(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号

特許第4132109号
(P4132109)

(45)発行日　平成20年8月13日(2008.8.13)　(24)登録日　平成20年6月6日(2008.6.6)

(51)Int.Cl.		F I		
G10L 19/08	(2006.01)	G10L	19/08	D
G10L 19/14	(2006.01)	G10L	19/14	400B
H03M 7/30	(2006.01)	G10L	19/14	400F
		H03M	7/30	B

請求項の数 4　(全 25 頁)

(21)出願番号　特願平8-270337
(22)出願日　平成8年10月11日(1996.10.11)
(65)公開番号　特開平9-190196
(43)公開日　平成9年7月22日(1997.7.22)
審査請求日　平成15年3月7日(2003.3.7)
審判番号　不服2005-15497(P2005-15497/J1)
審判請求日　平成17年8月11日(2005.8.11)
(31)優先権主張番号　特願平7-279410
(32)優先日　平成7年10月26日(1995.10.26)
(33)優先権主張国　日本国(JP)
(31)優先権主張番号　特願平7-280672
(32)優先日　平成7年10月27日(1995.10.27)
(33)優先権主張国　日本国(JP)

(73)特許権者　000002185
ソニー株式会社
東京都港区港南１丁目７番１号
(74)代理人　100067736
弁理士　小池　晃
(74)代理人　100086335
弁理士　田村　榮一
(74)代理人　100096677
弁理士　伊賀　誠司
(72)発明者　飯島　和幸
東京都品川区北品川６丁目７番３５号　ソニー株式会社内
(72)発明者　西口　正之
東京都品川区北品川６丁目７番３５号　ソニー株式会社内

最終頁に続く

(54)【発明の名称】音声信号の再生方法及び装置、並びに音声復号化方法及び装置、並びに音声合成方法及び装置

(57)【特許請求の範囲】

【請求項１】

入力音声信号を時間軸上で、所定の符号化単位で区分し、上記入力音声信号が有声音か無声音かを判別し、その判別結果に基づいて、有声音とされた部分ではサイン波合成符号化を行い、無声音とされた部分では合成による分析法を用いて最適ベクトルのクローズドループサーチによるベクトル量子化を行うことで求めた符号化パラメータに基づいて音声信号を再生する音声信号の再生方法であって、

上記符号化パラメータを補間処理して、時間軸圧縮または伸張した音声信号を求めるために、所望の時刻に対応する変更符号化パラメータを求め、

この変更符号化パラメータに基づいて有声音の場合には有声音用の合成フィルタにより有声音を合成し、無声音の場合には上記符号化パラメータを用いて無声音用の合成フィルタにより無声音を合成する際の線形予測残差に相当するノイズ信号成分を生成し、上記無声音用の合成フィルタにより無声音を合成し、上記それぞれの合成フィルタからの出力を加算して音声信号を再生するようにし、

無声音から無声音になる場合には、補間前の上記符号化パラメータを用いて上記無声音用の合成フィルタにより無声音を合成する際の線形予測残差に相当するノイズ信号成分を生成し、このノイズ信号成分のサンプルに対して補間位置を中心とする前後の範囲で上記所定の符号化単位のサンプルを切り出して変更符号化パラメータを作る

音声信号の再生方法。

【請求項２】

上記変更符号化パラメータの補間は、有声音から無声音または無声音から有声音に変わる場合には、上記所定の符号化単位毎に得られる符号化パラメータの内、補間位置に近い方の符号化パラメータを上記変更符号化パラメータとする請求項１記載の音声信号の再生方法。

【請求項３】

入力音声信号を時間軸上で、所定の符号化単位で区分し、上記入力音声信号が有声音か無声音かを判別し、その判別結果に基づいて、有声音とされた部分ではサイン波合成符号化を行い、無声音とされた部分では合成による分析法を用いて最適ベクトルのクローズドループサーチによるベクトル量子化を行うことで求めた符号化パラメータに基づいて音声信号を再生する音声信号の再生装置であって、

上記符号化パラメータを補間処理して、時間軸圧縮または伸張した音声信号を求めるために、所望の時刻に対応する変更符号化パラメータを求め、

この変更符号化パラメータに基づいて有声音の場合には有声音用の合成フィルタにより有声音を合成し、無声音の場合には上記符号化パラメータを用いて無声音用の合成フィルタにより無声音を合成する際の線形予測残差に相当するノイズ信号成分を生成し、上記無声音用の合成フィルタにより無声音を合成し、上記それぞれの合成フィルタからの出力を加算して音声信号を再生するようにし、

無声音から無声音になる場合には、補間前の上記符号化パラメータを用いて上記無声音を合成する際の線形予測残差に相当する上記ノイズ信号成分を生成し、このノイズ信号成分のサンプルに対して補間位置を中心とする前後の範囲で上記所定の符号化単位のサンプルを切り出して変更符号化パラメータを作る

音声信号の再生装置。

【請求項４】

上記変更符号化パラメータの補間は、有声音から無声音または無声音から有声音に変わる場合には、上記所定の符号化単位毎に得られる符号化パラメータの内、補間位置に近い方の符号化パラメータを上記変更符号化パラメータとする請求項３記載の音声信号の再生装置。

【発明の詳細な説明】

【０００１】

【発明の属する技術分野】

本発明は、音声信号をスピードコントロールして再生する音声信号の再生方法及び装置に関する。

【０００２】

また、本発明は、簡単な構成でピッチ変換が行えるような音声復号化方法及び装置並びに音声合成方法及び装置に関する。

【０００４】

【従来の技術】

オーディオ信号（音声信号や音響信号を含む）の時間領域や周波数領域における統計的性質と人間の聴感上の特性を利用して信号圧縮を行うような符号化方法が種々知られている。この符号化方法としては、大別して時間領域での符号化、周波数領域での符号化、分析合成符号化等が挙げられる。

【０００５】

音声信号等の高能率符号化の例として、ＭＢＥ（Multiband Excitation: マルチバンド励起）符号化、ＳＢＥ（Singleband Excitation:シングルバンド励起）あるいはサイン波合成符号化、ハーモニック（Harmonic）符号化、ＳＢＣ（Sub-band Coding:帯域分割符号化）、ＬＰＣ（Linear Predictive Coding: 線形予測符号化）、あるいはＤＣＴ（離散コサイン変換）、ＭＤＣＴ（モデファイドＤＣＴ）、ＦＦＴ（高速フーリエ変換）等が知られている。

【０００６】

【発明が解決しようとする課題】

ところで、符号励起線形予測（ＣＥＬＰ：Code Excited Linear Prediction ）符号化に代表されるような上記時間軸上の処理による音声高能率符号化方法では、時間軸のスピード変換（Modify）処理が困難であった。これは、復号化（デコーダ）出力の後にかなりの演算を行う必要があったためである。また、デコードした時間領域でスピードコントロールを行うため、例えばビットレートの変換などには使えなかった。

【０００７】

また、上記各種符号化方法で符号化された音声信号を復号化しようとする場合に、音声の音韻を変えずにピッチだけを可変としたいことがあるが、通常の音声復号化装置では、復号化された音声をピッチコントロールを用いてピッチ変換しなければならず、構成が複雑化し、価格が高騰するという欠点があった。

【０００８】

本発明は、上記実情に鑑みてなされたものであり、広いレンジにわたる任意のレートのスピードコントロールを、音韻，ピッチを不変として高品質に行うことのできる音声信号の再生方法及び装置の提供を目的とする。

【０００９】

また、本発明は、簡単な構成でピッチ変換あるいはピッチコントロールが行えるような音声復号化方法及び装置、並びに音声合成方法及び装置の提供を目的とする。

【００１１】

【課題を解決するための手段】

本発明に係る音声信号の再生方法は、入力音声信号を時間軸上で、所定の符号化単位で区分し、上記入力音声信号が有声音か無声音かを判別し、その判別結果に基づいて、有声音とされた部分ではサイン波合成符号化を行い、無声音とされた部分では合成による分析法を用いて最適ベクトルのクローズドループサーチによるベクトル量子化を行うことで求めた符号化パラメータに基づいて音声信号を再生する音声信号の再生方法であって、上記符号化パラメータを補間処理して、時間軸圧縮または伸張した音声信号を求めるために、所望の時刻に対応する変更符号化パラメータを求め、この変更符号化パラメータに基づいて有声音の場合には有声音用の合成フィルタにより有声音を合成し、無声音の場合には上記符号化パラメータを用いて無声音用の合成フィルタにより無声音を合成する際の線形予測残差に相当するノイズ信号成分を生成し、上記無声音用の合成フィルタにより無声音を合成し、上記それぞれの合成フィルタからの出力を加算して音声信号を再生するようにし、無声音から無声音になる場合には、補間前の上記符号化パラメータを用いて上記無声音用の合成フィルタにより無声音を合成する際の線形予測残差に相当するノイズ信号成分を生成し、このノイズ信号成分のサンプルに対して補間位置を中心とする前後の範囲で上記所定の符号化単位のサンプルを切り出して変更符号化パラメータを作る。

【００１２】

また、本発明に係る音声信号の再生装置は、入力音声信号を時間軸上で、所定の符号化単位で区分し、上記入力音声信号が有声音か無声音かを判別し、その判別結果に基づいて、有声音とされた部分ではサイン波合成符号化を行い、無声音とされた部分では合成による分析法を用いて最適ベクトルのクローズドループサーチによるベクトル量子化を行うことで求めた符号化パラメータに基づいて音声信号を再生する音声信号の再生装置であって、上記符号化パラメータを補間処理して、時間軸圧縮または伸張した音声信号を求めるために、所望の時刻に対応する変更符号化パラメータを求め、この変更符号化パラメータに基づいて有声音の場合には有声音用の合成フィルタにより有声音を合成し、無声音の場合には上記符号化パラメータを用いて無声音用の合成フィルタにより無声音を合成する際の線形予測残差に相当するノイズ信号成分を生成し、上記無声音用の合成フィルタにより無声音を合成し、上記それぞれの合成フィルタからの出力を加算して音声信号を再生するようにし、無声音から無声音になる場合には、補間前の上記符号化パラメータを用いて上記無声音用の合成フィルタにより無声音を合成する際の線形予測残差に相当するノイズ信号成分を生成し、このノイズ信号成分のサンプルに対して補間位置を中心とする前後の範囲で上記所定の符号化単位のサンプルを切り出して変更符号化パラメータを作る。

