

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4599558号
(P4599558)

(45) 発行日 平成22年12月15日(2010.12.15)

(24) 登録日 平成22年10月8日(2010.10.8)

(51) Int.Cl. F I
G 1 O L 19/12 (2006.01) G 1 O L 19/12 Z
G 1 O L 21/04 (2006.01) G 1 O L 21/04 1 2 O D

請求項の数 19 (全 35 頁)

(21) 出願番号	特願2005-125815 (P2005-125815)	(73) 特許権者	504174135 国立大学法人九州工業大学 福岡県北九州市戸畑区仙水町1番1号
(22) 出願日	平成17年4月22日(2005.4.22)	(74) 代理人	100121371 弁理士 石田 和人
(65) 公開番号	特開2006-301464 (P2006-301464A)	(74) 代理人	100133592 弁理士 山口 浩一
(43) 公開日	平成18年11月2日(2006.11.2)	(72) 発明者	佐藤 寧 福岡県北九州市若松区ひびきの2-4 九州工業大学 ヒューマンライフIT開発センター内
審査請求日	平成19年7月20日(2007.7.20)	審査官	山下 剛史

最終頁に続く

(54) 【発明の名称】 ピッチ周期等化装置及びピッチ周期等化方法、並びに音声符号化装置、音声復号装置及び音声符号化方法

(57) 【特許請求の範囲】

【請求項1】

入力される音声信号に対して有声音のピッチ周期を等化するピッチ周期等化装置であって、

音声信号のピッチ周波数の検出を行うピッチ検出手段；

前記ピッチ周波数から所定の基準周波数を差し引いた差分である残差周波数を演算する残差演算手段；

及び、前記残差周波数に基づいて、前記音声信号のピッチ周波数を前記基準周波数に近づける方向にシフトさせることにより、前記音声信号のピッチ周期を等化する周波数シフタ；

を備え、

前記周波数シフタは、

前記入力信号を所定の変調波で振幅変調し被変調波を生成する変調手段；

前記被変調波の単側波帯成分の信号のみを選択的に通過させるバンドパスフィルタ；

前記バンドパスフィルタでフィルタリングした被変調波に対して所定の復調波で復調を行い、出力音声信号として出力する復調手段；

及び、前記変調手段が変調に用いる変調波の周波数及び前記復調手段が復調に用いる復調波の周波数の何れか一方を所定の基本キャリア周波数とし、他方を前記基本キャリア周波数から前記残差周波数を差し引いた周波数に設定する周波数調整手段；

を備えていることを特徴とするピッチ周期等化装置。

【請求項 2】

前記ピッチ検出手段は、

前記周波数シフタに入力される入力音声信号のピッチ周波数（以下「入力ピッチ周波数」という。）を検出する入力ピッチ検出手段；

及び前記周波数シフタから出力される出力音声信号のピッチ周波数（以下「出力ピッチ周波数」という。）を検出する出力ピッチ検出手段；

を備え、

前記入力ピッチ周波数の時間平均である平均ピッチ周波数を演算するピッチ平均手段を備え、

前記残差演算手段は、前記平均ピッチ周波数を基準周波数として、前記出力ピッチ周波数と当該基準周波数との差分である残差周波数を演算すること
を特徴とする請求項 1 記載のピッチ周期等化装置。 10

【請求項 3】

前記ピッチ検出手段は、前記周波数シフタに入力される入力音声信号のピッチ周波数（以下「入力ピッチ周波数」という。）を検出する入力ピッチ検出手段であり、

前記入力ピッチ周波数の時間平均である平均ピッチ周波数を演算するピッチ平均手段を備え、

前記残差演算手段は、前記平均ピッチ周波数を基準周波数として、前記入力ピッチ周波数と当該基準周波数との差分である残差周波数を演算すること
を特徴とする請求項 1 記載のピッチ周期等化装置。 20

【請求項 4】

前記ピッチ検出手段は、前記周波数シフタから出力される出力音声信号のピッチ周波数（以下「出力ピッチ周波数」という。）を検出する出力ピッチ検出手段であり、

前記出力ピッチ周波数の時間平均である平均ピッチ周波数を演算するピッチ平均手段を備え、

前記残差演算手段は、前記平均ピッチ周波数を基準周波数として、前記出力ピッチ周波数と当該基準周波数との差分である残差周波数を演算すること
を特徴とする請求項 1 記載のピッチ周期等化装置。

【請求項 5】

前記ピッチ検出手段は、前記周波数シフタに入力される入力音声信号のピッチ周波数（以下「入力ピッチ周波数」という。）を検出する入力ピッチ検出手段であり、 30

前記基準周波数を出力する基準周波数発生手段を備え、

前記残差演算手段は、前記入力ピッチ周波数と前記基準周波数との差分である残差周波数を演算すること

を特徴とする請求項 1 記載のピッチ周期等化装置。

【請求項 6】

前記ピッチ検出手段は、前記周波数シフタから出力される出力音声信号のピッチ周波数（以下「出力ピッチ周波数」という。）を検出する出力ピッチ検出手段であり、

前記基準周波数を出力する基準周波数発生手段を備え、

前記残差演算手段は、前記出力ピッチ周波数と前記基準周波数との差分である残差周波数を演算すること 40

を特徴とする請求項 1 記載のピッチ周期等化装置。

【請求項 7】

入力される音声信号を符号化する音声符号化装置であって、

前記音声信号に対して有声音のピッチ周期を等化する請求項 1 乃至 6 の何れか一記載のピッチ周期等化装置；

及び、前記ピッチ周期等化装置が出力する音声信号（以下「ピッチ等化音声信号」という。）に対して、一定のピッチ数区間で直交変換を行い、各サブバンドの変換係数データを生成する直交変換手段；

を備えた音声符号化装置。 50

【請求項 8】

前記ピッチ周期等化装置が出力する前記ピッチ等化音声信号に対して、1ピッチ区間のサンプリング数が一定となるようにリサンプリングを行うリサンプリング手段を備えていることを特徴とする請求項 7 記載の音声符号化装置。

【請求項 9】

原音声信号に対しピッチ周波数が所定の基準周波数に等化され、直交変換によりサブバンド成分に分解されたピッチ等化音声信号、及び前記原音声信号のピッチ周波数から前記基準周波数を差し引いた差分である残差周波数信号に基づいて前記原音声信号を復号する音声復号装置であって、

一定のピッチ数区間で直交変換されたピッチ等化音声信号に対し逆直交変換を行うことによりピッチ等化音声信号を復元する逆直交変換手段；

及び、前記ピッチ等化音声信号のピッチ周波数を前記基準周波数に前記残差周波数を加えた周波数に近づける方向にシフトさせることにより、前記復元音声信号を生成する周波数シフタ；

を備え、

前記周波数シフタは、

前記ピッチ等化音声信号を所定の變調波で振幅變調し被變調波を生成する變調手段；

前記被變調波の単側波帯成分の信号のみを選択的に通過させるバンドパスフィルタ；

前記バンドパスフィルタでフィルタリングした被變調波に対して所定の復調波で復調を行い、復元音声信号として出力する復調手段；

及び、前記變調手段が變調に用いる變調波の周波数及び前記復調手段が復調に用いる復調波の周波数の何れか一方を所定の基本キャリア周波数とし、他方を前記基本キャリア周波数に前記残差周波数を加えた値に設定する周波数調整手段；

を備えていることを特徴とする音声復号装置。

【請求項 10】

入力される音声信号（以下「入力音声信号」という。）に対して有声音のピッチ周期を等化するピッチ周期等化方法であって、

前記入力音声信号を周波数シフタに入力し、前記周波数シフタからの出力信号（以下「出力音声信号」という。）を得る周波数シフトステップ；

前記出力音声信号のピッチ周波数（以下「出力ピッチ周波数」という。）を検出する出力ピッチ検出ステップ；

前記出力ピッチ周波数から所定の基準周波数を差し引いた差分である残差周波数を演算する残差周波数演算ステップ；

及び、前記出力ピッチ周波数と所定の基準周波数との差分である残差周波数を演算する残差周波数演算ステップ；

を有し、

前記周波数シフトステップにおいては、

變調に用いる變調波の周波数及び復調に用いる復調波の周波数の何れか一方を所定の基本キャリア周波数とし、他方を前記基本キャリア周波数から前記残差周波数演算ステップにおいて算出される前記残差周波数を差し引いた周波数に設定する周波数設定ステップ；

前記入力音声信号を前記變調波で振幅變調し被變調波を生成する變調ステップ；

前記被變調波の単側波帯成分のみを通過させるバンドパスフィルタにより、前記被變調波をフィルタリングする帯域縮小ステップ；

前記バンドパスフィルタでフィルタリングした被變調波に対して前記復調波で復調を行い、出力音声信号として出力する復調ステップ；

を有することを特徴とするピッチ周期等化方法。

【請求項 11】

前記出力ピッチ周波数の時間平均である平均ピッチ周波数を演算するピッチ平均化ステップ；

を有し、

10

20

30

40

50

前記残差周波数演算ステップにおいては、前記出力ピッチ周波数と前記平均ピッチ周波数との差分を演算し、これを前記残差周波数とすることを特徴とする請求項 10 記載のピッチ周期等化方法。

【請求項 12】

前記入力音声信号のピッチ周波数（以下「入力ピッチ周波数」という。）を検出する入力ピッチ検出ステップ；

前記入力ピッチ周波数の時間平均である平均ピッチ周波数を演算するピッチ平均化ステップ；
を有し、

前記残差周波数演算ステップにおいては、前記出力ピッチ周波数と前記平均ピッチ周波数との差分を演算し、これを前記残差周波数とすることを特徴とする請求項 10 記載のピッチ周期等化方法。

【請求項 13】

入力される音声信号（以下「入力音声信号」という。）に対して有声音のピッチ周期を等化するピッチ周期等化方法であって、

前記入力音声信号のピッチ周波数（以下「入力ピッチ周波数」という。）を検出する入力ピッチ検出ステップ；

前記入力音声信号を周波数シフトに入力し、前記周波数シフトからの出力信号（以下「出力音声信号」という。）を得る周波数シフトステップ；

及び、前記入力ピッチ周波数から所定の基準周波数を差し引いた差分である残差周波数を演算する残差周波数演算ステップ；

を有し、

前記周波数シフトステップにおいては、

変調に用いる変調波の周波数及び復調に用いる復調波の周波数の何れか一方を所定の基本キャリア周波数とし、他方を前記基本キャリア周波数から前記残差周波数演算ステップにおいて算出される前記残差周波数を差し引いた周波数に設定する周波数設定ステップ；

前記入力音声信号を前記変調波で振幅変調し被変調波を生成する変調ステップ；

前記被変調波の単側波帯成分のみを通過させるバンドパスフィルタにより、前記被変調波をフィルタリングする帯域縮小ステップ；

前記バンドパスフィルタでフィルタリングした被変調波に対して前記復調波で復調を行い、出力音声信号として出力する復調ステップ；

を有することを特徴とするピッチ周期等化方法。

【請求項 14】

前記入力ピッチ周波数の時間平均である平均ピッチ周波数を演算するピッチ平均化ステップ；

を有し、

前記残差周波数演算ステップにおいては、前記入力ピッチ周波数と前記平均ピッチ周波数との差分を演算し、これを前記残差周波数とすることを特徴とする請求項 13 記載のピッチ周期等化方法。

【請求項 15】

入力される音声信号を符号化する音声符号化方法であって、

請求項 10 乃至 14 の何れか一記載のピッチ周期等化方法により、前記音声信号に対して有声音のピッチ周期を等化するピッチ周期等化ステップ；

前記ピッチ周期等化ステップで等化された音声信号（以下「ピッチ等化音声信号」という。）に対して、一定のピッチ数区間で直交変換を行い、各サブバンドの変換係数データを生成する直交変換ステップ；

及び、前記変換係数データを符号化する波形符号化ステップ；

を備えた音声符号化方法。

【請求項 16】

前記ピッチ周期等化ステップにおいて等化された前記ピッチ等化音声信号に対して、1ピ

10

20

30

40

50

ッチ区間のサンプリング数が一定となるようにリサンプリングを行うリサンプリングステップ

を備えていることを特徴とする請求項 1 4 記載の音声符号化方法。

【請求項 1 7】

コンピュータで実行することにより、前記コンピュータを請求項 1 乃至 6 の何れか一記載のピッチ周期等化装置として機能させるプログラム。

【請求項 1 8】

コンピュータで実行することにより、前記コンピュータを請求項 7 又は 8 記載の音声符号化装置として機能させるプログラム。

【請求項 1 9】

コンピュータで実行することにより、前記コンピュータを請求項 9 記載の音声復号装置として機能させるプログラム。

10

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、ピッチ成分を含む音声信号のピッチ周期を等化するピッチ周期等化技術、及びそれを使用した音声符号化技術に関する。

【背景技術】

20

【0002】

音声符号化の分野において、10 k b p s 以下の低ビットレートでは、現在のところ、符号励振線形予測 (Code Excited Linear Prediction Coding Encoding: 以下「C E L P」という。) 符号化方式が広く用いられている (非特許文献 1 参照)。C E L P 符号化方式は、人間の音声発生機構を音源成分 (声帯) とスペクトル包絡成分 (声道) とによってモデル化し、それぞれのパラメータを符号化する方式である。

【0003】

符号化側においては、音声はフレームと呼ばれる単位に分割され、各フレームに対して、符号化が行われる。スペクトル包絡成分は、線形予測に基づく音声の A R モデル (Auto-Regressive model: 自己回帰モデル) に基づいて計算され、線形予測 (Linear Prediction Coding: 以下「L P C」という。) 係数という形で与えられる。また、音源成分は、予測残差という形で与えられる。この予測残差は、ピッチ情報を表す周期情報、音源情報である雑音情報、及びピッチと音源の混合比を表す利得情報に分離される。これら各情報は、符号帳に蓄えられた符号ベクトルにより構成される。符号ベクトルの決定は、各符号ベクトルをフィルタに通して音声を合成し、入力波形に最も近いものを探索する方法、いわゆる A b S (Analysis by Synthesis) 法による閉ループ探索によって行われる。

30

【0004】

また、復号側においては、符号化された各情報を復号し、L P C 係数、周期情報 (ピッチ情報)、雑音源情報、及び利得情報を復元する。雑音情報にピッチ情報を加えることにより励振源信号を生成する。この励振源信号を L P C 係数で構成される線形予測合成フィルタに通すことにより、合成音声を得る。

40

【0005】

図 1 6 は C E L P 符号化方式による音声符号化装置の基本構成例を表す (特許文献 1, 図 9 参照)。

【0006】

原音声信号は、所定のサンプル数のフレーム単位に分割され、入力端子 1 0 1 に入力される。入力端子 1 0 1 に入力された原音声信号は、線形予測分析部 1 0 2 において、周波数スペクトル包絡特性を表す L P C 係数が計算される。具体的には、フレームの自己相関関数を求め、Durbinの再帰解法などを用いて L P C 係数が計算される。

【0007】

50

LPC係数符号化部103は、このLPC係数を量子化し符号化することにより、LPC係数符号を生成する。この量子化は、量子化効率の優れた線スペクトル対(Line Spectrum Pair: LSP)パラメータ、偏自己相関方式(Partial auto-Correlation: PARCOR)パラメータ、反射係数等に変換して行われることが多い。LPC係数復号部104は、LPC係数符号を復号してLPC係数を再生する。この再生されたLPC係数に基づき、各フレームの予測残差成分(音源成分)の符号化のための符号帳探索が行われる。この符号帳探索は、フレームを更に分割した単位(以下「サブフレーム」という。)に対して行われることが多い。

