

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第3552201号

(P3552201)

(45) 発行日 平成16年8月11日(2004.8.11)

(24) 登録日 平成16年5月14日(2004.5.14)

(51) Int. Cl.<sup>7</sup>

F I

G 1 0 L 19/06  
 G 1 0 L 19/00  
 G 1 0 L 19/04  
 G 1 0 L 19/12  
 H 0 3 M 7/30

G 1 0 L 9/14 H  
 H 0 3 M 7/30 B  
 G 1 0 L 9/08 H  
 G 1 0 L 9/14 J  
 G 1 0 L 9/14 S

請求項の数 4 (全 14 頁)

(21) 出願番号 特願平11-185114  
 (22) 出願日 平成11年6月30日(1999.6.30)  
 (65) 公開番号 特開2001-13999(P2001-13999A)  
 (43) 公開日 平成13年1月19日(2001.1.19)  
 審査請求日 平成14年2月27日(2002.2.27)

(73) 特許権者 000003078  
 株式会社東芝  
 東京都港区芝浦一丁目1番1号  
 (74) 代理人 100058479  
 弁理士 鈴江 武彦  
 (74) 代理人 100084618  
 弁理士 村松 貞男  
 (74) 代理人 100068814  
 弁理士 坪井 淳  
 (74) 代理人 100092196  
 弁理士 橋本 良郎  
 (74) 代理人 100091351  
 弁理士 河野 哲  
 (74) 代理人 100088683  
 弁理士 中村 誠

最終頁に続く

(54) 【発明の名称】 音声符号化方法および装置

(57) 【特許請求の範囲】

## 【請求項1】

入力音声信号をスペクトル包絡を表すスペクトルパラメータと残差成分とで表し、これらスペクトルパラメータおよび残差成分を符号化する音声符号化方法において、前記入力音声信号から求められた第1の自己相関係数を修正して得られた第2の自己相関係数から前記スペクトルパラメータを算出して符号化し、  
 前記第2の自己相関係数を得る際の前記第1の自己相関係数の修正の度合いよりも小さな修正の度合で前記第1の自己相関係数を修正して求められた第3の自己相関係数から聴覚重み特性を求め、  
 前記スペクトルパラメータおよび前記聴覚重み特性を用いて前記残差成分を符号化すること

10

## 【請求項2】

入力音声信号をスペクトル包絡を表すスペクトルパラメータと残差成分とで表し、これらスペクトルパラメータおよび残差成分を符号化する音声符号化方法において、前記入力音声信号から求められた第1の自己相関係数を第1の自己相関窓を用いて修正して得た第2の自己相関係数から前記スペクトルパラメータを算出して符号化し、  
 前記第1の自己相関係数を前記第1の自己相関窓よりも前記第1の自己相関係数に与える修正の度合いが小さい第2の自己相関窓を用いて修正して得た第3の自己相関係数から聴覚重み特性を求め、  
 前記スペクトルパラメータおよび聴覚重み特性を用いて前記残差成分を符号化を行うこと

20

を特徴とする音声符号化方法。

【請求項 3】

入力音声信号をスペクトル包絡を表すスペクトルパラメータと残差成分とで表し、これらスペクトルパラメータおよび残差成分を符号化する音声符号化方法において、前記入力音声信号から所定の時間単位毎に第 1 の自己相関係数を算出するステップと、前記第 1 の自己相関係数に対し第 1 の自己相関窓を用いて窓掛けを行うことにより、前記第 1 の自己相関係数を修正して第 2 の自己相関係数を求めるステップと、前記第 2 の自己相関係数を用いて第 1 のスペクトルパラメータを算出するステップと、前記第 1 のスペクトルパラメータを符号化するステップと、前記第 1 の自己相関係数に対し前記第 1 の自己相関窓よりも前記第 1 の自己相関係数に与える修正の度合いが小さい、前記第 1 の自己相関係数とは形状の異なる第 2 の自己相関窓を用いて窓掛けを行うことにより、前記第 1 の自己相関係数を修正して第 3 の自己相関係数を求めるステップと、前記第 3 の自己相関係数を用いて第 2 のスペクトルパラメータを算出するステップと、前記第 2 のスペクトルパラメータに基づいて聴覚重み特性を設定するステップと、前記第 1 のスペクトルパラメータおよび前記聴覚重み特性を用いて前記残差成分を符号化するステップと

10

を有することを特徴とする音声符号化方法。

【請求項 4】

入力音声信号をスペクトル包絡を表すスペクトルパラメータと残差成分とで表し、これらスペクトルパラメータおよび残差成分を符号化する音声符号化装置において、前記入力音声信号から所定の時間単位毎に第 1 の自己相関係数を算出する自己相関算出手段と、前記第 1 の自己相関係数に対し第 1 の自己相関窓を用いて窓掛けを行うことにより、前記第 1 の自己相関係数を修正して第 2 の自己相関係数を求める第 1 の窓掛け手段と、前記第 2 の自己相関係数を用いて第 1 のスペクトルパラメータを算出する第 1 のスペクトルパラメータ算出手段と、前記第 1 のスペクトルパラメータ算出手段により算出されたスペクトルパラメータを符号化するスペクトルパラメータ符号化手段と、前記第 1 の自己相関係数に対し前記第 1 の自己相関窓よりも前記第 1 の自己相関係数に与える修正の度合いが小さい、前記第 1 の自己相関係数とは形状の異なる第 2 の自己相関窓を用いて窓掛けを行うことにより、前記第 1 の自己相関係数を修正して第 3 の自己相関係数を求める第 2 の窓掛け手段と、前記第 1 の自己相関係数に対し形状が前記第 1 の自己相関窓と異なる第 2 の自己相関窓を用いて窓掛けを行うことにより、第 3 の自己相関係数を求める第 2 の窓掛け手段と、前記第 3 の自己相関係数を用いて第 2 のスペクトルパラメータを算出する第 2 のスペクトルパラメータ算出手段と、前記第 2 のスペクトルパラメータに基づいて聴覚重み特性を設定する聴覚重み特性設定手段と、前記第 1 のスペクトルパラメータおよび前記聴覚重み設定手段により設定された聴覚重み特性を用いて前記残差成分を符号化する残差成分符号化手段と

