 3 の段階で得られた現在の長時間平滑化信号と、上記第 5 の段階で得られた予測合成された長時間平滑化信号を定数倍したものととの差を計算し、この差があるしきい値を越えた場合立ち上がりと判断することを特徴とする請求項 5 ないし 7 記載の音信号立ち上がり検出方法。

【発明の詳細な説明】

【0001】

【産業上の利用分野】この発明は楽音信号や音声信号の符号化、音声認識、音声通信などに利用され、音響信号や音声信号の立ち上がりを検出する方法に関する。

【0002】

【従来の技術】現在、聴覚特性を考慮に入れた音信号の立ち上がり検出方法として、例えばオーディオ符号化方式であるエムペグ・オーディオ (MPEG-Audio) 方式において聴覚的エン트로ピーを用いた検出方法が提案されている。これは、複数の周波数小帯域ごとに、聴覚モデルを用いて計算した入力信号に対する最小可聴ノイズの信号対雑音比 (SNR) (すなわちその小帯域において符号化時に確保する必要がある SNR...必要 SNR) を求め、その必要 SNR から聴覚的エン트로ピーと呼ばれる値を計算し、この値があるしきい値を越えた時を立ち上がりとして検出する。なお、エムペグ・オーディオ符号化方式における立ち上がり検出方法についての詳細は ISO/IEC 標準 IS - 11172 - 3 に述べられている。

【0003】

【発明が解決しようとする課題】しかし、この方法では、必要 SNR を計算する際に、入力信号の周波数特性を求める必要があり、全帯域で 1024 次の次元の大きい FFT を行うなど大きな演算量を要する処理が多くあり、この方法により立ち上がり検出装置単体として構成すると規模が大きすぎる。

【0004】従来においては平滑化処理は行われておら

ず、入力信号が接近したパルス列のような場合は、その最初のパルスのみに対応して、これだけを音信号立ち上がりとして検出すべきであるが、従来においてはその各パルスを音信号立ち上がりとして検出してしまふ。つまり小さい立ち上がりの連続は聴覚的には立ち上がりとして感じないが、従来方法ではこの場合、小さい各立ち上りを音信号の立ち上がりと検出する問題があった。

【0005】

【課題を解決するための手段】この発明によれば、第 1 の段階で入力音響信号に対応する強度を求め、第 2 の段階でその強度に対して短時間、つまり 5 ~ 40 ミリ秒程度の平滑化を行って短時間平滑化信号を得、第 3 の段階でその短時間平滑化信号に対して長時間、つまり 20 ~ 150 ミリ秒程度の平滑化を行って長時間平滑化信号を得、第 4 の段階でその長時間平滑化信号の不連続性を計算し、その得られた不連続点を立ち上がりとする。

【0006】上記入力音響信号に対応する強度としては、入力音響信号を線形予測分析し、その結果の線形予測係数と対応したフィルタ係数をもつ逆フィルタへこの入力音響信号を通してその残差信号を得、その残差信号の強度を用いてもよい。上記第 3 の段階の平滑化は、短時間平滑化信号にインパルス応答を畳み込むことで行う、あるいは短時間平滑化信号をデジタルの検波回路に通すことにより行う。

【0007】上記第 4 の段階は、第 5 の段階で過去の長時間平滑化信号から現在の長時間平滑化信号を予測合成し、その予測合成された長時間平滑化信号を第 6 の段階で実際の現在における長時間平滑化信号と比較し、後者が前者より著しく大きい場合に前記不連続点、つまり立ち上がりと判断する。前記予測合成は過去の長時間平滑化信号を線形和することによるか、過去の長時間平滑化信号を近似表現する関数によって補外することにより行う。

【0008】

【作用】人の聴覚では、音信号のパタンの微細構造と全体の構造を認識する機能がある。このため、例えば単一のインパルスは音の立ち上がりとして認識されるが、このインパルスが速い周期で繰り返されると立ち上がりとしては認識されない。このような聴覚特性をもつ神経の興奮パターンをこの発明では、短時間平滑化と、長時間平滑化を持つ聴覚モデルによって模擬し、パタンの不連続性を判定することによって立ち上がりを検出する。したがって、この発明では、聴覚特性に即した立ち上りの検出が行われる。

