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

特許第4005154号
(P4005154)

(45)発行日　平成19年11月7日(2007.11.7)　(24)登録日　平成19年8月31日(2007.8.31)

(51)Int.Cl.		F I		
G10L 19/04	(2006.01)	G10L	19/04	
G10L 19/08	(2006.01)	G10L	19/08	D
G10L 19/12	(2006.01)	G10L	19/12	Z
G10L 19/14	(2006.01)	G10L	19/14	400F

請求項の数 2　(全 43 頁)

(21)出願番号　特願平7-279412
(22)出願日　平成7年10月26日(1995.10.26)
(65)公開番号　特開平9-127989
(43)公開日　平成9年5月16日(1997.5.16)
審査請求日　平成14年10月28日(2002.10.28)
審判番号　不服2005-10041(P2005-10041/J1)
審判請求日　平成17年5月26日(2005.5.26)

(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　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
(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

最終頁に続く

(54)【発明の名称】音声復号化方法及び装置

(57)【特許請求の範囲】

【請求項１】

入力音声信号の有声音部分に対して短期予測残差を求め、上記求められた短期予測残差に基づいてサイン波分析符号化パラメータを求め、上記サイン波分析符号化パラメータに対して聴覚重み付きのベクトル量子化を施し、上記入力音声信号の無声音部分に対しては波形符号化により符号化することにより生成された符号化データを受け取って復号化する音声復号化方法であって、

上記ベクトル量子化されたサイン波分析符号化パラメータを逆ベクトル量子化して、サイン波合成により有声音部分の短期予測残差を求める工程と、

上記有声音部分の短期予測残差に基づいて有声音部分の時間軸波形を合成する第１の予測合成フィルタ処理工程と、

上記波形符号化された無声音部分について復号化を行い、無声音部分の短期予測残差を求める工程と、

上記無声音部分の短期予測残差に窓をかける窓かけ工程と、

上記窓をかけられた無声音部分の短期予測残差に基づいて無声音部分の時間軸波形を合成する第２の予測合成フィルタ処理工程と

を有することを特徴とする音声復号化方法。

【請求項２】

入力音声信号の有声音部分に対して短期予測残差を求め、上記求められた短期予測残差に基づいてサイン波分析符号化パラメータを求め、上記サイン波分析符号化パラメータに

対して聴覚重み付きのベクトル量子化を施し、上記入力音声信号の無声音部分に対しては波形符号化により符号化することにより生成された符号化データを受け取って復号化する音声復号化装置であって、

上記ベクトル量子化されたサイン波分析符号化パラメータを逆ベクトル量子化して、サイン波合成により有声音部分の短期予測残差を求める手段と、

上記有声音部分の短期予測残差に基づいて有声音部分の時間軸波形を合成する第１の予測合成フィルタ処理手段と、

上記波形符号化された無声音部分について復号化を行い、無声音部分の短期予測残差を求める手段と、

上記無声音部分の短期予測残差に窓をかける窓かけ手段と、

上記窓をかけられた無声音部分の短期予測残差に基づいて無声音部分の時間軸波形を合成する第２の予測合成フィルタ処理手段と

を有することを特徴とする音声復号化装置。

【発明の詳細な説明】

【０００１】

【発明の属する技術分野】

本発明は、入力音声信号をブロック単位で区分して、区分されたブロックを単位として符号化処理を施すことにより得られた信号を復号化する音声復号化方法及び装置に関するものである。

【０００２】

【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。

【０００３】

音声信号等の高能率符号化の例として、ハーモニック（Harmonic）符号化、ＭＢＥ（Multiband Excitation: マルチバンド励起）符号化等のサイン波分析符号化や、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。

【０００４】

この音声信号等の高能率符号化を用いた音声信号符号化装置では、入力音声信号の短期予測残差、例えばＬＰＣ（線形予測符号化）残差をマルチバンド励起（ＭＢＥ）符号化のようなサイン波分析（sinusoidal analysis ）符号化を行い、得られるスペクトルエンベロープ振幅データをベクトル量子化して、コードブックインデクスを出力する。

【０００５】

【発明が解決しようとする課題】

ところで、上記音声信号符号化装置において、上記ベクトル量子化により得られるコードブックインデクスを含む符号化データのビットレートは一定であり、符号化データのビットレートを可変にすることはできない。

【０００６】

また、上記符号化データが、例えばMビットである場合には、この符号化データを復号化する音声信号復号化装置は、Mビットの音声信号復号化装置でなければならない。即ち、音声信号復号化装置では、符号化データのビット数と同じビット数の復号化データしか得ることができず、復号化データのビット数を可変にすることができない。

【０００７】

そこで、本発明は上述の実情に鑑み、符号化データのビットレートを可変にすることができる音声符号化処理によって生成された符号化データを受け取って復号化する音声復号化方法及び装置を提供するものである。

【０００８】
【課題を解決するための手段】
本発明に係る音声復号化方法は、上述した課題を解決するために、入力音声信号の有声音部分に対して短期予測残差を求め、上記求められた短期予測残差に基づいてサイン波分析符号化パラメータを求め、上記サイン波分析符号化パラメータに対して聴覚重み付きのベクトル量子化を施し、上記入力音声信号の無声音部分に対しては波形符号化により符号化することにより生成された符号化データを受け取って復号化する音声復号化方法であって、上記ベクトル量子化されたサイン波分析符号化パラメータを逆ベクトル量子化して、サイン波合成により有声音部分の短期予測残差を求める工程と、上記有声音部分の短期予測残差に基づいて有声音部分の時間軸波形を合成する第１の予測合成フィルタ処理工程と、上記波形符号化された無声音部分について復号化を行い、無声音部分の短期予測残差を求める工程と、上記無声音部分の短期予測残差に窓をかける窓かけ工程と、上記窓をかけられた無声音部分の短期予測残差に基づいて無声音部分の時間軸波形を合成する第２の予測合成フィルタ処理工程とを有することを特徴とする。
また、本発明に係る音声復号化装置は、上述した課題を解決するために、入力音声信号の有声音部分に対して短期予測残差を求め、上記求められた短期予測残差に基づいてサイン波分析符号化パラメータを求め、上記サイン波分析符号化パラメータに対して聴覚重み付きのベクトル量子化を施し、上記入力音声信号の無声音部分に対しては波形符号化により符号化することにより生成された符号化データを受け取って復号化する音声復号化装置であって、上記ベクトル量子化されたサイン波分析符号化パラメータを逆ベクトル量子化して、サイン波合成により有声音部分の短期予測残差を求める手段と、上記有声音部分の短期予測残差に基づいて有声音部分の時間軸波形を合成する第１の予測合成フィルタ処理手段と、上記波形符号化された無声音部分について復号化を行い、無声音部分の短期予測残差を求める手段と、上記無声音部分の短期予測残差に窓をかける窓かけ手段と、上記窓をかけられた無声音部分の短期予測残差に基づいて無声音部分の時間軸波形を合成する第２の予測合成フィルタ処理手段とを有することを特徴とする。
【０００９】
【発明の実施の形態】
以下、本発明に係る好ましい実施の形態について説明する。
先ず、図１は、本発明に係る音声符号化方法及び音声信号符号化装置の実施の形態が適用された音声信号符号化装置の基本構成を示すブロック図であり、入力音声信号の短期予測残差を求める手段であるＬＰＣ逆フィルタ１１１と、上記短期予測残差からサイン波分析符号化パラメータを求める手段であるサイン波分析符号化部１１４と、上記サイン波分析符号化パラメータに対して聴覚重み付きのベクトル量子化を施す手段であるベクトル量子化部１１６と、上記入力音声信号を位相伝送を行う波形符号化により符号化する手段である第２の符号化部１２０とを有して成る。
【００１０】
また、図２は、本発明に係る音声復号化方法及び音声復号化装置の実施の形態が適用された、図１の音声信号符号化装置に対応する音声信号復号化装置の基本構成を示すブロック図であり、図３は、図１に示した音声信号符号化装置のより具体的な構成を示すブロック図であり、図４は、図２に示した音声信号復号化装置のより具体的な構成を示すブロック図である。
【００１１】
先ず、図１、図２、図３、及び図４に示す各ブロック図の構成について説明する。
【００１２】
ここで、図１の音声信号符号化装置の基本的な考え方は、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を用いるようにすることである。

【００１３】

上記第１の符号化部１１０には、例えばＬＰＣ残差をハーモニック符号化やマルチバンド励起（ＭＢＥ）符号化のようなサイン波分析符号化を行う構成が用いられる。上記第２の符号化部１２０には、例えば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用いた符号励起線形予測（ＣＥＬＰ）符号化の構成が用いられる。

【００１４】

図１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から得られた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の制御信号として送られており、上述した有声音（Ｖ）のとき上記インデクス及びピッチが選択されて各出力端子１０３及び１０４からそれぞれ取り出される。

【００１５】

図１の第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、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を聴覚重み付けフィルタ１２５を介して得られた音声との誤差を取り出し、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ベクトルを雑音符号帳１２１でサーチするような、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。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。

【００１６】

次に、図２は、本発明に係る音声復号化方法の一実施の形態が適用された音声信号復号化装置として、上記図１の音声信号符号化装置に対応する音声信号復号化装置の基本構成を示すブロック図である。

【００１７】

この図２において、入力端子２０２には上記図１の出力端子１０２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、上記図１の各出力端子１０３、１０４、及び１０５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、上記図１の出力端子１０７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。

【００１８】

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て有声音合成部２１１に送られる。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）残差を合成するものであり、この有声音合成部２１１には入力端子２０４及び２０５からのピッチ及びＶ／ＵＶ判定出力も供給されている。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。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残差が取り出される。このＬＰＣ残差もＬＰＣ合成フィルタ２１４に送られる。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残差とがそれぞれ独立に、ＬＰＣ合成処理が施される。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。ここで入力端子２０２からのＬＳＰのインデクスは、ＬＰＣパラメータ再生部２１３に送られて、ＬＰＣのαパラメータが取り出され、これがＬＰＣ合成フィルタ２１４に送られる。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。

【００１９】

次に、上記図１に示した音声信号符号化装置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１の各部と対応する部分には同じ指示符号を付している。

【００２０】

この図３に示された音声信号符号化装置において、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。

【００２１】

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。

【００２２】

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ー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。

【００２３】

α→ＬＳＰ変換回路１３３からのＬＳＰパラメータは、ＬＳＰ量子化器１３４によりマトリクスあるいは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２フレーム分まとめて、マトリクス量子化及びベクトル量子化している。

【００２４】

このＬＳＰ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。

【００２５】

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。

【００２６】

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。

【００２７】

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。

【００２８】

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。

【００２９】

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。

【００３０】

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。

【００３１】

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。

【００３２】
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。

【００３３】
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。

【００３４】
Ｖ／ＵＶ（有声音／無声音）判定部１１５では、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とに基づいて、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。

【００３５】
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１１９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。

【００３６】
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。

【００３７】
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１２

２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。

【００３８】

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。

【００３９】

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。

【００４０】

次に、図４は、上記図２に示した本発明に係る実施の形態としての音声信号復号化装置のより具体的な構成を示している。この図４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。

【００４１】

この図４において、入力端子２０２には、上記図１、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。

【００４２】

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。

【００４３】

また、図４の入力端子２０３には、上記図１、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１、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１、図３の端子１０５からのＶ／ＵＶ判定データが供給されている。

【００４４】

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、上記データ数

変換に対応する逆変換が施されて、スペクトルエンベロープのデータとなって、有声音合成部２１１のサイン波合成回路２１５に送られている。

【００４５】

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。

【００４６】

サイン波合成回路２１５には、入力端子２０４からのピッチ及び入力端子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、上述した図１、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。

【００４７】

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。

【００４８】

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。

【００４９】

次に、図４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。

【００５０】

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。

【００５１】

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。

【００５２】

上述した音声信号符号化装置では、出力データのビットレートが可変されて出力される。具体的には、出力データのビットレートを、低ビットレートと高ビットレートとに切り換えることができる。例えば、低ビットレートを２ｋbpsとし、高ビットレートを６ｋbpsと

する場合には、以下の表１に示す各ビットレートのデータが出力される。

【００５３】

【表１】

	2Kbps		6Kbps	
ﾋﾟｯﾁﾃﾞｰﾀ(有声音時)	8bits/20msec		8bits/20msec	
V/UV判定出力	1bit/20msec		1bit/20msec	
LSP量子化ｲﾝﾃﾞｸｽ	32bits/40msec		48bits/40msec	
有声音時 (V)	ｲﾝﾃﾞｸｽ 15bits/20msec		ｲﾝﾃﾞｸｽ 87bits/20msec	
	ｼｪｲﾌﾟ(1段目)	5+5bits/20msec	ｼｪｲﾌﾟ(1段目)	5+5bits/20msec
	ｹﾞｲﾝ	5bits/20msec	ｹﾞｲﾝ	5bits/20msec
			ｲﾝﾃﾞｸｽ(2段目)	72bits/20msec
無声音時 (UV)	ｲﾝﾃﾞｸｽ 11bits/10msec		ｲﾝﾃﾞｸｽ 23bits/5msec	
	ｼｪｲﾌﾟ(1段目)	7bits/10msec	ｼｪｲﾌﾟ(1段目)	9bits/5msec
	ｹﾞｲﾝ	4bits/10msec	ｹﾞｲﾝ	6bits/5msec
			ｼｪｲﾌﾟ(2段目)	5bits/5msec
			ｹﾞｲﾝ	3bits/5msec
有声音時	40bits/20msec		120bits/20msec	
無声音時	39bits/20msec		117bits/20msec	

