

(19) 日本国特許庁(JP)

(12) 公 開 特 許 公 報(A)

(11) 特許出願公開番号
特開2006-79089
(P2006-79089A)

(43) 公開日 平成18年3月23日 (2006.3.23)

(51) Int. Cl.	F I	テーマコード (参考)
G 1 O L 19/00 (2006.01)	G 1 O L 19/00 1 O O	5 D O 1 5
G 1 O L 15/00 (2006.01)	G 1 O L 3/00 5 5 1 A	
G 1 O L 15/28 (2006.01)	G 1 O L 15/28 2 1 O Z	

審査請求 有 請求項の数 26 O L (全 14 頁)

(21) 出願番号 特願2005-259381 (P2005-259381)	(71) 出願人 502032105
(22) 出願日 平成17年9月7日 (2005.9.7)	エルジー エレクトロニクス インコーポ レイティド
(31) 優先権主張番号 10-2004-0071327	大韓民国, ソウル 1 5 0 - 7 2 1 , ヨン ドンポーク, ヨイドードン, 2 O
(32) 優先日 平成16年9月7日 (2004.9.7)	(74) 代理人 100078282
(33) 優先権主張国 韓国 (KR)	弁理士 山本 秀策
	(74) 代理人 100062409
	弁理士 安村 高明
	(74) 代理人 100113413
	弁理士 森下 夏樹

最終頁に続く

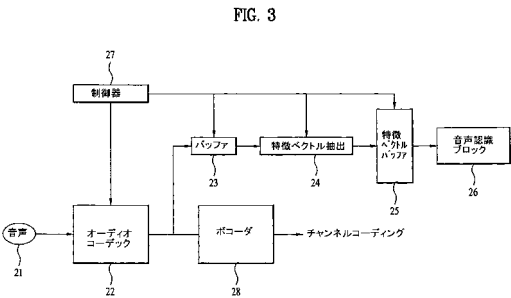
(54) 【発明の名称】 音声認識用ベースバンドのモデムと、それを用いる移動通信端末機

(57) 【要約】

【課題】 音声認識に適したサンプリングレートで調節することによって音声認識率を高められると共に、音声認識のためのプロセッシングユニットの一部の処理部分をハードウェア的に実現できる音声認識用ベースバンドモデムと、それを用いる移動通信端末機を提供する。

【解決手段】 音声認識のためのサンプリングレートと、音声通話のためのサンプリングレートのうち何れか一つを用いて、入力されたボイス信号に対して変調を実施するオーディオコーデックと、音声認識のためのサンプリングレートを用いて、変調されたオーディオコーデックの出力から特徴ベクトルを抽出する特徴ベクトル抽出ブロックと、特徴ベクトル抽出ブロックの出力を用いて、音声認識を行う音声認識ブロックと、音声通話のためのサンプリングレートを用いて、変調されたオーディオコーデックの出力をボコーディングするボコーダとを含んで構成される、ベースバンドモデム。

【選択図】 図 3



【特許請求の範囲】

【請求項 1】

音声認識のためのサンプリングレートと、音声通話のためのサンプリングレートのうち何れか一つを用いて、入力されたボイス信号に対して変調を実施するオーディオコーデックと、
前記音声認識のためのサンプリングレートを用いて、変調された前記オーディオコーデックの出力から特徴ベクトルを抽出する特徴ベクトル抽出ブロックと、
前記特徴ベクトル抽出ブロックの出力を用いて、音声認識を行う音声認識ブロックと、
前記音声通話のためのサンプリングレートを用いて、変調された前記オーディオコーデックの出力をボコーディグするボコーダとを含んで構成されることを特徴とするベースバンドモデム。 10

【請求項 2】

前記音声認識のためのサンプリングレートを用いて、変調された前記オーディオコーデックの出力を保存するバッファを更に含んで構成されることを特徴とする請求項 1 に記載のベースバンドモデム。

【請求項 3】

前記特徴ベクトル抽出ブロックは、前記バッファに保存されたデータから前記特徴ベクトルを抽出することを特徴とする請求項 2 に記載のベースバンドモデム。

【請求項 4】

前記バッファは、ピンポンバッファであることを特徴とする請求項 2 に記載のベースバンドモデム。 20

【請求項 5】

前記特徴ベクトル抽出ブロックから抽出された前記特徴ベクトルを保存する特徴ベクトルバッファを更に含んで構成されることを特徴とする請求項 1 に記載のベースバンドモデム。

【請求項 6】

前記入力されたボイス信号が前記音声認識のための信号であるか、前記音声通話のための信号であるかを区分する制御器を更に含んで構成されることを特徴とする請求項 1 に記載のベースバンドモデム。