【００１７】
【発明の実施の形態】
以下、先ず始めに本発明に係る音声信号の再生方法及び装置の実施例について図面を参照しながら説明する。この実施例は、入力音声信号が時間軸上で所定フレーム単位毎に区分されて符号化されることにより求められた符号化パラメータに基づいて音声信号を再生する図１の音声信号再生装置１である。
【００１８】
この音声信号再生装置１は、入力端子１０１から入力された音声信号をフレーム単位で符号化して、線スペクトル対（ＬＳＰ）パラメータや、ピッチや、有声音（Ｖ）／無声音（ＵＶ）や、スペクトル振幅Ａｍや、ＬＰＣ（線形予測符号化）残差のような符号化パラメータを出力する符号化部２と、この符号化部２からの上記符号化パラメータの出力周期を時間軸を圧縮伸長して変更する周期変更部３と、この周期変更部３により変更された周期で出力された上記符号化パラメータを補間処理して所望の時刻に対応する変更符号化パラメータを求め、この変更符号化パラメータに基づいて音声信号を合成して出力端子２０１から出力する復号化部４とを備えてなる。
【００１９】
先ず、符号化部２について図２及び図３を参照しながら説明する。符号化部２は、上記入力音声信号が有声音か無声音かを判別し、その判別結果に基づいて、有声音とされた部分ではサイン波合成符号化を行い、無声音とされた部分では合成による分析法を用いた最適ベクトルのクローズドループサーチによるベクトル量子化を行って符号化パラメータを求めている。つまり、入力音声信号の短期予測残差例えばＬＰＣ（線形予測符号化）残差を求めてサイン波分析（sinusoidal analysis ）符号化、例えばハーモニックコーディング（harmonic coding ）を行う第１の符号化部１１０と、入力音声信号に対して位相伝送を行う波形符号化により符号化する第２の符号化部１２０とを有し、入力信号の有声音（Ｖ：Voiced）の部分の符号化には第１の符号化部１１０を用い、入力信号の無声音（ＵＶ：Unvoiced）の部分の符号化には第２の符号化部１２０を用いる。
【００２０】
図２の例では、入力端子１０１に供給された音声信号が、第１の符号化部１１０のＬＰＣ逆フィルタ１１１及びＬＰＣ分析・量子化部１１３に送られている。ＬＰＣ分析・量子化部１１３で得られたＬＰＣ係数あるいはいわゆるαパラメータは、ＬＰＣ逆フィルタ１１１に送られて、このＬＰＣ逆フィルタ１１１により入力音声信号の線形予測残差（ＬＰＣ残差）が取り出される。また、ＬＰＣ分析・量子化部１１３からは、後述するようにＬＳＰ（線スペクトル対）の量子化出力が取り出され、これが出力端子１０２に送られる。ＬＰＣ逆フィルタ１１１からのＬＰＣ残差は、サイン波分析符号化部１１４に送られる。サイン波分析符号化部１１４では、ピッチ検出やスペクトルエンベロープ振幅計算が行われると共に、Ｖ（有声音）／ＵＶ（無声音）判定部１１５によりＶ／ＵＶの判定が行われる。サイン波分析符号化部１１４からのスペクトルエンベロープ振幅データがベクトル量子化部１１６に送られる。スペクトルエンベロープのベクトル量子化出力としてのベクトル量子化部１１６からのコードブックインデクスは、スイッチ１１７を介して出力端子１０３に送られ、サイン波分析符号化部１１４からのピッチ出力は、スイッチ１１８を介して出力端子１０４に送られる。また、Ｖ／ＵＶ判定部１１５からのＶ／ＵＶ判定出力は、出力端子１０５に送られると共に、スイッチ１１７、１１８の制御信号に使われる。スイッチ１１７、１１８は、上記制御信号により有声音（Ｖ）のとき上記インデクス及びピッチを選択して各出力端子１０３及び１０４からそれぞれ出力する。
【００２１】
また、上記ベクトル量子化部１１６でのベクトル量子化の際には、例えば、周波数軸上の有効帯域１ブロック分の振幅データに対して、ブロック内の最後のデータからブロック内の最初のデータまでの値を補間するようなダミーデータ，又は最後のデータ及び最初のデータを延長するようなダミーデータを最後と最初に適当な数だけ付加してデータ個数をN_F個に拡大した後、帯域制限型のO_S倍（例えば８倍）のオーバーサンプリングを施すこ

とによりＯ$_S$ 倍の個数の振幅データを求め、このＯ$_S$ 倍の個数（（m_{MX}＋１）×Ｏ$_S$ 個）の振幅データを直線補間してさらに多くのＮ$_M$ 個（例えば２０４８個）に拡張し、このＮ$_M$ 個のデータを間引いて上記一定個数Ｍ（例えば４４個）のデータに変換した後、ベクトル量子化している。

【００２２】

第２の符号化部１２０は、この例ではＣＥＬＰ（符号励起線形予測）符号化構成を有しており、合成による分析（Analysis by Synthesis ）法を用いたクローズドループサーチを用いた時間軸波形のベクトル量子化を行っている。具体的に、雑音符号帳１２１からの出力を、重み付きの合成フィルタ１２２により合成処理し、得られた重み付き合成音声を減算器１２３に送り、入力端子１０１に供給された音声信号の聴覚重み付けフィルタ１２５による音声との誤差を求める。距離計算回路１２４は、距離計算を行い、上記誤差が最小となるようなベクトルを雑音符号帳１２１でサーチする。このＣＥＬＰ符号化は、上述したように無声音部分の符号化に用いられており、雑音符号帳１２１からのＵＶデータとしてのコードブックインデクスは、上記Ｖ／ＵＶ判定部１１５からのＶ／ＵＶ判定結果が無声音（ＵＶ）のときオンとなるスイッチ１２７を介して、出力端子１０７より取り出される。

【００２３】

次に、上記図２に示した符号化部２のより具体的な構成について、図３を参照しながら説明する。なお、図３において、上記図２の各部と対応する部分には同じ指示符号を付している。

【００２４】

この図３に示された符号化部２において、入力端子１０１に供給された音声信号は、ハイパスフィルタ（ＨＰＦ）１０９にて不要な帯域の信号を除去するフィルタ処理が施された後、ＬＰＣ（線形予測符号化）分析・量子化部１１３のＬＰＣ分析回路１３２と、ＬＰＣ逆フィルタ回路１１１とに送られる。

【００２５】

ＬＰＣ分析・量子化部１１３のＬＰＣ分析回路１３２は、入力信号波形の２５６サンプル程度の長さを１ブロックとしてハミング窓をかけて、自己相関法により線形予測係数、いわゆるαパラメータを求める。データ出力の単位となるフレーミングの間隔は、１６０サンプル程度とする。サンプリング周波数ｆｓが例えば８ｋHzのとき、１フレーム間隔は１６０サンプルで２０msec となる。

【００２６】

ＬＰＣ分析回路１３２からのαパラメータは、α→ＬＳＰ変換回路１３３に送られて、線スペクトル対（ＬＳＰ）パラメータに変換される。これは、直接型のフィルタ係数として求まったαパラメータを、例えば１０個、すなわち５対のＬＳＰパラメータに変換する。変換は例えばニュートン－ラプソン法等を用いて行う。このＬＳＰパラメータに変換するのは、αパラメータよりも補間特性に優れているからである。

【００２７】

α→ＬＳＰ変換回路１３３からのＬＳＰパラメータは、ＬＳＰベクトル量子化器１３４によりベクトル量子化される。このとき、フレーム間差分をとってからベクトル量子化してもよく、複数フレーム分をまとめてマトリクス量子化してもよい。ここでは、２０msec を１フレームとし、２０msec 毎に算出されるＬＳＰパラメータをベクトル量子化している。

【００２８】

このＬＳＰベクトル量子化器１３４からの量子化出力、すなわちＬＳＰ量子化のインデクスは、端子１０２を介して復号化部３に取り出され、また量子化済みのＬＳＰベクトルは、ＬＳＰ補間回路１３６に送られる。

【００２９】

ＬＳＰ補間回路１３６は、上記２０msecあるいは４０msec毎に量子化されたＬＳＰのベクトルを補間し、８倍のレートにする。すなわち、２．５msec 毎にＬＳＰベクトルが更

新されるようにする。これは、残差波形をハーモニック符号化復号化方法により分析合成すると、その合成波形のエンベロープは非常になだらかでスムーズな波形になるため、ＬＰＣ係数が２０msec 毎に急激に変化すると異音を発生することがあるからである。すなわち、２．５msec 毎にＬＰＣ係数が徐々に変化してゆくようにすれば、このような異音の発生を防ぐことができる。

【００３０】

このような補間が行われた２．５msec 毎のＬＳＰベクトルを用いて入力音声の逆フィルタリングを実行するために、ＬＳＰ→α変換回路１３７により、ＬＳＰパラメータを例えば１０次程度の直接型フィルタの係数であるαパラメータに変換する。このＬＳＰ→α変換回路１３７からの出力は、上記ＬＰＣ逆フィルタ回路１１１に送られ、このＬＰＣ逆フィルタ１１１では、２．５msec 毎に更新されるαパラメータにより逆フィルタリング処理を行って、滑らかな出力を得るようにしている。このＬＰＣ逆フィルタ１１１からの出力は、サイン波分析符号化部１１４、具体的には例えばハーモニック符号化回路、の直交変換回路１４５、例えばＤＦＴ（離散フーリエ変換）回路に送られる。

【００３１】

ＬＰＣ分析・量子化部１１３のＬＰＣ分析回路１３２からのαパラメータは、聴覚重み付けフィルタ算出回路１３９にも送られて聴覚重み付けのためのデータが求められ、この重み付けデータが後述する聴覚重み付きのベクトル量子化器１１６と、第２の符号化部１２０の聴覚重み付けフィルタ１２５及び聴覚重み付きの合成フィルタ１２２とに送られる。

【００３２】

ハーモニック符号化回路等のサイン波分析符号化部１１４では、ＬＰＣ逆フィルタ１１１からの出力を、ハーモニック符号化の方法で分析する。すなわち、ピッチ検出、各ハーモニクスの振幅Ａmの算出、有声音（Ｖ）／無声音（ＵＶ）の判別を行い、ピッチによって変化するハーモニクスのエンベロープあるいは振幅Ａmの個数を次元変換して一定数にしている。

【００３３】

図３に示すサイン波分析符号化部１１４の具体例においては、一般のハーモニック符号化を想定しているが、特に、ＭＢＥ（Multiband Excitation: マルチバンド励起）符号化の場合には、同時刻（同じブロックあるいはフレーム内）の周波数軸領域いわゆるバンド毎に有声音（Voiced）部分と無声音（Unvoiced）部分とが存在するという仮定でモデル化することになる。それ以外のハーモニック符号化では、１ブロックあるいはフレーム内の音声が有声音か無声音かの択一的な判定がなされることになる。なお、以下の説明中のフレーム毎のＶ／ＵＶとは、ＭＢＥ符号化に適用した場合には全バンドがＵＶのときを当該フレームのＵＶとしている。

【００３４】

図３のサイン波分析符号化部１１４のオープンループピッチサーチ部１４１には、上記入力端子１０１からの入力音声信号が、またゼロクロスカウンタ１４２には、上記ＨＰＦ（ハイパスフィルタ）１０９からの信号がそれぞれ供給されている。サイン波分析符号化部１１４の直交変換回路１４５には、ＬＰＣ逆フィルタ１１１からのＬＰＣ残差あるいは線形予測残差が供給されている。オープンループピッチサーチ部１４１では、入力信号のＬＰＣ残差をとってオープンループによる比較的ラフなピッチのサーチが行われ、抽出された粗ピッチデータは高精度ピッチサーチ１４６に送られて、後述するようなクローズドループによる高精度のピッチサーチ（ピッチのファインサーチ）が行われる。また、オープンループピッチサーチ部１４１からは、上記粗ピッチデータと共にＬＰＣ残差の自己相関の最大値をパワーで正規化した正規化自己相関最大値ｒ(p) が取り出され、Ｖ／ＵＶ（有声音／無声音）判定部１１５に送られている。

【００３５】

直交変換回路１４５では例えばＤＦＴ（離散フーリエ変換）等の直交変換処理が施されて、時間軸上のＬＰＣ残差が周波数軸上のスペクトル振幅データに変換される。この直交変換回路１４５からの出力は、高精度ピッチサーチ部１４６及びスペクトル振幅あるいはエ

ンベロープを評価するためのスペクトル評価部１４８に送られる。

【００３６】

高精度（ファイン）ピッチサーチ部１４６には、オープンループピッチサーチ部１４１で抽出された比較的ラフな粗ピッチデータと、直交変換部１４５により例えばＤＦＴされた周波数軸上のデータとが供給されている。この高精度ピッチサーチ部１４６では、上記粗ピッチデータ値を中心に、0.2～0.5きざみで±数サンプルずつ振って、最適な小数点付き（フローティング）のファインピッチデータの値へ追い込む。このときのファインサーチの手法として、いわゆる合成による分析（Analysis by Synthesis)法を用い、合成されたパワースペクトルが原音のパワースペクトルに最も近くなるようにピッチを選んでいる。このようなクローズドループによる高精度のピッチサーチ部１４６からのピッチデータについては、スイッチ１１８を介して出力端子１０４に送っている。

