

(19)日本国特許庁(JP)

(12)特許公報 (B 2)

(11)特許番号

特許第 3 3 5 3 2 5 2 号

(P 3 3 5 3 2 5 2)

(45)発行日 平成14年12月3日(2002.12.3)

(24)登録日 平成14年9月27日(2002.9.27)

(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			J
	19/04		9/18	E

請求項の数 2

(全 6 頁)

(21)出願番号	特願平3-258936	(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10月7日(1991.10.7)	(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5-100697	(72)発明者	三樹 聡 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4月23日(1993.4.23)	(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
審査請求日	平成10年2月3日(1998.2.3)	(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】音声符号化方法

1

(57)【特許請求の範囲】

【請求項 1】 入力音声信号をフレーム単位で線形予測分析して予測係数を求め、ピッチ周期成分をもつ適応符号帳とL個(Lは2以上の整数)の系統の乱数符号帳とからの各ベクトルの利得付き和を励振源として、先に求められた予測係数に基づいて合成された音声信号と入力音声信号との誤差を最小とするようなピッチ周期、乱数ベクトルの符号、利得を決定する音声符号化法において、
 第1段階で適応符号帳からのベクトルPを、ピッチ周期を決めることにより決定し、
 第2段階でベクトルXとベクトルU_iの内積dを求め、dの大きいものから複数の乱数ベクトルの候補を、各系統の乱数符号帳ごとに予備選択し、
ただしXは現在のフレームの入力を0として合成フィル

2

タを動作させたときの応答信号を現在の入力音声信号から引いた信号、U_iはi系統(i=0,1,...,L-1)の乱数符号帳のベクトルC_iを合成フィルタH(Hはインパルス応答行列、又はこれと等価なフィルタ操作の処理を示す)に通した信号を、Pをその合成フィルタHに通した信号に直交化させたベクトルであり、第3段階で上記予備選択された候補から歪が最小となる乱数ベクトルを各系統ごとに1個ずつ本選択し、
第4段階で第1段階で決定されたベクトル及び第3段階で選択されたL個のベクトルを用いて、合成音声信号と入力音声信号との誤差を最小とする利得を決定することを特徴とする音声符号化方法。

【請求項 2】 $X^T H, P^T H^T H, H P^2$ をあらかじめ求め、その計算結果とC_iの計算により上記内積dを求めることを特徴とする請求項1記載の音声符号化

方法。

【発明の詳細な説明】

【0001】

【産業上の利用分野】この発明は音声信号をフレーム単位で線形予測分析して予測係数を求め、ピッチ周期成分をもつ適応符号帳と乱数符号帳とからの各ベクトルの利得付き和を励振源として、先に求めた予測係数に基づいて合成された音声と入力音声との誤差を最小とするようなピッチ周期、乱数ベクトルの符号、利得を決定する音符号化法に関する。

【0002】

【従来の技術】現在、音声を高能率に符号化する方法として、原音声をフレームと呼ばれる5～50ms程度の一定間隔の区間に分割し、その1フレームの音声を周波数スペクトルの包絡形状と、その包絡形状に対応する線形フィルタを駆動するための駆動音源信号という2つの情報に分離し、それぞれを符号化することが提案されている。その場合、駆動音源信号を符号化する方法として、駆動音源信号を音声の基本周波数(ピッチ周期)に対応すると考えられる周期成分と、それ以外の成分(言い換えれば非周期成分)とに分離して符号化する方法が知られている。この駆動音源情報の符号化法として符号駆動線形予測符号化(Code-Excited Linear Prediction Coding: CELP)およびベクトル和駆動線形予測符号化(Vector Sum Excited Linear Prediction Coding: VSELP)法がある。それぞれの技術については、M.R.Schroeder and B.S.Atal: "Code-Excited Linear Prediction (CELP): Highquality Speech at Very Low Bit Rates", Proc. ICASSP'85, 25.1.1, pp.937-940, 1985、およびI.A.Gerson and M.A.Jasiuk: "Vector Sum Excited* 30