【請求項 7】

前記制御器は、
前記入力されたボイス信号を、前記音声認識のための信号と、前記音声通話のための信号のうち一つで区分することにより、前記オーディオコーデックの変調に用いるサンプリングレートの可変を制御することを特徴とする請求項 6 に記載のベースバンドモデム。 30

【請求項 8】

前記制御器は、
前記入力されたボイス信号を、前記音声認識のための信号と、前記音声通話のための信号のうち一つで区分することにより、前記特徴ベクトル抽出ブロックと前記音声認識ブロックの活性化を制御することを特徴とする請求項 6 に記載のベースバンドモデム。

【請求項 9】

前記入力されたオーディオ信号が前記音声認識のための信号である場合に、前記制御器は、前記特徴ベクトル抽出ブロックと、前記音声認識ブロックを活性化させることを特徴とする請求項 8 に記載のベースバンドモデム。 40

【請求項 10】

前記制御器は、前記ベースバンドモデムに備えられた音声認識のためのレジスターをオン/オフして、前記特徴ベクトル抽出ブロックと前記音声認識ブロックの活性化を制御することを特徴とする請求項 6 に記載のベースバンドモデム。

【請求項 11】

前記音声認識のためのサンプリングレートは、12～32 kHz の範囲に含まれることを特徴とする請求項 1 に記載のベースバンドモデム。 50

【請求項 1 2】

前記音声認識のためのサンプリングレートは、16 kHzであることを特徴とする請求項 1 に記載のベースバンドモデム。

【請求項 1 3】

前記特徴ベクトル抽出ブロックは、ハードウェアで構成されることを特徴とする請求項 1 に記載のベースバンドモデム。

【請求項 1 4】

前記オーディオコーデックは、前記入力されたボイス信号に対してパルスコード変調を実施することを特徴とする請求項 1 に記載のベースバンドモデム。

【請求項 1 5】

前記ベースバンドモデムは、移動通信端末機に内装されることを特徴とする請求項 1 に記載のベースバンドモデム。

【請求項 1 6】

音声認識用サンプリングレートと、音声通話用サンプリングデータのうち何れか一つを用いて、入力されたボイス信号に対してパルスコード変調を実施するオーディオコーデックと、

前記音声認識用サンプリングレートで変調された前記オーディオコーデックの出力から特徴ベクトルを抽出する特徴ベクトル抽出ブロックと、前記特徴ベクトル抽出ブロックの出力を用いて音声認識アルゴリズムを行うプロセッサを備えた音声認識経路と、

前記音声通話用サンプリングレートで変調された前記オーディオコーデックの出力を音声コーディングするボコーダを備えた音声通話経路とを含んで構成されることを特徴とする移動通信端末機。

【請求項 1 7】

前記音声認識経路は、

前記音声認識用サンプリングレートで変調された前記オーディオコーデックの出力を保存するバッファを、特徴ベクトル抽出ブロックの入力端に更に備えることを特徴とする請求項 1 6 に記載の移動通信端末機。

【請求項 1 8】

前記バッファは、ピンポンバッファであることを特徴とする請求項 1 7 に記載の移動通信端末機。

【請求項 1 9】

前記音声認識経路は、

前記特徴ベクトル抽出ブロックから抽出された前記特徴ベクトルを保存する特徴ベクトルバッファを、前記プロセッサの入力端に更に備えることを特徴とする請求項 1 6 に記載の移動通信端末機。

【請求項 2 0】

前記入力されたボイス信号が前記音声認識用信号であるか、前記音声通話用信号であるかを区分して、前記経路の動作を制御する制御器を更に含んで構成されることを特徴とする請求項 1 6 に記載の移動通信端末機。

【請求項 2 1】

前記制御器は、

前記入力されたボイス信号を、前記音声認識用信号と、前記音声通話用信号のうち一つで区分し、それによって前記オーディオコーデックの変調に用いるサンプリングレートを可変させることを特徴とする請求項 2 0 に記載の移動通信端末機。

【請求項 2 2】

前記制御器は、

前記入力されたボイス信号が前記音声認識用信号である場合に、前記特徴ベクトル抽出ブロックと、前記音声認識ブロックの動作を活性化させることを特徴とする請求項 2 0 に記載の移動通信端末機。

【請求項 2 3】

前記制御器は、前記移動通信端末機に備えられた音声認識用レジスターをオン/オフして、前記特徴ベクトル抽出ブロックと、前記音声認識ブロックの動作を制御することを特徴とする請求項 20 に記載の移動通信端末機。