【0008】

ここで、符号帳は、適応符号帳105、雑音符号帳106、及び利得符号帳107から構成されている。

10

【0009】

適応符号帳105は、ピッチ周期とピッチパルスの振幅をピッチ周期ベクトルとして表し記憶した符号帳であり、音声のピッチ成分を表現する。ピッチ周期ベクトルは、以前のフレームまでの残差成分(既に量子化された直前の1~数フレーム分の駆動音源ベクトル)を、予め設定された周期だけ繰り返すことにより構成されたサブフレーム長のベクトルである。適応符号帳105は、かかるピッチ周期ベクトル群を記憶している。適応符号帳105は、これらピッチ周期ベクトル群の中から、音声の周期成分に対応して一つのピッチ周期ベクトルを選択し、時系列符号ベクトルの候補として出力する。

【0010】

20

雑音符号帳106は、残差信号からピッチ成分を除いた残りの波形である形状励振源成分を励振ベクトルとして表し記憶した符号帳であり、ピッチ以外の雑音的な成分(非周期的励振)を表現する。励振ベクトルは、入力音声とは独立に、白色雑音を基調として用意されたサブフレーム長のベクトルである。雑音符号帳106には、このような励振ベクトルが予め指定された数だけ記憶されている。雑音符号帳106は、これらピッチ励振ベクトル群の中から、音声の雑音成分に対応して一つの励振ベクトルを選択し、音声の非周期成分に対応する時系列符号ベクトルの候補として出力する。

【0011】

また、利得符号帳107は、音声のピッチ成分及びそれ以外の成分の利得を表現する。

【0012】

30

適応符号帳105及び雑音符号帳106から出力された各時系列符号ベクトルの候補は、それぞれ、利得部108、109においてピッチ利得 g_a 、形状利得 g_r が乗算される。利得 g_a 、 g_r は、利得符号帳107において選択され出力される。そして、両者は加算部110において加算され、駆動音源ベクトルの候補が生成される。

【0013】

合成フィルタ111は、LPC係数復号部104が出力するLPC係数をフィルタ係数とする線形フィルタである。合成フィルタ111は、加算部110から出力される駆動音源ベクトルの候補をフィルタリングして、再生音声候補ベクトルとして出力する。

【0014】

比較部112は、原音声信号ベクトルから上記再生音声候補ベクトルを減算し歪データを出力する。この歪データは、聴覚重み付けフィルタ113において、人間の聴覚の特性に対応した係数によって重み付けがされる。この聴覚重み付けフィルタ113は、通常、移動平均自己回帰型の10次程度のフィルタであり、フォルマントの山の部分をやや強調するように構成されている。この重み付けは、音声スペクトルの包絡の値が小さくなる谷の部分の周波数帯域では量子化雑音小さくなるように符号化を行うために行われる。

40

【0015】

距離最小化部114は、聴覚重み付けフィルタ113から出力された歪データの二乗誤差が最小となるような周期信号、雑音符号、及び利得符号を選択する。周期信号、雑音符号、及び利得符号は、それぞれ、適応符号帳105、雑音符号帳106、及び利得符号帳107に送られる。適応符号帳105は、入力される周期信号に基づいて、次の時系列符

50

号ベクトルの候補を出力する。雑音符号帳 106 は、入力される雑音符号に基づいて次の時系列符号ベクトルの候補を出力する。また、利得符号帳 107 は、入力される利得符号に基づいて、次の利得 g_a , g_r を出力する。

【0016】

距離最小化部 114 は、このような A b S ループを繰り返すことにより、聴覚重み付けフィルタ 113 が出力する歪データが最小化された時点で、上記周期信号、雑音符号、及び利得符号を、そのフレームにおける駆動音源ベクトルとして決定する。

【0017】

符号送部 115 は、距離最小化部 114 が決定する周期信号、雑音符号、及び利得符号と、LPC 係数符号化部 103 が出力する LPC 係数符号を、ビット系列の符号に変換し、さらに必要に応じて訂正符号を付加して出力する。

10

【0018】

図 17 は C E L P 符号化方式による音声復号装置の基本構成例を表す（特許文献 1 , 図 11 参照）。

【0019】

音声復号装置は、符号帳の検索を行わない点を除いて、音声符号化装置とほぼ同一の構成となる。符号受信部 121 は、LPC 係数符号、周期符号、雑音符号、及び利得符号を受信する。LPC 係数符号は、LPC 係数復号部 122 に送られる。LPC 係数復号部 122 は、LPC 係数符号を復号し LPC 係数（フィルタ係数）を生成する。

20

【0020】

適応符号帳 123 は、ピッチ周期ベクトル群を記憶している。ピッチ周期ベクトルは、以前のフレームまでの残差成分（既に復号された直前の 1 ~ 数フレーム分の駆動音源ベクトル）を、予め設定された周期だけ繰り返すことにより構成されたサブフレーム長のベクトルである。適応符号帳 123 は、符号受信部 121 から入力される周期符号に対応して一つのピッチ周期ベクトルを選択し、時系列符号ベクトルとして出力する。

【0021】

雑音符号帳 124 は、励振ベクトル群を記憶している。励振ベクトルは、入力音声とは独立に、白色雑音を基調として用意されたサブフレーム長のベクトルである。符号受信部 121 から入力される雑音符号に対応して一つの励振ベクトルを選択し、音声の非周期成分に対応する時系列符号ベクトルとして出力する。

30

【0022】

また、利得符号帳 125 は、音声のピッチ成分及びそれ以外の成分の利得（ピッチ利得 g_a , 形状利得 g_r ）群を記憶する。利得符号帳 125 は、符号受信部 121 から入力される利得符号に対応して一組のピッチ利得 g_a , 形状利得 g_r を選択して出力する。

【0023】

適応符号帳 123 及び雑音符号帳 124 から出力される時系列符号ベクトルは、それぞれ、利得部 126 , 127 においてピッチ利得 g_a , 形状利得 g_r が乗算される。そして、両者は加算部 128 において加算され、駆動音源ベクトルが生成される。

【0024】

合成フィルタ 129 は、LPC 係数復号部 122 が出力する LPC 係数をフィルタ係数とする線形フィルタである。合成フィルタ 129 は、加算部 128 から出力される駆動音源ベクトルの候補をフィルタリングして、再生音声として端子 130 に出力する。

40

【0025】

一方、MPEG 規格やオーディオ機器では、サブバンド符号化方式が多く用いられている。サブバンド符号化方式においては、音声信号を複数の周波数帯域（サブバンド）に分割し、各サブバンド内での信号エネルギーに応じたビット割り当てを行うことにより効率のよい符号化が行われる。サブバンド符号化方式を音声符号化に適用した技術としては、特許文献 2 ~ 4 に記載の技術が公知である。

【0026】

特許文献 2 ~ 4 に記載の音声符号化方式においては、基本的に次のような信号処理によ

50

って音声信号の符号化を行う。

【0027】

まず、入力された原音声信号からピッチを抽出する。そして、原音声信号をピッチ区間に分割する。次に、分割により得られた各ピッチ区間の音声信号について、それぞれのピッチ区間の標本化数が一定数となるようにリサンプリングを行う。そして、リサンプリングされた各ピッチ区間の音声信号に対し、DCT等の直交変換を施すことにより、 $(n+1)$ 個のデータから成るサブバンドデータを生成する。最後に、時系列的に得られる $(n+1)$ 個のデータのそれぞれに対して、フィルタリングを行うことにより、強度の時間変化のうち所定の周波数を超える成分を除去して平滑化し、 $(n+1)$ 個の音響情報データを生成する。また、サブバンドデータから高周波成分の割合を閾値判定することによって、原音声信号が摩擦音か否かを判定し、その判定結果を摩擦音情報として出力する。

10

【0028】

最終的に、原音声信号は、各ピッチ区間の元のピッチ長を表す情報(ピッチ情報)、 $(n+1)$ 個の音響情報データからなる音響情報、及び摩擦音情報に分割され符号化される。

【0029】

図18は、特許文献2記載の音声符号化装置(音声信号加工装置)の構成例を表す図である。原音声信号(音声データ)は、音声データ入力部141に入力される。ピッチ抽出部142は、音声データ入力部141に入力された音声データからピッチの基本周波数の信号(ピッチ信号)を抽出し、音声データをピッチ信号の単位周期(単位ピッチ区間)で区切る。そして、それぞれの単位ピッチ区間の音声データを、ピッチ信号との相関が最大となるように移相して調整し、ピッチ長固定部143に出力する。

20

【0030】

ピッチ長固定部143は、この各単位ピッチ区間の音声データを、各単位ピッチ区間の標本化数がほぼ等しくなるようにリサンプリングを行う。そして、リサンプリングされた単位ピッチ区間の音声データを、ピッチ波形データとして出力する。尚、このリサンプリングによって各単位ピッチ区間の長さ(ピッチ周期)に関する情報が除去されるので、ピッチ長固定部143は、各単位ピッチ区間における元のピッチ長を表す情報をピッチ情報として出力する。

【0031】

サブバンド分割部144は、このピッチ波形データにDCT等の直交変換を施してサブバンドデータを生成する。このサブバンドデータは、音声の基本周波数成分、及びこの音声の n 個の高調波成分の強度を表す $(n+1)$ 個のスペクトル強度データの時系列データからなる。

30

【0032】

帯域情報制限部145は、上記サブバンドデータを構成する $(n+1)$ 個のスペクトル強度データをそれぞれフィルタリングすることにより、 $(n+1)$ 個のスペクトル強度データの時間変化のうち、所定の周波数を超える成分を除去する。これは、ピッチ長固定部143におけるリサンプリングによって発生するエリアシングの影響を除去するために行われる処理である。

40

【0033】

帯域情報制限部145でフィルタリングされたサブバンドデータは、非線形量子化部146において非線形量子化され、辞書選択部147でコード化されて音響情報として出力される。

【0034】

一方、摩擦音検出部149は、サブバンドデータの全体のスペクトル強度に占める高周波成分の割合に基づき、入力された音声データが有声音か無声音(摩擦音)かを判別する。そして、この判別結果を、摩擦音情報として出力する。

【0035】

このように、原音声信号をサブバンドに分割する前にピッチの揺らぎを除去し、ピッチ

50

区間ごとに直交変換を行うことによってサブバンドに分割する。これにより、各サブバンドのスペクトル強度の時間変化が小さくなるため、音響情報に関して高い圧縮率が実現できる。

【0036】

【特許文献1】特許3199128号公報

【特許文献2】特開2003-108172号公報

【特許文献3】特開2003-108200号公報

【特許文献4】特開2004-12908号公報

【非特許文献1】Manfred R. Schroeder and Bishnu S. Atal, "Code-excited Linear Prediction (CELP): High-Quality Speech at Very Low Bit Rates", Proceedings of ICAS SP 85, pp. 25.1.1-25.1.4, 1985.

【非特許文献2】貴家仁志, 「デジタル信号処理シリーズ(第14巻)マルチレート信号処理」, 初版, 1995年10月6日, pp. 34-49, 78-79.

【発明の開示】

【発明が解決しようとする課題】

【0037】

上記従来のCELP符号化方式においては、残差信号のうちのピッチ成分は適応符号帳に用意されたピッチ周期ベクトル群の中から選択される。また、残差信号のうちの音源成分は雑音符号帳に用意された固定的な励振ベクトル群の中から選択される。そのため、入力音声を忠実に再現しようとした場合、適応符号帳のピッチ周期ベクトル群及び雑音符号帳の励振ベクトル群の中の候補数をできるだけ多く用意する必要がある。

【0038】

しかしながら、これら候補数を多くすると、適応符号帳及び雑音符号帳のメモリ容量が莫大となり、実装面積が増大する。また、候補数をあまり多くすると、候補数の対数に比例して周期符号や雑音符号の符号量も大きくなる。そのため、低ビットレートを実現するためには適応符号帳及び雑音符号帳の候補数をあまり多くすることはできない。

【0039】

したがって、限られた数のピッチ周期ベクトル及び励振ベクトルの中から候補を選択し、入力音声の音源成分を近似することとなり、歪みは或る程度以上は小さくできない。特に、音声信号において音源成分はかなりの割合を占める成分である半面、雑音的であるため予測が困難である。したがって、再生音声に或る程度の大きい歪みが生じ、更なる高音質化に限界がある。

【0040】

一方、特許文献2~4に記載の音声符号化方式では、音声信号をサブバンド符号方式により符号化するため、高音質で高圧縮率の符号化が可能である。

【0041】

しかしながら、この方式では、ピッチ長固定部において音声信号をリサンプリング(通常、ダウンサンプリング)する際にエリアシングの問題及びピッチの揺らぎにより音声信号が変調を受けるといった問題がある。

【0042】

前者は、ダウンサンプリングによりエリアシング成分が生じる現象であり、これは、通常のデシメータと同様にデシメーション・フィルタを使用することにより回避することが可能である(例えば、非特許文献2参照)。

【0043】

一方、後者は周期が揺らいだ信号をピッチ区間ごとに一定の標本化数に合わせ込むことによって、揺らぎにより音声信号が変調されることによって生じる。すなわち、ピッチ長固定部143は、ピッチ区間ごとに周期が揺らいでいる音声データを、それぞれのピッチ区間の標本化数が一定となるようにリサンプリングを行う。この場合、通常、ピッチの揺らぎの周期はピッチ周期に比べると1/10程度であり、かなり長い。従って、このようにピッチ周期が揺らいでいる音声信号を、それぞれのピッチ区間が同一の標本化数となる

10

20

30

40

50

ようにむりやりリサンプリングすると、ピッチの揺らぎの周波数によって周波数変調された状態となる。従って、ピッチの揺らぎ周波数によって周波数変調された音響情報から、音声信号を再び復元した際に、このピッチの揺らぎによる変調成分（以下「ピッチ揺らぎ変調成分」という。）がゴーストトーンとなって現れ、音声に歪みが生じる。

【0044】

この現象を防止するために、特許文献2, 3に記載の音声符号化装置においては、サブバンド分割部144が出力する各サブバンド成分のスペクトル強度データを、帯域情報制限部145によりフィルタリングすることによりスペクトル強度データの時間変化として現れるピッチ揺らぎ変調成分を除去しようとしている。

【0045】

しかしながら、帯域情報制限部145において通過帯域をあまり狭くしすぎると、ピッチ揺らぎ変調成分以外の原音声信号本来の時間変化成分まで平滑化されてしまうため、かえって音声信号の歪みを生じさせる結果となる。一方、帯域情報制限部145における通過帯域を広くすると、ピッチ揺らぎ変調成分が通過するためゴーストトーンが現れる。

【0046】

また、特許文献4に記載の音声符号化装置においては、サブバンド分割部144が出力する各サブバンド成分のスペクトル強度データを平均化することによってピッチ揺らぎ変調成分を除去しようとしている。しかし、この平均化により、ピッチ揺らぎ変調成分以外の原音声信号本来の時間変化成分まで失われるため、結局、音声信号の歪みを生じさせる結果となる。

【0047】

従って、特許文献2～4に記載の音声符号化方式では、ピッチ揺らぎ変調成分を除去することは困難であり、この変調成分による音声信号の歪みが避けられないという問題がある。

