

(19)日本国特許庁(JP)

(12)特許公報 ( B 2 )

(11)特許番号

第 2 7 6 1 5 5 3 号

(45)発行日 平成10年(1998)6月4日

(24)登録日 平成10年(1998)3月27日

(51) Int. Cl. <sup>6</sup>	識別記号	F I
G 1 0 L	9/14	J
H 0 3 M	7/30	Z
H 0 4 B	14/04	Z

請求項の数 1

(全 4 頁)

(21)出願番号	特願昭63-229624	(73)特許権者	999999999 日本電信電話株式会社 東京都新宿区西新宿3丁目19番2号
(22)出願日	昭和63年(1988)9月12日	(72)発明者	間野 一則 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
(65)公開番号	特開平2-75000	(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
(43)公開日	平成2年(1990)3月14日	(74)代理人	弁理士 草野 卓
審査請求日	平成6年(1994)9月20日	審査官	千葉 輝久
		(56)参考文献	特開 昭62 - 91024 ( J P , A )

(54)【発明の名称】音声符号化法

1

(57)【特許請求の範囲】

【請求項 1】音声信号を、一定区間毎に近接予測フィルタパラメータの符号と、長期予測フィルタパラメータの符号と、近接予測と長期予測後の残差を示す量子化符号とによって表現する音声符号化法において、各区間で各符号を一意に決定せず、それまでに残されている各符号組に対し、量子化歪が小さい順に各符号の組合せの候補をM組（Mは2以上の整数）選び、これら選ばれた各候補の量子化歪をその1区間前の対応する符号組の累積量子化歪に加算し、この加算した量子化歪が小さい順にM組の候補を残し、L区間（Lは2以上の整数）毎に、現区間での残されたM組の候補中の累積歪が最小となった符号組のL区間前の符号の組をその区間の符号化出力と決定することを特徴とする音声符号化法。

2

【発明の詳細な説明】

「産業上の利用分野」

この発明は、音声符号化法のうち9600bit/s以下の量子化歪最小化規準による低ビットレート波形符号化法に関するものである。

「従来の技術」

音声符号化法のうち、従来の9600bit/s以下の量子化歪最小化規準による低ビットレート波形符号化法の技術として、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 rates", IEEE Proc. of Inter'l. Conf. Acoust., Speech & Signal Process. 85, pp. 937 - 940 (1985), において示されている符号駆動型線形予測符号化（CELP符号化）が知られている。この符号化法は音声信号を近接予測フィルタパラメータ

と長期予測（またはピッチ予測）フィルタパラメータと、それらの予測後の主情報である残差波形のブロックに対応する雑音系列符号とによって音声を表現するものである。しかし、この方法は一分析区間毎の閉じた範囲内での量子化歪が最小となる符号化を行う手法であり、隣接または複数区間での最適性を保証するものではない。すなわち、近接予測フィルタや長期予測フィルタといった予測フィルタを使用している場合には一区間での量子化、符号化の結果が次の区間の予測フィルタにおいて利用されることになるために、一区間でのみ最適な符号であっても複数区間にわたって量子化歪を十分小さくするような符号化とはならない。また、かえって次の区間での歪最小化過程に悪影響を及ぼす可能性があるという欠点があった。

一方、J.B.Anderson,J.B.Bodielによる文献、(文献2)“Tree encoding of speech”,IEEE Trans.Information Theory,Vol.IT-21,pp.379-387(1975),では音声信号1点ではなく信号の複数点にわたる範囲で波形歪を最小とする符号選択法が示されている。この手法は1点毎に量子化する場合よりもかなり量子化歪を軽減できる。しかし、音声波形のサンプル単位の処理であるためサンプルあたり1bit以下の低いビットレート波形符号化の分野には直接応用することはできない。

この発明の目的は、新たな補助情報を必要とせずに従来できなかった複数区間全体の量子化歪の総和を小さくするという問題を解決した低ビットレート音声波形符号化法を提供することにある。

