

(19)日本国特許庁(JP)

(12)公開特許公報 ( A )

(11)特許出願公開番号

特開昭59 - 94797

(43)公開日 昭和59年(1984)5月31日

(51)Int. Cl. <sup>5</sup>	識別記号	庁内整理番号	F I	技術表示箇所
G 1 0 L		1/00		
H 0 4 B		1/66		
H 0 4 B		12/02		

審査請求 有 請求項の数2

(全6頁)

(21)出願番号 特願昭57-204850

(22)出願日 昭和57年(1982)11月22日

(71)出願人 999999999  
日本電信電話公社  
東 京

(72)発明者 守谷 健弘  
\*

(72)発明者 誉田 雅彰  
\*

(54)【発明の名称】音声の適応変換符号化方式

## (57)【要約】

〔目的〕入力音声信号のスペクトルをブロック単位に分割し、適応的情報割当てを行うことにより、音声品質の劣化を少なくして、量子化効率を高める。

〔構成〕入力信号は、直交変換部12で1フレームを単位にスペクトル変換され、平滑部17において、別に求められ、かつ量子化されたスペクトル包絡の情報で大域的に平坦化され、分割部21でp個ずつのブロックに分割される。なお、スペクトル包絡情報は、入力音声の線形予測分析によるスペクトル包絡の推定14後に、量子化18され、復号19されて得られる。また、補助情報の復号19出力に応じてブロックごとにビット数を割当23で、ベクトル量子化22を行う。そして補助情報と量子化22された波形情報は、合成されて復号器25へ送られる。

## 【特許請求の範囲】

【請求項1】(1) 音声信号サンプル値系列を一定の個数ごとに1フレームとし、1フレームごとに直交変換によりスペクトルを求め、適応的に量子化する符号化方式において、上記スペクトルの包絡を求めてパワとともに量子化して補助情報として符号化するスペクトル包絡抽出手段と、その補助情報を復号する局部復号化手段と、その復号された補助情報を使って上記スペクトルを周波数軸上で平坦化されたスペクトル信号系列とする平滑手

段と、その平坦化されたスペクトル復号系列を周波数軸上でブロックに分割するブロック分割手段と、その分割された各ブロックに対して上記補助情報を使って適応的に情報割当てを行う適応情報割当て手段と、その割当てによって上記分割されたスペクトル信号系列をベクトル量子化するベクトル量子化手段とを備えた音声の適応変換符号化方式。

【請求項2】(2) 音声信号のサンプル値系列を1フレームを単位として分析して符号化する符号化方式において上記音声信号を線形予測分析し、そのスペクトル包絡を求め、パワとともに量子化し補助情報として符号化する線形予測分析手段と、その補助情報を復号する局部復号化手段と、その復号された補助情報の線形予測係数によりフィルタ定数が制御されると共に上記音声信号のサンプル値系列が入力されて残差信号を出力する逆フィルタ手段と、その残差信号を1フレームごとに直交変換してスペクトルを求める直交変換手段と、このスペクトルを周波数軸上でブロックに分割するブロック分割手段と、その分割された各ブロックに対して上記補助情報を使って適応的に情報割当てを行う適応情報割当て手段と、その割当てに従って上記分割されたスペクトル残差信号系列をベクトル量子化するベクトル量子化手段とを備えた音声の適応変換符号化方式。

⑬ 日本国特許庁 (JP)

⑩ 特許出願公開

⑫ 公開特許公報 (A)

昭59-94797

⑨ Int. Cl. <sup>3</sup>	識別記号	庁内整理番号	⑭ 公開
G 10 L 1/00		7350-5D	昭和59年(1984)5月31日
H 04 B 1/66		7406-5K	発明の数 2
12/02		7830-5K	審査請求 有

(全 6 頁)

