

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4150798号
(P4150798)

(45) 発行日 平成20年9月17日(2008.9.17)

(24) 登録日 平成20年7月11日(2008.7.11)

(51) Int.Cl. F I
G 1 0 L 15/02 (2006.01) G 1 0 L 15/02 3 0 0 K
G 1 0 L 15/20 (2006.01) G 1 0 L 15/20 3 6 0 B

請求項の数 32 (全 37 頁)

<p>(21) 出願番号 特願2006-529254 (P2006-529254)</p> <p>(86) (22) 出願日 平成17年7月20日 (2005.7.20)</p> <p>(86) 国際出願番号 PCT/JP2005/013334</p> <p>(87) 国際公開番号 W02006/011405</p> <p>(87) 国際公開日 平成18年2月2日 (2006.2.2)</p> <p>審査請求日 平成18年12月25日 (2006.12.25)</p> <p>(31) 優先権主張番号 特願2004-220022 (P2004-220022)</p> <p>(32) 優先日 平成16年7月28日 (2004.7.28)</p> <p>(33) 優先権主張国 日本国(JP)</p> <p>早期審査対象出願</p>	<p>(73) 特許権者 304020292 国立大学法人徳島大学 徳島県徳島市新蔵町2丁目24番地</p> <p>(74) 代理人 100104949 弁理士 豊栖 康司</p> <p>(74) 代理人 100074354 弁理士 豊栖 康弘</p> <p>(72) 発明者 赤松 則男 徳島県徳島市南常三島町2丁目1番地 国立大学法人徳島大学工学部知能情報学科内</p> <p>審査官 間宮 嘉誉</p>
---	---

最終頁に続く

(54) 【発明の名称】 デジタルフィルタリング方法、デジタルフィルタ装置、デジタルフィルタプログラム及びコンピュータで読み取り可能な記録媒体

(57) 【特許請求の範囲】

【請求項1】

入力された信号波形に基づいて特徴量を抽出するデジタルフィルタリング方法であって、

信号波形を入力信号として入力し、信号波形を量子化して量子化データを取得し、メモリ手段に格納するステップと、

前記量子化データにおいて、任意の点に隣接する所定の範囲を加算区間として予め設定し、及び/又は加算区間に含まれるデータ数を予め設定しておくことにより、前記メモリ手段に格納された量子化データの各点毎に、加算区間に含まれるすべての量子化データのデータ値を加算して加算値を求め、前記加算値をデータ数で除算して除算値を求め、さらに必要に応じて該点の量子化データのデータ値と前記除算値との差を減算により減算値として演算し、前記除算値及び減算値を出力するステップと、

前記ステップを繰り返して各量子化データ毎に前記除算値及び減算値の演算を行い、各々出力される各除算値及び減算値に基づいて、入力された信号波形のフィルタリングを行うステップであって、前記減算値に基づいてハイパスフィルタリングを行う一方、前記除算値に基づいてローパスフィルタリングを行うステップと、を有することを特徴とするデジタルフィルタリング方法。

【請求項2】

請求項1に記載のデジタルフィルタリング方法であって、

前記加算区間が、任意の点に対して略対称となるよう所定の範囲が設定されることを特

徴とするデジタルフィルタリング方法。

【請求項 3】

請求項 1 又は 2 に記載のデジタルフィルタリング方法であって、
前記デジタルフィルタリングを行うステップが、
各点のデータのデータ値と各々の除算値とを比較し、比較結果の真偽に基づいて凹凸波形に変換するステップを含むことを特徴とするデジタルフィルタリング方法。

【請求項 4】

請求項 1 から 3 のいずれか一に記載のデジタルフィルタリング方法であって、
加算、減算、除算を整数演算で行うことを特徴とするデジタルフィルタリング方法。

【請求項 5】

請求項 1 から 4 のいずれか一に記載のデジタルフィルタリング方法であって、
前記所定の範囲の加算すべきデータの個数を 2 のべき乗とすることを特徴とするデジタルフィルタリング方法。

【請求項 6】

請求項 5 に記載のデジタルフィルタリング方法であって、
前記所定の範囲の加算したデータの個数での除算を、ビットシフト演算で行うことを特徴とするデジタルフィルタリング方法。

【請求項 7】

請求項 1 から 6 のいずれか一に記載のデジタルフィルタリング方法であって、
前記除算値の演算ステップにおいて、各々のデータにつき除算値を求めるために所定の範囲のデータのデータ値を加算した加算値を保持しておき、次のデータの加算値を求める際に、保持された加算値から、不要なデータ値を減算すると共に、必要なデータ値を加算することで、加算値を演算することを特徴とするデジタルフィルタリング方法。

【請求項 8】

請求項 7 に記載のデジタルフィルタリング方法であって、k 点を中心とする前後 n の区間 $N (= 2n)$ における除算値 α_k を、

【数 1】

$$\alpha_k = \frac{\left(\sum_{i=k+1}^{n+k} x_i \right) + \left(\sum_{i=k-n}^{k-1} x_i \right)}{N}$$

として表現する際、除算値の演算ステップにおいて、除算値 α_k を、その前段の位置である $(k-1)$ 点における除算値 α_{k-1} を用いて

【数 2】

$$\alpha_k = \alpha_{k-1} + (x_{k-1} + x_{n+k} - x_k - x_{k-1-n}) / N$$

で演算することを特徴とするデジタルフィルタリング方法。

【請求項 9】

請求項 1 から 8 のいずれか一に記載のデジタルフィルタリング方法であって、
 $N = 64$ 又は 256 であることを特徴とするデジタルフィルタリング方法。

【請求項 10】

請求項 1 から 8 のいずれか一に記載のデジタルフィルタリング方法であって、
特徴量が、 $N = 256$ として凹凸波形を参照し、母音波形の 1 周期の始まりから探索して得られる最初の凸部分の幅であることを特徴とするデジタルフィルタリング方法。

【請求項 11】

請求項 1 から 8 のいずれか一に記載のデジタルフィルタリング方法であって、
特徴量が、 $N = 256$ として凹凸波形を参照し、母音波形の 1 周期の始まりから探索し

10

20

30

40

50

て得られる最初の凸部分の幅に存在する振幅波形の面積であることを特徴とするデジタルフィルタリング方法。

【請求項 1 2】

請求項 1 から 8 のいずれかーに記載のデジタルフィルタリング方法であって、
特徴量が、 $N = 256$ として凹凸波形を参照し、母音波形の 1 周期の始まりから探索して得られる最初の凸部分の幅に存在する振幅波形の分散値であることを特徴とするデジタルフィルタリング方法。

【請求項 1 3】

請求項 1 から 8 のいずれかーに記載のデジタルフィルタリング方法であって、
特徴量が、 $N = 256$ として凹凸波形を参照し、母音波形の 1 周期の始まりから探索して得られる最初の凸部分の幅に存在する振幅波形を $0 \sim 1$ に正規化を行い、凸幅で生成した正弦波との類似性であることを特徴とするデジタルフィルタリング方法。

10

【請求項 1 4】

請求項 1 から 8 のいずれかーに記載のデジタルフィルタリング方法であって、
特徴量が、 $N = 64$ として 3 周期分の凹凸波形に存在する凸の数であることを特徴とするデジタルフィルタリング方法。

【請求項 1 5】

請求項 1 から 1 4 のいずれかーに記載のデジタルフィルタリング方法であって、
入力された信号波形が音信号であることを特徴とするデジタルフィルタリング方法。

【請求項 1 6】

請求項 1 から 1 4 のいずれかーに記載のデジタルフィルタリング方法であって、
入力された信号波形が電磁波であることを特徴とするデジタルフィルタリング方法。

20

【請求項 1 7】

請求項 1 から 1 4 のいずれかーに記載のデジタルフィルタリング方法であって、
入力された信号波形が生体信号であることを特徴とするデジタルフィルタリング方法。

【請求項 1 8】

入力された音声信号に基づいて音声認識を行うため音声信号から特徴量を抽出するデジタルフィルタリング方法であって、

音声信号を入力し、入力信号の振幅を量子化するステップと、

量子化された各点のデータにつき、隣接する所定の範囲のデータの振幅値を加算し、これを加算したデータの個数で除算して該データを中心とする除算値を求めるステップと、
各点のデータの振幅値と各々の除算値とを比較し、比較結果の真偽に基づいて凹凸波形に変換するステップと、

30

前記凹凸波形から特徴量を抽出して音声認識を行うステップと、

認識後の音声につき、子音部が認識されると音量を大きくし、子音部の後に母音が認識されると、母音部から所定時間で音量増幅を解除するステップと、
を備えることを特徴とするデジタルフィルタリング方法。

【請求項 1 9】

請求項 1 8 に記載のデジタルフィルタリング方法であって、さらに、

母音識別を行なうために母音部分の 1 周期波形であるピッチを参照して母音波形の定常とする区間を抽出するステップを備えることを特徴とするデジタルフィルタリング方法。

40

【請求項 2 0】

請求項 1 9 に記載のデジタルフィルタリング方法であって、

前記母音波形の定常区間の抽出ステップにおいて、全ピッチ数の約 $\frac{2}{3}$ の位置にあたるピッチを中心として前後の 1 周期分を抽出することを特徴とするデジタルフィルタリング方法。

【請求項 2 1】

請求項 1 9 又は 2 0 に記載のデジタルフィルタリング方法であって、

前記母音識別が 2 母音の組み合わせから得られる特徴量に基づいて行われることを特徴とするデジタルフィルタリング方法。

50

【請求項 22】

請求項 19 から 21 のいずれかーに記載のデジタルフィルタリング方法であって、
前記母音識別が、2つの母音の組み合わせによって得られる特徴量をマッピングした離散ボロノイ図に基づき境界線を求め、この境界線を基準として投票形式により行われることを特徴とするデジタルフィルタリング方法。

【請求項 23】

入力された信号波形に基づいて特徴量を抽出するデジタルフィルタ装置であって、
入力された信号波形の振幅を求め、前記振幅を量子化する振幅算出部と、
前記振幅算出部で量子化された各点のデータにつき、隣接する所定の範囲のデータの振幅値を加算し、これを加算したデータの個数で除算して該データを中心とする除算値を求め
10

除算値演算部と、
各データにつき、前記除算値演算部で演算された除算値と、前記振幅演算部で演算された振幅値との差を減算によって求める減算手段と、

加算値、減算値、除算値の少なくともいずれかを保持するメモリ手段と、

入力された信号波形に対し、前記除算値演算部で得られた除算値に基づいてローパスフィルタリングを行う一方、前記減算手段で得られた減算値に基づいてハイパスフィルタリングを行うフィルタリング手段と、

を備えることを特徴とするデジタルフィルタ装置。

【請求項 24】

請求項 23 に記載のデジタルフィルタ装置であって、さらに、
各データにつき、前記除算値演算部で演算された除算値と、前記振幅演算部で演算された振幅値とを比較し、比較結果の真偽を出力する比較部と、
20

前記比較部の出力に基づいて、信号波形を凹凸波形に変換する変換部と、
を備えることを特徴とするデジタルフィルタ装置。

【請求項 25】

請求項 23 又は 24 に記載のデジタルフィルタ装置であって、
前記除算値演算部が、所定の範囲の加算すべきデータの個数を2のべき乗として、前記所定の範囲の加算したデータの個数での除算を、ビットシフト演算で行うことを特徴とするデジタルフィルタ装置。

【請求項 26】

請求項 23 から 25 のいずれかーに記載のデジタルフィルタ装置であって、
前記除算値演算部が、各々のデータにつき除算値を求めるために所定の範囲のデータの振幅値を加算した加算値を保持しておき、次のデータの加算値を求める際に、保持された加算値から、不要な振幅値を減算すると共に、必要な振幅値を加算することで、加算値を演算することを特徴とするデジタルフィルタ装置。

【請求項 27】

請求項 23 から 26 のいずれかーに記載のデジタルフィルタ装置であって、
入力された信号波形が音信号であることを特徴とするデジタルフィルタ装置。

【請求項 28】

請求項 23 から 26 のいずれかーに記載のデジタルフィルタ装置であって、
入力された信号波形が電磁波であることを特徴とするデジタルフィルタ装置。

【請求項 29】

請求項 23 から 26 のいずれかーに記載のデジタルフィルタ装置であって、
入力された信号波形が生体信号であることを特徴とするデジタルフィルタ装置。

【請求項 30】

入力された音声信号に基づいて音声認識を行うため音声信号から特徴量を抽出するデジタルフィルタ装置であって、

入力された音声信号の振幅を求め、前記振幅を量子化する振幅算出部と、

前記振幅算出部で量子化された各点のデータにつき、隣接する所定の範囲のデータの振幅値を加算し、これを加算したデータの個数で除算して該データを中心とする除算値を求
50

める除算値演算部と、

各データにつき、前記除算値演算部で演算された除算値と、前記振幅演算部で演算された振幅値とを比較し、比較結果の真偽を出力する比較部と、

前記比較部の出力に基づいて、音声信号を凹凸波形に変換する変換部と、

入力された信号波形に対し、前記除算値演算部で得られた除算値に基づいてローパスフィルタリングを行う一方、前記減算手段で得られた減算値に基づいてハイパスフィルタリングを行うフィルタリング手段と、

を備えることを特徴とするデジタルフィルタ装置。

【請求項 3 1】

入力された信号波形に基づいて特徴量を抽出するデジタルフィルタプログラムであって

10

、
信号波形を入力し、入力信号の振幅を量子化する機能と、

量子化された各点のデータにつき、隣接する所定の範囲のデータの振幅値を加算し、これを加算したデータの個数で除算して該データを中心とする除算値を求める機能と、

各点のデータの振幅値と各々の除算値との差を演算する機能と、

演算された加算値、減算値、除算値の少なくともいずれかをメモリ手段に保持すると共に、これら加算値、減算値、除算値の少なくともいずれかの値に基づいて入力された信号波形のデジタルフィルタリングを行う機能であって、減算値に基づいてハイパスフィルタリングを行う一方、除算値に基づいてローパスフィルタリングを行う機能と、

をコンピュータに実現させることを特徴とするデジタルフィルタプログラム。

20

【請求項 3 2】

請求項 3 1 に記載されるプログラムを格納したコンピュータで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、アナログ信号をデジタル処理するデジタルフィルタに関するデジタルフィルタリング方法、デジタルフィルタ装置、デジタルフィルタプログラム及びコンピュータで読み取り可能な記録媒体並びに記録した機器に関し、例えば音声認識や補聴器等、音声の再現の他、画像処理等に利用可能なデジタルフィルタリング方法等に関する。

30

【背景技術】

【0002】

音声データを取得し、このデータを処理して発話された内容を認識する音声認識装置が実用化されている。例えば、発話によって文章入力を行うアプリケーションや発話によって操作する車載ナビゲーションシステム等には、音声認識エンジンが内蔵されている。

