

(19)日本国特許庁(JP)

(12)特許公報 (B 2)

(11)特許番号

特許第 3 2 7 5 2 4 9 号

(P 3 2 7 5 2 4 9)

(45)発行日 平成14年4月15日(2002.4.15)

(24)登録日 平成14年2月8日(2002.2.8)

(51)Int. Cl. ⁷	識別記号	F I
G 1 0 L	19/08	G 1 0 L 9/14
	19/04	
	19/12	

請求項の数 3

(全 1 0 頁)

(21)出願番号	特願平3-225843	(73)特許権者	000004226 日本電信電話株式会社 東京都千代田区大手町二丁目3番1号
(22)出願日	平成3年9月5日(1991.9.5)	(73)特許権者	392026693 株式会社エヌ・ティ・ティ・ドコモ 東京都千代田区永田町二丁目11番1号
(65)公開番号	特開平5-66800	(72)発明者	須田 博人 東京都千代田区内幸町1丁目1番6号 日本 電信電話株式会社内
(43)公開日	平成5年3月19日(1993.3.19)	(72)発明者	守谷 健弘 東京都千代田区内幸町1丁目1番6号 日本 電信電話株式会社内
審査請求日	平成9年9月18日(1997.9.18)	(74)代理人	100066153 弁理士 草野 卓
		審査官	山下 剛史

最終頁に続く

(54)【発明の名称】音声符号化・復号化方法

1

(57)【特許請求の範囲】

【請求項 1】 音声信号をフレームに分割し、そのフレームの特徴パラメータを抽出しこれを符号化し、その音声信号からその符号化された特徴パラメータにより表現される成分を取り除き残差を算出し、その残差を、第2の基本残差と、雑音符号帳から選択した第1、および第3の基本残差との和として表現し、これら各基本残差を量子化する音声符号化方法において、各フレームの第1の基本残差の量子化の前に、そのフレームの複数の仮想的な第2の基本残差を生成し、これら複数の仮想的な第2の基本残差の生成は、そのフレームの前フレーム以前の原音声からその特徴を取り除いた後に残された残差(第1の残差と呼ぶ)を生成し、その前フレームの量子化された第1、第2の基本残差の和を求め、この信号にその前フレーム以前の量子

2

化された残差(第1、第2および第3の基本残差の和)を時間的に接続した信号(第2の残差と呼ぶ)を生成し、これら第1の残差および第2の残差から線形な処理により信号(第3の残差と呼ぶ)を構成し、その第3の残差から切り出した一部分を直接あるいは繰り返して複数の波形を、上記複数の仮想的な第2の基本残差として生成し、上記フレームの残差から上記複数の仮想的な第2の基本残差をそれぞれ差し引いた複数の仮想残差を生成し、これら仮想残差をそれぞれそのフレームの上記第1の基本残差とし、その前フレームのその仮想残差を上記第3の基本残差として、これら第1の基本残差および第3の基本残差の量子化値を同時に決定し、その後、これら量子化された第1の基本残差および第3の基本残差を用い、上記複数の仮想的な第2の基本残差

10

中から上記第 2 の基本残差の量子化値を最終決定する、ことを特徴とする音声符号化方法。

【請求項 2】 上記フレームの第 1 の基本残差およびそのフレームの前のフレームの第 3 の基本残差を、ラップトランスフォーム処理を用いて生成し、これを量子化することを特徴とする請求項 1 記載の音声符号化方法。

【請求項 3】 包絡符号と、ピッチ符号と雑音符号とを入力し、

その雑音符号により雑音符号帳を讀出し、その讀出された雑音波形を 2 フレームに伸ばし、現フレームで伸ばされた雑音波形の前半を取出し、前フレームで伸ばされた雑音波形の後半を取出し、上記ピッチ符号で適応符号帳の波形を繰返し讀出し、その適応符号帳の出力と上記伸ばされた雑音波形の前半と、上記伸ばされた雑音波形の後半とを加算し、その加算出力を、上記包絡符号でフィルタ係数が制御される合成フィルタへ駆動信号として供給して復号音声信号を得る音声復号化方法。

【発明の詳細な説明】

【0001】

【産業上の利用分野】この発明は、音声を低速度で高品質で符号化する符号化方法、及びその符号化された音声符号を復号する復号化方法に関する。

【0002】

【従来の技術】低速度で高品質を達成する音声符号化方式として、M. R. Schroeder, B. S. Atal による文献(文献 1)、“Code-excited linear prediction (CELP): high-quality speech at very low bit rate”、IEEE Proc. of International Conf. Acoust., Speech & Signal Process., pp. 937-940 (1985) 方式や、T. Moriya, M. Honda による文献(文献 2)、“Transform Coding of Speech using a weighted vector quantizer”、IEEE J. Selected Areas in Communications, JSAC-6, pp. 425-431 (1988) などの方式が提案され検討されて来ている。これらの方式は CELP 方式と呼ばれ、4.8 kbps から 8 kbps 程度の符号化速度を対象にしたものであり、広く検討されている。CELP 方式は、音声のスペクトル包絡情報とピッチ情報および残差情報とにより音声を構成するモデルを用いる。そのため、符号化側ではこれら 3 種類の情報を抽出/量子化し、復号側では伝送されてきたこれら 3 種類の情報に基づいて音声波形を再生する。

【0003】これらの情報はいずれもフレームまたはフレームを複数に分割したサブフレーム単位で抽出/量子

化され、これが繰り返される。図 7 に、従来法における音声波形と、スペクトル包絡情報の更新フレームとピッチ情報の更新フレームと、残差情報の更新フレームとの関係例を示す。この例ではスペクトル包絡情報のフレーム更新周期が長く、他の情報のフレーム更新周期の 2 倍になっている。従来の CELP 方式の特徴の 1 つは、全ての情報の更新フレームの境界が一致する時刻が周期的に存在することであり、図 7 では、a, b, c 点がこの時刻に相当する。

10 【0004】

【発明が解決しようとする課題】しかし、このように更新フレームの境界が一致すると、この境界部分に歪みが集中することとなり、特に次のフレームの先頭部分での劣化が大きくなる可能性が高い。これは、量子化値決定における歪み最小の基準範囲がフレームまたはサブフレームに閉じているためである。フレームの中央付近の値の量子化値は前方や後方の値との関係からフレームの歪み最小基準のもとで決定される。ところが、フレームの最後のサンプルの量子化値は、前方の値との関係のみから決定されることになり、後方の値である次のフレームの先頭部分の値とは無関係に決定されてしまう。そのため、次のフレームの先頭部分の値から見ると、歪みを増大する結果となる値が量子化値として選択されることがあり、これが従来方法の問題点であった。

【0005】

【課題を解決するための手段】この発明においては以上の従来の CELP 方式の問題点を解決するため、例えば図 1 に図 7 と対応して示すように、残差情報の量子化の区切り(サブフレーム)が隣どうし重なり合うような構成とする。図 1 に示す同一サブフレームの全体を同時に量子化することで、いずれの部分においても信号の前後の関係を考慮した量子化が実現できる。重なり合うサブフレームの信号がなめらかにつながるように、例えばラップトランスフォーム処理(例えば、文献(文献 3) H. S. Malvar, “Lapped Transforms for Efficient Transform/Subband Coding”, IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-38, pp. 969-978, June 1990 に詳細が記述)を用いる。

【0006】従来の CELP 方式では、量子化雑音がフレームを重ねる毎に蓄積することを防ぐため、過去のフレームの量子化雑音の影響を除去するように現フレームの音声を符号化する構成をとる。具体的には、前フレームの量子化雑音の影響(現フレームの合成フィルタのゼロ入力応答)を現フレームの音声から差し引いて得られた信号を量子化する構成をとっている。この構成を実現するためには、現フレーム音声の量子化処理の前に前フレーム音声の量子化が全て完了していることが必要とな

る。フレーム毎に全ての処理が区切られている従来の CELP 方式においては、現フレーム音声の量子化処理の前に前フレーム音声の量子化が全て完了しているため現フレームの合成フィルタのゼロ入力応答が計算でき、前フレームの量子化雑音の影響を打消すことが可能である。しかし、この発明では例えば図 1 に示すように、現フレーム音声の量子化処理の前に前フレーム音声の量子化が全て完了することにならない。そのままでは、前フレームの量子化雑音の影響を次のフレームで打消すことができないため、以下のようにする。

【0007】あるフレームの処理を始める時点で、前フレームの情報でまだ量子化されていないものは残差である。この量子化されていない残差を適当な条件のもとで量子化されたと仮定してそのフレームの符号化処理を進める構成をとる。ただし、適応符号帳の周期の量子化および残差の形状の量子化（雑音符号帳のインデックス検索）処理において、最適候補以外に複数の準最適な候補を残しておき、この中から最終的に量子化雑音を最小にする組み合わせを選択する。従来の CELP は、適応符号帳の周期の量子化も残差の形状の量子化も 1 回で決定される。

【0008】以上この発明の具体的な主な特徴は、ラプトトランスフォーム処理を用いること、および適応符号帳の周期の量子化および残差の形状の量子化処理において複数の準最適な候補を残しておき最終的に量子化雑音を最小にする組み合わせを選択することである。

【0009】

【実施例】図 2 にこの発明による復号化方法の実施例を示す。受信符号、又はファイルの読出し出力符号としてスペクトル包絡情報を示す包絡符号と、ピッチ情報を示すピッチ符号と、残差情報を示す残差符号とが入力され、現 1 フレーム前の残差符号が第 1 のインデックスとして雑音符号帳 1 1 に与えられ、これより読出された雑音波形が変換器 1 2 で逆変形ラプトトランスフォーム変換され、その 2 フレーム長とされた変換出力は後半切り出し回路 1 3 で波形の後半が切り出されて利得付与回路 1 4 へ与えられ、その出力が第 1 基本残差として加算回路 1 5 へ供給される。

【0010】現フレームの残差符号が第 2 のインデックスとして雑音符号帳 1 1 と同一の雑音符号帳 1 6 に与えられ、これより読出された雑音波形が変換器 1 7 で逆変形ラプトトランスフォーム変換され、その 2 フレーム長とされた変換出力は前半切り出し回路 1 3 で波形の前半が切り出されて利得付与回路 1 9 へ与えられ、その出力が第 3 基本残差として加算回路 1 5 へ供給される。