⑮ 音声の適応変換符号化方式

武蔵野市緑町3丁目9番11号日  
本電信電話公社武蔵野電気通信  
研究所内

⑯ 特 願 昭57-204850

⑰ 発 明 者 菅田雅彰

⑱ 出 願 昭57(1982)11月22日

武蔵野市緑町3丁目9番11号日  
本電信電話公社武蔵野電気通信  
研究所内

特許法第30条第1項適用 昭和57年10月20日  
社団法人日本音響学会発行の「日本音響学会  
昭和57年度秋季研究発表会講演論文集1」に  
おいて発表

⑲ 出 願 人 日本電信電話公社

⑳ 発 明 者 守谷健弘

㉑ 代 理 人 弁理士 草野卓

明 細 書

1. 発明の名称

音声の適応変換符号化方式

2. 特許請求の範囲

(1) 音声信号のサンプル値系列を一定の個数ごとに1フレームとし、1フレームごとに直交変換によりスペクトルを求め、適応的に量子化する符号化方式において、上記スペクトルの包絡を求めてパワとともに量子化して補助情報として符号化するスペクトル包絡抽出手段と、その補助情報を復号する局部復号化手段と、その復号された補助情報を使つて上記スペクトルを周波数軸上で平坦化されたスペクトル信号系列とする平滑手段と、その平坦化されたスペクトル復号系列を周波数軸上でブロックに分割するブロック分割手段と、その分割された各ブロックに対して上記補助情報を使つて適応的に情報割当を行う適応情報割当手段と、その割当によつて上記分割されたスペクトル信号系列をベクトル量子化するベクトル量子化手段とを備えた音声の適応変換符号化方式。

(2) 音声信号のサンプル値系列を1フレームを単位として分析して符号化する符号化方式において上記音声信号を線形予測分析し、そのスペクトル包絡を求め、パワとともに量子化し補助情報として符号化する線形予測分析手段と、その補助情報を復号する局部復号化手段と、その復号された補助情報の線形予測係数によりフィルタ定数が制御されると共に上記音声信号のサンプル値系列が入力されて残差信号を出力する逆フィルタ手段と、その残差信号を1フレームごとに直交変換してスペクトルを求める直交変換手段と、このスペクトルを周波数軸上でブロックに分割するブロック分割手段と、その分割された各ブロックに対して上記補助情報を使つて適応的に情報割当を行う適応情報割当手段と、その割当に従つて上記分割されたスペクトル残差信号系列をベクトル量子化するベクトル量子化手段とを備えた音声の適応変換符号化方式。

3. 発明の詳細な説明

この発明は音声信号を周波数領域に変換し、そ

の量子化を適応的に変化させる適応変換符号化方式に関する。

#### <従来技術>

この種の音声符号化方式は例えば“特開昭55-57900号「音声信号処理回路」”に示されている。この方式は第1図に示すように、入力端子11よりの入力音声は例えば8 KHzでサンプリングされ、各サンプル値がデジタル信号として直交変換部12に輸入され、直交変換部12は例えば第2図Aに示す一定数の入力音声サンプル $S_1 \dots S_2 \dots S_n$ を離散的フーリエ変換により周波数領域の信号(スペクトル) $f_1 f_2 \dots f_n$ (第2図B)に変換されて適応量子化部13へ送られる。一方端子11の入力音声はスペクトル包絡抽出部14に輸入され、入力音声のスペクトルの包絡が線形予測分析により推定され、このスペクトル包絡及びピッチ周期は適応情報割当部15に供給される。適応情報割当部15は周波数領域の信号 $f_1 f_2 \dots f_n$ のそれぞれにおけるスペクトル包絡の瞬時レベルに応じて、このレベルが大きいければ割当

し、その単位でベクトル量子化をスペクトル包絡情報に応じて適応的に行うことにより、例えば符号化速度が9.6 Kbps以下においても音声品質の劣化を少なくするようにした音声の適応変換符号化方式を提供することにある。またスペクトルをブロック分割する前にスペクトルは平坦化しておくことによりベクトル量子化を効率的に行うことができる。

