

(19) 日本国特許庁 (JP)

(12) 特許公報 (B2)

(11) 特許番号

**特許第3616307号**

(P3616307)

(45) 発行日 平成17年2月2日 (2005.2.2)

(24) 登録日 平成16年11月12日 (2004.11.12)

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

F I

G 1 0 L 19/00  
G 1 0 L 19/02  
H 0 3 M 7/30

G 1 0 L 9/18 E  
H 0 3 M 7/30 B  
G 1 0 L 7/04 G

請求項の数 2

(全 7 頁)

(21) 出願番号 特願2000-150212(P2000-150212)  
(22) 出願日 平成12年5月22日 (2000.5.22)  
(65) 公開番号 特開2001-331198(P2001-331198A)  
(43) 公開日 平成13年11月30日 (2001.11.30)  
審査請求日 平成13年12月21日 (2001.12.21)

(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  
(74) 代理人 100066153  
弁理士 草野 卓  
(74) 代理人 100100642  
弁理士 稲垣 稔  
(72) 発明者 守谷 健弘  
東京都千代田区大手町二丁目3番1号 日  
本電信電話株式会社内  
(72) 発明者 岩上 直樹  
東京都千代田区大手町二丁目3番1号 日  
本電信電話株式会社内

最終頁に続く

(54) 【発明の名称】 音声・楽音信号符号化方法及びこの方法を実行するプログラムを記録した記録媒体

(57) 【特許請求の範囲】

【請求項 1】

音声・楽音信号をフレーム単位、またはフレームをさらに分割してサブフレーム単位とし

、前記音声・楽音信号をフレーム単位またはサブフレーム単位で周波数領域成分に変換し、前記周波数領域成分を変形させて量子化し、ベクトル単位で符号化出力を得る音声・楽音信号符号化方法であって、

前記周波数領域成分の変形は、

前記周波数領域成分を構成するベクトル毎の平均パワーを求め、前記ベクトル毎の平均パワーからフレーム毎の平均パワーを求め、

前記ベクトル毎に、当該ベクトルの平均パワー及びフレーム毎の平均パワーに基づいてビット配分を求め、

前記ビット配分が一定閾値以下となるベクトルに対応する周波数領域成分を減衰させる、ことを特徴とする、

音声・楽音信号符号化方法。

【請求項 2】

音声・楽音信号をフレーム単位、またはフレームをさらに分割してサブフレーム単位とする処理と、

前記音声・楽音信号をフレーム単位またはサブフレーム単位で周波数領域成分に変換する処理と、

周波数領域成分を変形させる処理と、  
 前記変形させた周波数領域成分を量子化する処理と、を有し、ベクトル単位で符号化出力  
 を得る音声・楽音信号符号化方法をコンピュータに実行させるプログラムを記録したコン  
 ピュータ読取可能な記録媒体であって、  
 前記周波数領域成分を変形させる処理は、  
 前記周波数領域成分を構成するベクトル毎の平均パワーを求め、前記ベクトル毎の平均パ  
 ワーからフレーム毎の平均パワーを求め、  
 前記ベクトル毎に、当該ベクトルの平均パワー及びフレーム毎の平均パワーに基づいてビット  
 配分を求め、  
 前記ビット配分が一定閾値以下となるベクトルに対応する周波数領域成分を減衰させる、  
 ことを特徴とする、

10

音声・楽音信号符号化方法をコンピュータに実行させるプログラムを記録したコンピ  
 ュータ読取可能な記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は楽音や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信号符  
 号化法に関するものである。

【0002】

【従来の技術】

20

従来の信号の変換符号化復号化器は図1に示されるような構成となっている。  
 符号器は、楽音や音声信号を時間窓で切り出したフレームを例えば、周波数変換としてM  
 D C T (Modified Discrete Cosine Transform)を使い、スペクトル推定部により、その変  
 換後の成分全体、すなわちスペクトルの包絡や全体の平均振幅を求め、平坦化・正規化部  
 でM D C T係数を正規化したあとで量子化部において量子化する。量子化には適応ビット  
 配分または適応重み付けベクトル量子化などを利用する。符号器は、スペクトル推定部で  
 生成された補助情報(スペクトル包絡成分情報)と量子化部で生成された主情報(周波数  
 領域成分情報)を出力する。

