

(19)日本国特許庁(JP)

(12)特許公報 ( B 2 )

(11)特許番号

特許第 3 4 6 8 3 3 5 号

( P 3 4 6 8 3 3 5 )

(45)発行日 平成15年11月17日(2003.11.17)

(24)登録日 平成15年9月5日(2003.9.5)

(51)Int. Cl. <sup>7</sup>	識別記号	F I		
G 1 0 L	19/04	H 0 3 M	7/30	A
	11/00	H 0 4 B	14/04	B
H 0 3 M	7/30	G 1 0 L	9/14	J
H 0 4 B	14/04		7/02	A

請求項の数 1

(全 4 頁)

(21)出願番号	特願平8-190375	(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8年7月19日(1996.7.19)	(72)発明者	守谷 健弘 東京都新宿区西新宿三丁目19番2号 日本 電信電話株式会社内
(65)公開番号	特開平10-39900	(72)発明者	神 明夫 東京都新宿区西新宿三丁目19番2号 日本 電信電話株式会社内
(43)公開日	平成10年2月13日(1998.2.13)	(72)発明者	岩上 直樹 東京都新宿区西新宿三丁目19番2号 日本 電信電話株式会社内
審査請求日	平成13年5月11日(2001.5.11)	(74)代理人	100066153 弁理士 草野 卓
		審査官	渡邊 聡

最終頁に続く

(54)【発明の名称】音響信号の周波数領域変換係数の包絡推定方法

1

(57)【特許請求の範囲】

【請求項 1】音響信号の重ね合わせ直交変換係数が入力され、その重ね合わせ直交変換係数のパワーの包絡形状を推定する方法において、

上記重ね合わせ直交変換係数に対して、三角関数にもとづく逆直交変換を行う第 1 の処理と、その第 1 の処理の結果に対して、線形予測分析して上記パワーの包絡形状を推定する第 2 の処理と、を有することを特徴とする音響信号の周波数領域変換係数の包絡推定方法。

【発明の詳細な説明】

【 0 0 0 1 】

【発明の属する技術分野】この発明は音声や音楽などの音響信号を例えば量子化して符号化する方法に適用され、音響信号を重ね合わせ直交変換 ( L O T : L a p p

2

ed Orthogonal Transform) し、かつその変換係数のパワー包絡形状を推定する方法に関する。

【 0 0 0 2 】

【従来の技術】音声や音楽信号を少ないビット数で符号化する際に周波数領域で量子化する手法がよく知られている。変換には D F T ( 離散フーリエ変換 )、D C T ( 離散コサイン変換 ) のような通常の直交変換、M D C T ( 変形離散コサイン変換 ) などの L O T ( 重ね合わせ直交変換 ) が使われる。量子化の前に周波数領域の係数を平坦化する目的で線形予測分析が有効であることも知られている。これらの技術を組み合わせて音響信号の広い範囲の信号に対して品質の高い符号化を実現する方法 ( 特願平 7 - 5 2 3 8 9 ) で提案されている。この処理

10

は図3に示される。

【0003】入力端子11からデジタル信号とされた音響信号が入力され、フレーム分割部12でそのN入力サンプルごとに過去2×Nサンプルの入力系列が抽出され、その2×N個のサンプルの系列をN個づつ重ね合わ\*

$$W(n) = \sin(n / (2N)) \dots (1)$$

この窓関数が掛けられた信号はLOTの一種であるN次のMDCT(Modified Discrete Cosine Transform:変形離散コサイン変換)部14で変形離散コサイン変換されて周波数領域係数(周波数軸上のそれぞれの点におけるサンプル値)に変換される。窓掛部13の出力は線形予測分析部15へも供給され、ここで線形予測分析され、P次の予測係数 $o_1, \dots, o_P$ が求められる。この予測係数 $o_1, \dots, o_P$ は量子化部16で例えばLSPパラメータに変換されてから量子化され、スペクトラム概形を示すインデックス $I_P$ が得られる。線形予測係数 $o_1, \dots, o_P$ のパワースペクトル包絡(電力伝達関数)の平方根をMDCT係数の振幅包絡の近似とする。この例では量子化部16よりのLSPパラメータがスペクトラム概形計算部17へ供給され、そのパワースペクトル包絡が計算される。この計算は例えば東海大出版、古井著「デジタル音声処理」の91頁に述べられている。この計算されたスペクトル包絡でMDCT部14よりのMDCT係数が平坦化部18により割算される。またスペクトラム概形計算部17で計算されたスペクトラム包絡は重み計算部19で聴覚特性に応じた重み付け係数を計算し、この重み付け係数を用いて、平坦化部18で平坦化されたMDCT係数が、重み付き量子化部21で聴覚重み付き量子化がなされ、その量子化インデックス $I_M$ が出力される。