【請求項 24】

前記音声認識用サンプリングレートは、12～32kHz の範囲に含まれることを特徴とする請求項 16 に記載の移動通信端末機。

【請求項 25】

前記音声認識用サンプリングレートは、16kHz であることを特徴とする請求項 1 に記載の移動通信端末機。

【請求項 26】

前記特徴ベクトルブロックは、ハードウェアで構成されることを特徴とする請求項 1 に記載の移動通信端末機。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、より高い音声認識率を保障する音声認識用ベースバンドのモデムと、それを用いる移動通信端末機に関する。

【背景技術】

【0002】

従来のベースバンドモデム（ベースバンドモデム）は、オーディオコーデックを含む。

【0003】

一方、従来の音声認識技術を移動通信端末機に適用する時は、大部分音声通話のための音声コーディングと音声認識で同一のサンプリングレートの音声を用いた。

【0004】

これは、16kHz マイクロフォンの入力を支援するベースバンドモデムが多くないためであり、現在、多くのベースバンドモデムがパルスコード変調（Pulse Code Modulation：以下、PCM）データを得ることが難しいからである。

【0005】

図 1 は従来技術に係るベースバンドモデムの構成図である。図 2 は、図 1 に示すベースバンドモデムを介した音声認識過程を説明した図面である。

【0006】

図 1 を参照すると、従来のベースバンドモデムは、オーディオコーデック 13 と、ボコーダ 15 と、プロセッサ 17 とで構成されている。

【0007】

オーディオコーデック 13 は、マイクロフォンからボイス信号が入力されると、そのボイス信号を所定のサンプリングレートで変調する。一例として、入力されたボイス信号を 8kHz のサンプリングレートで PCM を実施する。

【0008】

ボコーダ 15 は、オーディオコーデック 13 の出力に対してボコーディングを実施する。

【0009】

一例として、音声コーディング方式として Q U E L P (Qualcomm Code Excited Linear Prediction) や E V R C (Enhanced Variable Rate Coding) が用いられる。

【0010】

プロセッサ 17 は、ボコーダ 15 の出力に対して音声認識を実施する。より詳細に、プロセッサ 17 は、音声コーディングされたデータを復号化し、その復号化されたデータから特徴ベクトルを抽出する。そして、予め備えられた音声認識アルゴリズムに前記抽出された特徴ベクトルを用いて音声認識を実施する。

【0011】

前記プロセッサ 17 は、マイクロプロセッシングユニット (M P U : Micro Processing Unit) やデジタルシグナリングプロセッサ (D S P : Digital Signaling Processo

10

20

30

40

50

r)であることが好ましい。

【0012】

反面、前記入力されたボイス信号が音声通話のためのものである場合は、プロセッサ17は、ボコーダ15の出力に対してコンボルーションコードやターボコードなどを用いるチャンネルコーディングを実施する。

【0013】

上記の構成による音声認識過程を図2に基づいて以下に説明する。

【0014】

従来のベースバンドモデムは、マイクロフォンからボイス信号が入力されると、そのボイス信号を所定のサンプリングレートで変調する(S12)。一例として、入力されたボイス信号を8kHzのサンプリングレートでPCMを実施する。 10

【0015】

次いで、変調の結果に対してボイスコーディングを実施する(S14)。音声コーディング方式としては、QCELP(Qualcomm Code Excited Linear Prediction)やEVR Cが好ましい。

【0016】

次いで、音声コーディングの出力に対してMPU(Micro Processing Unit)やDSP(digital Signaling Processor)の内部で音声認識を実施する。前記の音声認識のために、まず、音声コーディングされたデータを復号化する(S16)。そして、復号化されたデータから特徴ベクトルを抽出する(S18)。 20

【0017】

最後に、音声認識アルゴリズムに前記抽出された特長ベクトルを用いて、音声認識を実施する(S20)。

【0018】

上記説明された従来技術で変調のためのサンプリングレートを8kHzにする理由は、4kHz以下のボイス成分でも十分に聞き取れるほどの音質水準を出すことができるからである。

【発明の開示】

【発明が解決しようとする課題】

【0019】

しかしながら、従来の移動通信端末機で音声認識を実施する場合には、大部分音声通話のためにサンプリングを経たデータを、音声認識用として用いる。そのため、従来技術では、満足すべき音声認識率を保障できなかった。 30