復号器は、主情報(周波数領域成分情報)を逆量子化部で逆量子化し、逆平坦化部におい  
 て、逆量子化信号を補助情報(スペクトル包絡成分情報)を用いて逆平坦化し、さらに、  
 逆M D C Tを行ない音声や楽音信号を出力する。

30

【0003】

符号化する帯域に対して、与えられたビット数が少ないと量子化雑音が多くなり、品質が  
 劣化する。このため、ビット数が少ないときには高い周波数成分を固定的に減衰させ、帯  
 域を狭めるかわりに量子化雑音を軽減する。音声などのようにもともと帯域が狭い場合は  
 この効果があるが、通常の音楽では量子化雑音が目立ちにくい反面、帯域が狭いことによ  
 る劣化が問題となり、量子化雑音と帯域の両立は難しかった。

【0004】

【発明が解決しようとする課題】

本発明の目的は、できるだけ高品質で楽音や音声を送信することであり、特に符号化に先  
 立って、量子化雑音を抑えきれない周波数成分を減衰させることで、再生信号の品質の低  
 下を抑える方法およびこの方法を実行するためのプログラムを記録した記録媒体を提供す  
 ることである。

40

【0005】

【課題を解決するための手段】

前記目的を達成するために、本発明は、与えられた情報量と信号の帯域に基づいて十分な  
 精度で量子化できない周波数成分を推定し、符号化の前にその成分を減衰させることで最  
 終的な聴覚上の歪を軽減する。すなわち、入力信号の特徴にあわせて適応的に帯域を制限  
 する点を特徴とする。

【0006】

50

## 【発明の実施の形態】

## 実施例

図2は、本発明の第1の実施例の符号器の構成図である。

音声・楽音信号を時間窓で切り出した1フレームについての処理を説明する。

この例ではフレーム単位でM D C T係数に変換し、M D C T係数をスペクトル包絡とパワーで正規化したあとで量子化を行なっている。この枠組みは従来法と同じである。

## 【0007】

典型的な数値としては1フレームは例えば1024サンプルからなり、ベクトルは8サンプルからなる。すなわち、1フレームあたりのベクトルの数Nは128である。

10

本発明ではM D C T係数を平坦化する前に前処理の変形を行なう。この前処理は、スペクトル算出部、減衰量算出部、前処理変形部によりM D C T係数のスペクトルを分析し、符号化条件により減衰量を算出し、それに基づいてM D C T係数を変形する。具体例としてはベクトルの絶対値または複数サンプルごとにベクトルとしてまとめたときの平均パワー $E_i$ を求める。

## 【0008】

## 【数1】

さらに全体の平均パワー（幾何平均） $\bar{E}$ を求める。

$$\bar{E} = \prod_{i=0}^{N-1} E_i^{(1/N)} \quad (1)$$

20

一方、フレームあたりのサンプル数とビット数からサンプルあたりのビット数を求め、後述するレート歪理論による割り当てビット数が0となるベクトルのパワー値を閾値とする

## 【0009】

ベクトルあたりの平均パワーが閾値以下のベクトルに対しては振幅を減衰させる。例えば量子化前のM D C T係数そのものに

## 【0010】

## 【数2】

$$\sqrt{E_i / \bar{E}}$$

30

をかける。さらにこの変形で信号全体のパワーが減衰するのでM D C T係数全体に係数をかけて、全体のパワーが保たれるように変形してもよい。この減衰係数は、符号化条件、すなわち、サンプリング周波数、与えられる量子化ビット数などに依存し、詳細なパラメータは実験的に調整したほうがよい。

## 【0011】

この変形されたM D C T係数を入力とみなして従来の平坦化と量子化を行なう。この前処理による変形の情報は復号器には伝えられないので、復号器で再生される信号は量子化精度がよくても元の入力信号とは異なる。ただし、この変形は低ビット量子化で生じる量子化歪より十分小さいように設定することで、品質を改善できる。