【００５４】

出力端子１０４からのピッチデータについては、有声音時に、常に８bits／２０msecで出力され、出力端子１０５から出力されるＶ／ＵＶ判定出力は、常に１bit／２０msecである。出力端子１０２から出力されるＬＳＰ量子化のインデクスは、３２bits／４０msecと４８bits／４０msecとの間で切り換えが行われる。また、出力端子１０３から出力される有声音時（Ｖ）のインデクスは、１５bits／２０msecと８７bits／２０msecとの間で切り換えが行われ、出力端子１０７ｓ、１０７ｇから出力される無声音時（ＵＶ）のインデクスは、１１bits／１０msecと２３bits／５msecとの間で切り換えが行われる。これにより、有声音時（Ｖ）の出力データは、２kbpsでは４０bits／２０msecとなり、６kbpsでは１２０bits／２０msecとなる。また、無声音時（ＵＶ）の出力データは、２kbpsでは３９bits／２０msecとなり、６kbpsでは１１７bits／２０msecとなる。

【００５５】

尚、上記ＬＳＰ量子化のインデクス、有声音時（Ｖ）のインデクス、及び無声音時（ＵＶ）のインデクスについては、後述する各部の構成と共に説明する。

【００５６】

次に、図５及び図６を用いて、ＬＳＰ量子化器１３４におけるマトリクス量子化及びベクトル量子化について詳細に説明する。

【００５７】

上述のように、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、ＬＳＰパラメータに変換される。例えば、ＬＰＣ分析回路１３２でＰ次のＬＰＣ分析を行う場合には、αパラメータはＰ個算出される。このＰ個のαパラメータは、ＬＳＰパラメータに変換され、バッファ６１０に保持される。

【００５８】

このバッファ６１０からは、２フレーム分のＬＳＰパラメータが出力される。２フレーム分のＬＳＰパラメータはマトリクス量子化部６２０でマトリクス量子化される。マトリクス量子化部６２０は、第１のマトリクス量子化部６２０_{1}と第２のマトリクス量子化部６２０_{2}とから成る。２フレーム分のＬＳＰパラメータは、第１のマトリクス量子化部６２０_{1}でマトリクス量子化され、これにより得られる量子化誤差が、第２のマトリクス量子化部６２０_{2}でさらにマトリクス量子化される。これらのマトリクス量子化により、時間軸方向の相関を取り除く。

【００５９】

マトリクス量子化部６２０_{2}からの２フレーム分の量子化誤差は、ベクトル量子化部６４０に入力される。ベクトル量子化部６４０は、第１のベクトル量子化部６４０_{1}と第２のベクトル量子化部６４０_{2}とから成る。さらに、第１のベクトル量子化部６４０_{1}は、２つのベクトル量子化部６５０、６６０から成り、第２のベクトル量子化部６４０_{2}は、２つのベクトル量子化部６７０、６８０から成る。第１のベクトル量子化部６４０_{1}のベクトル量子化部６５０、６６０で、マトリクス量子化部６２０からの量子化誤差が、それぞれ１フレーム毎にベクトル量子化される。これにより得られる量子化誤差ベクトルは、第２のベクトル量子化部６４０_{2}のベクトル量子化部６７０、６８０で、さらにベクトル量子化される。これらのベクトル量子化により、周波数軸方向の相関を処理する。

【００６０】

このように、マトリクス量子化を施す工程を行うマトリクス量子化部６２０は、第１のマトリクス量子化工程を行う第１のマトリクス量子化部６２０_{1}と、この第１のマトリクス量子化による量子化誤差をマトリクス量子化する第２のマトリクス量子化工程を行う第２のマトリクス量子化部６２０_{2}とを少なくとも有し、上記ベクトル量子化を施す工程を行うベクトル量子化部６４０は、第１のベクトル量子化工程を行う第１のベクトル量子化部６４０_{1}と、この第１のベクトル量子化の際の量子化誤差ベクトルをベクトル量子化する第２のベクトル量子化工程を行う第２のベクトル量子化部６４０_{2}とを少なくとも有する。

【００６１】

次に、マトリクス量子化及びベクトル量子化について具体的に説明する。

【００６２】

バッファ６１０に保持された、２フレーム分のＬＳＰパラメータ、すなわち１０×２の行列は、マトリクス量子化器６２０_{1}に送られる。上記第１のマトリクス量子化器６２０_{1}では、２フレーム分のＬＳＰパラメータが加算器６２１を介して重み付き距離計算器６２３に送られ、最小となる重み付き距離が算出される。

【００６３】

この第１のマトリクス量子化部６２０_{1}によるコードブックサーチ時の歪尺度d_{MQ1}は、ＬＳＰパラメータX_{1}、量子化値$\underline{X}_{1}$を用い、（１）式で示す。

【００６４】

【数１】

$$d_{MQ1}(X_1, X_1') = \sum_{t=0}^{1}\sum_{i=1}^{P} W(t,i)\left(X_1(t,i) - X_1'(t,i)\right)^2 \quad \cdots(1)$$

【００６５】
ここで、ｔはフレーム番号、ｉはＰ次元の番号を示す。
【００６６】
また、このときの、周波数軸方向及び時間軸方向に重みの制限を考慮しない場合の重みＷを（２）式で示す。
【００６７】
【数２】

$$W(t,i) = \frac{1}{X(t,i+1) - X(t,i)} + \frac{1}{X(t,i) - X(t,i-1)} \quad \cdots(2)$$

ただしｉ＝１のときは$X_{i-1}=0$
ｉ＝ｐのときは$X_{i+1}=\pi$

【００６８】
この（２）式の重みＷは、後段のマトリクス量子化及びベクトル量子化でも用いられる。
【００６９】
算出された重み付き距離はマトリクス量子化器（MQ_1）６２２に送られて、マトリクス量子化が行われる。このマトリクス量子化により出力される８ビットのインデクスは信号切換器６９０に送られる。また、マトリクス量子化による量子化値は、加算器６２１で、バッファ６１０からの次の２フレーム分のＬＳＰパラメータから減算される。重み付き距離計算器６２３では、加算器６２１からの出力を用いて、最小となる重み付き距離が算出される。このように、２フレーム毎に、順次、重み付き距離計算器６２３では重み付き距離が算出されて、マトリクス量子化器６２２でマトリクス量子化が行われる。また、加算器６２１からの出力は、第２のマトリクス量子化部620_2の加算器６３１に送られる。
【００７０】
第２のマトリクス量子化部620_2でも第１のマトリクス量子化部620_1と同様にして、マトリクス量子化を行う。上記加算器６２１からの出力は、加算器６３１を介して重み付き距離計算器６３３に送られ、最小となる重み付き距離が算出される。
【００７１】
この第２のマトリクス量子化部620_2によるコードブックサーチ時の歪尺度d_{MQ2}を、第１のマトリクス量子化部620_1からの量子化誤差X_2、量子化値$\underline{X_2}$により、（３）式で示す。
【００７２】
【数３】

$$d_{MQ2}(X_2, X_2') = \sum_{t=0}^{1}\sum_{i=1}^{P} W(t,i)\left(X_2(t,i) - X_2'(t,i)\right)^2 \quad \cdots(3)$$

【００７３】
この重み付き距離はマトリクス量子化器（MQ_2）６３２に送られて、マトリクス量子化が行われる。このマトリクス量子化により出力される８ビットのインデクスは信号切換器６９０に送られる。また、マトリクス量子化による量子化値は、加算器６３１で、次の２フレーム分の量子化誤差から減算される。重み付き距離計算器６３３では、加算器６３１からの出力を用いて、最小となる重み付き距離が順次算出される。また、加算器６３１からの出力は、第１のベクトル量子化部640_1の加算器６５１、６６１に１フレームずつ送られる。
【００７４】
この第１のベクトル量子化部640_1では、１フレーム毎にベクトル量子化が行われる。加算器６３１からの出力は、１フレーム毎に、加算器６５１、６６１を介して重み付き距離計算器６５３、６６３にそれぞれ送られ、最小となる重み付き距離が算出される。
【００７５】
量子化誤差X_2と量子化値$\underline{X}_2$との差分は、１０×２の行列であり、
$X_2 - X_2' = [\underline{X}_{3-1}, \underline{X}_{3-2}]$
と表すときの、この第１のベクトル量子化部640_1のベクトル量子化器６５２、６６２によるコードブックサーチ時の歪尺度d_{VQ1}、d_{VQ2}を、（４）、（５）式で示す。
【００７６】
【数４】

$$d_{VQ1}\left(\underline{X}_{3-1}, \underline{X}'_{3-1}\right) = \sum_{i=1}^{P} W(0,i)\left(X_{3-1}(0,i) - X'_{3-1}(0,i)\right)^2 \quad \cdots(4)$$

$$d_{VQ2}\left(\underline{X}_{3-2}, \underline{X}'_{3-2}\right) = \sum_{i=1}^{P} W(1,i)\left(X_{3-2}(1,i) - X'_{3-2}(1,i)\right)^2 \quad \cdots(5)$$

【００７７】
この重み付き距離はベクトル量子化器（VQ_1）６５２、ベクトル量子化器（VQ_2）６６２にそれぞれ送られて、ベクトル量子化が行われる。このベクトル量子化により出力される各８ビットのインデクスは信号切換器６９０に送られる。また、ベクトル量子化による量子化値は、加算器６５１、６６１で、次に入力される２フレーム分の量子化誤差ベクトルから減算される。重み付き距離計算器６５３、６６３では、加算器６５１、６６１からの出力を用いて、最小となる重み付き距離が順次算出される。また、加算器６５１、６６１からの出力は、第２のベクトル量子化部640_2の加算器６７１、６８１にそれぞれ送られる。
【００７８】
ここで、
$\underline{X}_{4-1} = X_{3-1} - \underline{X}'_{3-1}$
$\underline{X}_{4-2} = X_{3-2} - \underline{X}'_{3-2}$
と表すときの、この第２のベクトル量子化部640_2のベクトル量子化器６７２、６８２によるコードブックサーチ時の歪尺度d_{VQ3}、d_{VQ4}を、（６）、（７）式で示す。
【００７９】
【数５】

$$d_{VQ3}\left(\underline{X}_{4-1},\underline{X}'_{4-1}\right)=\sum_{i=1}^{P}W(0,i)\left(X_{4-1}(0,i)-X'_{4-1}(0,i)\right)^2 \quad \cdots(6)$$

$$d_{VQ4}\left(\underline{X}_{4-2},\underline{X}'_{4-2}\right)=\sum_{i=1}^{P}W(1,i)\left(X_{4-2}(1,i)-X'_{4-2}(1,i)\right)^2 \quad \cdots(7)$$