【0020】

また、従来技術においては、図2で説明したように、音声認識のためには必要のない声コーディングと復号化を経なければならない。

【0021】

勿論、音声認識のためのデジタル信号処理プロセッサチップや、音声認識用チップを移動通信端末機に用いる方法もあるが、これは端末機の単価を上昇させる要因となる。

【0022】

一方、従来の一部のベースバンドモデムは、音声認識のためにDTW(Dynamic Time Warping)のような方法を用いた。しかし、この方法も音声通話のためにサンプリングを経たデータを音声認識用に用いるため、十分な音声認識率を保障することはできなかった。 40

【0023】

また、従来は、音声認識のための動作時にも、ベースバンドモデムの内部に備えられたオーディオコーデックにサンプリングレートを上げたり、特徴ベクトルを抽出する部分がハードウェア的に具現されていなかった。

【0024】

さらに、音声認識のためのまた他の方法としては、ベースバンドモデムの外部に音声認識に適したサンプリングレートを有する別途のオーディオコーデックを設置する方法があ 50

る。しかし、これはハードウェア的な具現が容易ではない。

要すると、従来技術では、移動通信端末機で音声認識を実施するにおいて、音声通話と音声認識とを区分して、ベースバンドモデムのサンプリングレートを調節できるようにする技術がなかった。また、ベースバンドモデムがPCMデータを得ることが難しいという問題があった。

【0025】

本発明は上記の問題点を解決するために案出したもので、音声認識に適したサンプリングレートで調節して、音声認識率を高めることのできる音声認識用ベースバンドモデムと、それを用いる移動通信端末機を提供することにその目的がある。

【0026】

他の目的として、音声認識のためのプロセッシングユニットの一部の処理部分をハードウェア的に実現できる音声認識用ベースバンドモデムと、それを用いる移動通信端末機を提供することにある。

【課題を解決するための手段】

【0027】

上記目的を達成するために、本発明に係るベースバンドモデムは、音声認識のためのサンプリングレートと、音声通話のためのサンプリングレートのうち何れか一つを用いて、入力されたボイス信号に対して変調を実施するオーディオコーデックと、前記音声認識のためのサンプリングレートを用いて、変調された前記オーディオコーデックの出力から特徴ベクトルを抽出する特徴ベクトル抽出ブロックと、前記特徴ベクトル抽出ブロックの出力を用いて、音声認識を行う音声認識ブロックと、前記音声通話のためのサンプリングレートを用いて、変調された前記オーディオコーデックの出力を音声コーディングするボコーダとを含んで構成されることを特徴とする。

【0028】

上記目的を達成するために、本発明に係る移動通信端末機は、音声認識用サンプリングレートと、音声通話用サンプリングデータのうち何れか一つを用いて、入力されたボイス信号に対してパルスコード変調を実施するオーディオコーデックと、前記音声認識用サンプリングレートで変調された前記オーディオコーデックの出力から特徴ベクトルを抽出する特徴ベクトル抽出ブロックと、前記特徴ベクトル抽出ブロックの出力を用いて音声認識アルゴリズムを行うプロセッサを備えた音声認識経路と、前記音声通話用サンプリングレートで変調された前記オーディオコーデックの出力を音声コーディングするボコーダを備えた音声通話経路とを含んで構成されることを特徴とする。

【0029】

上記目的を達成するために、本発明は、例えば、以下の手段を提供する。

(項目1)

音声認識のためのサンプリングレートと、音声通話のためのサンプリングレートのうち何れか一つを用いて、入力されたボイス信号に対して変調を実施するオーディオコーデックと、
前記音声認識のためのサンプリングレートを用いて、変調された前記オーディオコーデックの出力から特徴ベクトルを抽出する特徴ベクトル抽出ブロックと、
前記特徴ベクトル抽出ブロックの出力を用いて、音声認識を行う音声認識ブロックと、
前記音声通話のためのサンプリングレートを用いて、変調された前記オーディオコーデックの出力をボコーディングするボコーダとを含んで構成されることを特徴とするベースバンドモデム。

(項目2)

前記音声認識のためのサンプリングレートを用いて、変調された前記オーディオコーデックの出力を保存するバッファを更に含んで構成されることを特徴とする項目1に記載のベースバンドモデム。

(項目3)

前記特徴ベクトル抽出ブロックは、前記バッファに保存されたデータから前記特徴ベクトル

10

20

30

40

50

ルを抽出することを特徴とする項目 2 に記載のベースバンドモデム。

(項目 4)

前記バッファは、ピンポンバッファであることを特徴とする項目 2 に記載のベースバンドモデム。

(項目 5)

前記特徴ベクトル抽出ブロックから抽出された前記特徴ベクトルを保存する特徴ベクトルバッファを更に含んで構成されることを特徴とする項目 1 に記載のベースバンドモデム。