20

30

40

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、音声信号をスペクトルパラメータと残差成分とで表して低ビットレートで高効率に符号化する音声符号化方法および装置に関する。

【0002】

【従来の技術】

音声信号を少ない情報量で蓄積したり伝送するための低ビットレート符号化方式として、

50

CELP ( Code Excited Linear Prediction , M . R . Schroeder and B . S . Atal , “ Code Excited Linear Prediction ( CELP ) : High Quality Speech at Very Low Bit Rates ” , Proc . ICA SSP , pp . 937 - 940 , 1985 ( 文献 1 ) が知られている。また、「音のコミュニケーション工学」日本音響学会編、1996年、コロナ社(文献2)のpp. 33 - 42においても、CELP方式が解説されている。

【0003】

CELP方式は線形予測分析に基づく符号化方式であり、音声信号を音韻情報となるスペクトル包絡を表すスペクトルパラメータと、音の高さ等を表す残差成分とで表して、両者を符号化する。音声信号のスペクトル包絡を表すスペクトルパラメータとしては様々なものがあるが、音声符号化の分野ではLPC係数(線形予測係数)が最も一般的に使用される。

10

【0004】

CELP方式では、LPC係数は音声信号の自己相関係数に窓掛けを行うことで修正された自己相関係数から求められる。自己相関係数からLPC係数を求めるには、Levinson-Durbin algorithmやDurbinの再帰的解法として知られている方法が用いられる。この方法の詳細については、例えば「デジタル音声処理」東海大学出版会、古井貞氏著(文献3)のpp. 75に記載されている。このようにして求められたLPC係数は、符号化に適したLSP係数などの等価なパラメータに変換される(文献3のpp. 89 - 92参照)。そして、これが符号化されることによりスペクトルパラメータの符号が求められる。

20

【0005】

一方、残差成分の符号化に当たっては、聴覚的に符号化歪みが聞こえにくくなるように聴覚重み付きの歪み尺度により符号選択が行われる。CELP方式などの従来の音声符号化技術では、符号化される前のLPC係数を聴覚重み付けにも利用することが特徴である。

【0006】

音声信号の復号に当たっては、スペクトルパラメータの符号と残差成分の符号が復号され、復号化されたスペクトルパラメータに従って、復号化された残差成分にスペクトル包絡を与えることにより音声信号を再生する。

30

【0007】

このように従来の音声符号化技術では、符号化対象にすることを主目的に求められたLPC係数を聴覚重み特性の設定にも流用しており、この制約のために必ずしも十分な聴覚重み特性を表現することができない。従って、例えば4 k b i t / s程度以下の低ビットレート符号化に従来の音声符号化技術を用いると、符号化歪みへの影響が大きい残差成分の劣化を聴覚重み付けによってマスクしきれなくなり、高品質の復号音声を得ることができなくなる。

【0008】

【発明が解決しようとする課題】

上述したように従来の音声符号化技術では、符号化対象にすることを主目的に求められたLPC係数を聴覚重み特性の設定にも流用することから、必ずしも十分な聴覚重み特性を表現することができないため、さらなる低ビットレート化を図ろうとすると、符号化歪みが大きい残差成分の劣化を聴覚重み付けによってマスクしきれなくなり、復号音声の品質が劣化するという問題があった。

40

【0009】

本発明は、低ビットレート化を図りつつ符号化歪みが知覚されにくい音声符号化方法および装置を提供することを目的とする。

【0010】

【課題を解決するための手段】

上記の課題を解決するため、本発明は入力音声信号をスペクトル包絡を表すスペクトルパ

50

ラメータと残差成分とで表し、これらスペクトルパラメータおよび残差成分を符号化する音声符号化に際して、入力音声信号から求められた第1の自己相関係数を修正して得られた第2の自己相関係数からスペクトルパラメータを算出して符号化し、また第2の自己相関係数を得る際の第1の自己相関係数の修正の度合いよりも小さな修正の度合で第1の自己相関係数を修正して求められた第3の自己相関係数から聴覚重み特性を求め、これらスペクトルパラメータおよび聴覚重み特性を用いて残差成分を符号化することを基本的な特徴とする。

【0011】

ここで、自己相関係数の修正は、例えば自己相関窓を用いて行われる。第1の自己相関係数に対して自己相関窓を用いて窓掛けを行うことにより、修正された第2または第3の自己相関係数が得られる。この場合、第2の自己相関係数を得る際に用いる第1の自己相関窓と、第3の自己相関係数を得る際に用いる第2の自己相関窓を形状の異なるものとする。

10

【0012】

より具体的には、本発明では入力音声信号から所定の時間単位毎に第1の自己相関係数が算出される。第1の自己相関係数に対し、第1の窓掛け部で第1の自己相関窓を用いて窓掛けを行うことにより第2の自己相関係数が求められ、同様に第2の窓掛け部で形状が第1の自己相関窓と異なる第2の自己相関窓を用いて窓掛けを行うことにより第3の自己相関係数が求められる。

