

(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 1 9 4 6 7 号

( P 3 2 1 9 4 6 7 )

(45)発行日 平成13年10月15日(2001.10.15)

(24)登録日 平成13年8月10日(2001.8.10)

(51)Int. Cl. <sup>7</sup>	識別記号	F I		
G 1 0 L	13/00	G 1 0 L	9/00	M
	19/00		9/14	S
	19/12		9/00	N

請求項の数 5

(全 1 2 頁)

(21)出願番号	特願平4-170890	(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4年6月29日(1992.6.29)	(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6-12095	(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6年1月21日(1994.1.21)	(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
審査請求日	平成11年4月9日(1999.4.9)	(74)代理人	100091959 弁理士 澤井 敬史
		審査官	渡邊 聡

最終頁に続く

(54)【発明の名称】音声復号化方法

1

(57)【特許請求の範囲】

【請求項 1】音声符号化情報と誤り訂正符号とをフレームごとに受信して前記音声符号化情報を復号して得られた現在、過去及び将来フレームのうち少なくともいずれかのパラメータに基づくサブフレームのパラメータ、を用いてサブフレームごとに適応符号帳及び雑音符号帳からのベクトルにそれぞれ重み付けられた後、加算した励振ベクトルを合成フィルタへ供給し、パワを乗じて音声再生する音声復号化方法において、  
音声符号化情報を復号する際、過去、現在及び将来フレームのそれぞれの符号誤り検出情報を用い、  
過去、現在、将来のフレームに対応するフラグをそれぞれS2、S1、S0としてフラグS1が1からM-1(Mはバースト性の誤りが加わっていると考えられるフレーム数)で、前記誤り訂正符号を用いて誤り検出されなかった場合、フ

2

ラグS0を0とし、誤り検出された場合フラグS0をS1+1とし、  
フラグS1がMの場合、前記誤り訂正符号を用いて誤り検出されなかった場合、フラグS0をM+1とし、誤り検出された場合フラグS0をMとし、  
フレーム更新の際にフラグS2をフラグS1で、フラグS1をフラグS0で更新し、  
フラグS1が0か、  
フラグS1が1からMのいずれかでフラグS0が0またはM+1か、  
フラグS1が1からMのいずれかでフラグS0が0でもM+1でもないか、  
フラグS1がM+1か、に現在フレーム内のサブフレームのパラメータを再生又は修復する処理が依存することを特徴とする音声復号化方法。

10

【請求項2】 請求項1記載の音声復号化方法において、  
 前記音声符号化情報は線形予測パラメータ符号を含み、  
 前記パラメータとして線スペクトル対について、  
 フラグS1が0の場合、フラグS2が0ならば過去に音声の再生に用いられたパラメータと現在フレームのパラメータからサブフレームのパラメータを補間し、  
 フラグS2が1からMのいずれかならば現在フレームのパラメータを用い、  
 もしくはフラグS0が0ならば将来フレームのパラメータと現在フレームのパラメータからサブフレームのパラメータを補間し、フラグS0が1からMのいずれかならば前記サブフレームのパラメータとして現在フレームのパラメータを用い、  
 フラグS1が1からMのいずれかでフラグS0が0またはM+1の場合、過去フレームのパラメータと将来フレームのパラメータを用いて前記サブフレームのパラメータを補間し、  
 フラグS1が1からMのいずれかでフラグS0が0でもM+1でもない場合、前記サブフレームのパラメータとして過去に音声の再生に用いられたパラメータを用い、  
 フラグS1がM+1の場合、過去に音声の再生に用いられたパラメータが平坦スペクトルとなるパラメータと現在フレームのパラメータを補間して前記サブフレームのパラメータを補間し、  
 もしくはフラグS0が0ならば将来フレームのパラメータと現在フレームのパラメータから前記サブフレームのパラメータを補間し、フラグS0がMならば前記サブフレームのパラメータとして現在フレームのパラメータを用いて前記合成フィルタの係数として設定することを特徴とする音声復号化方法。

【請求項3】 請求項1記載の音声復号化方法において、  
 前記音声符号化情報はパワパラメータ符号を含み、前記パラメータとしてパワについてフラグS1が0の場合、フラグS2が0ならば過去に音声の再生に用いられたパラメータと現在フレームのパラメータからサブフレームのパラメータを補間し、  
 フラグS2が1からMのいずれかならば現在フレームのパラメータを用い、  
 該パラメータが過去フレームで音声の再生に用いられたパラメータの定数倍より大きい場合に、該パラメータに1より小さい値を乗じたパラメータを用い、  
 もしくはフラグS0が0ならば将来フレームのパラメータと現在フレームのパラメータからサブフレームのパラメータを補間し、フラグS0が1からMのいずれかならば前記サブフレームのパラメータとして現在フレームのパラメータを用い、  
 フラグS1が1からMのいずれかでフラグS0が0またはM+1の場合、過去フレームで音声再生に用いられたパラメータと将来フレームのパラメータを用いて前記サブフレームのパラメータを補間し、該補間結果のパラメータが過去フレームで音声の再生に用いられたパラメータの一定の閾値倍より大きい場合に、前記補間結果のパラメータに1より小さい値を乗じ、  
 フラグS1が1からMのいずれかでフラグS0が0でもM+1でもない場合、前記サブフレームのパラメータとして過去に音声の再生に用いられたパラメータに1より小さい値を乗じ、  
 フラグS1がM+1の場合、過去に音声の再生に用いられたパラメータと現在フレームのパラメータから前記サブフレームのパラメータを補間し1より小さい値を乗じ、  
 もしくはフラグS0が0ならば将来フレームのパラメータと現在フレームのパラメータから前記サブフレームのパラメータを補間し、フラグS0がMならば前記サブフレームのパラメータとして現在フレームのパラメータに1より小さい値を乗じて音声を再生することを特徴とする音声復号化方法。

【請求項4】 請求項1記載の音声復号化方法において、  
 前記音声符号化情報はサブフレーム毎のピッチ周期符号を含み、  
 フラグS1が0の場合、現在フレームのピッチ周期符号で示されるピッチ遅延に基づき、  
 フラグS1が1からMのいずれかでフラグS0が0またはM+1の場合、過去フレームにおいて音声の再生に用いられたピッチ遅延と将来フレームのピッチ周期符号で示されるピッチ遅延に基づき、  
 フラグS1が1からMのいずれかでフラグS0が0でもM+1でもない場合、過去フレームで音声の再生に用いられたピッチ遅延に基づき、  
 フラグS1がM+1の場合、現在フレームのピッチ周期符号で示されるピッチ遅延に基づき、  
 過去の励振ベクトルを切り出し繰り返した時系列ベクトルを前記適応符号帳から出力することを特徴とする音声復号化方法。

【請求項5】 請求項1記載の音声復号化方法において、  
 前記音声符号化情報はサブフレーム毎のゲイン符号を含み、  
 フラグS1が0の場合、現在フレームのゲイン符号で示されるゲインの組を用い、フラグS1が1からMのいずれかでフラグS0が0またはM+1の場合、過去フレームにおいて音声の再生に用いられたゲインと将来フレームのゲイン符号で示されるゲインを用い、  
 フラグS1が1からMのいずれかでフラグS0が0でもM+1でもない場合、過去フレームで音声の再生に用いられたゲインの組を用い、  
 フラグS1がM+1の場合、現在フレームのゲイン符号で示されるゲインの組を用い、