【0005】

【発明が解決しようとする課題】MDCTはフレーム境界雑音が出ないという利点がある。しかし、MDCT変換操作と、その逆変換操作とで、時間領域の信号が折り返されているため、変換後の正確な包絡を求めることは容易でなかった。つまり従来においてはMDCT前の時間領域信号を線形予測分析し、その予測係数から、MDCT係数の包絡を近似しており、一方、MDCT操作において時間領域で信号の折り返しが生じており、この時間領域での折り返しが生じた変換係数(MDCT係数)の包絡を、前記時間領域で折り返しのない信号の線形予測係数から近似しても、比較的大きな近似誤差とならざるを得ない。

【0006】この発明の目的は、音響信号の周波数領域変換係数の包絡を線形予測係数から従来よりも高い精度で近似することができる方法を提供することにある。

【0007】

【課題を解決するための手段】本願発明の前段階において、窓関数を掛けたサンプルに対し、重ね合わせ直交変換により折り返されるべきサンプルを時間軸上で折り返

\*せ、直交変換処理用フレームとされる。このLOT処理用フレームは時間窓掛部13で2Nの窓関数(時間窓)が掛けられる。窓関数 $W(n)$ としては、例えば次式で示されるハニング窓が用いられる。

【0004】

し、その折り返し処理したものに対し、三角関数にもとづく直交変換を行って、重ね合わせ直交変換係数を得、また前記折り返し処理したものを線形予測分析して、パワー包絡形状を求められている。

【0008】

本願発明によれば、入力された重ね合わせ直交変換係数に対し、三角関数にもとづく逆直交変換を行い、その逆直交変換結果に対して、線形予測分析を行い、その予測係数から、前記入力された重ね合わせ直交変換係数のパワー包絡形状を推定する。

【0009】

【発明の実施の形態】図1に本願発明の前段階における処理手順を示し、図3と対応する部分に同一符号を付けてある。この前段階においても、従来と同様に、入力された音響信号のサンプル系列はフレーム分割部12で重ね合わされて、LOT用フレームとされ、その各LOT用フレームに対し、窓掛部13で窓関数が掛けられる。この窓関数としては従来用いられているものと同様のものを用いることができる。

【0010】

この段階においては、窓関数が掛けられたサンプルに対し、重ね合わせ直交変換により折り返されるべきサンプルを時間軸上(時間領域)で折り返し、その折り返し処理したサンプルに対し、直交変換部26で、三角関数にもとづく直交変換を行って、LOT係数を得る。また折り返し部25で処理されたサンプルに対し、線形予測分析部27で線形予測分析を行い、その求めた予測係数を用いて、包絡計算部28で包絡を計算して、直交変換部26で得たLOT係数のパワー包絡形状の近似値とする。

【0011】

以下、この段階を、LOTとしてMDCTを適用した場合につき、更に具体的に説明する。この場合は図3の場合と同様に入力音響信号サンプル系列から、Nサンプル(1フレーム)づつ重なった2NサンプルをLOT用フレームとし、その各2Nサンプルに対して窓関数をかける。次に折り返し部25で時間軸上の折り返し処理を行うが、MDCTにより折り返されるべきサンプルを以下により求める。即ち窓関数がかけられた系列を $x(n)$ とすると、この $x(n)$ は下記に定義されるMDCTを適用するとN点の係数 $Y(k)$ が得られる。

【0012】

$$Y(k) = \sum_{i=0}^{2N-1} x(i) \cos \{ (2i+1+N)(2k+1) / (4N) \} \dots (2)$$

このMDCTの変換関数、つまり $\cos \{ (2i+1+N)(2k+1) / (4N) \}$ は例えば $N=32$ とする

と、 $k = 0, k = 1, k = 2, k = 3, \dots$  のそれぞれについて、図 4 A, B, C, D に示すようになる。これらの曲線の前半 ( $0 \sim N, i/N = 0 \sim 1$ ) では  $N/2 (i/N = 0.5)$  を中心とする奇対称関数であり、後半 ( $N \sim 2N, i/N = 1 \sim 2$ ) では  $3N/2 (i/N = 1.5)$  を中心に偶対称関数であり、MDC T による時間領域でのサンプルは、 $i/N = 0 \sim 0.5, i/N = 1.5 \sim 2$  とそれぞれ乗算されるものである。そして MDC T は (2) 式で示したように  $x(i)$  との図 4 の関数との積和演算であるから、各フレームの先頭から  $N/2 (i/N = 0 \sim 0.5)$  の長さのサンプル  $x(i)$ , ( $i = 0, \dots, N/2 - 1$ ) を時間的に反転した上で次の  $N/2$  の長さのサンプル  $x(i)$ , ( $i = N/2, \dots, N - 1; i/N = 0.5 \sim 1$ ) から引き、またフレームの最後の  $N/n$  の長さのサンプル  $x(i)$ , ( $i =$

$$v(k) = \sum_{i=0}^{N-1} y(i) \cos \{ (2i + 1 + 2N)(2k + 1) / (4N) \} \dots (4)$$