「課題を解決するための手段」

この発明は、音声信号を一定分析区間として区別し、各分析区間で予測フィルタのパラメータを示す符号とフィルタを通したその区間の残差信号系列を示す符号とを少なくとも一方について複数個の候補を準備し、複数分析区間全体の予測フィルタパラメータ符号及び残差信号の符号の組合わせで量子化歪が最小となる組合わせを求めることを最も主要な特徴とする。従来の技術では、区間毎の量子化歪が小さくなるようにフィルタのパラメータ符号と残差の符号を決定したのに対して、複数区間での量子化歪の総和が最小となる符号探索を行うという点異なる。また、音声信号の量子化符号を複数サンプル単位で決定する別の従来の方法と異なり、この発明の符号化法は音声分析区間毎のパラメータ等の符号の決定にまで拡張されており、より低ビットレート符号化に適合した方法となっている。

「実施例」

ここでは一定区間内の基本的符号化法として符号駆動型線形予測符号化を取り上げてこの発明の具体的音声符号化法を説明する。伝送するパラメータは近接予測の符号、長期予測の符号、予測残差に対応する符号に大別できる。

第1図は、この発明の複数区間での与えられた符号候

補により総量子化歪の最小の符号決定の実施例を説明するための樹木図である。区間0...区間3は、音声信号の分析区間を示し、 $N_0, N_{10}, N_{11}, \dots, N_{311}$ は音声区間毎に階層化された樹木図の節点があり、各枝は、各区間の節点が前の区間の節点との接続を示し、0, 10, 11, ... 311は各節点におけるフィルタと符号の組合わせ候補による量子化歪であり、0, 10, 11, ... 311は各節点までの累積量子化歪であり、それぞれの下の数値はとの一例である。

この例では、区間0の符号がすでに決定されているとし、区間1の符号は、区間1より2区間あとの区間まで、すなわち、区間1から区間3までの3区間の全体の量子化歪を最小にする組を選択する。このためには、例えば(文献2)に示されているような、サンプル単位での木符号化法を新しく音声分析区間毎の処理に適用できるように区間単位の木符号化法に変更することにより実現できる。つまり、第1図に示す各節点には、残差信号の各サンプル点に対応するのではなく、音声信号の近接予測フィルタパラメータの符号と長期予測フィルタパラメータの符号とその区間内の主情報である残差信号系列を表現する符号とのうち少なくとも一つを対応させる。

いま、区間0の0, 0はともに0に設定する。区間0の節点の状態を初期状態として、区間1の音声信号について近接予測フィルタパラメータの符号化と長期予測フィルタパラメータの符号化と残差信号の量子化符号化とを行う。この区間毎の波形符号化法としては、例えば(文献1)に述べられているような方法を適用する。しかし、このとき各符号を区間1のみの量子化歪の小さい符号の組を一意にその区間の符号として決定せずに、フィルタパラメータと残差信号の符号との少なくとも一つに対応する符号に関して量子化歪の小さくなるものから順にM個(第1図では $M=2$ の場合に相当している)の符号候補を残しておく。この組は、残差信号の符号の候補だけでなく、近接予測フィルタのパラメータ符号、長期予測フィルタのパラメータ符号のいずれかが異なれば、次の区間での初期入力が変わってくるのでこの組合わせを変えて、量子化歪の小さくなるものをM個選ぶ。こうして選ばれた各組の区間1での量子化歪を10, 11とし、前区間の対応する節点 $N_{10}$ と $N_{11}$ からの累積量子化歪を10, 11とする。この場合、 $10 = 0 + 10 = 0 + 1 = 1$ ,  $11 = 0 + 11 = 0 + 0.5 = 0.5$ である。区間1での2組の符号化の結果は、次の区間2での符号化の近接予測フィルタと長期予測フィルタの初期入力として使用される。

区間2では、 $N_{10}, N_{11}$ のそれぞれに対して、まず、M個の符号の組の候補を残しておく。次にこの $M \times M$ 個の候補から、区間1と同様に区間2までの累積量子化歪を求め、それが小さくなる節点をM個残し(節点 $N_{200}$ と $N_{210}$ が対応)、その他の節点 $N_{201}$ と $N_{211}$ は候補からはずす。そして、残ったM個の節点情報を次の区間3での初