(項目 6)

前記入力されたボイス信号が前記音声認識のための信号であるか、前記音声通話のための信号であるかを区分する制御器を更に含んで構成されることを特徴とする項目 1 に記載のベースバンドモデム。 10

(項目 7)

前記制御器は、

前記入力されたボイス信号を、前記音声認識のための信号と、前記音声通話のための信号のうち一つで区分することにより、前記オーディオコーデックの変調に用いるサンプリングレートの可変を制御することを特徴とする項目 6 に記載のベースバンドモデム。

(項目 8)

前記制御器は、

前記入力されたボイス信号を、前記音声認識のための信号と、前記音声通話のための信号のうち一つで区分することにより、前記特徴ベクトル抽出ブロックと前記音声認識ブロックの活性化を制御することを特徴とする項目 6 に記載のベースバンドモデム。 20

(項目 9)

前記入力されたオーディオ信号が前記音声認識のための信号である場合に、前記制御器は、前記特徴ベクトル抽出ブロックと、前記音声認識ブロックを活性化させることを特徴とする項目 8 に記載のベースバンドモデム。

(項目 10)

前記制御器は、前記ベースバンドモデムに備えられた音声認識のためのレジスターをオン/オフして、前記特徴ベクトル抽出ブロックと前記音声認識ブロックの活性化を制御することを特徴とする項目 6 に記載のベースバンドモデム。

(項目 11)

前記音声認識のためのサンプリングレートは、12～32 kHz の範囲に含まれることを特徴とする項目 1 に記載のベースバンドモデム。 30

(項目 12)

前記音声認識のためのサンプリングレートは、16 kHz であることを特徴とする項目 1 に記載のベースバンドモデム。

(項目 13)

前記特徴ベクトル抽出ブロックは、ハードウェアで構成されることを特徴とする項目 1 に記載のベースバンドモデム。

(項目 14)

前記オーディオコーデックは、前記入力されたボイス信号に対してパルスコード変調を実施することを特徴とする項目 1 に記載のベースバンドモデム。 40

(項目 15)

前記ベースバンドモデムは、移動通信端末機に内装されることを特徴とする項目 1 に記載のベースバンドモデム。

(項目 16)

音声認識用サンプリングレートと、音声通話用サンプリングデータのうち何れか一つを用いて、入力されたボイス信号に対してパルスコード変調を実施するオーディオコーデックと、

前記音声認識用サンプリングレートで変調された前記オーディオコーデックの出力から特徴ベクトルを抽出する特徴ベクトル抽出ブロックと、前記特徴ベクトル抽出ブロックの出 50

力を用いて音声認識アルゴリズムを行うプロセッサを備えた音声認識経路と、前記音声通話用サンプリングレートで変調された前記オーディオコーデックの出力を音声コーディングするボコーダを備えた音声通話経路とを含んで構成されることを特徴とする移動通信端末機。

(項目 17)

前記音声認識経路は、前記音声認識用サンプリングレートで変調された前記オーディオコーデックの出力を保存するバッファを、特徴ベクトル抽出ブロックの入力端に更に備えることを特徴とする項目 16 に記載の移動通信端末機。

(項目 18)

前記バッファは、ピンポンバッファであることを特徴とする項目 17 に記載の移動通信端末機。

(項目 19)

前記音声認識経路は、前記特徴ベクトル抽出ブロックから抽出された前記特徴ベクトルを保存する特徴ベクトルバッファを、前記プロセッサの入力端に更に備えることを特徴とする項目 16 に記載の移動通信端末機。

(項目 20)

前記入力されたボイス信号が前記音声認識用信号であるか、前記音声通話用信号であるかを区分して、前記経路の動作を制御する制御器を更に含んで構成されることを特徴とする項目 16 に記載の移動通信端末機。

(項目 21)

前記制御器は、前記入力されたボイス信号を、前記音声認識用信号と、前記音声通話用信号のうち一つで区分し、それによって前記オーディオコーデックの変調に用いるサンプリングレートを可変させることを特徴とする項目 20 に記載の移動通信端末機。

(項目 22)

前記制御器は、前記入力されたボイス信号が前記音声認識用信号である場合に、前記特徴ベクトル抽出ブロックと、前記音声認識ブロックの動作を活性化させることを特徴とする項目 20 に記載の移動通信端末機。

(項目 23)

前記制御器は、前記移動通信端末機に備えられた音声認識用レジスタをオン/オフして、前記特徴ベクトル抽出ブロックと、前記音声認識ブロックの動作を制御することを特徴とする項目 20 に記載の移動通信端末機。