$$U_{ij} = HC_{ij} - (P^T H^T H C_{ij} H P) / H P^2 \quad (1)$$

上付添字Tは転置行列を示す。次に入力ベクトルX(前フレームから合成フィルタの応答分を除いたもの)とU

$$d = X - U_{ij} \quad (2)$$

はi=0からL-1までこの歪dを最小化するベクトル番号J(0)・・・J(L-1)を決める(S3)。

$$X - \{g_1 H P + g_2 H (C_{ij(0)})\}^2 \quad (3)$$

はi=0からL-1まで、ベクトルはすべてM次元としている。図4の右に各ステップにおける計算回数を示す。

【0007】

【発明が解決しようとする課題】VSELPの場合には割り当てられるビット数Bが例えば12の場合、これらの処理に関係するベクトルの総数Tも12であるため、現実的な演算量の範囲で処理が可能である。しかしながらCELPでは極性ビットを除いた11ビットに対応する符号帳ベクトルの個数Tは2¹¹個となり、膨大な演算量となり、実時間処理が困難となる。

【0008】この発明の目的は4kbit/s程度の低ビットで、できるだけ高品質の符号化音声をできるだけ

* Linear Prediction (VSELP)Speech Coding at 8 kbps", Proc. ICASSP'90, S9.3, pp.461-464, 1990、に述べられている。

【0003】これらの符号化では図2に示すように入力音声はフレームを単位として線形予測分析部11で線形予測分析されて予測係数が算出され、量子化されたのち、補助情報として伝送されると共にまた合成フィルタ12の係数として用いられる。励振源のうち適応符号帳13の出力ベクトルはピッチ周期を決めることで決定される。一方図2の場合、2系統(チャンネル)の乱数符号帳14, 15をもち、2つのベクトルの和で最終的な出力と入力音声の歪が最小となる組合せを符号帳から選択する。適応符号帳13からのベクトルと乱数符号帳14, 15からのベクトルはそれぞれに最適な利得(ゲイン)の組をゲイン符号帳16から探索して決定する。なお、図3に示したように乱数符号帳が1乃至複数系統であるのがCELPで、乱数符号帳の系統が割り当てられたビット数Bと同じであって、各乱数符号帳に唯一の基本ベクトルしか持たない場合がVSELPである。図2はこれらを一般化し、特に2系統の例を示している。

【0004】乱数符号帳の符号を決定するときには図4に示す処理を行なうことで、歪を小さくすることができる。まず適応符号帳13の出力ベクトルPを決定する(S1)。次に各乱数ベクトルC_{ij}(i=0,・・・L-1, j=0,・・・N-1, L:乱数符号帳の数、N:乱数符号帳あたりのベクトル数)を合成フィルタに通してHC_{ij}を作る(S2)。Hはインパルス応答行列である。次にHC_{ij}をHPと直交させてU_{ij}を作る。

【0005】

$$U_{ij} \text{ との歪 } d \text{ を求め、}$$

【0006】その後、次式を最小とする利得の和g₁, g₂を決定する(S4)。

$$X - \{g_1 H P + g_2 H (C_{ij(0)})\}^2 \quad (3)$$

少ない処理量で符号化する符号化法を提供することにある。

【0009】

【課題を解決するための手段】この発明によれば、複数の系統の乱数符号帳からの出力の和とピッチ周期の適応符号を励振源とするとき、ベクトルXとベクトルU_iとの内積dを求める。Xは現フレームの入力を0として合成フィルタを動作させたときの応答信号を現在の入力信号から引いた信号であり、U_iは乱数符号帳のベクトルC_iをフィルタ係数がHの合成フィルタ(Hはインパルス応答行列又はこれと等価なフィルタ操作の処理を表わす)を通し、更に適応符号帳からのベクトルPを合成フィルタに通したあとの信号に直交させたベクトルで