#### <第1実施例>

第3図はこの発明による音声符号化方式の実施例を示す。端子11からの入力信号は直交変換部12で1フレームを単位に、離散的フーリエ変換(DFT)、離散的余弦変換(DCT)などの直交変換により周波数領域の信号、即ちスペクトルに変換され、このスペクトルはスペクトル平滑部17において、別に求められ、量子化されたスペクトル包絡の情報で大域的に平坦化される。即ち端子11の入力音声はスペクトル包絡抽出部14において線形予測分析によりスペクトル包絡が推定され、このスペクトル包絡情報及び音声パワは

量子化部13における各信号 $f_1 f_2 \dots f_n$ に対する量子化ビットを適応的に変化する。このようにして量子化された情報と、ビット割当てを示す情報とが合成回路16で合成されて符号化出力として送られる。

この手法によつて8 KHzサンプリングの音声信号を16 Kbps程度の情報量で能率よく符号化でき、高品質の音声を得られる。しかし、ビット割当て情報に2 Kbps程度の情報量が必要であるため、全体で9.6 Kbps(一般に用いられている伝送速度の1つ)以下の情報量で符号化する際には信号 $f_1 f_2 \dots f_n$ を1ビット/サンプル以下で量子化する必要がある。この際、周波数の成分中の強さの小さい区間にはほとんど情報を割当てることができず、音声品質の大きな劣化を招く。

#### <発明の概要>

この発明は入力音声信号を周波数領域に変換し、その変換されたスペクトルをブロック単位に分割

量子化部18で補助情報として量子化され、この量子化出力は局部復号部19で復号され、その復号された補助情報によりスペクトル $f_1 f_2 \dots f_n$ がスペクトル平滑部17で割算される。

この平坦化されたスペクトルはブロック分割部21で第4図に示すように連続する $p$ 個ずつのブロックに $F_1=(f_{11} f_{12} \dots f_{1p}), F_2=(f_{21} f_{22} \dots f_{2p}), \dots, F_s=(f_{s1} f_{s2} \dots f_{sp})$ に分割される。スペクトルの各成分 $f_{ij}(i=1 \dots s, j=1 \dots p)$ はそれぞれ実部 $R(f_{ij})$ と虚部 $I(f_{ij})$ とよりなり、各ブロックごとにこれら実部を要素とするベクトル $R(F_1)=(R(f_{11}) R(f_{12}) \dots R(f_{1p})), R(F_2)=(R(f_{21}) R(f_{22}) \dots R(f_{2p})), \dots, R(F_s)=(R(f_{s1}) R(f_{s2}) \dots R(f_{sp}))$ と、同様に各虚部を要素とする $s$ 個のベクトル $I(F_i)=(I(f_{ij}))$ とを作る。

これらベクトルが、予め用意した辞書中の何れの標準ベクトルと最もよく対応するかを検出してベクトル量子化をベクトル量子化部22で行う。つまり辞書として予測される複数の標準的なベク

特開昭59-94797(3)

トルを記憶しておき、入力音声のベクトルが何れの標準ベクトルに近いかを検出し、その一致乃至類似した標準ベクトルを示す番号などの符号を出力する。従つて各スペクトル成分の濃さを量子化するよりも少ないビット数で符号化することができる。しかもこのベクトル量子化に対するビット割当てを適応的に変化する。

即ち局部復号化部19の出力である補助情報の復号出力に応じて前記ブロックごとにビット数を割当てて、一般的には強いスペクトルが含まれるブロックには多くのビットを割当て、弱いスペクトル量子化部22では多くのビット数が割当てられる時は、比較すべき標準ベクトルの数が多い辞書を参照し、少ないビット数が割当てられる時は、標準ベクトル数が少ない辞書を参照する。標準ベクトルの要素の数pは一定であるから標準ベクトル数が多い辞書は記憶されている標準ベクトルは微細なパターンをも表示していることになり、標準ベクトルの数が少ない辞書に記憶されている標準ベクトルは大ざっぱなパターンを示すに過ぎないと