【0011】入力されたピッチ符号が適応符号帳 2 1 へ与えられ、そのピッチ符号で示す周期で適応符号帳 2 1 が繰返し読出され、その読出された波形は利得付与回路 2 2 で利得が与えられて第 2 基本残差として加算回路 1 5 へ供給される。加算回路 1 5 の出力は適応符号帳に記

憶されると共に線形予測・合成フィルタ 2 3 に駆動信号として供給される。合成フィルタ 2 3 のフィルタ係数は入力された包絡符号により制御される。合成フィルタ 2 3 から復号音声信号が得られる。

【0012】図 3 にこの実施例の復号処理概要における信号のタイミング関係を示す。残差情報は前フレームと現フレームとが半分ずつ重ってそれぞれ符号化されているため、第 3 および第 1 の基本残差が時間的に連続し、かつ前後のフレームと重なりながら残差情報を構成する。このためこの発明ではフレームの先頭部分での歪みが小さく、従来の CELP 方式と大きく異なる。

【0013】図 4 にこの発明による符号化方法の実施例の処理概要を示す。通常の CELP 方式では、適応符号帳のラグ（ピッチ周期相当）を決定し、次に雑音符号帳のインデックスを決定する。しかしこの発明の方法では順番が逆転し、雑音符号帳のインデックスを決定してから適応符号帳のラグを決定する構成を取らざるをえない。この発明では、先に決定される雑音符号帳のインデックスを後に決定する適応符号帳（第 2 の基本残差）と組み合わせたときの mismatches の発生を防ぐため、雑音符号帳のインデックスの決定の前に仮想的第 2 の基本残差（第 2 の基本残差の推定値系列）をこの発明では導入する。

【0014】最適な仮想的第 2 の基本残差を解析的に求めることは難しいため、仮想的第 2 の基本残差の候補を複数作成し、復号音声信号の歪み最小を評価基準として、総当たりを行なう構成とした。ここで、仮想的第 2 の基本残差の候補数を N で表すこととする。さらに、仮想的第 2 の基本残差の各候補毎に、 M 個の雑音符号の候補を残す構成とした。即ち、 $N \times M$ 個の復号音声の候補を作成し、これから歪み最小となる候補を選択する構成をこの実施例はとる。

【0015】図 5 にこの発明の符号化方法を適用した符号化装置の例を示す。入力端子 2 5 からの入力音声は仮想的第 2 基本残差作成部 2 6、LPC 逆フィルタ 2 7 及び最適値検索部 2 8 へ供給される。仮想的第 2 基本残差作成部 2 6 から N 個の仮想的第 2 基本残差が作成されて雑音符号決定用音声合成部 2 9₁ ~ 2 9_N へ供給される。入力音声からスペクトル包絡の特徴を除去した残差波形が LPC 逆フィルタ 2 7 から得られ、この残差波形は雑音符号決定用音声合成部 2 9₁ ~ 2 9_N のすべてへ供給される。

【0016】雑音符号決定用音声合成部 2 9₁ において、第 1 の仮想的第 2 基本残差と LPC 逆フィルタ 2 7 からの残差波形との差が減算器 3 1 で求められ、これより雑音残差、つまり後半が第 1 の基本残差と、前半が第 3 の基本残差と対応したものが得られる。この減算器 3 1 の出力は変換部 3 2 で変形ラプトトランスフォーム変換され、その変換出力は重み付きベクトル量子化/逆量子化部 3 3 で図 2 中の雑音符号帳 1 1, 1 6 と同一の

雑音符号帳 3 4 の何れかの雑音符号に重み付き量子化される。この場合 M 個の候補が求められ、これにより各後半が第 1 の基本残差、前半が第 3 の基本残差の 2 フレーム分の残差波形に対する M 個のインデックスが各フレームごとに得られ、その各インデックスはそれぞれ雑音波形に逆量子化されて、音声波形合成部 3 5₁ ~ 3 5_Mへ供給され、M 個の音声波形に合成されて最適値検索部 2 8 へ供給される。雑音符号決定用音声合成部 2 9₂ ~ 2 9_N も同様に構成される。

【0 0 1 7】以上の処理を以下に更に詳細に説明する。
仮想的第 2 の基本残差の生成法

通常の C E L P 符号化方式においては、あるフレームの適応符号帳には、その前フレームの残差波形を適当な区間で切り出し繰り返した信号が書かれている。この発明の方法においては、あるフレームの前フレームの残差波形は、第 1, 第 2, および第 3 の基本残差の重み付き和に利得を掛けたものとなる。ところが、これら基本残差のうち第 3 の基本残差の量子化が終了(確定)していない。そのため、前フレームの残差波形、およびこれから生成されるそのフレームの第 2 の基本残差が確定できない。

【0 0 1 8】そこで、この第 2 の基本残差の推定系列は以下の方法により生成する。まず、L P C 逆フィルタ 2 7 で、そのフレームの前フレーム以前の原音声からその特徴を取り除いた後に残された残差(第 1 の残差と呼ぶ)を生成する。その第 1 の残差には、量子化雑音が含まれない。すなわち、第 1 の残差で合成フィルタ(量子化されていない係数)を駆動すると、原音声と一致した音声波形が得られることになる。次に、その前フレームの量子化された第 1, 第 2 の基本残差の和を求め、この信号にその前フレーム以前の量子化された残差(第 1, 第 2 および第 3 の基本残差の和)を時間的に連続するように接続した信号(第 2 の残差と呼ぶ)を生成する。その第 2 の残差と前記第 1 の残差との差が、第 3 の基本残差の目標波形(量子化前の第 3 の基本残差)である。さらに、前記第 1 の残差および前記第 2 の残差の重み付き線形和を求め、これを第 3 の残差と呼ぶ。次に、その第 3 の残差から切り出した一部分またはその第 3 の残差を時間的に補間してから切り出した一部分を直接あるいは繰り返して生成した波形を仮想的第 2 の基本残差とする。

【0 0 1 9】ここで、第 3 の残差から仮想的第 2 の基本残差波形を生成する過程を、詳細に述べる。第 3 の残差と現在のフレームの残差の相互相関を計算し、相互相関を最大にする時間差(ピッチラグ)を求める。第 3 の残差の時間的に新しい部分から、ピッチラグの長さの区間を切り出す。切り出した区間を繰り返しながら波形を延ばし、フレームと同じ長さにする。これが、仮想的第 2 の基本残差である。仮想的第 2 の基本残差を N 個生成するためには、前記相互相関を大きくするピッチラグを最

大にするものから N 個残し、それぞれのピッチラグに応じて上記の方法で仮想的第 2 の基本残差を生成する。

第 3 および第 1 の基本残差の同時量子化法

図 5 中の雑音符号決定用音声合成部 2 9₁ の詳細、つまり第 3 および第 1 の基本残差の同時量子化法および最適な合成音声波形の検索を図 6 に示す。前述のようにして生成した仮想的第 2 の基本残差を、そのフレームの残差波形から差し引いた波形を減算器 3 1 で生成する。この波形を第 1 の基本残差として、前フレームの量子化前の同様に得た基本残差、つまり第 3 の基本残差の後に時間的に連続するように合成回路 3 6 で接続する。この接続波形を、変換部 3 2 でモジュレイテッドラップトランスフォーム(Modulated Lapped Transform: MLT, 例えば、文献 3 に詳細が記述)アルゴリズムにより変換し 2 分の 1 のデータ数の周波数領域の波形を得る。この波形を、量子化部 3 3 a で重み付きベクトル量子化(Weighted Vector Quantization: WVQ, 文献 2 に詳細が記述)処理により雑音符号帳 3 4 の何れかの雑音符号に量子化する。この周波数領域の波形を雑音符号帳 3 4 の何れかの雑音波形で代表させることを各フレームごとに行う。このとき複数個(M 個)の候補を残す点にこの発明の特長がある。さらに、これら残された M 個の各候補について、以下の処理を行なう。

【0 0 2 0】(a) 逆量子化部 3 3 b で雑音符号帳 3 4 を用いて逆量子化し、(b) その逆量子化出力を変換部 3 7 で逆変形ラップトランスフォーム変換を行い、(c) その変換出力を音声波形合成部 5 5 で音声波形に合成する。

以上の処理により、M 個の候補に対する復号音声を生成する。

原音声波形との距離最小となる候補を選択

以上の処理の結果、合計で N × M 個の復号音声の候補が生成される。検索部 2 8 で各候補復号音声と原音声波形との距離を計算し、距離最小となる復号音声を選定する。

【0 0 2 1】最後に第 2 の基本残差を再決定する。再決定のため、図 5 中の第 2 の基本残差決定部 3 8 で距離最小となる復号音声を与える量子化された第 3 および第 1 の基本残差(図 4 の③の信号)を用い、第 2 の基本残差(図 4 の④の信号)を決定する。以上、この発明の符号化処理を図 4 にそってまとめる。

【0 0 2 2】(1) 原音声波形を L P C 逆フィルタリングした残差波形から、複数(N 個)の仮想的第 2 基本残差を生成する。

(2) 各仮想的第 2 基本残差に対して、複数(M 個)の量子化された第 3 および第 1 基本残差の組を求める。

(3) 復号音声と原音声波形との距離を選択の尺度として、N × M 個の復号音声の候補から最適候補を選定する。

【0023】(4) 決定された最適候補を与える、量子化された第3および第1基本残差を用いて、第2基本残差を決定する。

図2において、逆MLT処理は高速に行われるため、雑音符号帳11, 16、逆MLT回路12, 17を各1つとし、各時間領域に変換された2フレームの残差の前半を現フレームで、後半を前フレームで用いるようにしてもよい。上述では重複サブフレームを1としたが複数サブフレームずつ重複させてもよい。

【0024】

【発明の効果】以上述べたように、この発明によれば、残差情報、つまり雑音波形の量子化をフレームを一部重複させながら行っているため、フレームの境界における量子化雑音の増加を押さえ、かつ平均的な量子化雑音が低下し、しかも少ない量子化ビットで済む。