この実施例では折り返し処理結果  $y(i)$  に対して、線形予測分析部 2 7 で線形予測分析を行なって、P 次の予測係数  $a_1, \dots, a_P$  を求めて、MDC T 係数  $v(k)$  の包絡を推定する。図 3 に示したように、予測係数  $a_1, \dots, a_P$  を LSP パラメータに変換し、量子化した後に、包絡を計算する。

【0015】直交変換部 2 6 ではその変換処理において時間領域での折り返しは生じない。従って、線形予測分析部 2 7 より得た予測係数を用いてそのパワースペクトル包絡を求めると、直交変換部 2 6 の変換結果、つまり LOT 係数の包絡をよく近似したものが得られる。図 3 中の重み付き量子化部 2 1 で比較的粗く量子化を行い、その量子化誤差を求め、その量子化誤差を量子化するというように複数段階にわたって行うことがある。その場合は量子化部 2 1 の量子化結果を逆量子化し、それを逆平坦化し、その逆平坦化された復号 MDC T 係数と、MDC T 部 1 4 の MDC T 係数との差を求め、その差分 MDC T 係数を量子化する。この量子化のために差分 MDC T の包絡を求めるには、従来の手法からすると、その差分 MDC T に対し逆 MDC T を行って時間領域に戻し、その戻された時間領域信号を線形予測分析して、パワースペクトル包絡を求めることになる。この逆 MDC T において、時間領域での折り返しが生じるため、先の場合と同様に、包絡の近似精度はあまりよくならない。

【0016】請求項 1 の発明によればこのような問題を解決できる。請求項 1 の発明の実施例を図 2 に示す。入力端子 3 1 から、音響信号の MDC T 係数  $v(i)$ 、例えば前記差分 MDC T 係数を逆 DCT (逆離散コサイン変換) 部 3 2 を入力し、逆 DCT を行う。つまり下記の演算を行う。

$N + N/2, \dots, 2N - 1; i/N = 1.5 \sim 2$ ) を時間的に反転した上でその直前の  $N/2$  の長さのサンプル  $x(i)$ , ( $i = N, \dots, N + N/2 - 1; i/N = 1, \dots, 1.5$ ) に加えることにより、折り返し部 2 5 の処理を行うことができる。

【0013】つまり折り返し部 2 5 の折り返し処理は次式で表わせる。

$$y(i) = x(N/2 + i) - x(N/2 - 1 - i), (i = 0, \dots, N/2 - 1)$$

$$y(i) = x(N/2 + i) + x(5N/2 - 1 - i), (i = N/2, \dots, N - 1)$$

この折り返し処理結果  $y(i)$  に対して、直交変換部 2 6 で下記の N 点の DCT (離散コサイン変換) を行う。この結果は通常の MDC T 係数と同一のものとなる。

【0014】

$$z(i) = \sum_{k=0}^{N-1} v(k) \cos \{ (2i + 1 + 2N)(2k + 1) / (4N) \} \dots (5)$$

この逆 DCT により時間領域信号とされたものは入力

$v(i)$  を直接逆 MDC T したのものと見ると折り返し処理を行なった波形、つまり図 1 中の折り返し部 2 5

の処理結果の波形と同一となっている。従って、この逆 DCT された信号に対して線形予測分析部 3 3 で線形予測分析を行い、その予測係数のパワースペクトルに包絡を包絡計算部 3 4 で求めて、入力 MDC T 係数  $v(i)$  のパワー包絡の近似を得る。なお、逆直交変換部 3 2 の出力に対して、DCT を行えば入力端子 3 1 の入力 MDC T 係数となる。

【0017】上述においては、この発明を MDC T、つまり DCT を基礎とした LOT について適用したが、DST (離散サイン変換) や DFT などの一般の直交変換を基礎とする LOT にも共通に利用できる。

【0018】

【発明の効果】この発明によれば、時間領域での重ね合わせによるフレーム境界での連続性を維持するという LOT の特徴を生かしながら、LOT 係数の包絡を精度よく求めることができ、符号化に適用することにより量子化歪を小さくすることができる。また LOT 係数を多段で量子化する場合に、前段で量子化された LOT 係数の誤差の包絡を精度よく推定することができる。