【0003】

一般的に、音声認識では入力音声信号からその特徴を良く表す少数のパラメータ（音声特徴量という）を抽出し、これを予め登録してあるモデルの特徴量と照合して、最も近いモデルを認識結果とする方法が採用される。この方法において、音声特徴量の果たす役割は重要であり、より少ないパラメータ数で精度良く音声の特徴を表すようにするための様々な研究が行われている。

40

【0004】

音声認識に使用される代表的な特徴ベクトルとしては、バンドパスフィルタ又はフーリエ変換によって求めることができるパワースペクトラムや、LPC（線形予測）分析によって求められるケプストラム係数等がよく知られている。特徴ベクトル抽出手段により抽出された特徴ベクトルの時系列は、パターン照合手段に送出され、認識される（例えば特許文献 1 参照）。

【0005】

しかしながら、このような方法ではパターン照合を行う際の特徴量の抽出のための処理

50

演算が極めて複雑になり、高い処理能力が要求される。その結果、音声認識のための構造が複雑になり、また処理時間も長くなるという問題があった。

【0006】

一方で、携帯電話の普及や携帯電子端末の小型化に伴い、このような機器における入出力のためのインターフェースとして音声認識技術が注目されている。最近では、隠れマルコフモデル(HMM: hidden Markov model)を用いたワード・スポッティングや連続音声認識等、発声者の発声方式に制約を要求しない、自然発話認識のための研究や開発が盛んに行われている。

【0007】

また一方で、難聴者のための補聴器や音声を再生するスピーカ等においても、音声が聞き取りやすくなるように音声波形を処理する技術も注目を集めている。このような携帯型機器への応用においては、できる限りシンプルなアルゴリズムであり、しかも高速に処理できる音声認識技術が求められる。しかしながら、実用可能なレベルで音声認識のための演算処理を簡素化した補聴器は未だ開発されていなかった。

【0008】

また音声認識に限らず、アナログ信号の解析処理一般において、スペクトル分析等にフーリエ変換が利用されることが多い。しかしながらフーリエ変換や逆フーリエ変換を実現するには演算が複雑で処理量が多くなるため、高い処理能力が求められ、これを実現するハードウェアも回路構成が複雑となり、高価になるという問題があった。

【特許文献1】特開2003-271190号公報

【発明の開示】

【発明が解決しようとする課題】

【0009】

本発明は、このような問題点を解決するためになされたものである。本発明の主な目的は、波形信号処理において演算を簡素化し、高速かつ低負荷の処理とすることで実装や組み込みを容易にしたデジタルフィルタリング方法、デジタルフィルタ装置、デジタルフィルタプログラム及びコンピュータで読み取り可能な記録媒体並びに記録した機器を提供することにある。

【課題を解決するための手段】

【0010】

上記の目的を達成するために、本発明の第1の側面に係るデジタルフィルタリング方法は、入力された信号波形に基づいて特徴量を抽出するデジタルフィルタリング方法であって、信号波形を入力信号として入力し、信号波形を量子化して量子化データを取得し、メモリ手段に格納するステップと、前記量子化データにおいて、任意の点に隣接する所定の範囲を加算区間として予め設定し、及び/又は加算区間に含まれるデータ数を予め設定しておくことにより、前記メモリ手段に格納された量子化データの各点毎に、加算区間に含まれるすべての量子化データのデータ値を加算して加算値を求め、前記加算値をデータ数で除算して除算値を求め、さらに必要に応じて該点の量子化データのデータ値と前記除算値との差を減算により減算値として演算し、前記除算値及び減算値を出力するステップと、前記ステップを繰り返して各量子化データ毎に前記除算値及び減算値の演算を行い、各々出力される各除算値及び減算値に基づいて、入力された信号波形のフィルタリングを行うステップであって、前記減算値に基づいてハイパスフィルタリングを行う一方、前記除算値に基づいてローパスフィルタリングを行うステップとを有する。

【0011】

また、本発明の第2の側面に係るデジタルフィルタリング方法は、前記加算区間が、任意の点に対して略対称となるよう所定の範囲が設定される。

【0012】

【0013】

さらに、本発明の第3の側面に係るデジタルフィルタリング方法は、前記デジタルフィルタリングを行うステップが、各点のデータのデータ値と各々の除算値とを比較し、比較

10

20

30

40

50

結果の真偽に基づいて凹凸波形に変換するステップを含む。これにより、整数の加算と比較演算で特徴量の抽出が可能となり、演算処理を極めて簡素化して高速、安価に行うことができる。

【0014】

さらに、本発明の第4の側面に係るデジタルフィルタリング方法は、加算、減算、除算を整数演算で行う。これによって浮動小数点演算を排し、固定小数点演算のみで演算を行うことができ、高速かつ低負荷で演算処理を行うことができる。

【0015】

さらに、本発明の第5の側面に係るデジタルフィルタリング方法は、前記所定の範囲の加算すべきデータの個数を2のべき乗とする。これにより、除算の際にビット・シフト演算が可能となり、さらに演算処理を簡素化でき、高速化に寄与し得る。

10

【0016】

さらに、本発明の第6の側面に係るデジタルフィルタリング方法は、前記所定の範囲の加算したデータの個数での除算を、ビットシフト演算で行う。これにより、除算の際にビット・シフト演算が可能となり、さらに演算処理を簡素化でき、高速化に寄与し得る。

【0017】

さらに、本発明の第7の側面に係るデジタルフィルタリング方法は、前記除算値の演算ステップにおいて、各々のデータにつき除算値を求めるために所定の範囲のデータのデータ値を加算した加算値を保持しておき、次のデータの加算値を求める際に、保持された加算値から、不要なデータ値を減算すると共に、必要なデータ値を加算することで、加算値を演算する。これにより、各除算値演算において加算値を求める際に、前回のデータについて演算した加算値を利用して、必要なデータの入れ替えによって所望の加算値とすることができ、加算演算を大幅に簡素化でき、演算処理をさらに高速化することが可能となる。

20

【0018】

さらに、本発明の第8の側面に係るデジタルフィルタリング方法は、k点を中心とする前後nの区間N(=2n)における除算値 α_k を、

【0019】

【数3】

$$\alpha_k = \frac{\left(\sum_{i=k+1}^{n+k} x_i \right) + \left(\sum_{i=k-n}^{k-1} x_i \right)}{N}$$

30

として表現する際、除算値の演算ステップにおいて、除算値 α_k を、その前段の位置である(k-1)点における除算値 α_{k-1} を用いて

【0020】

【数4】

$$\alpha_k = \alpha_{k-1} + (x_{k-1} + x_{n+k} - x_k - x_{k-1-n}) / N$$

40

で演算する。これにより、各除算値を前段の除算値を利用して逐次的に求めることができ、演算処理量を大幅に低減して高速かつ低負荷な信号波形特徴量抽出処理を実現する。

【0021】

さらに、本発明の第9の側面に係るデジタルフィルタリング方法は、N=64又は256である。

【0022】

さらに、本発明の第10の側面に係るデジタルフィルタリング方法は、特徴量が、N=256として凹凸波形を参照し、母音波形の1周期の始まりから探索して得られる最初の

50

凸部分の幅である。

【0023】

さらに、本発明の第11の側面に係るデジタルフィルタリング方法は、特徴量が、 $N = 256$ として凹凸波形を参照し、母音波形の1周期の始まりから探索して得られる最初の凸部分の幅に存在する振幅波形の面積である。

【0024】

さらに、本発明の第12の側面に係るデジタルフィルタリング方法は、特徴量が、 $N = 256$ として凹凸波形を参照し、母音波形の1周期の始まりから探索して得られる最初の凸部分の幅に存在する振幅波形の分散値である。

【0025】

さらに、本発明の第13の側面に係るデジタルフィルタリング方法は、特徴量が、 $N = 256$ として凹凸波形を参照し、母音波形の1周期の始まりから探索して得られる最初の凸部分の幅に存在する振幅波形を0～1に正規化を行い、凸幅で生成した正弦波との類似性である。

【0026】

さらに、本発明の第14の側面に係るデジタルフィルタリング方法は、特徴量が、 $N = 64$ として3周期分の凹凸波形に存在する凸の数である。

【0027】

さらに、本発明の第15の側面に係るデジタルフィルタリング方法は、入力された信号波形が音信号である。

【0028】

さらに、本発明の第16の側面に係るデジタルフィルタリング方法は、入力された信号波形が電磁波である。

【0029】

さらに、本発明の第17の側面に係るデジタルフィルタリング方法は、入力された信号波形が生体信号である。

【0030】

さらに、本発明の第18の側面に係るデジタルフィルタリング方法は、入力された音声信号に基づいて音声認識を行うため音声信号から特徴量を抽出するデジタルフィルタリング方法であって、音声信号を入力し、入力信号の振幅を量子化するステップと、量子化された各点のデータにつき、隣接する所定の範囲のデータの振幅値を加算し、これを加算したデータの個数で除算して該データを中心とする除算値を求めるステップと、各点のデータの振幅値と各々の除算値とを比較し、比較結果の真偽に基づいて凹凸波形に変換するステップと、前記凹凸波形から特徴量を抽出して音声認識を行うステップと、認識後の音声につき、子音部が認識されると音量を大きくし、子音部の後に母音が認識されると、母音部から所定時間で音量増幅を解除するステップとを備える。これにより、整数の加算と比較演算で特徴量の抽出が可能となり、演算処理を極めて簡素化して高速、安価に行うことができる。また、聴感上は音声出力を大きくしたように聴取され、音声の認識が容易になる一方で、常に大音量の音声を聞いている訳ではないので、大音量による不快感は極めて少なくなる。特に子音部の直後に表れる母音部の音量は大きいので、子音とその直後の母音により音声は認識され易くなる。この際に母音部は、1周期から数周期の短時間の波形により認識される。聴覚と脳機能による短時間の音声認識が完了すると、それ以後は母音の継続であるから、その期間において音声の出力信号を小さくでき、再び子音が入力される時は、音声出力を大きくして音声認識を容易にする。

【0031】

さらに、本発明の第19の側面に係るデジタルフィルタリング方法はさらに、母音識別を行なうために母音部分の1周期波形であるピッチを参照して母音波形の定常とする区間を抽出するステップを備える。

【0032】

さらに、本発明の第20の側面に係るデジタルフィルタリング方法は、前記母音波形の

10

20

30

40

50

定常区間の抽出ステップにおいて、全ピッチ数の約3分の2の位置にあたるピッチを中心として前後の1周期分を抽出する。

【0033】

さらに、本発明の第21の側面に係るデジタルフィルタリング方法は、前記母音識別が2母音の組み合わせから得られる特徴量に基づいて行われる。

【0034】

さらに、本発明の第22の側面に係るデジタルフィルタリング方法は、前記母音識別が、2つの母音の組み合わせによって得られる特徴量をマッピングした離散ポロノイ図に基づき境界線を求め、この境界線を基準として投票形式により行われる。

【0035】

【0036】

さらに、本発明の第23の側面に係るデジタルフィルタ装置は、入力された信号波形に基づいて特徴量を抽出するデジタルフィルタ装置であって、入力された信号波形の振幅を求め、前記振幅を量子化する振幅算出部と、前記振幅算出部で量子化された各点のデータにつき、隣接する所定の範囲のデータの振幅値を加算し、これを加算したデータの個数で除算して該データを中心とする除算値を求める除算値演算部と、各データにつき、前記除算値演算部で演算された除算値と、前記振幅演算部で演算された振幅値との差を減算によって求める減算手段と、加算値、減算値、除算値の少なくともいずれかを保持するメモリ手段と、入力された信号波形に対し、前記除算値演算部で得られた除算値に基づいてローパスフィルタリングを行う一方、前記減算手段で得られた減算値に基づいてハイパスフィルタリングを行うフィルタリング手段とを備える。

【0037】

さらに、本発明の第24の側面に係るデジタルフィルタ装置はさらに、各データにつき、前記除算値演算部で演算された除算値と、前記振幅演算部で演算された振幅値とを比較し、比較結果の真偽を出力する比較部と、前記比較部の出力に基づいて、信号波形を凹凸波形に変換する変換部とを備える。これにより、整数の加算と比較演算で特徴量の抽出が可能となり、演算処理を極めて簡素化して高速、安価に行うことができる。

【0038】

さらに、本発明の第25の側面に係るデジタルフィルタ装置は、前記除算値演算部が、所定の範囲の加算すべきデータの個数を2のべき乗として、前記所定の範囲の加算したデータの個数での除算を、ビットシフト演算で行う。

【0039】

さらに、本発明の第26の側面に係るデジタルフィルタ装置は、前記除算値演算部が、各々のデータにつき除算値を求めるために所定の範囲のデータの振幅値を加算した加算値を保持しておき、次のデータの加算値を求める際に、保持された加算値から、不要な振幅値を減算すると共に、必要な振幅値を加算することで、加算値を演算する。これによって計算量を減らし、処理の軽減と高速化を図ることができる。

【0040】

さらに、本発明の第27の側面に係るデジタルフィルタ装置は、入力された信号波形が音信号である。

【0041】

さらに、本発明の第28の側面に係るデジタルフィルタ装置は、入力された信号波形が電磁波である。

【0042】

さらに、本発明の第29の側面に係るデジタルフィルタ装置は、入力された信号波形が生体信号である。

【0043】

さらに、本発明の第30の側面に係るデジタルフィルタ装置は、入力された音声信号に基づいて音声認識を行うため音声信号から特徴量を抽出するデジタルフィルタ装置であって、入力された音声信号の振幅を求め、前記振幅を量子化する振幅算出部と、前記振幅算

10

20

30

40

50

出部で量子化された各点のデータにつき、隣接する所定の範囲のデータの振幅値を加算し、これを加算したデータの個数で除算して該データを中心とする除算値を求める除算値演算部と、各データにつき、前記除算値演算部で演算された除算値と、前記振幅演算部で演算された振幅値とを比較し、比較結果の真偽を出力する比較部と、前記比較部の出力に基づいて、音声信号を凹凸波形に変換する変換部と、入力された信号波形に対し、前記除算値演算部で得られた除算値に基づいてローパスフィルタリングを行う一方、前記減算手段で得られた減算値に基づいてハイパスフィルタリングを行うフィルタリング手段とを備える。これにより、整数の加算と比較演算で特徴量の抽出が可能となり、演算処理を極めて簡素化して高速、安価に行うことができる。

【0044】

さらに、本発明の第31の側面に係るデジタルフィルタプログラムは、入力された信号波形に基づいて特徴量を抽出するデジタルフィルタプログラムであって、信号波形を入力し、入力信号の振幅を量子化する機能と、量子化された各点のデータにつき、隣接する所定の範囲のデータの振幅値を加算し、これを加算したデータの個数で除算して該データを中心とする除算値を求める機能と、各点のデータの振幅値と各々の除算値との差を演算する機能と、演算された加算値、減算値、除算値の少なくともいずれかをメモリ手段に保持すると共に、これら加算値、減算値、除算値の少なくともいずれかの値に基づいて入力された信号波形のデジタルフィルタリングを行う機能であって、減算値に基づいてハイパスフィルタリングを行う一方、除算値に基づいてローパスフィルタリングを行う機能とをコンピュータに実現させる。これにより、整数の加算と比較演算で特徴量の抽出が可能となり、演算処理を極めて簡素化して高速、安価に行うことができる。