【図面の簡単な説明】

【図1】この発明におけるフレーム更新例の概要を示すタイムチャート。

【図2】この発明の復号化方法を適用した復号化器の一例を示すブロック図。

【図3】図2の復号化器の処理の流れを示すタイムチャート。

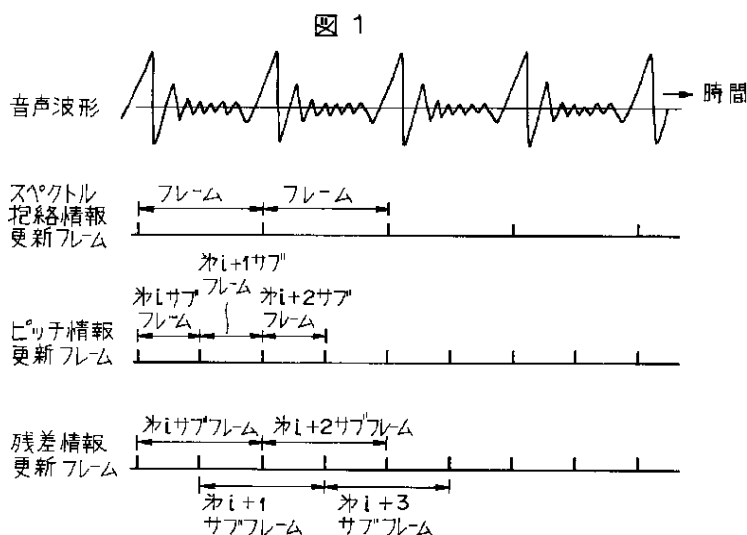
【図4】この発明による符号化方法の例を示すタイムチャート。

10 【図5】この発明の符号化方法を適用した符号化器の一例を示すブロック図。

【図6】図5中の雑音符号決定用音声合成部の具体例を示すブロック図。

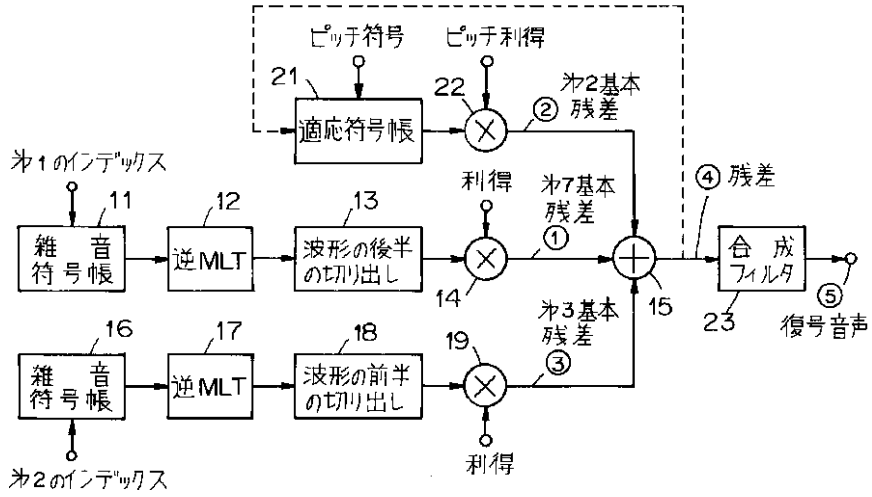
【図7】従来の音声符号化方法におけるフレーム更新を示すタイムチャート。

【図1】

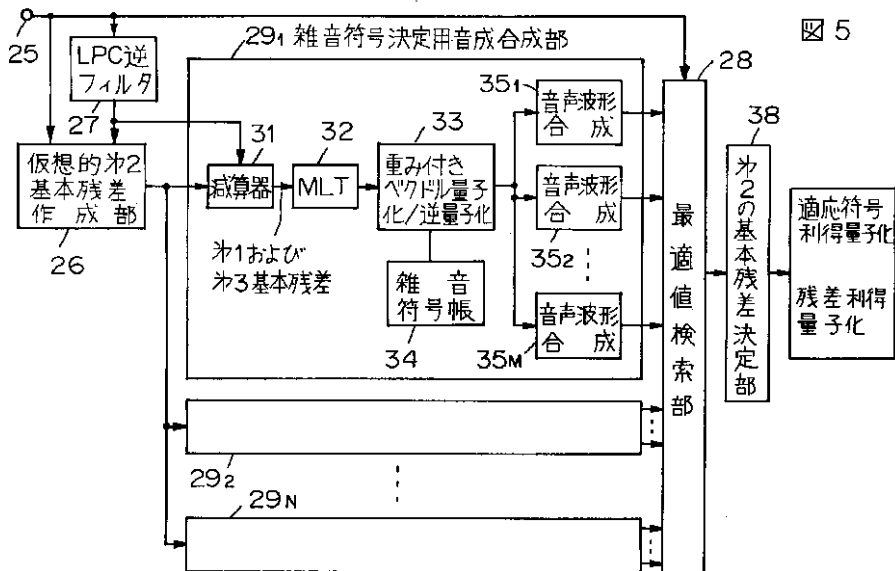


【図2】

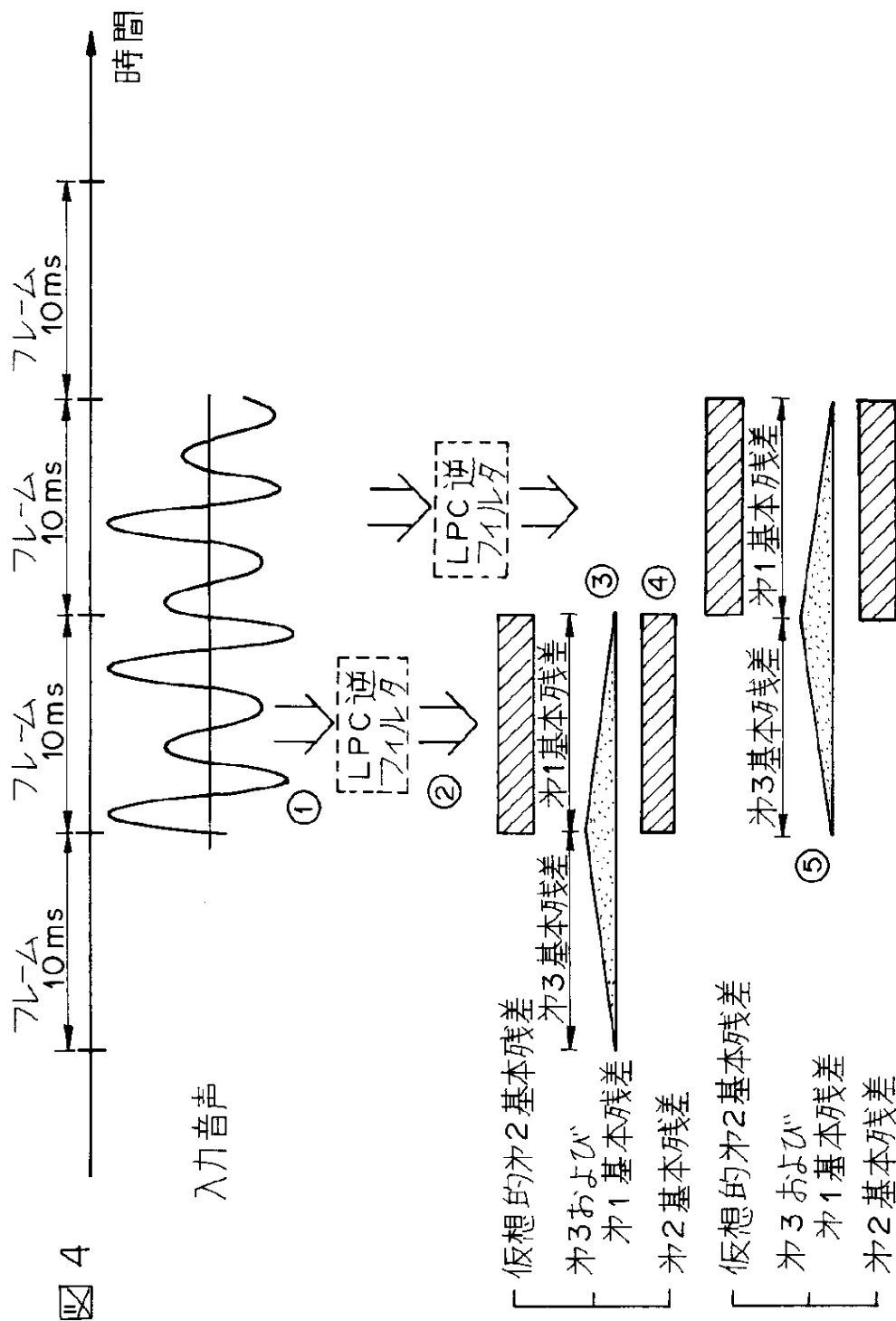
図2



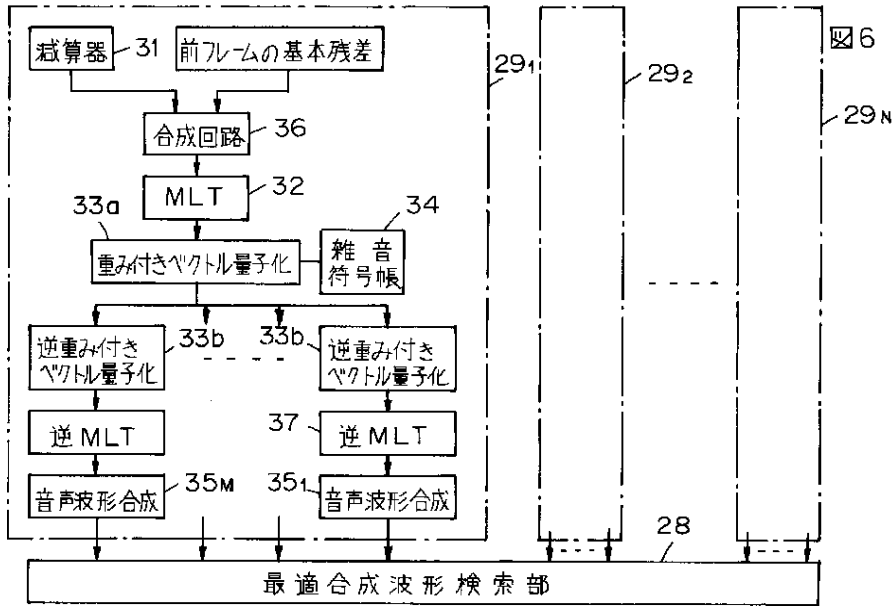
【図5】



【図 4】

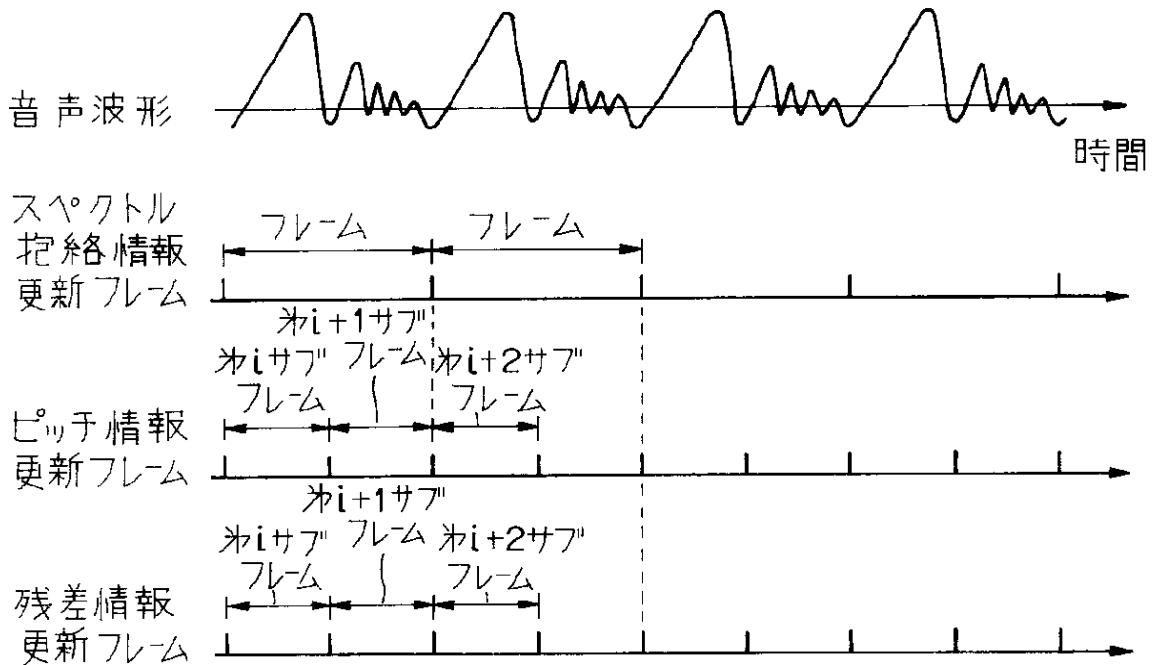


【図 6】



【図 7】

図 7



フロントページの続き

(58)調査した分野(Int.Cl.⁷, DB名)

G10L 19/00 - 19/14

H03M 7/30

H04B 14/04

(19)日本国特許庁(JP)

(12)特許公報 (B 2)

(11)特許番号

特許第 3 2 9 6 3 6 3 号

(P 3 2 9 6 3 6 3)

(45)発行日 平成14年6月24日(2002.6.24)

(24)登録日 平成14年4月12日(2002.4.12)

(51)Int. Cl.⁷

識別記号

F I

G 1 0 L 19/04
19/00

G 1 0 L 9/14 J
9/18 E

請求項の数 5

(全 1 2 頁)

(21)出願番号 特願平3-99173
(22)出願日 平成3年4月30日(1991.4.30)
(65)公開番号 特開平4-328800
(43)公開日 平成4年11月17日(1992.11.17)
審査請求日 平成9年12月12日(1997.12.12)

(73)特許権者 000004226
日本電信電話株式会社
東京都千代田区大手町二丁目3番1号
(72)発明者 大室 伸
東京都千代田区内幸町一丁目1番6号 日本
電信電話株式会社内
(72)発明者 守谷 健弘
東京都千代田区内幸町一丁目1番6号 日本
電信電話株式会社内
(72)発明者 誉田 雅彰
東京都千代田区内幸町一丁目1番6号 日本
電信電話株式会社内
(74)代理人 100066153
弁理士 草野 卓

審査官 山下 剛史

最終頁に続く

(54)【発明の名称】音声の線形予測パラメータ符号化方法

1

(57)【特許請求の範囲】

【請求項 1】 音声スペクトル包絡特性を表す線形予測パラメータを符号化する符号化方法において、
上記線形予測パラメータを粗い時間間隔 mT (m は2以上の整数)で算出し、その算出した線形予測パラメータを予め決められた複数の代表値の中の一つで表現する前段符号化過程と、
上記粗い時間間隔 mT よりも狭い時間間隔 T における上記線形予測パラメータを上記粗い時間間隔 mT ごとの代表値より推定する過程と、
その推定値と、上記狭い時間間隔 T の各時刻で算出した真の線形予測パラメータとの誤差が最小となる代表値を、予め決められた複数の代表値の中から選択する後段符号化過程とを有し、
上記線形予測パラメータを上記前段符号化過程で得られ

2

た代表値と上記後段符号化過程で得られた代表値とによって表現することを特徴とする音声の線形予測パラメータ符号化方法。

【請求項 2】 上記前段符号化過程で線形予測パラメータを表現する代表値として誤差の小さい順に複数の候補を残し、

上記後段符号化過程では、これら候補の各々に対して誤差最小の代表値を決め、これら上記各候補とその対応する後段符号化過程とによって表現される線形予測パラメータと真の線形予測パラメータとの誤差が最小となる組合せを前段符号化過程の代表値と後段符号化過程の代表値とを最終的に決定することを特徴とする請求項 1 記載の音声の線形予測パラメータ符号化方法。

【請求項 3】 上記狭い時間間隔 T の推定値と、同時刻における真の線形予測パラメータとの誤差を、狭い時間

10

間隔工の各時刻でパラメータ次元方向にまとめてベクトル形式で表現し、これらベクトルを連続する時間間隔工ごとの複数の時刻でまとめた行列形式で上記後段符号化過程における代表値を表現することを特徴とする請求項 1 または 2 記載の音声の線形予測パラメータ符号化方法。