【0013】

第2の自己相関係数を用いて符号化対象となる第1のスペクトルパラメータが算出され、この第1のスペクトルパラメータが符号化される。一方、第3の自己相関係数を用いて別の第2のスペクトルパラメータが算出され、この第2のスペクトルパラメータから聴覚重み特性が設定され、第1のスペクトルパラメータおよび聴覚重み特性を用いて残差成分が符号化される。

20

【0014】

このような本発明によると、第1の自己相関窓を符号化対象となる第1のスペクトルパラメータ（例えばLPC係数）を求めるために最適化した形状とし、第2の自己相関窓を聴覚重み特性の設定に用いる第2のスペクトルパラメータを求めるために最適化した形状とすることにより、符号化対象の第1のスペクトルパラメータと聴覚重み特性のそれぞれを精度よく求めることができるようになる。従って、非常に低い符号化ビットレートでも、復号時に符号化歪みが知覚されにくく、高品質の復号音声を再生できる音声符号化が可能となる。

30

【0015】

【発明の実施の形態】

以下、図面を参照して本発明の実施の形態を説明する。

【0016】

(第1の実施形態)

図1は、本発明の第1の実施形態に係る音声符号化装置の構成を示すブロック図である。この音声符号化装置は自己相関算出部101、第1窓掛け部102、第2窓掛け部103、符号化用スペクトルパラメータ算出部104、聴覚重み用スペクトルパラメータ算出部105、スペクトルパラメータ符号化部106、聴覚重み設定部107、残差成分符号化部108および多重化部109からなる。

40

【0017】

自己相関算出部101では、所定のサンプリング周波数でサンプリングされデジタル化された入力音声信号から、所定の時間単位毎に次式で示すように第1の自己相関係数 $r_i$  ( $r_0, r_1, \dots, r_N$ )が算出される。

【0018】

【数1】

$$r_i = \sum_{n=0}^{L-1-i} x_n x_{n+i} \quad (1)$$

## 【0019】

ここで、 $\{x_n\}$ は入力音声信号に長さLの時間窓をかけて切り出した入力音声信号系列、Nは自己相関の次数をそれぞれ表しており、入力音声信号のサンプリング周波数が8k Hzの場合、典型的なNの値はN=10である。 10

## 【0020】

次に、自己相関算出部101で求められた第1の自己相関係数 $r_i$ に対して、第1窓掛け部102で第1の自己相関窓による窓掛けが行われることにより、第1の自己相関係数 $r_i$ が修正され、第2の自己相関係数 $i(0, 1, \dots, N)$ が求められる。この第1の自己相関窓による窓掛け処理の一例は、次式に示される。

$$i = r_i \times w_i \quad (i = 0, 1, \dots, N) \quad (2)$$

ここで、 $w_i$ は第1の自己相関窓を表す。

## 【0021】

次に、符号化用スペクトルパラメータ算出部104において、第2の自己相関係数 $i$ を用いて符号化対象となるスペクトルパラメータが求められる。スペクトルパラメータとしてはパワースペクトル、LPCケプストラム、メル尺度のスペクトルパラメータ、サブバンドエネルギーなど様々なものが知られているが、ここではLPC係数(線形予測係数)の例について示す。LPC係数は、次の線形方程式を解くことにより算出される。 20

$$= \quad (3)$$

ここで、は次式に示すように、第2の自己相関係数 $i$ から構成される自己相関行列である。

## 【0022】

## 【数2】

$$\Phi = \begin{bmatrix} \phi_0 & \phi_1 & \dots & \phi_{N-1} \\ \phi_1 & \phi_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \phi_1 \\ \phi_N & \dots & \phi_1 & \phi_0 \end{bmatrix}$$

$$\alpha = [\alpha_1, \alpha_2, \dots, \alpha_N]^T$$

$$\psi = [\phi_1, \phi_2, \dots, \phi_N]^T$$

(4)

30

40

## 【0023】

式(3)の方程式からLPC係数 $\{i\}$ を求める方法としては、例えばLevinson-Durbin algorithmや、Durbinの再帰的解法を用いることができ、これらの方法は先の文献3のpp.75に記載されているので、詳細な説明は省略する。

## 【0024】

こうして求められた符号化対象となるスペクトルパラメータ(以下、符号化用スペクトルパラメータという)(ここではLPC係数 $\{i\}$ )は、スペクトルパラメータ符号化部 50

106によって符号化され、量子化されたスペクトルパラメータとこれを表すスペクトルパラメータの符号Aが出力される。

【0025】

このスペクトルパラメータの符号化に当たっては、例えばスペクトルパラメータがLPC係数である場合、LPC係数をLSP(線スペクトル対)係数などの等価なパラメータに変換し(文献3のpp.89-92参照)、これをベクトル量子化法を用いて符号化することにより、同一ビット数の下でより量子化歪みの少ないスペクトルパラメータの符号化を行うことができる。