【0048】

そこで、本発明の目的は、低ビットレートを実現するとともに、ピッチの揺らぎによる周波数変調のような歪みを生じることなく、従来よりも再生音声の歪みを小さく抑えることを可能とする音声符号化技術、及びそれに使用するのに適したピッチ周期等化技術を提供することにある。

【課題を解決するための手段】

【0049】

ピッチ成分を含む音声信号は、同一の音素内においては、隣接するピッチ区間の波形は比較的類似している。従って、それぞれのピッチ区間において又は一定の数のピッチ区間において変換符号化を行うことができれば、隣り合うピッチ区間でのスペクトルは類似し、冗長性の大きいスペクトルの時間系列が得られると考えられる。そして、これを符号化すれば符号化効率は向上すると考えられる。この場合、符号帳は使用する必要がない。また、原音声の波形をそのまま符号化するため、歪みの少ない再生音声を得ることも可能である。

【0050】

しかしながら、原音声信号の各ピッチ周波数は男女差、個人差、音素、感情及び会話内容によって様々である。また、同一の音素においても、それぞれのピッチ周期は揺らいだり変化したりする。従って、そのままでは各ピッチ区間で変換符号化を行ったとしても、得られるスペクトル列は時間的変化が大きく、高い符号化効率は期待できない。

【0051】

そこで、本発明の音声符号化方法においては、ピッチ成分を含む原音声に含まれる情報を、ピッチの基本周波数の情報、ピッチ周期の揺らぎの情報、及び各ピッチ区間内における波形の情報に分離する方法を採用する。ピッチの基本周波数の情報及びピッチ周期の揺らぎ情報が除去された原音声信号は、ピッチ周期は一定となり、各ピッチ区間又は一定数のピッチ区間における変換符号化は容易である。そして、隣接するピッチ区間の波形の相関は大きいいため、変換符号化によって得られるスペクトルは、等化されたピッチ周波数及

10

20

30

40

50

びその高調波成分に集約され高い符号化効率を得ることが期待できる。

【 0 0 5 2 】

上記本発明の音声符号化方法において、原音声信号からピッチの基本周波数の情報及びピッチ周期の揺らぎの情報を抽出し除去するのに、ピッチ周期等化技術が用いられる。そこで、以下、本発明に係るピッチ周期等化装置及び方法、並びに音声符号化装置及び方法の構成及び作用について説明する。

【 0 0 5 3 】

〔本発明の構成及び作用〕

本発明に係るピッチ周期等化装置の第1の構成は、入力される音声信号に対して有声音のピッチ周期を等化するピッチ周期等化装置において、音声信号のピッチ周波数の検出を行うピッチ検出手段；前記ピッチ周波数から所定の基準周波数を差し引いた差分である残差周波数を演算する残差演算手段；及び、前記残差周波数に基づいて、前記音声信号のピッチ周波数を前記基準周波数に近づける方向にシフトさせることにより、前記音声信号のピッチ周期を等化する周波数シフト；を備え、前記周波数シフトは、前記入力信号を所定の變調波で振幅變調し被變調波を生成する變調手段；前記被變調波の単側波帯成分の信号のみを選択的に通過させるバンドパスフィルタ；前記バンドパスフィルタでフィルタリングした被變調波に対して所定の復調波で復調を行い、出力音声信号として出力する復調手段；及び、前記變調手段が變調に用いる變調波の周波数及び前記復調手段が復調に用いる復調波の周波数の何れか一方を所定の基本キャリア周波数とし、他方を前記基本キャリア周波数から前記残差周波数を差し引いた値に設定する周波数調整手段；を備えていることを特徴とする。

【 0 0 5 4 】

この構成によれば、音声信号のピッチ周期を基準周期（基準周波数の逆数）に等化する際に、一旦、入力された音声信号を變調波で振幅變調し、被變調波をバンドパスフィルタに通して下側波帯を除去する。そして、復調波を用いて単側波帯の被變調波を復調する。この際、残差周波数が0の場合は變調波及び復調波はともに基本キャリア周波数とされるが、残差周波数が0でない場合には、變調波又は復調波の何れか一方は、周波数調整手段によって基本キャリア周波数から残差周波数を差し引いた値に設定される。これにより、入力された音声信号のピッチの基本周波数と基準周波数の差分がキャンセルされ、出力音声信号のピッチ周期は基準周期に等化される。

【 0 0 5 5 】

このように、ピッチ周期を所定の基準周期に等化することにより、音声信号に含まれるピッチの男女差、個人差、音素、感情及び会話内容によって変化するピッチ周波数のジッタ成分や変化成分が除去される。

【 0 0 5 6 】

また、音声信号のピッチ周期を基準周期に等化する際に、単側波帯變調を用いるため、エリアシングの問題は生じない。また、ピッチ周期の等化を行う際にリサンプリングを用いないため、上記従来方式（特許文献2～4）のようにピッチの揺らぎにより音声信号が變調を受けるという問題を生じない。従って、ピッチ周期が等化された出力音声信号は、等化による歪みが生じることがない。

【 0 0 5 7 】

また、入力音声信号に含まれる情報は、ピッチの基準周波数の情報、ピッチ毎のピッチ周波数の變動に関する情報、及びピッチに重畳する波形成分の情報に分離される。これらはそれぞれ、基準周波数、残差周波数、及び等化後の音声信号の1ピッチ区間内の波形として得られる。基準周波数は、音素毎にほぼ一定であるため、符号化を行う場合の符号化効率が高い。また、一般的に各音素内ではピッチ周波数の揺らぎの變動幅は小さいため、ビン周波数（bin-frequency）は狭レンジであり、残差周波数は符号化を行う場合の符号化効率が高い。また、等化後の音声信号の1ピッチ区間内の波形は、ピッチの揺らぎ成分が除去されているため、各ピッチ区間で同一の標本化数となる。また、同一音素内の各ピッチ区間の波形は相似性が強いため、各ピッチ区間で同一の標本化数に等化することによ

って各ピッチ区間における波形は類似性の高いものとなる。したがって、一乃至所定数のピッチ区間で変換符号化することによって、符号量を大幅に圧縮することが可能である。したがって、音声信号の符号化効率を向上させることが可能となる。

【0058】

尚、上記本発明の構成は、音声信号の中で、ピッチを含む有声音に対してピッチ周期を等化するものである。したがって、ピッチを含まない無声音やノイズについては、別途、公知のケプストラム解析やスペクトル形状の特徴分析を用いた方法により分別するようによればよい。

【0059】

また、このピッチ周期等化装置は、音声符号化以外にも音声検索等の音声マッチング技術に適用することも可能である。すなわち、各ピッチ区間を同一周期に等化することによって各ピッチ区間における波形は類似性の高いものとなり、音声信号の対比が容易となるため、音声検索に応用した場合には、音声のマッチング精度を向上させることが可能となる。

【0060】

本発明に係るピッチ周期等化装置の第2の構成は、前記第1の構成において、前記ピッチ検出手段は、前記周波数シフタに入力される入力音声信号のピッチ周波数（以下「入力ピッチ周波数」という。）を検出する入力ピッチ検出手段；及び前記周波数シフタから出力される出力音声信号のピッチ周波数（以下「出力ピッチ周波数」という。）を検出する出力ピッチ検出手段；を備え、前記入力ピッチ周波数の時間平均である平均ピッチ周波数を演算するピッチ平均手段を備え、前記残差演算手段は、前記平均ピッチ周波数を基準周波数として、前記出力ピッチ周波数と当該基準周波数との差分である残差周波数を演算することを特徴とする。

【0061】

この構成によれば、基準周波数として入力ピッチ周波数の時間平均を使用することにより、各音素内のピッチ周波数に男女差、個人差、音素による差、感情、又は会話内容による差があっても、それらの差異に適応して最適な周波数を基準周波数とすることが可能となる。

【0062】

また、出力ピッチ周波数と基準周波数との差分を残差周波数として、この周波数を周波数シフタのシフト量にフィードバックすることによって、周波数シフタによるピッチ周期の等化誤差を低減させ、ピッチ毎のピッチ周波数の変動に関する情報と、ピッチに重畳する波形成分の情報との分離を効率的に行うことが可能となる。

【0063】

ここで、ピッチ平均手段による時間平均は、単純な幾何学平均、加重平均等を問わない。また、ピッチ平均手段としてローパスフィルタを使用することができる。この場合、ピッチ平均手段による時間平均は加重平均となる。

【0064】

本発明に係るピッチ周期等化装置の第3の構成は、前記第1の構成において、前記ピッチ検出手段は、前記周波数シフタに入力される入力音声信号のピッチ周波数（以下「入力ピッチ周波数」という。）を検出する入力ピッチ検出手段であり、前記入力ピッチ周波数の時間平均である平均ピッチ周波数を演算するピッチ平均手段を備え、前記残差演算手段は、前記平均ピッチ周波数を基準周波数として、前記入力ピッチ周波数と当該基準周波数との差分である残差周波数を演算することを特徴とする。

【0065】

このように、基準周波数として入力ピッチ周波数の時間平均を使用することにより、上述のように最適な周波数を基準周波数とすることが可能となる。

【0066】

また、入力ピッチ周波数と基準周波数との差分を残差周波数として、この周波数を周波数シフタのシフト量にフィードフォワードすることによって、周波数シフタによるピッ

10

20

30

40

50

子周期の等化誤差を低減させ、ピッチ毎のピッチ周波数の変動に関する情報と、ピッチに重畳する波形成分の情報との分離を効率的に行うことが可能となる。

【0067】

本発明に係るピッチ周期等化装置の第4の構成は、前記第1の構成において、前記ピッチ検出手段は、前記周波数シフトから出力される出力音声信号のピッチ周波数（以下「出力ピッチ周波数」という。）を検出する出力ピッチ検出手段であり、前記出力ピッチ周波数の時間平均である平均ピッチ周波数を演算するピッチ平均手段を備え、前記残差演算手段は、前記平均ピッチ周波数を基準周波数として、前記出力ピッチ周波数と当該基準周波数との差分である残差周波数を演算することを特徴とする。

【0068】

このように、基準周波数として出力ピッチ周波数の時間平均を使用することにより、上述のように最適な周波数を基準周波数とすることが可能となる。

【0069】

また、入力ピッチ周波数と基準周波数との差分を残差周波数として、この周波数を周波数シフトのシフト量にフィードバックすることによって、周波数シフトによるピッチ周期の等化誤差を低減させ、ピッチ毎のピッチ周波数の変動に関する情報と、ピッチに重畳する波形成分の情報との分離を効率的に行うことが可能となる。

【0070】

本発明に係るピッチ周期等化装置の第5の構成は、前記第1の構成において、前記ピッチ検出手段は、前記周波数シフトに入力される入力音声信号のピッチ周波数（以下「入力ピッチ周波数」という。）を検出する入力ピッチ検出手段であり、前記基準周波数を出力する基準周波数発生手段を備え、前記残差演算手段は、前記入力ピッチ周波数と前記基準周波数との差分である残差周波数を演算することを特徴とする。

【0071】

このように、基準周波数として、基準周波数発生手段が出力する決められた周波数を用いることによって、入力音声信号に含まれる音声の情報のうち、ピッチの基本周波数の情報及びピッチ毎のピッチ周波数の変動に関する情報は残差周波数として分離される。また、ピッチに重畳する波形成分の情報は、等化後の音声信号の1ピッチ区間内の波形として分離される。

【0072】

ピッチの基本周波数の男女差、個人差、音素による差、又は会話内容による差は一般的に狭く、また、ピッチ毎のピッチ周波数の変動も一般的に小さい。したがって、残差周波数は狭レンジとなり、符号化を行う場合の符号化効率が高い。また、等化後の音声信号の1ピッチ区間内の波形は、ピッチの揺らぎ成分が除去されているため、変換符号化により符号量を大幅に圧縮することが可能である。したがって、音声信号の符号化効率を向上させることが可能となる。

【0073】

本発明に係るピッチ周期等化装置の第6の構成は、前記第1の構成において、前記ピッチ検出手段は、前記周波数シフトから出力される出力音声信号のピッチ周波数（以下「出力ピッチ周波数」という。）を検出する出力ピッチ検出手段であり、前記基準周波数を出力する基準周波数発生手段を備え、前記残差演算手段は、前記出力ピッチ周波数と前記基準周波数との差分である残差周波数を演算することを特徴とする。

【0074】

このように、基準周波数として、基準周波数発生手段が出力する決められた周波数を用いることによって、上述の第5の構成の場合と同様に、音声信号の符号化効率を向上させることが可能となる。

【0075】

本発明に係る音声符号化装置の第1の構成は、入力される音声信号を符号化する音声符号化装置であって、前記音声信号に対して有声音のピッチ周期を等化する前記第1乃至6の何れか一の構成のピッチ周期等化装置；及び、前記ピッチ周期等化装置が出力する音声

10

20

30

40

50

信号（以下「ピッチ等化音声信号」という。）に対して、一定のピッチ区間数で直交変換を行い、各サブバンドの変換係数データを生成する直交変換手段；を備えたことを特徴とする。

【0076】

この構成によれば、上述のように、ピッチ周期等化装置において、入力音声信号に含まれる、ピッチの基本周波数の情報、ピッチ毎のピッチ周波数の変動に関する情報、及びピッチに重畳する波形成分の情報は、それぞれ、基準周波数、残差周波数、及び等化後の音声信号（ピッチ等化音声信号）の1ピッチ区間内の波形にそれぞれ分離される。

【0077】

ここで得られるピッチ等化音声信号の1ピッチ区間内の波形（以下「単位ピッチ区間波形」という。）は、基本ピッチ周波数に重畳した音声波形から、ピッチ毎のピッチ周期の変動（ジッタ）や変化が除去されたものである。したがって、直交変換する際に、各ピッチ区間を同一のサンプリング間隔で同一の分解能で直交変換することができるため、ピッチ区間毎の変換符号化を容易に実行できる。また、同じ音素内で隣接するピッチ区間の単位ピッチ区間波形の相関は大きくなる。

【0078】

したがって、このピッチ等化音声信号に対して一定のピッチ区間数で直交変換を行い、各サブバンドの変換係数データとすることにより、高い符号化効率を得ることが可能となる。

【0079】

ここで、直交変換手段により直交変換を行う「一定のピッチ区間数」としては、1ピッチ区間又は2以上の整数倍のピッチ区間とすることができる。しかし、各サブバンドの変換係数データの時間的変化を最小とし、高い符号化効率を得るには1ピッチ区間とするのが好ましい。2以上のピッチ区間とすると各サブバンドの周波数は基準周波数の高調波成分以外の周波数が含まれてくるのに対し、1ピッチ区間とすると、各サブバンドの周波数はすべて基準周波数の高調波成分となるため、各サブバンドの変換係数データの時間的変化が最小となるからである。

【0080】

また、ピッチの基本周波数の情報、及びそれぞれのピッチ区間ごとのピッチ周波数の変動に関する情報の符号化については、それぞれ、ピッチ検出手段が出力するピッチ周波数、及び残差演算手段が出力する残差周波数を符号化することにより行うことができる。ピッチの基本周波数は、音素毎にほぼ一定であるため、符号化を行う場合の符号化効率が高い。また、一般的に各音素内ではピッチの揺らぎの変動幅は小さいため、残差周波数は狭いレンジであり、符号化を行う場合の符号化効率が高い。したがって、全体としての符号化効率も高くなる。