【請求項 4】 上記線形予測パラメータを線スペクトル対と呼ばれるパラメータで表現し、上記推定過程で推定される上記推定値に、上記後段符号化過程の代表値を加算することによって線形予測パラメータを表現することを特徴とする請求項 1 乃至 3 の何れかに記載の音声の線形予測パラメータ符号化方法。

【請求項 5】 上記推定値を、前段符号化過程で mT ごとに得られた隣接する代表値から線形補間法によって決定することを特徴とする請求項 1 乃至 3 の何れかに記載の音声の線形予測パラメータ符号化方法。

【発明の詳細な説明】

【0001】

【産業上の利用分野】この発明は音声の伝送、蓄積などに用いられ、情報量を減少するための音声スペクトル包絡特性を表す線形予測パラメータを符号化する方法に関する。

【0002】

【従来の技術】従来の音声符号化系において、音声スペクトル包絡特性を表す線形フィルタの係数は、線形予測分析を一定時間間隔で行って計算され、偏自己相関 (PARCOR) 係数や、線スペクトル対 (LSP) などのパラメータに変換、量子化され、デジタル符号に変換した後、記憶、または伝送された。これらの方法の詳細は、例えば、古井貞熙著“デジタル音声処理”(東海大学出版会)に記載されている。

【0003】このとき、係数を更新する時間間隔を長くすれば、記憶または伝送のための情報量を減らすことができるが、あまり長くしすぎるともとの音声に合成するときに記憶または伝送しなかった時間におけるフィルタ係数の推定精度が低下し、再生音声の品質劣化につながった。そこで、一般には上記時間間隔を 20 ミリ秒程度以下に設定している。

【0004】

【発明が解決しようとする課題】この問題を解決し、より高効率な符号化を行うための方法として、マトリクス量子化と呼ばれる方法がある。これは、線形予測分析は短い時間間隔、例えば 10 ミリ秒から 20 ミリ秒程度の間隔で行っておき、数組の分析結果をひとまとめにして符号化する方法である。しかしこの方法によって、直接数組のフィルタ係数を量子化しようすると、非常に多くの計算量と、符号帳のための多くの記憶とが必要になり、現実的なハードウェア規模のもとではそれほど効率をあげることができない。

【0005】この発明の目的は、マトリクス量子化が量

子化歪を小さくする観点で高効率であるという利点を生かしつつ、現実的な計算量と記憶量とのもとの高効率な予測パラメータ(フィルタ係数)の符号化(量子化)方法を提供することにある。

【0006】

【課題を解決するための手段】この発明は、音声の符号化において、線形予測パラメータ(線形フィルタの係数)を、最初に粗い時間間隔、例えば 40 ミリ秒毎に符号化(量子化)し(前段の符号化又は前段量子化と呼ぶ)、この前段量子化値と、過去の時間における前段量子化値とから、前段量子化の時間間隔よりも狭い時間間隔の値を、適当な推定方法、例えば補間によって適当な時間間隔、例えば 10 ミリ秒毎に推定する。

【0007】次に、上記推定値と上記狭い時間間隔の真の線形予測パラメータ(線形フィルタ係数)との誤差系列を、適当な時間の組、例えば 4 つをひと組にして符号化(量子化)する(後段符号化又は後段量子化と呼ぶ)。この後段量子化の方法には、例えばマトリクス量子化などが用いられる。このように、この発明では、前段量子化の後に後段量子化を行うが、前段量子化の最適量子化値が、前段量子化と後段量子化とを組み合わせたときの最適値になっているとは限らない。そこで、請求項 2 の発明では前段量子化では、一意に量子化値を決定してしまうのではなく、数個の候補を残し、それぞれの候補に対してそれぞれ後段量子化を行って、両者を組み合わせた時の最適量子化値を求めて全体の性能を高める。この方法は、一般にディレイドディジションと呼ばれる。

【0008】再生(復号)側では、前段量子化値からこれら間の値を狭い時間間隔で推定し、この推定値の系列と後段量子化値の系列とから再生される値を線形フィルタ係数とする。音声信号を再生する場合には、上記再生されたフィルタ係数をもつ線形フィルタを駆動音源信号によって駆動して音声を合成する。なお、この量子化における最小量子化時間単位、例えば 10 ミリ秒以下の時間間隔のフィルタ係数については、再生側で最も近い時刻の係数をそのまま用いてもよいし、任意の時刻(通常は前後の時刻)の量子化値から補間によって推定してもよい。

【0009】

【作用】この発明において、前段量子化と後段量子化とに割り当てるビット数を適当に定めることにより、高効率でかつ、現実的な計算量、記憶量のハードウェアで実現可能とする。すなわち、前段量子化に線形フィルタ係数を量子化するための全ビットを割り当てると、後段量子化を使用しない従来と同様の量子化方法となり、後段量子化にマトリクス量子化を用い、これに全ビットを割り当てると、数組のフィルタ係数に直接マトリクス量子化を適用する従来方法になる。既に指摘したように、前者は計算量や記憶量の点で優れているが、性能的には限