【００８０】
この重み付き距離はベクトル量子化器（ＶＱ$_3$）６７２、ベクトル量子化器（ＶＱ$_4$）６８２にそれぞれ送られて、ベクトル量子化が行われる。このベクトル量子化により出力される各８ビットのインデクスは信号切換器６９０に送られる。また、ベクトル量子化による量子化値は、加算器６７１、６８１で、次に入力される２フレーム分の量子化誤差ベクトルから減算される。重み付き距離計算器６７３、６８３では、加算器６７１、６８１からの出力を用いて、最小となる重み付き距離が順次算出される。
【００８１】
また、コードブックの学習時には、上記各歪尺度をもとにして、一般化ロイドアルゴリズム（ＧＬＡ）により学習を行う。
【００８２】
尚、コードブックサーチ時と学習時の歪尺度は、異なる値であっても良い。
【００８３】
上記マトリクス量子化器６２２、６３２、ベクトル量子化器６５２、６６２、６７２、６８２からの各８ビットのインデクスは、信号切換器６９０で切り換えられて、出力端子６９１から出力される。
【００８４】
具体的には、低ビットレート時には、上記第１のマトリクス量子化工程を行う第１のマトリクス量子化部６２０$_1$、上記第２のマトリクス量子化工程を行う第２のマトリクス量子化部６２０$_2$、及び上記第１のベクトル量子化工程を行う第１のベクトル量子化部６４０$_1$での出力を取り出し、高ビットレート時には、上記低ビットレート時の出力に上記第２のベクトル量子化工程を行う第２のベクトル量子化部６４０$_2$での出力を合わせて取り出す。
【００８５】
これにより、２ｋbps時には、３２bits／４０msecのインデクスが出力され、６ｋbps時には、４８bits／４０msecのインデクスが出力される。
【００８６】
また、上記マトリクス量子化部６２０及び上記ベクトル量子化部６４０では、上記ＬＰＣ係数を表現するパラメータの持つ特性に合わせた、周波数軸方向又は時間軸方向、あるいは周波数軸及び時間軸方向に制限を持つ重み付けを行う。
【００８７】
先ず、ＬＳＰパラメータの持つ特性に合わせた、周波数軸方向に制限を持つ重み付けについて説明する。例えば、次数Ｐ＝１０とするとき、ＬＳＰパラメータＸ（ｉ）を、低域、中域、高域の３つの領域として、
Ｌ$_1$＝｛Ｘ（ｉ）｜１≦ｉ≦２｝
Ｌ$_2$＝｛Ｘ（ｉ）｜３≦ｉ≦６｝
Ｌ$_3$＝｛Ｘ（ｉ）｜７≦ｉ≦１０｝
とグループ化する。そして、各グループＬ$_1$、Ｌ$_2$、Ｌ$_3$の重み付けを１／４、１／２、１／４とすると、各グループＬ$_1$、Ｌ$_2$、Ｌ$_3$の周波数軸方向のみに制限を持つ重みは、（８）、（９）、（１０）式となる。
【００８８】
【数６】

$$W'(i)=\frac{W(i)}{\sum_{j=1}^{2}W(j)}\times\frac{1}{4}\quad\cdots(8)$$

$$W'(i)=\frac{W(i)}{\sum_{j=3}^{6}W(j)}\times\frac{1}{2}\quad\cdots(9)$$

$$W'(i)=\frac{W(i)}{\sum_{j=7}^{10}W(j)}\times\frac{1}{4}\quad\cdots(10)$$

【００８９】
これにより、各ＬＳＰパラメータの重み付けは、各グループ内でのみ行われ、その重みは各グループに対する重み付けで制限される。

【００９０】
ここで、時間軸方向からみると、各フレームの重み付けの総和は、必ず１となるので、時間軸方向の制限は１フレーム単位である。この時間軸方向のみに制限を持つ重みは、（１１）式となる。

【００９１】
【数７】

$$W'(i,t)=\frac{W(i,t)}{\sum_{j=1}^{10}\sum_{s=0}^{1}W(j,s)}\quad\cdots(11)\quad\left(1\leqq i\leqq 10,0\leqq t\leqq 1\right)$$

【００９２】
この（１１）式により、周波数軸方向での制限のない、フレーム番号ｔ＝０，１の２つのフレーム間で、重み付けが行われる。この時間軸方向にのみ制限を持つ重み付けは、マトリクス量子化を行う２フレーム間で行う。

【００９３】
また、学習時には、学習データとして用いる全ての音声フレーム、即ち全データのフレーム数Ｔについて、（１２）式により、重み付けを行う。

【００９４】
【数８】

$$W'(i,t)=\frac{W(i,t)}{\sum_{j=1}^{10}\sum_{s=0}^{T}W(j,s)}\quad\cdots(12)\quad\left(1\leqq i\leqq 10,0\leqq t\leqq T\right)$$

【００９５】
また、周波数軸方向及び時間軸方向に制限を持つ重み付けについて説明する。例えば、次数Ｐ＝１０とするとき、ＬＳＰパラメータＸ（ｉ，ｔ）を、低域、中域、高域の３つの領域として、
L_1＝｛Ｘ（ｉ，ｔ）｜１≦ｉ≦２，０≦ｔ≦１｝
L_2＝｛Ｘ（ｉ，ｔ）｜３≦ｉ≦６，０≦ｔ≦１｝
L_3＝｛Ｘ（ｉ，ｔ）｜７≦ｉ≦１０，０≦ｔ≦１｝
とグループ化する。各グループL_1、L_2、L_3の重み付けを１／４、１／２、１／４とすると、各グループL_1、L_2、L_3の周波数軸方向及び時間軸方向に制限を持つ重み付けは、（１３）、（１４）、（１５）式となる。

【００９６】
【数９】

$$W'(i,t)=\frac{W(i,t)}{\sum_{j=1}^{2}\sum_{s=0}^{1}W(j,s)}\times\frac{1}{4}\quad\cdots(13)$$

$$W'(i,t)=\frac{W(i,t)}{\sum_{j=3}^{6}\sum_{s=0}^{1}W(j,s)}\times\frac{1}{2}\quad\cdots(14)$$

$$W'(i,t)=\frac{W(i,t)}{\sum_{j=7}^{10}\sum_{s=0}^{1}W(j,s)}\times\frac{1}{4}\quad\cdots(15)$$

【００９７】
この（１３）、（１４）、（１５）式により、周波数軸方向では３つの帯域毎に、時間軸方向ではマトリクス量子化を行う２フレーム間に重み付けの制限を加えた重み付けを行う。これは、コードブックサーチ時及び学習時共に有効となる。

【００９８】
また、学習時においては、全データのフレーム数について重み付けを行う。ＬＳＰパラメータＸ（ｉ，ｔ）を、低域、中域、高域の３つの領域として、
L_1＝｛Ｘ（ｉ，ｔ）｜１≦ｉ≦２，０≦ｔ≦Ｔ｝
L_2＝｛Ｘ（ｉ，ｔ）｜３≦ｉ≦６，０≦ｔ≦Ｔ｝
L_3＝｛Ｘ（ｉ，ｔ）｜７≦ｉ≦１０，０≦ｔ≦Ｔ｝
とグループ化し、各グループL_1、L_2、L_3の重み付けを１／４、１／２、１／４とすると、各グループL_1、L_2、L_3の周波数軸方向及び時間軸方向に制限を持つ重み付けは、（１６）、（１７）、（１８）式となる。

【００９９】
【数１０】

$$W'(i,t) = \frac{W(i,t)}{\sum_{j=1}^{2}\sum_{s=0}^{T} W(j,s)} \times \frac{1}{4} \quad \cdots(16)$$

$$W'(i,t) = \frac{W(i,t)}{\sum_{j=3}^{6}\sum_{s=0}^{T} W(j,s)} \times \frac{1}{2} \quad \cdots(17)$$

$$W'(i,t) = \frac{W(i,t)}{\sum_{j=7}^{10}\sum_{s=0}^{T} W(j,s)} \times \frac{1}{4} \quad \cdots(18)$$

【０１００】
この（１６）、（１７）、（１８）式により、周波数軸方向では３つの帯域毎に重み付けを行い、時間軸方向では全フレーム間で重み付けを行うことができる。
【０１０１】
さらに、上記マトリクス量子化部６２０及び上記ベクトル量子化部６４０では、上記ＬＳＰパラメータの変化の大きさに応じて重み付けを行う。音声フレーム全体においては少数フレームとなる、Ｖ→ＵＶ、ＵＶ→Ｖの遷移（トランジェント）部において、子音と母音との周波数特性の違いから、ＬＳＰパラメータは大きく変化する。そこで、（１９）式に示す重みを、上述の重みＷ’（ｉ，ｔ）に乗算することにより、上記遷移部を重視する重み付けを行うことができる。
【０１０２】
【数１１】

$$Wd(t) = \sum_{i=1}^{10} \left| X_1(i,t) - X_1(i,t-1) \right|^2 \quad \cdots(19)$$

【０１０３】
尚、（１９）式の代わりに、（２０）式を用いることも考えられる。
【０１０４】
【数１２】

$$Wd(t) = \sum_{i=1}^{10} \sqrt{\left| X_1(i,t) - X_1(i,t-1) \right|} \quad \cdots(20)$$

【０１０５】
このように、ＬＳＰ量子化器１３４では、２段のマトリクス量子化及び２段のベクトル量子化を行うことにより、出力するインデクスのビット数を可変にすることができる。
【０１０６】
次に、ベクトル量子化部１１６の基本構成を図７、図７のベクトル量子化部１１６のより

具体的な構成を図８に示し、ベクトル量子化器１１６におけるスペクトルエンベロープ（Ａｍ）の重み付きベクトル量子化の具体例について説明する。

【０１０７】

先ず、図３の音声信号符号化装置において、スペクトル評価部１４８の出力側あるいはベクトル量子化器１１６の入力側に設けられたスペクトルエンベロープの振幅のデータ数を一定個数にするデータ数変換の具体例について説明する。

【０１０８】

このデータ数変換には種々の方法が考えられるが、本実施の形態においては、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を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）のデータに変換している。

【０１０９】

図７の重み付きベクトル量子化を行うベクトル量子化器１１６は、第１のベクトル量子化工程を行う第１のベクトル量子化部５００と、この第１のベクトル量子化部５００における第１のベクトル量子化の際の量子化誤差ベクトルを量子化する第２のベクトル量子化工程を行う第２のベクトル量子化部５１０とを少なくとも有する。この第１のベクトル量子化部５００は、いわゆる１段目のベクトル量子化部であり、第２のベクトル量子化部５１０は、いわゆる２段目のベクトル量子化部である。

【０１１０】

第１のベクトル量子化部５００の入力端子５０１には、スペクトル評価部１４８の出力ベクトル$\underline{X}$、即ち一定個数Ｍのエンベロープデータが入力される。この出力ベクトル$\underline{X}$は、ベクトル量子化器５０２で重み付きベクトル量子化される。これにより、ベクトル量子化器５０２から出力されるシェイプインデクスは出力端子５０３から出力され、また、量子化値$\underline{X}_0$’は出力端子５０４から出力されると共に、加算器５０５、５１３に送られる。加算器５０５では、出力ベクトル$\underline{X}$から量子化値$\underline{X}_0$’が減算されて、複数次元の量子化誤差ベクトル$\underline{Y}$が得られる。

【０１１１】

この量子化誤差ベクトル$\underline{Y}$は、第２のベクトル量子化部５１０内のベクトル量子化部５１１に送られる。このベクトル量子化部５１１は、複数個のベクトル量子化器で構成され、図７では、２個のベクトル量子化器511_1、511_2から成る。量子化誤差ベクトル$\underline{Y}$は次元分割されて、２個のベクトル量子化器511_1、511_2で、それぞれ重み付きベクトル量子化される。これらのベクトル量子化器511_1、511_2から出力されるシェイプインデクスは、出力端子512_1、512_2からそれぞれ出力され、また、量子化値$\underline{Y}_1$’、$\underline{Y}_2$’は次元方向に接続されて、加算器５１３に送られる。この加算器５１３では、量子化値$\underline{Y}_1$’、$\underline{Y}_2$’と量子化値$\underline{X}_0$’とが加算されて、量子化値$\underline{X}_1$’が生成される。この量子化値$\underline{X}_1$’は出力端子５１４から出力される。

【０１１２】

これにより、低ビットレート時には、上記第１のベクトル量子化部５００による第１のベクトル量子化工程での出力を取り出し、高ビットレート時には、上記第１のベクトル量子化工程での出力及び上記第２の量子化部５１０による第２のベクトル量子化工程での出力を取り出す。

【０１１３】

具体的には、図８に示すように、ベクトル量子化器１１６内の第１のベクトル量子化部５００のベクトル量子化器５０２は、Ｌ次元、例えば４４次元の２ステージ構成としている。

【０１１４】

すなわち、４４次元でコードブックサイズが３２のベクトル量子化コードブックからの出

力ベクトルの和に、ゲインg_iを乗じたものを、４４次元のスペクトルエンベロープベクトル$\underline{X}$の量子化値$\underline{X}_0$’として使用する。これは、図８に示すように、２つのシェイプコードブックをＣＢ０、ＣＢ１とし、その出力ベクトルを$\underline{s}_{0i}$、$\underline{s}_{1j}$、ただし０≦ｉ，ｊ≦３１、とする。また、ゲインコードブックＣＢｇの出力をg_l、ただし０≦ｌ≦３１、とする。g_lはスカラ値である。この最終出力$\underline{X}_0$’は、$g_i(\underline{s}_{0i}+\underline{s}_{1j})$となる。

【０１１５】

ＬＰＣ残差について上記ＭＢＥ分析によって得られたスペクトルエンベロープＡｍを一定次元に変換したものを$\underline{X}$とする。このとき、$\underline{X}$をいかに効率的に量子化するかが重要である。

【０１１６】

ここで、量子化誤差エネルギＥを、

$E=\|W\{H\underline{X}-Hg_l(\underline{s}_{0i}+\underline{s}_{1j})\}\|^2$　　・・・　（２１）