(項目 24)

前記音声認識用サンプリングレートは、12～32 kHz の範囲に含まれることを特徴とする項目 16 に記載の移動通信端末機。

(項目 25)

前記音声認識用サンプリングレートは、16 kHz であることを特徴とする項目 1 に記載の移動通信端末機。

(項目 26)

前記特徴ベクトルブロックは、ハードウェアで構成されることを特徴とする項目 1 に記載の移動通信端末機。

【発明の効果】

【0030】

本発明に係る音声認識用ベースバンドモデムと、それを用いる移動通信端末機によれば次のような効果がある。

【0031】

第一に、オーディオコーデックによる変調時に音声認識に適したサンプリングレートを

10

20

30

40

50

用いるので、音声認識率が向上する。

【0032】

第二に、特徴ベクトルの抽出のための部分をハードウェアで実現することで、音声認識のためのプロセッシングユニットの演算量を減らすことができ、また、電力消費面でも経済的である。

【0033】

第三に、音声認識アルゴリズムで固定点の移行やアルゴリズムの改善のための部分を、上述したMPUまたはDSPを介して実現することで、以後、必要による修正が容易である。

【発明を実施するための最良の形態】

10

【0034】

以下、本発明に係る音声用ベースバンドモデムと、それを用いる移動通信端末機についての好適な実施例を添付の図面に基づいて詳細に説明する。

【0035】

図3は、本発明の一実施例によるベースバンドモデムの構成を示すブロックダイアグラムである。特に、図3のベースバンドモデムは、移動通信端末機に装着されることが好ましい。図2を参照すると、本発明に係るベースバンドモデムは、オーディオコーデック22、制御器27、ボコーダ28、特徴ベクトル抽出ブロック24、多数のバッファ23、25、そして音声認識ブロック26で構成されている。

【0036】

20

オーディオコーデック22は、マイクロフォンからボイス信号が入力されると、そのボイス信号を所定のサンプリングレートで変調する。特に、オーディオコーデック22は、入力されたボイス信号を、可変的なサンプリングレートでPCMを実施する。

【0037】

上記でマイクロフォンは、入力されるユーザーの音声を経電的な信号に変換する。

【0038】

オーディオコーデック22は、入力されるボイス信号が音声認識のための信号であるか、音声通話のための信号であるかによってサンプリングレートを変更して、PCMを行う。

【0039】

30

より詳細には、オーディオコーデック22は、音声通話のための信号に対しては、8kHz程度のサンプリングレートをPCMに用いる。反面、音声認識のための信号に対しては、より高い12～32kHz範囲のサンプリングレートをPCMに用いる。

【0040】

より好ましくは、音声認識のための信号に対しては16kHzのサンプリングレートをPCMに用いる。これは、16kHzのサンプリングレートが音声認識率を向上させると知られているためである。

【0041】

一方、入力される信号が音声認識のための信号であるか、音声通話のための信号であるかの区分は、ユーザーがどんなアプリケーションを選択するかによって実現される。

40

【0042】

即ち、ユーザーが音声通話のためのアプリケーションを選択すると、以後オーディオコーデック22に入力される信号は、音声通話のためのボイス信号である。そして、ユーザーが音声認識のためのアプリケーションを選択すると、以後オーディオコーデック22に入力される信号は、音声認識のためのボイス信号である。

【0043】

そのため、移動通信端末機の立場で入力される信号が、音声認識のための信号であるか音声通話のための信号であるかを区分する。

【0044】

本発明では、制御器27が現在のユーザーが選択したアプリケーションが何であることを

50

把握して、音声通話のための信号伝達経路と、音声認識のための信号伝達経路のうち、一つの経路を活性化させる。より正確には、制御器 27 は、音声認識のための信号伝達経路の要素 23、24、25 を活性化させたり、不活性化させる。もし、ユーザーが音声認識のためのアプリケーションを選択すると、制御器 27 は、音声認識のための信号伝達経路の要素 23、24、25 を活性化させる。

【0045】

しかし、ユーザーが音声認識のためのアプリケーションを選択しない時は、制御器 27 は、音声認識のための信号伝達経路の要素 23、24、25 を不活性化させ、オーディオコーデック 22 の出力がボコーダ 28 に伝達されるように制御する。

【0046】

また、制御器 27 は、オーディオコーデック 22 のサンプリングレートを制御する。即ち、制御器 27 は、ユーザーがどんなアプリケーションを選択するかによって、オーディオコーデック 22 に入力される信号が音声通話のためのものであるか、音声認識のためのものであるかが分かるので、各状況によってオーディオコーデック 22 が適切なサンプリングレートを用いて PCM を実行するように制御する。