界がある。一方、後者は性能の点では優れていると考えられるが、そのままではハードウェア的な見地から好ましくない。そこで、この発明では、前者と後者とに適当なビットを配分して、組み合わせることにより、現実的にかつ高能率な量子化器を設計することができるようにしたものである。

【0010】

【実施例】以下、この発明の一実施例を図面を参照して説明する。図1にこの発明による音声の線形予測パラメータ符号化法の構成例を示し、図2にその再生側の構成を示す。図1において、入力端子1からは、標本化されたデジタル信号 $s(t)$ が入力される。線形予測分析部2では、音声信号 $s(t)$ のN個のサンプルをいったんデータバッファに蓄えた後、これらのサンプルに対して線形予測分析を行って、一組の線形予測係数 $a_i(k)$, ($i = 1, 2, \dots, p$) を算出する。ここで、パラメータ p は分析次数と呼ばれ、一般には10~16程度の値が用いられる。また上記Nサンプルの単位はフレーム更新周期または単にフレーム周期と呼ばれる。この結果、N個の入力サンプルに対して p 個の線形

ラメータ

$$f_i(k-j), i=1, 2, \dots, p$$

$$j=m-1, m-2, \dots, 1, 0$$

を蓄え、これらの値を前段量子化部5と後段量子化部としてのマトリクス量子化部8とに供給する。図1では説明を簡単にするため、 $m = 4$ の場合について示した。

【0013】前段量子化部5では、バッファ部4によって供給された $m \times p$ 個のパラメータのうち、最後のフレームの p 個のパラメータ $f_i(k)$ を予め決められたビット数 B_1 で量子化する。つまり予め決めた複数の代表値の中の一つで表現する。なお、ここで用いる量子化法は任意の方法を用いてよい。例えば図3に示す、ディレイドディシジョンを用いない多段ベクトル量子化法を用いる。図は簡単のため $N = 3$ 段の場合について示す。多段ベクトル量子化器は、符号帳 23_n ($n = 1, 2, 3, \dots, N$)、ベクトル加算器 27_n 、歪判定部 28_n を1組とする量子化段の直列的な組み合わせからなる。ただし、1段目についてはベクトル加算器 27_1 は不要

$$f_i(k) \text{ と } q_i^{(1)} + c_i^{(2)}(r) \text{ との歪}$$

が最小となる r を r_2 として、

$$q_i^{(2)} = q_i^{(1)} + c_i^{(2)}(r_2)$$

を2段目までの量子化値として決定する。同様にして $n - 1$ 段目までの量子化値 $q_i^{(n-1)}$ が決まったとする

$$f_i(k) \text{ と } q_i^{(n)} = q_i^{(n-1)} + c_i^{(n)}(r_n) \text{ との歪}$$

が最小となるように決定する。こうして得られたN段目までの量子化値(ベクトル) $q_i^{(N)}$ を $f_i(k)$ の量子化値 $f_i^*(k)$ として出力する。

【0015】少ない計算量でしかも量子化能率が高い点から、図4に示すディレイドディシジョンにつき多段ベ

* 予測係数が出力されることになる。そこでここでは、線形予測係数の時刻を表す単位として、N個のサンプルを単位とした時間単位を k で表し、「第 k フレームの i 次の線形予測係数 $a_i(k)$ 」などと呼ぶことにする。なお、これらの処理方法の詳細は、前述の古井の著書に記載されている。

【0011】線スペクトル対(LSP)算出部3では、 p 個の線形予測係数を同じく p 個の線スペクトル対、 $f_i(k)$ に変換する。この実施例において線形予測係数を線スペクトル対に変換したのは、線スペクトル対の性質を用いると、より高品質な量子化結果が得られるからである。しかし、この線スペクトル対算出部3は必ずしも必要ではなく、線形予測係数 $a_i(k)$ を直接バッファ部4に入力してもよい。また、線スペクトル対のほか、偏自己相関(PARCOR)係数など、線形予測係数と相互変換が可能な任意のパラメータに変換してもよい。

【0012】バッファ部4では、 p 個の線スペクトル対を m フレーム分、即ち $m \times p$ 個の p

で省略されてある。

【0014】まず、符号帳 $1(23_1)$ の保持する R_1 個の各コードベクトルのうち r 番目のベクトルを $c_i^{(1)}(r)$ として順に歪判定部 28_1 に送る。歪判定部 28_1 では予め定義された歪尺度に基づいて入力ベクトル $f_i(k)$, ($i = 1, 2, \dots, p$) と $c_i^{(1)}(r)$ との歪を計算し、最も歪の小くなったときのコードベクトル $c_i^{(1)}(r_1)$ を1段目の量子化値 $q_i^{(1)}$ として出力する。次に2段目の量子化部では、符号帳 $2(23_2)$ の保持する R_2 個のコードベクトルのうち、 r 番目のベクトルを $c_i^{(2)}(r)$ とし、これをベクトル加算器 27_2 で1段目の量子化値 $q_i^{(1)}$ に加えて $q_i^{(1)} + c_i^{(2)}(r)$ を求める。歪判定部 28_2 では、1段目と同様にして

と、 n 段目の量子化値 $q_i^{(n)}$ は、

クトル量子化が好ましい。この量子化法の図3に示したものと違いは、各量子化段で最適量子化値 $q_i^{(n)}$ を一意に決定してしまうのではなく、歪の小さい順にいくつかの量子化値を量子化候補保持部 35_n に保持しておく。そして次の段の歪判定部 28_{n+1} において前段の量

量子化値としてどの候補を用いると歪が最も小さくなるかを判断して前段の量子化値を決定する。具体的には、 $n - 1$ 段目までの量子化値の第 u 番目候補を $q_i^{(n-1)}(u)$ とする。

$$f_i^{(n)}(k) \text{ と } q_i^{(n-1)}(u) + c_{ij}^{(n)}(r) \text{ との歪}$$

が最小となるような u を u_r として決定する。こうしてすべての r について上記 u_r を求めた後、 r を変化させ

$$f_i^{(n)}(k) \text{ と } q_i^{(n-1)}(u_r) + c_{ij}^{(n)}(r) \text{ との歪}$$

が小さい順に n 段目の u 番目の量子化候補 $q_i^{(n)}(u)$ として保持する。

【0016】一般にディレイドディシジョンつき多段ベクトル量子化は、最後段では量子化候補を保持する必要がなく、最も歪の小さい量子化値に決定する。しかし請求項 2 の発明では、上記ディレイドディシジョンつき多段ベクトル量子化を、前段量子化器 5 として用いた場合、後段量子化器であるマトリクス量子化においてもディレイドディシジョンを用いるため、最後段の後にも、図 4 における最後段の歪判定部 2 8 3 より複数の量子化候補を出力させ、図 1 中の量子化候補保持部 6 に保持させる。即ち、請求項 2 の発明では前段量子化部 5 として図 4 のベクトル量子化器を用いる場合、その図 4 のベクトル量子化の最後段目でも一意に量子化値を決定してし

$$f_i^*(k-j) = f_i^*(k) + j \cdot f_i^*$$

ただし、

$$f_i^* = \{ f_i^*(k-m) - f_i^*(k) \} / m$$

で表される。

【0018】補間部 8 によって推定された推定値 $f_i^*(k-j)$ ($j = m-1, m-2, \dots, 2, 1$) と $f_i^*(k)$ との各パラメータは、マトリクス量子化部 9 に供給される。マトリクス量子化部 9 は上記 $f_i^*(k-j)$ ($j = m-1, m-2, \dots, 1, 0$) の値と、バッファ部 4 から出力されている真の値 $f_i(k-j)$ とから、最適量子化値 $F_{ij}^*(k)$ と歪尺度 d_{min} とを出力する。このマトリクス量子化部 9 の構成例を図 5 に示す。符号帳 4 1 はマトリクス量子化に割り当てられたビット数に対応する数のコードマトリクスを保持している。割り当てられるビット数を BMQ とすると、符号帳 4 1 に保持するコードマトリクスの数は、 $R_{MQ} = 2^{BMQ}$ である。符号帳 4 1 は、上記 R_{MQ} 個のコードマトリクスを順に行列加算器 4 2 に供給する。このとき、符号帳 4 1 から供給される r 番目のコードマトリクスを $C_{ij}^{(MQ)}(r)$ とする。

$$d(r) = \sum \sum \{ w_{ij} \{ f_i(k-j) - U_{ij}(r) \}^2 \}$$

$$d_{min} = \min d(r)$$

ただし、最初の $j = 0$ から $m-1$ まで、次の $i = 1$ から p まで、 w_{ij} は重み係数で、適当に定めること

$$w_{ij} = 2 \cdot 1 / \{ \{ f_{i+1}(k-j) - f_i(k-j) \} + \{ f_i(k-j) - f_{i-1}(k-j) \} \}$$

$$\text{ただし、} f_0(k-j) = 0, 0, f_{p+1}(k-j) = \pi$$

とする。また、ケプストラム距離を用いる場合には、 $f_i(k-j)$ ($i = 1, 2, \dots, p$) と $U_{ij}(r)$ をそ

$f_i^{(n-1)}(u)$ とする。 n 段目の符号帳の r 番目のコードベクトルを $c_{ij}^{(n)}(r)$ とすると、まず

て

まうのではなく、量子化候補保持部 6 にいくつかの候補を残した後、図 1 におけるマトリクス量子化まで含めて最適値となるように前段量子化値を決定する。そのための帰還を図 1 において歪判定部 1 0 から量子化候補保持部 6 に対して行う。

【0017】図 1 において前段量子化部 5 で $f_i(k)$ が量子化された値 $f_i^*(k)$ は、バッファ部 7 と補間部 8 とに供給される。バッファ部 7 では、既に量子化された $f_i^*(k)$ の m フレーム前の値 $f_i^*(k-m)$ を保持し、これを補間部 8 へ供給する。補間部 8 では、 $f_i^*(k-m)$ と $f_i^*(k)$ との値とから、線形補間によってこれら間の各フレームにおける量子化値 $f_i^*(k-j)$ ($j = m-1, m-2, \dots, 2, 1$) を推定する。即ち、