【0081】

さらに、CELP方式と比較すると、本発明に係る音声符号化装置は、符号帳を使用することなく低ビットレートの音声符号化を達成できる点に特徴がある。符号帳を使用しないため、音声符号化装置及び音声復号装置において符号帳を用意する必要がない。そのため、ハードウェアで構成する際の実装面積を小さくすることができる。

【0082】

また、上述したように、符号帳を利用する場合、入力音声と符号帳の候補とのマッチングの度合いにより音声の歪み度合いが定まる。したがって、符号帳の候補と大きく異なる音声が入力された場合には大きな歪みが現れることとなる。この現象を回避しようとした場合、符号帳中の候補数をできるだけ多く用意する必要がある。しかし、候補数を増やすと候補数の対数に比例して全体の符号量も大きくなる。したがって、低ビットレートを実現するためには符号帳の候補数をあまり多くすることはできないため、歪みは或る程度以上は小さくできない。

【0083】

しかし、本発明に係る音声符号化装置は、入力音声を変換符号化によって直接符号化す

10

20

30

40

50

るため、常に入力音声に適合した最適な符号化が行われる。したがって、符号化による音声の歪みは最小限に抑えることが可能であり、S/N比の高い音声符号化を達成することが可能となる。

【0084】

本発明に係る音声符号化装置の第2の構成は、前記第1の構成において、前記ピッチ周期等化装置が出力する前記ピッチ等化音声信号に対して、1ピッチ区間の標本化数が一定となるようにリサンプリングを行うリサンプリング手段を備えていることを特徴とする。

【0085】

この構成により、基準周波数として入力ピッチ周波数の平均あるいは出力ピッチ周波数の平均である平均ピッチ周波数を用いた場合において、基準周波数が時間的に緩やかに変化する場合、リサンプリングによってピッチ区間を常に一定数の標本化数としておくことにより、直交変換手段を簡単に構成することができる。すなわち、直交変換手段は、実際にはPFB (Polyphase Filter Bank: 多層フィルタ・バンク) が用いられるが、ピッチ区間の標本化数が変化すると使用可能なフィルタ数(サブバンドの数)が変化するため、使用されないフィルタ(サブバンド)が生じ、無駄が生じる。従って、リサンプリングによってピッチ区間を常に一定数の標本化数としておくことにより、このような無駄をなくすることができる。

10

【0086】

ここで、リサンプリング手段によるリサンプリングは、特許文献2~4において用いられているリサンプリングとは異なることを注意しておく。特許文献2~4におけるリサンプリングは、揺らぎのあるピッチ周期を一定のピッチ周期とするために行われるリサンプリングである。従って、ピッチ周期の揺らぎの周期(およそ 10^{-3} sec程度)に応じて各ピッチ区間のリサンプリング間隔は振動するものとなる。従って、リサンプリングの結果、ピッチ周期の揺らぎの周期による周波数変調の効果が顕著である。一方、本発明におけるリサンプリングは、既にピッチ周期が等化された音声信号に対して、基準周波数の変化によりピッチ区間ごとの標本化数が異なるのを防止するために行われるリサンプリングである。基準周波数の変化は通常極めて緩やかであり(およそ100 msec程度)、リサンプリングによる周波数変調の影響が問題となることはない。

20

【0087】

本発明に係る音声復号装置は、原音声信号に対しピッチ周波数が所定の基準周波数に等化され、直交変換によりサブバンド成分に分解されたピッチ等化音声信号、及び前記原音声信号のピッチ周波数から前記基準周波数を差し引いた差分である残差周波数信号に基づいて前記原音声信号を復号する音声復号装置であって、一定のピッチ数区間で直交変換されたピッチ等化音声信号に対し逆直交変換を行うことによりピッチ等化音声信号を復元する逆直交変換手段; 及び、前記ピッチ等化音声信号のピッチ周波数を前記基準周波数に前記残差周波数を加えた周波数に近づける方向にシフトさせることにより、前記復元音声信号を生成する周波数シフト; を備え、前記周波数シフトは、前記ピッチ等化音声信号を所定の変調波で振幅変調し被変調波を生成する変調手段; 前記被変調波の単側波帯成分の信号のみを選択的に通過させるバンドパスフィルタ; 前記バンドパスフィルタでフィルタリングした被変調波に対して所定の復調波で復調を行い、復元音声信号として出力する復調手段; 及び、前記変調手段が変調に用いる変調波の周波数及び前記復調手段が復調に用いる復調波の周波数の何れか一方を所定の基本キャリア周波数とし、他方を前記基本キャリア周波数に前記残差周波数を加えた値に設定する周波数調整手段; を備えていることを特徴とする。

30

40

【0088】

この構成により、上記第1又は第2の構成の音声符号化装置により符号化された音声信号を復号することができる。

【0089】

本発明に係るピッチ周期等化方法の第1の構成は、入力される音声信号(以下「入力音声信号」という。)に対して有声音のピッチ周期を等化するピッチ周期等化方法であって

50

、前記入力音声信号を周波数シフトに入力し、前記周波数シフトからの出力信号（以下「出力音声信号」という。）を得る周波数シフトステップ；前記出力音声信号のピッチ周波数（以下「出力ピッチ周波数」という。）を検出する出力ピッチ検出ステップ；前記出力ピッチ周波数から所定の基準周波数を差し引いた差分である残差周波数を演算する残差周波数演算ステップ；及び、前記出力ピッチ周波数と所定の基準周波数との差分である残差周波数を演算する残差周波数演算ステップ；を有し、前記周波数シフトステップにおいては、変調に用いる変調波の周波数及び復調に用いる復調波の周波数の何れか一方を所定の基本キャリア周波数とし、他方を前記基本キャリア周波数から前記残差周波数演算ステップにおいて算出される前記残差周波数を差し引いた周波数に設定する周波数設定ステップ；前記入力音声信号を前記変調波で振幅変調し被変調波を生成する変調ステップ；前記被変調波の単側波帯成分のみを通過させるバンドパスフィルタにより、前記被変調波をフィルタリングする帯域縮小ステップ；前記バンドパスフィルタでフィルタリングした被変調波に対して前記復調波で復調を行い、出力音声信号として出力する復調ステップ；を有することを特徴とする。

10

【0090】

本発明に係るピッチ周期等化方法の第2の構成は、前記第1の構成において、前記出力ピッチ周波数の時間平均である平均ピッチ周波数を演算するピッチ平均化ステップ；を有し、前記残差周波数演算ステップにおいては、前記出力ピッチ周波数と前記平均ピッチ周波数との差分を演算し、これを前記残差周波数とすることを特徴とする。

【0091】

20

本発明に係るピッチ周期等化方法の第3の構成は、前記第1の構成において、前記入力音声信号のピッチ周波数（以下「入力ピッチ周波数」という。）を検出する入力ピッチ検出ステップ；前記入力ピッチ周波数の時間平均である平均ピッチ周波数を演算するピッチ平均化ステップ；を有し、前記残差周波数演算ステップにおいては、前記出力ピッチ周波数と前記平均ピッチ周波数との差分を演算し、これを前記残差周波数とすることを特徴とする。

【0092】

本発明に係るピッチ周期等化方法の第4の構成は、入力される音声信号（以下「入力音声信号」という。）に対して有声音のピッチ周期を等化するピッチ周期等化方法であって、前記入力音声信号のピッチ周波数（以下「入力ピッチ周波数」という。）を検出する入力ピッチ検出ステップ；前記入力音声信号を周波数シフトに入力し、前記周波数シフトからの出力信号（以下「出力音声信号」という。）を得る周波数シフトステップ；及び、前記入力ピッチ周波数から所定の基準周波数を差し引いた差分である残差周波数を演算する残差周波数演算ステップ；を有し、前記周波数シフトステップにおいては、変調に用いる変調波の周波数及び復調に用いる復調波の周波数の何れか一方を所定の基本キャリア周波数とし、他方を前記基本キャリア周波数から前記残差周波数演算ステップにおいて算出される前記残差周波数を差し引いた周波数に設定する周波数設定ステップ；前記入力音声信号を前記変調波で振幅変調し被変調波を生成する変調ステップ；前記被変調波の単側波帯成分のみを通過させるバンドパスフィルタにより、前記被変調波をフィルタリングする帯域縮小ステップ；前記バンドパスフィルタでフィルタリングした被変調波に対して前記復調波で復調を行い、出力音声信号として出力する復調ステップ；を有することを特徴とする。

30

40

【0093】

本発明に係るピッチ周期等化方法の第5の構成は、前記第4の構成において、前記入力ピッチ周波数の時間平均である平均ピッチ周波数を演算するピッチ平均化ステップ；を有し、前記残差周波数演算ステップにおいては、前記入力ピッチ周波数と前記平均ピッチ周波数との差分を演算し、これを前記残差周波数とすることを特徴とする。

【0094】

本発明に係る音声符号化方法の第1の構成は、入力される音声信号を符号化する音声符号化方法であって、前記第1乃至5の何れか一の構成のピッチ周期等化方法により、前記

50

音声信号に対して有声音のピッチ周期を等化するピッチ周期等化ステップ；前記ピッチ周期等化ステップで等化された音声信号（以下「ピッチ等化音声信号」という。）に対して、一定のピッチ区間数で直交変換を行い、各サブバンドの変換係数データを生成する直交変換ステップ；及び、前記変換係数データを符号化する波形符号化ステップ；を備えたことを特徴とする。

【0095】

本発明に係る音声符号化方法の第2の構成は、前記第1の構成において、前記ピッチ周期等化ステップにおいて等化された前記ピッチ等化音声信号に対して、1ピッチ区間の標本化数が一定となるようにリサンプリングを行うリサンプリングステップを備えていることを特徴とする。

10

【0096】

本発明に係るプログラムは、コンピュータで実行することにより、前記コンピュータを請求項1乃至6の何れか一記載のピッチ周期等化装置として機能させることを特徴とする。

【0097】

また、本発明に係るプログラムは、コンピュータで実行することにより、前記コンピュータを請求項7又は8記載の音声符号化装置として機能させることを特徴とする。

【0098】

また、本発明に係るプログラムは、コンピュータで実行することにより、前記コンピュータを上記本発明の音声復号装置として機能させることを特徴とする。

20

【発明の効果】

【0099】

以上のように、本発明に係るピッチ周期等化装置によれば、入力音声信号に含まれる情報を、ピッチの基本周波数の情報、ピッチ毎のピッチ周波数の変動に関する情報、及びピッチに重畳する波形成分の情報に分離する。これらの情報は、それぞれ、基準周波数、残差周波数、及び等化後の音声信号の1ピッチ区間内の波形として取り出される。

【0100】

このように、分離された情報から、ピッチの基本周波数の情報及びピッチに重畳する波形成分の情報のみを用いれば、マッチング誤差が小さく精度のよい音声検索が可能となる。

30

【0101】

また、各情報を分離して、それぞれの情報を個別に最適な符号化方法で符号化することにより、入力音声信号の符号化効率を向上させることが可能となる。

【0102】

故に、精度のよい音声検索を可能とし、入力音声信号の符号化効率を向上させることを可能とするピッチ周期等化装置を提供することができる。

【0103】

また、本発明に係る音声符号化装置によれば、入力音声信号に含まれる情報は、ピッチ周期等化装置において、ピッチの基本周波数の情報、ピッチ毎のピッチ周波数の変動に関する情報、及びピッチに重畳する波形成分の情報に分離され、それぞれ、基準周波数、残差周波数、及びピッチ等化音声信号の1ピッチ区間内の波形として得られる。そして、このピッチ等化音声信号に対して一定のピッチ区間数で直交変換を行うことで、ピッチに重畳する波形成分の情報を効率よく符号化することができる。

40

【発明を実施するための最良の形態】

【0104】

以下、本発明を実施するための最良の形態について、図面を参照しながら説明する。

【実施例1】

【0105】

図1は、本発明の実施例1に係るピッチ周期等化装置1の構成を表すブロック図である。ピッチ周期等化装置1は、入力ピッチ検出手段2、ピッチ平均手段3、周波数シフタ4

50

、出力ピッチ検出手段 5、残差演算手段 6、及び P I D コントローラ 7 を備えている。

【 0 1 0 6 】

入力ピッチ検出手段 2 は、入力端子 In より入力される入力音声信号 $x_{in}(t)$ から、当該音声信号に含まれるピッチの基本周波数を検出する。ピッチの基本周波数を検出する方法は、現在までに種々の方法が考案されているが、本実施例ではその代表的なものを示す。この入力ピッチ検出手段 2 は、ピッチ検出手段 1 1、バンドパスフィルタ (Band Pass Filter: 以下「B P F」という。) 1 2、及び周波数カウンタ 1 3 を備えている。

【 0 1 0 7 】

ピッチ検出手段 1 1 は、入力音声信号 $x_{in}(t)$ から、ピッチの基本周波数 f_0 を検出する。例えば、入力音声信号 $x_{in}(t)$ が図 2 (a) のような波形であったとする。ピッチ検出手段 1 1 は、まずこの波形に対して短時間フーリエ変換を行い、図 2 (b) のようなスペクトル波形 $X(f)$ を導出する。

10

【 0 1 0 8 】

通常、音声波形は、ピッチ以外にも多くの周波数成分を含み、ここで得られるスペクトル波形は、ピッチの基本周波数及びピッチの高調波成分以外にも、付加的に多くの周波数成分を有する。したがって、このスペクトル波形 $X(f)$ からピッチの基本周波数 f_0 を抽出するのは一般に困難である。そこで、ピッチ検出手段 1 1 は、このスペクトル波形 $X(f)$ に対し再度フーリエ変換を行う。これにより、スペクトル波形 $X(f)$ に含まれるピッチの高調波の間隔 f_0 の逆数 $F_0 = 1 / f_0$ の点に鋭いピークを持つスペクトル波形が得られる (図 2 (c) 参照)。ピッチ検出手段 1 1 は、このピークの位置 F_0 を検出することによって、ピッチの基本周波数 $f_0 = 1 / F_0$ を検出する。

20

【 0 1 0 9 】

また、ピッチ検出手段 1 1 は、スペクトル波形 $X(f)$ から、入力音声信号 $x_{in}(t)$ が有声音か無声音かを判別する。有声音の場合には、ノイズフラグ信号 V_{noise} として 0 を出力する。無声音の場合にはノイズフラグ信号 V_{noise} として 1 を出力する。なお、有声音と無声音の判別は、スペクトル波形 $X(f)$ の傾き検出によって行われる。図 5 は有声音「あ」のフォルマント特性を示す図であり、図 6 は無声音「す」の自己相関及びケプストラム波形並びに周波数特性を示す図である。有声音は、図 5 のように、スペクトル波形 $X(f)$ は、全体的に低周波側が大きく高周波側に向かって小さくなるようなフォルマント特性を示す。それに対して、無声音は、図 6 のように、全体的に高周波側に向かって大きくなるような周波数特性を示す。したがって、スペクトル波形 $X(f)$ の全体的な傾きを検出することによって、入力音声信号 $x_{in}(t)$ が有声音か無声音かを判別することができる。

30

【 0 1 1 0 】

尚、入力音声信号 $x_{in}(t)$ が無声音の場合、ピッチが存在しないので、ピッチ検出手段 1 1 が出力するピッチの基本周波数 f_0 は無意味な値となる。