【0026】

次に、自己相関係数から聴覚重み特性を設定するために、まず自己相関算出部101で求められた第1の自己相関係数 $r_i$ に対して第2窓掛け部103で第2の自己相関窓による窓掛けが行われることにより、第1の自己相関係数 $r_i$ が修正され、第3の自己相関係数 $i(0, 1, \dots, N)$ が求められる。この第2の自己相関窓による窓掛け処理の一例は、次式に示される。

$$i = r_i \times v_i \quad (i = 0, 1, \dots, N) \quad (5)$$

ここで、 $v_i$ は第2の自己相関窓を表す。

【0027】

第2の自己相関窓 $v_i$ は、聴覚重み特性を設定するために用いる窓であり、第1の自己相関窓 $w_i$ とは形状が異なる。より具体的には、第1の自己相関窓 $w_i$ よりも第2の自己相関窓 $v_i$ の方が自己相関係数に与える修正の度合いが小さくなるような関係になるように、これら2つの自己相関窓 $w_i, v_i$ を設定することが望ましい。この理由は、次の通りである。

【0028】

まず、第1の自己相関窓 $w_i$ については、これを用いて最終的に符号化用スペクトルパラメータ算出部104で求められる符号化用スペクトルパラメータは、スペクトルパラメータ符号化部106内で量子化された後に音声信号を生成するための合成フィルタのフィルタ特性として用いられるので、周波数特性に過度に強い共振点を持たないように、自己相関係数に対する修正の度合いが比較的強い窓形状とすることが望ましい。

【0029】

一方、第2の自己相関窓 $v_i$ は、最終的に音声信号のスペクトルの形状に対応した周波数マスキング効果を反映するための聴覚重み特性の設定に使用することから、過度に強い共振点を持たないようにすることが必要ではあるが、合成フィルタのフィルタ特性として用いられることはないので、第1の自己相関窓 $w_i$ よりも自己相関係数に対する修正の度合いは小さくなるような窓形状とすることが望ましい。

【0030】

次に、聴覚重み用スペクトルパラメータ算出部105において、第2窓掛け部103で得られた第3の自己相関係数 $i$ を用いて聴覚重みの設定に必要なスペクトルパラメータ(以下、聴覚重み用スペクトルパラメータという)が求められる。この聴覚重み用スペクトルパラメータとして、LPC係数を用いる場合には、前述の符号化用パラメータとして用いられるLPC係数の算出方法と同じアルゴリズムを用いることができることは言うまでもない。このときの線形方程式は、次のようになる。

$$= \quad (6)$$

ここで、 $i$ は次式に示すように、第3の自己相関係数 $i$ から構成される自己相関行列である。

【0031】

【数3】

10

20

30

40

$$\Phi' = \begin{bmatrix} \phi'0 & \phi'1 & \cdots & \phi'N-1 \\ \phi'1 & \phi'0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \phi'1 \\ \phi'N & \cdots & \phi'1 & \phi'0 \end{bmatrix}$$

$$\beta = [\beta_1, \beta_2, \dots, \beta_N]^T$$

$$\psi' = [\phi'1, \phi'2, \dots, \phi'N]^T$$

(7)

10

## 【0032】

第2の自己相関窓  $v_i$  は第1の自己相関窓  $w_i$  と異なるため、第2の自己相関窓  $v_i$  により修正される第3の自己相関係数  $i$  を用いて算出される聴覚重み用スペクトルパラメータとしてのLPC係数  $\{i\}$  は、第1の自己相関窓  $w_i$  により修正された符号化用スペクトルパラメータとしてのLPC係数  $\{i\}$  とはスペクトルの特性が異なる。従って、第2の自己相関窓  $v_i$  を聴覚重み用に適切に設定することによって、より精度の高い聴覚重み特性を残差成分の符号化に用いることができるという効果がある。

20

## 【0033】

聴覚重み設定部107は、聴覚重み用スペクトルパラメータ（この例ではLPC係数  $\{i\}$ ）を用いて、残差成分符号化部108での聴覚重み付けに用いる聴覚重み特性の設定を行う。残差成分符号化部108において時間領域で聴覚重み付けを行って残差成分の符号化を行う場合には、聴覚重み付けは  $W(z)$  なる特性の重みフィルタによるフィルタリング処理として実現される。LPC係数  $\{i\}$  を用いた聴覚重み付けフィルタ特性  $W(z)$  の典型的な例は、次式に示される。

## 【0034】

【数4】

$$W(z) = \frac{B(z/\gamma_1)}{B(z/\gamma_2)} \quad (8)$$

30

## 【0035】

ここで、 $B(z)$  は次式となる。

## 【0036】

【数5】

$$B(z) = 1 + \sum_{i=1}^N \beta_i z^{-i} \quad (9)$$

40

## 【0037】

$\beta_1, \beta_2$  は残差成分符号化部108での聴覚重み付け特性を設定するパラメータであり、 $\beta_1 > \beta_2 > 0$  の関係が必要である。典型的な例としては、例えば  $\beta_1 = 0.94$ 、 $\beta_2 = 0.6$  を用いることができる。

## 【0038】

50

残差成分符号化部 108 は、入力音声信号とスペクトルパラメータ符号化部 106 からの量子化されたスペクトルパラメータと聴覚重みの情報を入力し、量子化されたスペクトルパラメータと共に音声信号を表すに必要な残差成分の符号化を行い、得られた残差成分の符号 B を出力する。

【0039】

以上のようにしてスペクトルパラメータ符号化部 106 で得られたスペクトルパラメータの符号 A と、残差成分符号化部 108 で得られた残差成分の符号 B は多重化部 109 で多重化され、入力音声信号を表す符号化データとして出力される。この符号化データは、蓄積系または伝送系へ送出される。