【0045】

また本発明の第32の側面に係るコンピュータで読み取り可能な記録媒体は、上記プログラムを格納するものである。記録媒体には、CD-ROM、CD-R、CD-RWやフレキシブルディスク、磁気テープ、MO、DVD-ROM、DVD-RAM、DVD-R、DVD+R、DVD-RW、DVD+RW、Blue-ray、HD DVD(AOD)等の磁気ディスク、光ディスク、光磁気ディスク、半導体メモリその他のプログラムを格納可能な媒体が含まれる。またプログラムには、上記記録媒体に格納されて配布されるものの他、インターネット等のネットワーク回線を通じてダウンロードによって配布される形態のものも含まれる。さらに記録した機器には、上記プログラムがソフトウェアやファームウェア等の形態で実行可能な状態に実装された汎用もしくは専用機器を含む。さらにまたプログラムに含まれる各処理や機能は、コンピュータで実行可能なプログラムソフトウェアにより実行してもよいし、各部の処理を所定のゲートアレイ(FPGA、ASIC)等のハードウェア、又はプログラム・ソフトウェアとハードウェアの一部の要素を実現する部分的ハードウェア・モジュールとが混在する形式で実現してもよい。

【発明の効果】

【0046】

本発明のデジタルフィルタリング方法、デジタルフィルタ装置、デジタルフィルタプログラム及びコンピュータで読み取り可能な記録媒体並びに記録した機器によれば、整数の加算と比較演算にて特徴量の抽出が可能となり、信号波形処理における演算処理量を大幅に少なくして高速化を図ることができる。特にCPUやメモリ量が制限され、限られた演算処理能力の携帯型電子機器にも実装可能な低負荷の波形信号処理を実現できる。

【発明を実施するための最良の形態】

【0047】

以下、本発明の実施の形態を図面に基づいて説明する。ただし、以下に示す実施の形態は、本発明の技術思想を具体化するためのデジタルフィルタリング方法、デジタルフィルタ装置、デジタルフィルタプログラム及びコンピュータで読み取り可能な記録媒体並びに記録した機器を例示するものであって、本発明はデジタルフィルタリング方法、デジタルフィルタ装置、デジタルフィルタプログラム及びコンピュータで読み取り可能な記録媒体並びに記録した機器を以下のものに特定しない。また、本明細書は特許請求の範囲に示さ

10

20

30

40

50

れる部材を、実施の形態の部材に特定するものでは決してない。特に実施の形態に記載されている構成部品の寸法、材質、形状、その相対的配置等は特に特定の記載がない限りは、本発明の範囲をそれのみに限定する趣旨ではなく、単なる説明例にすぎない。なお、各図面が示す部材の大きさや位置関係等は、説明を明確にするため誇張していることがある。さらに以下の説明において、同一の名称、符号については同一もしくは同質の部材を示しており、詳細説明を適宜省略する。さらに、本発明を構成する各要素は、複数の要素を同一の部材で構成して一の部材で複数の要素を兼用する態様としてもよいし、逆に一の部材の機能を複数の部材で分担して実現することもできる。

【0048】

本明細書においてデジタルフィルタ装置とこれに接続される操作、制御、入出力、表示、その他の処理等のためのコンピュータ、プリンタ、外部記憶装置その他の周辺機器との接続は、例えばIEEE1394、RS-232x、RS-422、RS-423、RS-485、USB等のシリアル接続、パラレル接続、あるいは10BASE-T、100BASE-TX、1000BASE-T等のネットワークを介して電氣的に接続して通信を行う。接続は有線を使った物理的な接続に限られず、IEEE802.1x、OFDM方式等の無線LANやBluetooth等の電波、赤外線、光通信等を利用した無線接続等でもよい。さらに信号波形処理の対象となる信号波形データや処理後のデータの保存や設定の保存等を行うための記録媒体には、メモリカードや磁気ディスク、光ディスク、光磁気ディスク、半導体メモリ等が利用できる。

【0049】

本発明のデジタルフィルタリング手法は、現在フーリエ変換などが利用されている画像処理システムのマッチングや音声のリアルタイム分析・合成、熱伝導、応用力学、電磁場の解法といった工学分野で、フーリエ変換に代わるデジタルフィルタリングとして使用できる。特にフーリエ変換等のデジタル信号処理で必須となる浮動小数点演算のような複雑、高度な演算を行うことなく、四則演算のみで処理できるので、高度な処理能力を要せず比較的安価なLSIなどで実現でき、しかも低負荷であるため高速な処理が可能でリアルタイム処理にも適している。

【0050】

デジタルフィルタリングを行うデジタルフィルタ装置は、入力された入力信号の信号波形を量子化する量子化手段と、量子化データを格納可能なメモリ手段と、任意の量子化データについて対象点に隣接する所定の範囲を加算区間として予め設定し、あるいは/なおかつ加算区間に含まれるデータ数を予め設定する設定手段と、メモリ手段に格納された量子化データの各点毎に、加算区間に含まれるすべての量子化データのデータ値を加算して加算値を求め、加算値をデータ数で除算して除算値を求め、さらに必要に応じて該点の量子化データのデータ値と除算値との差を減算により減算値として演算し、これら除算値及び減算値を出力する演算手段とを備える。量子化手段は、例えば入力信号の信号波形の振幅を求め、振幅を量子化する振幅算出部で構成できる。またメモリ手段は、量子化データその他、加算値、減算値、除算値の少なくともいずれかを保持可能なメモリが利用できる。さらに演算部は、振幅算出部で量子化された各点のデータにつき、隣接する所定の範囲のデータの振幅値を加算し、これを加算したデータの個数で除算して該データを中心とする除算値を求める除算値演算部や、各データにつき除算値演算部で演算された除算値と、振幅演算部で演算された振幅値との差を減算によって求める減算手段とで構成できる。このようにして得られた除算値や減算値を用いて、各種のデジタルフィルタリング処理を行う。

【0051】

デジタルフィルタリングの手順の一例を、図25のフローチャートに基づいて説明する。まずステップS1で入力信号の信号波形を入力する。次にステップS2でこの信号波形を量子化する。必要に応じて標本化や符号化の処理を行うことも可能であることは言うまでもない。量子化された離散的な各量子化データは、ステップS3でメモリに格納される。

10

20

30

40

50

【 0 0 5 2 】

次に、ステップ S 4 でフィルタリング処理点の決定を行う。ここでは、フィルタリング処理は量子化データの各点すべてについて順次行われる。例えばインクリメント処理などによって、フィルタリング処理を行う対象点を順次切り替えながら、各点の加算値、除算値、減算値などを演算していく。ステップ S 4 でフィルタリング処理点が決まると、設定されたフィルタリング処理条件に基づいてフィルタリング処理が行われる。フィルタリング処理を行うための条件の設定は、信号波形の入力前に予め設定手段から設定しておくことが好ましい。設定すべき項目としては、量子化データのデータ値を加算する範囲、すなわちフィルタリング処理点に隣接する一定範囲である加算区間、及び/又は加算区間に含まれるデータ数がある。図 25 の例では、ステップ S 5 で加算区間の設定が呼び出され、これに基づいてステップ S 6 で加算区間における全量子化データのデータ値の加算値 S を算出し、一方ステップ S 7 で加算区間に含まれるデータ数 N を算出している。なおデータ数は、量子化データのサンプリング周期（あるいは標本化の周波数）が既知の場合、加算区間によって一義的に決まるため、毎回演算する必要はなく、単に N を既定値として呼び出すだけでよい。逆にデータ数 N とサンプリング周期が決まっている場合は、加算区間が一義的に決まるため、この場合は加算区間を単に規定値として呼び出すだけで済む。いずれの場合も、予め設定手段で設定されたフィルタリング処理条件に基づいて、演算に必要な情報が取得される。

10

【 0 0 5 3 】

このようにして加算値 S と加算区間のデータ数 N が演算されると、ステップ S 8 で加算値 S をデータ数 N で除算して除算値 A を算出できる。さらに必要に応じてステップ S 9 でフィルタリング処理点のデータ値から除算値 A を減算（あるいはその逆）し、減算値 B を算出する。そしてステップ S 10 で除算値 A 及び/又は減算値 B をフィルタリング処理の結果として出力する。このようにして一のフィルタリング処理点におけるフィルタリング処理値が演算されると、フィルタリング処理値をメモリに一時的に格納するとともに、ステップ S 11 に戻り、フィルタリング処理点を変更して、上記の演算を繰り返す。最終的にすべての量子化データについて得られたフィルタリング処理値を用いて、所望のデジタルフィルタリング処理を行う。このデジタルフィルタリング処理は、本発明が適用される用途に応じて適宜選択される。

20

【 0 0 5 4 】

図 1 に、本発明の一実施の形態として、デジタルフィルタ装置を音声認識装置に適用した例のブロック図を示す。図 1 (a) に示す音声認識装置 100 は、マイクロフォン 10、雑音除去装置 12、特徴量抽出部 14、類似度算出部 16、標準モデル辞書 20、判定処理部 18 を備える。

30

【 0 0 5 5 】

なお図 1 (a) では、発声者が発声した音声をマイクロホンで入力して、A / D (アナログ / デジタル) 変換器でデジタルの電気信号 (デジタル音声信号) に変換しているが、直接音声データを入力する他、ネットワーク接続された外部機器から音声データを入力する方式とすることもできる。

【 0 0 5 6 】

マイクロフォン 10 から取り込まれた音声は、雑音除去装置 12 に入力され、ここで 10 m s 程度の周期でフレーム分析され、周囲環境の雑音や、マイクや伝送路が有する伝送特性雑音を除去される。その後、特徴量抽出部 14 により音声認識で使用される特徴量が抽出される。そして、あらかじめ標準モデル辞書 20 に登録されている複数の単語の特徴量との類似度が類似度算出部 16 で算出される。その類似度がある閾値よりも大きいかなかを判定処理部 18 で判定し、大きければその単語を認識結果として出力する。これらの特徴量抽出部 14、類似度算出部 16、判定処理部 18 等の各種演算処理部材は、マイクロプロセッサ (M P U) や C P U、L S I、F P G A や A S I C 等のゲートアレイといった論理回路や中央演算処理装置等のハードウェアやソフトウェア、あるいはこれらの混在により実現できる。また必ずしも各構成要素が図 1 に示した構成と同一でなくてもよく、

40

50

その機能が実質的に同一であるもの、あるいは一つの要素が図1に示す構成における複数の要素の機能を備えるものは、本発明に含まれる。

【0057】

特徴量としては、一般にはケプストラム係数が利用され、対数的な変換処理により対数スペクトルを求め、逆フーリエ変換や逆コサイン変換をすることにより算出して抽出することが行われていた。ただ、この方法では周波数スペクトル等への演算が必要となり処理量の負担が大きいので、本実施の形態では各母音（5母音）の特徴を振幅波形から抽出する方式とする。振幅波形から特徴を抽出することで周波数スペクトル等への変換のための演算を省くことが可能であり、演算回数も比較的少ない計算量で済む。得られた特徴量を用いて離散ポロノイ図により領域を分割し、異なるカテゴリの境界座標を算出して最小2乗法による識別境界線を決定する。

10

【0058】

標準モデルとしては、複数の認識対象語彙毎の特徴量の時系列を確率的な遷移として表現する隠れマルコフモデル（HMM）と呼ばれる方法がある。HMMとは、あらかじめ個人差による音韻や単語の特徴量の時系列をHMMモデルに学習させておき、入力音声モデルに確率値としてどのくらい近いかを捉えて認識する方法である。また、標準モデルとしては、複数の認識対象語彙毎の特徴量の時系列の中の代表的な特徴量の時系列をモデルとしても良いし、さらに特徴量の時系列を時間的あるいは周波数的に正規化（伸縮）することで得られる特徴量の正規化時系列を用いてもよい。例えば、時間軸上で任意の長さに正規化する方法としDPマッチング（動的計画法）があり、あらかじめ決定した対応付けの規則に従って、時間的特徴量の時系列を正規化することが可能である。

20

【0059】

本実施形態では、このようにいずれの場合の標準モデルを使用することができる。ただし、いずれの標準モデルを作成する場合でも、標準モデルを作成するための複数の音声データをあらかじめ用意しておき、入力音声の振幅に対して同様の処理を行い凹凸波形に変換して登録しておく必要がある。

（特徴量抽出部）

【0060】

特徴量抽出部14は、図1(b)に示すように、振幅算出部22と、除算値演算部24と、比較部26と、変換部28とを備える。振幅算出部22は、入力波形の振幅波形に基づいて量子化を行う。量子化されたデータは除算値演算部24で、各点の除算値を求める。ここでは、図2(a)に示すようにサンプル点を中心とするN個のデータについて除算値を求める。すなわち、サンプル点kを中心として、その前のn(=N/2)個のデータ及びその後のn個のデータについて、振幅値を加算してNで除算する。具体的には、N=2×nとし、x_kの前後n点のサンプル値より得られた除算値α_kを以下の数5にて演算する。

30

【0061】

【数5】

$$\alpha_k = \frac{\left(\sum_{i=k+1}^{n+k} x_i \right) + \left(\sum_{i=k-n}^{k-1} x_i \right)}{N}$$

40

【0062】

ここで、添え字のkは現在参照しているサンプル点である。x_kはk点における振幅値を表している。またNは凹凸波形を算出するために用いる閾値を計算するための幅を表している。ここでNを2のべき乗とすることで、除算の際に2進数であればビット・シフト演算が利用でき、除算値演算部24での演算を簡単にすることができるので好ましい。

【0063】

50

このようにして除算値演算部 2 4 で各点につき得られた除算値を、比較部 2 6 で振幅値と比較する。具体的には、各点の振幅値 x_k とその除算値 α_k とを比較し、以下の数 6 のように比較結果を出力する。

【 0 0 6 4 】

【数 6】

$$B_k = \begin{cases} a (\alpha_k < x_k) \\ b (\alpha_k > x_k) \end{cases}$$

10

【 0 0 6 5 】

このように、比較部 2 6 は振幅波形の各点につき除算値を演算し、 k 点のサンプル値 (振幅値) が各除算値より大か小により得られる信号を出力する。比較結果として x_k が除算値以上であれば a 、除算値以下であれば b を比較部 2 6 は出力する。この結果を、変換部 2 8 で凹凸状の波形として出力する。例えば $a = 1$ 、 $b = 0$ とすると、振幅波形は谷か山 (0 か 1) の凹凸波形で表現できる。この凹凸波形は、除算値を演算した N 個の領域において、 x_k が除算値以上すなわち波形が凸形になっているのか、あるいは x_k が除算値以下すなわち凹形になっているかを表現している。よって、 N の値を変化させれば元の振幅波形の凹凸を粗く、あるいは細かく表現することができ、 N をパラメータとして変化させることにより複数個の特徴量を抽出できる。このように振幅波形の特徴を凹凸のみで簡素化して表現することにより、必要な特徴量を抽出でき音声認識モデルに適用することができる。特に、凹凸のみで表現される特徴量は、あまり高度でない情報処理機能においても音声波形から子音部と母音部の切り出し (Segmentation) が実現できる。また認識された音声に基づいて音声出力に処理を加える場合も、実用的な時間間隔で音声出力の大きさを変化させることが可能となる。

