

(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 0 9 2 4 8 号

(P 3 2 0 9 2 4 8)

(45)発行日 平成13年9月17日(2001.9.17)

(24)登録日 平成13年7月13日(2001.7.13)

(51)Int. Cl. ⁷	識別記号	F I
G 1 0 L	19/12	G 1 0 L 9/14 S
	19/00	9/18 E

請求項の数 1

(全 7 頁)

(21)出願番号	特願平5-165582	(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5年7月5日(1993.7.5)	(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
(65)公開番号	特開平7-20896	(72)発明者	片岡 章俊 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7年1月24日(1995.1.24)	(74)代理人	100066153 弁理士 草野 卓
審査請求日	平成11年11月10日(1999.11.10)	審査官	渡邊 聡
		(56)参考文献	特開 平5 - 100697 (J P , A) 特開 平1 - 319799 (J P , A)

最終頁に続く

(54)【発明の名称】音声の励振信号符号化法

1

(57)【特許請求の範囲】

【請求項 1】 ピッチ周期成分ベクトルと雑音励振ベクトルを線形予測合成フィルタの励振源とする音声合成モデルを持ち、合成音声と入力音声の誤差を最小化するように雑音励振ベクトルを符号帳のなかから選択する音声の励振信号符号化法において、雑音励振ベクトル毎にその合成後の波形と入力音声波形の内積（第 1 の内積値）を求め、合成フィルタのインパルス応答の各係数の 2 項積またはその線形和をパラメータベクトルとし、予め、雑音励振ベクトル毎に作られた回帰係数ベクトルと前記パラメータベクトルの内積計算（第 2 の内積値）で雑音励振ベクトルの合成後のエネルギー又はその逆数を推定し、前記第 1 の内積値を前記第 2 の内積値でわった値を基準

2

に雑音励振ベクトルの選択または予備選択を行なうことを特徴とする音声の励振信号符号化法。

【発明の詳細な説明】

【 0 0 0 1 】

【産業上の利用分野】この発明は音声信号をできるだけ少ない情報量でデジタル符号化する高能率音声符号化法で、特に線形予測合成フィルタに供給する励振信号を決定する励振信号符号化方法に関するものである。

【 0 0 0 2 】

10 【従来の技術】 8 k b i t / s 程度以下の音声符号化法として、C E L P (C o d e E x c i t e d L i n e a r P r e d i c t i o n : 符号励振線形予測) 符号化が良く知られている。この符号化法は図 5 に示すように、フレーム単位で雑音符号帳 1 1 と適応符号帳 1 2 から励振ベクトルを選択して、それぞれ利得調整手段 1

3, 14で利得を調整して線形予測合成フィルタ15に入力して音声を合成する。この合成音声と入力音声との差を差回路16で求め、その差出力を聴感補正フィルタ17を通し、そのフィルタ出力から歪計算部18で歪を計算し、その歪が小さくなるように励振信号を選択し利得調整する。励振ベクトルの選択基準は合成された信号と入力信号の聴感上の誤差を最小化することである。このように最終的に合成される波形をフィードバックして励振ベクトルを決定するため、高い品質が得られる。

【0003】この半面、大量の雑音励振ベクトルを記憶しておく必要があることや、励振ベクトルの探索の演算量が膨大となるという難点がある。この為記憶容量や演算量を削減するためにいくつかの手法が考えられているが、以下で探索の基本原則を示し、この発明と関連のある代表的な演算量削減法について整理して紹介する。*

$$d = \| X - gHC^j \|^2 \tag{1}$$

$$H = \begin{bmatrix} h_0 & 0 & \dots & \dots & 0 \\ h_1 & h_0 & 0 & \dots & 0 \\ h_2 & h_1 & h_0 & \dots & 0 \\ \dots & \dots & \dots & \dots & 0 \\ h_{n-1} & h_{n-2} & h_{n-3} & \dots & h_0 \end{bmatrix} \tag{2}$$

ここで、最適な利得が与えられ、雑音励振ベクトル決定が最適な利得となり (X^T は X の転置行列)、このときとなる。結局第2項の

$$g = X^T HC^j / ((HC^j)^T (HC^j)) \tag{3}$$

$$d = X^T X - (X^T HC^j)^2 / ((HC^j)^T (HC^j)) \tag{4}$$

$$f = (X^T HC^j)^2 / ((HC^j)^T (HC^j)) \tag{5}$$

を最大とする j を捜せばよいことになる。ここで HC^j の合成演算には $n^2 / 2$ 程度の積和演算が必要で、 j について C^j を合成するには莫大な演算量を必要とする。そこで種々の提案がなされている。