【0040】

次に、図 2 に示すフローチャートを用いて、本実施形態による音声符号化装置と同様の音声符号化処理をソフトウェアにより実現する場合の処理手順を説明する。

【0041】

まず、入力音声信号から所定の時間単位毎に第 1 の自己相関係数  $r_i (r_0, r_1, \dots, r_N)$  を求める (ステップ S1)。

次に、この自己相関係数  $r_i$  に第 1 の自己相関窓  $w_i (w_0, w_1, \dots, w_N)$  による窓掛けを行い、修正された第 2 の自己相関係数  $i (0, 1, \dots, N)$  を求める (ステップ S2)。

次に、第 2 の自己相関係数  $i$  を用いて符号化対象となる符号化用スペクトルパラメータを求める (ステップ S3)。

次に、符号化用スペクトルパラメータを符号化し、この符号化の過程で得られる量子化されたスペクトルパラメータとそれを表すスペクトルパラメータの符号を求める (ステップ S4)。

【0042】

一方、ステップ S1 で求められた第 1 の自己相関係数  $r_i$  から、聴覚重み特性を設定するまでの処理を以下のようにして行う。

すなわち、自己相関係数  $r_i$  に第 2 の自己相関窓  $v_i (v_0, v_1, \dots, v_N)$  による窓掛けを行い、修正された第 3 の自己相関係数  $i (0, 1, \dots, N)$  を求める (ステップ S5)。

次に、第 3 の自己相関係数  $i$  を用いて、聴覚重みの設定に必要な聴覚重み用スペクトルパラメータを求める (ステップ S6)。

次に、この聴覚重み用スペクトルパラメータを用いて残差成分符号化で用いる聴覚重み特性の設定を行う (ステップ S7)。

次に、入力音声信号と量子化されたスペクトルパラメータと聴覚重み特性の情報を用いて、量子化されたスペクトルパラメータと共に音声信号を表すために必要な残差成分の符号化を行う (ステップ S8)。

そして、ステップ S4 と S8 の処理により得られたスペクトルパラメータの符号と残差成分の符号を多重化して音声信号の符号データとして出力する (ステップ S9)。

【0043】

以上のステップ S1 ~ S9 の処理を終えると、一つの時間単位 (典型的には、入力音声信号が 8 kHz でサンプリングされているとき 20 msec) の音声信号の符号化処理が終了する。この一連の処理をステップ S10 で次の時間単位の処理を行わないと判定されるまで時間単位毎に連続して行うことにより、連続して入力される音声信号の符号化を行うことができる。

【0044】

(第 2 の実施形態)

図 3 は、本発明を CELP 方式に適用した音声符号装置の構成を示すブロック図である。この図では、特に CELP 方式の特徴である残差成分符号化部について、図 1 よりも詳細に示している。CELP 方式の詳細については、前述したように文献 1 や文献 2 に記載されている。

10

20

30

40

50



## 【 0 0 4 5 】

この音声符号化装置は自己相関算出部 3 0 1、第 1 窓掛け部 3 0 2、第 2 窓掛け部 3 0 3、符号化用 L P C 係数算出部 3 0 4、聴覚重み用 L P C 係数算出部 3 0 5、L P C 係数符号化部 3 0 6、聴覚重み設定部 3 0 7、残差成分符号化部 3 0 8 および多重化部 3 0 9 からなる。

## 【 0 0 4 6 】

ここで、自己相関算出部 3 0 1、第 1 窓掛け部 3 0 2、第 2 窓掛け部 3 0 3、符号化用 L P C 係数算出部 3 0 4、聴覚重み用 L P C 係数算出部 3 0 5、L P C 係数符号化部 3 0 6 および聴覚重み設定部 3 0 7 については、第 1 の実施形態における自己相関算出部 1 0 1、第 1 窓掛け部 1 0 2、第 2 窓掛け部 1 0 3、符号化用スペクトルパラメータ算出部 1 0 4、聴覚重み用スペクトルパラメータ算出部 1 0 5、スペクトルパラメータ符号化部 1 0 6 および聴覚重み設定部 1 0 7 と同様であるので、説明を省略する。

## 【 0 0 4 7 】

残差成分符号化部 3 0 8 は目標信号生成部 3 1 1、適応音源符号化部 3 1 2、雑音音源符号化部 3 1 3、ゲイン符号化部 3 1 4、駆動信号生成部 3 1 5 および重み付き合成フィルタ 3 1 6 から構成される。以下、残差成分符号化部 3 0 8 の各部の構成について詳細に説明する。

## 【 0 0 4 8 】

目標信号生成部 3 1 1 は、聴覚重み設定部 3 0 7 により聴覚重み特性が設定される聴覚重みフィルタを有し、この聴覚重みフィルタを用いて入力音声信号に対しフィルタリングを行うことにより聴覚重み付けがなされた音声信号を生成すると共に、この聴覚重み付けがなされた音声信号から前の時間単位での符号化の影響を差引くことにより、残差成分の符号化の目標となる目標信号  $\{ f_n \}$  を生成する。

## 【 0 0 4 9 】

適応音源符号化部 3 1 2 は、C E L P 方式の音声符号化でよく知られている適応符号帳を有し、目標信号  $\{ f_n \}$  (目標ベクトル  $f$ ) を用いて次式の誤差ベクトル  $e_0$  の大きさをより小さくする、好ましくは最小化する最適な適応符号ベクトル  $c_0$  を適応符号帳の中から探索する。

$$e_0 = f - H w c_0(i) \quad (10)$$

ここで、 $i$  は適応符号ベクトルの候補となるコードベクトルのインデックスを示す。また、 $H w$  は聴覚重み付けられた音声のスペクトル包絡特性 (聴覚重み付けられた合成フィルタの特性)  $H w(z)$  を有するフィルタのインパルス応答から構成されるインパルス応答行列である。