20

【 0 0 6 6 】

さらに、除算値の演算において、振幅波形の加算値を求める際に、近接する位置での算出値を利用して演算を簡素化することができる。すなわち、あるデータにつき除算値を求めるために $N (= 2n)$ 個のデータの振幅値を加算した加算値を保持しておく。このとき、除算値 α_k は以下の数 7 のように演算できる。

30

【 0 0 6 7 】

【数 7】

$$\alpha_k = \frac{\sum_{i=k-n}^{k-1} x_i + \sum_{i=k+1}^{n+k} x_i}{N}$$

$$= \frac{\sum_{i=k-n}^{k-2} x_i + x_{k-1} + \sum_{i=k+1}^{n+k-1} x_i + x_{n+k}}{N}$$

40

【 0 0 6 8 】

ここで、除算値 α_k は以下の数 8 のように変形できる。

【 0 0 6 9 】

【数 8】

$$\alpha_k = \frac{\sum_{i=k-n}^{k-2} x_i + \sum_{i=k+1}^{n+k-1} x_i}{N} + \frac{x_{k-1} + x_{n+k}}{N}$$

【0070】

一方、除算値 α_k は上記数 5 で表記されるので、以下の数 9 のようにも表記できる。

【0071】

10

【数 9】

$$\alpha_k = \frac{\sum_{i=k-n}^{k-1} x_i + \sum_{i=k+1}^{n+k} x_i}{N}$$

【0072】

したがって、 k を $k-1$ に置き換えると、 $k-1$ すなわち k より 1 つ前のデータに関する除算値 α_{k-1} は以下の数 10 のように表記できる。

20

【0073】

【数 10】

$$\alpha_{k-1} = \frac{x_{k-1-n} + \sum_{i=k-n}^{k-2} x_i + \sum_{i=k}^{n+k-1} x_i}{N}$$

【0074】

さらに、 α_{k-1} を変形すると、次式数 11 のようになる。

30

【0075】

【数 11】

$$\alpha_{k-1} - \frac{x_{k-1-n}}{N} - \frac{x_k}{N} = \frac{\sum_{i=k-n}^{k-2} x_i + \sum_{i=k+1}^{n+k-1} x_i}{N}$$

【0076】

上記数 11 で α_{k-1} に代わって除算値 α_k を求めると、以下の数 12 のようになる。

40

【0077】

【数 12】

$$\alpha_k = \frac{\sum_{i=k-n}^{k-2} x_i + \sum_{i=k+1}^{n+k-1} x_i}{N} + \frac{x_{k-1} + x_{n+k}}{N}$$

【0078】

さらに、 α_{k-1} を再記すると、数 13 のようになる。

【0079】

50

【数 1 3】

$$\alpha_{k-1} - \frac{x_{k-1-n}}{N} - \frac{x_k}{N} = \frac{\sum_{i=k-n}^{k-2} x_i + \sum_{i=k+1}^{n+k-1} x_i}{N}$$

【0080】

上記数 1 2、数 1 3 を整理すると、以下の数 1 4 のようになる。

【0081】

【数 1 4】

$$\begin{aligned} \alpha_k &= \alpha_{k-1} + \frac{x_{k-1} + x_{n+k}}{N} - \frac{x_k + x_{k-1-n}}{N} \\ &= \alpha_{k-1} + \frac{x_{k-1} + x_{n+k} - x_k - x_{k-1-n}}{N} \end{aligned}$$

【0082】

数 1 4 から、除算値 α_k はその前段で演算した除算値 α_{k-1} を用いて容易に演算することが可能となる。すなわち、 α_{k-1} が求まると、数 1 4 の 2 項目である以下の数 1 5

【0083】

【数 1 5】

$$\alpha_k = \alpha_{k-1} + (x_{k-1} + x_{n+k} - x_k - x_{k-1-n}) / N$$

に α_{k-1} に加算することで、除算値 α_k を求めることができる。これにより、逐次的に α_{k+1} 、 α_{k+2} 、 \dots 、 α_{k+n} を求めることができる。上記の演算を図で表すと、図 2 (b) のようになる。この図に示すように、 α_k と α_{k-1} を算出する際のデータには共通領域があることが分かる。したがって、これらの共通領域のデータに対しては、一度の計算で演算は終了し、この演算結果をメモリなどの記憶手段に格納しておけば、次の演算に利用することができるので、全体の計算時間が短縮される。以上のように、演算を簡素化し、さらに演算量も減らすことで演算処理を極めて低負荷にすることができ、非常に簡単な演算によってすべての除算値 α_k を求めるアルゴリズムが得られる。これによって、音声波形から音声認識のための特徴抽出を高速かつ簡単に求めることが可能となり、実用性が極めて高い。

【実施例 1】

【0084】

次に、実施例 1 としてコンピュータ・シミュレーションにより抽出した特徴量を用いた母音の認識実験の結果を図 3 ~ 図 10 に示す。

(音声データ)

【0085】

実施例 1 では、音声データから特徴量を抽出する特徴量抽出部 1 4 として、音声の特徴を生かしアナログ処理とデジタル処理部を組み合わせで作られた専用の集積回路 (IC) を使用した。音声の特徴とは、音声を波形で表したとき一般に正領域と負領域において非対称であること、声帯から送り出される圧力はパルス信号の発生、減衰に基づいていることである。これらの点を考慮してマイクロフォン 10 から得られる音声信号を電圧値で測定すると、同時に正負の電力値の最大を一定時間保持させながら、次の正負の電圧値を検出するまでの時間を検出し、ピッチを検出している。この IC を用いることで、音声波形

10

20

30

40

50

とピッチを検出することが可能である。図3に、ICにより取得した音声波形とピッチ情報を示す。

【0086】

また分類は5母音から得られる2母音の組み合わせを用いて、投票形式で行なわれる。そして音声の振幅波形から認識に有効な特徴の抽出と解析を行なう。実施例1では、成人男性1名から67音素の17セットの音声データを取得する。サンプリング周波数は81.92kHzである。さらに、音声は自然発話と意識発話のデータを取得し、自然発話は一日の時間を問わず発生した音声であり、意識発話は夜の静かな時間帯に音素をはっきりと発声した音声である。音声の長さとして意識発話のピッチ数は自然発話の約1.5倍の長さで取得されている。

(母音波形抽出と前処理)

【0087】

母音識別を行なうためにICにより得られているピッチを参照して音素データから母音の定常とする区間を抽出する。そこで、全ピッチ数の3分の2の位置にあたるピッチを中心に前後の1周期分を抽出し、全部で3周期分の信号を定常な母音波形として特徴を抽出するために用いる。そして、抽出した3周期の母音波形から凹凸波形を生成する。ここで、元の振幅波形である定常とした3周期の母音波形を図4に、 $N = 256$ として図4から抽出した1周期分の凹凸波形を図5に、 $N = 64$ として図4から抽出した1周期分の凹凸波形を図6に、それぞれ示す。これらの図においては、上記数6の $a = 0.8$ 、 $b = 0.2$ としている。図5と図6を比較すると、 N の小さい図6の方が細かな凹凸波形となっており、振幅波形の山谷を細かく抽出していること、及び N の大きい図5は、振幅波形の山谷を粗く抽出していることが判る。実施例1では、このように N を256と64の2つに設定して得られた凹凸波形を用いて認識実験を行う。なお、 N の値として凹凸波形を生成する幅である256と64は経験的に決定しているが、これ以外の値とすることもできることはいうまでもない。

【0088】

このようにして得られた振幅波形と凹凸波形を用いて時間軸上で母音の特徴量を抽出する。特徴量は主に1周期分の波形から抽出している。その1周期分の波形は3周期の波形の始点とした点から類似度(ユークリッド距離)を計算し、距離の近い2周期を選択する。そして時間的に早く存在している波形を選択する。これは中心より後半での母音波形の抽出を行っているために、後半に位置する母音波形よりも前半に位置する母音波形のほうが、より母音の特徴を保った波形であると考えられるためである。このように選択された1周期分の波形部分と3周期分の凹凸波形から特徴を抽出する。次に、母音を識別するための特徴量について述べる。抽出する特徴量の数は全部で5個である。母音を識別するための提案システムは、5母音の内の2母音の各組み合わせから選ばれた母音の投票数により識別したい母音であると決定している。2母音の組み合わせから選択すべき母音を識別するために、各2母音を識別する特徴は組み合わせごとに異なる。2母音を識別しやすい特徴量を各組み合わせごとに選択することで比較的高い識別率が得られると考えられる。抽出する5つの特徴量は、以下の通りである。

【0089】

(1) 256凹凸波形を参照し、母音波形の1周期の始まりから探索して得られる最初の凸部分の幅

【0090】

(2) 256凹凸波形を参照し、母音波形の1周期の始まりから探索して得られる最初の凸部分の幅に存在する振幅波形の面積

【0091】

(3) 256凹凸波形を参照し、母音波形の1周期の始まりから探索して得られる最初の凸部分の幅に存在する振幅波形の分散値

【0092】

(4) 256凹凸波形を参照し、母音波形の1周期の始まりから探索して得られる最初

10

20

30

40

50

の凸部分の幅に存在する振幅波形を 0 ~ 1 に正規化を行い、凸幅で生成した正弦波との類似性

(5) 3 周期分の 6 4 凹凸波形に存在する凸の数である。

【 0 0 9 3 】

ここで特徴量 4 について詳述する。まず特徴量 4 を抽出するために正弦波を生成する。その正弦波は、

(1) 凸区間の振幅値を 0 ~ 1 に正規化を行う。その振幅値の最大値の位置の検出を行う。

(2) 始点から最大値位置までに 0 から / 2 までの正弦波を生成する。最大値位置から終点までに / 2 から までの正弦波を生成する。

という処理で行なう。正弦波により得られた波形と振幅波形の例を図 7 と図 8 に示す。図 7 は、凸部分の / e / 音素波形と生成した正弦波であり、図 8 は、凸部分の / o / 音素波形と生成した正弦波である。特徴量 4 はこの二つの波形から逐次 4 点で角度が計算され、その差の合計が特徴量とされる。これらの特徴量の大半は 1 周期の始めに存在する凸部分の特徴量で構成されている。この部分は母音波形 (定常波形) で最も変化があり、また各母音の特徴が現われていると考えられる。この始めに位置する凸部分の検出は 1 周期分の 2 5 6 凹凸波形から検出され、凹幅が最も大きい箇所の次に現われる凸部分である。

(母音認識実験)

【 0 0 9 4 】

次に、抽出した特徴量を用いて離散ポロノイ図と最小 2 乗法を適用し母音の識別を行なう。離散ポロノイ図とは空間中に配置された多数の母点の勢力範囲を表す図である。離散ポロノイ図は母点の存在する空間を離散化し、各画素がどの母点に近いかによって空間を分割したものである。本実施例で用いた離散ポロノイ図は逐次添加法を適用している。これは離散化された空間に母点を一つずつ追加していき、新しいポロノイ領域だけを効率よく作成することで、高速に離散ポロノイ図を作成する方法である。本実施例で用いた離散ポロノイ領域は 5 1 6 0 x 5 1 6 0 の領域であり、各母点との対応は得られている母点の最大値を用いて 0 から最大値までを均等に分割している。

(母音認識アルゴリズム)

【 0 0 9 5 】

識別アルゴリズムは 2 母音の組み合わせ ($_5C_2$) によりなされる。そして、各組み合わせに適した特徴量を用いて母音を選択し、全組み合わせから得られた投票数により識別する母音を決定する。その各組み合わせに適した特徴量 (識別のために使用した) を表 1 に示す。

【 0 0 9 6 】

【表 1】

/a/-/i/	/a/-/u/	/a/-/e/	/a/-/o/	/i/-/u/
1-2	1-2	1-2	1-2	1-5
/i/-/e/	/i/-/o/	/u/-/e/	/u/-/o/	/e/-/o/
1-2	1-2	1-4	1-2	3-4

【 0 0 9 7 】

ここで、識別のための特徴量の組み合わせを示す表 1 において、上段の / a / - / i / 等は各母音の組み合わせを表しており、下段の 1 - 2 等は使用した特徴量の番号 (1) 及び (2) を示している。これらの 2 母音の識別を行なう特徴の選択は全データから抽出した特徴量の分布を調査し、目視による特徴量の分布を確認した後に経験的に決定している。さらに母音の識別を行なうための特徴量の数が 2 個であるのは、単一の特徴量を用いて

識別を行なえば簡単な閾値処理になるため処理時間は少なく済むが、同じ母音であっても分布の広がりのために境界付近のデータに関して単一の特徴で識別を行なうと誤認識を起こしやすいと考えられるからである。また2個の特徴量を用いて目視によるデータ分布の確認を行ないやすいことや、直線的な閾値処理ではなく非線形的な識別が可能であり演算回数をできるだけ少なくすることが目的である。

【0098】

本手法で母音識別は離散ポロノイ図による2クラスの母音の組み合わせの境界線を求めるものである。得られている特徴量をマッピングしポロノイ領域に分ける。そして、同カテゴリによる領域の統合を図り、2クラス間の境界部分の座標を得る。そして、その座標により構成される境界線を最小2乗法により境界線の関数を算出する。その境界線の関数を用いて母音の識別は2クラス間の投票形式により行なわれる。選択された母音の投票数が単独で最も多く存在した場合のみ、その母音に識別されたとする。一番多い投票数が等しく複数の母音に見られる場合は破棄する。ここで、/i/と/u/以外の組み合わせに対する識別は抽出した特徴量をそのまま用いている。しかし、/i/と/u/の識別に対しては特徴量1の常用対数をとることで数値を変換している。これは特徴量5に対して特徴量1の値のスケールが大きいためであり、最小2乗法を適用して境界線の関数を求めやすくするためである。また比較実験としてマハラノビス距離を用いた認識実験を行なう。これは各組み合わせから母音の選択はマハラノビス距離に基づいて行なうものである。マハラノビス距離は各グループの中心から分散を考慮した距離を示す。本実施例で抽出された母音の特徴は分布に偏りのある特徴量である。そのためマハラノビス距離による識別が有効であると考えられる。意識発話による各母音データ数は188個であり、自然発話による各母音データは178個である。そして各識別に適した特徴量を用いてポロノイ図とマハラノビスにより得られた自然発話による識別結果と意識発話による識別結果を表2と表3に示す。表2は、離散ポロノイ図を用いた自然発話と意識発話の単独第一候補による認識率を、表3はマハラノビス距離による自然発話と意識発話の単独第一候補による認識率を、それぞれ示している。