ある。得られた内積 d の大きいものから複数、例えば 1 つの乱数符号帳内の乱数ベクトルの数の 4 分の 1 の数の候補を予備選択し、その選択した候補の組合せだけから最終的に歪を最小とする乱数ベクトルの組を決める。このようにしてほぼ最適な乱数符号を非常に少ない演算量*

$$D = X - g U_{ij} \quad (4)$$

$$= X^T X - 2 g X^T U_{ij} + g^2 U_{ij}^2$$

ここで i = 0 から L - 1 まで、L は乱数符号帳の系統 (チャンネル) 数、M はベクトルの次元数、N は乱数符号帳の系統あたりのベクトル数である。利得 g は励振ベクトル決定後に量子化するので、g は任意の値をとることができる。(4) 式を g で偏微分した式をゼロと

$$E = (X^T U_{ij})^2 / U_{ij}^2 \quad (5)$$

i = 0 から L - 1 までこの E の演算には、E の分子の内積に M L N 回、分母のエネルギーの演算に M N^L 回の積和演算が必要であり、さらに N^L 回の加減算、除算、比較が必要である。このほか合成部で M² N L、直交部で 2 M N L 回程度の積和演算が必要である。なお、H P の演算は適応符号ベクトル P を決定するときに求め

$$E = (X^T U_{0j} + X^T U_{1j})^2 / (U_{0j} + U_{1j})^2 \quad (6)$$

例えば B = 1 2 で、極性を除いて各系統に 5 ビット割り当てられるとすると、N = 2^{(12/2)-1} = 2⁵ = 3 2 となる。このとき分子の積和演算の数は 6 4 M 回に対し、分母のエネルギーは 1 0 2 4 M 回の演算を要する。従って、内積演算のみでその大きなものから複数を予備選択

$$D_2 = X^T U_{0j} + X^T U_{1j} = d_{0j} + d_{1j} \quad (7)$$

$$d_{0j} = X^T H \{ C_{0j} - (P^T H^T H C_{0j} P) / HP^2 \} \quad (8)$$

$$d_{1j} = X^T H \{ C_{1j} - (P^T H^T H C_{1j} P) / HP^2 \} \quad (9)$$

上記のなかで H は行列であるため H C の合成演算に多くの演算を必要とする。しかし、D₂ の計算の範囲では X^T H と P^T H^T H および H P² を先に一回だけ計算しておけば数多くの C に対する合成演算 (フィルタの畳み込み) H C をする必要がなくなる。そこで各系統独立に d_{0j}、d_{1j} を高速に計算し、それぞれの系統で内積の大きなものから候補を残し、その残した少ないベクトルの組合せで最終的な歪が最小となるものを選択する。こ

$$d_{0j} = X^T H \{ C_{0j} - (P^T H^T H C_{0j} P) / HP^2 \}$$

を計算し (S₄)、d_{0j} のうち大きいものから n 個を残す (S₅)。同様にして他方の乱数符号帳のベクトル C_{1j} について計算して d_{1j} を求め、大きいものから n 個残す (S₆)。これら残された 2 n 個の d_{0j}、d_{1j} に対す*

$$X - \{ g_1 HP + g_2 H (C_{0j(0)} + C_{1j(1)}) \}^2$$

を最小とする g₁、g₂ の組を決定する (S₉)。

【0 0 1 5】一般の L について X^T H と P^T H^T H および H P² の計算に M² + M² + M、また P^T H^T H C の演算に L M N、d_{ij} の演算に L M N 回の積和が必要である。さらに N のなかから n を選ぶソーティングも L 回必要である。以上が予備選択で、本選択は見かけ上、乱数符号帳のベクトル数が小さくなって従来どおりの距離計算を行なうことになる。

*で求めることができる。

【0 0 1 0】

【実施例】図 4 に示した従来の符号決定法をもとにこの発明による処理を説明する。図 4 の距離部の計算、つまり (2) 式を展開する。

し、その時の g を求め、この g の値を (4) 式に代入すると D = X^T X - 2 E + E となる。従って D の最小化は E の最大化と等価となる。

【0 0 1 1】

られているのでここでの演算量には含めない。

【0 0 1 2】簡単のため、特に L = 2 の場合についてこの発明の処理の背景を説明する。また L = 1 の場合にはこの発明のメリットは少なく、L = B すなわち V S E L P の場合にはこの発明を適用することができない。E を L = 2 としてもう一度以下のように書き直す。

し、その選択された少数の候補ベクトルのみについてエネルギーを計算することで演算量を削減できる。(6) 式の分子の項のかつこ内を D₂ とし、D₂ を d_{0j} と d_{1j} にわけて展開する。