【 0 1 1 1 】

B P F 1 2 は、中心周波数が可変な狭帯域の F I R (Finite Impulse Response) 型のフィルタが使用される。B P F 1 2 は、ピッチ検出手段 1 1 により検出されるピッチの基本周波数 f_0 を通過帯域の中心周波数として設定する (図 2 (d) 参照)。そして、B P F 1 2 は、入力音声信号 $x_{in}(t)$ をフィルタリングし、ピッチの基本周波数 f_0 のほぼ正弦波状の波形を出力する (図 2 (e) 参照)。

40

【 0 1 1 2 】

周波数カウンタ 1 3 は、B P F 1 2 が出力するほぼ正弦波状の波形のゼロクロス点の単位時間当たりの数をカウントすることにより、ピッチの基本周波数 f_0 を出力する。この検出されたピッチの基本周波数 f_0 が入力ピッチ検出手段 2 の出力信号 (以下「基本周波数信号」) V_{pitch} として出力される (図 2 (f) 参照)。

【 0 1 1 3 】

ピッチ平均手段 3 は、ピッチ検出手段 1 1 が出力するピッチの基本周波数信号 V_{pitch} を平均化するものであり、通常のローパスフィルタ (Low Pass Filter: 以下「L P F」

50

という。)が使用される。ピッチ平均手段3により、基本周波数信号 V_{pitch} が平滑化され、音素内では時間的にほぼ一定の信号となる(図2(g)参照)。この平滑化された基本周波数が基準周波数 f_s として使用される。

【0114】

周波数シフト4は、入力音声信号 $x_{in}(t)$ のピッチ周波数を基準周波数 f_0 に近づける方向にシフトさせることにより、音声信号のピッチ周期を等化する。

【0115】

出力ピッチ検出手段5は、周波数シフト4より出力される出力音声信号 $x_{out}(t)$ から、当該出力音声信号 $x_{out}(t)$ に含まれるピッチの基本周波数 f_0' を検出する。この出力ピッチ検出手段5も、基本的に入力ピッチ検出手段2と同様の構成とすることができる。本実施例の場合、出力ピッチ検出手段5は、BPF15及び周波数カウンタ16を備えている。

10

【0116】

BPF15は、中心周波数が可変な狭帯域のFIR型のフィルタが使用される。BPF15は、ピッチ検出手段11により検出されるピッチの基本周波数 f_0 を通過帯域の中心周波数として設定する。そして、BPF15は、出力音声信号 $x_{out}(t)$ をフィルタリングし、ピッチの基本周波数 f_0' のほぼ正弦波状の波形を出力する。周波数カウンタ16は、BPF15が出力するほぼ正弦波状の波形のゼロクロス点の単位時間当たりの数をカウントすることにより、ピッチの基本周波数 f_0' を出力する。この検出されたピッチの基本周波数 f_0' が出力ピッチ検出手段5の出力信号 V_{pitch}' として出力される。

20

【0117】

残差演算手段6は、出力ピッチ検出手段5が出力する基本周波数 f_0' からピッチ平均手段3が出力する基準周波数 f_s を引いた残差周波数 f_{pitch} を出力する。この残差周波数 f_{pitch} は、PIDコントローラ7を介して周波数シフト4に入力される。周波数シフト4は、残差周波数 f_{pitch} に比例して、入力音声信号のピッチ周波数を基準周波数 f_0 に近づける方向にシフトさせる。

【0118】

尚、PIDコントローラ7は、直列接続されたアンプ18及び抵抗20、並びに、アンプ18に対して並列接続されたコンデンサ19から構成されている。このPIDコントローラ7は、周波数シフト4、出力ピッチ検出手段5、及び残差演算手段6からなるフィードバックループの発振を防止するためのものである。

30

【0119】

尚、図1では、PIDコントローラ7は、アナログ回路表示しているが、デジタル回路で構成してもよい。

【0120】

図3は周波数シフト4の内部構成を表す図である。周波数シフト4は、発信器21、変調器22、BPF23、電圧制御発信器(Voltage Controlled Oscillator:以下「VCO」という。)24、及び復調器25を備えている。

【0121】

発信器21は、入力音声信号 $x_{in}(t)$ の振幅変調を行うための一定周波数の変調キャリア信号C1を出力する。通常、音声信号の帯域は8kHz程度である(図3(a)参照)。したがって、発信器21が発生する変調キャリア信号C1の周波数(以下「キャリア周波数」という。)としては、通常は20kHz程度のものが使用される。

40

【0122】

変調器22は、発信器21が出力する変調キャリア信号C1を入力音声信号 $x_{in}(t)$ で振幅変調し、被変調信号を生成する。この被変調信号は、キャリア周波数を中心として、その両側に音声信号の帯域と同じバンド幅の側波帯(上側波帯及び下側波帯)を有する信号である(図3(b)参照)。

【0123】

BPF23は、被変調信号の上側波帯成分のみを通過させるBPFである。これにより

50

、BPF23から出力される被変調信号は、下側波帯のみがカットされた単側波帯信号となる。

【0124】

VCO24は、発信器21が出力する変調キャリア信号C1と同じキャリア周波数の信号を、PIDコントローラ7を介して残差演算手段6から入力される残差周波数 f_{pitch} の信号（以下「残差周波数信号」という。） V_{pitch} により周波数変調して得られる信号（以下「復調キャリア信号」という。）を出力する。復調キャリア信号の周波数は、キャリア周波数から残差周波数を差し引いた周波数となる。

【0125】

復調器25は、BPF23が出力する上側波帯のみの被変調信号を、VCO24が出力する復調キャリア信号により復調し、音声信号を復元する（図3（d）参照）。このとき、復調キャリア信号は、残差周波数信号 V_{pitch} で変調されている。そのため、被変調信号を復調する際に、入力音声信号 $x_{in}(t)$ のピッチ周波数の基準周波数 f_s からのずれが消去される。すなわち、入力音声信号 $x_{in}(t)$ のピッチ周期は、基準周期 $1/f_s$ に等化される。

10

【0126】

図4は、周波数シフタ4の内部構成の他の例を表す図である。図4においては、図3の発信器21とVCO24とを入れ替えた構成とされている。この構成によっても、図3の場合と同様に、入力音声信号 $x_{in}(t)$ のピッチ周期を基準周期 $1/f_s$ に等化することができる。

20

【0127】

以上のように構成された実施例1のピッチ周期等化装置1について、以下その動作を説明する。

【0128】

まず、入力音声信号 $x_{in}(t)$ が入力端子Inから入力されると、入力ピッチ検出手段2は、入力音声信号 $x_{in}(t)$ が有声音か無声音かを判別してノイズフラグ信号 V_{noise} を出力端子OUT_4へ出力するとともに、入力音声信号 $x_{in}(t)$ からピッチ周波数を検出し、基本周波数信号 V_{pitch} をピッチ平均手段3に出力する。ピッチ平均手段3は、基本周波数信号 V_{pitch} を平均化し（この場合、LPFを使用するので加重平均となる。）、これを基準周波数信号 AV_{pitch} として出力する。この基準周波数信号 AV_{pitch} は、出力端子OUT_3から出力されるとともに、残差演算手段6に入力される。

30

【0129】

一方、周波数シフタ4は、入力音声信号 $x_{in}(t)$ の周波数をシフトさせ、出力音声信号 $x_{out}(t)$ として出力端子Out_1へ出力する。初期状態においては、残差周波数信号 V_{pitch} は0（リセット状態）であり、周波数シフタ4は、入力音声信号 $x_{in}(t)$ がそのまま出力音声信号 $x_{out}(t)$ として出力端子Out_1へ出力される。

【0130】

次に、出力ピッチ検出手段5は、周波数シフタ4が出力する出力音声信号のピッチ周波数 f_0' を検出する。検出されたピッチ周波数 f_0' は、ピッチ周波数信号 V_{pitch}' として残差演算手段6に入力される。

40

【0131】

残差演算手段6は、ピッチ周波数信号 V_{pitch}' から基準周波数信号 AV_{pitch} を差し引くことにより、残差周波数信号 V_{pitch} を生成する。この残差周波数信号 V_{pitch} は、出力端子Out_2へ出力されるとともに、PIDコントローラ7を介して周波数シフタ4へ入力される。

【0132】

周波数シフタ4は、PIDコントローラ7を介して入力される残差周波数信号 V_{pitch} に比例して、周波数のシフト量を設定する。この場合、残差周波数信号 V_{pitch} が正値であれば、残差周波数信号 V_{pitch} に比例した量だけ周波数を下げるようにシフト量が設定される。残差周波数信号 V_{pitch} が負値であれば、残差周波数信号 V_{pitch} に比例

50

した量だけ周波数を上げるようにシフト量が設定される。

【0133】

このようなフィードバック制御により、入力音声信号 $x_{in}(t)$ のピッチ周期は、常に基準周期 $1/f_s$ に維持され、出力音声信号 $x_{out}(t)$ のピッチ周期は等化される。

【0134】

以上のように、本実施例1のピッチ周期等化装置1によれば、入力音声信号 $x_{in}(t)$ に含まれる情報は、

- (a) 有声音か無声音かを示す情報；
- (b) 1ピッチ区間の音声波形を表す情報；
- (c) 基準ピッチ周波数の情報；
- (d) 各ピッチ区間のピッチ周波数の基準ピッチ周波数からの偏倚量を表す残差周波数情報；

10

に分離される。(a)~(d)の情報は、それぞれ、ノイズフラグ信号 V_{noise} 、ピッチ周期が基準周期 $1/f_s$ (入力音声信号の過去のピッチ周波数の加重平均の逆数) に等化された出力音声信号 $x_{out}(t)$ 、基準周波数信号 AV_{pitch} 、及び残差周波数信号 V_{pitch} として出力される。

【0135】

出力音声信号 $x_{out}(t)$ は、男女差、個人差、音素、感情及び会話内容によって変化するピッチ周波数のジッタ成分や変化成分が除去された音声信号であり、抑揚のない平坦的・機械的な音声信号である。したがって、同じ有声音の出力音声信号 $x_{out}(t)$ は、男女差、個人差、音素、感情又は会話内容に無関係にほぼ同じ波形が得られるため、出力音声信号 $x_{out}(t)$ を比較することによって有声音についてのマッチングを精度よく行うことが可能となる。すなわち、ピッチ周期等化装置1を音声検索装置に応用すれば、検索精度を向上させることが可能となる。

20

【0136】

また、有声音の出力音声信号 $x_{out}(t)$ はピッチ周期が基準周期 $1/f_s$ に等化されているので、一定数のピッチ区間でサブバンド符号化を行うことにより、出力音声信号 $x_{out}(t)$ の周波数スペクトル $X_{out}(f)$ は、基準周波数の高調波成分のサブバンド成分に集約される。音声はピッチ間の波形相関が大きいので、各サブバンド成分のスペクトル強度の時間変化は緩やかである。したがって、各サブバンド成分を符号化し、その他の雑音成分を省略することにより、高効率の符号化が可能となる。また、基準周波数信号 AV_{pitch} 、及び残差周波数信号 V_{pitch} は、音声の性質上、同一音素内で狭レンジでしか変動しないため、高効率の符号化が可能である。したがって、全体として入力音声信号 $x_{in}(t)$ の有声音成分を高効率で符号化することが可能となる。

30

【実施例2】

【0137】

図7は本発明の実施例2に係るピッチ周期等化装置1'の構成を表す図である。実施例1のピッチ周期等化装置1が残差周波数 f_{pitch} のフィードバック制御によりピッチ周期を等化する構成としたのに対し、実施例2のピッチ周期等化装置1'は、残差周波数 f_{pitch} のフィードフォワード制御によりピッチ周期を等化する構成とされている。

40

【0138】

図7において、入力ピッチ検出手段2、ピッチ平均手段3、周波数シフト4、残差演算手段6、ピッチ検出手段11、BPF12、及び周波数カウンタ13は、図1のものと同様であるため、同一符号を付して説明は省略する。

【0139】

ピッチ周期等化装置1'では、残差演算手段6は、入力ピッチ検出手段2が出力する基本周波数信号 V_{pitch} から基準周波数信号 AV_{pitch} を差し引くことによって残差周波数信号 V_{pitch} を生成する。また、フィードフォワード制御であるため、発振対策は不要であり、PIDコントローラ7は省略されている。また、フィードフォワード制御であるため、出力ピッチ検出手段5も省略されている。その他の構成は実施例1と同様である

50

。

【0140】

このような構成によっても、実施例1の場合と同様に、入力音声信号 $x_{in}(t)$ をノイズフラグ信号 V_{noise} 、出力音声信号 $x_{out}(t)$ 、基準周波数信号 AV_{pitch} 、及び残差周波数信号 V_{pitch} に分離することが可能である。

【実施例3】

【0141】

図8は本発明の実施例3に係る音声符号化装置30の構成を表す図である。音声符号化装置30は、ピッチ周期等化装置1、1'、リサンプラ31、アナライザ32、量子化器33、ピッチ等化波形符号化器34、差分ビット演算器35、及びピッチ情報符号化器36を備えている。

10

【0142】

ピッチ周期等化装置1、1'は、実施例1、2に示したピッチ周期等化装置である。リサンプラ31は、ピッチ周期等化装置1、1'の出力端子Out_1から出力される出力音声信号 $x_{out}(t)$ の各ピッチ区間について、同一の標本化数となるようにリサンプリングを行い、等標本数音声信号 $x_{eq}(t)$ として出力する。

【0143】

アナライザ32は、等標本数音声信号 $x_{eq}(t)$ について、一定のピッチ区間数で変形離散コサイン変換(Modified Discrete Cosine Transform:以下「MDCT」という。)を行い、 n 個のサブバンド成分の周波数スペクトル信号 $X(f) = \{X(f_1), X(f_2), \dots, X(f_n)\}$ を生成する。量子化器33は、周波数スペクトル信号 $X(f)$ を所定の量子化曲線に従って量子化する。ピッチ等化波形符号化器34は、量子化器33が出力する周波数スペクトル信号 $X(f)$ を符号化し、符号化波形データとして出力する。この符号化には、ハフマン符号化法や算術符号化法等のエントロピ符号化法が使用される。

20

【0144】

差分ビット演算器35は、ピッチ等化波形符号化器34が出力する符号化波形データの符号量から目的ビット数を減算し差分(以下「差分ビット数」という。)を出力する。量子化器33は、この差分ビット数によって量子化曲線を平行移動させ、符号化波形データの符号量が目的ビット数の範囲内となるように調整する。

30

【0145】

ピッチ情報符号化器36は、ピッチ周期等化装置1、1'が出力する残差周波数信号 V_{pitch} 及び基準周波数信号 AV_{pitch} を符号化し、符号化ピッチデータとして出力する。この符号化には、ハフマン符号化法や算術符号化法等のエントロピ符号化法が使用される。

【0146】

以上のように構成された本実施例に係る音声符号化装置30について、以下その動作を説明する。

【0147】

まず、入力音声信号 $x_{in}(t)$ が入力端子Inから入力される。ピッチ周期等化装置1、1'は実施例1で説明したように、入力音声信号 $x_{in}(t)$ の波形情報を、
(a) 有声音か無声音かを示す情報；
(b) 1ピッチ区間の音声波形を表す情報；
(c) 基準ピッチ周波数の情報；
(d) 各ピッチ区間のピッチ周波数の基準ピッチ周波数からの偏倚量を表す残差周波数情報；