【0099】

【表2】

表2: 離散ポロノイ図を用いた自然発話と意識発話の単独第一候補による認識率

自然発話						意識発話					
/a/	/i/	/u/	/e/	/o/	平均	/a/	/i/	/u/	/e/	/o/	平均
99.4	91.5	94.9	91.5	90.4	93.5	98.9	97.3	97.8	98.9	98.4	98.2

【0100】

【表3】

表3: マハラノビス距離による自然発話と意識発話の単独第一候補による認識率

自然発話						意識発話					
/a/	/i/	/u/	/e/	/o/	平均	/a/	/i/	/u/	/e/	/o/	平均
95.2	91.4	82.4	98.4	88.2	91.1	98.4	97.3	90.4	100.0	90.9	95.4

【0101】

以上のように、離散ポロノイ図を適用し境界線を求めることで、マハラノビス距離を用いて行なった認識実験より認識率が向上している。これは離散ポロノイ図により明確な境界線の導出が可能であることで認識率が向上したと考えられ良好な結果を得ることができた。また表3に示すように、マハラノビス距離を用いた認識結果から自然発話と意識発話の両方とも/u/と/e/の識別結果が他の母音に比べて低い。これは/i/と/u/、

/ e / と / o / の抽出した特長が境界を超えて存在するためである。この境界付近に存在する特徴を持つ母音は、マハラノビス距離により分散を考慮され境界を得ているが、正しい識別が行なえていないことが考えられる。この理由として、現在用いているデータから正確な各母音に関する分散が計算できないこと、あるいは音圧データを用いることで類似している母音の特徴の分布状態に偏りが存在すること等が考えられる。

【 0 1 0 2 】

ここで、意識発話に用いた特徴の分布を図 9 と図 1 0 に示す。図 9 は / i / と / u / に関する特徴分布を示し、図 1 0 は / e / と / o / に関する特徴分布を示している。この結果から、自然発話より意識発話の結果が良好であることがわかる。これは意識して声を発声しているために、母音部分の 1 周期波形（ピッチ）が多く現われ、安定した母音波形を抽出できていると考えられる。このように意識発話を行なうことで比較的良好な識別結果が得られる。さらに簡易な演算のみで特徴量の抽出やマハラノビス距離による識別を行なっているため小規模なハードウェア構成で実現可能であると考えられる。しかし、認識精度において離散ボロノイ図を用いた認識精度には及ばない。そのために離散ボロノイ空間の設定を適切に行なうことができれば、離散ボロノイ図を用いた小規模なハードウェアが構築できると考えられる。このように、本実施例 1 を適用することで比較的良好な認識精度が得られると考えられる。また、境界線を求めるのではなく勢力図を保持することで特徴量をその勢力図に照らし合わせ、母音の決定も可能である。

10

【 0 1 0 3 】

以上のように、本実施例によってモバイル機器のための音声認識システムが実現可能となる。特に小型のハードウェアのために比較的簡単な演算による母音の振幅波形からの特徴量抽出と 5 母音の識別を行ない、その有効性を検証した。

20

【 0 1 0 4 】

さらに、除算値等による特徴量のスケールの正規化を行なうことで、より小スペースで離散ボロノイ図が適用可能となり、演算時間を少なくすることができる。また、より明確な識別が可能である特徴の抽出にも適用可能である。例えば、識別境界線を 2 分類を行なうのに適した手法である SVM を用いることができる。また、母音波形の 1 周期ごとにも違いがあるために定常であるとする母音波形の特定や母音に対して音声の始まりから終わりまでのピッチ幅の変動等を調査し、標準モデルとして登録しておくことで識別に有効な特徴の抽出や演算時間がさらに改善される。

30

【実施例 2】

【 0 1 0 5 】

（補聴機能）

さらに、本発明は音声信号処理の前処理として広い応用範囲を備えており、入力された音声信号の認識のみならず、音声の出力に際しての処理、例えば音声をより聞き取りやすい形に加工することもできる。この機能を応用すれば、音声をより聞き取りやすく加工する補聴機能として、補聴器等に利用できる。

【 0 1 0 6 】

従来の補聴器では単純に検知した音声信号をすべて増幅させて音量を大きくするものであった。しかしながら、補聴器から出力される音量を増加すると、目的とする音以外の雑音も大きく聞こえるため、耳にガンガンと音が入り頭痛障害等を引き起こしたり不快感を伴うことがあった。また、小さなスピーカでは音が割れて音質が劣化する。さらに音量を大きくすると消費電力が大きくなり、長時間の使用が困難になる。さらにまたイヤースピーカの口径が大きくなり、重量も増加し、補聴器全体の形状も大きくなるといった欠点があった。

40

【 0 1 0 7 】

これに対して、本発明の音声信号の特徴量抽出機能を応用し、音声の聞き取り難い成分を聞き取りやすく加工することができる。すなわち、音声は子音と母音に分かれるが、母音は比較的音量が大きく聞き取りやすいのに対し、子音は時間的に短く音量も小さくなりがちで聞き取り難い部分となる。そのため、子音の部分を強調すれば、聞き取りやすくす

50

ることができる。検出された音声信号から、子音と母音とを区別するには、上述した音声認識技術を適用してもよい。この場合は、具体的な発生音の識別までは不要で、母音と子音の区別ができれば十分であるため、より精度を向上させることができ、しかも処理をさらに簡素化できる。これによって子音と母音とを区別し、さらに再生前に音声信号を加工する。

【0108】

ここでは、再生中に音量をほぼリアルタイムに調整するアルゴリズムとして、聞き取り難く時間的にも短い子音部がきたときに音量を大きくし、それ以外の部分は通常の音量とする。子音部の直後に表れる母音部の音量は一般に大きいので、子音とその直後の母音により音声は認識される。この際に母音部は1周期から数周期の短時間波形で音声認識は完了する。聴覚と脳機能による短時間の音声認識が完了すると、それ以後は母音の継続であり、その期間において音声の出力信号を非常に小さくする。そして再び子音が入力されると、同様に音声出力を大きくする。この操作により、音声出力を聴感上は大きくしたことになるが、常に大音量の音声を聞いている訳ではないので、不快感は極めて少なくなる。また必要に応じて、母音の音量を絞るように調整してもよい。

10

【0109】

上記の音声信号処理のアルゴリズムにおいて重要な点は、音声の出力を通常とする、あるいは低下させた状態から、子音が入力される時点で音声出力を増加させる処理である。従来のシステムでは音声認識自体の処理量が多く、さらにこのような音量変化の処理を加えると、音声の子音の検出に要する時間と処理量が増大し、実用化は容易でない。これに対して、本発明のアルゴリズムでは整数演算のみで子音部と母音部を検出することが可能であるので、非常に短時間の簡単な信号処理により子音の検出が可能となる。すなわち、高速化が可能な演算処理量に抑えることができるので、補聴器や携帯電話のような携帯型電気機器等の小型のシステムへの組み込みや実装が容易となる。特に専用の処理を行うICを用意し、上記のデジタル信号処理を行わせることにより、携帯機器が保有するあまり高度でない情報処理機能においても音声波形から子音部と母音部の切り出し (Segmentation) が実現され、実用的な処理速度で音声出力の大きさを変化させることが可能となる。このように、本実施例2によれば長い処理時間と大規模集積回路を必要とする不動小数点演算を用いることなく、整数型の演算のみで高速かつシンプルに実行できるので、実用的な価値は高い。また、音声信号の処理には通常DSP (Digital Signal Processor) 等を用いるが、浮動小数点演算を行うと長時間を要し、しかも集積回路の規模が増加するので、携帯電話などに搭載することは困難である。これに対し以上のアルゴリズムは整数型演算とビット・シフト演算のみで高速に処理することができ、これを回路として組み込むには小規模の集積回路で実現可能である。携帯機器や補聴器などの小型装置に搭載することができることは、実用化に際して大きな特長となる。

20

30

【0110】

さらに、補聴器の使用者に応じた周波数特性に調整することもできる。補聴器のスピーカから出力される音声の周波数特性を、使用者の耳の聴覚特性に合致させるよう調整することにより、聞こえ難い周波数に対して補正をかけることが可能となり、使用者に応じて聞き取りやすい適切な音声で聴くことができるようになる。

40

【0111】

このように、本発明によれば音声の聞き取りを容易にできる高性能な携帯型補聴器が実現できる。特に小型の補聴器は使用できる電池の大きさも制限され、小消費電力とする必要があるため、本発明のように演算処理量が少なく必要な消費電力も少なくて済むデジタルフィルタリング方法は、理想的である。

【0112】

また、本発明は補聴器以外の携帯機器への実装も容易に行える。すなわち、携帯電話やPHS等、音声を再生するスピーカを備える機器において、上記の処理を適用することで通話内容をより聞き取りやすくすることができる。特に携帯電話は小型、軽量化や連続駆動時間の長時間化の要求が強く、それでいて高性能化が求められているため、本発明のよ

50

うに低消費電力で音声をクリアに再生できるという優れた機能は非常に実用性がある。また、上記と同様に携帯電話の使用者に応じた周波数特性に再生設定を調整しておくことで、使用者に応じた最適な状態で音声を聞くことのできる携帯電話が実現可能である。これによって、補聴器を利用しない者であっても携帯電話の音声を聞き取りやすくして便利に使用できる。特に、通常の携帯電話ではスピーカから出力される音声等は出力も小さく、健全な聴覚特性の使用者向けに作成されているので、聴覚特性が低下した高齢者には使用し難いことがあったが、本発明を利用して音声を聞き取りやすくした高齢者やシニア向けの携帯電話を実現することが可能となる。

(音声信号のデジタルフィルタリング)

【0113】

ここで本発明の実施の形態に係るデジタルフィルタリングを適用して音声を聞き取り易くする手法を、聴覚補助装置に適用する一例を説明する。図11は、聴覚補助装置の一例として補聴器のブロック図を示している。図11(a)に示す聴覚補助装置200は、音声入力部10A、A/D(アナログ/デジタル)変換器12A、凹凸波形変換部14A、声成分抽出部16A、強調処理部18A、音声出力部20Aを備える。音声入力部10Aとしては、マイクロホンや音声入力端子などが利用できる。特に補聴器として利用する場合は、マイクロホンで入力した音声信号をA/D変換器12Aでデジタル音声信号に変換して声成分抽出部16Aに送出する。ただ、直接デジタルの音声データを音声入力部10Aから入力する形態や、ネットワーク接続された外部機器から音声データを入力する方式も適宜採用できる。マイクロホンから取り込まれた音声信号は、アナログフィルタなどの雑音除去装置に入力され、ここで10ms程度の周期でフレーム分析され、周囲環境の雑音や、マイクや伝送路が有する伝送特性雑音を除去される。その後、A/D変換器12Aでデジタル音声信号に変換して凹凸波形変換部14Aにより量子化され、凹凸波形に変換される。さらにこの凹凸波形から声成分抽出部16Aで人の声に対応する声成分を抽出し、これを強調処理部18Aで強調声波形に補正し、音声出力部20Aから出力される。音声出力部20Aはスピーカや音声出力端子などである。凹凸波形変換部14A、声成分抽出部16A、強調処理部18A等の各種演算処理部材は、マイクロプロセッサ(MPU)やCPU、LSI、FPGAやASIC等のゲートアレイといった論理回路や中央演算処理装置等のハードウェアやソフトウェア、あるいはこれらの混在により実現できる。また必ずしも各構成要素が図11に示した構成と同一でなくてもよく、その機能が実質的に同一であるもの、あるいは一つの要素が図11に示す構成における複数の要素の機能を備えるものは、本発明に含まれる。

【0114】

次に、図11に示す補聴器にデジタルフィルタリングを適用して、音声を聞き取り易くする手順の一例を、図12のフローチャートに基づいて説明する。図12に示すように、音声入力部10で音声信号波形を取得した後、声成分抽出部16で人の声成分を抽出し、さらに強調処理部18で凹凸波形を強調して強調声波形を生成して、この強調声波形に基づいて音声出力部20で補正された声を出力する。以下、図13~図14を参照して凹凸波形変換部14及び声成分抽出部16で音声信号波形から声成分を抽出する手順を説明する。図13~図14は、取得した音声信号波形をそれぞれ示している。図13は、音声の最も低い周波数に対応する波形を示しており、その周期は T_1 である。一方、図14は、音声の最も高い周波数に対応する波形を示しており、その周期は T_2 である。人の声に含まれる周波数成分は、男性でも最低の周波数 $f_1(=1/T_1)$ は通常100Hz止まりであり、100Hz以下の周波数成分は殆ど無い。このとき、周期 T_1 は $1/100\text{Hz}=0.01\text{s}=10\text{ms}$ である。そしてこの一周期の正又は負の半サイクルは、その $1/2$ 、すなわち $10\text{ms}/2=5\text{ms}$ となる。したがって、音声信号波形のサンプリングは、 $5\text{ms}(0.005\text{s})$ 程度の幅(窓)でサンプリング周期を設定すれば、音声の最も低い周波数 T_1 にも対応してサンプリングが可能となる。本実施の形態では、周波数が100Hz近傍にある場合の余裕を考慮して、6msに設定する。言い換えると、半波長が6msよりも長い成分は人の声でない成分(すなわちノイズ)であると峻別でき、これによ

10

20

30

40

50

って音声信号から人の声（の低音領域）に相当する成分を抽出することができる。

【0115】

例えば、音楽用CDと同等の音質でサンプリングする場合、 $f = 44.1 \text{ kHz}$ すなわち1秒間に44100サンプルを取得する必要がある。この周期でサンプリングする場合、 0.6 ms の区間（窓）で取得されるサンプル数は、 $44100 \text{ サンプル} \times 0.006 \text{ ms} = 264.6 \text{ サンプル}$ となる。したがって、上記の手法でサンプリングする場合は、 $N = 264.6$ に設定すればよいことになる。ここで、上述したように N を2のべき乗に設定すれば、ビットシフトでの演算が可能となり、浮動小数点演算を使用しない整数型演算で高速且つ低負荷に処理することができる。ここで $N = 2^n$ とすると、 $N = 2^8 = 256$ とすれば、上記とほぼ等しい設定が実現できる。この場合は、 $n = 8$ を採用する。以上から、低音の場合は $n = 8$ を採用することで、低い周波数でのサンプリングすなわち音声信号取得に対応できることが判る。