期入力として利用し、区間 2 と同様に各節点から M 個の候補 (N300, N301, N310, N311) を残す。

こうして、区間 1 からあらかじめ設定しておいた L 区間あとの (第 1 図で L = 2 の場合に相当する) 区間まで到達したら、そこまでの累積量子化歪の最も小さくなる節点を選ぶ。ここでは、300, 301, 310, 311 では 300 が最小なので節点 N300 が選ばれる。従って、区間 1 での符号の組は、節点 N300 に接続している節点 N10 に属する符号となる。以下同様にして、区間 2 以降についても L 区間先までの累積量子化歪が小さくなる節点を選択していけばよい。

この実施例でわかるように区間 1 での量子化歪だけでは、N11 の方が N10 よりも小さいが、区間 3 までの累積量子化歪では、N10 の符号の組を選択した方が歪は小さくなる。この発明では、このような作用があるので従来の区間だけの符号化法に比べて音声信号系列全体の量子化歪を小さくする効果がある。

上記説明では簡単のため複数種類の符号の更新周期は同一としたが、歪を最小化する意味で更新周期はパラメータ毎に変えた方がよい。この場合でも、それぞれのパラメータについてその更新周期を単位として上記符号化法を適用することができる。

「発明の効果」

第 2 図は、実際の音声における低ビット波形符号化として従来法と本発明法の比較をしたものである。この場合、近接予測と長期予測の残差の符号についてのみに本発明を適用した。総ビットレートを 4.8 kbit/s 固定として、横軸 L は、対象区間からの累積区間の区間数 (すなわち L 区間後までの累積を考える) を示す。一区間は 5 [ms] であり、信号を 8 [KHz サンプリングとすれば 40

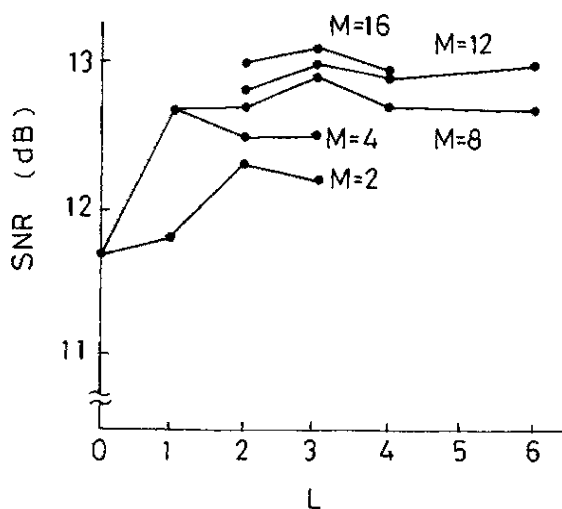
点である。縦軸は、原音声と符号化再生音声との信号対雑音比 (SNR) を示す。M は、各区間で残す候補の数である。第 2 図で L = 0 の場合 (つまり対象区間のみの量子化歪で符号を決定する従来法の場合) の SNR は 11.7 dB であったに対して、L が 1 以上で SNR の改善があった。特に M = 16, L = 3 では、SNR = 13.1 dB であり、従来法に比べて 1.4 dB の性能向上が得られた。

以上説明したように、本発明は、新たな補助情報を必要とせず音声波形符号化の性能を改善することができ、特に区間毎のパラメータと残差系列を表現する符号の組にして符号化を行うので、サンプル単位の処理ではなくブロック単位の処理形式になっており、低ビットレートでの符号化法として適している。特に長期予測係数、つまり音声波形のピッチ周期性を捉える際に抽出誤りが生じやすく、フレーム毎に 2 倍周期あるいは半周期のピッチに基づく長期予測係数を符号化してしまい、音質劣化が大きくなることがある。この劣化は、フレーム内に留まらず、長期予測フィルタあるいは適応符号長によって、ピッチ周期性のおよぶ後続フレームにも重大な悪影響を与える。しかし、この発明では、長期予測係数も、線形予測係数、残差信号と一緒に複数候補を残し、複数フレームに渡る累積波形歪みを計算し、3 パラメータの最適な符号の組を求めることにより、フレーム間で正しい連続的なピッチ周期性を与える長期予測係数を符号値とすることができる。

【図面の簡単な説明】

第 1 図はこの発明の符号化に用いる複数区間にわたる符号選択手順の動作原理を示す図、第 2 図はこの発明の一例における性能向上を示す図である。

【第 2 図】



【第 1 図】