40

に分離し、それぞれ、ノイズフラグ信号 V_{noise} 、出力音声信号 $x_{out}(t)$ 、基準周波数信号 AV_{pitch} 、及び残差周波数信号 V_{pitch} として出力する。ノイズフラグ信号 V_{noise} は出力端子Out_4から出力され、出力音声信号 $x_{out}(t)$ は出力端子Out_1から出力され、基準周波数信号 AV_{pitch} は出力端子Out_3から出力され、残差周波数信号 V_{pitch}

50

は出力端子Out_2 から出力される。

【0148】

次に、リサンプラ31は、各ピッチ区間において、基準周波数信号 AV_{pitch} を一定のリサンプリング数 n で除算することによりリサンプリング周期を計算する。そして、出力音声信号 $x_{out}(t)$ をそのリサンプリング周期によりリサンプリングし、等標本数音声信号 $x_{eq}(t)$ として出力する。これにより、出力音声信号 $x_{out}(t)$ の1ピッチ区間の標本化数が一定の値とされる。

【0149】

次に、アナライザ32は、等標本数音声信号 $x_{eq}(t)$ を、一定のピッチ区間数のサブフレームに区分する。そして、サブフレーム毎に変形離散コサイン変換を行うことによ

10

【0150】

ここで、1つのサブフレームの長さは、1ピッチ周期の整数倍とされる。本実施例では、サブフレームの長さは1ピッチ周期(標本化数 n)とする。従って、 n 個の周波数スペクトル信号 $\{X(f_1), X(f_2), \dots, X(f_n)\}$ が出力される。周波数 f_1 は基準周波数の第1高調波、周波数 f_2 は基準周波数の第2高調波、周波数 f_n は基準周波数の第 n 高調波である。

【0151】

このように、1ピッチ周期の整数倍のサブフレームに分割して各サブフレームを直交変換することによりサブバンド符号化を行うことで、音声波形データの周波数スペクトル信

20

【0152】

図10に各サブバンドのスペクトル強度の時間変化の例を示す。図10(a)は日本語の母音に対する各サブバンドのスペクトル強度の時間変化を示している。下から基準周波数の第1高調波、第2高調波、...、第8高調波の順に示している。図10(b)は「アラユルゲンジツラスベテジブンノハウヘネジマゲタノダ」という音声信号に対する各サブバンドのスペクトル強度の時間変化を示している。これも、下から基準周波数の第1高調波、第2高調波、...、第8高調波の順に示している。図10(a)(b)は横軸が時間であり、縦軸がスペクトル強度である。これから分かるように、有声音の各ピッチ区間では各サブバンドのスペクトル強度は平坦な(DC的な)特性を示す。従って、符号化した場合に符号化効率が高いことが容易に分かる。

30

【0153】

次に、量子化器33は、周波数スペクトル信号 $X(f)$ を量子化する。ここで、量子化器33はノイズフラグ信号 V_{noise} を参照し、ノイズフラグ信号 V_{noise} が0(有声音)の場合と1(無声音)の場合とで量子化曲線を切り換える。

【0154】

ノイズフラグ信号 V_{noise} が0(有声音)の場合、量子化曲線は、図8(a)に示したように、周波数が高くなるに従って量子化ビット数が減少するような量子化曲線とされる

40

これは、有声音の周波数特性は、図5に示したように低周波数域で大きく高周波域にいくに従って減少する特性を有することに対応させたものである。

一方、ノイズフラグ信号 V_{noise} が1(無声音)の場合、量子化曲線は、図8(b)に示したように、周波数が高くなるに従って量子化ビット数が増加するような量子化曲線とされる。これは、無声音の周波数特性は、図6に示したように高周波域にいくに従って増加する特性を有することに対応させたものである。

【0155】

この量子化曲線の切り換えにより、有声音が無声音かに対応して最適な量子化曲線が選択される。

50

【0156】

【 0 1 5 7 】

尚、補足として、量子化ビット数について説明する。量子化器 3 3 による量子化のデータフォーマットは図 9 (a) (b) に示したように、小数点以下の実数部 (F L) 及び 2 の冪乗を表す指数部 (E X P) によって表現される。但し、0 以外の数を表す場合において、実数部 (F L) の先頭の 1 ビットは必ず 1 であるように指数部 (E X P) が調整されるものとする。

【 0 1 5 8 】

例えば、実数部 (F L) が 4 ビット、指数部 (E X P) が 2 ビットの場合において、4 ビットで量子化する場合、及び 2 ビットで量子化する場合は、次のようになる (図 9 (c) , (d) 参照)。

【 0 1 5 9 】

(1) 4 ビットで量子化する場合

(例 1) $X(f) = 8 = [1 0 0 0]_2$ (但し、 $[]_2$ は 2 進数表記を表す。) は、
 $FL = [1 0 0 0]_2$, $EXP = [1 0 0]_2$

(例 2) $X(f) = 7 = [0 1 0 0]_2$ は、
 $FL = [1 1 1 0]_2$, $EXP = [0 1 1]_2$

(例 3) $X(f) = 3 = [1 0 0 0]_2$ は、
 $FL = [1 1 0 0]_2$, $EXP = [0 1 0]_2$

【 0 1 6 0 】

(2) 2 ビットで量子化する場合

(例 1) $X(f) = 8 = [1 0 0 0]_2$ は、
 $FL = [1 0 0 0]_2$, $EXP = [1 0 0]_2$

(例 2) $X(f) = 7 = [0 1 0 0]_2$ は、
 $FL = [1 1 0 0]_2$, $EXP = [0 1 1]_2$

(例 3) $X(f) = 3 = [1 0 0 0]_2$ は、
 $FL = [1 1 0 0]_2$, $EXP = [0 1 0]_2$

【 0 1 6 1 】

すなわち、n ビットで量子化する場合は、実数部 (F L) の先頭から n ビットを残し、残りのビットは 0 とするものとする (図 9 (d) 参照)。

【 0 1 6 2 】

次に、ピッチ等化波形符号化器 3 4 は、量子化器 3 3 が出力する量子化された周波数スペクトル信号 $X(f)$ をエントロピ符号化法により符号化し、符号化波形データを出力する。また、ピッチ等化波形符号化器 3 4 は、符号化波形データの符号量 (ビット数) を差分ビット演算器 3 5 に出力する。差分ビット演算器 3 5 は、符号化波形データの符号量から所定の目的ビット数を減算し、差分ビット数を出力する。量子化器 3 3 は、差分ビット数に応じて、有声音に対する量子化曲線を平行移動的に上下させる。

【 0 1 6 3 】

例えば、 $\{ f_1, f_2, f_3, f_4, f_5, f_6 \}$ に対する量子化曲線が $\{ 6, 5, 4, 3, 2, 1 \}$ であったとし、差分ビット数として 2 が入力されたとすると、量子化器 3 3 は、量子化曲線を下方に 2 だけ平行移動する。その結果、量子化曲線は $\{ 4, 3, 2, 1, 0, 0 \}$ となる。また、差分ビット数として - 2 が入力されたとすると、量子化器 3 3 は、量子化曲線を上方に 2 だけ平行移動する。その結果、量子化曲線は $\{ 8, 7, 6, 5, 4, 3 \}$ となる。

【 0 1 6 4 】

このように有声音の量子化曲線を上下に変化させることによって、各サブフレームの符号化波形データの符号量が目的ビット数程度に調整される。

【 0 1 6 5 】

一方、これに並行して、ピッチ情報符号化器 3 6 は、基準周波数信号 AV_{pitch} 及び残差周波数信号 V_{pitch} を符号化する。

【 0 1 6 6 】

10

20

30

40

50

以上のように、本実施例の音声符号化装置 30 によれば、有声音のピッチ周期を等化し、1 ピッチ周期の整数倍の長さのサブフレームに分割し、これら各サブフレームを直交変換してサブバンド符号化することにより、時間的に変化が少ないサブフレームの周波数スペクトルが時系列的に得られる。従って、高い符号化効率で符号化することができる。

【実施例 4】

【0167】

図 11 は、本発明の実施例 4 に係る音声復号装置 50 の構成を表すブロック図である。音声復号装置 50 は、実施例 3 の音声符号化装置 30 により符号化された音声信号を復号する装置である。音声復号装置 50 は、ピッチ等化波形復号器 51、逆量子化器 52、シンセサイザ 53、ピッチ情報復号器 54、ピッチ周波数検出手段 55、差分器 56、加算器 57、及び周波数シフト 58 を備えている。

10

【0168】

音声復号装置 50 には、符号化波形データ及び符号化ピッチデータが入力される。符号化波形データは、図 9 のピッチ等化波形符号化器 34 から出力される符号化波形データである。符号化ピッチデータは、図 9 のピッチ情報符号化器 36 から出力される符号化ピッチデータである。

【0169】

ピッチ等化波形復号器 51 は、符号化波形データを復号し、量子化後の各サブバンドの周波数スペクトル信号（以下「量子化周波数スペクトル信号」という。）を復元する。逆量子化器 52 は、この量子化周波数スペクトル信号を逆量子化し、 n 個のサブバンドの周波数スペクトル信号 $X(f) = \{X(f_1), X(f_2), \dots, X(f_n)\}$ を復元する。

20

【0170】

シンセサイザ 53 は、周波数スペクトル信号 $X(f)$ を逆変形離散コサイン変換（Inverse Modified Discrete Cosine Transform：以下「IMDCT」という。）し、1 ピッチ区間の時系列データ（以下「等化音声信号」という。） $x_{eq}(t)$ を生成する。ピッチ周波数検出手段 55 は、この等化音声信号 $x_{eq}(t)$ のピッチ周波数を検出し等化ピッチ周波数信号 V_{eq} として出力する。

【0171】

一方、ピッチ情報復号器 54 は、符号化ピッチデータを復号することにより、基準周波数信号 $A V_{pitch}$ 及び残差周波数信号 V_{pitch} を復元する。差分器 56 は、基準周波数信号 $A V_{pitch}$ から等化ピッチ周波数信号 V_{eq} を差し引いた差分を基準周波数変化信号 $A V_{pitch}$ として出力する。加算器 57 は、残差周波数信号 V_{pitch} と基準周波数変化信号 $A V_{pitch}$ とを加算してこれを修正残差周波数信号 V_{pitch} として出力する。

30

【0172】

周波数シフト 58 は、図 3 又は図 4 に示した周波数シフト 4 と同様の構成を有する。この場合、入力端子 In には等化音声信号 $x_{eq}(t)$ が入力され、VCO 24 には修正残差周波数信号 V_{pitch} が入力される。VCO 24 は発信器 21 が出力する変調キャリア信号 C1 と同じキャリア周波数の信号を、加算器 57 から入力される修正残差周波数信号 V_{pitch} により周波数変調して得られる信号（以下「復調キャリア信号」という。）を出力するが、この場合、復調キャリア信号の周波数は、キャリア周波数に残差周波数を加えた周波数となる。

40

【0173】

これにより、周波数シフト 58 において等化音声信号 $x_{eq}(t)$ の各ピッチ区間のピッチ周期に揺らぎ成分が加えられ、音声信号 $x_{res}(t)$ が復元される。

【実施例 5】

【0174】

図 12 は本発明の実施例 5 に係るピッチ周期等化装置 41 の構成を表す図である。本実施例に係るピッチ周期等化装置 41 の基本構成は、実施例 2 に係るピッチ周期等化装置 1' と略同様であるが、基準周波数として一定の周波数を使用する点において異なる。

50

【 0 1 7 5 】

ピッチ周期等化装置 4 1 は、入力ピッチ検出手段 2、周波数シフタ 4、残差演算手段 6、及び基準周波数発生器 4 2 を備えている。入力ピッチ検出手段 2、周波数シフタ 4、残差演算手段 6 は図 7 と同様のものであるため説明は省略する。

【 0 1 7 6 】

基準周波数発生器 4 2 は、予め定められた一定の基準周波数信号を発生する。残差演算手段 6 は、入力ピッチ検出手段 2 が出力する基本周波数信号 V_{pitch} からこの基準周波数信号 V_s を差し引いて、残差周波数信号 V_{pitch} を生成する。この残差周波数信号 V_{pitch} は周波数シフタ 4 にフィードフォワードされる。あとの構成及び動作は実施例 2 と同様である。

10

【 0 1 7 7 】

この構成によれば、ピッチ周期等化装置 4 1 は、入力音声信号 $x_{in}(t)$ の波形情報を

- (a) 有声音か無声音かを示す情報；
- (b) 1 ピッチ区間の音声波形を表す情報；
- (c) 各ピッチ区間のピッチ周波数の基準ピッチ周波数からの偏倚量を表す残差周波数情報；

に分離し、それぞれ、ノイズフラグ信号 V_{noise} 、出力音声信号 $x_{out}(t)$ 、及び残差周波数信号 V_{pitch} として出力する。実施例 2 と異なるのは、基準ピッチ周波数の情報を各ピッチ区間のピッチ周波数の基準ピッチ周波数からの偏倚量を表す残差周波数情報の中に繰り込んだ点である。一般にピッチ周波数はあまり大きな変化はしないため、このように残差周波数情報の中に繰り込んで残差周波数信号 V_{pitch} のレンジはさほど大きくはならない。従って、これによっても高い符号化効率を可能とするピッチ周期等化装置 4 1 が得られる。

20

【 実施例 6 】

【 0 1 7 8 】

図 1 3 は本発明の実施例 6 に係るピッチ周期等化装置 4 1 ' の構成を表す図である。本実施例に係るピッチ周期等化装置 4 1 ' の基本構成は、実施例 1 に係るピッチ周期等化装置 1 と略同様であるが、基準周波数として一定の周波数を使用する点において異なる。

【 0 1 7 9 】

ピッチ周期等化装置 4 1 ' は、周波数シフタ 4、出力ピッチ検出手段 5 "、残差演算手段 6、PID コントローラ 7、及び基準周波数発生器 4 2 を備えている。周波数シフタ 4、出力ピッチ検出手段 5 "、残差演算手段 6 は図 8 と同様のものであるため説明は省略する。また、基準周波数発生器 4 2 は、図 1 2 のものと同様である。

30

【 0 1 8 0 】

基準周波数発生器 4 2 は、予め定められた一定の基準周波数信号を発生する。残差演算手段 6 は、出力ピッチ検出手段 5 " が出力する基本周波数信号 V_{pitch}' からこの基準周波数信号 V_s を差し引いて、残差周波数信号 V_{pitch} を生成する。この残差周波数信号 V_{pitch} は PID コントローラ 7 を介して周波数シフタ 4 にフィードバックされる。あとの構成及び動作は実施例 1 と同様である。

40

【 0 1 8 1 】

この構成によれば、ピッチ周期等化装置 4 1 ' は、入力音声信号 $x_{in}(t)$ の波形情報を、

- (a) 有声音か無声音かを示す情報；
- (b) 1 ピッチ区間の音声波形を表す情報；
- (c) 各ピッチ区間のピッチ周波数の基準ピッチ周波数からの偏倚量を表す残差周波数情報；

に分離し、それぞれ、ノイズフラグ信号 V_{noise} 、出力音声信号 $x_{out}(t)$ 、及び残差周波数信号 V_{pitch} として出力する。実施例 3 と異なるのは、基準ピッチ周波数の情報を各ピッチ区間のピッチ周波数の基準ピッチ周波数からの偏倚量を表す残差周波数情報の中

50

に繰り込んだ点である。一般にピッチ周波数はあまり大きな変化はしないため、このように残差周波数情報の中に繰り込んで残差周波数信号 V_{pitch} のレンジはさほど大きくはならない。従って、これによっても高い符号化効率を可能とするピッチ周期等化装置 41' が得られる。