$=\|WH\{\underline{X}-g_l(\underline{s}_{0i}+\underline{s}_{1j})\}\|^2$

と定義する。この（２１）式において、ＨはＬＰＣの合成フィルタの周波数軸上での特性であり、Ｗは聴覚重み付けの周波数軸上での特性を表す重み付けのための行列である。

【０１１７】

現フレームのＬＰＣ分析結果によるαパラメータを、α_i（１≦ｉ≦Ｐ）として、

【０１１８】

【数１３】

$$H(z)=\frac{1}{1+\sum_{i=1}^{P}\alpha_i z^{-i}} \quad \cdots(22)$$

【０１１９】

の周波数特性からＬ次元、例えば４４次元の各対応する点の値をサンプルしたものである。

【０１２０】

算出手順としては、一例として、１、α_1、α_2、・・・、α_pに０詰めして、すなわち、１、α_1、α_2、・・・、α_p、０、０、・・・、０として、例えば２５６点のデータにする。その後、２５６点ＦＦＴを行い、$(r_e^2+I_m^2)^{1/2}$を０～πに対応する点に対して算出して、その逆数をとる。それをＬ点、すなわち例えば４４点に間引いたものを対角要素とする行列を、

【０１２１】

【数１４】

$$H=\begin{bmatrix} h(1) & & 0 \\ & h(2) & \\ & & \ddots \\ 0 & & h(L) \end{bmatrix}$$

【０１２２】

とする。

【０１２３】

聴覚重み付け行列Ｗは、

【０１２４】
【数１５】

$$W(z)=\frac{1+\sum_{i}^{P}\alpha_i\lambda_b^i z^{-i}}{1+\sum_{i}^{P}\alpha_i\lambda_a^i z^{-i}} \quad \cdots(23)$$

【０１２５】
とする。この（２３）式で、α_i は入力のＬＰＣ分析結果である。また、λa、λbは定数であり、一例として、λa＝０．４、λb＝０．９が挙げられる。
【０１２６】
行列あるいはマトリクスWは、上記（２３）式の周波数特性から算出できる。一例として、１、$\alpha_1\lambda b$、$\alpha_2\lambda b^2$、・・・、$\alpha_p\lambda b^p$、０、０、・・・、０として２５６点のデータとしてＦＦＴを行い、０以上π以下の区間に対して$(r_e^2[i]+I_m^2[i])^{1/2}$、０≦ｉ≦１２８、を求める。次に、１、$\alpha_1\lambda a$、$\alpha_2\lambda a^2$、・・・、$\alpha_p\lambda a^p$、０、０、・・・、０として分母の周波数特性を２５６点ＦＦＴで０～πの区間を１２８点で算出する。これを$(r_e'^2[i]+I_m'^2[i])^{1/2}$、０≦ｉ≦１２８、とする。
【０１２７】
【数１６】

$$w_0[i]=\frac{\sqrt{re^2[i]+\mathrm{Im}^2[i]}}{\sqrt{re'^2[i]+\mathrm{Im}'^2[i]}} \quad (0\le i\le 128)$$

【０１２８】
として、上記（２３）式の周波数特性が求められる。
【０１２９】
これをＬ次元、例えば４４次元ベクトルの対応する点について、以下の方法で求める。より正確には、直線補間を用いるべきであるが、以下の例では最も近い点の値で代用している。
【０１３０】
すなわち、
$\omega[i]=\omega_0[\mathrm{nint}(128i/L)]$　　$1\le i\le L$
ただし、nint（Ｘ）は、Ｘに最も近い整数を返す関数
である。
【０１３１】
また、上記Ｈに関しても同様の方法で、h(1)、h(2)、・・・、h(L)を求めている。すなわち、
【０１３２】
【数１７】

$$H=\begin{bmatrix} h(1) & & 0 \\ & h(2) & \\ & & \ddots \\ 0 & & h(L) \end{bmatrix} \quad W=\begin{bmatrix} w(1) & & 0 \\ & w(2) & \\ & & \ddots \\ 0 & & w(L) \end{bmatrix}$$

$$WH=\begin{bmatrix} h(1)w(1) & & 0 \\ & h(2)w(2) & \\ & & \ddots \\ 0 & & h(L)w(L) \end{bmatrix} \quad \cdots(24)$$

【０１３３】
となる。
【０１３４】
ここで、他の例として、ＦＦＴの回数を減らすのに、Ｈ(ｚ)Ｗ(ｚ)を先に求めてから、周波数特性を求めてもよい。すなわち、
【０１３５】
【数１８】

$$H(z)W(z)=\frac{1}{1+\sum_{i=1}^{P}\alpha_i z^{-i}}\cdot\frac{1+\sum_{i=1}^{P}\alpha_i\lambda_b^i z^{-i}}{1+\sum_{i=1}^{P}\alpha_i\lambda_a^i z^{-i}} \quad \cdots(25)$$

【０１３６】
この（２５）式の分母を展開した結果を、
【０１３７】
【数１９】

$$\left(1+\sum_{i=1}^{P}\alpha_i z^{-i}\right)\left(1+\sum_{i=1}^{P}\alpha_i\lambda_a^i z^{-i}\right)=1+\sum_{i=1}^{2P}\beta_i z^{-i}$$

【０１３８】
とする。ここで、１、β_1、β_2、・・・、β_{2p}、０、０、・・・、０として、例えば２５６点のデータにする。その後、２５６点ＦＦＴを行い、振幅の周波数特性を、
【０１３９】
【数２０】

$$rms[i]=\sqrt{re''^{2}[i]+\mathrm{Im}''^{2}[i]}\quad 0\le i\le 128$$

【０１４０】
とする。これより、
【０１４１】
【数２１】

$$wh_0[i]=\frac{\sqrt{re^{2}[i]+\mathrm{Im}^{2}[i]}}{\sqrt{re''^{2}[i]+\mathrm{Im}''^{2}[i]}}\quad 0\le i\le 128$$

【０１４２】
これをＬ次元ベクトルの対応する点について求める。上記ＦＦＴのポイント数が少ない場合は、直線補間で求めるべきであるが、ここでは最寄りの値を使用している。すなわち、
【０１４３】
【数２２】

$$wh[i]=wh_0\left[n\,\mathrm{int}\left(\frac{128}{L}\cdot i\right)\right]\quad 1\le i\le L$$

【０１４４】
である。これを対角要素とする行列をＷ’とすると、
【０１４５】
【数２３】

$$W'=\begin{bmatrix} wh(1) & & 0\\ & wh(2) & \\ & & \ddots\\ 0 & & wh(L)\end{bmatrix}\quad\cdots(26)$$

【０１４６】
となる。（２６）式は上記（２４）式と同一のマトリクスとなる。
【０１４７】
このマトリクス、すなわち重み付き合成フィルタの周波数特性を用いて、上記（２１）式を書き直すと、
【０１４８】
【数２４】

$$E = \left\| W'\left(\underline{x} - g_l\left(\underline{s}_{0i} + \underline{s}_{1j}\right)\right)\right\|^2 \quad \cdots(27)$$

【０１４９】
となる。
【０１５０】
ここで、シェイプコードブックとゲインコードブックの学習法について説明する。
【０１５１】
先ず、ＣＢ０に関しコードベクトル$\underline{s}_{0c}$を選択する全てのフレームｋに関して歪の期待値を最小化する。そのようなフレームがＭ個あるとして、
【０１５２】
【数２５】

$$J = \frac{1}{M}\sum_{k=1}^{M}\left\| W_k'\left(\underline{x} - g_k\left(\underline{s}_{0c} + \underline{s}_{1k}\right)\right)\right\|^2 \quad \cdots(28)$$

【０１５３】
を最小化すればよい。この（２８）式中で、W'_kはｋ番目のフレームに対する重み、$\underline{X}_k$はｋ番目のフレームの入力、g_kはｋ番目のフレームのゲイン、$\underline{s}_{1k}$はｋ番目のフレームについてのコードブックＣＢ１からの出力、をそれぞれ示す。
【０１５４】
この（２８）式を最小化するには、
【０１５５】
【数２６】

$$\begin{aligned} J &= \frac{1}{M}\sum_{k=1}^{M}\left\{\left(\underline{x}_k^T - g_k\left(\underline{s}_{0c}^T + \underline{s}_{1k}^T\right)\right)W_k'^T W_k'\left(\underline{x}_k - g_k\left(\underline{s}_{0c} + \underline{s}_{1k}\right)\right)\right\} \\ &= \frac{1}{M}\sum_{k=1}^{M}\left\{\underline{x}_k^T W_k'^T W_k'\underline{x}_k - 2g_k\left(\underline{s}_{0c}^T + \underline{s}_{1k}^T\right)W_k'^T W_k'\underline{x}_k \right. \\ &\qquad \left. + g_k^2\left(\underline{s}_{0c}^T + \underline{s}_{1k}^T\right)W_k'^T W_k'\left(\underline{s}_{0c} + \underline{s}_{1k}\right)\right\} \\ &= \frac{1}{M}\sum_{k=1}^{M}\left\{\underline{x}_k^T W_k'^T W_k'\underline{x}_k - 2g_k\left(\underline{s}_{0c}^T + \underline{s}_{1k}^T\right)W_k'^T W_k'\underline{x}_k \right. \\ &\qquad + g_k^2\underline{s}_{0c}^T W_k'^T W_k'\underline{s}_{0c} + 2g_k^2\underline{s}_{0c}^T W_k'^T W_k'\underline{s}_{1k} \\ &\qquad \left. + g_k^2\underline{s}_{1k}^T W_k'^T W_k'\underline{s}_{1k}\right\} \quad \cdots(29) \end{aligned}$$

【０１５６】
【数２７】

$$\frac{\partial J}{\partial \underline{s}_{0c}} = \frac{1}{M}\sum_{k=1}^{M}\{-2g_k W_k'^T W_k \underline{x}_k + 2g_k^2 W_k'^T W_k \underline{s}_{0c} + 2g_k^2 W_k'^T W_k \underline{s}_{1k}\} = 0 \quad \cdots(30)$$

より $\sum_{k=1}^{M}\left(g_k W_k'^T W_k' \underline{x}_k - g_k^2 W_k'^T W_k \underline{s}_{1k}\right) = \sum_{k=1}^{M} g_k^2 W_k'^T W_k \underline{s}_{0c}$

よって

$$s_{0c} = \left\{\sum_{k=1}^{M} g_k^2 W_k'^T W_k\right\}^{-1} \cdot \left\{\sum_{k=1}^{M} g_k W_k'^T W_k \left(\underline{x} - g_k \underline{s}_{1k}\right)\right\} \quad \cdots(31)$$

{ } $^{-1}$は逆行列を意味する。

$W_k'^T$は W_k' の転置行列である。

【０１５７】
次に、ゲインに関しての最適化を考える。
【０１５８】
ゲインのコードワードｇ$_c$を選択するｋ番目のフレームに関しての歪の期待値Ｊ$_g$は、
【０１５９】
【数２８】

$$J_g = \frac{1}{M}\sum_{k=1}^{M}\left\| W_k'\left(x_k - g_c\left(\underline{s}_{0k} + \underline{s}_{1k}\right)\right)\right\|^2$$

$$= \frac{1}{M}\sum_{k=1}^{M}\left\{ x_k^T W_k'^T W_k' x_k - 2 g_c x_k^T W_k'^T W_k'\left(\underline{s}_{0k} + \underline{s}_{1k}\right) - g_c^2\left(\underline{s}_{0k}^T + \underline{s}_{1k}^T\right) W_k'^T W_k'\left(\underline{s}_{0k} + \underline{s}_{1k}\right)\right\}$$

$$\frac{\partial J_g}{\partial g_c} = \frac{1}{M}\sum_{k=1}^{M}\left\{ -2 x_k^T W_k'^T W_k'\left(\underline{s}_{0k} + \underline{s}_{1k}\right) - 2 g_c\left(\underline{s}_{0k}^T + \underline{s}_{1k}^T\right) W_k'^T W_k'\left(\underline{s}_{0k} + \underline{s}_{1k}\right) = 0 \text{を解いて}\right.$$

$$\sum_{k=1}^{M} x_k^T W_k'^T W_k'\left(\underline{s}_{0k} + \underline{s}_{1k}\right) = \sum_{k=1}^{M} g_c\left(\underline{s}_{0k}^T + \underline{s}_{1k}^T\right) W_k'^T W'\left(\underline{s}_{0k} + \underline{s}_{1k}\right)$$

$$g_c = \frac{\sum_{k=1}^{M} x_k^T W_k'^T W_k'\left(\underline{s}_{0k} + \underline{s}_{1k}\right)}{\sum_{k=1}^{M}\left(\underline{s}_{0k}^T + \underline{s}_{1k}^T\right) W_k'^T W'\left(\underline{s}_{0k} + \underline{s}_{1k}\right)} \quad \cdots(32)$$