【0047】

制御器 27 の制御動作に対する一例を説明する。即ち、ユーザーが音声認識を用いる自動ダイアリング、メニュー選択、名前の呼出など、音声認識のためのアプリケーションを選択すると、制御器 27 は、ベースバンドモデムに備えられたレジスターのうち、音声認識モードを示す部分をオンにして、オーディオコーデック 22 のサンプリングレートを音声認識用サンプリングレート（例えば、16 kHz）に設定する。そして、バッファ 23 と、特徴ベクトル抽出ブロック 24 と、特徴ベクトルバッファ 25 で前記音声認識モードを示す部分をオン/オフにして制御する。

【0048】

要すると、制御器 27 は、オーディオコーデック 22 が用いるサンプリングレートの可変を制御する。そして、制御器 27 は、ユーザーが選択したアプリケーションによって、オーディオコーデック 22 の出力が伝達された経路を決定する。

【0049】

一方、音声認識のための信号伝達経路上で、特徴ベクトル抽出ブロック 24 の入力端に備えられたバッファ 23 は、音声認識のための音声信号（PCM データ）を保存する。前記バッファ 23 は、ピンポンバッファであることが好ましい。ここで、ピンポンバッファは、二重バッファリングの構造を用いる。即ち、二つの保存領域に区分して、その二つの保存領域のうち一方の領域にデータを満たすと同時に、他方の領域では、保存されていたデータを出力する構造である。

【0050】

一方、本発明では、上記の二重バッファリングの構造を用いたり、三つ以上の保存領域に区分して、その保存領域をリングの形態で構成した構造を用いる。ここで、前記バッファ 23 は、20～40 ms バッファである。

【0051】

特徴ベクトル抽出ブロック 24 は、入力端に位置したバッファ 23 から PCM データの伝達を受け、そこから特徴ベクトルを抽出する。前記特徴ベクトル抽出ブロック 24 は、MFCC (mel-frequency cepstral coefficients)、PLP (perceptual linear prediction)、LPC (Linear Predictive Coding)、又は LPCC (Linear Predictive Cepstral Coefficients) のうち、何れかの方式を用いる。

【0052】

そして、特徴ベクトル抽出ブロック 24 の出力端に備えられた特徴ベクトル 25 は、特徴ベクトル抽出ブロック 24 から抽出された特徴ベクトルを保存する。

【0053】

本発明では、20～40 ms 程度の短時間単位で反復的に特徴ベクトルを抽出し、その特徴ベクトルをアレイ形態で特徴ベクトルバッファ 25 に保存する。

10

20

30

40

50

【0054】

一般的に、特徴ベクトルを抽出する時は、フィルタバンク、フィルタリング、FFT (Fast Fourier Transform)、DCT (Discrete Cosine Transform)、IFFT (Inverse Fast Fourier Transform)などの過程を経なければならない。そのため、多くの演算量が要求される。また、特徴ベクトルを抽出する過程は規則性が強い。したがって、本発明では、特徴ベクトルの抽出のためのブロック24をハードウェアとして実現する。

【0055】

別途の例として、特徴ベクトル抽出のためにソフトウェア的な実現も可能である。

【0056】

音声認識ブロック26は、特徴ベクトルバッファ25に保存された特徴ベクトルを用いて、音声認識を実施する。前記音声認識ブロック26は、音声認識用アルゴリズムを備えたMPUやDSPであることが好ましい。

【0057】

音声認識アルゴリズム自体は変化可能性が多い。そして、トレーニングファイル及びパラメーターによって固定点の移行において差が生じ得る。そして、アルゴリズムの改善のためにビタビ復号化や言語モデリングや文法に当たる部分が用いられる。それにより、本発明では、音声認識アルゴリズムで固定点の移行やアルゴリズムの改善のための部分を、上述したMPU又はDSPを介して実現する。

【0058】

別途の例として、本発明では、MPU又はDSPを介した音声認識のために、予めノイズ除去を実施する。そのノイズの除去も上述したMPU又はDSPを介して行うことが好ましい。

【0059】

ボコーダ28は、音声通話のためのオーディオコーデック22の出力(8kHzのサンプリングレートを用いたPCMデータ)に対して音声コーディングを実施する。即ち、ボコーダ28は、音声通話のための音声信号が入力されると、そのPCMデータをQCELP (Qualcomm Code Excited Linear Prediction)やEVRC (Enhanced Variable Rate Coding)やVSELP (Vector Sum Excited Linear Prediction)やRPE-LTP (residual pulse excitation/long term prediction)などを用いて音声コーディングする。