40

## 【0012】

図3は、この処理によるM D C T係数の変形例である。

(A)の実線で表わされるもとのスペクトル（ベクトル毎の平均パワー）の閾値Tより小さいベクトルに対して(B)のような減衰係数をかける。係数が1ということは変形なしということになる。この結果(A)の破線で表わされるスペクトルに変形される。

この処理の原理はレート歪理論にある。

## 【0013】

## 【数3】

ベクトルごとの平均歪を最小にするビット配分  $b_i$  は  $\bar{b}$  をベクトルあたりの平均ビット数としたとき以下のようになる。

$$b_i = \bar{b} + \frac{1}{2} \log_2 \frac{E_i}{\bar{E}} \quad (2)$$

ベクトルあたりの平均ビット数が少ない時には  $b_i$  が負の値になる場合がある。実際に配分するビット数  $b_i$  は負にはなりえないから  $b_i$  が負の場合は 0 とする。すなわち情報を送らずに復号器でそのベクトルの値を 0 とする。情報量が少ないときにはパワーの小さい成分の伝送を放棄することになる。

【0014】

実際の符号化ではビット配分を示す補助情報を細かく送ることは情報量が増えてしまい得策ではないし、すでに標準化された方法では補助情報を送ることはできない。本発明では、あらかじめ量子化ビットを配分できないような帯域の信号を減衰させることでその帯域の量子化雑音を小さくする。

【0015】

図4は、同じフレームのなかで時間的に分割して周波数領域に変換する場合の実施例である。

この実施例の場合、一つのフレームを4つのサブフレームに分割して、それぞれのサブフレームごとにMDC T係数を求めたものである。フレーム全体のパワーから閾値Tを基準に減衰量を決定する。

【0016】

図5は、聴覚のマスク効果を利用した減衰量を決定する実施例である。

ここまでの実施例では(A)のようにもとのスペクトルに対して量子化雑音が周波数軸上で平坦になるように最適量子化を行なうことを前提としてきた。実際の符号化では(B)のようにマスク効果を考慮してもとのスペクトルと相関をもつように量子化雑音を整形することがよく行なわれる。そこで(B)の量子化雑音が平坦となるように逆特性をもとの信号にかけることで(C)の変形したスペクトルを作る。このスペクトルを入力のス

ペクトルとみなしてこれまでの実施例に適用すればよい。また、スペクトルのパワーを算出するベクトルは聴覚の分解能を考慮して、低域ほど幅が狭く、広域が広がる周波数目盛り尺度(例えば、バーク尺度)で分割することもできる。

【0017】

なお、本発明の符号器をCPUやメモリ等を有するコンピュータで構成し、記録媒体に記録されたプログラムをコンピュータに読取り、コンピュータの動作を制御して前述の実施の形態の各構成要素を実現する。

【0018】

【発明の効果】

本発明により、量子化歪が避けられない周波数成分を適応的に減衰させているので、その成分に対する量子化誤差を相対的に小さくすることができ、品質を改善できる。音声信号のように低域にパワーが集中している場合には量子化する帯域が狭められ、低域の量子化誤差が小さくなる。また、一般の音楽ではパワーが少なく聴覚的に重要でない帯域の信号が減衰し、その他の重要な帯域の歪が小さくなる。

【0019】

周波数領域の量子化を行なう符号化には処理量の大きな増加はなく簡単に組み入れることができ、時間領域の符号化と組み合わせることも可能である。

【図面の簡単な説明】

【図1】従来のMDC T係数を量子化する符号器及び復号器の構成を示す図。

【図2】本発明の第1実施例である符号器の構成を示す図。

- 【図3】本発明の第1実施例の符号化のMDC T係数の変形例を示す説明図。
- 【図4】1つのフレームを4つのサブフレームに分割したMDC T係数の例を示す図。
- 【図5】聴覚特性の補正を行なう場合の説明図。