行列加算器 4 2 は、この $C_{ij}^{(MQ)}(r)$ と、前段量子化値と推定値 $f_i^*(k-j)$ ($j = m-1, m-2, \dots, 2, 1, 0$) とを加算して、マトリクス $U_{ij}(r)$ を生成する。即ち

$$U_{ij}(r) = f_i^*(k-j) + C_{ij}^{(MQ)}(r)$$

$$i = 1, 2, \dots, p$$

$$j = 0, 1, \dots, m-1$$

歪判定部 4 3 では、上記 $U_{ij}(r)$ とバッファ部 4 からの入力される真の値 $f_i(k-j)$ との歪をはかり、歪が最も小さくなるコード番号 r_{MQ} を決定し、そのときの $C_{ij}^{(MQ)}(r_{MQ})$ と歪 d_{min} とを出力する。歪尺度には、LSP ユークリッド距離、重み付き LSP ユークリッド距離、ケプストラム距離など、任意の尺度を用いてよい。例えば、重み付き LSP ユークリッド距離ならば、

によって量子化性能を高めることができる。例えば、

$f_i(k-j)$ ($i = 1, 2, \dots, p$) と $U_{ij}(r)$ をそ

れぞれケプストラム $f_{i,c}(k-j)$ と $U_{i,c}(r)$ ($i = 1, 2, \dots, n_c$) に変換した後ユークリッド距離を求めればよい。ここで、 n_c はケプストラム次数と呼ばれ、通常は分析次数 $p \sim 30$ 程度の間の値が用いられる。

【0019】前段量子化部5で1つの量子化値 $f_{i^*}(k)$ のみを出力する場合は d_{min} を出力することなく、マトリクス量子化部9で量子化した $F_{ij^*}(k)$ を出力すればよい。前段量子化部5で複数の候補を出力する場合は、マトリクス量子化部9の出力 d_{min} は、前段の量子化部の量子化候補保持部6へ帰還される。量子化候補保持部6は、図3又は図4中の歪判定部28₃における歪の小さい順にM個量子化候補を保持し、マトリクス量子化部9から帰還する d_{min} の値が最小となる $f_{i^*}(k; u)$ を最適な前段の量子化値 $f_{i^*}(k)$ として出力する。これを前述のように、ディレイドディシジョンと呼ぶ。なお、量子化候補保持部6における候補の数Mは、前述のように大きいほどこの発明の性能を高めることができるが、逆に計算量が増大し*

$$f_{i^*}(k-j) = f_{i^*}(k) + j \cdot \Delta f_{i^*}$$

ただし、

$$\Delta f_{i^*} = [f_{i^*}(k-m) - f_{i^*}(k)] / m$$

で表される。行列加算器17は、上記

$$\begin{aligned} & f_{i^*}(k-j) \\ & j = m-1, m-2, \dots, 1, 0 \\ & i = 1, 2, \dots, p \end{aligned}$$

と

$$\begin{aligned} & F_{ij^*}(k') \\ & j = m-1, m-2, \dots, 1, 0 \\ & i = 1, 2, \dots, p \end{aligned}$$

を加算し、

$$\begin{aligned} & f'_{i^*}(k-j) \\ & j = m-1, m-2, \dots, 1, 0 \\ & i = 1, 2, \dots, p \end{aligned}$$

を得る。ここで $f_{i^*}(k-j)$ は線スペクトル対 $f_{i^*}(k-j)$ の量子化値に相当する。この $f_{i^*}(k-j)$ の組は、バッファ部18に送られる。バッファ部18では $m \times p$ 個のパラメータ $f_{i^*}(k-j)$ を蓄え、 $f_{i^*}(k-(m-1))$ 、 $f_{i^*}(k-(m-2))$ 、...、 $f_{i^*}(k)$ の順に1フレーム(p個)ずつフィルタ係数変換部19に出力する。フィルタ係数変換部19は受け取った $f_{i^*}(k)$ 、即ち線スペクトル対を線形予測係数 $a_{i^*}(k)$ に変換する。

【0021】以上の手順によって線スペクトル対が量子化される様子を図6に模式的に示す。図6(a)は音声波形を表す。線形予測分析は音声波形の下に枠で示すようなフレームと呼ばれる一定の長さを単位として行われ、Nサンプルずつフレームの位置をずらしながら分析

*ハードウェアに負担がかかる。そこで、Mは性能があまり劣化しない範囲でできるだけ小さい数がよく、通常4~8程度の値を用いる。なお、M=1ならばディレイドディシジョンを用いない場合に等しい。

【0020】こうして、線形フィルタ係数分析部12から前段の量子化値 $f_{i^*}(k)$ とマトリクス量子化値 $F_{ij^*}(k)$ とが出力される。これと前段量子化値 $f_{i^*}(k)$ とが符号化部13で符号化され、伝送又は蓄積などされる。再生側では、図2に示すように符号化側によって伝送または蓄積された符号を復号化部14で復号し、それぞれ $f_{i^*}(k)$ と $F_{ij^*}(k)$ を得る。このうち $f_{i^*}(k)$ は、補間部16とバッファ部15とに供給される。バッファ部15は、バッファ部7と同様に、 $f_{i^*}(k)$ のmフレーム前の値 $f_{i^*}(k-m)$ を蓄え、補間部16に供給する。補間部16補間部8と同様の動作をし、 $f_{i^*}(k)$ と $f_{i^*}(k-m)$ とから線形補間によって、 $f_{i^*}(k-j)$ ($j = m-1, m-2, \dots, 1$) を推定する。即ち、

を続ける。各分析フレーム毎に得られたp個の線形予測係数をp個の線スペクトル対(LSP)に変換したものを図6(b)に黒丸印で示す。周波数の低い方から順に1次、2次、...、p次のLSPと呼ぶ。ここでは簡単のため、 $p = 4$ として図示したが、実際の分析では、分析次数pとして10~16の値が用いられる。図のように、線スペクトル対(LSP)の値は時間とともに緩やかに変化し、異なる次元のパラメータが互いに交差することはない。この実施例ではこの線スペクトル対の緩やかに変化する性質を利用して高品質な量子化を実現する。

【0022】まずはじめに、mフレームに1回のp個の線スペクトル対をひとまとめにしてベクトル量子化(VQ)する。これを図6(b)に白丸印で示す。図はm=3の場合である。次に上記ベクトル量子化(VQ)した

結果の量子化値を線形に補間する。これを図 6 (c) に破線で示す。このような線形補間だけでは図 6 (c) に示すように、ベクトル量子化しなかったフレームの線スペクトル対は誤差が大きくなってしまふ。そこで真の線スペクトル対と破線で示される補間値との誤差を p 個ずつ m フレーム分、図 6 (c) ではそれぞれ黒丸印と斜線丸印とで示される 1 2 個ずつのデータと破線との誤差をひとまとめにしてマトリクス量子化 (MQ) する。

【 0 0 2 3 】通信路または蓄積装置では、上記ベクトル量子化値とマトリクス量子化値とを適当な符号で表現して伝送または蓄積する。再生側ではまず m フレーム毎にベクトル量子化値を再生し、線形補間した後、マトリクス量子化値を線形補間値に加算して量子化された p 個の線スペクトル対の組が再生される。

【 0 0 2 4 】なお、この発明は、全部を論理回路の組み合わせによるハードウェアで実現してもよいし、一部または全部をマイクロプログラムを含めた、ソフトウェアによって実現してもよい。変形例ベクトル量子化値を線*

$$f_i^*(k-j), j = m-1, m-2, \dots, 1$$

を推定しているが、補間する事を前提として前段の量子化値を決定してもよい。即ち、

$$f_i(k-j), j = m-1, m-2, \dots, 1, 0$$

と

$$f_i^*(k-j), j = m-1, m-2, \dots, 1, 0$$

との歪が最小になるように $f_i^*(k)$ を決定する。なお、 $f_i(k-j)$ と $f_i^*(k-j)$ との歪尺度は適

$$d = \sum \sum [w_{i,j} \{f_i(k-j) - f_i^*(k-j)\}^2]$$

$$w_{i,j} = 2.0 / [\{f_{i+1}(k-j) - f_i(k-j)\} + \{f_i(k-j) - f_{i-1}(k-j)\}]$$

ただし、最初の $j = 0$ から $m - 1$ まで、次の $i = 1$ から p まで、 f_0

$$(k-j) = 0.0, f_{p+1}(k-j) =$$

のようにする。この場合は若干計算量が增大するが、量子化性能は良い。上記効果はディレイドディシジョンの候補数を十分多くとることによって包含されるが、ディレイドディシジョンの候補数を増やすよりは計算量の増大は少ない。

【 0 0 2 7 】このほか前記実施例では、ベクトル量子化をしたフレームにおける誤差も含めてマトリクス量子化をしたが、ベクトル量子化にある程度十分なビットを割り当てれば、ベクトル量子化したフレームはそのままに

$$f_i^*(k-j), j = m-1, m-2, \dots, 2, 1$$

の値を推定するため、もし $f_i^*(k-m)$ が伝送誤りによって重大な劣化を生じた場合には、送信側と受信側*

$$f_i^*(k-j), j = m-1, m-2, \dots, 2, 1$$

が異なる値に推定され、重大な劣化が後のフレームまで伝搬するおそれがある。そこでこのような重大な劣化のおそれのある場合には、 $f_i^*(k)$ と $f_i^*(k-m)$ の値を用いて間の値を補間によって推定することをやめ、これから伝送すべきフレームの値のみを用いて量

*形補間する際に、m フレーム前のベクトル量子化値と線形補間するのではなく、m フレーム前のマトリクス量子化値を加えた値と現在のフレームのベクトル量子化値とを線形補間すれば、符号帳作成手順は多少複雑になるが、補間がより正確に行われマトリクス量子化の効率がよい。

【 0 0 2 5 】また、補間部 8 および補間部 1 6 において、 $f_i^*(k)$ と $f_i^*(k-m)$ とから線形補間によって $f_i^*(k-j), j = m-1, m-2, \dots, 1$ を推定すると推定が十分でない場合がある。そこで $f_i^*(k)$ とマトリクス量子化値とを加えて既に量子化された $f_i(k-m), f_i(k-(m+1))$ の 3 点から二次曲線で補間してもよいし、更に前の値を使って最小二乗法で推定してもよい。

【 0 0 2 6 】前記実施例では、前段量子化値として $f_i(k)$ と $f_i^*(k)$ との歪が最小になるように $f_i^*(k)$ を決定した後、補間によって

を推定しているが、補間する事を前提として前段の量子化値を決定してもよい。即ち、

$$f_i(k-j), j = m-1, m-2, \dots, 1, 0$$

と

$$f_i^*(k-j), j = m-1, m-2, \dots, 1, 0$$

当に定める。例えば重み付き L S P ユークリッド距離ならば、

$$d = \sum \sum [w_{i,j} \{f_i(k-j) - f_i^*(k-j)\}^2]$$

$$w_{i,j} = 2.0 / [\{f_{i+1}(k-j) - f_i(k-j)\} + \{f_i(k-j) - f_{i-1}(k-j)\}]$$

して、純粋に補間されたフレーム (m - 1 フレーム分)