【０１６０】
上記（３１）式及び（３２）式は、シェイプ$\underline{s}_{0i}$、$\underline{s}_{1i}$及びゲインg_i、０≦ｉ≦３１の最適なセントロイドコンディション(Centroid Condition)、すなわち最適なデコーダ出力を与えるものである。なお、$\underline{s}_{1i}$に関しても$\underline{s}_{0i}$と同様に求めることができる。
【０１６１】
次に、最適エンコード条件（Nearest Neighbour Condition）を考える。
【０１６２】
歪尺度を求める上記（２７）式、すなわち、$E = \| W'(\underline{X} - g_l(\underline{s}_{0i} + \underline{s}_{1j})) \|^2$を最小化する$\underline{s}_{0i}$、$\underline{s}_{1j}$を、入力$\underline{X}$、重みマトリクスW'が与えられる毎に、すなわち毎フレームごとに決定する。
【０１６３】
本来は、総当り的に全てのg_l（０≦ｌ≦３１）、$\underline{s}_{0i}$（０≦ｉ≦３１）、$\underline{s}_{1j}$（０≦ｊ≦３１）の組み合せの、３２×３２×３２＝３２７６８通りについてＥを求めて、最小のＥを与えるg_l、$\underline{s}_{0i}$、$\underline{s}_{1j}$の組を求めるべきであるが、膨大な演算量となるので、本実施の形態では、シェイプとゲインのシーケンシャルサーチを行っている。なお、$\underline{s}_{0i}$と$\underline{s}_{1j}$との組み合せについては、総当りサーチを行うものとする。これは、３２×３２＝１０２４通りである。以下の説明では、簡単化のため、$\underline{s}_{0i} + \underline{s}_{1j}$を$\underline{s}_m$と記す。
【０１６４】
上記（２７）式は、$E = \| W'(\underline{X} - g_l \underline{s}_m) \|^2$となる。さらに簡単のため、$\underline{X}_w = W'\underline{X}$、$\underline{s}_w = W'\underline{s}_m$とすると、
【０１６５】
【数２９】

$$E = \left\| \underline{x}_w - g_l \underline{s}_w \right\|^2 \quad \cdots(33)$$

$$E = \left\| \underline{x}_w \right\|^2 + \left\| \underline{s}_w \right\|^2 \left(g_l - \frac{\underline{x}_w^T \cdot \underline{s}_w}{\left\| \underline{s}_w \right\|^2} \right)^2 - \frac{\left(\underline{x}_w^T \cdot \underline{s}_w \right)^2}{\left\| \underline{s}_w \right\|^2} \quad \cdots(34)$$

【０１６６】
となる。従って、g_1 の精度が充分にとれると仮定すると、
【０１６７】
【数３０】

$$(1) \quad \frac{\left(\underline{x}_w^T \cdot \underline{s}_w \right)^2}{\left\| \underline{s}_w \right\|^2} \text{ の最大となる} \underline{s}_w \text{をサーチ}$$

$$(2) \quad \frac{\underline{x}_w^T \cdot \underline{s}_w}{\left\| \underline{s}_w \right\|^2} \quad \text{にもっとも近い} g_l \text{をサーチ}$$

【０１６８】
という２つのステップに分けてサーチすることができる。元の表記を用いて書き直すと、
【０１６９】
【数３１】

$$(1)' \quad \frac{\left(\underline{x}^T W'^T W' \left(\underline{s}_{0i} + \underline{s}_{1j} \right) \right)^2}{\left\| W' \left(\underline{s}_{0i} + \underline{s}_{1j} \right) \right\|^2} \text{を最大化する}$$

$$\underline{s}_{0i} \ \underline{s}_{1j} \text{ の組をサーチ} \cdots(35)$$

$$(2)' \quad \frac{\underline{x}^T W'^T W' \left(\underline{s}_{0i} + \underline{s}_{1j} \right)}{\left\| W' \left(\underline{s}_{0i} + \underline{s}_{1j} \right) \right\|^2}$$

$$\text{にもっとも近い} g_l \text{をサーチ}$$

【０１７０】
となる。この（３５）式が最適エンコード条件(Nearest Neighbour Condition) である。
【０１７１】
ここで上記（３１）、（３２）式の条件（Centroid Condition）と、（３５）式の条件を用いて、ＬＢＧ(Linde-Buzo-Gray)アルゴリズム、いわゆる一般化ロイドアルゴリズム（Generalized Lloyd Algorithm:ＧＬＡ）によりコードブック（ＣＢ０、ＣＢ１、ＣＢｇ）を同時にトレーニングできる。
【０１７２】
ところで、ベクトル量子化器１１６でのベクトル量子化の際の聴覚重み付けに用いられる

重みW’については、上記（２６）式で定義されているが、過去のW’も加味して現在のW’を求めることにより、テンポラルマスキングも考慮したW’が求められる。

【０１７３】

上記（２６）式中のwh(1),wh(2),・・・,wh(L)に関して、時刻ｎ、すなわち第ｎフレームで算出されたものをそれぞれ$wh_n(1),wh_n(2),\cdots,wh_n(L)$とする。

【０１７４】

時刻ｎで過去の値を考慮した重みを$A_n(i)$、１≦ｉ≦Ｌ　と定義すると、

$$A_n(i)=\lambda A_{n-1}(i)+(1-\lambda)\,wh_n(i) \qquad (wh_n(i)\leqq A_{n-1}(i))$$
$$=wh_n(i) \qquad (wh_n(i)>A_{n-1}(i))$$

とする。ここで、λは例えばλ＝０．２とすればよい。このようにして求められた$A_n(i)$、１≦ｉ≦Ｌ　について、これを対角要素とするマトリクスを上記重みとして用いればよい。

【０１７５】

このように重み付きベクトル量子化により得られたシェイプインデクス$\underline{s}_{0i}$、$\underline{s}_{1j}$は、出力端子５２０、５２２からそれぞれ出力され、ゲインインデクスg_1は、出力端子５２１から出力される。また、量子化値$\underline{X}_0$’は、出力端子５０４から出力されると共に、加算器５０５に送られる。

【０１７６】

この加算器５０５では、出力ベクトル$\underline{X}$から量子化値$\underline{X}_0$’が減算されて、量子化誤差ベクトル$\underline{Y}$が生成される。この量子化誤差ベクトル$\underline{Y}$は、具体的には、８個のベクトル量子化器５１１_1～５１１_8から成るベクトル量子化部５１１に送られて、次元分割され、各ベクトル量子化器５１１_1～５１１_8で重み付きのベクトル量子化が施される。

【０１７７】

第２のベクトル量子化部５１０では、第１のベクトル量子化部５００と比較して、かなり多くのビット数を用いるため、コードブックのメモリ容量及びコードブックサーチのための演算量（Complexity）が非常に大きくなり、第１のベクトル量子化部５００と同じ４４次元のままでベクトル量子化を行うことは、不可能である。そこで、第２のベクトル量子化部５１０内のベクトル量子化部５１１を複数個のベクトル量子化器で構成し、入力される量子化値を次元分割して、複数個の低次元ベクトルとして、重み付きのベクトル量子化を行う。

【０１７８】

ベクトル量子化器５１１_1～５１１_8で用いる各量子化値$\underline{Y}_0$～$\underline{Y}_7$と、次元数と、ビット数との関係を、表２に示す。

【０１７９】

【表２】

量子化値	次元	ビット数(bits)
$\underline{Y}_0$	4	１０
$\underline{Y}_1$	4	１０
$\underline{Y}_2$	4	１０
$\underline{Y}_3$	4	１０
$\underline{Y}_4$	4	9
$\underline{Y}_5$	8	8
$\underline{Y}_6$	8	8
$\underline{Y}_7$	8	7

【０１８０】
ベクトル量子化器５１１$_1$～５１１$_8$から出力されるインデクスＩｄｖｑ$_0$～Ｉｄｖｑ$_7$は、各出力端子５２３$_1$～５２３$_8$からそれぞれ出力される。これらのインデクスの合計は７２ビットである。
【０１８１】
また、ベクトル量子化器５１１$_1$～５１１$_8$から出力される量子化値$\underline{Y}_0$’～$\underline{Y}_7$’を次元方向に接続した値を$\underline{Y}$’とすると、加算器５１３では、量子化値$\underline{Y}$’と量子化値$\underline{X}_0$’とが加算されて、量子化値$\underline{X}_1$’が得られる。よって、この量子化値$\underline{X}_1$’は、
$\underline{X}_1$’＝$\underline{X}_0$’＋$\underline{Y}$’
＝$\underline{X}$－$\underline{Y}$＋$\underline{Y}$’
で表される。すなわち、最終的な量子化誤差ベクトルは、$\underline{Y}$’－$\underline{Y}$となる。
【０１８２】
尚、音声信号復号化装置側では、この第２のベクトル量子化部５１０からの量子化値$\underline{X}_1$’を復号化するときには、第１のベクトル量子化部５００からの量子化値$\underline{X}_0$’は不要であるが、第１のベクトル量子化部５００及び第２のベクトル量子化部５１０からのインデクスは必要とする。
【０１８３】
次に、上記ベクトル量子化部５１１における学習法及びコードブックサーチについて説明する。
【０１８４】
先ず、学習法においては、量子化誤差ベクトル$\underline{Y}$及び重みＷ’を用い、表２に示すように、８つの低次元ベクトル$\underline{Y}_0$～$\underline{Y}_7$及びマトリクスに分割する。このとき、重みＷ’は、例えば４４点に間引いたものを対角要素とする行列、
【０１８５】
【数３２】

$$W' = \begin{bmatrix} wh(1) & & & 0 \\ & wh(2) & & \\ & & \ddots & \\ 0 & & & wh(44) \end{bmatrix} \quad \cdots(36)$$

【０１８６】
とすると、以下の８つの行列に分割される。
【０１８７】
【数３３】

$$W_1' = \begin{bmatrix} wh(1) & & 0 \\ & \ddots & \\ 0 & & wh(4) \end{bmatrix}$$

$$W_2' = \begin{bmatrix} wh(5) & & 0 \\ & \ddots & \\ 0 & & wh(8) \end{bmatrix}$$

$$W_3' = \begin{bmatrix} wh(9) & & 0 \\ & \ddots & \\ 0 & & wh(12) \end{bmatrix}$$

$$W_4' = \begin{bmatrix} wh(13) & & 0 \\ & \ddots & \\ 0 & & wh(16) \end{bmatrix}$$

$$W_5' = \begin{bmatrix} wh(17) & & 0 \\ & \ddots & \\ 0 & & wh(20) \end{bmatrix}$$

$$W_6' = \begin{bmatrix} wh(21) & & 0 \\ & \ddots & \\ 0 & & wh(28) \end{bmatrix}$$

$$W_7' = \begin{bmatrix} wh(29) & & 0 \\ & \ddots & \\ 0 & & wh(36) \end{bmatrix}$$

$$W_8' = \begin{bmatrix} wh(37) & & 0 \\ & \ddots & \\ 0 & & wh(44) \end{bmatrix}$$

【０１８８】
このように、Ｙ及びＷ’の低次元に分割されたものを、それぞれ
$\underline{Y}_i$、W_i'　（１≦ｉ≦８）
とする。
【０１８９】
ここで、歪尺度Ｅを、
$E = \| W_i' (\underline{Y}_i - \underline{s}) \|^2$　・・・（３７）
と定義する。このコードベクトル$\underline{s}$は$\underline{Y}_i$の量子化結果であり、歪尺度Ｅを最小化する、コードブックのコードベクトル$\underline{s}$がサーチされる。
【０１９０】
尚、W_i’は、学習時には重み付けがあり、サーチ時には重み付け無し、すなわち単位行列とし、学習時とコードブックサーチ時とでは異なる値を用いるようにしてもよい。

【０１９１】
また、コードブックの学習では、一般化ロイドアルゴリズム（ＧＬＡ）を用い、さらに重み付けを行っている。先ず、学習のための最適なセントロイドコンディションについて説明する。コードベクトル$\underline{s}$を最適な量子化結果として選択した入力ベクトル$\underline{Y}$がＭ個ある場合に、トレーニングデータを$\underline{Y}_k$とすると、歪の期待値Ｊは、全てのフレームｋに関して重み付け時の歪の中心を最小化するような（３８）式となる。
【０１９２】
【数３４】

$$J = \frac{1}{M}\sum_{k=1}^{M}\left\|W_k'(\underline{Y}_k - \underline{s})\right\|^2$$

$$= \frac{1}{M}\sum_{k=1}^{M}(\underline{Y}_k - \underline{s})^T W_k'^T W_k'(\underline{Y}_k - \underline{s})$$

$$= \frac{1}{M}\sum_{k=1}^{M}\underline{Y}_k^T W_k'^T W_k'\underline{Y}_k - 2\underline{Y}_k^T W_k'^T W_k'\underline{s} + \underline{s}^T W_k'^T W_k'\underline{s} \quad \cdots(38)$$

$$\frac{\partial J}{\partial \underline{s}} = \frac{1}{M}\sum_{k=1}^{M}\left(-2\underline{Y}_k^T W_k'^T W_k' + 2\underline{s}^T W_k'^T W_k'\right) = 0$$

を解いて

$$\sum_{k=1}^{M}\underline{Y}_k^T W_k'^T W_k' = \sum_{k=1}^{M}\underline{s}^T W_k'^T W_k'$$

両辺の転値Ｔをとって

$$\sum_{k=1}^{M}W_k'^T W_k'\underline{Y}_k = \sum_{k=1}^{M}W_k'^T W_k'\underline{s}$$