## 【 0 0 5 0 】

聴覚重み付けられたスペクトル包絡特性  $H w(z)$  は、次式で表される。

## 【 0 0 5 1 】

## 【 数 6 】

$$Hw(z) = \frac{W(z)}{Aq(z)} = \frac{1}{Aq(z)} \cdot \frac{B(z/\gamma_1)}{B(z/\gamma_2)} \quad (11)$$

## 【 0 0 5 2 】

ここで、 $W(z)$  は式 (4) に示した聴覚重みフィルタ特性、また  $Aq(z)$  は次式で表される。

## 【 0 0 5 3 】

## 【 数 7 】

10

20

30

40

$$Aq(z) = 1 + \sum_{i=1}^N \alpha_{qi} z^{-i} \quad (12)$$

【 0 0 5 4 】

ただし、 $q_i$  は量子化された L P C 係数である。

【 0 0 5 5 】

こうして適応符号ベクトルの候補の中から選択された適応符号ベクトルのインデックス I と、これに対応する適応符号ベクトル  $c_0(I)$  が適応音源符号化部 3 1 2 から出力される。

【 0 0 5 6 】

次に、雑音音源符号化部 3 1 3 においては、C E L P 方式の音声符号化でよく知られている所定の方法で構成される雑音符号帳または擬似的に雑音を表現することのできるパルス音源等を用いて、適応音源符号化部 3 1 2 で表しきれなかった成分の符号化を行う。この際に用いる目標ベクトル  $d$  は、 $d = f - c_0(I)$  とすることができる。この目標ベクトル  $d$  を用いて、次式の誤差ベクトル  $e_1$  の大きさをより小さくする、好ましくは最小化する最適な雑音符号ベクトル  $c_1$  を雑音符号ベクトル候補の中から探索する。

$$e_1 = d - H w c_1(j) \quad (13)$$

ここで、 $j$  は雑音符号ベクトルの候補となる符号ベクトルのインデックスを示す。

【 0 0 5 7 】

こうして雑音符号ベクトルの候補の中から選択された雑音符号ベクトルのインデックス J と、これに対応する雑音符号ベクトル  $c_1(J)$  が雑音音源符号化部 3 1 1 から出力される。

【 0 0 5 8 】

次に、ゲイン符号化部 3 1 4 は、C E L P 方式の音声符号化でよく知られている所定の方法で構成されるゲイン符号化帳を有し、適応音源符号化部 3 1 2 から出力される適応符号ベクトル  $c_0(I)$  と雑音音源符号化部 3 1 3 から出力される雑音符号ベクトル  $c_1(J)$  にそれぞれ乗じるためのゲインを符号化する。符号化に際しては、次式に示す誤差ベクトル  $e_g$  の大きさをより小さくする、好ましくは最小化する最適なゲインをゲイン符号化帳に格納されたゲインベクトルの候補  $g_0(k)$ 、 $g_1(k)$  (ただし、 $k$  はゲインベクトルのインデックス)の中から探索する。

$$e_g = f - g_0(k) H w c_0(I) - g_1(k) H w c_1(J) \quad (14)$$

こうしてゲインベクトルの候補  $g_0(k)$ 、 $g_1(k)$ の中から探索されたゲインのインデックス K と、これに対応するゲインベクトル  $g_0(K)$ 、 $g_1(K)$  がゲイン符号化部 3 1 4 から出力される。

【 0 0 5 9 】

適応音源符号化部 3 1 2 から出力される適応符号化ベクトル  $c_0(I)$ 、雑音音源符号化部 3 1 3 から出力される雑音符号ベクトル  $c_1(J)$  およびゲイン符号化部 3 1 4 から出力されるゲインベクトル  $g_0(K)$ 、 $g_1(K)$  は、駆動信号生成部 3 1 5 に入力される。駆動信号生成部 3 1 5 は、次式に示すように適応符号ベクトル  $c_0(I)$ 、雑音符号ベクトル  $c_1(J)$  をそれぞれにゲインベクトル  $g_0(K)$ 、 $g_1(K)$  を乗じた後に加算することにより、量子化された残差ベクトル  $e_x$  を求める。この残差ベクトル  $e_x$  は、適応音源符号化部 3 1 2 に入力されて適応符号帳に格納されるとともに、重み付き合成フィルタ 3 1 6 に駆動信号として入力される。

【 0 0 6 0 】

$$e_x = g_0(K) c_0(I) + g_1(K) c_1(J) \quad (15)$$

そして最後に、残差ベクトル  $e_x$  と重み付き合成フィルタの特性  $W(z)$  および  $A_q(z)$  を用いて、入力音声信号の次の時間単位の符号化に及ぶ影響を求めるための重み付き合成フィルタの内部状態を求め、これを目標信号生成部 311 に供給する。

【0061】

最後に、以上のようにして得られたスペクトルパラメータ (LPC 係数) の符号 A と、図 1 における残差成分の符号 B に相当する適応符号ベクトルのインデックス I、雑音符号ベクトルのインデックス J およびゲインベクトルのインデックス K が多重化部 309 で多重化され、入力音声信号を表す符号化データとして出力される。この符号化データは、蓄積系または伝送系に送出される。

【0062】

次に、本実施形態に係る音声復号化装置について説明する。図 4 は、同実施形態に係る図 3 に示した音声符号化装置に対応する音声復号化装置の構成を示すブロック図である。