ータと将来フレームのパラメータを用いて前記サブフレームのパラメータを補間し、該補間結果のパラメータが過去フレームで音声の再生に用いられたパラメータの一定の閾値倍より大きい場合に、前記補間結果のパラメータに1より小さい値を乗じ、  
 フラグS1が1からMのいずれかでフラグS0が0でもM+1でもない場合、前記サブフレームのパラメータとして過去に音声の再生に用いられたパラメータに1より小さい値を乗じ、  
 フラグS1がM+1の場合、過去に音声の再生に用いられたパラメータと現在フレームのパラメータから前記サブフレームのパラメータを補間し1より小さい値を乗じ、  
 もしくはフラグS0が0ならば将来フレームのパラメータと現在フレームのパラメータから前記サブフレームのパラメータを補間し、フラグS0がMならば前記サブフレームのパラメータとして現在フレームのパラメータに1より小さい値を乗じて音声を再生することを特徴とする音声復号化方法。

【請求項4】 請求項1記載の音声復号化方法において、  
 前記音声符号化情報はサブフレーム毎のピッチ周期符号を含み、  
 フラグS1が0の場合、現在フレームのピッチ周期符号で示されるピッチ遅延に基づき、  
 フラグS1が1からMのいずれかでフラグS0が0またはM+1の場合、過去フレームにおいて音声の再生に用いられたピッチ遅延と将来フレームのピッチ周期符号で示されるピッチ遅延に基づき、  
 フラグS1が1からMのいずれかでフラグS0が0でもM+1でもない場合、過去フレームで音声の再生に用いられたピッチ遅延に基づき、  
 フラグS1がM+1の場合、現在フレームのピッチ周期符号で示されるピッチ遅延に基づき、  
 過去の励振ベクトルを切り出し繰り返した時系列ベクトルを前記適応符号帳から出力することを特徴とする音声復号化方法。

【請求項5】 請求項1記載の音声復号化方法において、  
 前記音声符号化情報はサブフレーム毎のゲイン符号を含み、  
 フラグS1が0の場合、現在フレームのゲイン符号で示されるゲインの組を用い、フラグS1が1からMのいずれかでフラグS0が0またはM+1の場合、過去フレームにおいて音声の再生に用いられたゲインと将来フレームのゲイン符号で示されるゲインを用い、  
 フラグS1が1からMのいずれかでフラグS0が0でもM+1でもない場合、過去フレームで音声の再生に用いられたゲインの組を用い、  
 フラグS1がM+1の場合、現在フレームのゲイン符号で示されるゲインの組を用い、

適応符号帳及び雑音符号帳からのベクトルにそれぞれ重み付けることを特徴とする音声復号化方法。

【発明の詳細な説明】

【0001】

【産業上の利用分野】この発明は、例えば伝送路上で符号誤りが頻発する移動無線システムへ適用され、符号誤り制御技術（誤り訂正／検出符号化処理等）が施された音声符号化情報と符号誤り検出情報から、符号誤りによる品質劣化の少ない音声を再生するための音声復号化方法に関する。

【0002】

【従来の技術】ディジタル移動無線や音声蓄積サービスのための8kbit/s以下の高能率音声符号化方法としては、符号励振線形予測符号化（Code-Excited Linear Prediction Coding：CELP）、ベクトル和励振線形予測符号化（Vector Sum Excited Linear Prediction Coding：VSELP）といった手法が代表的である。それぞれの技術については、M.R.Schroeder and B.S.Atal：“Code-Excited Linear Prediction（CELP）：High-quality Speech at Very Low Bit Rates”，Proc. ICASSP-85, 25. 1.1, pp.937-940，（1985年）、および、I.A.Gerson and M.A. Jasiuk：“Vector Sum Excited Linear Prediction（VSELP）Speech Coding at 8kbps”，Proc. ICASSP-90, S9.3, pp.461-464,（1990年）に述べられている。

【0003】ここでは、CELP符号化をとりあげて説明する。CELP符号化では、音声符号化情報として、線形予測パラメータ符号と、そのフレームのワを示すワパラメータ符号とピッチ周期符号と雑音符号と音源のゲインパラメータ符号を、5～50ms程度を1フレームとして送信する。CELP復号器を図1に示す。CELP復号器では、復号された線形予測パラメータAを線形予測合成フィルタ12の係数として設定する。それまでに得られた直前（過去）のフレームの励振ベクトル（残差波形）と、復号されたピッチ周期符号Lを用いて適応符号帳10からその周期で、過去の励振ベクトルを切り出し、これをフレーム分繰り返した時系列ベクトルが出力され、また、入力された雑音符号Cが示す符号ベクトルが雑音符号帳11からベクトルとして読みだされる。これら時系列ベクトルは、乗算器21、22で入力されたゲイン符号G(=g1, g2)に応じて、それぞれ重み付けされた後、加算器23で加算され、その加算出力が励振ベクトルとして合成フィルタ12へ供給され、ワゲインPを乗算器24で乗じて、再生音声を得られる。さらに、ポストフィルタ13により、ホルマント強調、ピッチ強調により雑音感を減らした最終的な音声出力を得る。

【0004】符号誤りの発生する応用分野に適用される音声伝送方式では、誤り訂正符号化技術を用いて符号誤りによる伝送音声の品質劣化を押さえている。しかし、バースト誤りが頻発する移動無線のような適用領域では、強力な誤り訂正符号を用いたとしても十分な適用効

果が得られない。これは、バースト内での誤りの密度が高いため、たとえ強力な誤り訂正符号を用いてもすべての誤りを除去することが難しいことに起因する。

【0005】訂正しきれない誤りを残したままの符号化情報から音声を復号すると、復号音声は大きく歪む。歪みを小さく押さえるため、誤り検出情報として、フレームの符号が正しく復号できたかどうかをチェックする誤り検出機能を持たせておき、誤りが検出された場合には、通常の復号法とは異なる処理、すなわち、波形修復処理を行ない、符号誤りの影響を極力押さえる工夫がされている。

【0006】これまでの修復方法としては、次の2種類がある。一つは、例えば、M.J.McLaughlin：“Channel coding for digital speech transmission in the Japanese digital cellular system”，電子情報通信学会無線通信システム研究会、RCS90-27, pp.41-45, 1990年、に示すように、現在のフレームが符号誤りのあるフレームの場合に、過去のフレームのパラメータを繰り返す方法や、パラメータを繰り返してさらに、過去のフレームの状態に応じて現フレームの利得を制御する方法である。

【0007】もう一つの修復方法は、特開平4-30200「音声復号化方式」に、示されているように、誤りがあった場合に、過去の正しいフレームと将来の正しいフレームから、現在の誤ったフレームを補間する方法がある。この場合には、過去の正しいフレームと将来の正しいフレームが時間的に近ければ、歪みの小さい補間を行なうことができる。

【0008】

【発明が解決しようとする課題】しかし、過去のみからの修復は、結局、過去の音声を繰り返して、誤りのあるフレームのワを小さくして歪みを聞こえにくくするものであり、この場合には、過去の情報のみから現在の誤りフレームパラメータを決定するので、修復のために現フレーム以上の遅延は必要ないが、次のフレームとの連続性が悪くなり不連続性が新たな品質劣化を招くという欠点がある。