における誤差のみをマトリクス量子化してもよい (マトリクス量子化の負担が軽減される) 。

【 0 0 2 8 】この発明を例えばデジタル自動車電話など非常に品質の悪い通信路に応用する場合を考える。上記通信路では伝送符号誤りが頻繁に生じ、再生された音声に重大な劣化を招く。前記実施例では補間部 8 及び 1 6 において、過去に伝送した $f_i^*(k-m)$ を用いて

*とで推定値

量子化した方がよい。図 1 における線形フィルタ係数分析部 1 2 の変形例を図 7 に示す。図 1 と対応する部分に同一符号を付けてある。バッファ部 4 より得られた m フレーム分の線スペクトル対の値、

$$f_i(k-j), i=1, 2, 3, \dots, p$$

$$j=m-1, m-2, \dots, 1, 0$$

は平均 L S P 算出部 4 6 に供給され、j についての平均*

$$f_i^{(ave)}, i=1, 2, \dots, p$$

を得る。これを前段量子化部 5 と同様に例えばベクトル量子化する。上記量子化値の第 u 番目の候補を $f_i^{(ave)*}(u)$ とし

$$f_i^*(k-j), j=m-1, m-2, \dots, 1$$

と $f_i^*(k)$ をマトリクス量子化部 9 に供給したが、ここでは

$$f_i^*(k-j) = f_i^{(ave)*}(u)$$

$$j=m-1, m-2, \dots, 1, 0$$

$$f_i^*(k-j), j=m-1, m-2, \dots, 1, 0$$

を行列加算器 1 7 に入力する代わりに

$$f_i^*(k-j) = f_i^{(ave)*}(u)$$

$$j=m-1, m-2, \dots, 1, 0$$

として入力する。図 7 に示す方法の場合、符号誤りのないところでは図 1 に比べて性能が劣化するが、ある時刻における符号誤りによる劣化が後のフレームに伝搬することはない。

$$f_i^*(k-j), j=m-1, m-2, \dots, 1, 0$$

の推定値として図 7 と同様にしてマトリクス量子化部 9 に供給する。この構成では図 7 に示した構成よりも線スペクトル対の平均を求めない分、ハードウェア的に簡単にできるうえ、図 1 に示す前段量子化部 5 と図 8 に示す前段量子化部 5 とを同一にできるメリットがある。

【0030】ここまではこの発明の音声符号化への適用例を示したが、この他に音声合成、音声認識、話者認識等の音声の音響的特徴を表現する一手法としても応用できることは明らかである。

【0031】

【発明の効果】この発明による線形予測パラメータ符号化方法の効果調べるために、以下の条件で音声分析合成実験を行った。0.3~3.4kHz 帯域(電話帯域)の音声を標準化周波数 8kHz で標準化した後、入力端子 1 に入力する。線形予測分析部 2 では、音声信号に分析窓長 32ms のハミング窓を乗じ、分析次数を $p=10$ 〔次〕として自己相関法による線形予測分析を行う。フレーム更新周期は 10ms、即ち $N=80$ 〔サンプル〕、前段量子化の周期とマトリクス量子化の単位とは 40ms、即ち $m=4$ 〔フレーム〕とした。前段量子化には、ディレイドディシジョン付き多段ベクトル量子化を用い、ディレイドディシジョンの候補数は 4 つとした。段数は 3 段で、1、2 段目に 8 ビット、3 段目には 6 ビットを割当て、前段量子化のための合計ビット数は 22 ビットとした。ただし、前段量子化のビット数を 22 ビット以外にする場合には、3 段目を可変にし、24

*された線スペクトル対
くつかを量子化候補保持部 6 に保持して、マトリクス量子化部 9 に供給する。このとき図 1 では補間された値
のように各フレームを同じ値として供給する。再生側でも同様に補間部 1 6 の出力

【0029】図 7 に示す例の更に簡単な構成の一例を図 8 に示す。図 7 では平均 L S P 算出部 4 6 によってこれから量子化すべき m フレーム分の平均 L S P を求め、前段量子化部 4 7 で例えばベクトル量子化しているが、図 8 ではこれから量子化すべき m フレーム分の線スペクトル対のうち、あるフレーム j の線スペクトル対のみ前段量子化部 5 で量子化する。この量子化値(ベクトル)を

ビットを越える場合には 1 段増やして 4 段とし、3 段目を 8 ビットにして 4 段目を可変とした。

【0032】前段量子化及びマトリクス量子化における歪判定部に用いる距離尺度には、重み付き L S P ユークリッド距離とケプストラム距離を用い、各性能を比較した。性能そのものは後者の方が優れていることがわかったが、後者は L S P をいったんケプストラムに変換してユークリッド距離をとらなければならない分計算量が多く、性能にあまり大きな差がなければ前者の方がハードウェア的な面で優れている。

【0033】前段量子化であるベクトル量子化と、マトリクス量子化のための各符号帳は、重み付き L S P ユークリッド距離を用いた L B G アルゴリズムにより別途作成した。なお L B G アルゴリズムについては、文献 Y . Linde , A . Buzo and R . M . Gray : " An Algorithm for Vector Quantizer Design " , IEEE Trans . Comm . COM - 28 , p . p . 84 - 95 1980 に記載されている。

【0034】この発明の効果は、再生された音声の品質によって評価されるべきであるが、この発明による線形予測パラメータの量子化性能が採用した音源パラメータの表現方式(性質)によって左右されることを防ぎ、純粋にスペクトル再現性の効果を評価するため、今回の実験では、評価尺度としてケプストラム歪尺度を用いた。この尺度は、スペクトルの歪具合をデシベル単位で表し

たものに相当し、ケプストラム歪の値が小さいほど良好な性能を有することを示す。

【0035】図9は従来方法である前段量子化のみ(マトリクス量子化なし)で、前段量子化の周期を10ms(1フレーム)から60ms(6フレーム)まで変えたときのケプストラム歪の変化を示したものである。この図からわかるように、前段量子化の周期を広げるほど伝送または蓄積のための情報を削減することができるが、逆に歪が増大していくことがわかる。

【0036】図10は、この発明を用いた結果で、前段量子化は前記22ビットのまま、この他に、マトリクス量子化に6~10ビットを割り当てた場合の結果である。図9と図10とより、マトリクス量子化MQに6ビット程度割り当てただけでも歪が1.8dB以下となり、30ms(3フレーム)に1回前段の量子化のみをする場合の歪1.95dBよりもはるかに優れ、マトリクス量子化に10ビット割り当てれば歪が1.6dB程度に小さくなり、20ms(2フレーム)に1回前段の量子化のみをするのに匹敵する性能が得られることがわかる。即ち、この発明を用いれば、22ビットで20ms毎に多段ベクトル量子化をするのに比べて、同程度の品質が40ミリで(22+10)ビットでよく、従来方式に対し44分の32の情報量(ビットレート)で実現でき、28%の情報量削減になる。

【0037】図11は、線形予測パラメータの量子化のために合計30ビットが割り当てられたと仮定した場合に、前段量子化(ベクトル量子化)と後段量子化(マトリクス量子化)とに割り当てるビットの割合を変えて、歪を求めたものである。全ビットを前段量子化に割り当てた場合に比べて、いくらかをマトリクス量子化に割り当てる方がはるかに良い結果が得られ、この発明の有効性が示される。図より、マトリクス量子化に割り当てるビット数を増やすと歪が徐々に減少するが、8~10ビットのあたりでは減少率は小さくなる。マトリクス量子化にあまり多量のビットを割り当てることは、計算量の

増大につながり、今回の実験は10ビットまでとしたが、図11の結果よりそれほど多量のビットをマトリクス量子化に割当なくとも、8~10ビット程度を割り当てるだけで、十分高性能が維持できることが推測される。

【0038】この発明の効果は、線形予測パラメータを量子化する場合に、ベクトル量子化などの従来方法を一定周期毎に適用する方法に比べて、それほど計算量を増大させることなく、同一情報量の場合にはより高品質な量子化結果を与える。また、同程度の品質を実現するためにはより少ない情報量で実現することができる。

【図面の簡単な説明】

【図1】この発明による線形予測パラメータ符号化法の一例を示す構成図。

【図2】図1に示した符号化法に対する再生側の復号化法の構成を示すブロック図。

【図3】前段量子化法としての多段ベクトル量子化法を示すブロック図。

【図4】前段量子化法としてのディレイドディシジョン付き多段ベクトル量子化法の構成例を示すブロック図。

【図5】マトリクス量子化法の一例を示す構成図。

【図6】この発明によって線スペクトル対が量子化される過程を模式的に説明するための図。

【図7】補間をしない変形例を示す構成図。

【図8】図7より更に簡単にした変形例を示す構成図。

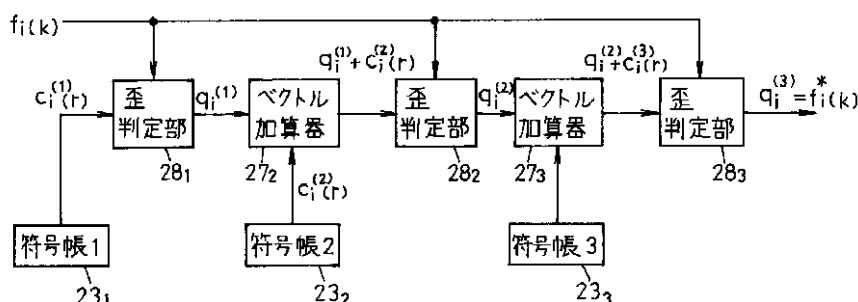
【図9】従来法の範囲内で情報量を削減しようとしたときの品質の劣化を調べた実験結果を示すグラフ。