【0006】まず式(5)の分子は $X^T H$ を1回だけ先に計算し、そのあとで C^j と内積をとれば近似誤差を発生させずに演算量を削減する手法が知られている。内積は n 回の演算ですむためである。ところが、分母の計算で演算量を削減しようとすると、近似誤差が生ずる。例えば分母の項を $(C^j)^T H^T HC^j$ として、 $H^T H$ をテプリッツ型の相関行列で近似し、 H の自己相関関数と C^j の自己相関関数の内積で計算する方法が知られている。この方法については実施例のなかで詳しく紹介するが、これにより、分母の項を n 回の積和で計算できる。しかしながら、近似誤差が大きく、歪の増加と品質の劣化を招くことや、演算量の削減効果は必ずしも大きくない。また雑音励振ベクトル C^j 毎に相関値を記憶しておくために、符号帳のための記憶容量と同程度の記憶容量の追加が必要となる。

【0007】これとは別に C にローパスフィルタを掛けてダウンサンプルした励振ベクトルをもち、分母の項をこのベクトルで近似する手法も知られている。ダウンサ

*【0004】CELP符号化における雑音励振ベクトルの探索では、前述したように線形予測合成フィルタ15の出力合成波形と入力音声との差をとり、その差出力を聴感補正フィルタ17を通したあとの歪が最小となるベクトルを選択する。具体的には、あるフレーム(フレーム長は n サンプル)で量子化の目標とする n 次元の縦ベクトル(入力音声から、前のフレームからの応答成分や適応符号ベクトルからの合成成分を差し引いた現フレームの入力音声を、聴感補正フィルタに通した信号ベクトル)を X 、符号帳中の m 種類の雑音励振ベクトルの n 次元縦ベクトルを C^j ($j = 0, \dots, m - 1$)、聴感補正フィルタ17も一括した合成フィルタ15のインパルス応答行列 ($n \times n$) を H 、雑音励振ベクトルの利得を g とすると、下記の歪み d を最小とする j を探索する。

【0005】

後にその利得を量子化することとすると、

の歪は

サンプルした比率で演算量を削減できるが、この手法でも近似誤差やダウンサンプルしたベクトルのための記憶容量が増加するという問題がある。また探索を2段階にわけ、第1段では近似計算で複数の C^j を候補として予備選択し、その候補についてのみ第2段で式(5)を最大とするものを選択する方法も知られている。第1段の近似法としては分子の項だけを用いる方法や、上記で紹介した近似法を用いる方法などが考えられる。第1段の近似誤差を小さくすることや候補数を増やすことで歪を小さくできるが、演算量削減効果が小さくなってしまふ。

【0008】

【発明が解決しようとする課題】この発明の目的は音声を少ない情報量で符号化するとき、できるだけ演算量を抑えたまま符号化音声の品質を向上させ、特に上記従来法の中で紹介した2段階の探索と組合せ、歪の劣化が小さい割に大幅な演算量の削減ができる音声の励振信号符号化方法を提供することにある。

【0009】

【課題を解決するための手段】この発明では雑音励振ベクトルを選択する(または予備選択する)際に式(5)の分母の項、すなわちベクトルのエネルギーの項も考慮するが、そのエネルギーの項は、インパルス応答の係数

の 2 項積、またはその線形和をパラメータベクトルとし、このパラメータベクトルと、それに対応して予め蓄えておいた雑音励振ベクトル毎に作られた回帰係数ベクトルとの簡単な内積で求めることが特徴である。この回帰係数ベクトルは雑音励振ベクトル毎に平均推定誤差が小さくなるように設計し、予め蓄えておく。

【0010】図 1 に 2 段階の探索を行なう場合について、この発明方法と従来法との比較を示す。従来においては図 1 B に示すようにエネルギー項を全く無視し、式 (5) の分子の計算結果のみで予備選択するか、図 1 C に示すようにインパルス応答の自己相関関数と雑音励振ベクトルの自己相関関数を使った近似演算でエネルギー項 (式 (5) の分母の項) を求め、これを用いて式 (5) を近似計算して予備選択を行っているが、この発明では図 1 A に示すようにインパルス応答の 2 項積のまたその線形和と回帰係数とでエネルギー項を求めて予備選択、または本選択を行い演算量が少なく、精度の高い励振信号の符号化が可能となる。