云える。

この適応的情報割当て(ビット割当て)は入力信号と出力信号のフレームごとのSN比を最大化することを目的として行われる。直交変換してもSN比は不変であるから符号化器24のスペクトル平滑部17の出力と受信側の復号化器25のスペクトル再生出力との歪を最小とするようにすればよく、歪尺度はユークリッド距離とする。1フレームあたりの歪Dは次式である。

$$D = \sum_{j=1}^s \sum_{\ell=1}^p (g_{(p-1)j+\ell})^2 \cdot 2^{-2b_j/p} \cdot C$$

- { b<sub>j</sub> : ベクトルあたりに割当てる量子化ビット数
- g<sub>i</sub> : 周波数iにおけるパワースペクトル包絡値
- C : 定数

また全サンプル(スペクトル)数はp・sであつてサンプルあたりの平均情報量(平均量子化ビット数)Bは、

$$B = \sum_{j=1}^s b_j / (p \cdot s)$$

である。Bは一定に保持するから歪Dを最小化する

る量子化ビット数b<sub>j</sub>は次式となる。

$$b_j = B \cdot p + \frac{p}{2} \log_2 \left\{ \frac{\sum_{\ell=1}^p g_{(p-1) \cdot j + \ell}}{\sum_{m=1}^s \left( \sum_{\ell=1}^p g_{(p-1) \cdot m + \ell} \right)^{1/s}} \right\}$$

このb<sub>j</sub>を整数値化し、2<sup>b<sub>j</sub></sup>個からなる辞書から歪最小となるものを選択することで量子化が実行される。

なお量子化部18における量子化もベクトル量子化することができる。このスペクトル包絡の量子化出力、つまり補助情報と、ベクトル量子化部22の出力である波形情報は合成されて符号化出力として復号化器25へ送られる。

復号化器25では入力された波形情報が平滑化スペクトル再生部26で、符号化器24におけるベクトル量子化部22で用いた辞書と同一のものを用いて標準ベクトルを各ブロックの量子化符号により読出して、平滑化スペクトルを再生する。一方入力された補助情報はスペクトル包絡再生部27でスペクトル包絡が再生され、これとパワとを再生された平滑化スペクトルに対してスペクトル再生部28で乗算してスペクトルを再生する。

この再生されたスペクトルを逆変換部29で時間領域に逆変換して出力端子31に再生音声信号を得る。

<第2実施例>

上述においては直交変換を行つた後にスペクトル平滑化を行つたが、入力音声を逆フィルタに通した後に、直交変換を行つてもよい。例えば第5図に第3図と対応する部分に同一符号を付けて示すように入力端子11からの入力音声信号は逆フィルタ32を通して直交変換部1.2へ供給される。一方入力音声信号は線形予測分析器33でスペクトル包絡が分析され、その分析予測係数は量子化部18でベクトル量子化され、その量子化出力は局部復号化部19で復号化され、その復号出力、つまり線形予測係数により逆フィルタ32のフィルタ定数が制御される。この逆フィルタ32の出力は残差信号であり、これを直交変換して前述と同様に符号化して送出する。復号化器25ではスペクトル再生部26でベクトル量子化された符号を復号して残差信号のスペクトルを再生し、これ

を時間領域に逆変換して線形予測合成フィルタ部34へ送出する。この合成フィルタ部34のフィルタ定数は、スペクトル包絡再生部27で再生された予測係数により制御され、フィルタ部34より音声信号が再生される。

第6図Aに、入力音声信号の波形 $a_1$ 、その直交変換出力の実部の波形 $a_2$ 、虚部の波形 $a_3$ を示し、第6図Bに入力音声信号の波形 $a_1$ を逆フィルタ部32に通した後の残差信号波形 $b_1$ を、この残差信号の直交変換出力の実部の波形 $b_2$ を、虚部の波形 $b_3$ をそれぞれ示す。音声入力波形 $a_1$ のスペクトル包絡線 $b_4$ と、各ブロックに対する割当ビット $b_5$ をそれぞれ示す。ただし $p=6$ 、 $B=1.0$ の例である。