【図10】この発明を適用したときの実験結果のグラフで、前段量子化のビット数を22ビットに固定し、マトリクス量子化のビット数を可変にした図。

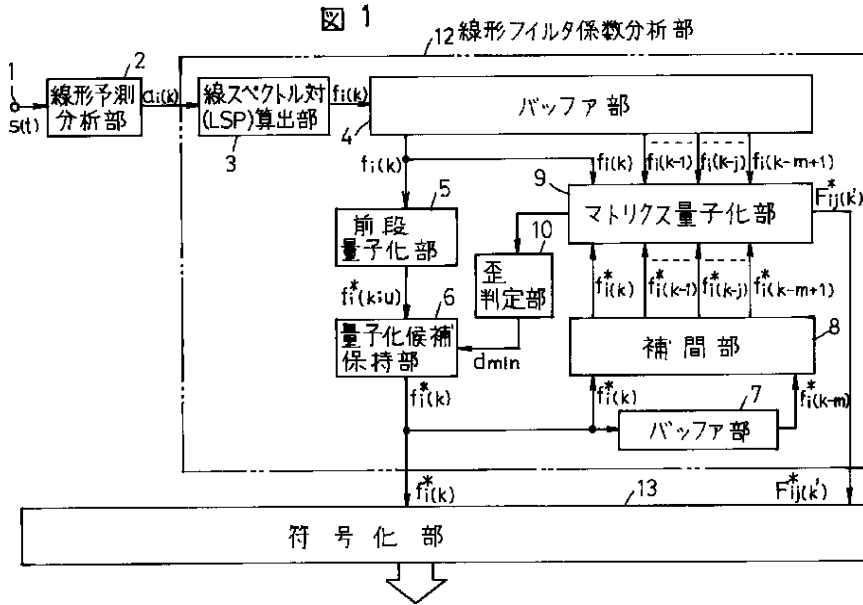
【図11】この発明を適用したときの実験結果のグラフで、全ビットを30ビットに固定した場合に前段量子化とマトリクス量子化とに割り当てるビット数の割合を変えたときの歪の変化を示した図。

【図3】

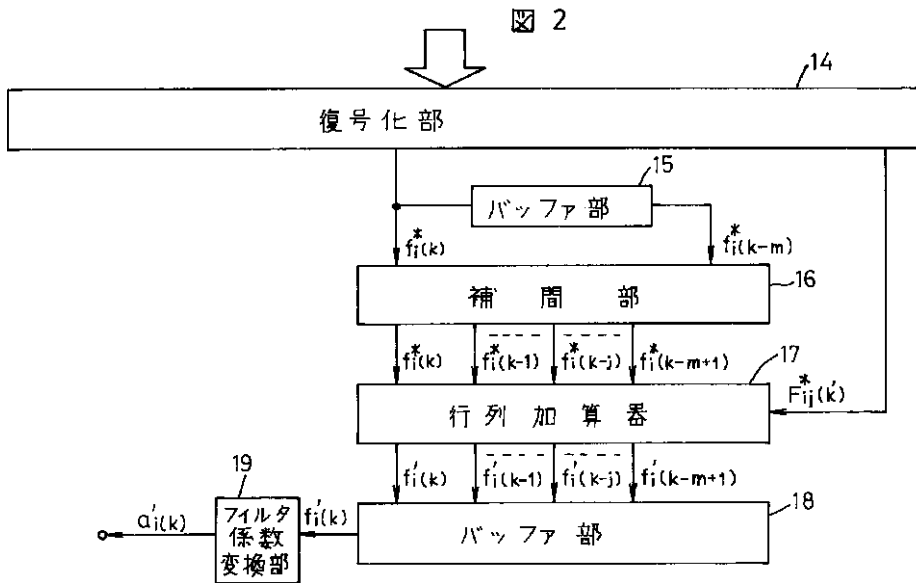
図 3



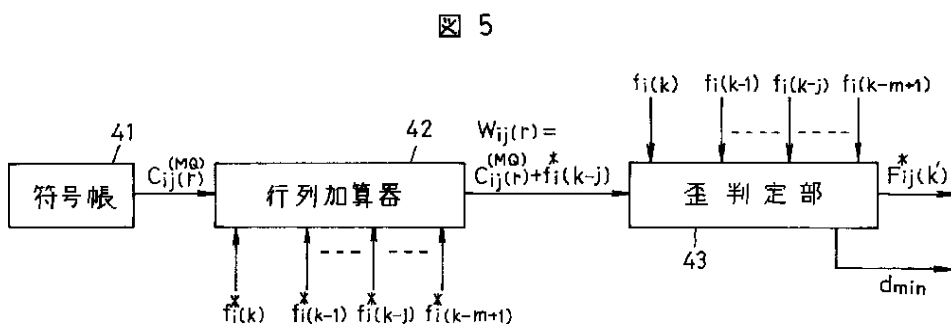
【図 1】



【図 2】

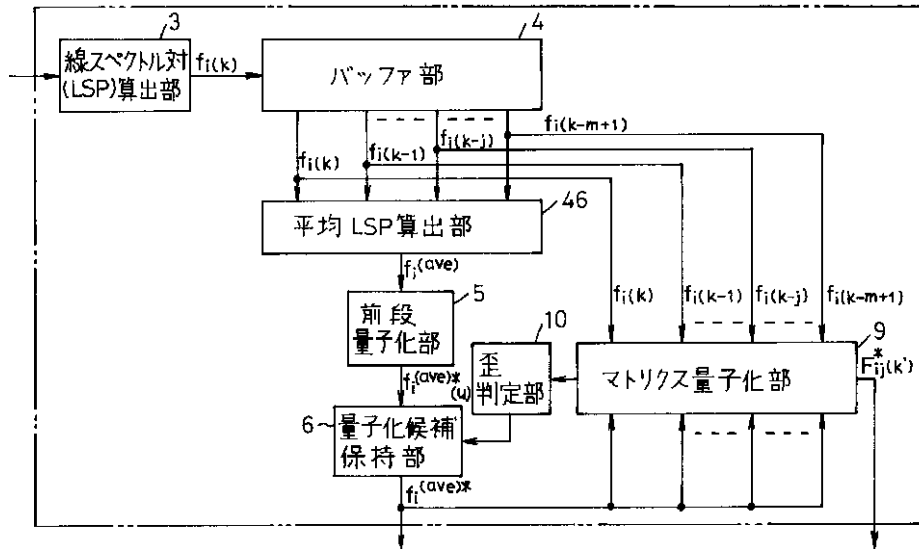


【図 5】



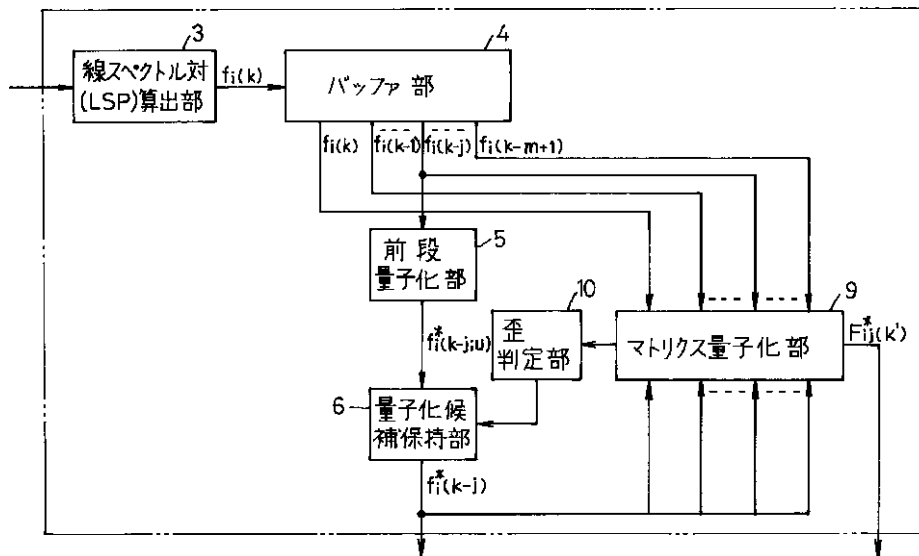
【図 7】

図 7



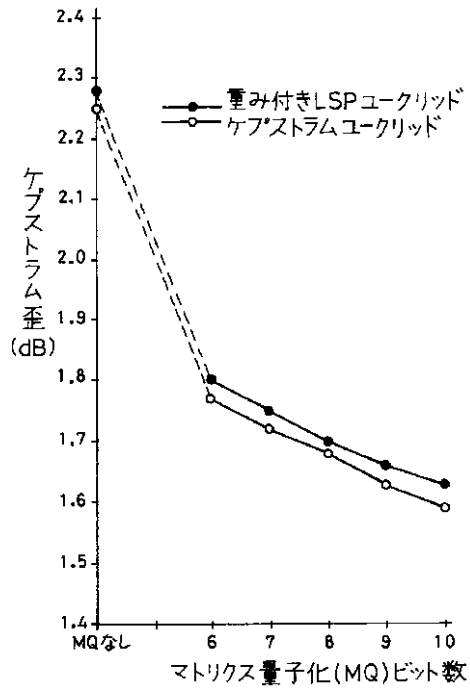
【図 8】

図 8



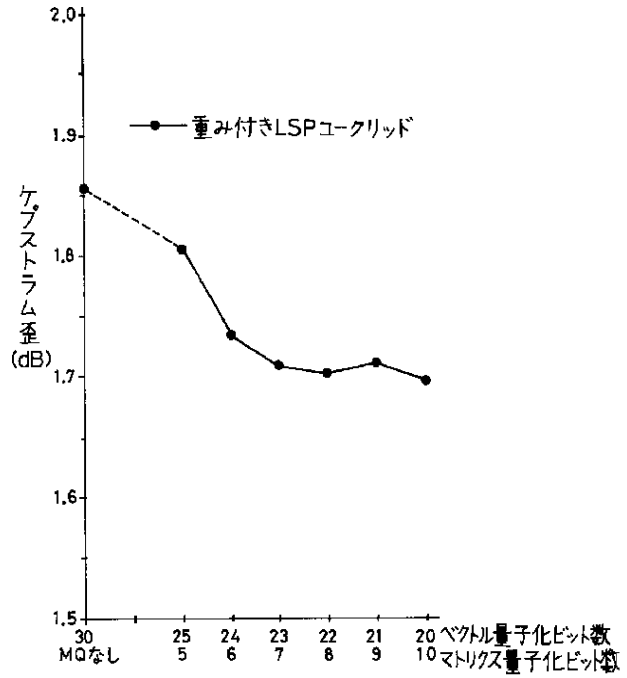
【図10】

図 10



【図11】

図 11



フロントページの続き

- (56)参考文献 特開 平1 - 215141 (JP, A)
- 特開 平1 - 240031 (JP, A)

(58)調査した分野(Int.Cl.⁷, DB名)

- G10L 19/00 - 19/14
- H03M 7/30
- H04B 14/04