【実施例 7】

【0182】

図 14 は本発明の実施例 7 に係る音声符号化装置 30' の構成を表す図である。音声符号化装置 30' は、ピッチ周期等化装置 41, 41'、アナライザ 32、量子化器 33、ピッチ等化波形符号化器 34、差分ビット演算器 35、及びピッチ情報符号化器 36' を備えている。

10

【0183】

アナライザ 32、量子化器 33、ピッチ等化波形符号化器 34、及び差分ビット演算器 35 については実施例 3 と同様のものである。また、ピッチ周期等化装置 41, 41' は、実施例 5 又は実施例 6 に係る音声符号化装置 30' である。

【0184】

ピッチ周期等化装置 41, 41' においては、ピッチ周期は常に一定の基準周期 $1/f_s$ に等化される。従って、1ピッチ区間の標本化数は常に一定であり、実施例 3 の音声符号化装置 30 におけるリサンブラ 31 は必要ないため省略されている。また、ピッチ周期は常に一定の基準周期 $1/f_s$ に等化されるので、ピッチ周期等化装置 41, 41' は、基準周波数信号 AV_{pitch} を出力しない。従って、ピッチ情報符号化器 36' は残差周波数信号 V_{pitch} のみを符号化する。

20

【0185】

以上のような構成により、ピッチ周期等化装置 41, 41' を用いた音声符号化装置 30' を実現することができる。この音声符号化装置 30' を実施例 3 の音声符号化装置 30 と比較すると、以下の点で異なる。

【0186】

(1) 実施例 3 の音声符号化装置 30 では、基準周波数信号 AV_{pitch} が時間的に多少変化するため、出力音声信号 $x_{out}(t)$ のリサンプリングが必要であったのに対し、音声符号化装置 30' は基準周波数信号 V_s が常に一定であるためリサンプリングが必要ない。そのため、装置構成を簡略化し処理時間を高速化することができる。

30

【0187】

(2) 実施例 3 の音声符号化装置 30 では、ピッチ情報は、基準周期情報(基準周波数信号 AV_{pitch})と残差周波数情報(残差周波数信号 V_{pitch})とに分離され、それぞれの情報について符号化が行われていたのに対し、音声符号化装置 30' では、基準周期情報は残差周波数情報(残差周波数信号 V_{pitch})の中に取り込まれ、残差周波数情報のみを符号化している。このように基準周期情報(すなわち、平均ピッチ周波数の時間変化情報)と残差周波数情報とを分離しない場合、残差周波数信号 V_{pitch} のレンジは実施例 3 の場合に比べて多少大きくなる。しかしながら、平均ピッチ周波数の時間変化は小さいため、残差周波数信号 V_{pitch} のレンジが若干大きくなったとしても残差周波数信号 V_{pitch} は依然狭レンジの信号であるため、符号化効率が極端に低下することはない。従

40

【0188】

(3) 音声符号化装置 30' では、各ピッチ区間のピッチ周期を一定の基準周期に強制的に等化するため、場合によっては入力音声信号 $x_{in}(t)$ のピッチ周期と基準周期との差が大きい場合がある。かかる場合、等化によって若干の歪みが生じる場合がある。従って、実施例 3 の音声符号化装置 30 に比べ、符号化による SN の低下は多少大きくなる。

【実施例 8】

【0189】

図 15 は、本発明の実施例 8 に係る音声復号装置 50' の構成を表すブロック図である。音声復号装置 50' は、実施例 7 の音声符号化装置 30' により符号化された音声信号

50

を復号する装置である。音声復号装置 50' は、ピッチ等化波形復号器 51、逆量子化器 52、シンセサイザ 53、ピッチ情報復号器 54'、及び周波数シフタ 58 を備えている。このうち、実施例 4 と同様のものについては同符号を付してある。

【0190】

音声復号装置 50' には、符号化波形データ及び符号化ピッチデータが入力される。符号化波形データは、図 14 のピッチ等化波形符号化器 34 から出力される符号化波形データである。符号化ピッチデータは、図 14 のピッチ情報符号化器 36' から出力される符号化ピッチデータである。

【0191】

本実施例の音声復号装置 50' は実施例 4 の音声復号装置 50 に対してピッチ周波数検出手段 55、差分器 56、及び加算器 57 が省略されている。ピッチ情報復号器 54' は、符号化ピッチデータを復号することにより、残差周波数信号 V_{pitch} を復元する。周波数シフタ 58 は、シンセサイザ 53 が出力する等化音声信号 $x_{eq}(t)$ の各ピッチ区間のピッチ周波数を当該ピッチ周波数に残差周波数信号 V_{pitch} を加えたものに変換し、音声信号 $x_{res}(t)$ として復元する。その他の動作に関しては実施例 4 と同様である。

【0192】

尚、実施例 1 ~ 8 までのピッチ周期等化装置 1, 1'、音声符号化装置 30, 30'、及び音声復号装置 50, 50' はハードウェア的に構成する例を示したが、各機能ブロックをプログラムとして構成し、コンピュータに実行させることによってコンピュータを各装置として機能させる構成としてもよい。

【図面の簡単な説明】

【0193】

【図 1】本発明の実施例 1 に係るピッチ周期等化装置 1 の構成を表すブロック図である。

【図 2】ピッチ検出手段 11 における信号処理の概略を説明する図である。

【図 3】周波数シフタ 4 の内部構成を表す図である。

【図 4】周波数シフタ 4 の内部構成の他の例を表す図である。

【図 5】有声音「あ」のフォルマント特性を示す図である。

【図 6】無声音「す」の自己相関及びケプストラム波形並びに周波数特性を示す図である。

【図 7】本発明の実施例 2 に係るピッチ周期等化装置 1' の構成を表す図である。

【図 8】本発明の実施例 3 に係る音声符号化装置 30 の構成を表す図である。

【図 9】量子化ビット数についての説明図である。

【図 10】各サブバンドのスペクトル強度の時間変化の例である。

【図 11】本発明の実施例 4 に係る音声復号装置 50 の構成を表すブロック図である。

【図 12】本発明の実施例 5 に係るピッチ周期等化装置 41 の構成を表す図である。

【図 13】本発明の実施例 6 に係るピッチ周期等化装置 41' の構成を表す図である。

【図 14】本発明の実施例 7 に係る音声符号化装置 30' の構成を表す図である。

【図 15】本発明の実施例 8 に係る音声復号装置 50' の構成を表すブロック図である。

【図 16】CELP 符号化方式による音声符号化装置の基本構成例を表す図である。

【図 17】CELP 符号化方式による音声復号装置の基本構成例を表す図である。

【図 18】特許文献 2 記載の音声符号化装置の構成例を表す図である。

【符号の説明】

【0194】

- 1, 1' ピッチ周期等化装置
- 2 入力ピッチ検出手段
- 3 ピッチ平均手段
- 4 周波数シフタ
- 5, 5' 出力ピッチ検出手段
- 6 残差演算手段
- 7 PID コントローラ

10

20

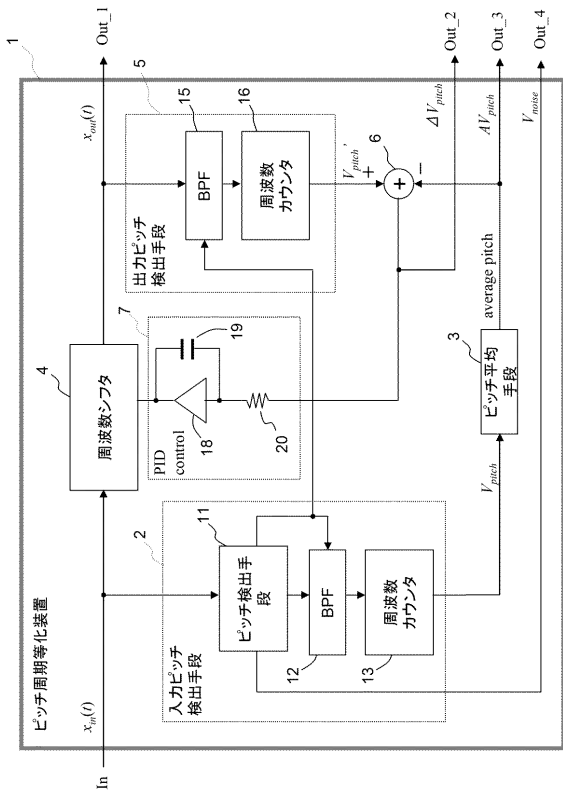
30

40

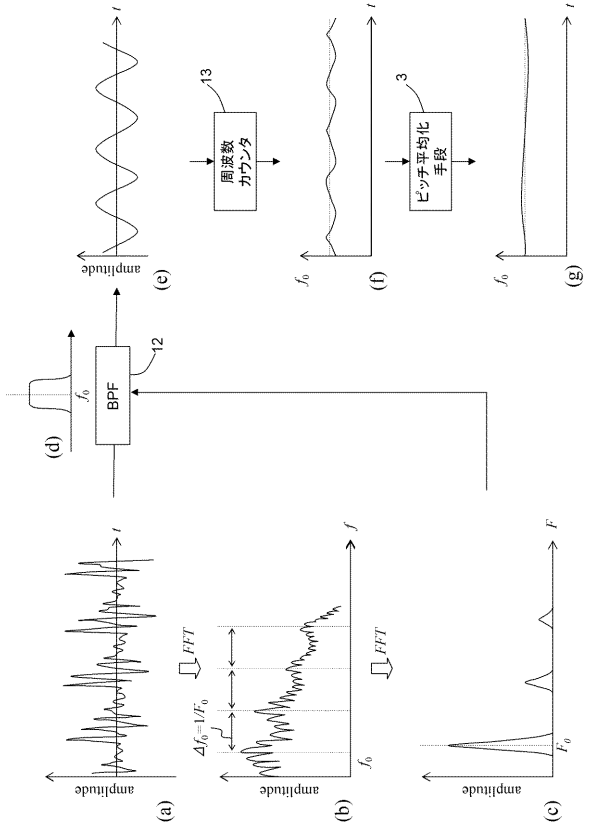
50

1 1	ピッチ検出手段	
1 2 , 1 5	バンドパスフィルタ (B P F)	
1 3	周波数カウンタ	
1 6	周波数カウンタ	
1 8	アンプ	
1 9	コンデンサ	
2 0	抵抗	
2 1	発信器	
2 2	変調器	
2 3	B P F	10
2 4	電圧制御発信器 (V C O)	
2 5	復調器	
3 0 , 3 0 '	音声符号化装置	
3 1	リサンブラ	
3 2	アナライザ	
3 3	量子化器	
3 4	ピッチ等化波形符号化器	
3 5	差分ビット演算器	
3 6 , 3 6 '	ピッチ情報符号化器	
4 1 , 4 1 '	ピッチ周期等化装置	20
4 2	基準周波数発生器	
5 0 , 5 0 '	音声復号装置	
5 1	ピッチ等化波形復号器	
5 2	逆量子化器	
5 3	シンセサイザ	
5 4 , 5 4 '	ピッチ情報復号器	
5 5	ピッチ周波数検出手段	
5 6	差分器	
5 7	加算器	
5 8	周波数シフタ	30