【0009】また、線形予測分析フィルタを入力信号に通すと、音の立ち上がりのような過去の信号から予測できない不連続点が強調された残差信号が得られる。したがって、この残差信号の強さを入力音響信号と対応する強さとして用いるとより、いっそう検出感度が高くなる。

【0010】

【実施例】図1に、この発明の実施例における処理手順を示す。第1の段階11では、入力音響信号は3.2 KHzまたは4.8 KHzなどの標準化デジタル信号として入力され、その入力音響信号の音強度として、例えば入力音響信号の絶対値、二乗値を用いることができる。あるいは、入力音響信号を線形予測分析し、その線形予測係数と対応したフィルタ係数をもつフィルタ、いわゆる逆フィルタに入力音響信号を通して残差信号を得、この残差信号の絶対値、二乗値、またはこれを入力信号の振幅の平均値で正規化した信号を用いてもよい。

【0011】第2の段階12では、第1の段階11からの強度信号を例えば約5ms程度の頻度で、過去20ms程度の時間窓をかけ、窓内の平均を計算して短時間平滑化して短時間平滑化信号を得る。前記時間窓の窓形状は、例えば、矩形窓、ハミング窓、ハンニング窓、指数関数窓などを用い、時間窓は5～40ms程度とされる。第3の段階13では、第2の段階12からの短時間平滑化信号に、立ち上がりが速く立ち下りが遅いインパルス応答を畳み込むことによって長時間の平滑化を行う。この平滑化は20～150msにわたって連続的に行われる。このインパルス応答の波形としては図2Aに示すように何れも急に立ち上がり、徐々に立ち下がり、その立ち上がり、立ち下がりはいずれも直線的、指数関数的、三角関数的などである。あるいは図2Bに等価回路で示されるような立ち上がりの時定数が小さく、立ち下がりの時定数が大きい（立ち上がりの時定数は0でも良い）デジタルの検波回路に短時間平滑化信号を通して長時間の平滑化を行っても良い。このデジタル検波回路の処理は例えば図2Cに示すように入力された短時間平滑化信号系列 $x_0, x_1, \dots, x_{n-1}, x_n, x_{n+1}, \dots$ のそれぞれを、その直前に得られた出力の長時間平滑化信号系列 $y_0, y_1, \dots, y_{n-1}, y_n, y_{n+1}, \dots$ と比較し (21)、 $x_n > y_{n-1}$ ならば x_n を y_n として出力し (22)、 $x_n > y_{n-1}$ でなければ y_{n-1} に R ($0 < R < 1$) を乗算した値を y_n として出力する (23)。

【0012】第4の段階14では、過去における第3の段階13からの長時間平滑化信号より現在における長時間平滑化信号を予測する予測合成段階15と、その予測

合成出力と現在における第3の段階13の出力とを比較する比較段階16とにわかれる。予測合成段階15では、例えば過去の信号の線形和による予測合成、最小二乗法や最尤推定法による関数の当てはめなどの手法を用いて過去の信号から現在の信号の予測値を合成する。つまり、長時間平滑化信号 y_n の系列を、各信号時間だけの単位遅延回路25の直列接続に入力し、その各遅延回路25の出力に乗算器26で定数を掛算し、その掛算結果を加算器27で加算して予測値 y_{n+1} として出力する。あるいは図3Bに示すように長時間平滑化信号 y_n の系列を時間に対する関数28として近似し、その関数で補外して予測値 y_{n+1} を得る。

【0013】比較段階16では、現在における長時間平滑化信号 y_{n+1} とその予測合成信号 y_{n+1} との比を y_{n+1} / y_{n+1} を計算し、この値があるしきい値を越えていたら立ち上がりとする、あるいは、現在における信号 y_{n+1} と予測合成信号 y_{n+1} との差 ($y_{n+1} - y_{n+1}$) = y を計算し、この値 y があるしきい値を超えていたら立ち上がりとする。あるいは、現在における信号 y_{n+1} と予測合成信号を定数倍したものの $A y_{n+1}$ との差を計算し、この値があるしきい値を超えていたら立ち上がりとする方法を用いる。

【0014】

【発明の効果】以上述べたように、この発明では、短時間平滑化と、長時間平滑化とにより神経の興奮パターンを模擬し、この信号を立ち上がりを判定に用いるので、より聴覚特性に即した立ち上がりの検出が可能である。