【0 0 1 3】

の計算手順をまとめて図 1 に示す。

【0 0 1 4】まず適応符号帳の出力ベクトル P を決定する (S₁)。この時 H P が計算されている。次に X^T H、P^T H^T H、H P² を計算する (S₂)。次に一方の乱数符号帳のベクトル C_{0j} について C_{0j} - (P^T H^T H C_{0j} P) / H P² を計算し (S₃)、更に

* 乱数ベクトル C_{0j}、C_{1j} についてのみ U_{0j}、U_{1j} を計算する (S₇)。U_{0j} + U_{1j}² を含めて (6) 式の E の最大値となる C_{0j}、C_{1j} を探す (S₈)。その C_{0j(0)}、C_{1j(1)} について

【0 0 1 6】上述では合成フィルタの伝達関数としてインパルス応答行列 H を用いたが、これと等価なフィルタ操作の処理を示す伝達関数を用いてもよい。

【0 0 1 7】

【発明の効果】上記実施例で示したように、この発明によれば乱数ベクトルの距離計算部において、数多くの乱数ベクトルに対する合成フィルタ演算をまったく行わずに各系統ごとの内積演算を行なうことができ、さらに

内積演算で絞られた候補についてのみエネルギーの計算を行なうため、演算量を大幅に削減することができる。

【0018】M = 80、L = 2、N = 32の場合に概算すると予備選択は本選択の数十分の1の演算量となる。一方、本選択の演算量は乱数ベクトルに比例する演算と乱数ベクトルの2乗以上に比例する演算とからなるため、予備選択の候補数が減少するとそれに比例する以上に演算量が減少する。例えば予備選択で1/4に絞ると、予備選択の分を含めても演算量は1/4以下となる。またこのようにしても歪の増大は非常に少なく、最終的な出力音声のSNRの違いは0.5 dB以下であ