【0009】また、特開平4-30200による過去と将来の正しいフレームからの補間は、バースト誤りの頻発する伝送路に適用した場合には、現在のフレームの次のフレームが正しいフレームであるとは限らず、過去と将来の正しいフレームが時間的に大きく離れてしまう場合があり、補間によってかえって大きな歪みを生じてしまう欠点がある。また、現在フレーム以降何フレームも誤りが続く場合、将来の正しいフレーム情報を使用するためには、その分だけ大きな遅延時間が必要となり、移動無線のように実時間性を重視する用途には適さない。

【0010】さらに、CELP符号化では、適応符号帳や合成フィルタのメモリが過去のフレームの復号結果に依存するので、現在フレームで符号誤りがなくても過去のフレームで誤りがあれば、再生音声に歪みが生じてしまう

という問題がある。本発明の目的は、このような問題点を解決し、バースト誤りのある伝送路でも品質劣化の少ない音声復号化方法を提供することにある。

【0011】

【課題を解決するための手段】 本発明では、音声符号化情報と誤り訂正符号とをフレームごとに受信して前記音声符号化情報を復号して得られた現在、過去及び将来フレームのうち少なくともいずれかのパラメータに基づくサブフレームのパラメータ、を用いてサブフレームごとに適応符号帳及び雑音符号帳からのベクトルにそれぞれ重み付けられた後、加算した励振ベクトルを合成フィルタへ供給し、パワを乗じて音声を再生する音声復号化方法において、音声符号化情報を復号する際、過去、現在及び将来フレームのそれぞれの符号誤り検出情報を用い、過去、現在、将来のフレームに対応するフラグをそれぞれ $S_2, S_1, S_0$ としてフラグ $S_1$ が1から $M-1$  ( $M$ はバースト性の誤りが加わっていると考えられるフレーム数)で、前記誤り訂正符号を用いて誤り検出されなかった場合、フラグ $S_0$ を0とし、誤り検出された場合フラグ $S_0$ を $S_{1+1}$ とし、フラグ $S_1$ が $M$ の場合、前記誤り訂正符号を用いて誤り検出されなかった場合、フラグ $S_0$ を $M+1$ とし、誤り検出された場合フラグ $S_0$ を $M$ とし、フレーム更新の際にフラグ $S_2$ をフラグ $S_1$ で、フラグ $S_1$ をフラグ $S_0$ で更新し、フラグ $S_1$ が0か、フラグ $S_1$ が1から $M$ のいずれかでフラグ $S_0$ が0または $M+1$ か、フラグ $S_1$ が1から $M$ のいずれかでフラグ $S_0$ が0でも $M+1$ でもないか、フラグ $S_1$ が $M+1$ か、に現在フレーム内のサブフレームのパラメータを再生又は修復する処理が依存することを特徴とする音声復号化方法。ここで、さらにフラグ $S_1$ が0の場合、フラグ $S_2$ が0ならば過去に音声の再生に用いられたパラメータと現在フレームのパラメータからサブフレームのパラメータを補間し、フラグ $S_2$ が1から $M$ のいずれかならば現在フレームのパラメータを用い、もしくはフラグ $S_0$ が0ならば将来フレームのパラメータと現在フレームのパラメータからサブフレームのパラメータを補間し、フラグ $S_0$ が1から $M$ のいずれかならば前記サブフレームのパラメータとして現在フレームのパラメータを用い、フラグ $S_1$ が1から $M$ のいずれかでフラグ $S_0$ が0または $M+1$ の場合、過去フレームのパラメータと将来フレームのパラメータを用いて前記サブフレームのパラメータを補間し、フラグ $S_1$ が1から $M$ のいずれかでフラグ $S_0$ が0でも $M+1$ でもない場合、前記サブフレームのパラメータとして過去に音声の再生に用いられたパラメータを用い、フラグ $S_1$ が $M+1$ の場合、過去に音声の再生に用いられたパラメータが平坦スペクトルとなるパラメータと現在フレームのパラメータを補間して前記サブフレームのパラメータを補間し、もしくはフラグ $S_0$ が0ならば将来フレームのパラメータと現在フレームのパラメータから前記サブフレームのパラメータを補間し、フラグ $S_0$ が $M$ ならば前記サブフレームのパラメータとして現在フレームのパラメ

ータを用いて前記合成フィルタの係数として設定する。あるいは、さらに前記音声符号化情報はパワパラメータ符号を含み、前記パラメータとしてパワについてフラグ $S_1$ が0の場合、フラグ $S_2$ が0ならば過去に音声の再生に用いられたパラメータと現在フレームのパラメータからサブフレームのパラメータを補間し、フラグ $S_2$ が1から $M$ のいずれかならば現在フレームのパラメータを用い、該パラメータが過去フレームで音声の再生に用いられたパラメータの定数倍より大きい場合に、該パラメータに1より小さい値を乗じたパラメータを用い、もしくはフラグ $S_0$ が0ならば将来フレームのパラメータと現在フレームのパラメータからサブフレームのパラメータを補間し、フラグ $S_0$ が1から $M$ のいずれかならば前記サブフレームのパラメータとして現在フレームのパラメータを用い、フラグ $S_1$ が1から $M$ のいずれかでフラグ $S_0$ が0または $M+1$ の場合、過去フレームで音声再生に用いられたパラメータと将来フレームのパラメータを用いて前記サブフレームのパラメータを補間し、該補間結果のパラメータが過去フレームで音声の再生に用いられたパラメータの一定の閾値倍より大きい場合に、前記補間結果のパラメータに1より小さい値を乗じ、フラグ $S_1$ が1から $M$ のいずれかでフラグ $S_0$ が0でも $M+1$ でもない場合、前記サブフレームのパラメータとして過去に音声の再生に用いられたパラメータに1より小さい値を乗じ、フラグ $S_1$ が $M+1$ の場合、過去に音声の再生に用いられたパラメータと現在フレームのパラメータから前記サブフレームのパラメータを補間し1より小さい値を乗じ、もしくはフラグ $S_0$ が0ならば将来フレームのパラメータと現在フレームのパラメータから前記サブフレームのパラメータを補間し、フラグ $S_0$ が $M$ ならば前記サブフレームのパラメータとして現在フレームのパラメータに1より小さい値を乗じる。もしくは、さらに前記音声符号化情報はサブフレーム毎のピッチ周期符号を含み、フラグ $S_1$ が0の場合、現在フレームのピッチ周期符号で示されるピッチ遅延に基づき、フラグ $S_1$ が1から $M$ のいずれかでフラグ $S_0$ が0または $M+1$ の場合、過去フレームにおいて音声の再生に用いられたピッチ遅延と将来フレームのピッチ周期符号で示されるピッチ遅延に基づき、フラグ $S_1$ が1から $M$ のいずれかでフラグ $S_0$ が0でも $M+1$ でもない場合、過去フレームで音声の再生に用いられたピッチ遅延に基づき、フラグ $S_1$ が $M+1$ の場合、現在フレームのピッチ周期符号で示されるピッチ遅延に基づき、過去の励振ベクトルを切り出し繰り返した時系列ベクトルを前記適応符号帳から出力する。または、さらに前記音声符号化情報はサブフレーム毎のゲイン符号を含み、フラグ $S_1$ が0の場合、現在フレームのゲイン符号で示されるゲインの組を用い、フラグ $S_1$ が1から $M$ のいずれかでフラグ $S_0$ が0または $M+1$ の場合、過去フレームにおいて音声の再生に用いられたゲインと将来フレームのゲイン符号で示されるゲインを用い、フラグ $S_1$ が1から $M$ のいずれかでフラグ $S_0$ が0