【0060】

ボコーダ28の出力に対しては、コンボルーションコードやターボコードなどを用いるチャンネルコーディングを実施する。そして、チャンネルコーディングの以後に無線変調などの過程を更に経る。

【0061】

上述した前記ベースバンドモデムは、移動通信端末機の出荷時に内部の構成要素として予め装着させることが好ましい。しかし、本発明に係るベースバンドモデムを独立的なモジュールに実現して、移動通信端末機に後で装着するように構成してもよい。即ち、本発明による技術的な範囲は、前記両者の場合に共に及ぶことができる。

【0062】

以上で説明した内容を通じて当業者であれば本発明の技術思想を逸脱しない範囲で多様な変更及び修正が可能なが分かる。したがって、本発明の技術的な範囲は明細書の詳細な説明に記載された内容に限定されるものではなく、特許請求の範囲によって定められなければならない。

【0063】

以上のように、本発明の好ましい実施形態を用いて本発明を例示してきたが、本発明は、この実施形態に限定して解釈されるべきものではない。本発明は、特許請求の範囲によってのみその範囲が解釈されるべきであることが理解される。当業者は、本発明の具体的な好ましい実施形態の記載から、本発明の記載および技術常識に基づいて等価な範囲を実施することができることが理解される。

10

20

30

40

50

【0064】

音声認識に適したサンプリングレートで調節することによって音声認識率を高められると共に、音声認識のためのプロセッシングユニットの一部の処理部分をハードウェア的に実現できる音声認識用ベースバンドモデムと、それを用いる移動通信端末機を提供する。

【0065】

本発明に係るベースバンドモデムは、音声認識のためのサンプリングレートと、音声通話のためのサンプリングレートのうち何れか一つを用いて、入力されたボイス信号に対して変調を実施するオーディオコーデックと、前記音声認識のためのサンプリングレートを用いて、変調された前記オーディオコーデックの出力から特徴ベクトルを抽出する特徴ベクトル抽出ブロックと、前記特徴ベクトル抽出ブロックの出力を用いて、音声認識を行う音声認識ブロックと、前記音声通話のためのサンプリングレートを用いて、変調された前記オーディオコーデックの出力をボコーディングするボコーダとを含んで構成されることを特徴とする。

10

【図面の簡単な説明】

【0066】

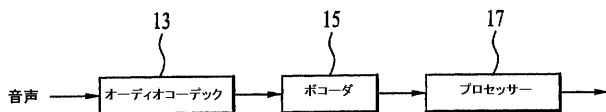
【図1】従来技術に係るベースバンドモデムの構成図である。

【図2】図1に示したベースバンドモデムを介した音声認識過程を説明した図面である。

【図3】本発明の一実施例によるベースバンドモデムの構成を示すブロックダイアグラムである。

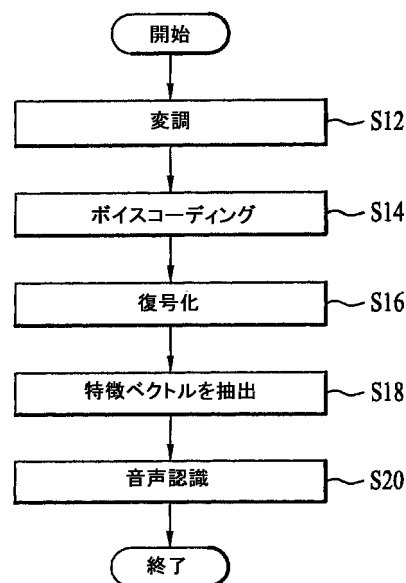
【図1】

FIG. 1



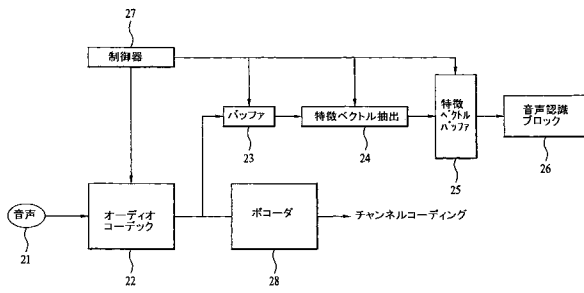
【図2】

FIG. 2



【図 3】

FIG. 3



フロントページの続き

(72)発明者 キム チャン ウ

大韓民国, キョンギード, コヤンーシ, イルサンーグ, テファードン, テファ メウル
エルジー ビレッジ, 104-902

Fターム(参考) 5D015 KK00