【0011】

【実施例】図 2 にこの発明の最も基本的な実施例の要部の構成を示す。通常 m 個の n 次元雑音励振ベクトルからなる符号帳 11 のほかに、この実施例では雑音励振ベクトル毎に、回帰係数の表 21、インパルス応答系列の係数ごとの 2 項積の線形結合を演算してエネルギー推定パラメータを作る手段 22 が設けられ、この推定パラメ*

$$f_d = (C^T)^T H^T H C^T = h^T D^T D h \tag{6}$$

ただし

$$h^T = \{h_0, \dots, h_{n-1}\} \tag{7}$$

$$(C^T)^T = \{c_0, \dots, c_{n-1}\} \tag{8}$$

$$D = \begin{bmatrix} c_0 & 0 & \dots & \dots & 0 \\ c_1 & c_0 & 0 & \dots & 0 \\ c_2 & c_1 & c_0 & \dots & 0 \\ \dots & \dots & \dots & \dots & 0 \\ c_{n-1} & c_{n-2} & c_{n-3} & \dots & c_0 \end{bmatrix} \tag{9}$$

である。さらに を用いて書きかえる。

【0016】

* ータと表 21 中の回帰係数との内積でエネルギー値を求める。

【0012】2 段階の探索を行なうとしたときの探索手順は以下ようになる。

1. あるフレームで合成フィルタや聴感補正フィルタの特性が与えられると、合成フィルタのインパルス応答が決まる。このインパルス応答の係数ごとの 2 項積からエネルギー推定パラメータベクトルを作成する。

2. 次に式 (5) の分子の項を全ての雑音励振ベクトルに対して求める。

【0013】3. 各雑音励振ベクトルに対して、パラメータベクトルと、対応する回帰係数ベクトルとの内積の計算からエネルギー推定値、さらにそのエネルギー推定値を用いて (分子の項 / エネルギー推定値) を算出し、その算出値の大きいものから設定された数の候補を残す。

4. 予備選択した候補についてのみ式 (5) の計算を行ない、最適な雑音励振ベクトルの符号を決定する。

【0014】この処理で問題となるのが、どのようなパラメータベクトルを作るかという点とそのベクトルに対しどのように回帰係数ベクトルを用意するかという点である。まずパラメータベクトルの構成を示すため、式 (5) の分母を f_d とし、以下のように書きかえる。

【0015】

【数 1】

音励振ベクトルの探索の前に、すべての雑音励振ベクトルをピッチ成分と直交化させる手法や、複数チャンネルの雑音励振ベクトルをもつ場合にもこの発明を適用できる。

【0020】すなわち第2の実施例として雑音励振ベク

$$f = \frac{(X^T H C^j - H P)^2}{((H C^j)^T (H C^j) - (H C^j)^T (H P))} \quad (21)$$

ここで

$$= \frac{(H C^j)^T (H P)}{(H P)^T (H P)} \quad (22)$$

である。式(21)でも分子の項は式(5)の場合と同様に比較的少ない演算で正確に計算できる。また分母の

以下の第2項も積和の計算も同様である。従って分母の第1項を第1の実施例と同じ手順で表から読み出せば演算量を削減できる。また分母の第2項の計算を全く省略して、第1項のみで評価しても歪の増加はほとんどない。

【0021】第3の実施例として2チャンネルの雑音励

$$f = \frac{(X^T H C^j + X^T H C'^k)^2}{(H C^j)^T (H C^j) + 2(H C^j)^T (H C'^k) + (H C'^k)^T (H C'^k)} \quad (23)$$

しかしながら各チャンネル毎の予備選択ではそれぞれ、第1の実施例と同じ処理を行なう。すなわち分母のエネルギー項の概算値を表から読みだし、式(5)を最大とする候補を複数個選択する。そして、各チャンネルからの候補の組合せで式(23)が最大となるjとkの組合せを探索する。

【0023】上述において、パラメータベクトル β_j は1個の2項積で構成してもよい。また雑音励振ベクトルの実際のエネルギーの逆数値と、その推定値との差が小さくなるように回帰係数を求め記憶しておき、パラメータベクトルと推定値との内積でエネルギーの逆数値を推定してもよい。更に予備選択のみならず、場合によっては式(5)の正確な演算を行うことなく、前記予備選択を最終選択としてもよい。

【0024】

【発明の効果】この発明を用いた場合の入力音声と符号化歪とのSNRと、音声符号化処理全体の演算量(単位MOPS: Mega Operation Per Second: 実時間処理に必要な1秒あたり100万回単位の演算処理回数)との関係を従来法と比較したものを図4に示す。この場合、雑音励振ベクトルは2チャンネルからなり、各チャンネル毎に7ビット(極性を除く)を配分した場合である。雑音励振ベクトルの次元数nは40でピッチ成分との直交化処理を併用している。またエネルギー推定パラメータの個数は6とした。ここでの従来法Aは式(5)の分子の項のみで予備選択を行なう方法(図1Bに示した方法)であり、従来法Bは図