【００３７】

スペクトル評価部１４８では、ＬＰＣ残差の直交変換出力としてのスペクトル振幅及びピッチに基づいて、各ハーモニクスの大きさ及びその集合であるスペクトルエンベロープが評価され、高精度ピッチサーチ部１４６、Ｖ／ＵＶ（有声音／無声音）判定部１１５及び聴覚重み付きのベクトル量子化器１１６に送られる。

【００３８】

Ｖ／ＵＶ（有声音／無声音）判定部１１５は、直交変換回路１４５からの出力と、高精度ピッチサーチ部１４６からの最適ピッチと、スペクトル評価部１４８からのスペクトル振幅データと、オープンループピッチサーチ部１４１からの正規化自己相関最大値ｒ(p) と、ゼロクロスカウンタ１４２からのゼロクロスカウント値及び、該当フレームのr.m.sであるlevに基づいて、当該フレームのＶ／ＵＶ判定が行われる。さらに、ＭＢＥの場合の各バンド毎のＶ／ＵＶ判定結果の境界位置も当該フレームのＶ／ＵＶ判定の一条件としてもよい。このＶ／ＵＶ判定部１１５からの判定出力は、出力端子１０５を介して取り出される。

【００３９】

ところで、スペクトル評価部１４８の出力部あるいはベクトル量子化器１１６の入力部には、データ数変換（一種のサンプリングレート変換）部が設けられている。このデータ数変換部は、上記ピッチに応じて周波数軸上での分割帯域数が異なり、データ数が異なることを考慮して、エンベロープの振幅データ｜A_m｜ を一定の個数にするためのものである。すなわち、例えば有効帯域を３４００kHzまでとすると、この有効帯域が上記ピッチに応じて、８バンド～６３バンドに分割されることになり、これらの各バンド毎に得られる上記振幅データ｜A_m｜ の個数m_{MX}＋１も８～６３と変化することになる。このためデータ数変換部では、この可変個数m_{MX}＋１の振幅データを一定個数Ｍ個、例えば４４個、のデータに変換している。

【００４０】

このスペクトル評価部１４８の出力部あるいはベクトル量子化器１１６の入力部に設けられたデータ数変換部からの上記一定個数Ｍ個（例えば４４個）の振幅データあるいはエンベロープデータが、ベクトル量子化器１１６により、所定個数、例えば４４個のデータ毎にまとめられてベクトルとされ、重み付きベクトル量子化が施される。この重みは、聴覚重み付けフィルタ算出回路１３９からの出力により与えられる。ベクトル量子化器１１６からの上記エンベロープのインデクスは、スイッチ１１７を介して出力端子１０３より取り出される。なお、上記重み付きベクトル量子化に先だって、所定個数のデータから成るベクトルについて適当なリーク係数を用いたフレーム間差分をとっておくようにしてもよい。

【００４１】

次に、第２の符号化部１２０について説明する。第２の符号化部１２０は、いわゆるＣＥＬＰ（符号励起線形予測）符号化構成を有しており、特に、入力音声信号の無声音部分の符号化のために用いられている。この無声音部分用のＣＥＬＰ符号化構成において、雑音符号帳、いわゆるストキャスティック・コードブック（stochastic code book）１２１か

らの代表値出力である無声音のＬＰＣ残差に相当するノイズ出力を、ゲイン回路１２６を介して、聴覚重み付きの合成フィルタ１２２に送っている。重み付きの合成フィルタ１２２では、入力されたノイズをＬＰＣ合成処理し、得られた重み付き無声音の信号を減算器１２３に送っている。減算器１２３には、上記入力端子１０１からＨＰＦ（ハイパスフィルタ）１０９を介して供給された音声信号を聴覚重み付けフィルタ１２５で聴覚重み付けした信号が入力されており、合成フィルタ１２２からの信号との差分あるいは誤差を取り出している。この誤差を距離計算回路１２４に送って距離計算を行い、誤差が最小となるような代表値ベクトルを雑音符号帳１２１でサーチする。このような合成による分析（Analysis by Synthesis ）法を用いたクローズドループサーチを用いた時間軸波形のベクトル量子化を行っている。

【００４２】

このＣＥＬＰ符号化構成を用いた第２の符号化部１２０からのＵＶ（無声音）部分用のデータとしては、雑音符号帳１２１からのコードブックのシェイプインデクスと、ゲイン回路１２６からのコードブックのゲインインデクスとが取り出される。雑音符号帳１２１からのＵＶデータであるシェイプインデクスは、スイッチ１２７ｓを介して出力端子１０７ｓに送られ、ゲイン回路１２６のＵＶデータであるゲインインデクスは、スイッチ１２７ｇを介して出力端子１０７ｇに送られている。

【００４３】

ここで、これらのスイッチ１２７ｓ、１２７ｇ及び上記スイッチ１１７、１１８は、上記Ｖ／ＵＶ判定部１１５からのＶ／ＵＶ判定結果によりオン／オフ制御され、スイッチ１１７、１１８は、現在伝送しようとするフレームの音声信号のＶ／ＵＶ判定結果が有声音（Ｖ）のときオンとなり、スイッチ１２７ｓ、１２７ｇは、現在伝送しようとするフレームの音声信号が無声音（ＵＶ）のときオンとなる。

【００４４】

符号化部２の出力した上記符号化パラメータは、周期変更部３に供給される。周期変更部３は、上記符号化パラメータの出力周期を時間軸を圧縮伸長して変更する。この周期変更部３により変更された周期で出力された上記符号化パラメータは、復号化部４に供給される。

【００４５】

復号化部４は、周期変更部３により時間軸が例えば圧縮された上記符号化パラメータを補間処理して所定フレーム毎の時刻に対応する変更符号化パラメータを生成するパラメータ変更処理部５と、上記変更符号化パラメータに基づいて有声音部分と無声音部分を合成する音声合成処理部６とを備えてなる。

【００４６】

この復号化部４について図４及び図５を参照しながら説明する。図４において、入力端子２０２には周期変更部３からの上記ＬＳＰ（線スペクトル対）の量子化出力としてのコードブックインデクスが入力される。入力端子２０３、２０４、及び２０５には、上記周期変更部３からの各出力、すなわちエンベロープ量子化出力としてのインデクス、ピッチ、及びＶ／ＵＶ判定出力がそれぞれ入力される。また、入力端子２０７には、上記周期変更部３からのＵＶ（無声音）用のデータとしてのインデクスが入力される。

【００４７】

入力端子２０３からのエンベロープ量子化出力としてのインデクスは、逆ベクトル量子化器２１２に送られて逆ベクトル量子化され、ＬＰＣ残差のスペクトルエンベロープが求められる。このＬＰＣ残差のスペクトルエンベロープは、有声音合成部２１１に送られる前に、図中矢印で示す$Ｐ_1$付近において一旦パラメータ変更処理部５に取り出されて後述するパラメータ変更処理が施された後、有声音合成部２１１に送られる。

【００４８】

有声音合成部２１１は、サイン波合成により有声音部分のＬＰＣ（線形予測符号化）残差を合成する。この有声音合成部２１１には、入力端子２０４及び２０５から入力されて$Ｐ_2$及び$Ｐ_3$の位置で一旦パラメータ変更処理部５に取り出され、パラメータ変更処理が施さ

れたピッチ及びＶ／ＵＶ判定出力も供給される。有声音合成部２１１からの有声音のＬＰＣ残差は、ＬＰＣ合成フィルタ２１４に送られる。

【００４９】

また、入力端子２０７からのＵＶデータのインデクスは、無声音合成部２２０に送られる。このＵＶデータのインデクスは、無声音合成部２２０内にて雑音符号帳を参照することにより無声音部分のＬＰＣ残差とされる。この際、ＵＶデータのインデクスは無声音合成部２２０内からP_4に示すようにパラメータ変更処理部５に一旦取り出され、パラメータ変更処理が施される。パラメータ変更処理が施された後のＬＰＣ残差もＬＰＣ合成フィルタ２１４に送られる。

【００５０】

ＬＰＣ合成フィルタ２１４では、上記有声音部分のＬＰＣ残差と無声音部分のＬＰＣ残差とにそれぞれ独立に、ＬＰＣ合成処理を施す。あるいは、有声音部分のＬＰＣ残差と無声音部分のＬＰＣ残差とが加算されたものに対してＬＰＣ合成処理を施すようにしてもよい。

【００５１】

また、入力端子２０２からのＬＳＰのインデクスは、ＬＰＣパラメータ再生部２１３に送られる。このＬＰＣパラメータ再生部２１３では、最終的にＬＰＣのαパラメータが取り出されるが、その途中にあってＬＳＰの逆ベクトル量子化データが矢印P_5に示すように一旦パラメータ変更処理部５に取り出されてパラメータ変更処理される。

【００５２】

パラメータ変更処理された逆量子化データは、このＬＰＣパラメータ再生部２１３に戻されＬＳＰの補間を行った後、ＬＰＣのαパラメータとされてＬＰＣ合成フィルタ１４に供給される。ＬＰＣ合成フィルタ２１４によりＬＰＣ合成されて得られた音声信号は、出力端子２０１より取り出される。

【００５３】

この図４に示した音声合成処理部６は、上述したようにパラメータ変更処理部５で算出された変更符号化パラメータを受け取って合成音声を出力している。実際には、図５に示すような構成となる。この図５において、上記図４の各部と対応する部分には、同じ指示符号を付している。

【００５４】

この図５において、入力端子２０２を介して入力されたＬＳＰのインデクスは、ＬＰＣパラメータ再生部２１３のＬＳＰの逆ベクトル量子化器２３１に送られてＬＳＰ（線スペクトル対）データに逆ベクトル量子化されてからパラメータ変更処理部５に送られる。

【００５５】

入力端子２０３からのスペクトルエンベロープＡｍのベクトル量子化されたインデクスデータは、逆ベクトル量子化器２１２に送られて逆ベクトル量子化が施され、スペクトルエンベロープのデータとなって、パラメータ変更処理部５に送られる。

【００５６】

また、入力端子２０４、２０５からのピッチ、Ｖ／ＵＶ判定データもパラメータ変更処理部５に送られる。

【００５７】

また、図５の入力端子２０７ｓ及び２０７ｇには、周期変更部３を介した上記図３の出力端子１０７ｓ及び１０７ｇからのＵＶデータとしてのシェイプインデクス及びゲインインデクスがそれぞれ供給され、無声音合成部２２０に送られている。端子２０７ｓからのシェイプインデクスは、無声音合成部２２０の雑音符号帳２２１に、端子２０７ｇからのゲインインデクスはゲイン回路２２２にそれぞれ送られている。雑音符号帳２２１から読み出された代表値出力は、無声音のＬＰＣ残差に相当するノイズ信号成分であり、これがゲイン回路２２２で所定のゲインの振幅となりパラメータ変更処理部５に送られる。

【００５８】

パラメータ変更処理部５は、符号化部２が出力し、周期変更部３で出力周期が変更された

上記符号化パラメータに補間処理を施して変更符号化パラメータを生成し、音声合成処理部６に供給する。ここで、パラメータ変更処理部５は、上記符号化パラメータをスピード変換している。このため、音声信号再生装置１はデコーダ出力後のスピード変換処理が不要で、かつ同様のアルゴリズムで異なるレートでの固定レートに容易に対応することもできる。

【００５９】

以下、図６及び図８のフローチャートを参照しながら周期変更部３とパラメータ変更処理部５の動作について説明する。

【００６０】

先ず、図６のステップＳ１に示すように、周期変更部３はＬＳＰ，ピッチ，有声音／無声音Ｖ／ＵＶ，スペクトルエンベロープＡm，ＬＰＣ残差のような符号化パラメータを受け取る。ここで、ＬＳＰを$l_{sp}[n][p]$，ピッチを$p_{ch}[n]$，Ｖ／ＵＶを$vu_{v}[n]$，Ａmを$a_{m}[n][k]$，ＬＰＣ残差を$r_{es}[n][i][j]$とする。