10

【0116】

同様に、高い周波数 $f_2 (= 1/T_2)$ に対しては、 $n = 1, 2, 3$ のいずれか一を採用できる。これによって、人の声の高い周波数成分よりもさらに高い周波数成分は、人の声でないノイズであると峻別できる。この結果、人の声が含まれない低域と高域をカットして人の声のみを抽出できる。この例では、音声信号の高周波域信号、中周波域信号、低周波域信号に分けて、それぞれ低周波域に対しては $n = 8$ 、中周波域に対しては $n = 5$ 又は6、高周波域に対しては $n = 1, 2, 3$ のいずれかに設定している。高域、中域、低域のそれぞれにおいて、 n の値をいずれに設定するかは、使用される環境やユーザの聴力特性などに応じて設定される。また、高域、中域、低域の3つで分ける他、高域と低域のみの2つを利用する方法や4以上に区別する方法も用途やユーザなどに応じて適宜採用できる。

20

【0117】

以上のように、取得した音声信号に対して、 n の値を高域、中域、低域の3つでそれぞれ変化させて人の声を含む音声信号として高周波域信号、中周波域信号、低周波域信号をそれぞれ抽出することができる。いいかえると、サンプリングの際に n の値を変化させることで、人の声にあたる成分のみを抽出できる。言い換えると、以上の工程では、音声信号波形から低音と高音をカットして人の声の成分を抽出し、これをデジタル処理して凹凸波形としている。このようにして得られた凹凸パターンは人の声成分を含んでいるため、これを明瞭にするための補正を行う。

30

【実施例3】

【0118】

（補聴機能付携帯電話）

また、補聴器機能を携帯電話等の携帯電気機器に組み込むことも可能である。このような補聴機能付携帯電話を、例えば、携帯電話の動作モードを切り替えて補聴器として使用可能とできる。これによって、携帯電話を使用するのと同じ姿勢で、すなわち耳元に電話機をおく状態で補聴器として使用することができるので、使用時の不自然さが少なく、周囲の人間も補聴器を使用しているとは気が付きにくく、使用時の抵抗感を和らげることができる。特に近年、携帯電話が普及し、年齢・男女を問わず携帯電話を使用することが通常の行動パターンとして定着した結果、携帯電話を耳に当てることに対して、周囲の人々は特別な反応を示さなくなり、当たり前に行きとして認識されている。他方、特別な形状をした補聴器を取り出して耳に当てることは通常では不自然さがつきまとい、しゃべり方にも多少の変化が表れ、補聴器の利用者も周囲の人々も特別な心理状態に陥りやすく、自然な会話が阻害される可能性がある。このような状況に鑑み、携帯電話に補聴器を組み込むことで、携帯電話を耳に当てて、補聴器として利用していることは外面的には判別できず、ごく普通の日常風景に埋没させることができ、補聴器の利用者の心理的なプレッシャは軽減される。

40

【0119】

また補聴器を携帯電話と一体とすることで、多くのメリットが生まれる。例えば、携帯

50

電話のスピーカを補聴器のスピーカとして共用することができるので、補聴器専用のイヤースピーカを持ち歩く必要がない。また携帯電話を持ち運ぶ感覚で補聴器を常時携帯できるので、特別に補聴器を持ち歩いているという感覚が少なくなり、補聴器の使用における抵抗感を軽減できる。さらに補聴器を使用しないときには、身体に着けておく必要がないので、使用者の身体的負担が少ない。さらにまた、携帯電話のデジタル処理の集積回路に補聴器機能をもつ集積回路を組み込むことが可能であり、補聴器専用の集積回路を別個に製作する必要がなく、補聴器と携帯電話を個別に持つ場合に比較してコストを安価に抑えることができる。さらに近年は携帯電話用に長期使用可能な充電式電池が開発されており、これら最新のものを使用することができるので、補聴器の電源供給不足の心配が少なくなる。さらにまた、電池に限らず本体も最新の携帯電話をベースに製作できるので、機能的、デザイン的にも最新の携帯電話モデルを採用でき、利便性が高く、コスト的にも安く製作することが可能である。特に、補聴器専用の金型を作製すれば高価になるが、携帯電話と兼用することにより金型代の節約にもなる。

10

【0120】

この際、携帯電話にはマイク等の集音機を、使用者が通話するために使用するものと、周囲の音を集音するためのものとで2個以上を設けることが好ましい。

【0121】

さらに、補聴機能等のように音声信号の再生側で聞き取りやすく処理する他、音声信号の入力側で同様の処理を行うこともできる。例えば携帯電話において、スピーカのみならずマイク側に本発明を適用し、通話の相手側に向けて送信される音声信号に、上記のような子音強調処理を行えば、携帯電話の使用者のみならず通話先の相手においても、会話内容を聞きとりやすくできる。さらにまた、電話のみならず、TVやラジオ、TV電話、TV会議システム等、スピーカやイヤフォンで音声を再生する機器にも同様に適用できる。

20

【実施例4】

【0122】

(外国語の再生)

さらに、英会話等の外国語の発音、聞き取りにおいても、本発明を適用できる。例えば英語の場合は、振幅波形の山谷を強調することで、より聞き取りやすくできる。これにより、同時通訳用のマイクやスピーカ等の通信システムや外国語学習用教材のテープやCD、DVD等の再生装置等に適用して、より外国語の発音を聞き取りやすく処理することができる。

30

【0123】

このように本発明のアルゴリズムを適用すると、音声信号の特徴抽出を高速に行うことができるので、リアルタイムな音声信号処理が可能となる。その結果として、音声認識、高音質化補聴器、多量の音声データを含むCDの高音質化等が実現される。

【実施例5】

【0124】

(デジタルフィルタ回路)

また本発明は、上述のような音声信号の処理に限られず、他のアナログ波形にも適用できる。すなわち、アナログ波形を凹凸波形に変換するという、一種のデジタル化によってDSPなどでの扱いを容易とできる。例えば、本発明の実施の形態に係るデジタルフィルタリング方法は、デジタルフィルタ回路に適用できる。デジタルフィルタ回路は、時間の関数として与えられる連続な量に対する変換を、アナログ素子によってではなく、離散化時点でA/D変換器でデジタル化し、デジタル信号処理回路で数値演算を行う回路である。本発明は、画像などのように2次元の空間座標の関数として与えられる量や、地震波やレーダなどのように空間座標と時間座標の両者の関数として与えられる量に対するデジタルフィルタリングのいずれにも適用できる。特に本実施例5によれば、位相シフトの無いデジタルフィルタ回路が実現できる。

40

【実施例6】

【0125】

50

(スピーカへの適用例)

次に、本発明に係るデジタルフィルタリングを、スピーカに適用した実施例 6 について、図 15 ~ 図 21 に基づき説明する。

【0126】

可聴周波数はおよそ 20 Hz から 20 kHz までであり、1 個のスピーカだけで全帯域に渡って良好に音を出すことはできない。従って、通常は複数個のスピーカを用いることにより可聴帯域をカバーする。これをマルチ・ウェイ方式のスピーカ・システムと呼び、その低音専用スピーカをウーファと呼び、中音専用のスピーカをスコカと呼び、高音専用のスピーカをツイータと呼ぶ。ウーファとスコカとツイータの 3 種類のスピーカを用いるスピーカ・システムを 3 ウエイ方式と呼ぶ。同様に、2 種類のスピーカを用いるスピーカ・システムを 2 ウエイ方式と呼び、4 種類のスピーカを用いるスピーカ・システムを 4 ウエイ方式と呼ぶ。

10

【0127】

説明を簡略化するために、2 ウエイ方式のスピーカ・システムを考察する。2 ウエイ方式のスピーカ・システムにおいては、高音専用のスピーカと低音専用のスピーカを用いる。一方、通常のアンプは可聴周波数の全帯域に渡って、殆どフラットな特性で音信号を出力することが可能である。従って、アンプの出力信号を高音専用スピーカと低音専用スピーカに送り込むには、音信号を高音帯域と低音帯域に分割する必要がある。信号を分割するには、フィルタを適用する。複数個のスピーカを駆動するには、ネットワーク・フィルタと呼ばれる周波数分割器を用いる。通常マルチ・ウェイ方式のスピーカ・ボックス内には、ネットワーク・フィルタが内蔵されている。ネットワーク・フィルタによって高音帯域の信号と低音帯域の信号に分割され、高音信号が高音専用スピーカに送られ、低音信号が低音専用スピーカに送られ、音波の合成は空間で行われる。通常スピーカ・システムのネットワーク・フィルタは、コイル (L) とコンデンサ (キャパシタ; C) 及び抵抗 (R) によって構成される。

20

【0128】

図 15 に、コイル (L) と抵抗 (R) を用いて構成した低音通過フィルタ回路を示す。図 15 に示す回路は L - R 回路と呼ばれ、その周波数特性を図 16 に示す。入力信号の周波数が変化すると、出力信号の振幅と位相は同時に変化する。図 16 (a) は、周波数が大きくなると振幅が減衰する現象を示している。振幅が約 70% に低下する周波数 f_T を遮断周波数 (カット・オフ周波数) と呼ぶ。図 16 (b) には、周波数が高くなると位相が変化する現象を示す。周波数が十分に高くなると、出力信号は入力信号に対して 90 度の位相遅れとなる。遮断周波数 f_T においては 45 度の位相遅れが発生する。複数の周波数成分を含む音楽信号においては、位相遅れが周波数に依存するので、フィルタを通過した波形は位相歪みを発生する。この L - R フィルタを通過することにより発生する位相歪みは、一度発生すると、これを修復することも元の状態に回復することもできないので、このフィルタを非可逆フィルタと呼ぶ。マルチ・ウェイ方式のスピーカ・システムにおいては、低音専用スピーカを駆動する信号として図 16 (a) に示す振幅特性は必要であるが、図 16 (b) に示すように周波数に依存する位相誤差が付随して発生するので、複数の周波数を含む音信号のフィルタとして通常のアナログ・フィルタは不向きである。しかしながら、これに代わるフィルタが存在しないため、やむを得ず使用しているのが現状である。また位相補償回路を付加する必要があり、複雑で高価になるという問題があった。

30

40

【0129】

これに対して、本発明を適用したローパスフィルタリングとハイパスフィルタリングを適用する無位相シフト・フィルタにおいては、振幅の減衰は発生するものの、位相の変化は全く発生しないので、オーディオには理想のフィルタを構成することが可能となる。

【0130】

図 15 に示す L - R 回路を適用すると、信号の高い周波数成分がカットされ、これを低音専用のスピーカに入力する。低音専用のスピーカは口径も大きく、振動慣性が大きく、高い周波数に追従してコーン紙を自由に動かすことが出来ないため、そこから高音を出力

50

すると、聞きづらい音質になる。従って、高い周波数成分を減衰させて低音のみを低音専用スピーカに入力する。一方、図17に示すC-R回路を用いて高音通過フィルタを構成し、その出力を高音専用のスピーカに入力する。C-Rフィルタの周波数特性を図18に示す。図18(a)に示す信号の振幅特性は、高い周波数のみがC-Rフィルタを通過することを示す。図18(b)に示す位相特性は、周波数が十分に低い帯域では90度付近の位相進みであるが、周波数が高くなると、位相進みが0度に近づくことを示している。入力信号の低い周波数成分が減衰しているため、このフィルタの出力を高音専用スピーカに入力してもスピーカが破壊されることはない。

【0131】

図15に示すL-R回路はアナログ積分回路とも呼ばれ、図17に示すC-R回路はアナログ微分回路とも呼ばれる。アナログの微分・積分回路においては、信号の振幅が周波数に依存して減衰する特性は得られるが、同時に位相も周波数に依存して変化する欠点がある。また信号の分割にはチャンネル・ディバイダを用いる方法もあるが、この方法でも依然として位相の周波数依存性は発生する。

【0132】

これに対して、デジタルの微分・積分を適用するフィルタにおいては、信号の振幅は周波数に依存して減衰するが、位相が周波数に依存して変化することは皆無であるため、オーディオ用のフィルタとして理想的な特性を発揮することが可能になる。

【0133】

2ウェイ方式のスピーカ・システムのネットワーク・フィルタを図19に示す。この図において、C-R回路の出力を高音専用のスピーカに入力し、L-R回路の出力を低音専用のスピーカに入力する。これにより低音と高音とに分割されるが、総合的な周波数特性として信号の振幅は殆どフラット(平坦)に調整することができる。しかしながら、位相の遅れと進みにより発生する位相歪みは修復されない。従って、通常のスピーカ・システムに用いられるネットワーク・フィルタにおいては位相歪みが発生し、これが音質劣化の原因の1つになっている。すなわち、アナログの微分・積分回路を適用するネットワーク・フィルタは非可逆性であるため、これを用いる限り過渡現象と位相歪みによる音質劣化を回避することはできない。

【0134】

これに対して、本発明に係る無位相シフト・フィルタは可逆性フィルタであり、過渡現象も発生しないため、これを適用しても音質を劣化させる要因は殆ど無い。

【0135】

上述の可逆性と非可逆性について、以下詳細に説明する。図20に、C-R回路を用いる高音通過フィルタとL-R回路を用いる低音通過フィルタの出力の加算を示す。加算演算においては、高音通過フィルタの出力と低音通過フィルタの出力の単純な加算であり、アナログ加算器を用いても加算結果は得られるが、デジタルの単純加算を用いても加算結果は得られる。上記のアナログ・フィルタの加算結果は元の入力信号とは異なる。その原因はフィルタを通過する際に位相差が発生すること、及び過渡現象にある。従って、通常のスピーカ・システムに用いられているネットワーク・フィルタの出力波形を加算して元の入力波形と比較すると、互いに異なり、非可逆性になる。これはスピーカ・システムの音質劣化の原因の1つになる。

【0136】

次に、図21に基づき、本発明の実施例6に係る無位相シフト用フィルタの可逆性を検証する。この図に示すハイパスフィルタリングを用いる高音通過部の出力信号を高音専用スピーカに送り、ローパスフィルタリングを用いる低音通過部の出力信号を低音専用スピーカに送ることによりマルチ・ウェイのスピーカ・システムを構成する。可逆性を検証するために、ハイパスフィルタリングの出力とローパスフィルタリングの出力の加算を求める。加算演算においては、ハイパスフィルタリングの出力とローパスフィルタリングの出力の単純な加算であり、アナログ加算器を用いても加算結果は得られるが、デジタルの単純加算を用いても加算結果は得られる。上記の加算結果は元の入力信号と完全に一致する

10

20

30

40

50

ので、可逆性が保証される。すなわち、音信号は低い周波数成分と高い周波数成分に分割されるが、位相歪みも過渡現象も発生しないので、分割した信号を合成すると、分割前の信号に完全に一致する。この可逆性は、オーディオ装置においては音再現性の面で極めて重要である。なお図 2 1 において、ハイパスフィルタリングは減算値に相当し、ローパスフィルタリングは除算値に相当する。

【 0 1 3 7 】

このように、マルチ・ウェイ方式のスピーカ・システムにおいては、本発明に係るハイパスフィルタリングを用いて高音帯域通過部（ハイパスフィルタリング）を構成し、その出力を高音専用スピーカに入力し、デジタル積分を用いて低音帯域通過部（ローパスフィルタリング）を構成し、その出力を低音専用スピーカに入力することによりネットワーク・フィルタを構成すると、アナログのフィルタがもつ全ての音質劣化の要因を回避することができるので、オーディオ装置を用いて音質を楽しむことができるようになる。