でもM+1 でもない場合、過去フレームで音声の再生に用いられたゲインの組を用い、ラグS1がM+1の場合、現在フレームのゲイン符号で示されるゲインの組を用い、適応符号帳及び雑音符号帳からのベクトルにそれぞれ重み付ける。

【0012】

【0013】

【実施例】以下では、高能率音声符号化方式としてCELP方式をとりあげ、CELP方式へのこの発明の実施例を説明する。本実施例では、CELP方式の復号器について説明するが、これと類似の方式または、類似の構成を含む方式に対しても適用が可能である。

【0014】図2は、本発明の実施例におけ音声復号処理のブロック図を示す。本実施例では、過去、現在フレームと将来の1フレームまでの情報を用いた復号化について示す。まず、伝送路より送られた符号をフレーム毎に誤り訂正回路30において誤り訂正処理を行なう。30より音声符号化情報をバッファ31に送り、誤り検出情報をバッファ34に送る。31、34のバッファには、過去および現在フレームとさらに少なくとも1フレーム先の情報まで取り込む。バッファ34からの誤り検出情報は、誤り状態管理部35に入り、過去、現在、将来フレームの誤り状態が決定される。

【0015】バッファ31の出力である音声符号化情報と誤り状態管理部35からの誤り状態をパラメータ復号部32に送り、現在フレームのパラメータを復号・補間する。パラメータ復号部32の出力により、音声復号器33で音声を再生する。誤り状態管理部35では、図3に示すように、図2のバッファ34に蓄えられている次のフレームまでの誤り検出情報をもとに、過去、現在、将来フレームの状態を表すフラグ群 352 (S2, S1, S0) を更新する。S2が過去のフレーム、S1が現在フレーム、S0が将来フレームに対応する。フラグS2とS1は、音声符号化のフレーム更新の際に、前のフレームでのフラグS1、S0の内容をシフトしたものである。S0フラグの内容は、状態フラグ更新部351 において対応するバッファ34からの入力フレームの誤り検出情報と、S1フラグ(これは、前フレームではS0フラグに相当する)の内容に従って決定される。

【0016】図4にS0の更新表を示す。図4において、バッファ34の入力 = 0 のとき、その入力フレームは、誤り訂正が完全にできたことを示し、バッファ34の入力 = 1 のとき、誤り訂正ができなかった、つまり、誤りが検出されたフレームであることを意味する。S1が0 のとき、入力フレームの誤りが検出されなければ、S0 = 0 とし、誤りが検出されれば、S0 = 1 とする。S1が1 からあらかじめ与えられた数M-1 までの場合には、入力フレームの誤りが検出されなければS0 = 0、誤りが検出されれば、S0 = S1+1として、対応するフレームまで、誤りフレームの連続して続いたフレーム数をフラグの内容として保持する。S1がM の場合には、入力フレームで誤りが検

出されなければ、S0=M+1とし、誤りが検出されれば、S0=Mとする。ここで、M は、連続するフレームにバースト性の誤りが加わっていると考えられる長さ、例えば6、とする。S1がM+1 の場合には、入力フレームで誤りが検出されなければ、S0=0、誤りが検出されれば、S0=Mとする。従って、S0=M+1の場合には、対応フレームの誤り検出がない場合であるが、バースト誤りが続いた後であるので、0 とはせず、さらにその次のフレームで、誤りが検出されなくなると0にする。

10 【0017】このような状態フラグの遷移に関しては、従来の技術で述べたM.J.McLaughlinの文献(電子情報通信学会無線通信システム研究会、RCS90-27, pp.41-45, 1990年)において、現在フレームを入力として、S0を現在フレームのフラグとして、過去のフラグとともに用いているが、本発明では、現在フレームがS1、将来のフレームをS0として、これから復号しようとするフレームの次に来るフレームのフラグを用いる点が大きく異なる。

20 【0018】次に、こうして得られた誤り管理制御部35の出力S(S2, S1, S0) を用いて、図2のパラメータ復号部32で過去、現在、将来フレームの誤り状態を考慮した各音声符号化情報のパラメータ復号を行なう。パラメータ復号部32での誤り状態フラグに依存した処理手順のフロー図の例を図5に示す。まず、図5のステップ101 において、現在フレームの誤り状態フラグS1が0 であるかどうかを判断し、0 ならば、ステップ111 の処理1 を実行する。処理1 では、S1=0 であるから、現在フレームの音声符号化情報に基づいて復号したパラメータを基本とし、S0, S2 を見て、S0またはS2が0 でない場合には、連続的にパラメータが接続するように修正する。

30 【0019】次に、ステップ102 において、S1が0 でなく、フラグが1 からM のいずれかの場合には、現在フレームにおいて、誤り検出されたことを示している。従って、この場合は、現在フレームの音声符号化情報には、誤りが含まれている。そこで、過去フレーム及び将来フレームの情報を使うことになる。ステップ103 において、S0=0またはM+1 ならば将来フレームが正しいので、ステップ112 の処理2 として、過去フレームで復号に使われたパラメータと将来フレームの復号パラメータから現在フレームのパラメータを両側から内挿補間する。その際、S1のフラグの誤り継続の度合いにより、内挿の重みを変えることもできる。例えば、S1がMに近ければ、過去フレームのパラメータは、バースト誤りが継続して、次に述べるステップ113 処理3 の外挿的な手法によって得られたパラメータであるので、過去のパラメータの重みは小さく、将来フレームの重みを大きくすることにより、より歪の小さい現在フレームのパラメータを得ることができる。ステップ103において、S0 = 0 かつS0 M+1 の場合、将来フレームには誤りが検出されている。その場合、ステップ113 の処理3 では、過去フレームで復号に使われたパラメータの現在フレームへの外挿

による補間処理を行なう。その際、S1フラグの誤り継続の度合いにより、外挿結果への重みを変えることもできる。

【0020】ステップ102において、N0の場合には、 $S1 = M+1$ であり、これは、過去のフレームにM フレーム以上のバースト誤りがあった後に現在フレームが正しいことを示している。従って、ステップ114の処理4では、現在フレームの音声符号化情報に基づいて復号したパラメータを基本とし、さらに、 $S2=M$ であるから、過去フレームにバースト誤りがあったこと、及び、 $S0$ が0またはM

になるので、その値によって、連続的にパラメータが接続するように修正する。

【0021】次に、実際のパラメータについて、図5の処理1～処理4の例について述べる。CELP符号化における各フレームの音声符号化情報は、図1に示したように、線形予測パラメータ符号A、パワパラメータ符号P、ピッチ周期符号Lと雑音符号帳の雑音符号ベクトル符号C、音源のゲインパラメータ符号Gからなる。まず、線形予測パラメータ符号Aについての実施例を示す。線形予測パラメータとしては、補間の際には、線形補間特性の良い線スペクトル対(LSP : Line Spectrum Pair)を用いる。線形予測パラメータは、今1フレームを20msとすれば、実際には、5ms～10ms程度のサブフレーム単位毎にパラメータを必要とする。その場合、前フレーム(過去フレーム)または、後フレーム(将来フレーム)とのパラメータ間の補間によって、サブフレームのパラメータを求める。すると、前(または、後)フレームに符号誤りがある場合には、誤ったパラメータと直接補間すると、大きな歪が生じる。また、現在フレームのパラメータが誤っている場合でもすぐ前後のパラメータが誤りのない状態あれば、前後フレームから、現在フレームを補間することができる。