【００６１】

なお、パラメータ変更処理部５で最終的に算出される変更符号化パラメータを$mod_l_{sp}[m][p]$，$mod\_p_{ch}[m]$，$mod\_vu_{v}[m]$，$mod\_a_{m}[m][k]$，$mod_r_{es}[m][i][j]$とする。ここで、ｋはハーモニクス数、ｐはＬＳＰ次数である。ｎ，ｍは、時間軸のインデクスに相当するフレームナンバーに対応する。ｎは時間軸変更前、ｍは時間軸変更後である。ｎ，ｍともに例えば２０msecをフレームインターバルとするフレームのインデクスである。また、ｉはサブフレーム番号、ｊはサンプル番号である。

【００６２】

次に、周期変更部３は、ステップＳ２に示すようにオリジナルの時間長となるフレーム数をN_1とし、変更後の時間長となるフレーム数をN_2としてから、ステップＳ３に示すようにN_1の音声をN_2の音声に時間軸圧縮する。すなわち、周期変更部３での時間軸圧縮の比をspdとすると、spdを$N_2／N_1$として求める。ここで、$0 \leqq n < N_1$，$0 \leqq m < N_2$である。

【００６３】

次に、パラメータ変更処理部５は、ステップＳ４に示すように、時間軸変更後の時間軸のインデクスに相当するフレームナンバーに対応するｍを２とする。

【００６４】

そして、パラメータ変更処理部５は、ステップＳ５に示すように、二つのフレームf_{r0}，f_{r1}と、該二つのフレームf_{r0}，f_{r1}とm／spdとの差left，rightとを求める。

【００６５】

上記符号化パラメータのl_{sp}，p_{ch}，vu_{v}，a_{m}，r_{es}を＊とするときmod_＊[m]は、

$mod\_*[m] = *[m／spd]$　　（$0 \leqq m < N_2$）

という一般式で表せる。しかし、m／spdは、整数にはならないので、

f_{r0}＝「m／spd」

$f_{r1} = f_0 + 1$

の２フレームから補間して、m／spdにおける変更符号化パラメータを作る。

【００６６】

ここで、フレームf_{r0}とm／spdとフレームf_{r1}との間には、図７に示すような関係、すなわち、

left＝m／spd－f_{r0}

right＝f_{r1}－m／spd

が成立する。

【００６７】

この図７におけるm／spdのときの符号化パラメータ、すなわち変更符号化パラメータをステップＳ６に示すように、補間処理によって作ればよい。

【００６８】

単純に直線補間により求めると、

mod_＊[m]＝＊[f_{r0}]×right＋＊[f_{r1}]×left
となる。
【００６９】
しかし、２つのフレームf_{r0}，f_{r1}間での補間では、それらのフレームが有声音（Ｖ）と，無声音（ＵＶ）というように異なる場合には、上記一般式を適用できない。このため、２つのフレームf_{r0}，f_{r1}間における有声音（Ｖ）と，無声音（ＵＶ）との関係によって、パラメータ変更処理部５は、図８のステップＳ１１以降に示すように、上記符号化パラメータの求め方を変える。
【００７０】
先ず、ステップＳ１１に示すように２つのフレームf_{r0}，f_{r1}が有声音（Ｖ），有声音（Ｖ）であるか否かを判断する。ここで、２つのフレームf_{r0}，f_{r1}が共に、有声音（Ｖ）であると判断すると、ステップＳ１２に進み、全てのパラメータを線形補間して以下のように表す。
【００７１】
mod_p_{ch}[m]＝p_{ch}[f_{r0}]×right＋p_{ch}[f_{r1}]×left
mod_a_m[m][k]＝a_m[f_{r0}][k]×right＋a_m[f_{r1}][k]×left
ただし、０≦ｋ＜Ｌである。ここで、Ｌはハーモニクスとしてとりうる最大の数である。また、a_m[n][k]は、ハーモニクスの存在しない位置では０を入れておく。フレームf_{r0}とフレームf_{r1}とで、ハーモニクスの数が異なる時には、余った方のハーモニクスは、相方を０として補間する。または、デコーダ側でデータ数変換器を通す前であれば、０≦ｋ＜ＬのＬ＝４３といった固定の値でもよい。
【００７２】
mod_l_{sp}[m][p]＝l_{sp}[f_{r0}][p]×right＋l_{sp}[f_{r1}][p]×left
ただし、０≦ｐ＜Ｐである。ここで、ＰはＬＳＰの次数であり、通常は１０を使用する。
【００７３】
mod_vu_v[m]＝１
Ｖ／ＵＶの判定で１は有声音（Ｖ）を、０は無声音（ＵＶ）を意味する。
【００７４】
次に、ステップＳ１１で２つのフレームf_{r0}，f_{r1}が共に有声音（Ｖ）でないと判断すると、ステップＳ１３に示すような判断、すなわち、２つのフレームf_{r0}，f_{r1}が共に無声音（ＵＶ）であるか否かを判断する。
【００７５】
ここで、ＹＥＳ（共に無声音である。）となると、補間処理部５は、ステップＳ１４に示すように、p_{ch}を最大値とし、ｍ／ｓｐｄを中心にr_{es}の前後８０サンプルづつ切り出してmod_r_{es}を作る。
【００７６】
実際、このステップＳ１４においては、left＜rightであるときに、ｍ／ｓｐｄを中心に図９の（Ａ）に示すようにr_{es}の前後８０サンプルづつ切り出してmod_r_{es}に入れる。すなわち、
ｆｏｒ（ｊ＝０；ｊ＜FRM×（１／２－ｍ／ｓｐｄ＋f_{r0}）；ｊ$^{++}$）{mod_r_{es}[m][０][ｊ]＝r_{es}[f_{r0}][０][ｊ＋（ｍ／ｓｐｄ－f_{r0}）×FRM]；}；
ｆｏｒ（ｊ＝FRM×（１／２－ｍ／ｓｐｄ＋f_{r0}）；ｊ＜FRM／２；ｊ$^{++}$）{mod_r_{es}[m][０][ｊ]＝r_{es}[f_{r0}][１][ｊ－FRM×（１／２－ｍ／ｓｐｄ＋f_{r0}）]；}；
ｆｏｒ（ｊ＝０；ｊ＜FRM×（１／２－ｍ／ｓｐｄ＋f_{r0}）；ｊ$^{++}$）{mod_r_{es}[m][１][ｊ]＝r_{es}[f_{r0}][１][ｊ＋（ｍ／ｓｐｄ－f_{r0}）×FRM]；}；
ｆｏｒ（ｊ＝FRM×（１／２－ｍ／ｓｐｄ＋f_{r0}）；ｊ＝FRM／２；ｊ$^{++}$）{mod_r_{es}[m][１][ｊ]＝r_{es}[f_{r0}][０][ｊ－FRM×（１／２－ｍ／ｓｐｄ＋f_{r0}）]；}；
とする。ここで例えばFRMは１６０である。
【００７７】
一方、このステップＳ１４においては、left≧rightであるときに、ｍ／ｓｐｄを中心に

図９の（Ｂ）に示すようにr_{es}の前後８０サンプルづつ切り出してmod_r_{es}とする。

【００７８】

ステップＳ１３の条件を満たさない場合、ステップＳ１５に進み、フレームf_{r0}が有声音（Ｖ）で，f_{r1}が無声音（ＵＶ）であるか否かを判断する。ここでＹＥＳ（フレームf_{r0}が有声音（Ｖ）で，f_{r1}が無声音（ＵＶ）である。）となると、ステップＳ１６に進み、ＮＯ（フレームf_{r0}が無声音（ＵＶ）であり、f_{r1}が有声音（Ｖ）である。）となると、ステップＳ１７に進む。

【００７９】

ステップＳ１５以降の処理では、二つのフレームf_{r0}，f_{r1}が、例えば有声音（Ｖ），無声音（ＵＶ）のように、異なった場合について説明している。これは、例えば有声音（Ｖ），無声音（ＵＶ）のように、異なった２つのフレームf_{r0}，f_{r1}間でパラメータを補間すると意味のないものになってしまうためである。

【００８０】

ステップＳ１６では、図７に示す上記left（＝m／spd－f_{r0}）と上記right（＝f_{r1}－m／spd）の大きさを比較している。これにより、m／spdに対してフレームf_{r0}が近いのか否かを判断している。

【００８１】

フレームf_{r0}が近い場合には、ステップＳ１８に示すように、このフレームf_{r0}側のパラメータを用いて、

$mod\_p_{ch}[m]=p_{ch}[f_{r0}]$

$mod\_a_{m}[m][k]=a_{m}[f_{r0}][k]$，（ただし、０≦ｋ＜Ｌである。）

$mod\_l_{sp}[m][p]=l_{sp}[f_{r0}][p]$，（ただし、０≦ｐ＜Ｉである。）

$mod\_vu_{v}[m]=1$

とする。

【００８２】

また、ステップＳ１６でＮＯと判断したときには、left≧rightでありフレームf_{r1}の方が近いので、ステップＳ１９に進み、ピッチを最大値にすると共に、図９の（Ｃ）に示すようにf_{r1}側のr_{es}をそのまま使用してmod_r_{es}とする。すなわち、$mod_r_{es}[m][i][j]=r_{es}[f_{r1}][i][j]$とする。これは、有声音である$f_{r0}$ではＬＰＣ残差$r_{es}$が伝送されないためである。

【００８３】

次に、ステップＳ１７では、ステップＳ１５で２つのフレームf_{r0}，f_{r1}が無声音（ＵＶ），有声音（Ｖ）であるという判断を受けて、上記ステップＳ１６と同様の判断を行う。すなわち、図７に示す上記left（＝m／spd－f_{r0}）と上記right（＝f_{r1}－m／spd）の大きさを比較している。これにより、m／spdに対してフレームf_{r0}が近いのか否かを判断している。

【００８４】

フレームf_{r0}が近い場合には、ステップＳ１８に進み、ピッチを最大値にすると共に、図９の（Ｄ）に示すようにf_{r0}側のr_{es}をそのまま使用してmod_r_{es}とする。すなわち、$mod_r_{es}[m][i][j]=r_{es}[f_{r0}][i][j]$とする。これは、有声音である$f_{r1}$ではＬＰＣ残差$r_{es}$が伝送されないためである。

【００８５】

また、ステップＳ１７でＮＯと判断したときには、left≧rightでありフレームf_{r1}の方が近いので、ステップＳ２１に進み、このフレームf_{r1}側のパラメータを用いて、

$mod\_p_{ch}[m]=p_{ch}[f_{r1}]$

$mod\_a_{m}[m][k]=a_{m}[f_{r1}][k]$，（ただし、０≦ｋ＜Ｌである。）

$mod\_l_{sp}[m][p]=l_{sp}[f_{r1}][p]$，（ただし、０≦ｐ＜Ｉである。）

$mod\_vu_{v}[m]=1$

とする。

【００８６】

このように２つのフレームf_{r0}，f_{r1}間における有声音（Ｖ）と，無声音（ＵＶ）との関係によって、補間処理部５は、図８に詳細を示した図６のステップＳ６の補間処理を異ならせる。このステップＳ６の補間処理が終了すると、ステップＳ７に進み、ｍをインクリメントする。そして、このｍがN_2に等しくなるまで、ステップＳ５，ステップＳ６の処理を繰り返す。

【００８７】

ここで、周期変更部３とパラメータ変更処理部５の動作について図１０を参照しながらまとめて説明しておく。図１０の（Ａ）に示すように、符号化部２が例えば周期２０msec毎に抽出している符号化パラメータの該周期を、周期変更部３は、図１０の（Ｂ）に示すように、時間圧縮して１５msecとする。そして、上述したように、パラメータ変更処理部５が二つのフレームf_{R0}，f_{r1}のＶ／ＵＶの状態に応じた補間処理により、図１０の（Ｃ）に示すように周期２０msec毎に変更符号化パラメータを算出する。