【 実施例 7 】

【 0 1 3 8 】

（脳波測定への適用例）

次に、本発明に係るデジタルフィルタリングを脳波測定に適用した実施例 7 について、図 2 2 に基づき説明する。図 2 2 は、頭部前頭葉から採取した脳波のスペクトル波形を示している。この図において、細線で示す曲線は採取した脳波に何の処理も加えずにフーリエ変換により求めたスペクトルである。なお、脳波採取の際には、60 Hz のハム・フィルタをツイン T ブリッジ・フィルタにより施したが、電源から電磁波として 60 Hz 及びその高調波成分である 120 Hz と 180 Hz の雑音が混入している。脳波の源信号に対して、 $N = 16$ として、デジタルフィルタリングを施し、デジタル積分を求めると、図 2 2 においての太線で示すスペクトルが得られる。さらに、脳波の源信号にハイパスフィルタリングを施すと、図 2 2 の波線で示すスペクトルが得られる。この図において、太線で示すデジタル積分のスペクトルは電源からのノイズが完全に除去されており、10 Hz 付近にあるアルファ波も顕著に検出される。

【 0 1 3 9 】

一方、従来のフィルタを脳波に施すと、脳波の位相が変化するので、位相に着目した情報が失われる。これに対し、本実施例 7 に係るデジタルフィルタリングを施すと、位相を完全に保つことができ、しかも電源からのハム・ノイズを完全に除去することが可能となり、非常に良好に信号を採取し、処理することが可能となる。複数の電極を用いて脳電位図を描く際には、複数の脳波信号の位相を保持して雑音を除去する必要があるので、デジタルフィルタリングは最適であり、従来のアナログ及びデジタルのフィルタはこの目的のためには適していない。従って、デジタルフィルタリングは脳波や心電波形や筋電波形などの生体信号の計測の際に混入するノイズの除去には最適であり、心電計、脳波計及び筋電計などの生体計測装置のフィルタに用いると、その無位相シフトのフィルタ特性が有効に活用される。

【 実施例 8 】

【 0 1 4 0 】

（電磁波の受信）

さらに、本発明に係るデジタルフィルタリングを、電磁波の受信に適用した実施例 8 を図 2 3 ~ 図 2 4 に基づいて説明する。図 2 3 は、2.4 GHz 帯域の電磁波を発信し、その反射波をアンテナで受信し、A/D 変換した信号波形を示している。信号波形から対物の振動状態を測定するために用いる CT (Computerized Tomography) などの装置においては、ノイズの除去は精度を向上させるために必要な処理である。通常のノイズ除去にはフィルタを用いるが、周波数が高くなり、周波数が変動する場合には適切なフィルタのパラメータを設定することが困難となる。

【 0 1 4 1 】

これに対して図 2 3 に示す電磁波の信号波形にデジタルフィルタリングを施し、デジタル積分を施すと、図 2 4 に示す波形になる。この図に示すように、デジタルフィルタリン

10

20

30

40

50

グによって図 2 3 に含まれているノイズ成分が完全に除去される、しかも、図 2 3 の波形と図 2 4 の波形を厳密に比較すると、両者の位相の差異は全く発生していない。従来のアナログ・フィルタやデジタル・フィルタでは、これらを適用すると必ず信号の位相の変化を伴い、後の処理に誤差を発生させる原因になる。これに対して、実施例 8 に係るデジタルフィルタリングを施すと、位相の変化を伴わずに、ノイズ成分を完全に除去することが可能であるので、その有効性は高い。特に、RF (Radio Frequency) 信号を用いて計測する CT や MRI などの計測には、無位相シフト特性を持つデジタルフィルタリングを施すと、フィルタリング特性は良好であり、後の信号処理に位相誤差が発生しないので、デジタルフィルタリングは極めて有効な手法であると言える。

【産業上の利用可能性】

10

【0142】

本発明のデジタルフィルタリング方法、デジタルフィルタ装置、デジタルフィルタプログラム及びコンピュータで読み取り可能な記録媒体並びに記録した機器は、音声認識等の音声信号の前処理や後処理に好適に適用でき、例えば個人適応型の音声認識システムとして携帯電話等のモバイル機器に適用できる。また音声認識に限らず、スピーカの再現品質向上、脳波測定、電波の受信等においても、高品質で低負荷の信号処理を実現できる。

【図面の簡単な説明】

【0143】

【図 1】本発明の一実施の形態に係る音声認識装置の電子線撮像部の構成を示すブロック図である。

20

【図 2】振幅波形の N 個のデータについて除算値 k_k を求める様子を示す説明図である。

【図 3】入力される音声波形とそのピッチ情報を示す説明図である。

【図 4】定常とした 3 周期の母音波形を示すグラフである。

【図 5】 $N = 256$ として図 4 から抽出した 1 周期分の凹凸波形を示すグラフである。

【図 6】 $N = 64$ として図 4 から抽出した 1 周期分の凹凸波形を示すグラフである。

【図 7】凸部分の / e / 音素波形と生成した正弦波の波形を示すグラフである。

【図 8】凸部分の / o / 音素波形と生成した正弦波の波形を示すグラフである。

【図 9】意識発話に用いた / i / と / u / に関する特徴分布を示すグラフである。

【図 10】意識発話に用いた / e / と / o / に関する特徴分布を示すグラフである。

【図 11】本発明の実施例 2 に係る聴覚補助装置の構成を示すブロック図である。

30

【図 12】音声信号処理方法の手順を示すフローチャートである。

【図 13】音声の最も低い周波数に対応する波形を示すグラフである。

【図 14】音声の最も高い周波数に対応する波形を示すグラフである。

【図 15】低音通過フィルタ回路を示す回路図である。

【図 16】図 15 の L - R 回路の周波数特性を示すグラフである。

【図 17】高音通過フィルタ回路を示す回路図である。

【図 18】図 17 の C - R 回路の周波数特性を示すグラフである。

【図 19】2 ウエイ方式のスピーカ・システムのネットワーク・フィルタを示す回路図である。

【図 20】図 17 の高音通過フィルタと図 16 の低音通過フィルタの出力を加算した回路を示すブロック図である。

40

【図 21】本発明の実施の形態に係る無位相シフト用フィルタを示すブロック図である。

【図 22】頭部前頭葉から採取した脳波のスペクトル波形を示すグラフである。

【図 23】2 . 4 GHz 帯域の電磁波の反射波をアンテナで受信し、A / D 変換した信号波形を示すグラフである。

【図 24】図 2 3 に示す電磁波の信号波形にデジタルフィルタリングを施し、デジタル積分を施した波形を示すグラフである。

【図 25】本発明の一実施の形態に係るデジタルフィルタリング方法の手順を示すフローチャートである。

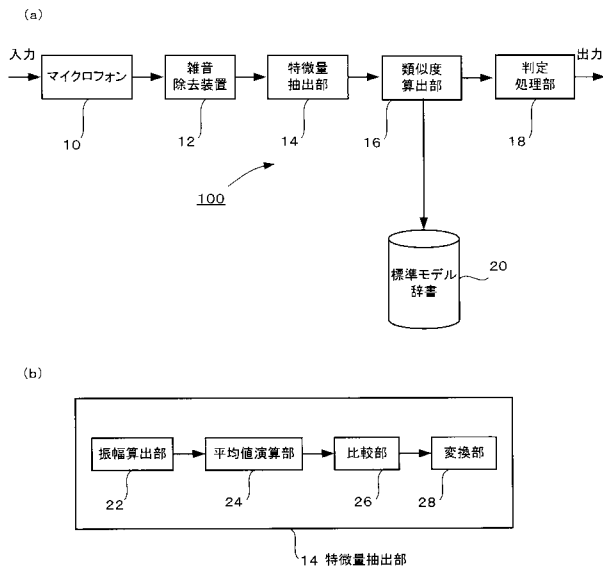
【符号の説明】

50

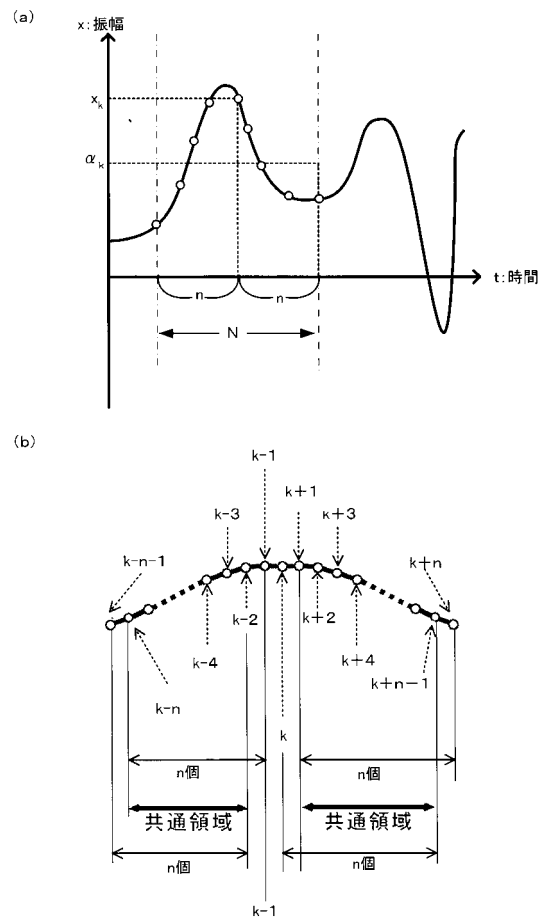
【 0 1 4 4 】

- 1 0 0 ... 音声認識装置
- 2 0 0 ... 聴覚補助装置
- 1 0 ... マイクフォン ; 1 2 ... 雑音除去装置
- 1 4 ... 特徴量抽出部 ; 1 6 ... 類似度算出部 ; 1 8 ... 判定処理部 ; 2 0 ... 標準モデル辞書
- 2 2、2 2 A ... 振幅算出部
- 2 4、2 4 A ... 除算値演算部
- 2 6、2 6 A ... 比較部
- 2 8、2 8 A ... 変換部
- 1 0 A ... 音声入力部 ; 1 2 A ... A / D 変換器
- 1 4 A ... 凹凸波形変換部 ; 1 6 A ... 声成分抽出部
- 1 8 A ... 強調処理部 ; 2 0 A ... 音声出力部