【0022】過去(時刻 $n-1$ とする)、現在(時刻 $n$ )、将来(時刻 $n+1$ )フレームでの誤り訂正直後で補間をしないLSPパラメータをそれぞれ、 $(n-1)$ 、 $(n)$ 、 $(n+1)$ とし、誤り状態フラグを考慮して補間したあとの時刻 $n-1$ 、 $n$ に対応するパラメータを $(n-1)$ 、 $(n)$ で表す。また、 $a$ 、 $t$ 、 $b$ の範囲の $t$ に対して、時刻 $a$ で $X_a$ 、時刻 $b$ で $X_b$ のパラメータであるときの時刻 $t$ での内挿パラメータ $(t)$ を、

$$(t) = f\_A(t; (a, X_a), (b, X_b))$$

で定義する。関数 $f\_A$ は、2点 $(a, X_a)$ 、 $(b, X_b)$ 間を補間する関数で、時刻 $t$ の補間値を出力するものである。例えば、直線補間でも良い。

【0023】図5ステップ111の処理1では、現在フレームの誤り状態フラグ $S1=0$ であり、 $(n)$ には誤りがないので、 $S2, S0$ の誤り状態に応じて次のようにすれば良い。現在フレーム内の $t$ において、 $n-1 \leq t \leq n$ では、 $S2=0$ ならば、 $(t) = f\_A(t; (n-1), (n-1)), (n, (n))$ ,

$S2=1, \dots, M$ ならば、 $(t) = f\_A(t; (n-1), (n)), (n, (n))$

とする。また、 $n \leq t \leq n+1$ のとき、

$S0=0$ ならば、 $(t) = f\_A(t; (n), (n)), (n+1, (n+1))$ ,

$S0=1, \dots, M$ ならば、 $(t) = f\_A(t; (n), (n)), (n+1, (n))$

とする。

【0024】図5のステップ112の処理2では、現在フレームパラメータに誤りがあるので $(n)$ は使用せず、 $S0=0$ または $M+1$ なので、 $(n-1)$ 、 $(n+1)$ を用いて、現フレーム内の時刻 $t$ のパラメータを、

$(t) = f\_A(t; (n-1), (n-1)), (n+1, (n+1))$ により補間する。ステップ113の処理3では、 $S0=1, \dots, M$ なので、過去のパラメータを用いて、

$$(t) = f\_A(t; (n-1), (n-1)), (n+1, (n-1))$$

により、現フレーム内の時刻 $t$ のパラメータを補間する。この場合、 $S2, S1, S0$ の値に応じて $(t)$ のスペクトルのバンド幅を拡張するような関数を用いて出力のLSPパラメータの共振特性を弱めるようにして、歪を防ぐような補間をしても良い。

【0025】図5のステップ114の処理4では、 $(n)$ は使えるが、前フレームがバースト誤りなので、 $(n-1)$ を使うか、または $0$ として、平坦スペクトルとなるようなパラメータのどちらかを用いる。これを $'$ とする。現フレーム内の $t$ において、 $n-1 \leq t \leq n$ のとき、

$$(t) = f\_A(t; (n-1, ')), (n, (n))$$

$n \leq t \leq n+1$ では、

$S0=0$ ならば、 $(t) = f\_A(t; (n), (n)), (n+1, (n+1))$ ,

$S0=M$ ならば、 $(t) = f\_A(t; (n), (n)), (n+1, (n))$ とする。

【0026】次に、パワパラメータ符号の図5の各処理について考える。基本的には、線形予測パラメータの補間の場合と同様である。パワの場合もサブフレーム単位のパワを計算する場合について示す。過去(時刻 $n-1$ とする)、現在(時刻 $n$ )、将来(時刻 $n+1$ )フレームでの誤り訂正直後で補間をしないパワパラメータをそれぞれ、 $p(n-1)$ 、 $p(n)$ 、 $p(n+1)$ とし、誤り状態フラグを考慮して補間したあとの時刻 $n-1$ 、 $n$ に対応するパラメータを $P(n-1)$ 、 $P(n)$ とする。

【0027】時刻 $a$ で $X_a$ 、時刻 $b$ で $X_b$ であるときの時刻 $t$ でのパワの内挿パラメータ $P(t)$ を、 $P(t) = f\_P(t; (a, X_a), (b, X_b))$ で定義する。関数 $f\_P$ は、2点 $(a, X_a)$ 、 $(b, X_b)$ 間の直線補間、または、 $X_a, X_b$ の対数表現の直線補間といった関数でも良い。

【0028】図5のステップ111の処理1では、 $S1=0$ なので、 $p(n)$ には、誤りがないので、現フレーム内の $t$ において、 $n-1 \leq t \leq n$ では、 $S2=0$ ならば、 $P(t) = f\_P(t;$

$(n-1, P(n-1)), (n, p(n)), S_2=1, \dots, M$ ならば、 $P(t) = f\_P(t; (n-1, p(n)), (n, p(n)))$ とする。そして、結果の  $P(t)$  が、 $P(n-1)$  よりも一定の定数倍 (例えば 1) よりも大きい場合には、 $' P(t), (0 \underline{\hspace{1cm}} < 1)$ として、パワを小さく制限することにより、急激なパワ変化による歪を聞こえないようにする。また、 $n \quad t \quad n+1$ では、 $S_0=0$ ならば、 $P(t) = f\_P(t; (n, p(n)), (n+1, p(n+1))), S_0=1, \dots, M$ ならば、 $P(t) = f\_P(t; (n, p(n)), (n+1, p(n)))$ とする。

【0029】図5のステップ112の処理2では、現在フレームパラメータに誤りがあるので、 $p(n)$ は使用せずに、 $S_0=0$ または $M+1$ なので、 $P(n-1)$ 、 $p(n+1)$ を用いて、現フレーム内の時刻 $t$ のパラメータを、 $P(t) = f\_P(t; (n-1, P(n-1)), (n+1, p(n+1)))$ により補間する。ここで、一定の閾値 (例えば (1)) に対して、 $P(t) > P(n-1)$  であり、パワ変化が大きすぎる場合には、 $' P(t), (0 \underline{\hspace{1cm}} < 1)$ のようにパワを置換し、パワを小さく制限することにより、歪が聞こえないようにする。 $'$ は、 $S_1$ のバースト誤りの継続数によって変更しても良い。

【0030】ステップ113の処理3では、 $S_0=1, \dots, M$ なので、過去のパラメータを用いて、 $P(t) = f\_P(t; (n-1, P(n-1)), (n+1, P(n-1)))$ とし、さらに、誤りフレームの音量を弱めて $P(t)$ を $' P(t), (0 \underline{\hspace{1cm}} < 1)$ とする。 $'$ は、 $S_1$ のバースト誤りの継続数によって変更しても良い。例えば、 $S_1=M$ ならば、現在フレームまでに、 $M$ フレーム以上の大きなバースト誤りが生じていることを意味するので、 $' = 0$ として無音区間とし、歪みのある音声を出力しないようにする。