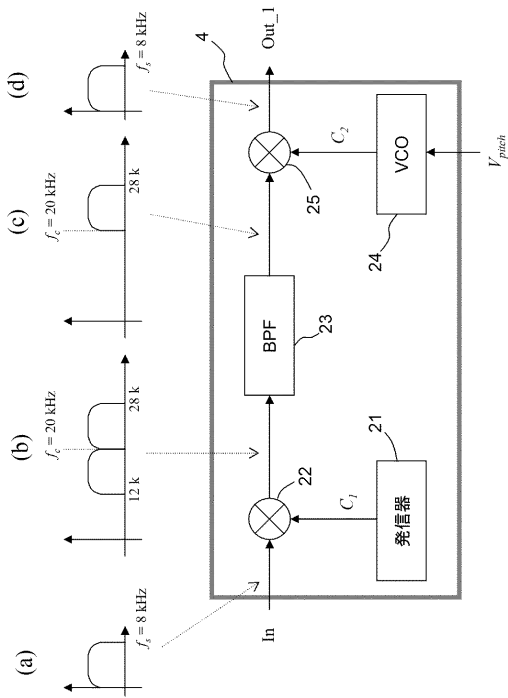
【図1】



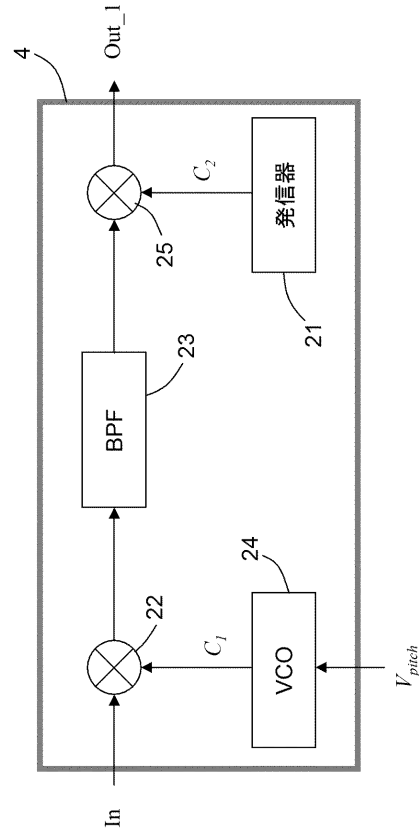
【図2】



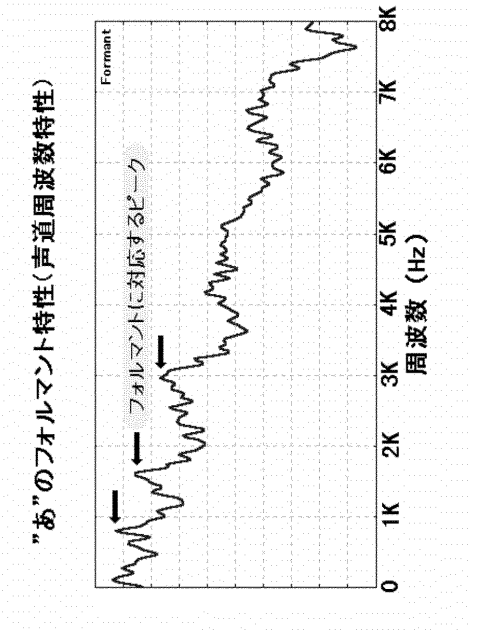
【図3】



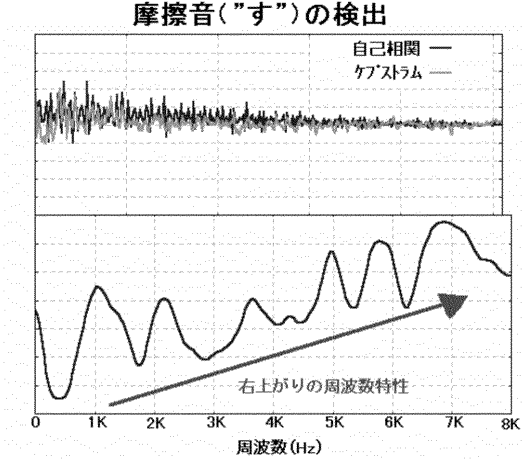
【図4】



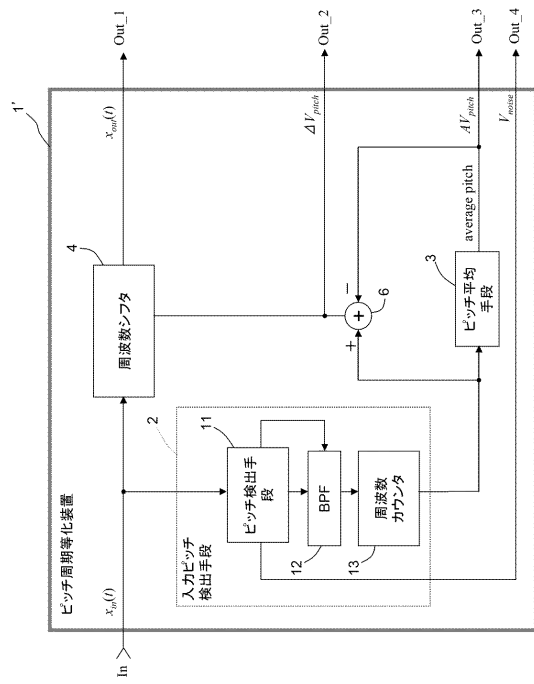
【図5】



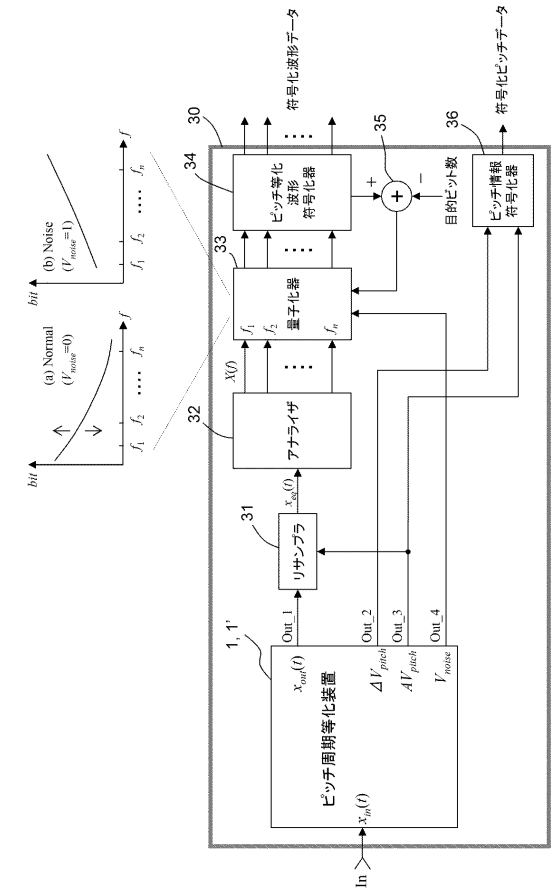
【図6】



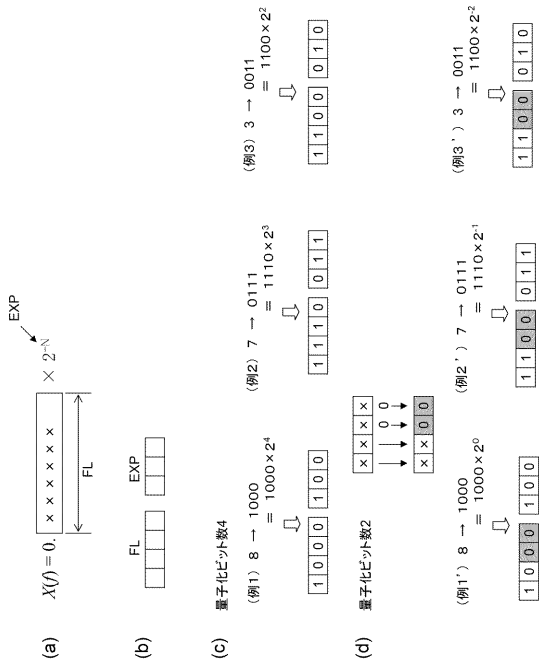
【図7】



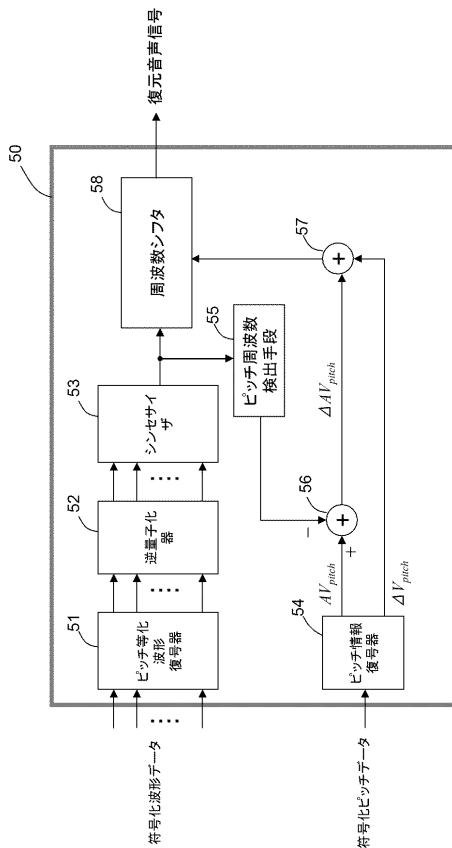
【図8】



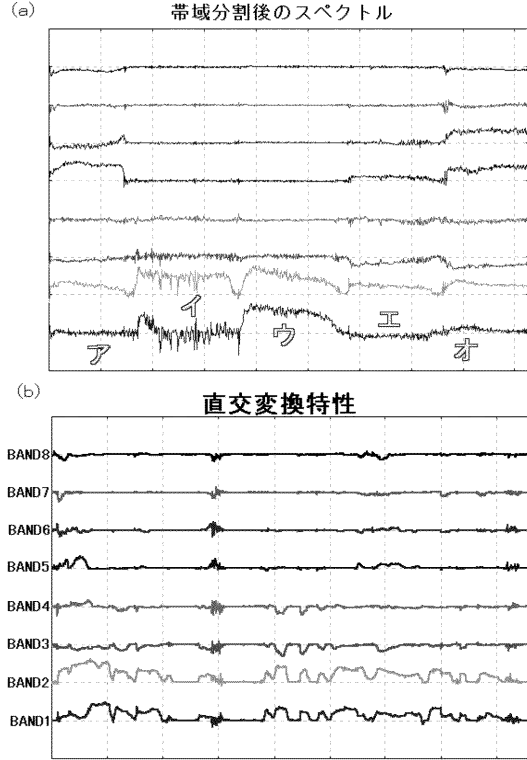
【図9】



【図11】

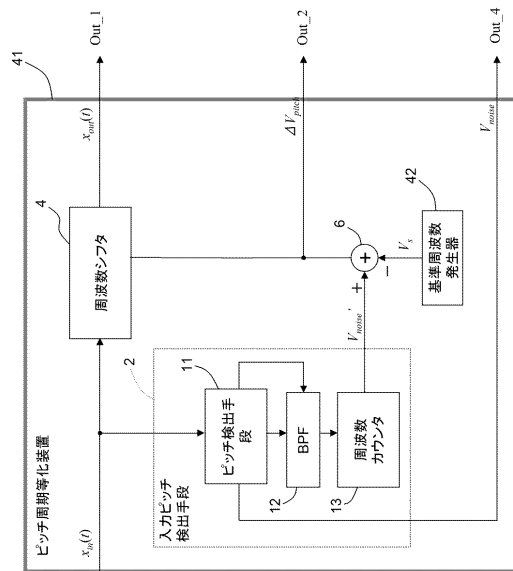


【図10】

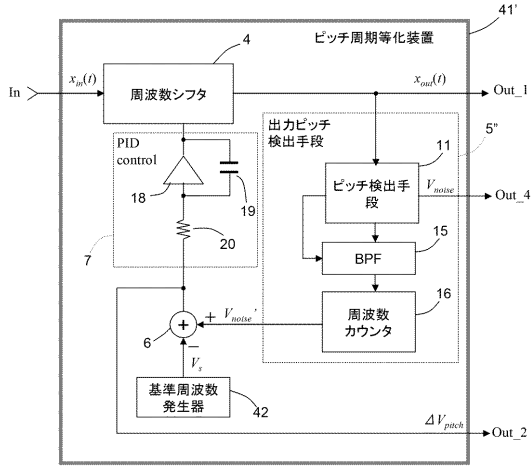


アラユル ゲンジツ ラ スペテ ジボンノホウヘ ネジマゲタノダ

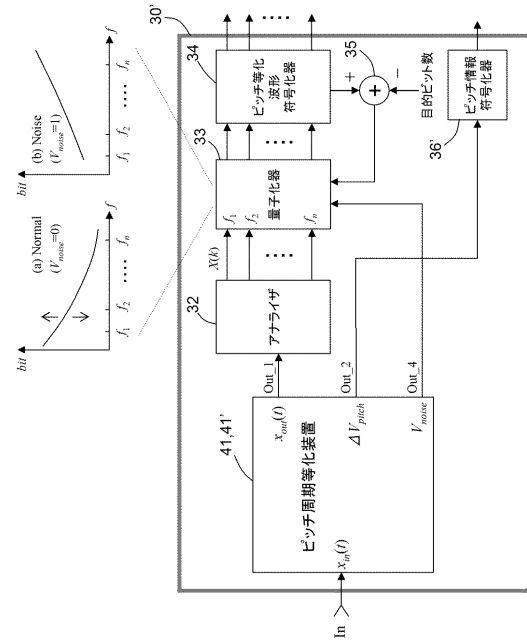
【図12】



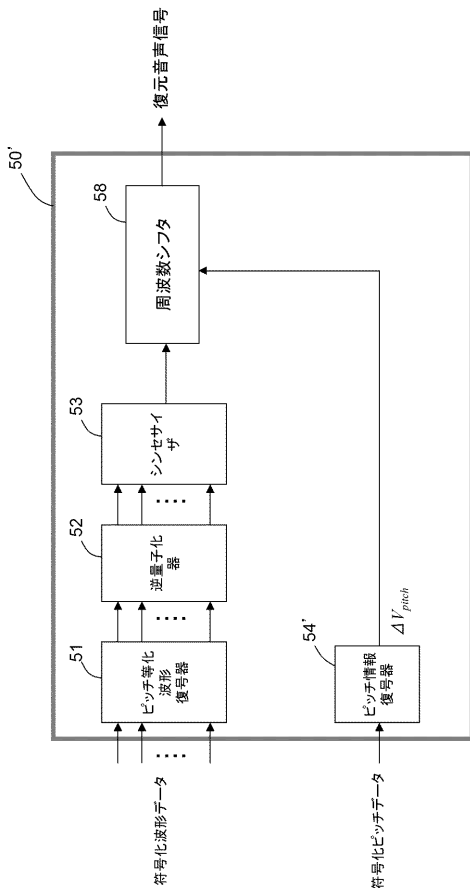
【図13】



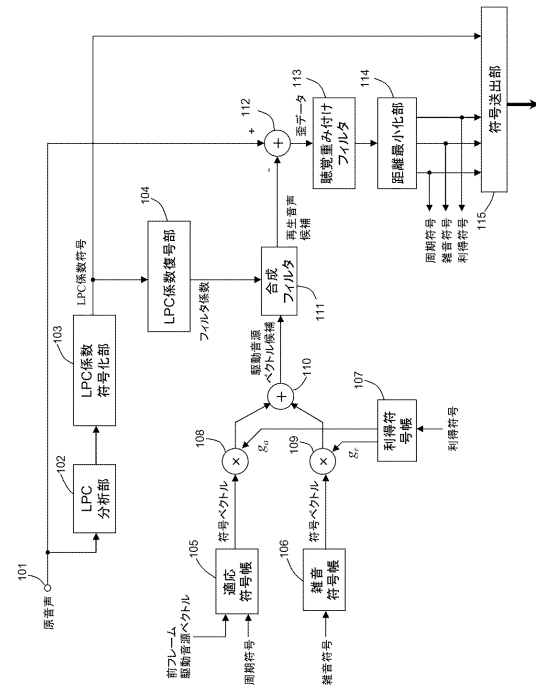
【図14】



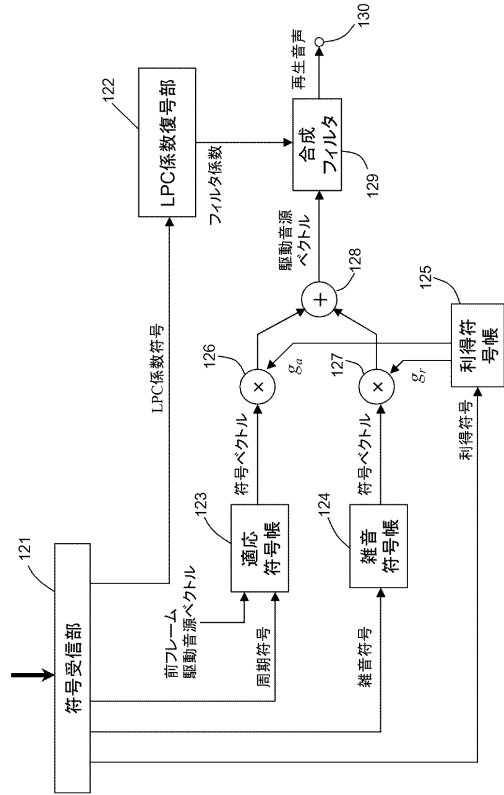
【図15】



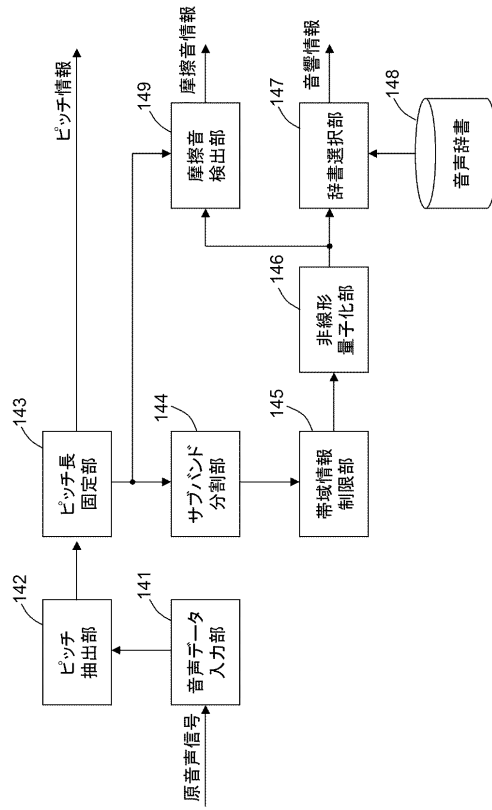
【図16】



【図 17】



【図 18】



フロントページの続き

- (56)参考文献 特開平 8 - 2 0 2 3 9 5 (J P , A)
特開平 3 - 8 0 3 0 0 (J P , A)
望月亮他, " 駆動点同期型ピッチ波形抽出法 ", 日本音響学会誌, Vol.53, No.10 (1997-10),
pp.772-778
新居康彦他, " ピッチ波形抽出位置の検討 ", 電子情報通信学会技術研究報告, Vol.95, No.41
, SP95-8 (1995-05), pp.53-59
- (58)調査した分野(Int.Cl., DB名)
G 1 0 L 1 1 / 0 0 - 2 1 / 0 6
H 0 4 L 2 7 / 0 2