【００８８】

また、周期変更部３とパラメータ変更処理部５を逆の順番として、図１１の（Ａ）に示す符号化パラメータを先ず図１１の（Ｂ）に示すように補間してから、図１１の（Ｃ）に示すように圧縮して変更符号化パラメータを算出してもよい。

【００８９】

ここで、図５に戻る。パラメータ変更処理部５で算出されたＬＳＰデータに関する変更符号化パラメータ$mod_l_{sp}[m][p]$は、ＬＳＰ補間回路232_v、232_uに送られてＬＳＰの補間処理が施された後、ＬＳＰ→α変換回路234_v、234_{uv}でＬＰＣ（線形予測符号）のαパラメータに変換され、このαパラメータがＬＰＣ合成フィルタ２１４に送られる。ここで、ＬＳＰ補間回路232_v及びＬＳＰ→α変換回路234_vは有声音（Ｖ）用であり、ＬＳＰ補間回路232_u及びＬＳＰ→α変換回路234_uは無声音（ＵＶ）用である。またＬＰＣ合成フィルタ２１４は、有音声部分のＬＰＣ合成フィルタ２３６と、無声音部分のＬＰＣ合成フィルタ２３７とを分離している。すなわち、有音声部分と無音声部分とでＬＰＣの係数補間を独立に行うようにして、有音声から無音声への遷移部や、無音声から有音声への遷移部で、全く性質の異なるＬＳＰ同士を補間することによる悪影響を防止している。

【００９０】

パラメータ変更処理部５で算出されたスペクトルエンベロープデータに関する変更符号化パラメータ$mod_a_m[m][k]$は有声音合成部２１１のサイン波合成回路２１５に送られている。このサイン波合成回路２１５には、パラメータ変更処理部５で算出されたピッチに関する変更符号化パラメータ$mod_p_{ch}[m]$及び上記Ｖ／ＵＶ判定データに関する変更符号化パラメータ$mod_vu_v[m]$も供給されている。サイン波合成回路２１５からは、上述した図３のＬＰＣ逆フィルタ１１１からの出力に相当するＬＰＣ残差データが取り出され、これが加算器２１８に送られている。

【００９１】

また、パラメータ変更処理部５で算出されたスペクトルエンベロープデータに関する変更符号化パラメータ$mod_a_m[m][k]$と、ピッチに関する変更符号化パラメータ$mod_p_{ch}[m]$及び上記Ｖ／ＵＶ判定データに関する変更符号化パラメータ$mod_vu_v[m]$とは、有声音（Ｖ）部分のノイズ加算のためのノイズ合成回路２１６に送られている。このノイズ合成回路２１６からの出力は、重み付き重畳加算回路２１７を介して加算器２１８に送られている。これは、サイン波合成によって有声音のＬＰＣ合成フィルタへの入力となるエクサイテイション（Excitation：励起、励振）を作ると、男声等の低いピッチの音で鼻づまり感がある点、及びＶ（有声音）とＵＶ（無声音）とで音質が急激に変化し不自然に感じる場合がある点を考慮し、有声音部分のＬＰＣ合成フィルタ入力すなわちエクサイテイションについて、音声符号化データに基づくパラメータ、例えばピッチ、スペクトルエンベロープ振幅、フレーム内の最大振幅、残差信号のレベル等を考慮したノイズをＬＰＣ残差信号の有声音部分に加えているものである。

【００９２】

加算器２１８からの加算出力は、ＬＰＣ合成フィルタ２１４の有声音用の合成フィルタ２３６に送られてＬＰＣの合成処理が施されることにより時間波形データとなり、さらに有声音用ポストフィルタ２３８ｖでフィルタ処理された後、加算器２３９に送られる。

【００９３】

ここで、ＬＰＣ合成フィルタ２１４は、上述したように、Ｖ（有声音）用の合成フィルタ２３６と、ＵＶ（無声音）用の合成フィルタ２３７とに分離されている。すなわち、合成フィルタを分離せずにＶ／ＵＶの区別なしに連続的にＬＳＰの補間を２０サンプルすなわち２．５msec 毎に行う場合には、Ｖ→ＵＶ、ＵＶ→Ｖの遷移（トランジェント）部において、全く性質の異なるＬＳＰ同士を補間することになり、Ｖの残差にＵＶのＬＰＣが、ＵＶの残差にＶのＬＰＣが用いられることにより異音が発生するが、このような悪影響を防止するために、ＬＰＣ合成フィルタをＶ用とＵＶ用とで分離し、ＬＰＣの係数補間をＶとＵＶとで独立に行わせたものである。

【００９４】

また、パラメータ変更処理部５で算出されたＬＰＣ残差に関する変更符号化パラメータmod_r_{es}[m][ｉ][ｊ]は、窓かけ回路２２３に送られて、上記有声音部分とのつなぎを円滑化するための窓かけ処理が施される。

【００９５】

窓かけ回路２２３からの出力は、無音声合成部２２０からの出力として、ＬＰＣ合成フィルタ２１４のＵＶ（無音声）用の合成フィルタ２３７に送られる。合成フィルタ２３７では、ＬＰＣ合成処理が施されることにより無音声部分の時間波形データとなり、この無声音部分の時間波形データは無声音用ポストフィルタ２３８ｕでフィルタ処理された後、加算器２３９に送られる。

【００９６】

加算器２３９では、有声音用ポストフィルタ２３８ｖからの有声音部分の時間波形信号と、無声音用ポストフィルタ２３８ｕからの無声音部分の時間波形データとが加算され、出力端子２０１より取り出される。

【００９７】

このように、この音声信号再生装置１は、変更符号化パラメータmod_＊[m]の配列（０≦m＜N_2）を本来の配列＊[ｎ]（０≦ｎ＜N_1）のかわりにデコードしている。デコード時のフレーム間隔は従来通り例えば２０msecのように固定である。このため、N_2＜N_1の時には、時間軸圧縮となり、スピードアップとなる。他方、N_2＞N_1の時には、時間軸伸長となり、スピードダウンとなる。

【００９８】

上記時間軸変更を行っても、瞬時スペクトル、ピッチが不変である為、０．５≦spd≦２程度以上の広い範囲の変更を行っても劣化が少ない。

【００９９】

この方式では、最終的に得られたパラメータ列を本来のスペーシング（２０msec）に並べてデコードするため、任意のスピードコントロール（上下）が簡単に実現できる。又、スピードアップとスピードダウンが区別なしに、同一の処理で可能である。

【０１００】

このため、例えば固体録音した内容をリアルタイムの倍のスピードで再生できる。このとき、再生スピードを高速にしてもピッチ、音韻が不変であるため、かなりの高速再生を行っても内容を聞きとることができる。

ここで、N_2＜N_1のとき、すなわち再生スピードを遅くした場合、無声音フレームにおいては同じＬＰＣ残差r_{es}から複数のmod_r_{es}が作られるので再生音が不自然になることがある。そこで、mod_r_{es}に対し、ノイズを適量加えることにより不自然さを改善する事が可能である。また、ノイズを加える以外にも、mod_r_{es}を適当に生成したガウシアンノイズなどで置き換えたり、コードブックよりランダムに選択した励起ベクトルを用いることも考えられる。

【０１０１】

なお、上記音声信号再生装置１では、周期変更部３によって符号化部２からの上記符号化パラメータの出力周期の時間軸を圧縮してスピードアップさせていたが、復号化部４にてフレーム長を可変にしてスピードをコントロールしてもよい。

【０１０２】

この場合、復号化部４を構成するパラメータ変更処理部５は、上記フレーム長が可変となるためパラメータ生成前後でフレーム番号ｎを変化させない。

【０１０３】

先ず、パラメータ変更処理部５は、該当フレームが有声音、無声音に拘らず、$l_{sp}[n][p]$、$vu_v[n]$を、$mod\_l_{sp}[n][p]$、$mod_vu_v[n]$とする。

【０１０４】

$p_{ch}[n]$，$a_m[n][k]$については、$mod_vu_v[n]$が１、すなわち該当フレームが有声音（Ｖ）である場合、$mod\_p_{ch}[n]$，$mod_a_m[n][k]$とする。

【０１０５】

$r_{es}[n][i][j]$については、$mod_vu_v[n]$が０、すなわち該当フレームが無声音（ＵＶ）である場合、$mod_r_{es}[n][i][j]$とする。

【０１０６】

ここで、パラメータ変更処理部５は、各パラメータの変換を、$l_{sp}[n][p]$，$p_{ch}[n]$，$vu_v[n]$，$a_m[n][k]$についてはそのまま、$mod\_l_{sp}[n][p]$，$mod\_p_{ch}[n]$，$mod\_vu_v[n]$，$mod_a_m[n][k]$とするが残差信号$r_{es}[n][i][j]$についてはスピードspdによって、$mod_r_{es}[n][i][j]$を異ならせる。

【０１０７】

スピードspd＜１．０のとき、すなわちスピードが速い場合、図１２に示すように、元のフレームの残差信号を中央から切り出す。元フレーム長をorgFrmLとしたとき、元フレーム$r_{es}[n][i]$から$(orgFrmL-frmL)/2 \leqq j \leqq (orgFrmL+frmL)/2$の部分を切り出し、$mod_r_{es}[n][i]$とする。なお、元フレームの先頭から切り出すことも可能である。

【０１０８】

一方、スピードspd＞１．０のとき、すなわちスピードが遅い場合、図１３に示すように、元のフレームを用い、不足分は元のフレームにノイズ成分を加えたものを用いる。なお、不足分として、コードブックよりランダムに選んだ励起ベクトルを用いてもよい。また、デコードされた励起ベクトルに適当に生成したノイズ成分を付加してもよい。さらに、ガウシアンノイズを発生し、それを励起ベクトルとして用いてよい。これは同じ波形形状のフレームが連続することにより生じる違和感を軽減するためである。また、元フレームの両端に上記のようなノイズ成分等を付加してもよい。

【０１０９】

このため、音声合成処理部６は、フレーム長を変更することによりスピードコントロールを実現する音声信号再生装置１にあっては、ＬＳＰ補間部232_v、232_uと、サイン波合成部２１５と、窓かけ部２２３の動作を時間軸圧縮伸長によりスピードをコントロールする場合に対して異ならせる。

【０１１０】

先ず、ＬＳＰ補間部232_vでは、該当フレームが有声音（Ｖ）ならば$frmL/p \leqq 20$を満たす最小の整数ｐを、また、ＬＳＰ補間部232_uでは、該当フレームが無声音（ＵＶ）ならば$frmL/p \leqq 80$を満たす最小の整数ｐを求め、ＬＳＰ補間のためのサブフレームsubl[ｉ][ｊ]の範囲を、以下の式により定める。

【０１１１】

$$nint(frmL/p \times i) \leqq j \leqq nint(frmL/p \times (i+1)),(0 \leqq i \leqq p-1)$$

ここで、nint(x)は小数第１位を四捨五入することにより、xに最も近い整数を返す関数である。ただし、有声音、無声音いずれの場合もfrmLが２０、８０以下となった場合はｐ＝１とする。

例えば、ｉ番目のサブフレームについて、サブフレームの中心は$frmL \times (2i+1)/2p$であるから、$frmL \times (2p-2i-1)/2p : frmL \times (2i+1)/2p$の割合でＬ

ＳＰの補間を行う。

【０１１２】

なお、この他にも、サブフレーム数をある定数に固定してしまい、つねに同じ比で各サブフレームのＬＳＰ補間を行ってもよい。サイン波合成部２１５では、フレーム長frmLに応じてサンプル数を発生する。サイン波合成の具体的な方法としては本件出願人が先に提案した特願平６ー１９８４５１号の明細書及び図面に開示したものを挙げることができる。窓かけ部２２３では、フレーム長frmLに合わせて、窓長を変更する。