【0031】図5のステップ114の処理4では、 $p(n)$ は使えるが、前フレームがバースト誤りフレームである。そこで、現フレーム内の $t$ において、 $n-1 \quad t \quad n$ のとき、 $P(t) = f\_P(t; (n-1, P(n-1)), (n, p(n)))$ そして、 $P(t)$ を $' P(t), (0 \underline{\hspace{1cm}} < 1)$ に置き換える。 $n \quad t \quad n+1$ では、 $S_0=0$ ならば、 $P(t) = f\_P(t; (n, p(n)), (n+1, p(n+1))), S_0=M$ ならば、 $P(t) = f\_P(t; (n, p(n)), (n+1, p(n)))$ とする。 $S_0=M$ の場合は、さらに、 $P(t)$ を $' P(t), (0 \underline{\hspace{1cm}} < 1)$ とする。

【0032】次に、ピッチ周期符号 $L$ のパラメータ復号の例を示す。ピッチ周期符号 $L$ は、図1において、適応符号帳10のピッチ遅延を示す遅延値である。これは、1フレームを $k$ 個 (例えば2個) のサブフレームに分け、その個数だけ $L$ の値をピッチ情報として伝送する。ピッチ遅延は、CELPではピッチ周期だけでなく2倍周期、3倍周期といったとびとびの値をとることもある。従って、本実施例では、 $k=2$ とした場合について、前フレームと後のフレームとのピッチ周期の値を現フレームが誤ったときに置換するという手法について示す。もちろん、ピッチ周期の変動が小さく、前後フレームのピッチ周期を補間可能な場合には、補間によって求めてもよ

い。

【0033】過去 (時刻 $n-1$ )、現在 (時刻 $n$ )、将来 (時刻 $n+1$ ) フレームでの誤り訂正直後ピッチ遅延をそれぞれ、 $l k(n-1)$ 、 $l k(n)$ 、 $l k(n+1)$ 、( $k=1, 2$ )とし、誤り状態フラグを考慮したあとの時刻 $n-1$ 、 $n$ に対応するパラメータを $L k(n-1)$ 、 $L k(n)$ 、( $k=1, 2$ )とする。図5のステップ111の処理1では、 $S_1=0$ なので、 $l k(n)$ に誤りがないので、

$$L k(n) = l k(n), (k=1, 2)$$

とする。

【0034】図5のステップ112の処理2では、現在フレームのパラメータに誤りがあるので、 $l k(n)$ は使用せずに、

$$L 1(n) = L 2(n-1), L 2(n) = l 1(n+1)$$

とする。ステップ113の処理3では、 $S_0=1, \dots, M$ なので過去のパラメータを用いて、

$$L k(n) = L 2(n-1), (k=1, 2)$$

とする。

【0035】ステップ114の処理4では、現在フレームが正しいので、処理1と同じく、

$$L k(n) = l k(n), (k=1, 2)$$

とする。次に、音源ゲインパラメータ $G$ の復号例を示す。 $G$ は、図1において、適応符号帳10と雑音符号帳11で選ばれた各音源信号に乗ずるゲイン $g_1, g_2$ の組を総称している。この $G$ に関しては、ピッチ遅延と同様にパラメータ復号を行なう。過去 (時刻 $n-1$ )、現在 (時刻 $n$ )、将来 (時刻 $n+1$ ) フレームでの誤り訂正直後音源ゲインをそれぞれ、 $k(n-1)$ 、 $k(n)$ 、 $k(n+1)$ 、( $k=1, 2$ )とし、誤り状態フラグを考慮したあとの時刻 $n-1$ 、 $n$ に対応するパラメータを $G k(n-1)$ 、 $G k(n)$ 、( $k=1, 2$ )とする。

【0036】図5のステップ111の処理1では、 $S_1=0$ であり  $k(n)$ に誤りがないので、

$$G k(n) = k(n), (k=1, 2)$$

とする。図5のステップ112の処理2では、現在フレームパラメータに誤りがあるので、 $k(n)$ は使用せずに、 $G 1(n) = G 2(n-1)$ 、 $G 2(n) = l(n+1)$

とする。

【0037】ステップ113の処理3では、 $S_0=1, \dots, M$ なので過去のパラメータを用いて、

$$G k(n) = G 2(n-1), (k=1, 2)$$

とする。ステップ114の処理4では、現在フレームが正しいので、処理1と同じく、

$$G k(n) = k(n), (k=1, 2)$$

とする。

【0038】雑音符号帳の雑音ベクトル符号 $C$ の復号に関しては、フレーム間の相関は、大きくないので、本実施例では、誤り状態フラグに関係なく現在フレームに送られた符号をそのまま再生に使用する。こうして、図2のパラメータ復号部32で、過去、現在、将来の誤り状態

フラグに基づき、音声符号化情報のパラメータが復号・修復される。その結果が、次の音声復号器33に渡され、音声再合成される。

【0039】音声復号器33においても、誤り状態管理部35からの過去、現在、将来のフレーム誤り状態フラグS (=S2,S1,S0)を用いて、復号音声の再生音声を制御することにより伝送誤りによる歪みを小さくできる。すなわ

$$H(z) = \frac{(1 - \eta z^{-1})A(z/\gamma 1)}{((1 - \delta B(z))A(z/\gamma 2))}$$

$$A(z) = \sum_{i=0}^n \alpha_i z^{-i}$$

$$B(z) = \sum_{i=-1}^1 \beta_i z^{-i}$$

【0041】ここで、 $\alpha_i$  は、各フレームの線形予測係数であり、 $\beta_i$  は長期予測係数、 $\tau$  はピッチ遅延である。これらは、伝送パラメータあるいは、復号音声から適応的に決定される。

$A(z/\gamma 1)/A(z/\gamma 2)$

でホルマント強調を行い、 $1/(1 - B(z))$ でピッチの強調を、さらに、

【0042】

【数2】

$$1 - \eta z^{-1}$$

【0043】でスペクトル傾斜補正の効果をもたせている。本実施例では、 $\eta$ 、 $\delta$ 、 $\gamma 1$ 、 $\gamma 2$ を誤り状態フラグS (=S2,S1,S0)の関数としてH(z)を各フレーム毎に適応的に変化させる。S2,S1,S0がいずれも0またはM+1の場合には、通常値を使用するが、S2,S1,S0のいずれかが0とM+1以外で誤りフレームである場合には、適応符号帳や、雑音符号ベクトル符号に誤りが含まれている可能性が高いので、 $\eta$ と $\delta$ を小さくして、高域のスペクトル傾斜と音源のピッチ強調を弱め、 $\gamma 1$ と $\gamma 2$ を調整してホルマント強調の度合いを強めることにより、出力音声の雑音感を抑えることができる。ここで、S0の将来フレームまでの誤り状態フラグを用いることにより、S2,S1だけの場合よりも、 $\eta$ 、 $\delta$ 、 $\gamma 1$ 、 $\gamma 2$ の切替えを連続的に行うことが可能となり、フレーム間での切替えによる音質変化による劣化を抑えることができる。

ち、図1のポストフィルタ13において、Sを用いてポストフィルタの係数を制御して聴覚的な歪みを抑える。通常、ポストフィルタは、次のようなH(z)によって構成されている。