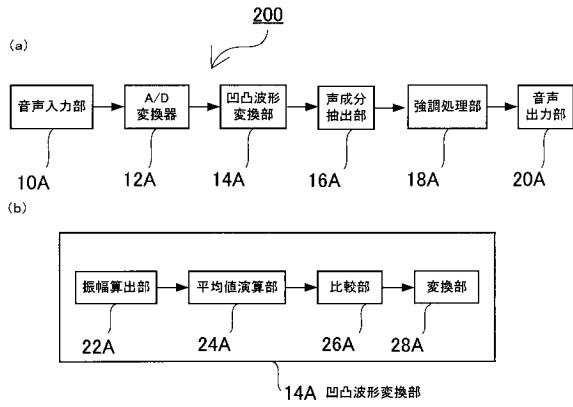
【 図 1 】



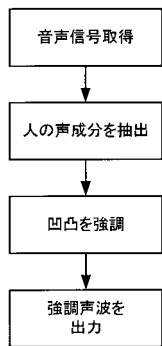
【 図 2 】



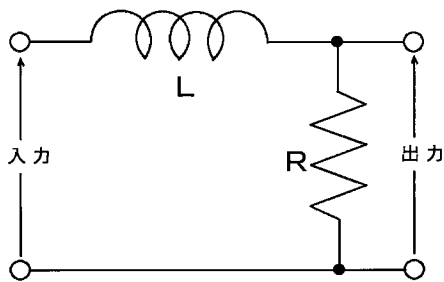
【図11】



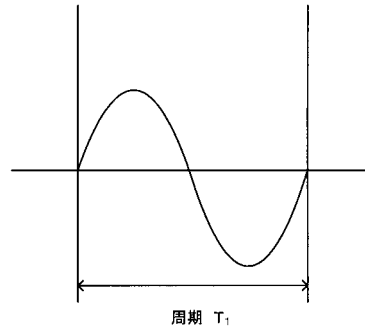
【図12】



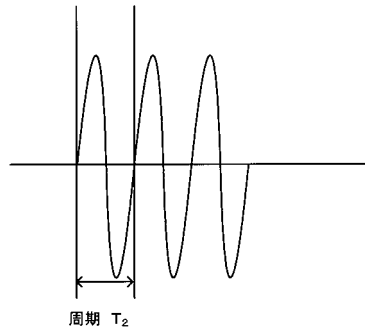
【図15】



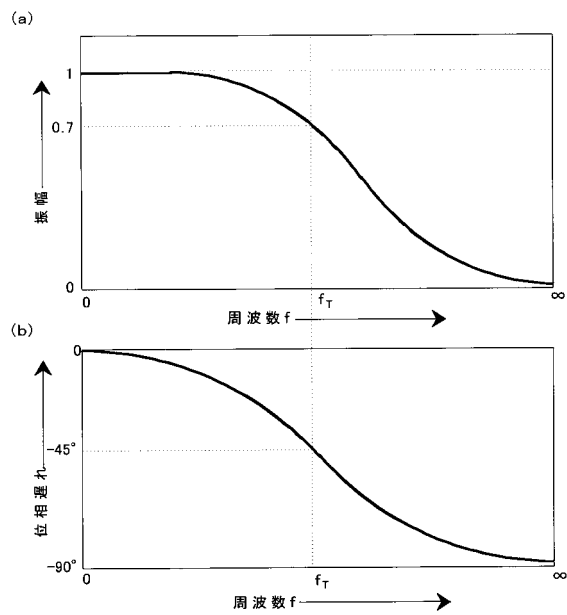
【図13】



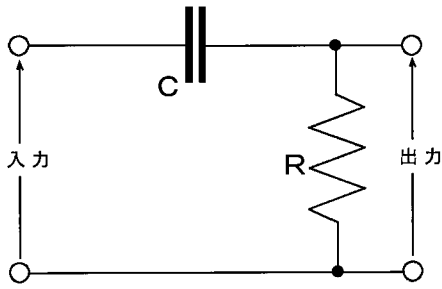
【図14】



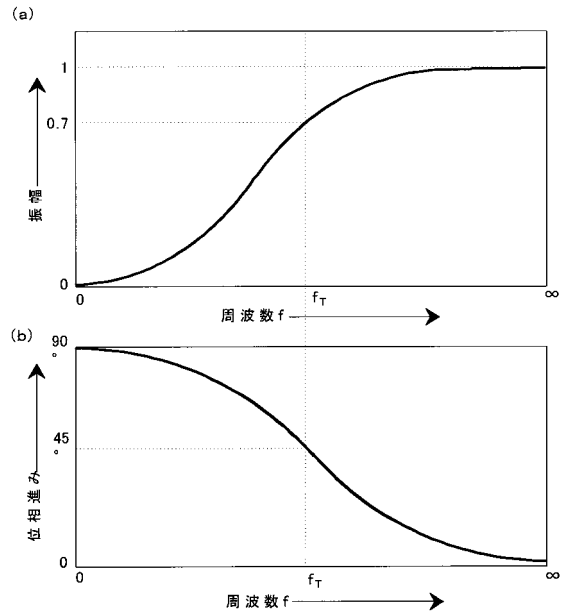
【図16】



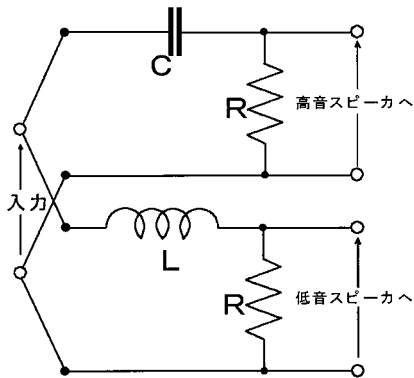
【図 17】



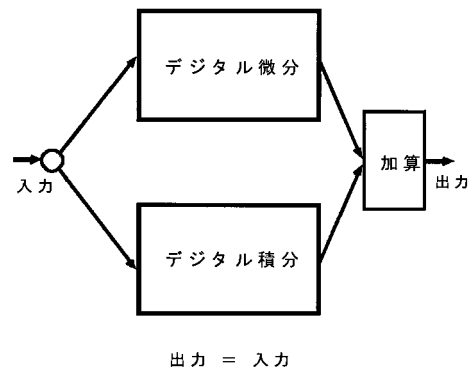
【図 18】



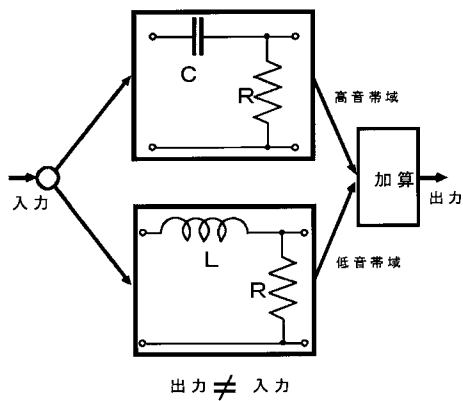
【図 19】



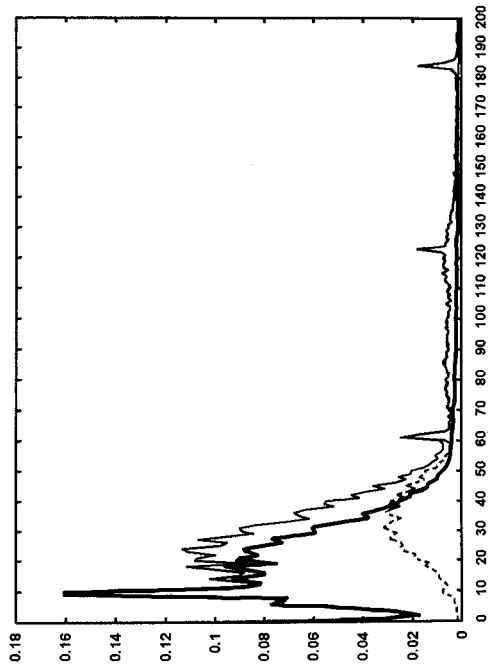
【図 21】



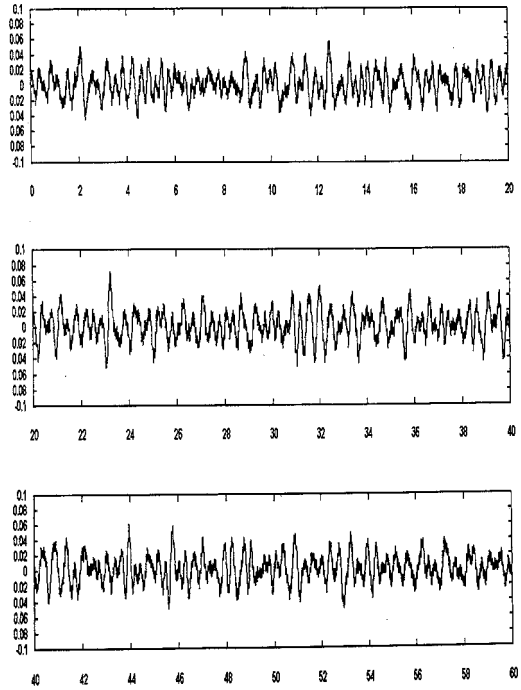
【図 20】



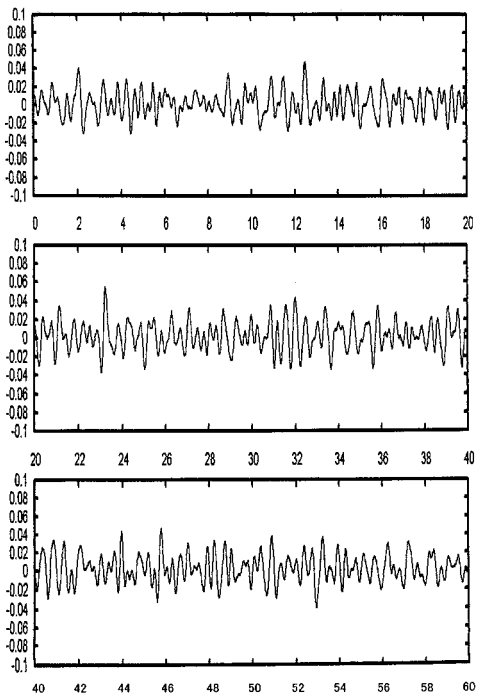
【 図 2 2 】



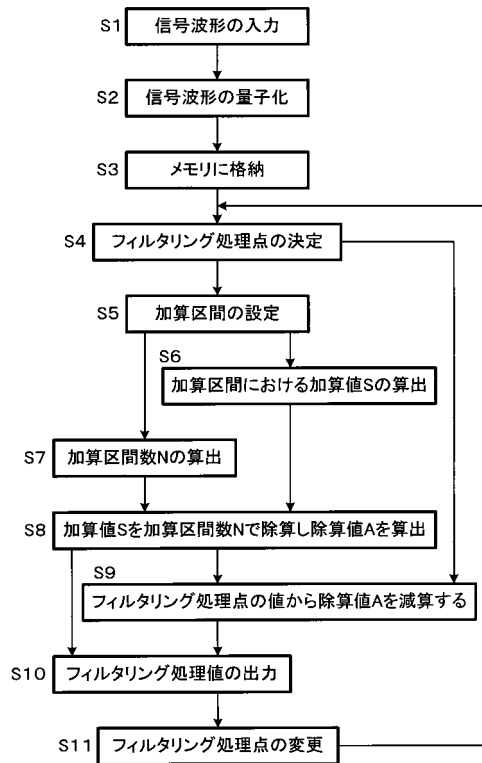
【 図 2 3 】



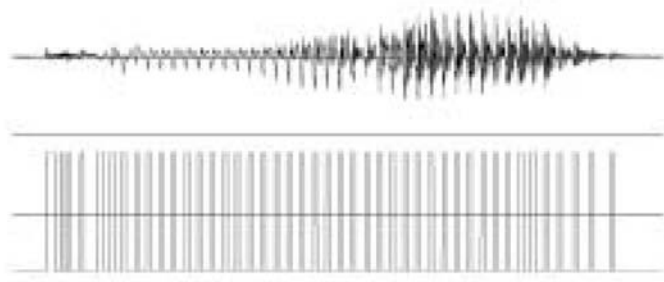
【 図 2 4 】



【 図 2 5 】

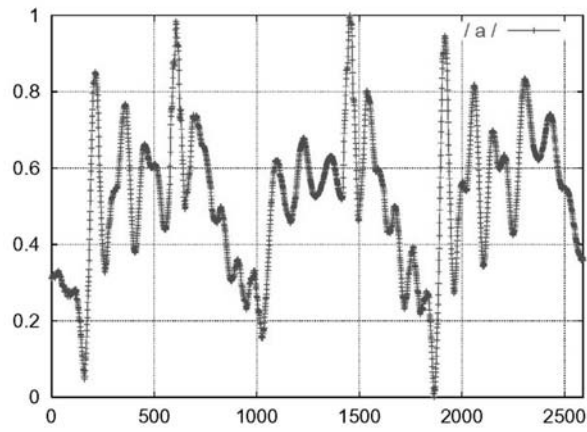


【図3】



取得した音声波形とピッチ

【図4】



定常とした3周期の母音波形(/a/)

【 図 5 】

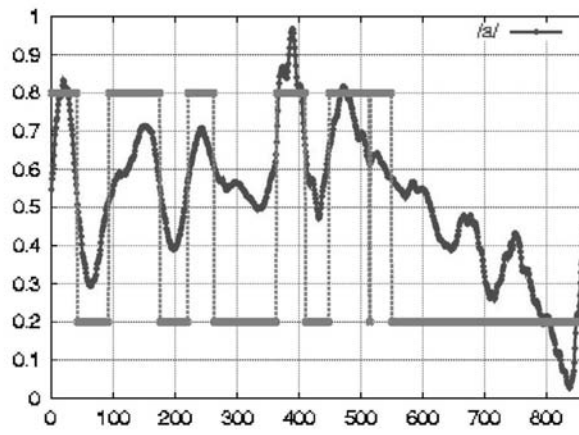


図4から抽出した1周期分の256凹凸波形

【 図 6 】

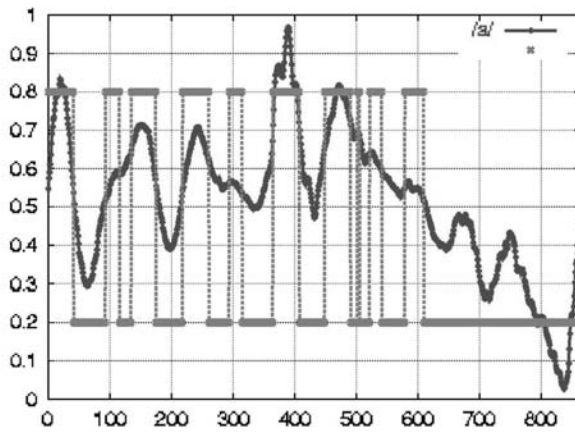
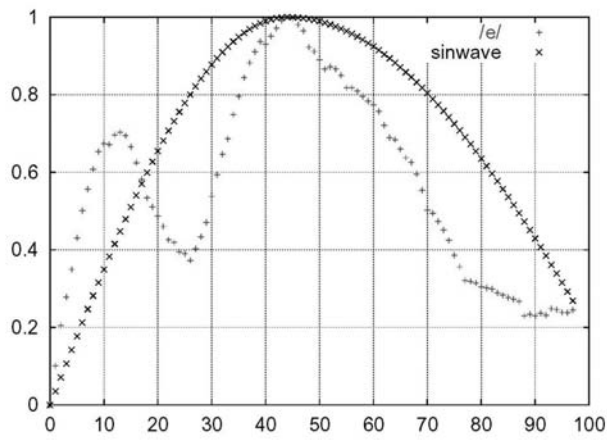


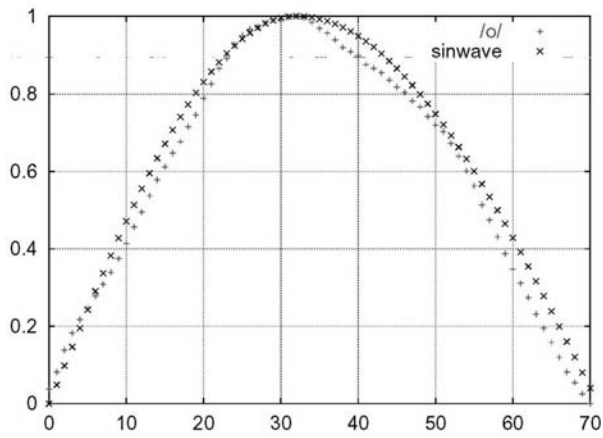
図4から抽出した1周期分の64凹凸波形

【 図 7 】



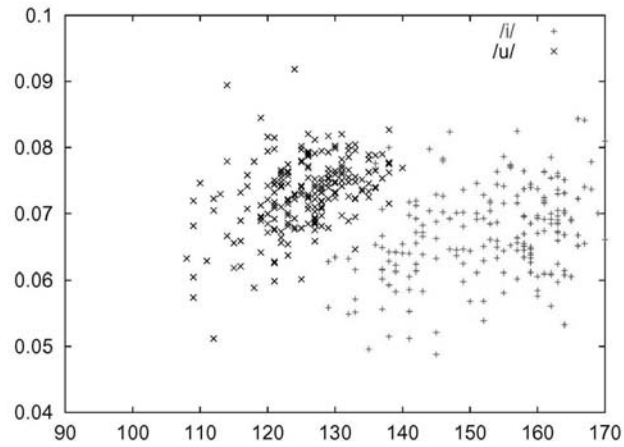
凸部分の/e/音素波形と生成した正弦波

【 図 8 】



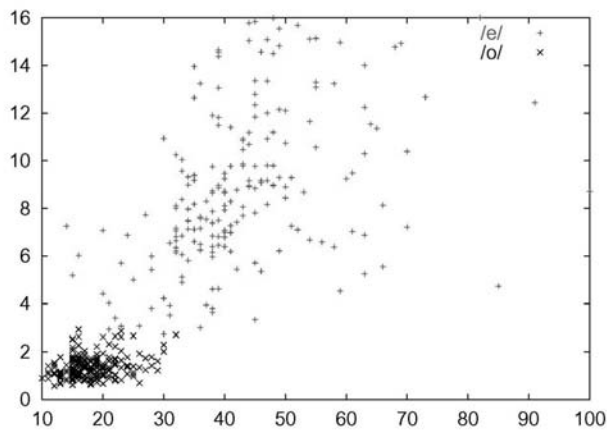
凸部分の/o/音素波形と生成した正弦波

【 図 9 】



/i/と/u/に関する特徴分布

【 図 10 】



/e/と/o/に関する特徴分布

フロントページの続き

- (56)参考文献 特開平6 - 140856 (JP, A)
特開2001 - 350488 (JP, A)
特開2003 - 316378 (JP, A)
特開2002 - 14689 (JP, A)
特開2000 - 152394 (JP, A)
矢間優希, 満倉靖恵, 福見稔, 赤松則男, モバイルインタフェースのための離散ポロノイ図と統計的処理による音圧情報を用いた母音識別, 電子情報通信学会技術研究報告, 日本, 社団法人電子情報通信学会, 2004年 7月 2日, Vol. 104, No. 169, p. 1 - 6

(58)調査した分野(Int.Cl., DB名)

G10L 11/00-21/06
H04R 25/00-25/04
IEEE
JSTPlus(JDreamII)
JST7580(JDreamII)