上述において量子化の単位となるベクトルの次元 $P$ を入力音声のピッチ周波数に適応させ、1フレームの長さをピッチ周期の整数倍とすることで量子化の効率をさらに高めることができる。この場合はピッチ周波数は時間的に変化するためピッチ周波数も補助情報に含める。また、ベクトルを

KHz、線形予測分析部33の分析次数を8次、分析長(変換長)を26~31ms、分析の重複2ms(台形窓で接続)、ベクトル次元数を6~12(ピッチ適応)とした場合の情報量 $B$ (ビット/サンプル)に対するSN比を第7図に示す。第6図において曲線41は均一量子化で、各1サンプルごとに符号化した場合、曲線42は均一量子化で6次元固定ベクトル符号化した場合、曲線43は均一量子化でベクトルの次元をピッチ周波数に応じて変化させて符号化した場合、曲線44は適応量子化で各サンプルごとに符号化する場合(従来方式)、曲線45はこの発明の方式で6次元固定ベクトル量子化による符号化する場合、曲線46はこの発明の方式でベクトルの次元をピッチ周波数に応じて適応的に変化させて符号化する場合である。これらより、均一量子化(曲線41~43)よりも適応量子化(曲線44~46)の方が優れ、適応量子化でも従来方式(曲線44)よりもこの発明方式(曲線45, 46)の方が優れていることが理解される。0.5~1.1ビット/サンプル領

#### 特開昭59-94797(4)

実部、虚部独立とせず、複素数のままの単位として処理することも可能である。また上述における各部はそれぞれ独立した或は共通の電子計算機で処理することができる。

#### <効果>

以上説明したように、周波数領域で平坦化された信号をブロックに分割し適応的情報割当をすることで量子化効率を高めることができ、特に9.6Kbps以下でスカラ量子化の従来の適応変換符号化方式より高いSN比を持つ音声を再生することができる。周波数領域の平坦化によりベクトル量子化の標準ベクトルの数が少なく済む。また1ブロックあたりに割当てられる情報量が整数であればよく、1サンプルあたりの情報量は $1/P$ ビットの単位で細かく割当てられる。このことにより従来方式の欠点であつた情報量がまづたく割当てられない周波数成分が存在し、かつそれが適応的に変化することに起因する聴覚的劣化を避けることができる。

次に実験例を述べる。サンプリング周波数を8

域で、このSN比の向上は学習サンプル外でも女声で2.5dB、男声で1.0dB程度得られた。スペクトル包絡もベクトル量子化することによりピッチ、パワなどを含めて補助情報は800bps程度と見積ることができるから、残差信号1サンプル当りの情報量 $B$ が0.5で4.8Kbps、1.1で9.6Kbpsの符号化が可能である。

#### 4.図面の簡単な説明

第1図は従来の適応変換符号化方式を示すブロック図、第2図はその動作の説明に供する図、第3図はこの発明による適応変換符号化方式の一例を示すブロック図、第4図はそのブロック分割の例を示す図、第5図はこの発明の他の例を示すブロック図、第6図はその動作例を示す図、第7図は各種符号化方式のSN比-情報量 $B$ との関係を示す図である。

- 11: 音声入力、12: 直交変換部、14: スペクトル包絡抽出器、17: スペクトル平滑部、18: ベクトル量子化器、19: 局部復号化器、21: ブロック分割部、22: ベクトル量子化

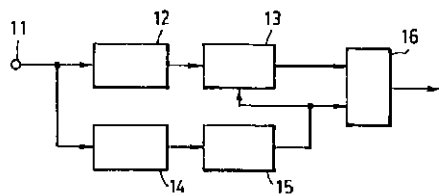
特開昭59-94797(5)