よって

$$\underline{s} = \left(\sum_{k=1}^{M}W_k'^T W_k'\right)^{-1}\sum_{k=1}^{M}W_k'^T W_k'\underline{Y}_k \quad \cdots(39)$$

【０１９３】
上記（３９）式で示す$\underline{s}$は最適な代表ベクトルであり、最適なセントロイドコンディションである。
【０１９４】
また、最適エンコード条件は、$\|W_i'(\underline{Y}_i - \underline{s})\|^2$ の値を最小化する$\underline{s}$をサーチすればよい。ここで、サーチ時のW_i'は、必ずしも学習時と同じW_i'である必要はなく、重み無しで
【０１９５】
【数３５】

$$\begin{bmatrix} 1 & & 0 \\ & 1 & \\ & \ddots & \\ 0 & & 1 \end{bmatrix}$$

【０１９６】
のマトリクスとしてもよい。
【０１９７】
このように、音声信号符号化装置内のベクトル量子化部１１６を２段のベクトル量子化部から構成することにより、出力するインデクスのビット数を可変にすることができる。
【０１９８】
次に、本発明の前記ＣＥＬＰ符号化構成を用いた第２の符号化部１２０は、より具体的には図９に示すような、多段のベクトル量子化処理部（図９の例では２段の符号化部１２０_1と１２０_2）の構成を有するものとなされている。なお、当該図９の構成は、伝送ビットレートを例えば前記２ｋｂｐｓと６ｋｂｐｓとで切り換え可能な場合において、６ｋｂｐｓの伝送ビットレートに対応した構成を示しており、さらにシェイプ及びゲインインデクス出力を２３ビット／５ｍｓｅｃと１５ビット／５ｍｓｅｃとで切り換えられるようにしているものである。また、この図９の構成における処理の流れは図１０に示すようになっている。
【０１９９】
この図９において、例えば、図９のＬＰＣ分析回路３０２は前記図３に示したＬＰＣ分析回路１３２と対応し、図９のＬＳＰパラメータ量子化回路３０３は図３の前記α→ＬＳＰ変換回路１３３からＬＳＰ→α変換回路１３７までの構成と対応し、図９の聴覚重み付けフィルタ３０４は図３の前記聴覚重み付けフィルタ算出回路１３９及び聴覚重み付けフィルタ１２５と対応している。したがって、この図９において、端子３０５には前記図３の第１の符号化部１１３のＬＳＰ→α変換回路１３７からの出力と同じものが供給され、また、端子３０７には前記図３の聴覚重み付けフィルタ算出回路１３９からの出力と同じものが、端子３０６には前記図３の聴覚重み付けフィルタ１２５からの出力と同じものが供給される。ただし、この図９の聴覚重み付けフィルタ３０４では、前記図３の聴覚重み付けフィルタ１２５とは異なり、前記ＬＳＰ→α変換回路１３７の出力を用いずに、入力音声データと量子化前のαパラメータとから、前記聴覚重み付けした信号（すなわち前記図３の聴覚重み付けフィルタ１２５からの出力と同じ信号）を生成している。
【０２００】
また、この図９に示す２段構成の第２の符号化部１２０_1及び１２０_2において、減算器３１３及び３２３は図３の減算器１２３と対応し、距離計算回路３１４及び３２４は図３の距離計算回路１２４と、ゲイン回路３１１及び３２１は図３のゲイン回路１２６と、ストキャスティックコードブック３１０，３２０及びゲインコードブック３１５，３２５は図３の雑音符号帳１２１とそれぞれ対応している。
【０２０１】
このような図９の構成において、先ず、図１０のステップＳ１に示すように、ＬＰＣ分析回路３０２では、端子３０１から供給された入力音声データ$\underline{x}$を前述同様に適当なフレームに分割してＬＰＣ分析を行い、αパラメータを求める。ＬＳＰパラメータ量子化回路３０３では、上記ＬＰＣ分析回路３０２からのαパラメータをＬＳＰパラメータに変換して量子化し、さらにこの量子化したＬＳＰパラメータを補間した後、αパラメータに変換する。次に、当該ＬＳＰパラメータ量子化回路３０３では、当該量子化したＬＳＰパラメー

タを変換したαパラメータ、すなわち量子化されたαパラメータから、ＬＰＣ合成フィルタ関数１／Ｈ（ｚ）を生成し、これを端子３０５を介して１段目の第２の符号化部１２０$_1$の聴覚重み付き合成フィルタ３１２に送る。

【０２０２】

一方、聴覚重み付けフィルタ３０４では、ＬＰＣ分析回路３０２からのαパラメータ（すなわち量子化前のαパラメータ）から、前記図３の聴覚重み付けフィルタ算出回路１３９によるものと同じ聴覚重み付けのためのデータを求め、この重み付けのためのデータが端子３０７を介して、１段目の第２の符号化部１２０$_1$の聴覚重み付き合成フィルタ３１２に送られる。また、当該聴覚重み付けフィルタ３０４では、図１０のステップＳ２に示すように、入力音声データと量子化前のαパラメータとから、前記聴覚重み付けした信号（前記図３の聴覚重み付けフィルタ１２５からの出力と同じ信号）を生成する。すなわち、先ず、量子化前のαパラメータから聴覚重み付けフィルタ関数Ｗ（ｚ）を生成し、さらに入力音声データ$\underline{x}$に当該フィルタ関数Ｗ（ｚ）をかけて$\underline{x}_W$を生成し、これを上記聴覚重み付けした信号として、端子３０６を介して１段目の第２の符号化部１２０$_1$の減算器３１３に送る。

【０２０３】

１段目の第２の符号化部１２０$_1$では、９ビットシェイプインデクス出力のストキャスティックコードブック（stochastic code book）３１０からの代表値出力（無声音のＬＰＣ残差に相当するノイズ出力）がゲイン回路３１１に送られ、このゲイン回路３１１にて、ストキャスティックコードブック３１０からの代表値出力に６ビットゲインインデクス出力のゲインコードブック３１５からのゲイン（スカラ値）を乗じ、このゲイン回路３１１にてゲインが乗じられた代表値出力が、１／Ａ（ｚ）＝（１／Ｈ（ｚ））・Ｗ（ｚ）の聴覚重み付きの合成フィルタ３１２に送られる。この重み付きの合成フィルタ３１２からは、図１０のステップＳ３のように、１／Ａ（ｚ）のゼロ入力応答出力が減算器３１３に送られる。当該減算器３１３では、上記聴覚重み付き合成フィルタ３１２からのゼロ入力応答出力と、上記聴覚重み付けフィルタ３０４からの上記聴覚重み付けした信号$\underline{x}_W$とを用いた減算が行われ、この差分或いは誤差が参照ベクトル$\underline{r}$として取り出される。図１０のステップＳ４に示すように、１段目の第２の符号化部１２０$_1$でのサーチ時には、この参照ベクトル$\underline{r}$が、距離計算回路３１４に送られ、ここで距離計算が行われ、量子化誤差エネルギＥを最小にするシェイプベクトル$\underline{s}$とゲインｇがサーチされる。ただし、ここでの１／Ａ（ｚ）はゼロ状態である。すなわち、コードブック中のシェイプベクトル$\underline{s}$をゼロ状態の１／Ａ（ｚ）で合成したものを$\underline{s}_{syn}$とするとき、式（４０）を最小にするシェイプベクトル$\underline{s}$とゲインｇをサーチする。

【０２０４】

【数３６】

$$E = \sum_{n=0}^{N-1} \left(\underline{r}(n) - g\underline{s}_{syn}(n) \right)^2 \quad \cdots(40)$$

【０２０５】

ここで、量子化誤差エネルギＥを最小とする$\underline{s}$とｇをフルサーチしてもよいが、計算量を減らすために、以下のような方法をとることができる。

【０２０６】

第１の方法として、以下の式（４１）に定義するＥ$_s$を最小とするシェイプベクトル$\underline{s}$をサーチする。

【０２０７】

【数３７】

$$E_s = \frac{\sum_{n=0}^{N-1} \underline{r}(n)\underline{s}_{syn}(n)}{\sqrt{\sum_{n=0}^{N-1} \underline{s}_{syn}(n)^2}} \quad \cdots(41)$$

【０２０８】
第２の方法として、第１の方法により得られた$\underline{s}$より、理想的なゲインは、式（４２）のようになるから、式（４３）を最小とするｇをサーチする。
【０２０９】
【数３８】

$$g_{ref} = \frac{\sum_{n=0}^{N-1} \underline{r}(n)\underline{s}_{syn}(n)}{\sum_{n=0}^{N-1} \underline{s}_{syn}(n)^2} \quad \cdots(42)$$

【０２１０】
$E_g = (g_{ref} - g)^2$　　（４３）
ここで、Ｅはｇの二次関数であるから、E_gを最小にするｇはＥを最小化する。
【０２１１】
上記第１，第２の方法によって得られた$\underline{s}$とｇより、量子化誤差ベクトル$\underline{e}$（ｎ）は次の式（４４）のように計算できる。
【０２１２】
$\underline{e}(n) = \underline{r}(n) - g\underline{s}_{syn}(n)$　　（４４）
これを、２段目の第２の符号化部１２０$_2$のリファレンス入力として１段目と同様にして量子化する。
【０２１３】
すなわち、上記１段目の第２の符号化部１２０$_1$の聴覚重み付き合成フィルタ３１２からは、端子３０５及び端子３０７に供給された信号がそのまま２段目の第２の符号化部１２０$_2$の聴覚重み付き合成フィルタ３２２に送られる。また、当該２段目の第２の符号化部１２０$_2$減算器３２３には、１段目の第２の符号化部１２０$_1$にて求めた上記量子化誤差ベクトル$\underline{e}$（ｎ）が供給される。
【０２１４】
次に、図１０のステップＳ５において、当該２段目の第２の符号化部１２０$_2$でも１段目と同様に処理が行われる。すなわち、５ビットシェイプインデクス出力のストキャスティックコードブック３２０からの代表値出力がゲイン回路３２１に送られ、このゲイン回路３２１にて、当該コードブック３２０からの代表値出力に３ビットゲインインデクス出力のゲインコードブック３２５からのゲインを乗じ、このゲイン回路３２１の出力が、聴覚重み付きの合成フィルタ３２２に送られる。当該重み付きの合成フィルタ３２２からの出力は減算器３２３に送られ、当該減算器３２３にて上記聴覚重み付き合成フィルタ３２２からの出力と１段目の量子化誤差ベクトル$\underline{e}$（ｎ）との差分が求められ、この差分が距離計算回路３２４に送られてここで距離計算が行われ、量子化誤差エネルギＥを最小にするシェイプベクトル$\underline{s}$とゲインｇがサーチされる。

【０２１５】
上述したような１段目の第２の符号化部１２０$_1$のストキャストコードブック３１０からのシェイプインデクス出力及びゲインコードブック３１５からのゲインインデクス出力と、２段目の第２の符号化部１２０$_2$のストキャストコードブック３２０からのインデクス出力及びゲインコードブック３２５からのインデクス出力は、インデクス出力切り換え回路３３０に送られるようになっている。ここで、当該第２の符号化部１２０から２３ビット出力を行うときには、上記１段目と２段目の第２の符号化部１２０$_1$及び１２０$_2$のストキャストコードブック３１０，３２０及びゲインコードブック３１５，３２５からの各インデクスを合わせて出力し、一方、１５ビット出力を行うときには、上記１段目の第２の符号化部１２０$_1$のストキャストコードブック３１０とゲインコードブック３１５からの各インデクスを出力する。

【０２１６】
その後は、ステップＳ６のようにフィルタ状態がアップデートされる。

【０２１７】
ところで、本実施の形態では、２段目の第２の符号化部１２０$_2$のインデクスビット数が、シェイプベクトルについては５ビットで、ゲインについては３ビットと非常に少ない。このような場合、適切なシェイプ、ゲインがコードブックに存在しないと、量子化誤差を減らすどころか逆に増やしてしまう可能性がある。

【０２１８】
この問題を防ぐためには、ゲインに０を用意しておけばよいが、ゲインは３ビットしかなく、そのうちの一つを０にしてしまうのは量子化器の性能を大きく低下させてしまう。そこで、比較的多いビット数を割り当てたシェイプベクトルに、要素が全て０のベクトルを用意する。そして、このゼロベクトルを除いて、前述のサーチを行い、量子化誤差が最終的に増えてしまった場合に、ゼロベクトルを選択するようにする。なお、このときのゲインは任意である。これにより、２段目の第２の符号化部１２０$_2$が量子化誤差を増すことを防ぐことができる。

【０２１９】
なお、図９の例では、２段構成の場合を例に挙げているが、２段に限らず複数段構成とすることができる。この場合、１段目のクローズドループサーチによるベクトル量子化が終了したら、Ｎ段目（２≦Ｎ）ではＮ－１段目の量子化誤差をリファレンス入力として量子化を行い、さらにその量子化誤差をＮ＋１段目のリファレンス入力とする。