る。

【図面の簡単な説明】

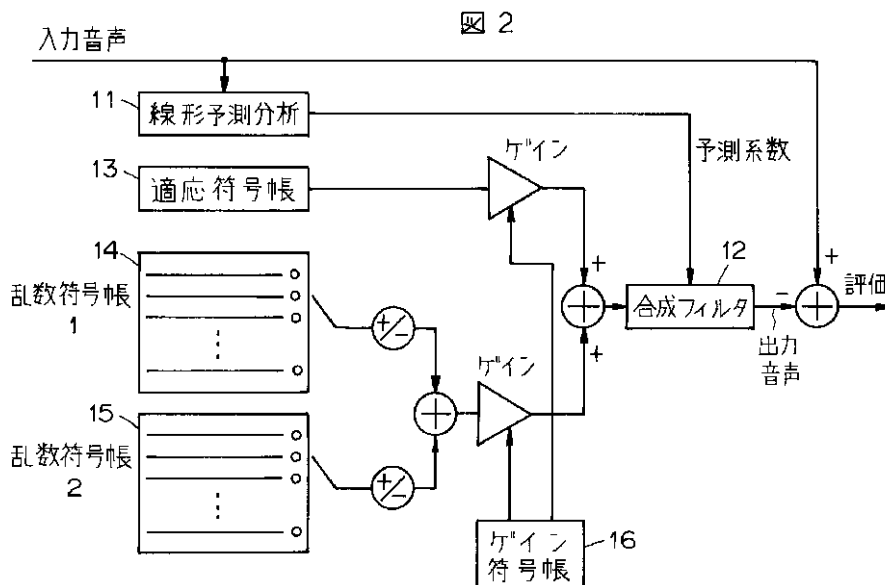
【図1】この発明方法の要部である乱数ベクトルの決定手順を示す流れ図。

【図2】音声符号化における励振信号符号化法の一般的構成の要部を示すブロック図。

【図3】乱数符号帳の系統数（チャンネル数）と総ベクトル数との関係を示す。

【図4】乱数符号を最適に決定する従来の手順を示す流れ図。

【図2】

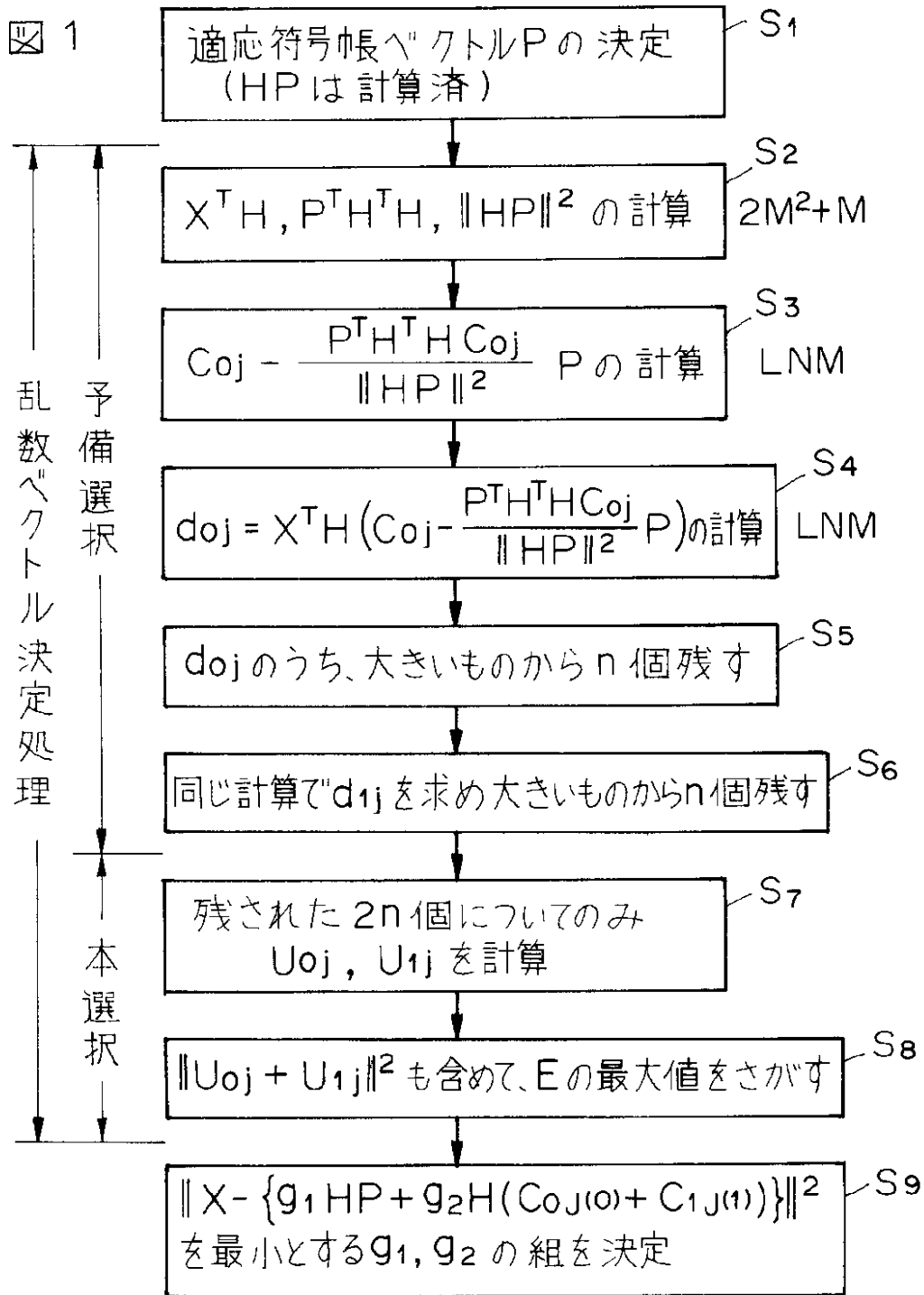


【図3】

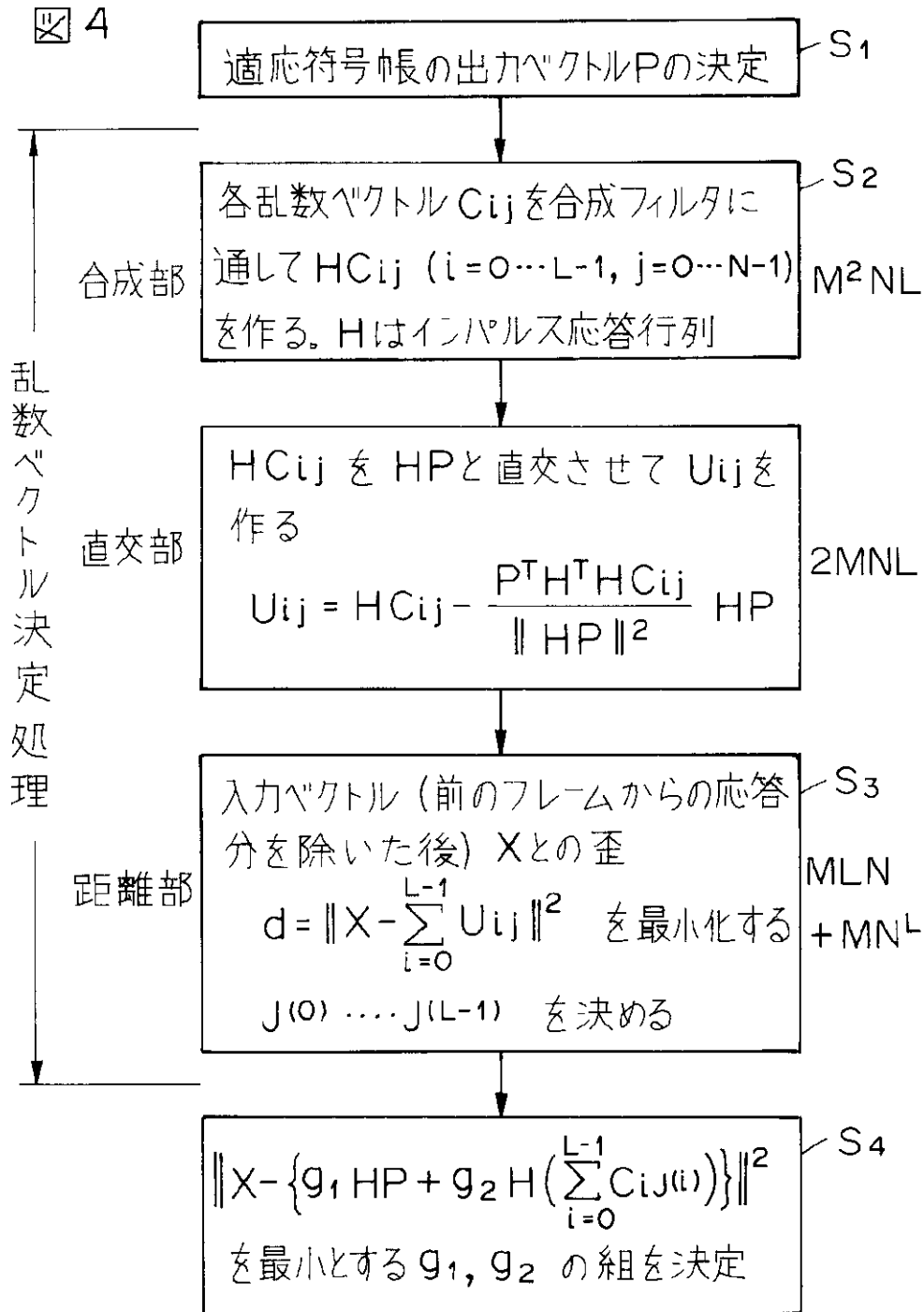
図 3

	系統(チャンネル)の数 L	各系統内のベクトル数 N	ベクトル総数 T
(CELP)	1	2^{B-1}	2^{B-1}
	2	$2^{B/2-1}$	$2^{B/2}$
	⋮		
(VSELP)	B	1	B

【図1】



【図 4】



フロントページの続き

(56)参考文献 特開 平 3 - 130800 (J P , A)
 特開 平 4 - 270400 (J P , A)
 特開 平 4 - 51199 (J P , A)

(58)調査した分野(Int.Cl.7, DB名)
 G10L 19/00 - 19/14
 H03M 7/30
 H04B 14/04