【図1】

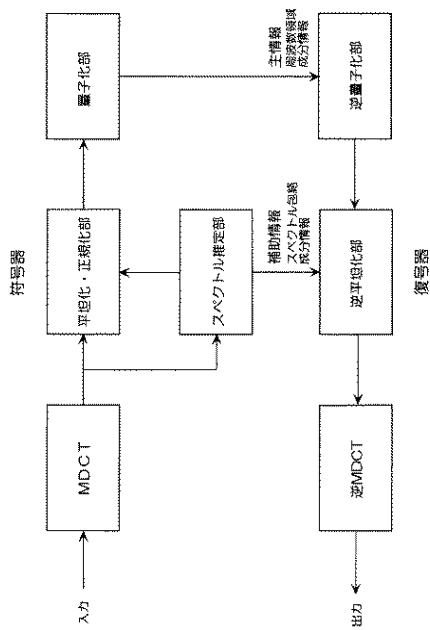


図1

【図2】

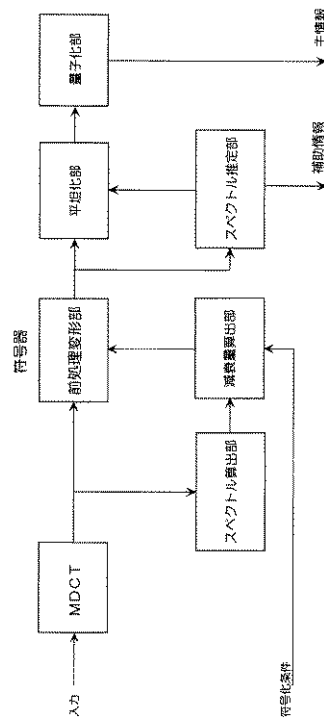
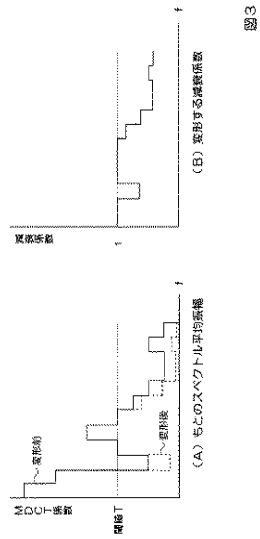


図2

【図3】



【図4】

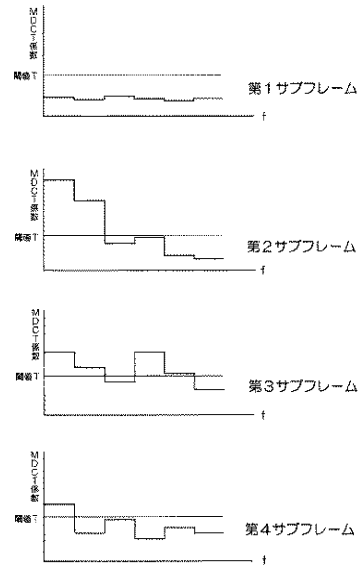


図4

【図5】

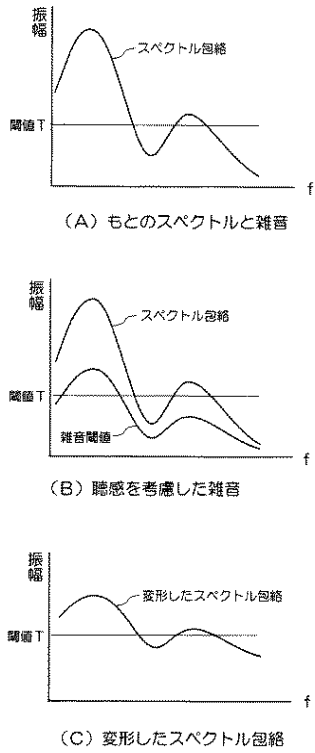


図5

-----  
フロントページの続き

- (72)発明者 神 明夫  
東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内  
(72)発明者 森 岳至  
東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内

審査官 山下 剛史

(56)参考文献 特開平08-328600(JP,A)

(58)調査した分野(Int.Cl.<sup>7</sup>, DB名)  
G10L 19/00  
G10L 19/02  
H03M 7/30  
JSTPlusファイル(JOIS)