【０２２０】
上述したように、図９及び図１０から、第２の符号化部に多段のベクトル量子化器を用いることにより、従来のような同じビット数のストレートベクトル量子化や共役コードブックなどを用いたものと比較して、計算量が少なくなる。特に、ＣＥＬＰ符号化では、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ため、サーチの回数が少ないことが重要である。また、２段の第２の符号化部１２０$_1$と１２０$_2$の両インデクス出力を用いる場合と、１段目の第２の符号化部１２０$_1$のインデクス出力のみを用いる（２段目の第２の符号化部１２０$_2$の出力インデクスを用いない）場合とを切り換えることにより、簡単にビット数を切り換えることが可能となっている。さらに上述したように、１段目と２段目の第２の符号化部１２０$_1$と１２０$_2$の両インデクス出力を合わせて出力するようなことを行えば、後のデコーダ側において例えば何れかを選ぶようにすることで、デコーダ側でも容易に対応できることになる。すなわち例えば６ｋｂｐｓでエンコードしたパラメータを、２ｋｂｐｓのデコーダでデコードするときに、デコーダ側で容易に対応できることになる。またさらに、例えば２段目の第２の符号化部１２０$_2$のシェイプコードブックにゼロベクトルを含ませることにより、割り当てられたビット数が少ない場合でも、ゲインに０を加えるよりは少ない性能劣化で量子化誤差が増加することを防ぐことが可能となっている。

【０２２１】
次に、上記ストキャスティックコードブックのコードベクトル（シェイプベクトル）は例

えば以下のようにして生成することができる。

【０２２２】

例えば、ストキャスティックコードブックのコードベクトルは、いわゆるガウシアンノイズのクリッピングにより生成することができる。具体的には、ガウシアンノイズを発生させ、これを適当なスレシホールド値でクリッピングし、それを正規化することで、コードブックを構成することができる。

【０２２３】

ところが、音声には様々な形態があり、例えば「さ，し，す，せ，そ」のようなノイズに近い子音の音声には、ガウシアンノイズが適しているが、例えば「ぱ，ぴ，ぷ，ぺ，ぽ」のような立ち上がりの激しい子音（急峻な子音）の音声については、対応しきれない。

【０２２４】

そこで、本発明では、全コードベクトルのうち、適当な数はガウシアンノイズとし、残りを学習により求めて上記立ち上がりの激しい子音とノイズに近い子音の何れにも対応できるようにする。例えば、スレシホールド値を大きくとると、大きなピークを幾つか持つようなベクトルが得られ、一方、スレシホールド値を小さくとると、ガウシアンノイズそのものに近くなる。したがって、このようにクリッピングスレシホールド値のバリエーションを増やすことにより、例えば「ぱ，ぴ，ぷ，ぺ，ぽ」のような立ち上がりの激しい子音や、例えば「さ，し，す，せ，そ」のようなノイズに近い子音などに対応でき、明瞭度を向上させることができるようになる。なお、図１１には、図中実線で示すガウシアンノイズと図中点線で示すクリッピング後のノイズの様子を示している。また、図１１の（Ａ）はクリッピングスレシホールド値が１．０の場合（すなわちスレシホールド値が大きい場合）を、図１１の（Ｂ）にはクリッピングスレシホールド値が０．４の場合（すなわちスレシホールド値が小さい場合）を示している。この図１１の（Ａ）及び（Ｂ）から、スレシホールド値を大きくとると、大きなピークを幾つか持つようなベクトルが得られ、一方、スレシホールド値を小さくとると、ガウシアンノイズそのものに近くなることが判る。

【０２２５】

このようなことを実現するため、先ず、ガウシアンノイズのクリッピングにより初期コードブックを構成し、さらに予め適当な数だけ学習を行わないコードベクトルを決めておく。この学習しないコードベクトルは、その分散値が小さいものから順に選ぶようにする。これは、例えば「さ，し，す，せ，そ」のようなノイズに近い子音に対応させるためである。一方、学習を行って求めるコードベクトルは、当該学習のアルゴリズムとしてＬＢＧアルゴリズムを用いるようにする。ここで最適エンコード条件（Nearest Neighbour Condition）でのエンコードは固定したコードベクトルと、学習対象のコードベクトル両方を使用して行う。セントロイドコンディション（Centroid Condition）においては、学習対象のコードベクトルのみをアップデートする。これにより、学習対象となったコードベクトルは「ぱ，ぴ，ぷ，ぺ，ぽ」などの立ち上がりの激しい子音に対応するようになる。

【０２２６】

なお、ゲインは通常通りの学習を行うことで、これらのコードベクトルに対して最適なものが学習できる。

【０２２７】

上述したガウシアンノイズのクリッピングによるコードブックの構成のための処理の流れを図１２に示す。

【０２２８】

この図１２において、ステップＳ１０では、初期化として、学習回数ｎ＝０とし、誤差$D_0=\infty$とし、最大学習回数n_{max}を決定し、学習終了条件を決めるスレシホールド値εを決定する。

【０２２９】

次のステップＳ１１では、ガウシアンノイズのクリッピングによる初期コードブックを生成し、ステップＳ１２では学習を行わないコードベクトルとして一部のコードベクトルを固定する。

【０２３０】
次にステップＳ１３では上記コードブックを用いてエンコードを行い、ステップＳ１４では誤差を算出し、ステップＳ１５では（$D_{n-1}-D_n$）／$D_n<\varepsilon$、若しくは$n=n_{max}$か否かを判断し、Ｙｅｓと判断した場合には処理を終了し、Ｎｏと判断した場合にはステップＳ１６に進む。
【０２３１】
ステップＳ１６ではエンコードに使用されなかったコードベクトルの処理を行い、次のステップＳ１７ではコードブックのアップデートを行う。次にステップＳ１８では学習回数ｎを１インクリメントし、その後ステップＳ１３に戻る。
【０２３２】
以上説明したような信号符号化装置及び信号復号化装置は、例えば図１３及び図１４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。
【０２３３】
すなわち、図１３は、上記図１、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１３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１、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１、図３と共に説明したような符号化処理が行われ、図１、図２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。
【０２３４】
また、図１４は、上記図２、図４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１４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２、図４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２、図４と共に説明したような復号化処理が施され、図２、図４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。
【０２３５】
【発明の効果】
以上の説明からも明らかなように、本発明に係る音声符号化方法及び音声符号化装置は、入力音声信号の有声音部分に対して短期予測残差を求め、上記求められた短期予測残差に基づいてサイン波分析符号化パラメータを求め、上記サイン波分析符号化パラメータに対して聴覚重み付きのベクトル量子化を施し、上記入力音声信号の無声音部分に対して波形符号化により符号化するものであり、上記聴覚重み付きのベクトル量子化では、第１のベクトル量子化を行い、この第１のベクトル量子化の際の量子化誤差ベクトルを量子化する第２のベクトル量子化を行うことにより、データ伝送路の容量に応じて、出力する符号化データのビット数の切り換えを簡易に行うことができ、複数のデータビットレートに対応させることができる。また、符号化側と復号化側とでビットレートが異なっている場合であっても復号化側で容易に対応可能な符号化データ列を生成することができる。　また、本発明に係る音声復号化方法及び音声復号化装置は、入力音声信号の有声音部分に対して短期予測残差を求め、上記求められた短期予測残差に基づいてサイン波分析符号化パラメータを求め、上記サイン波分析符号化パラメータに対して聴覚重み付きのベクトル量子化

を施し、上記入力音声信号の無声音部分に対して波形符号化により符号化することにより生成された符号化データを受け取って復号化するものであり、上記ベクトル量子化されたサイン波分析符号化パラメータを逆ベクトル量子化して、サイン波合成により有声音部分の短期予測残差を求め、上記有声音部分の短期予測残差に基づいて有声音部分の時間軸波形を合成し、また、上記波形符号化された無声音部分について復号化を行い、無声音部分の短期予測残差を求め、上記求められた無声音部分の短期予測残差に基づいて無声音部分の時間軸波形を合成することにより、上述した符号化処理によって生成された符号化データ列を復号化することができる。

【図面の簡単な説明】

【図１】本発明に係る音声符号化方法及び音声符号化装置の実施の形態が適用される音声信号符号化装置の基本構成を示すブロック図である。

【図２】本発明に係る音声復号化方法及び音声復号化装置の実施の形態が適用される、図１に対応する音声信号復号化装置の基本構成を示すブロック図である。

【図３】図１の音声信号符号化装置のより具体的な構成を示すブロック図である。

【図４】図２の音声信号復号化装置のより具体的な構成を示すブロック図である。

【図５】ＬＳＰ量子化部の基本構成を示すブロック図である。

【図６】ＬＳＰ量子化部のより具体的な構成を示すブロック図である。

【図７】ベクトル量子化部の基本構成を示すブロック図である。

【図８】ベクトル量子化部のより具体的な構成を示すブロック図である。

【図９】本発明の音声信号符号化装置のＣＥＬＰ符号化部分（第２の符号化部）の具体的構成を示すブロック回路図である。

【図１０】図９の構成における処理の流れを示すフローチャートである。

【図１１】ガウシアンノイズと、異なるスレシホールド値でのクリッピング後のノイズの様子を示す図である。

【図１２】学習によってシェイプコードブックを生成する際の処理の流れを示すフローチャートである。

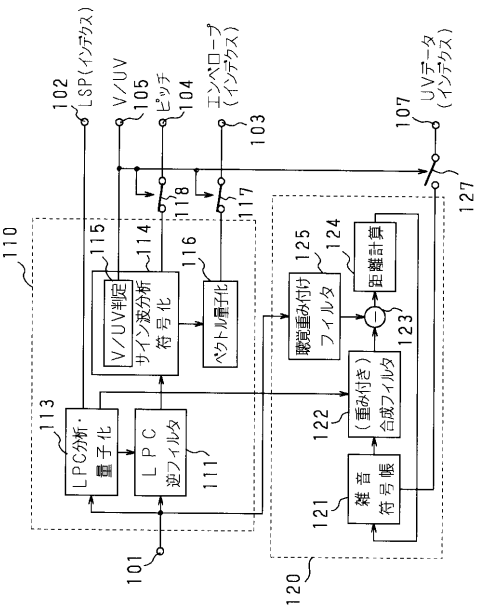
【図１３】本発明の実施の形態となる音声信号符号化装置が用いられる携帯端末の送信側構成を示すブロック図である。

【図１４】本発明の実施の形態となる、図１３に対応する音声信号復号化装置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。

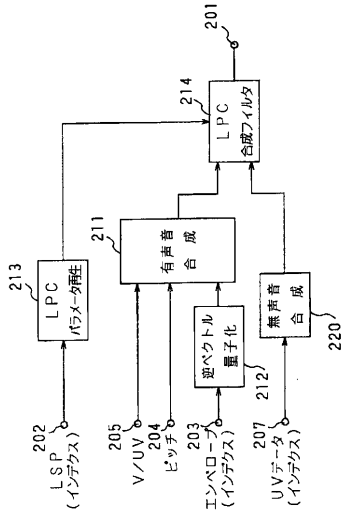
【符号の説明】

５００　第１のベクトル量子化部
５１０　第２のベクトル量子化部
５０２　ベクトル量子化器
５１１　ベクトル量子化部
５０５、５１３　加算器

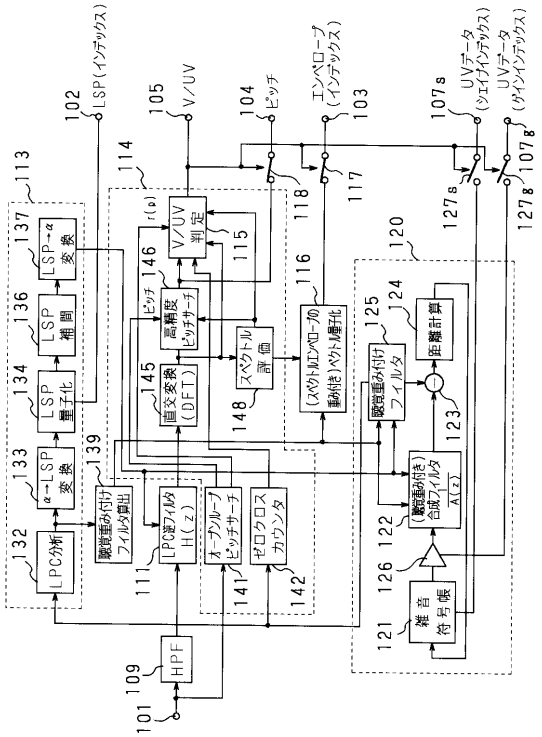
【図１】


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分析・量子化
LPC逆フィルタ
V／UV判定 サイン波分析符号化
ベクトル量子化
聴覚重み付けフィルタ
距離計算
（重み付き）合成フィルタ
雑音符号帳