トルをピッチ成分(適応符号帳からの出力)n次元ベクトルPと直交化させる場合について説明する。この場合には最終的に最大化する式(5)が次のように変形される。

振ベクトルをもつ場合について示す。ここでも2段階の探索を行なうと図3に示されるような処理を行なうことになる。第1チャンネルの励振ベクトルを C^j 、第2チャンネルの励振ベクトルを C^k とすると、最終的な評価尺度は以下ようになる。

【0022】

【数2】

1Cに示したエネルギー項を近似計算して予備選択を行なう方法である。図中の数字は各チャンネル毎に128個の中から残す候補数である。この図からこの発明により従来法Aと比較して、残す候補数が同一であれば、演算量は少し増加するが、それ以上にSNRの改善効果があり、つまり同一SNRであればこの発明の方が演算量が著しく少なくなることがわかる。また従来法Bと比較して予備選択の候補数が同じ場合に同じようなSNRの改善効果があるが、従来法Bの演算量の増加が大きいため、この発明の方が明らかに有利であることがわかる。

【0025】従来法Aに対し、この発明で増加する演算量の内分けは、 β_j の計算がフレームあたり1回、 β_k の内積が各雑音励振ベクトル毎に1回ずつで、予備選択の個数を1個増やす程度の比較的少ない量である。またこの発明では β_j のための記憶容量が増加するが、例えば β_j が6次元で、40次元の雑音励振ベクトルを用いる時には、記憶容量の増加する割合は約15%である。

【図面の簡単な説明】

【図1】2段階の探索におけるこの発明の処理を従来法と対比させて示したブロック図。

【図2】この発明の実施例の要部を示すブロック図。

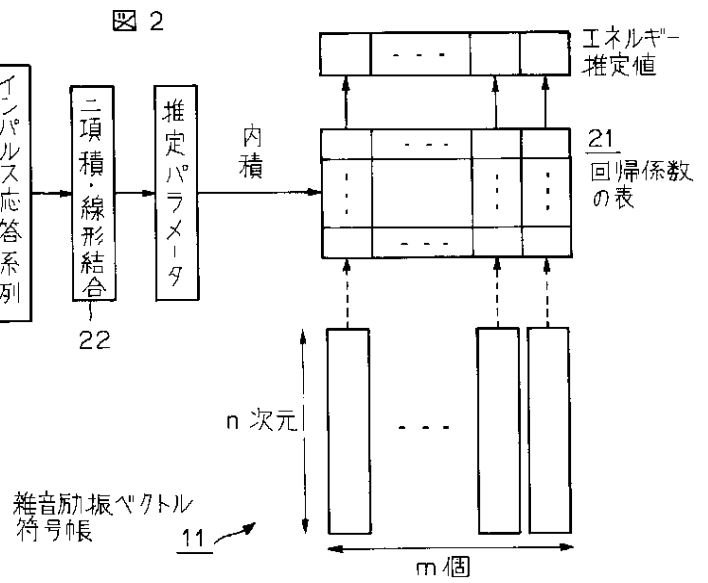
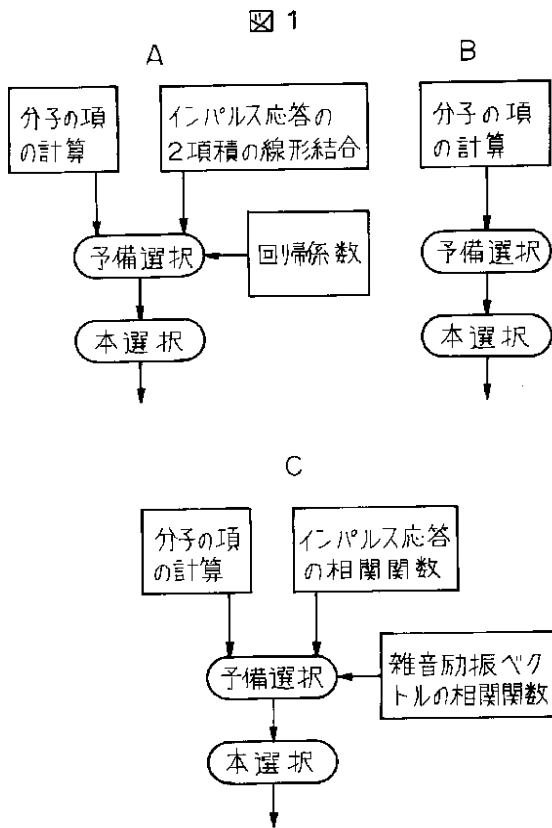
【図3】この発明を2チャンネルの雑音励振ベクトルの場合への適用した実施例の要部を示すブロック図。