【0040】

【数1】

【0044】

【発明の効果】以上説明したように、本発明によれば、音声符号化ビットに伝送路の符号誤りが残留した場合に、復号音声に重畳する耳障りな歪みを抑えることが可能となる。

【図面の簡単な説明】

【図1】 CELP音声復号器のブロック図である。

【図2】本発明の実施例における音声復号処理のブロック図である。

【図3】図2の誤り状態管理部35の拡大図である。

30 【図4】図3の状態フラグ更新部351における、S0フラグの更新表である。

【図5】誤り状態フラグに応じた処理フローの概略図である。

【符号の説明】

1 0 は適応符号帳

1 1 は雑音符号帳

1 2 は線形予測合成フィルタ

1 3 はポストフィルタ

2 1 , 2 2 , 2 4 は乗算器

40 2 3 は加算器

3 0 は誤り訂正回路

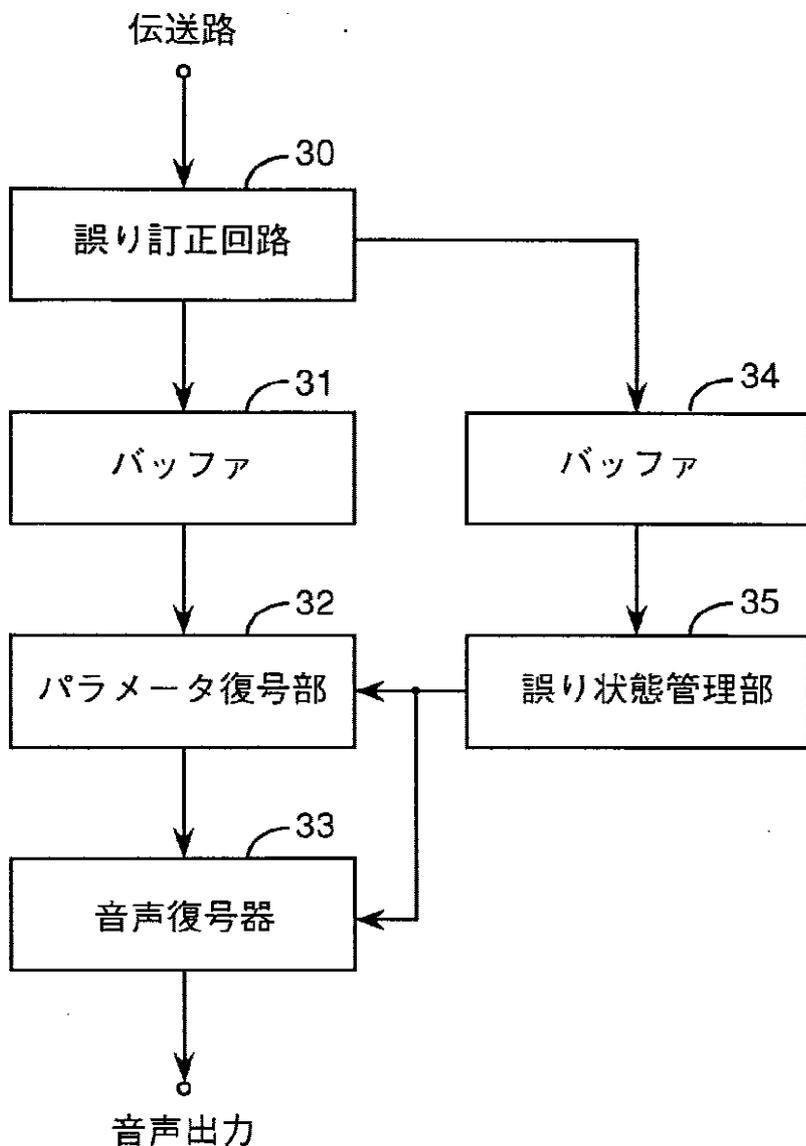
3 1 , 3 4 はバッファ

3 2 はパラメータ復号部