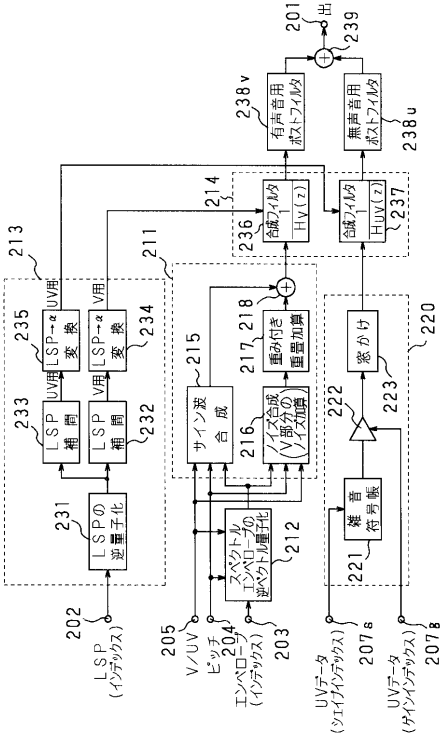
【図２】


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

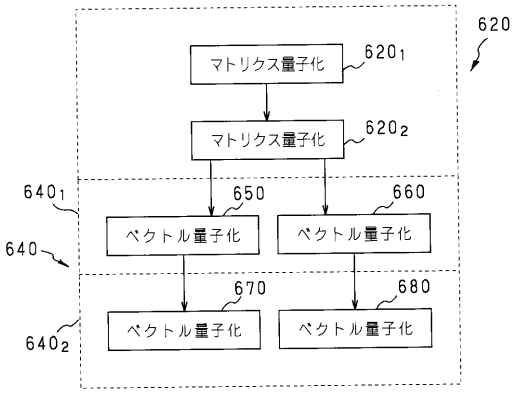
【図３】


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（ツエイインデックス）
UVデータ（ウインインデックス）
HPF
LPC分析
α→LSP変換
LSP量子化
LSP補間
LSP→α変換
聴覚重み付けフィルタ算出
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）
（重み付き）ベクトル量子化
聴覚重み付けフィルタ
距離計算
（重み付き）合成フィルタ A(z)
雑音符号帳

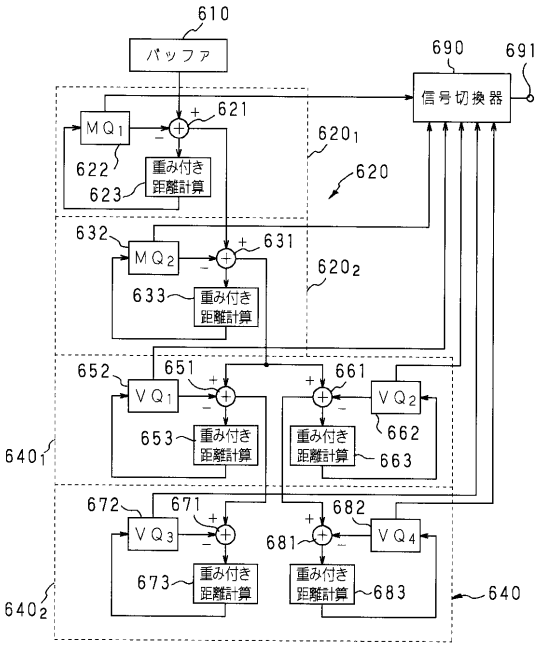
【図４】


出力
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（ツエイインデックス）
UVデータ（ウインインデックス）
LSPの逆量子化
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）
重み付き重畳加算
窓かけ
雑音符号帳
逆ベクトル量子化
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

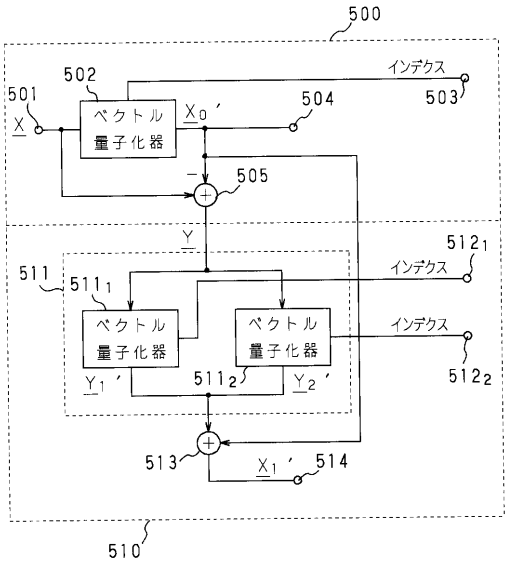
【図５】


620
マトリクス量子化
620_1
マトリクス量子化
620_2
640_1
650
660
ベクトル量子化
ベクトル量子化
640
670
680
ベクトル量子化
ベクトル量子化
640_2

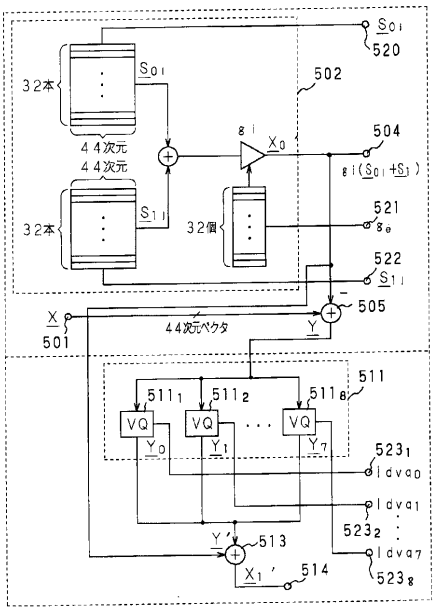
【図６】


610
バッファ
690
691
信号切換器
MQ1
621
622
重み付き距離計算
623
620_1
620
632
MQ2
631
633
重み付き距離計算
620_2
652
651
661
VQ1
VQ2
653
重み付き距離計算
重み付き距離計算
662
663
640_1
672
671
682
VQ3
681
VQ4
673
重み付き距離計算
重み付き距離計算
683
640_2
640

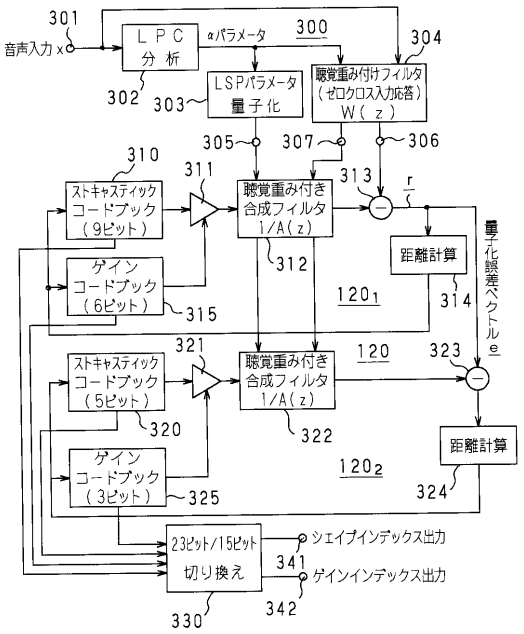
【図７】


500
502
インデクス
501
X
ベクトル量子化器
X_0'
504
503
505
Y
512_1
511
511_1
インデクス
ベクトル量子化器
ベクトル量子化器
インデクス
Y_1'
511_2
Y_2'
512_2
513
X_1'
514
510

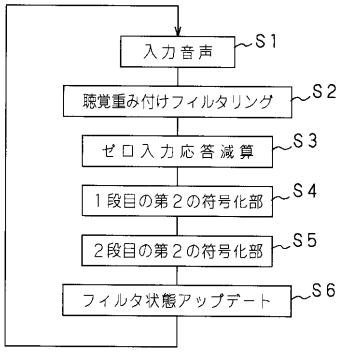
【図８】


S_{01}
520
32本
S_{01}
502
44次元
44次元
g_i
X_0'
504
$g_i(S_{01}+S_{11})$
32本
S_{11}
32個
521
g_0
522
S_{11}
505
X
501
44次元ベクタ
Y
511
511_1
511_2
511_8
VQ
VQ
VQ
Y_0
Y_1
Y_7
523_1
Idva0
Idva1
523_2
Idva7
523_8
Y'
513
X_1'
514

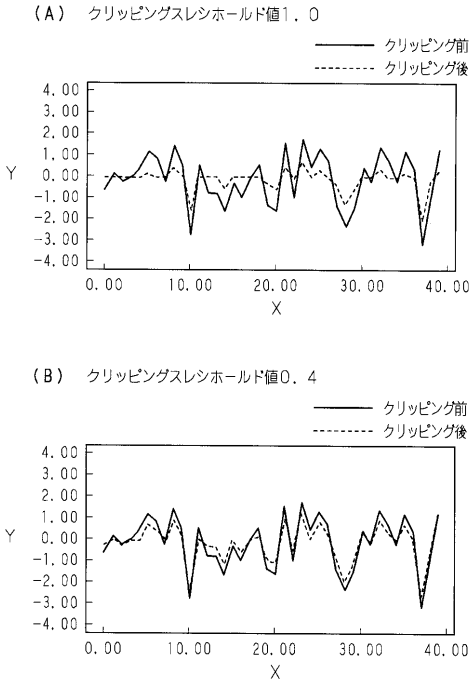
【図９】



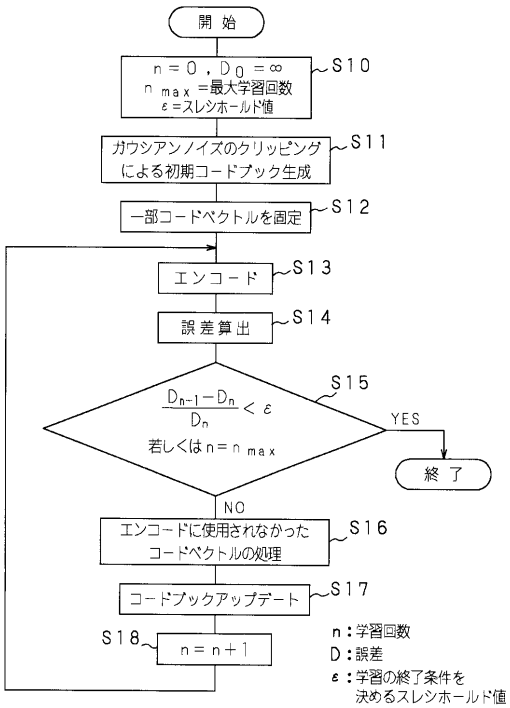
【図１０】



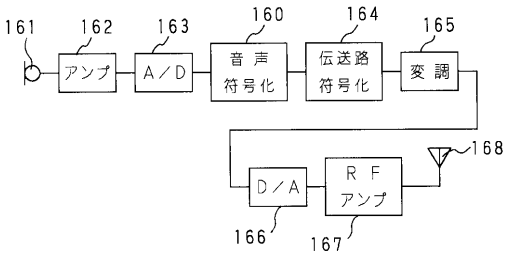
【図１１】



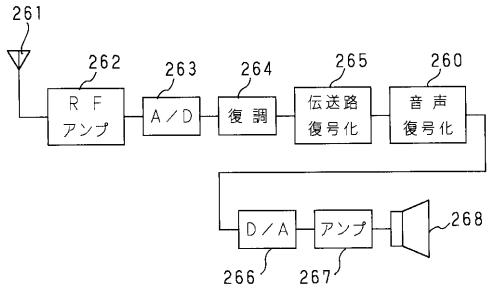
【図１２】



【図１３】


161
162
163
160
164
165
アンプ
A／D
音 声
符号化
伝送路
符号化
変 調
168
D／A
R F
アンプ
166
167

【図１４】


261
262
263
264
265
260
R F
アンプ
A／D
復 調
伝送路
復号化
音 声
復号化
D／A
アンプ
268
266
267

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 原 光明
審判官 加藤 恵一
審判官 脇岡 剛

(56)参考文献 特許第３５５７６６２（ＪＰ，Ｂ２）
特開平５－２８１９９８（ＪＰ，Ａ）
特開平６－２２２７９４（ＪＰ，Ａ）
M. Ｎｉｓｈｉｇｕｃｈｉ, ｅｔ ａｌ, ”Ｈａｒｍｏｎｉｃ ａｎｄ Ｎｏｉｓｅ Ｃｏｄｉｎｇ ｏｆ ＬＰＣ Ｒｅｓｉｄｕａｌｓ ｗｉｔｈ Ｃｌａｓｓｉｆｉｅｄ Ｖｅｃｔｏｒ Ｑｕａｎｔｉｚａｔｉｏｎ”Ｐｒｏｃ. ｏｆ ＩＣＡＳＳＰ－９５, ｖｏｌ. １, １９９５. ５, ｐ. ４８４－４８７
神明夫，他，”スケーラブルな階層構造を持つ楽音・音声の符号化”，日本音響学会平成７年度秋季研究発表会講演論文集，３－１－５，ｐ．２７７－２７８，平成7年9月

(58)調査した分野(Int.Cl.，ＤＢ名)
G10L19/00-19/14