【図4】この発明によるSNRと演算量の関係を従来法と比較した図。

【図5】CELP符号化法の基本原理を示すブロック図。

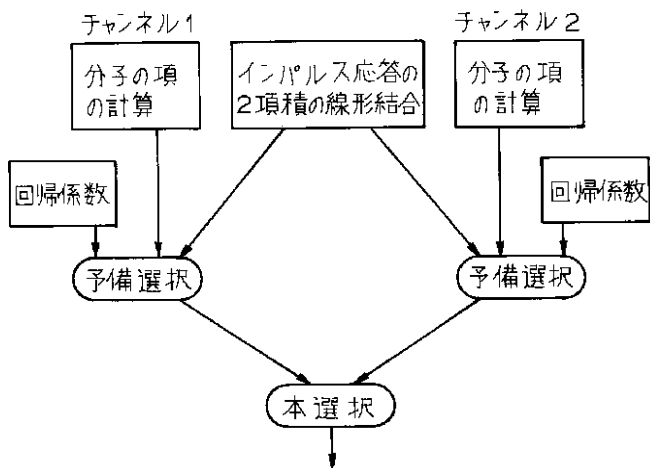
【図1】

【図2】

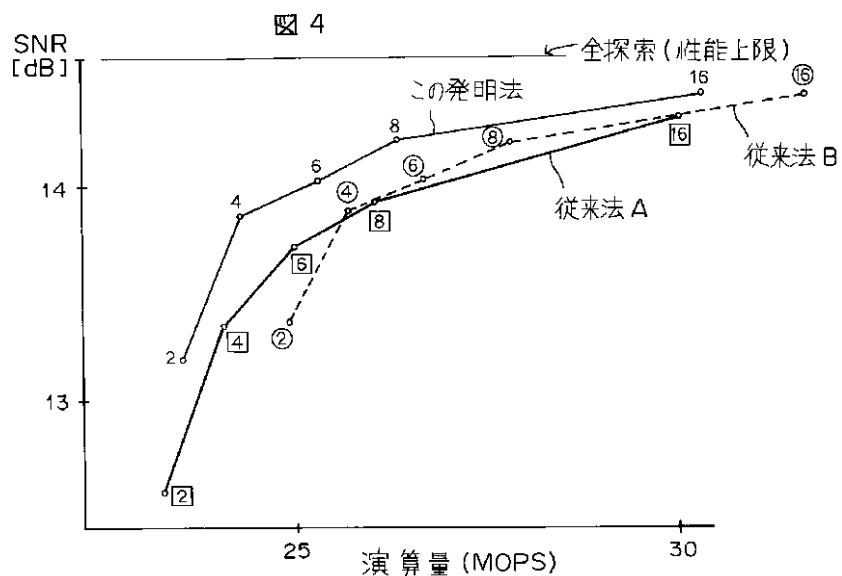


【図3】

図 3



【図 4】



【図 5】

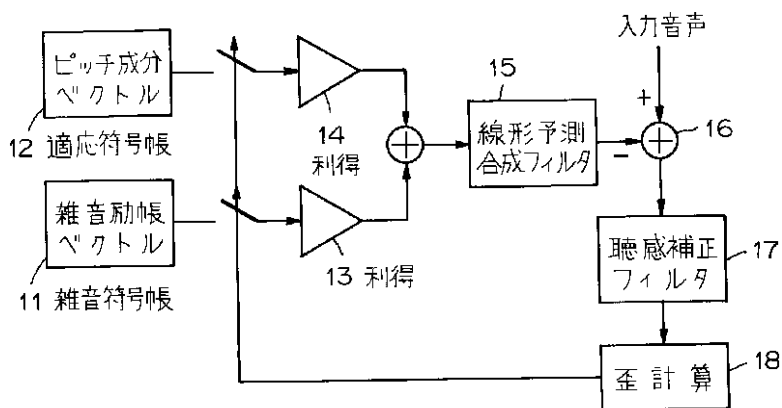


図 5

フロントページの続き

(58)調査した分野(Int.Cl.⁷, DB名)

G10L 19/12