【0063】

本発明は、基本的に符号化側におけるスペクトルパラメータ (例えば LPC 係数) および残差成分の抽出法に特徴を有するものであり、図 3 に示した音声符号化装置から出力される符号化データそのものは、従来の CELP 方式のそれと基本的に変わらない。従って、音声復号化装置の構成は、従来の CELP 方式のそれと同様でよい。

【0064】

図 4 に示す音声復号化装置は、分離部 400、LPC 係数復号化部 401、適応音源復号化部 402、雑音音源復号化部 403、ゲイン復号化部 404、駆動信号生成部 405、合成フィルタ 406 およびポストフィルタ 407 から構成される。

【0065】

分離部 400 では、図 3 に示した音声符号化装置より蓄積系または伝送系を経て入力された符号化データから、スペクトルパラメータ (LPC 係数) の符号 A と、残差成分の符号に相当する適応符号ベクトルのインデックス I、雑音符号ベクトルのインデックス J およびゲインベクトルのインデックス K が分離され、それぞれ LPC 係数復号化部 401、適応音源復号化部 402、雑音音源復号化部 403 およびゲイン復号化部 404 に入力される。

【0066】

LPC 係数復号化部 401 では、音声符号化装置と同様にしてスペクトルパラメータ A の符号に対応する量子化された LPC 係数を再生し、これを合成フィルタ 406 およびポストフィルタ 407 に供給する。

【0067】

適応音源復号化部 402 は、図 3 の適応音源符号化部 312 と同様に適応符号帳を有し、インデックス I に対応する適応符号ベクトル  $c_0(I)$  を求めて駆動信号生成部 405 に供給する。雑音音源復号化部 403 は、図 3 の雑音音源符号化部 313 と同様に雑音符号帳を有し、インデックス J に対応する雑音符号ベクトル  $c_1(J)$  を求めて駆動信号生成部 405 に供給する。さらに、ゲイン復号化部 404 は、図 3 のゲイン符号化部 314 と同様にゲイン符号帳を有し、インデックス K に対応するゲインベクトル  $g_0(K)$ 、 $g_1(K)$  を求めて駆動信号生成部 405 に供給する。

【0068】

駆動信号生成部 405 は、図 3 の駆動信号生成部 315 と同様に式 (15) に従って適応符号ベクトル  $c_0(I)$ 、雑音符号ベクトル  $c_1(J)$  およびゲインベクトル  $g_0(K)$ 、 $g_1(K)$  から量子化された残差ベクトル  $e_x$  を求める。この残差ベクトル  $e_x$  は、適応音源復号化部 402 に入力されて適応符号帳に格納されるとともに、合成フィルタ 406 に駆動信号として入力される。

【0069】

合成フィルタ 406 は、LPC 係数復号化部 401 で求められた量子化された LPC 係数  $q_i$  を用いて式 (12) と逆特性  $1/A_q(z)$  のフィルタリングを駆動信号 (残差ベクトル  $e_x$ ) に対して行うことにより、復号された音声信号を合成する。この合成フィル

10

20

30

40

50

タ 4 0 6 の出力信号は、L P C 係数復号化部 4 0 1 で求められた量子化された L P C 係数  $q_i$  を用いて特性が設定されたポストフィルタ 4 0 7 によってスペクトル形状が強調されることにより、最終的な復号音声信号が生成される。

【 0 0 7 0 】

以上、本発明の実施形態について説明したが、本発明はこれに限られるものでなく、種々変形して実施することができる。

例えば、上記の実施形態においては、自己相関係数の修正法として自己相関係数に自己相関窓を乗じて窓掛け処理を行う方法を例にとり説明したが、自己相関係数の修正手法はこれに限られるものではない。要するに、符号化の対象とするスペクトルパラメータに供する自己相関係数と、聴覚重み特性の設定に供する自己相関係数がそれぞれに適した異なる条件で修正され、共通の自己相関係数に端を発して求められる手法であればよい。

10

【 0 0 7 1 】

また、自己相関係数の定義に上記実施形態での説明と多少違いがある場合や、自己相関係数の代りに正規化自己相関係数を用いた場合についても、本発明を適用できることは言うまでもない。

【 0 0 7 2 】

【発明の効果】

以上説明したように、本発明では入力音声信号から求められた第 1 の自己相関係数をそれぞれ異なる条件で修正して得られた第 2、第 3 の自己相関係数をそれぞれ用いて、符号化対象のスペクトルパラメータと残差成分の符号化に用いる聴覚重み特性を個別に求めることによって、符号化対象のスペクトルパラメータおよび聴覚重み特性をいずれも精度よく求めることが可能となる。

20

【 0 0 7 3 】

従って、本発明によると 4 k b i t / s 程度以下というような低ビットレート符号化においても、符号化歪みが知覚されにくい高品質の復号音声を得られる音声符号化を実現することができる。