【図面の簡単な説明】

【図 1】本願発明の前段階の処理手順を示す図。

【図 2】請求項 1 の発明の実施例の処理手順を示す図。

【図 3】従来の音響信号変換符号化法の処理手順を示す図。

【図 4】MDC T の変換係数の例を示す図。

【図 1】

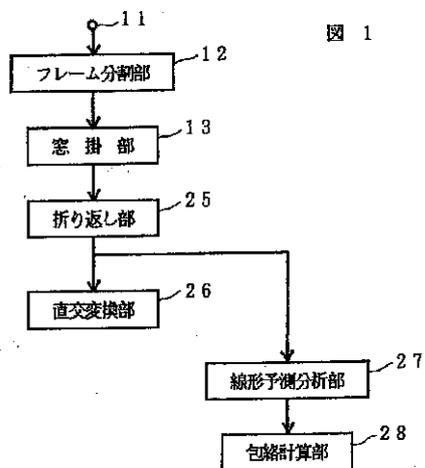


図 1

【図 2】

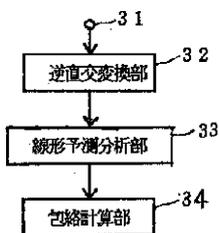


図 2

【図 3】

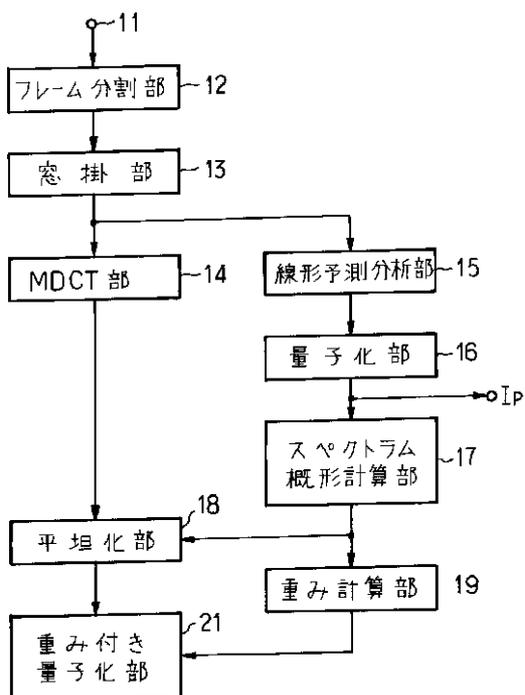


図 3

【図 4】

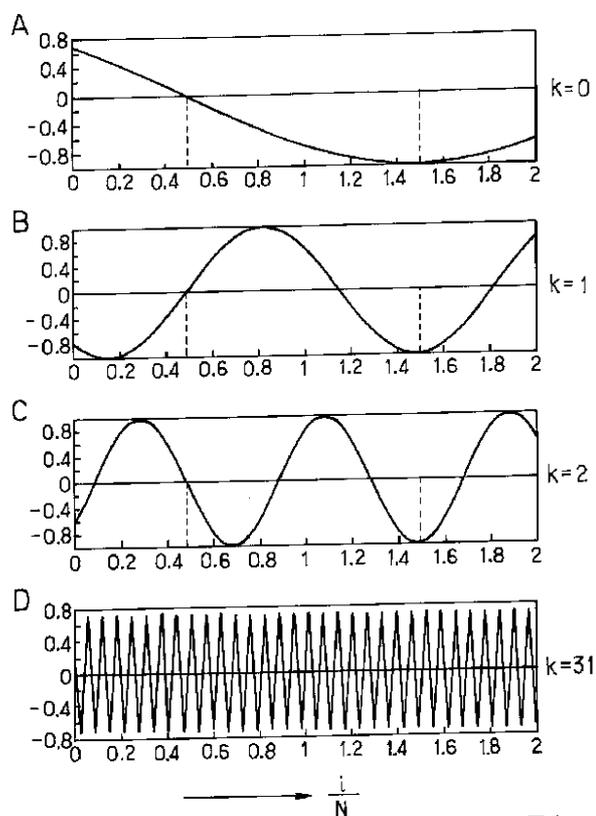


図 4

フロントページの続き

(72)発明者 三樹 聡

東京都新宿区西新宿三丁目19番2号 日  
本電信電話株式会社内

(56)参考文献 特開 平 6 - 232824 ( J P , A )  
特開 平 4 - 44099 ( J P , A )

(72)発明者 池田 和永

東京都新宿区西新宿三丁目19番2号 日  
本電信電話株式会社内

(58)調査した分野(Int.Cl.<sup>7</sup>, DB名)

G10L 19/04  
G10L 11/00