【０１１３】

なお、上記音声信号再生装置１では、周期変更部３、及びパラメータ変更処理部５を用いて、出力周期を時間軸上で圧縮伸長した符号化パラメータを変更することによって、ピッチ、音韻を不変としながらも再生スピードを可変としているが、周期変更部３を省略して符号化部２からの符号化データを図１４に示す復号化部８のデータ数変換部２７０により処理して音韻を変えずにピッチを可変とすることもできる。図１４において、上記図４の各部と対応する部分には、同じ指示符号を付している。

【０１１４】

この復号化部８の基本的な考え方は、符号化部２から入力された音声符号化データのハーモニクスの基本周波数と所定の帯域内における個数をデータ変換手段となるデータ数変換部２７０により変換して、復号化処理を施すことにより、音韻を変えずにピッチのみを変更するものである。データ数変換部２７０は、入力された各ハーモニクスにおけるスペクトルの大きさを表すデータの個数を補間処理により変更することによってピッチを変更する。

【０１１５】

図１４において、入力端子２０２には、上記図２、３の出力端子１０２からの出力に相当するＬＳＰのベクトル量子化出力、いわゆるコードブックのインデクスが供給されている。

【０１１６】

このＬＳＰのインデクスは、ＬＰＣパラメータ再生部２１３のＬＳＰの逆ベクトル量子化器２３１に送られてＬＳＰ（線スペクトル対）データに逆ベクトル量子化され、ＬＳＰ補間回路２３２、２３３に送られてＬＳＰの補間処理が施された後、ＬＳＰ→α変換回路２３４、２３５でＬＰＣ（線形予測符号）のαパラメータに変換され、このαパラメータがＬＰＣ合成フィルタ２１４に送られる。ここで、ＬＳＰ補間回路２３２及びＬＳＰ→α変換回路２３４は有声音（Ｖ）用であり、ＬＳＰ補間回路２３３及びＬＳＰ→α変換回路２３５は無声音（ＵＶ）用である。またＬＰＣ合成フィルタ２１４は、有声音部分のＬＰＣ合成フィルタ２３６と、無声音部分のＬＰＣ合成フィルタ２３７とを分離している。すなわち、有声音部分と無声音部分とでＬＰＣの係数補間を独立に行うようにして、有声音から無声音への遷移部や、無声音から有声音への遷移部で、全く性質の異なるＬＳＰ同士を補間することによる悪影響を防止している。

【０１１７】

また、図１４の入力端子２０３には、上記図２、図３のエンコーダ側の端子１０３からの出力に対応するスペクトルエンベロープ（Ａm）の重み付けベクトル量子化されたコードインデクスデータが供給され、入力端子２０４には、上記図２、図３の端子１０４からのピッチのデータが供給され、入力端子２０５には、上記図２、図３の端子１０５からのＶ／ＵＶ判定データが供給されている。

【０１１８】

入力端子２０３からのスペクトルエンベロープＡmのベクトル量子化されたインデクスデータは、逆ベクトル量子化器２１２に送られて逆ベクトル量子化が施される。この逆ベクトル量子化されたエンベロープの振幅データの個数は、上述したように一定個数、例えば４４個とされており、基本的には、ピッチデータに応じた本数のハーモニクスとなるようにデータ数変換する。これに対して本例のように、ピッチを変更したい場合には、逆ベクトル量子化器２１２からのエンベロープデータをデータ数変換部２７０に送って、変更し

たいピッチに応じて補間処理等によりエンベロープの振幅データの個数を変更している。

【０１１９】

また、データ数変換部２７０には、入力端子２０４からのピッチデータも供給されており、エンコード時のピッチが、変更したいピッチに変換されて出力される。このデータ数変換部２７０からのＬＰＣ残差のスペクトルエンベロープの変更ピッチに応じた個数の振幅データと、変更されたピッチデータとが有声音合成部２１１のサイン波合成回路２１５に送られている。

【０１２０】

ここで、データ数変換部２７０でのＬＰＣ残差のスペクトルエンベロープの振幅データの個数を変換するには、種々の補間方法が考えられるが、例えば、周波数軸上の有効帯域１ブロック分の振幅データに対して、ブロック内の最後のデータからブロック内の最初のデータまでの値を補間するようなダミーデータを付加してデータ個数をN_F 個に拡大した後あるいはブロック内の左端及び右端（最初と最後）のデータを延長してダミーデータとして、帯域制限型のO_S 倍（例えば８倍）のオーバーサンプリングを施すことによりO_S 倍の個数の振幅データを求め、このO_S 倍の個数（（m_{MX}＋１）×O_S 個）の振幅データを直線補間してさらに多くのN_M 個（例えば２０４８個）に拡張し、このN_M 個のデータを間引いて、変更したいピッチに応じた個数Ｍのデータに変換すればよい。

【０１２１】

データ数変換部２７０においては、スペクトルエンベロープの形状を変えないで、ハーモニクスの立っている位置だけを変更するようにしている。このため、音韻は不変である。

【０１２２】

ここで、上記データ数変換部２７０における動作の一例として、ピッチラグＬのときの周波数F_0＝ｆｓ／Ｌを、Ｆｘ に変換する場合について説明する。ｆｓ はサンプリング周波数であり、例えば ｆｓ＝８ｋHz＝８０００Hzとする。

【０１２３】

このとき、ピッチ周波数F_0＝８０００／Ｌ であり、ハーモニクスは４０００Hzまでの間にｎ＝Ｌ／２本立っている。通常の音声帯域の３４００Hz幅では、約（Ｌ／２）×（３４００／４０００）である。これを、上述したデータ数変換あるいは次元変換により一定の本数、例えば４４本に変換した後、ベクトル量子化を行う。なお、単にピッチ変換を行うのであれば、量子化を行う必要はない。

【０１２４】

ベクトル逆量子化後に、データ数変換部２７０において、４４本のハーモニクスを次元変換で任意の本数、すなわち任意のピッチ周波数Ｆｘ に変更できる。ピッチ周波数Ｆｘ （Hz）に対応するピッチラグＬｘ は、Ｌｘ＝８０００／Ｆｘであり、３４００Hzまでの間には、

$(Lx/2)\times(3400/4000) = (4000/Fx)\times(3400/4000) = 3400/Fx$

すなわち、３４００／Ｆｘ 本のハーモニクスが立っている。すなわち、データ数変換部２７０内での次元変換あるいはデータ数変換で、４４点→３４００／Ｆｘ への変換を行えばよい。

【０１２５】

なお、エンコード時にスペクトルのベクトル量子化に先だってフレーム間差分をとっている場合には、ここでの逆ベクトル量子化後にフレーム間差分の復号を行ってからデータ数変換を行い、スペクトルエンベロープのデータを得る。

【０１２６】

サイン波合成回路２１５には、データ数変換部２７０からのＬＰＣ残差のスペクトルエンベロープ振幅データやピッチデータの他にも、入力端子２０５からの上記Ｖ／ＵＶ判定データが供給されている。サイン波合成回路２１５からは、ＬＰＣ残差データが取り出され、これが加算器２１８に送られている。

【０１２７】

また、逆ベクトル量子化器２１２からのエンベロープのデータと、入力端子２０４、２０５からのピッチ、Ｖ／ＵＶ判定データとは、有声音（Ｖ）部分のノイズ加算のためのノイ

ズ合成回路２１６に送られている。このノイズ合成回路２１６からの出力は、重み付き重畳加算回路２１７を介して加算器２１８に送っている。これは、サイン波合成によって有声音のＬＰＣ合成フィルタへの入力となるエクサイテイション（Excitation：励起、励振）を作ると、男声等の低いピッチの音で鼻づまり感がある点、及びＶ（有声音）とＵＶ（無声音）とで音質が急激に変化し不自然に感じる場合がある点を考慮し、有声音部分のＬＰＣ合成フィルタ入力すなわちエクサイテイションについて、音声符号化データに基づくパラメータ、例えばピッチ、スペクトルエンベロープ振幅、フレーム内の最大振幅、残差信号のレベル等を考慮したノイズをＬＰＣ残差信号の有声音部分に加えているものである。

【０１２８】

加算器２１８からの加算出力は、ＬＰＣ合成フィルタ２１４の有声音用の合成フィルタ２３６に送られてＬＰＣの合成処理が施されることにより時間波形データとなり、さらに有声音用ポストフィルタ２３８ｖでフィルタ処理された後、加算器２３９に送られる。

【０１２９】

次に、図１４の入力端子２０７ｓ及び２０７ｇには、上記図３の出力端子１０７ｓ及び１０７ｇからのＵＶデータとしてのシェイプインデクス及びゲインインデクスがそれぞれ供給され、無声音合成部２２０に送られている。端子２０７ｓからのシェイプインデクスは、無声音合成部２２０の雑音符号帳２２１に、端子２０７ｇからのゲインインデクスはゲイン回路２２２にそれぞれ送られている。雑音符号帳２２１から読み出された代表値出力は、無声音のＬＰＣ残差に相当するノイズ信号成分であり、これがゲイン回路２２２で所定のゲインの振幅となり、窓かけ回路２２３に送られて、上記有声音部分とのつなぎを円滑化するための窓かけ処理が施される。

【０１３０】

窓かけ回路２２３からの出力は、無声音合成部２２０からの出力として、ＬＰＣ合成フィルタ２１４のＵＶ（無声音）用の合成フィルタ２３７に送られる。合成フィルタ２３７では、ＬＰＣ合成処理が施されることにより無声音部分の時間波形データとなり、この無声音部分の時間波形データは無声音用ポストフィルタ２３８ｕでフィルタ処理された後、加算器２３９に送られる。

【０１３１】

加算器２３９では、有声音用ポストフィルタ２３８ｖからの有声音部分の時間波形信号と、無声音用ポストフィルタ２３８ｕからの無声音部分の時間波形データとが加算され、出力端子２０１より取り出される。

【０１３２】

以上説明したように、スペクトルエンベロープの形状を変えないでハーモニクスの本数を変えることにより、音声の音韻を変えることなくピッチを変えることができる。従って、１つの音声パターンの符号化されたデータすなわちエンコーデッドビットストリームを持っていれば、そのピッチを任意に変更して合成することができる。

【０１３３】

すなわち、図１５において、符号化データ出力部３０１からは、上述した図２や図３のエンコーダ等により符号化されることによって得られたエンコーデッドビットストリームあるいは符号化データが出力され、これらのデータの内、少なくともピッチデータ及びスペクトルエンベロープデータがデータ変換部３０２を介して波形合成部３０３に送られ、またＶ／ＵＶ（有声音／無声音）判定データのようなピッチ変換に無関係のデータは直接的に波形合成部３０３に送られる。

【０１３４】

波形合成部３０３は、スペクトルエンベロープデータやピッチデータに基づいて音声波形を合成するものであり、上記図４や図５のような方式の合成装置の場合には、ＬＳＰデータやＣＥＬＰ用のデータ等も符号化データ出力部３０１から取り出されて供給されることは勿論である。

【０１３５】

この図１５のような構成において、少なくともピッチデータやスペクトルエンベロープデータが、上述したように、データ変換部３０２で変更したいピッチに応じて変換された後、波形合成部３０３に送られて音声波形が合成されることにより、音韻を変化させることなくピッチが変更された音声信号を、出力端子３０４から取り出すことができる。

【０１３６】

また、このような技術を、規則合成、テキスト合成等と組み合わせることもできる。

【０１３７】

図１６は、音声のテキスト合成に本発明を適用した例を示すものであり、上述したような音声圧縮符号化のデコーダと、テキスト音声合成の音声合成器とを兼用させることができる。また、図１６の例では、音声データの再生も組み合わせて使用している。