【図面の簡単な説明】

【図1】この発明の実施例の手順を示す流れ図。

【図2】Aは長時間平滑化において畳み込みを行う場合に用いるインパルス応答の形状の例を示す図、Bは長時間平滑化に用いるデジタル検波回路の等価回路を示す図、Cはこの処理を離散信号に施す手順を示す流れ図である。

【図3】Aは連続性検出における予測合成を過去の長時間平滑化信号の線形和により求める場合の構成例を示す図、Bは過去の長時間平滑化信号を近似表現する関数の例を示す図である。

【図 1】

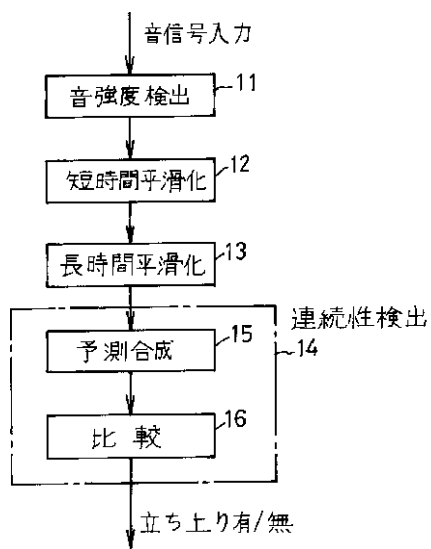


図 1

【図 2】

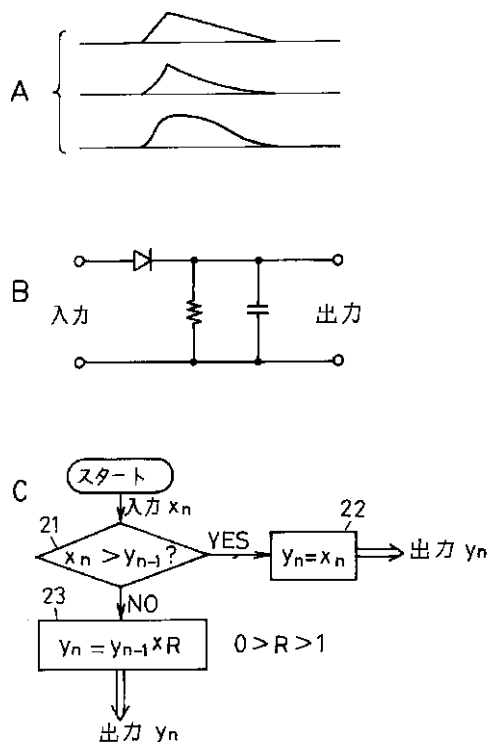


図 2

【図 3】

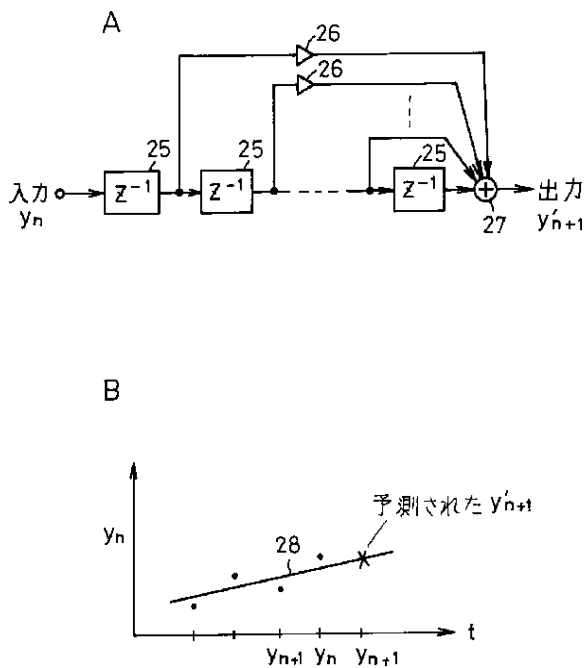


図 3

フロントページの続き

(56)参考文献 特開 昭61 - 46999 (J P , A)
特開 平 3 - 253899 (J P , A)
特開 平 4 - 67200 (J P , A)
特開 平 6 - 19498 (J P , A)
特開 平 6 - 27995 (J P , A)

(58)調査した分野(Int.Cl.⁷, D B 名)
G10L 11/02, 21/02