【図面の簡単な説明】

【図 1】本発明の第 1 の実施形態に係る音声符号化装置の構成を示すブロック図

【図 2】同第 2 の実施形態に係る音声符号化の処理手順を示すフローチャート

【図 3】本発明の第 2 の実施形態に係る音声符号化装置の構成を示すブロック図

30

【図 4】同第 2 の実施形態に係る音声符号化装置に対応する音声復号化装置の構成を示すブロック図

【符号の説明】

1 0 1 , 3 0 1 ... 自己相関算出部

1 0 2 , 3 0 2 ... 第 1 窓掛け部

1 0 3 , 3 0 3 ... 第 2 窓掛け部

1 0 4 , 3 0 4 ... 符号化用スペクトルパラメータ算出部

1 0 5 , 3 0 5 ... 聴覚重み用スペクトルパラメータ算出部

1 0 6 ... スペクトルパラメータ符号化部

3 0 6 ... L P C 係数符号化部

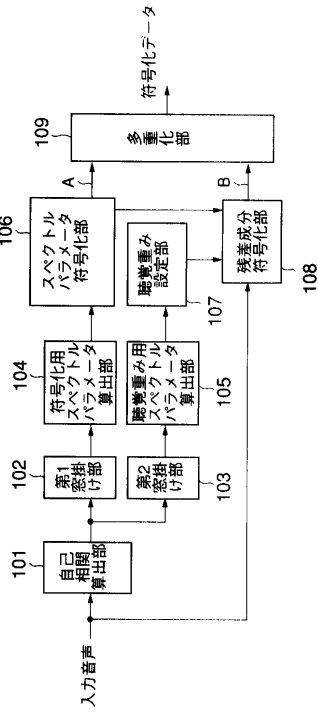
40

1 0 7 , 3 0 7 ... 聴覚重み設定部

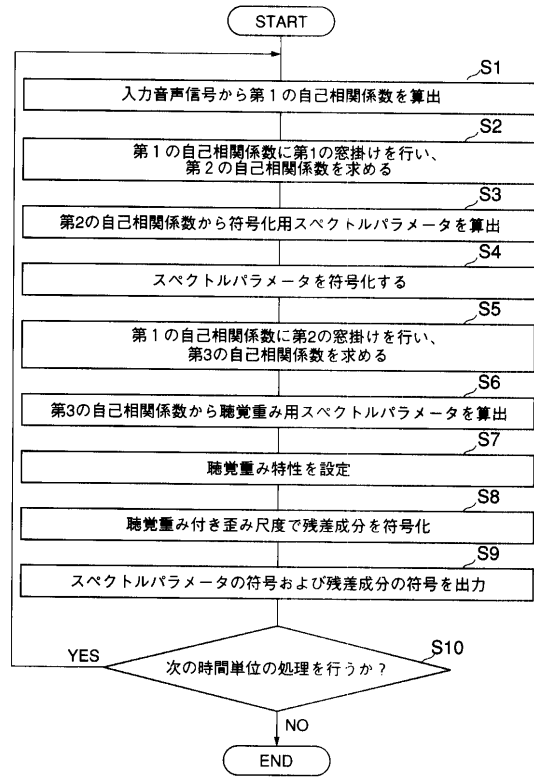
1 0 8 , 3 0 8 ... 残差成分符号化部

1 0 9 , 3 0 9 ... 多重化部

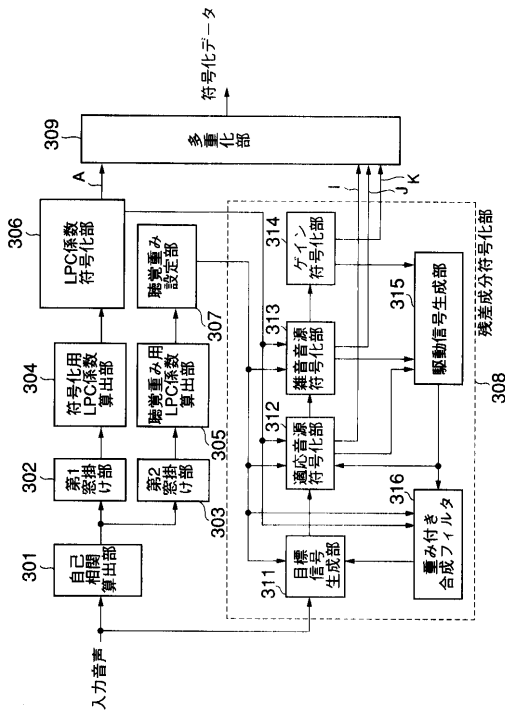
【 図 1 】



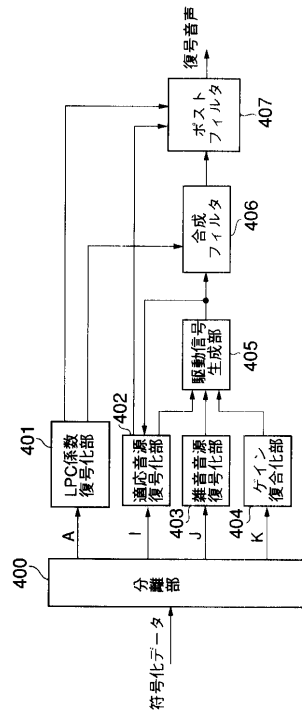
【 図 2 】



【 図 3 】



【 図 4 】



---

フロントページの続き

(74)代理人 100070437

弁理士 河井 将次

(72)発明者 三関 公生

兵庫県神戸市東灘区本山南町8丁目6番26号 株式会社東芝関西研究センター内

(72)発明者 押切 正浩

兵庫県神戸市東灘区本山南町8丁目6番26号 株式会社東芝関西研究センター内

審査官 渡邊 聡

(56)参考文献 特開平10-143198(JP,A)

特開平09-230896(JP,A)

特開2002-341889(JP,A)

(58)調査した分野(Int.Cl.<sup>7</sup>, DB名)

JICSTファイル(JOIS)

G10L 19/00

G10L 19/04

G10L 19/06

G10L 19/12

H03M 7/30