【０１３８】

すなわち、図１６において、規則音声合成部３００内に、音声の規則合成部と、上述したようなピッチ変更のためのデータ変換を伴った音声合成部とが含まれており、テキスト解析部３１０からデータが入力されて、合成された所望のピッチの音声信号が出力され、この合成音声信号は切換スイッチ３３０の被選択端子ａに送られる。また、音声再生部３２０は、必要に応じて圧縮処理が施されてＲＯＭ等に記憶された音声データを読み出し、圧縮処理に対応する伸長処理が施して、音声信号を出力するものである。この再生音声信号は切換スイッチ３３０の被選択端子ｂに送られる。切換スイッチ３３０で上記合成音声信号、再生音声信号の一方が選択されて、出力端子３４０より取り出される。

【０１３９】

この図１６に示すような装置は、例えば自動車等のナビゲーション装置に適用することができる。このナビゲーション装置に適用する場合において、例えば「右に曲がってください。」といった方向指示等の定形の発話には、音声再生部３２０からの高品質でクリアな再生音声を用い、地名や建物名等のように数が膨大でＲＯＭ等に音声情報として蓄えることが難しい特殊な名称等の発話には、規則音声合成部３００からの合成音声を用いることが挙げられる。

【０１４０】

また、本発明を用いることで同一のハードウェアが、３００と３２０の両方に使用できるメリットがある。

【０１４１】

なお、本発明は上記実施の形態のみに限定されるものではなく、例えば上記図１、図３の音声分析側（エンコード側）の構成や、図１４の音声合成側（デコード側）の構成については、各部をハードウェア的に記載しているが、いわゆるＤＳＰ（ディジタル信号プロセッサ）等を用いてソフトウェアプログラムにより実現することも可能である。また、上記ベクトル量子化の代わりに、複数フレームのデータをまとめてマトリクス量子化を施してもよい。さらに、本発明は、種々の音声分析／合成方法に適用でき、用途としても、伝送や記録再生に限定されず、ピッチ変換やスピード変換、規則音声合成、あるいは雑音抑圧のような種々の用途に応用できることは勿論である。

【０１４２】

以上説明したような符号化部及び復号化部は、例えば図１７及び図１８に示すような携帯通信端末あるいは携帯電話機等に使用される音声コーデックとして用いることができる。

【０１４３】

すなわち、図１７は、上記図２、図３に示したような構成を有する音声符号化部１６０を用いて成る携帯端末の送信側構成を示している。この図１７のマイクロホン１６１で集音された音声信号は、アンプ１６２で増幅され、Ａ／Ｄ（アナログ／ディジタル）変換器１６３でディジタル信号に変換されて、音声符号化部１６０に送られる。この音声符号化部１６０は、上述した図２、図３に示すような構成を有しており、この入力端子１０１に上記Ａ／Ｄ変換器１６３からのディジタル信号が入力される。音声符号化部１６０では、上記図２、図３と共に説明したような符号化処理が行われ、図２、図３の各出力端子からの出力信号は、音声符号化部１６０の出力信号として、伝送路符号化部１６４に送られる。

伝送路符号化部１６４では、いわゆるチャネルコーディング処理が施され、その出力信号が変調回路１６５に送られて変調され、Ｄ／Ａ（ディジタル／アナログ）変換器１６６、ＲＦアンプ１６７を介して、アンテナ１６８に送られる。

【０１４４】

また、図１８は、上記図５、図１４に示したような構成を有する音声復号化部２６０を用いて成る携帯端末の受信側構成を示している。この図１８のアンテナ２６１で受信された音声信号は、ＲＦアンプ２６２で増幅され、Ａ／Ｄ（アナログ／ディジタル）変換器２６３を介して、復調回路２６４に送られ、復調信号が伝送路復号化部２６５に送られる。２６４からの出力信号は、上記図５、図１４に示すような構成を有する音声復号化部２６０に送られる。音声復号化部２６０では、上記図５、図１４と共に説明したような復号化処理が施され、図５、図１４の出力端子２０１からの出力信号が、音声復号化部２６０からの信号としてＤ／Ａ（ディジタル／アナログ）変換器２６６に送られる。このＤ／Ａ変換器２６６からのアナログ音声信号がスピーカ２６８に送られる。

【０１４５】

【発明の効果】

本発明に係る音声信号の再生方法及び装置は、符号化パラメータを補間処理して所望の時刻に対応する変更符号化パラメータを求め、この変更符号化パラメータに基づいて有声音の場合には有声音用の合成フィルタにより有声音を合成し、無声音の場合には無声音用の合成フィルタにより無声音を合成し、上記それぞれの合成フィルタからの出力を加算して音声信号を再生するので、広いレンジにわたる任意のレートのスピードコントロールを簡単にかつ音韻，ピッチを不変として高品質に行うことができる。

【０１４６】

また、本発明に係る音声信号の再生方法によれば、入力音声信号を時間軸上で所定のブロック単位毎に区分して得た符号化パラメータを用い、符号化時とは異なる長さのブロックで音声を再生するので、広いレンジにわたる任意のレートのスピードコントロールを簡単にかつ音韻，ピッチを不変として高品質に行うことができる。

【０１４７】

また、本発明に係る音声復号化方法及び装置は、入力されたデータのハーモニックスの基本周波数と所定の帯域内における個数を変換し、上記入力された各ハーモニックスにおけるスペクトルの大きさを表すデータの個数を補間処理することによってピッチを変更するので、簡単な構成で任意のピッチに変更することができる。

【０１４８】

この場合、音声圧縮のデコーダとテキスト音声合成の音声合成器を兼用させることが挙げられる。ここで、定型の発話には圧縮・伸張によりクリアな再生音を得て、特殊な合成にはテキスト合成あるいは規則合成を用いることにより、効率的な音声出力システムを構成することができる。

【図面の簡単な説明】

【図１】本発明に係る音声信号の再生方法及び装置の実施の形態となる音声信号再生装置の基本構成を示すブロック図である。

【図２】上記音声信号再生装置の符号化部の概略構成を示すブロック図である。

【図３】上記符号化部の詳細な構成を示すブロック図である。

【図４】上記音声信号再生装置の復号化部の概略構成を示すブロック図である。

【図５】上記復号化部の詳細な構成を示すブロック図である。

【図６】上記復号化部の変更符号化パラメータ算出部の動作を説明するためのフローチャートである。

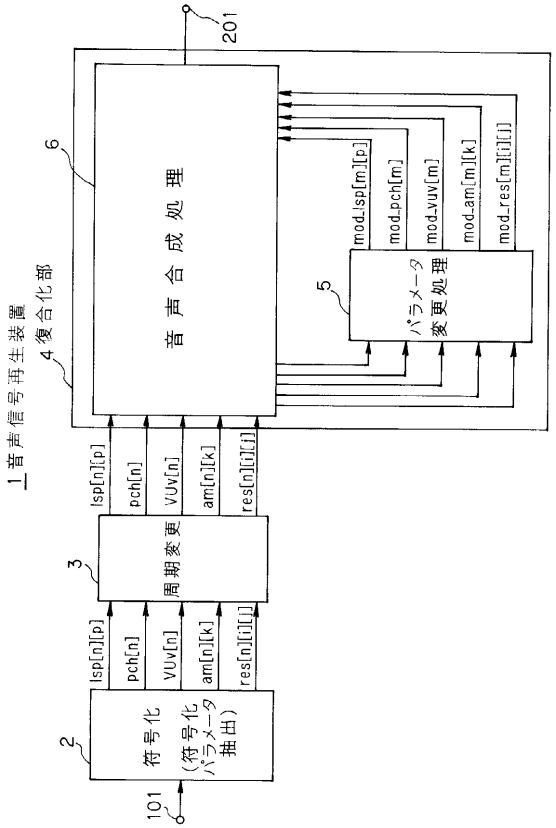
【図７】上記変更符号化パラメータ算出部で得られる変更符号化パラメータを時間軸上で表現した模式図である。

【図８】上記変更符号化パラメータ算出部の補間処理の動作を詳細に説明するためのフローチャートである。

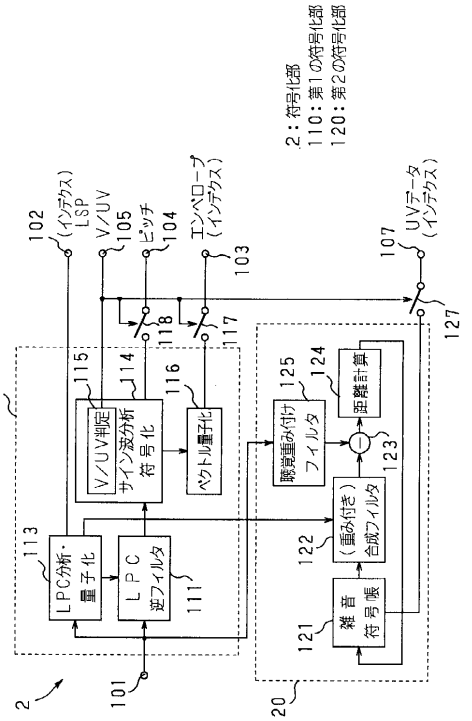
【図９】上記補間処理の動作を説明するための模式図である。

【図１０】上記変更符号化パラメータ算出部の動作例を説明するための模式図である。
【図１１】上記変更符号化パラメータ算出部の他の動作例を説明するための模式図である。
【図１２】復号化部にてフレーム長を可変にしてスピードを速くコントロールする場合の動作を説明するための図である。
【図１３】復号化部にてフレーム長を可変にしてスピードを遅くコントロールする場合の動作を説明するための図である。
【図１４】上記復号化部の詳細な他の構成を示すブロック図である。
【図１５】音声合成装置への適用例を示すブロック図である。
【図１６】テキスト音声合成装置への適用例を示すブロック図である。
【図１７】上記符号化部が用いられる携帯端末の送信側構成を示すブロック図である。
【図１８】上記復号化部が用いられる携帯端末の受信側構成を示すブロック図である。
【符号の説明】
１　音声信号再生装置、２　符号化部、３　周期変更部、４　復号化部、５　パラメータ変更処理部、６　音声合成処理部