3 3 は音声復号器

3 5 は誤り状態管理部

【図 2】

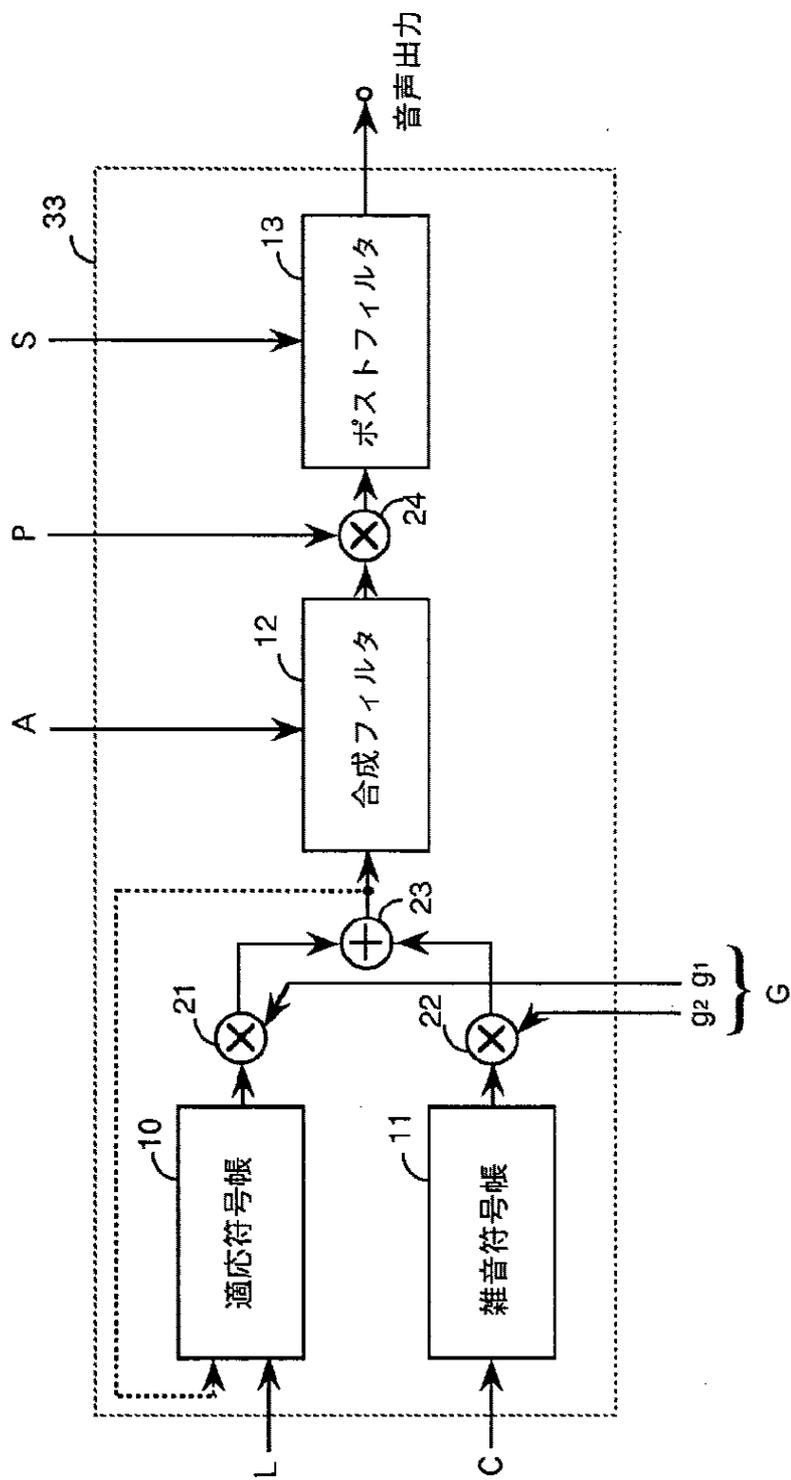


【図 4】

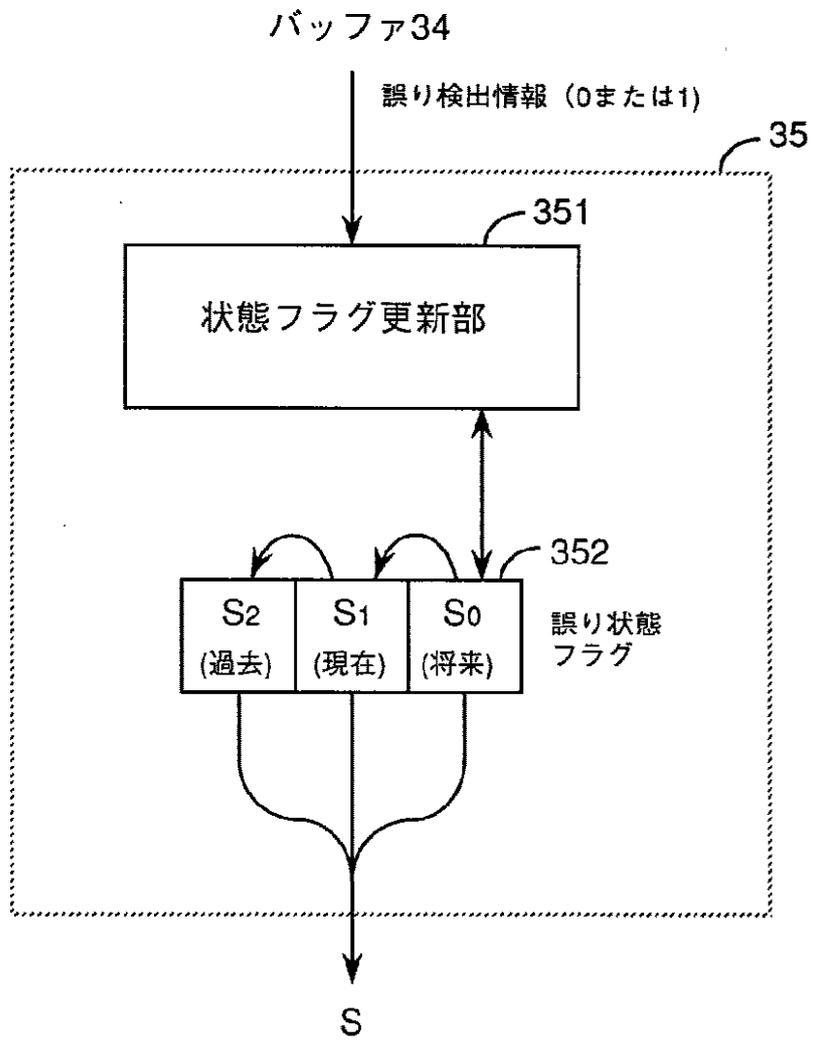
Soの更新

新Si (旧So)	バッファ34 入力	新So
0	0	0
1	0	0
2	0	0
⋮	⋮	⋮
M-1	0	0
M	0	M+1
M+1	0	0
0	1	1
1	1	2
2	1	3
⋮	⋮	⋮
M-1	1	M
M	1	M
M+1	1	M

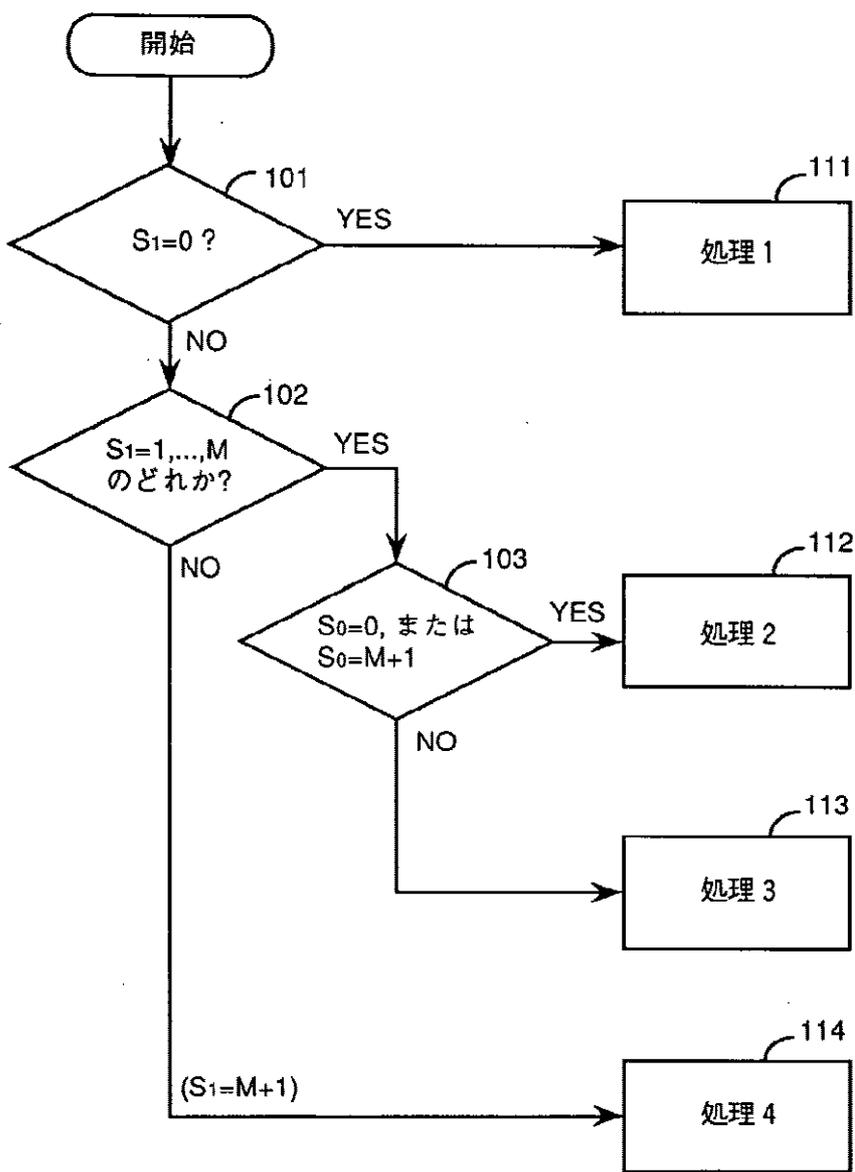
【図 1】



【図3】



【図 5】



フロントページの続き

(56) 参考文献 特開 平 2 - 176800 ( J P , A )  
特開 平 4 - 30200 ( J P , A )

(58) 調査した分野 ( Int . Cl . 7 , D B 名 )

G10L 13/00  
G10L 19/00