器、23：適応情報割当部、24：符号化器。

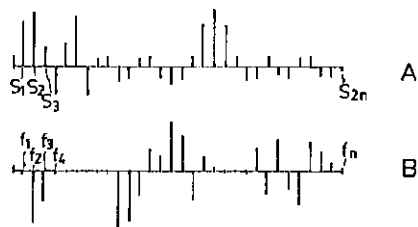
特許出願人 日本電信電話公社

代理人 草野 卓

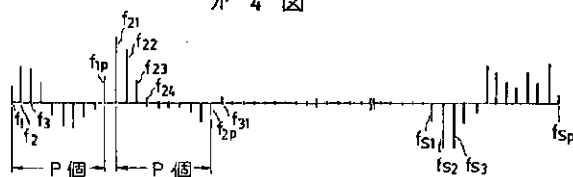
カ 1 図



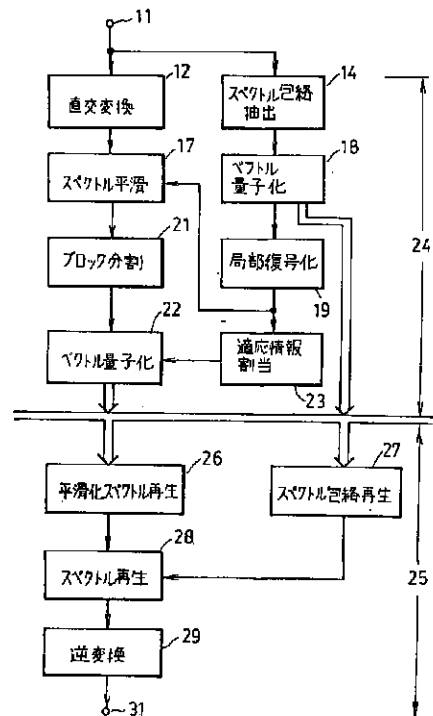
カ 2 図



カ 4 図

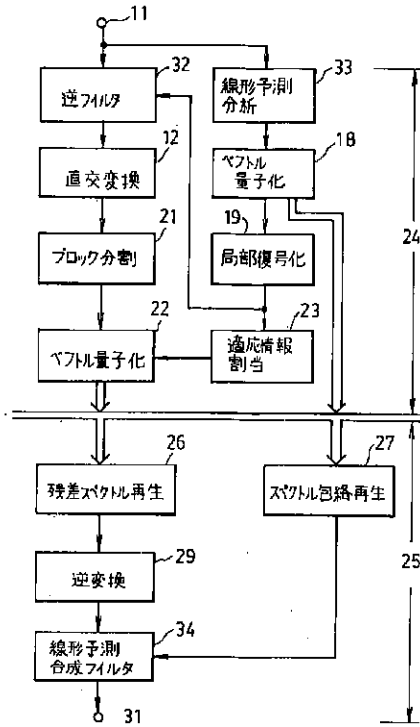


カ 3 図

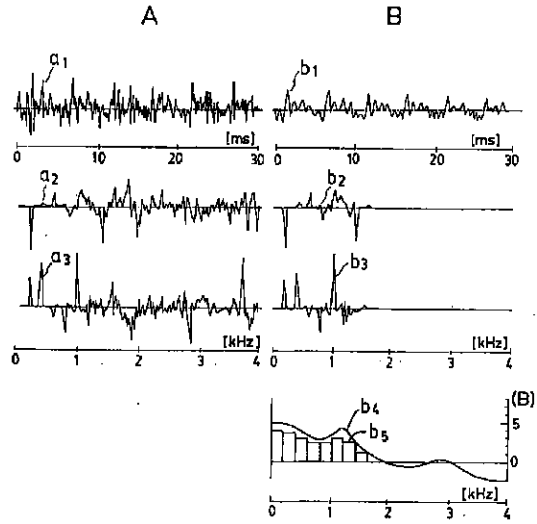


特開昭59-94797(6)

カ 5 図



カ 6 図



カ 7 図