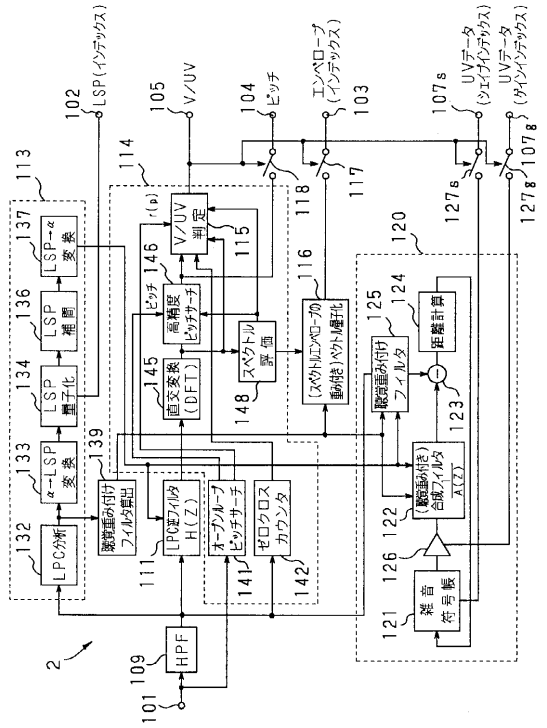
【図１】



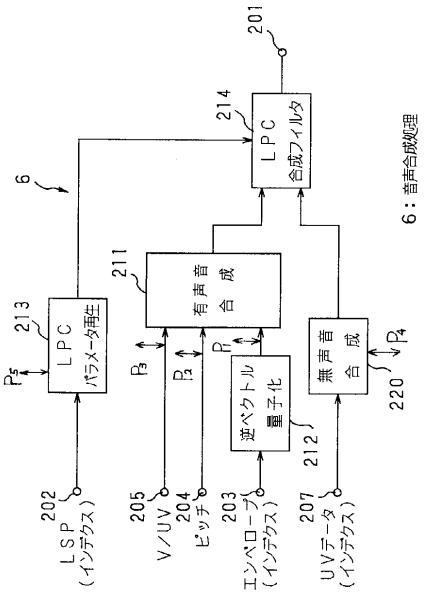
【図２】



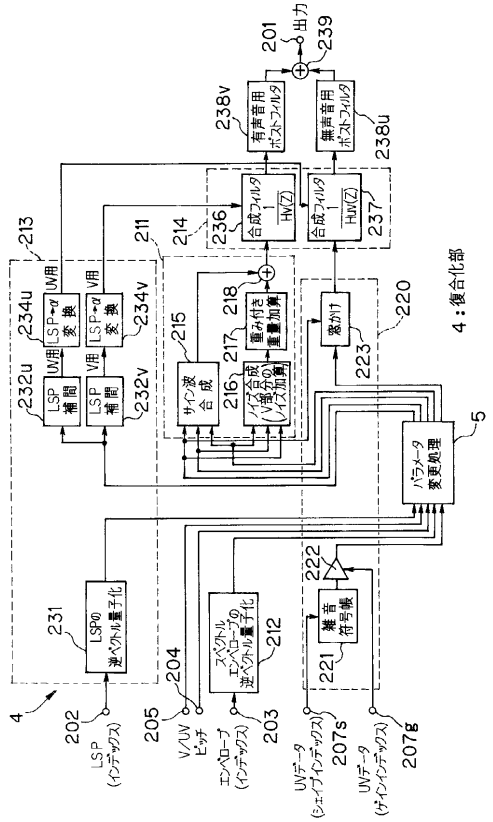
【図３】


LSP(インデックス)
V/UV
ピッチ
エンベロープ(インデックス)
UVデータ(シェイプインデックス)
UVデータ(ゲインインデックス)
HPF
LPC分析
α→LSP変換
LSP量子化
LSP補間
LSP→α変換
LPC逆フィルタ H(Z)
オープンループピッチサーチ
ゼロクロスカウンタ
直交変換(DFT)
高精度ピッチサーチ
V/UV判定
スペクトル評価
(スペクトルエンベロープの動的きパクトル量子化)
聴覚重み付けフィルタ
距離計算
(聴覚重み付き)合成フィルタ A(Z)
雑音符号帳
2
101
109
111
113
114
115
116
117
118
120
121
122
123
124
125
126
127s
127g
132
133
134
136
137
139
141
142
145
146
148
107s
107g

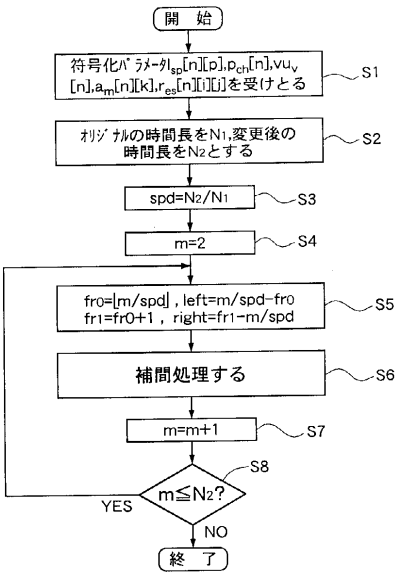
【図４】


LSP(インデックス)
V/UV
ピッチ
エンベロープ(インデックス)
UVデータ(インデックス)
LPC パラメータ再生
有声音合成
逆ベクトル量子化
無声音合成
LPC合成フィルタ
6：音声合成処理
6
201
202
203
204
205
207
211
212
213
214
220

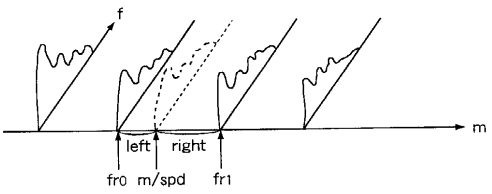
【図５】


LSP(インデックス)
V/UV
ピッチ
エンベロープ(インデックス)
UVデータ(シェイプインデックス)
UVデータ(ゲインインデックス)
LSP補間
LSP→α変換
サイン波合成
ノイズ合成(V部分のノイズ加算)
重み付き重量加算
窓かけ
パラメータ変更処理
合成フィルタ 1 Hv(Z)
合成フィルタ 1 Huv(Z)
有声音用ポストフィルタ
無声音用ポストフィルタ
出力
4：復合化部
5
201
202
203
204
205
207s
207g
211
212
213
214
215
216
217
218
220
221
222
223
231
232u
232v
234u
234v
236
237
238u
238v
239

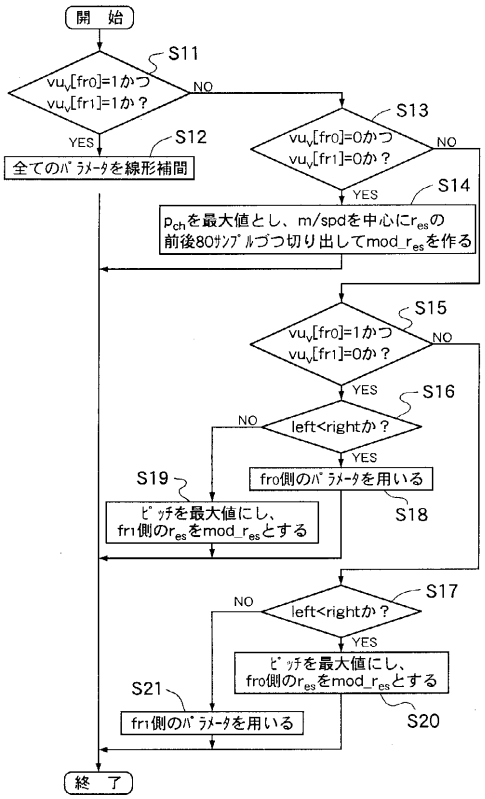
【図６】


開始
符号化パラメータl_{sp}[n][p],p_{ch}[n],vu_v[n],a_m[n][k],r_{es}[n][i][j]を受けとる
S1
オリジナルの時間長をN_1,変更後の時間長をN_2とする
S2
spd=N_2/N_1
S3
m=2
S4
fr0=⌊m/spd⌋, left=m/spd−fr0
fr1=fr0+1, right=fr1−m/spd
S5
補間処理する
S6
m=m+1
S7
m≦N_2?
S8
YES
NO
終了

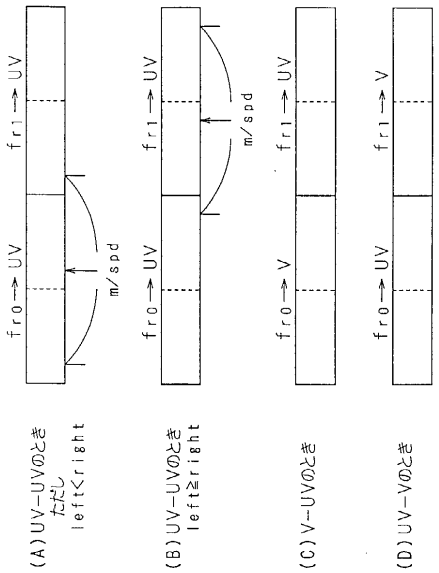
【図７】


f
m
left
right
fr0
m/spd
fr1

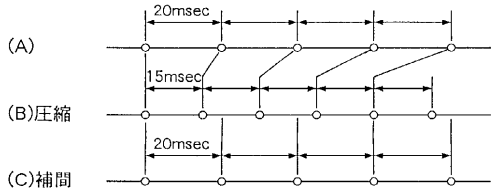
【図８】



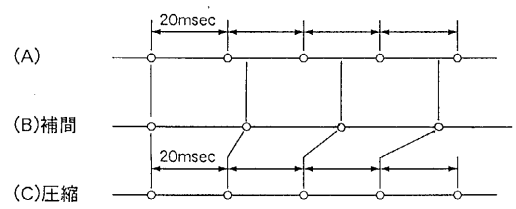
【図９】



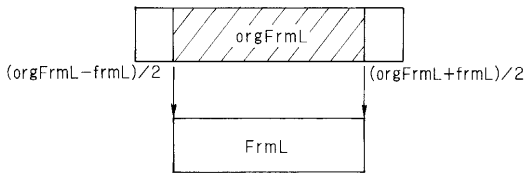
【図１０】



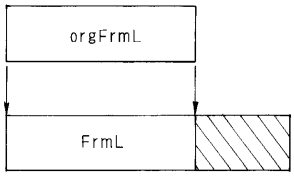
【図１１】



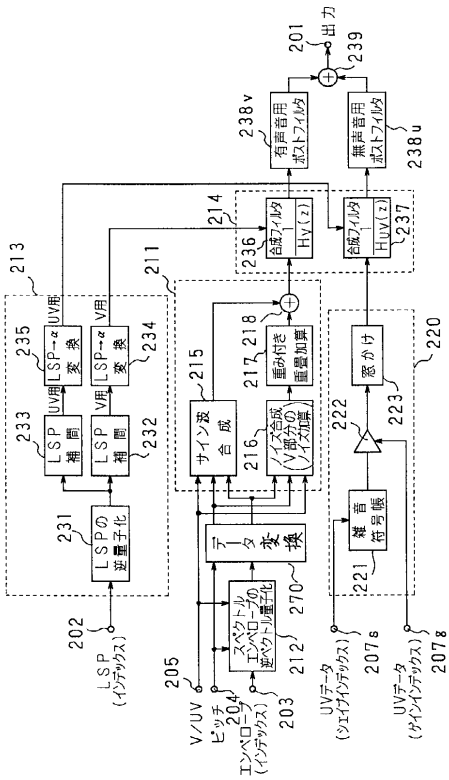
【図１２】



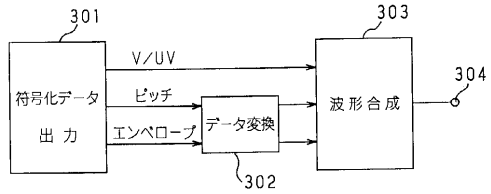
【図１３】



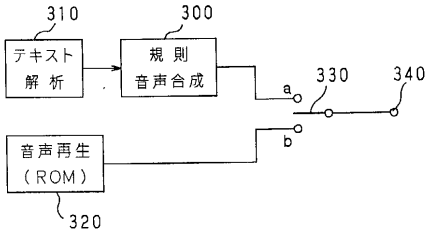
【図１４】



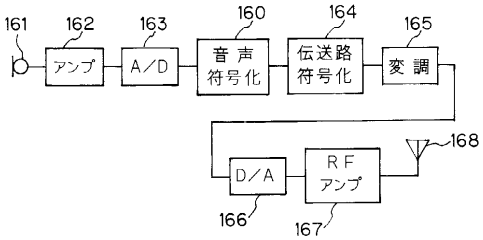
【図１５】



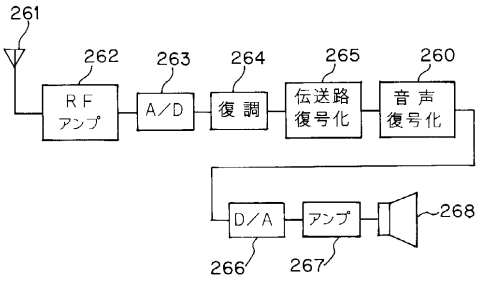
【図１６】



【図１７】



【図１８】



フロントページの続き

(72)発明者　松本　淳
東京都品川区北品川６丁目７番３５号　ソニー株式会社内
(72)発明者　大森　士郎
東京都品川区北品川６丁目７番３５号　ソニー株式会社内

合議体
審判長　板橋　通孝
審判官　廣川　浩
審判官　加藤　恵一

(56)参考文献　特開平０７－０４４１９３（ＪＰ，Ａ）
特開平０３－２１１５９９（ＪＰ，Ａ）
特開平０７－２２５５９９（ＪＰ，Ａ）