

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4273207号  
(P4273207)

(45) 発行日 平成21年6月3日(2009.6.3)

(24) 登録日 平成21年3月13日(2009.3.13)

(51) Int.Cl. F I  
**HO4N 7/26 (2006.01)** HO4N 7/13 Z

請求項の数 6 (全 13 頁)

(21) 出願番号 特願平9-530572  
 (86) (22) 出願日 平成9年2月22日(1997.2.22)  
 (65) 公表番号 特表2000-505619(P2000-505619A)  
 (43) 公表日 平成12年5月9日(2000.5.9)  
 (86) 国際出願番号 PCT/EP1997/000863  
 (87) 国際公開番号 WO1997/032436  
 (87) 国際公開日 平成9年9月4日(1997.9.4)  
 審査請求日 平成16年2月2日(2004.2.2)  
 (31) 優先権主張番号 96103269.5  
 (32) 優先日 平成8年3月2日(1996.3.2)  
 (33) 優先権主張国 欧州特許庁(EP)

(73) 特許権者  
 ドイチェ トムソン—ブランド ゲーエム  
 ベーハー  
 ドイツ連邦共和国, 78048 ヴィリン  
 ゲン—シュヴェニンゲン, ヘルマン—シュ  
 ヴェアー—シュトラ—セ 3  
 (74) 代理人  
 弁理士 伊東 忠彦  
 (74) 代理人  
 弁理士 片山 修平  
 (72) 発明者  
 ブラヴァト, マイノルフ  
 ドイツ連邦共和国, 30161 ハノーヴ  
 ァー, イゼルンハーゲネナー—シュトラ—  
 セ 18

最終頁に続く

(54) 【発明の名称】 デジタルビデオデータを符号化又は復号化する方法及び装置

## (57) 【特許請求の範囲】

## 【請求項 1】

瞬時記憶ビットレートが可変である記憶媒体上の所定の記憶容量を占める記憶されるべきデジタルビデオデータのシーケンスを符号化する方法であって、

全体的な平均ビットレートは上記所定の記憶容量を上記シーケンスの時間的長さで割ったものであり、

上記所定の記憶容量の第1の部分内で、上記データは、上記全体的な平均ビットレートよりも低い平均ビットレートと、フレームタイプの符号化において大きな仮想バッファサイズを用いることによるゆるい瞬時ビットレート制御とを用いて符号化され、

上記所定の記憶容量の残りの部分内で、上記データは、上記全体的な平均ビットレートよりも高い平均ビットレートと、上記フレームタイプの符号化においてより小さな仮想バッファサイズを用いることによるきつい瞬時ビットレート制御とを用いて符号化されることを特徴とする方法。

## 【請求項 2】

請求項1記載の方法であって、上記所定の記憶容量の上記第1の部分で、上記平均ビットレートは、上記全体的な平均ビットレートの略95%であり、

上記所定の記憶容量の上記第2の部分で、上記平均ビットレートは、上記全体的な平均ビットレートの略150%である方法。

## 【請求項 3】

請求項1又は2記載の方法であって、上記第1の部分は上記所定の記憶容量の略80%で

10

20

あり、上記第 2 の部分は上記所定の記憶容量の略 20% である方法。

【請求項 4】

請求項 1 乃至 3 のうちいずれか 1 項記載の方法であって、上記所定の記憶容量の上記第 1 の部分と上記第 2 の部分との間に、上記ビットレート制御の特性が滑らかに変化する別の部分がある方法。

【請求項 5】

瞬時記憶ビットレートが可変である記憶媒体上の所定の記憶容量を占める記憶されるべきデジタルビデオデータのシーケンスを符号化する装置であって、  
全体的な平均ビットレートは上記所定の記憶容量を上記シーケンスの時間的長さでわったものであり、

10

離散コサイン変換手段と、量子化手段と、逆量子化手段と、逆離散コサイン変換手段と、画像記憶手段と、符号化された出力信号を送出し上記量子化手段の出力に接続されたエン

トローピー符号化手段とを含む符号化ループを使用し、  
上記量子化手段と、上記逆量子化手段と、上記エントローピー符号化手段とは、  
上記所定の記憶容量の第 1 の部分内で、上記データが、上記全体的な平均ビットレートよりも低い平均ビットレートと、フレームタイプの符号化において大きな仮想バッファサイズを用いることによるゆるい瞬時ビットレート制御とを用いて符号化され、

上記所定の記憶容量の残りの部分内で、上記データが、上記全体的な平均ビットレートよりも高い平均ビットレートと、上記フレームタイプの符号化においてより小さな仮想バッファサイズを用いることによるきつい瞬時ビットレート制御とを用いて符号化されるよう

20

に制御されることを特徴とする装置。

【請求項 6】

特に光ディスクである記憶媒体にデジタルビデオデータのシーケンスを記憶する方法であって、

請求項 1 に記載の方法により前記デジタルビデオデータのシーケンスを符号化する段階と、

上記記憶媒体に上記符号化したデジタルビデオデータのシーケンスを記録する段階と、  
を有することを特徴とする方法。

【発明の詳細な説明】

本発明は、一定のデータ記憶容量を有する記憶媒体のためデジタルビデオデータを符号化し、一定のデータ記憶容量を有する記憶媒体からのデジタルビデオデータを復号化する方法並びに装置に関する。

30

発明の背景

ISO/IEC 13818 Information Technology- Generic coding of moving pictures and associated audio information: Video (MPEG2) には、符号化デジタルビデオ及びオーディオデータストリームに関する国際標準が記載されている。画像内容の変化に起因して、生成されるコードの量は時間に関して変化する。しかし、固定チャネル容量の場合、一般的に、送信側でエンコーダバッファを使用し、受信側でデコーダバッファを使用することにより実現され得る固定ビットレート (CBR) MPEG エンコーダコントロールが望ましい。これについては、上記国際標準の付属書類 C 及び D. 4 に詳細に説明されている。他のモードは可変ビットレート (VBR) 符号化である。

40

発明の概要

可変ビットレート動作が可能である DVD (デジタルビデオディスク) のようなデジタル記憶媒体が存在する場合、固定ビットレート符号化の代わりに可変ビットレート符号化を使用することにより符号化品質を改良することが可能である。

本発明の目的は、符号化品質が固定ビットレート符号化の場合の品質よりも基本的に改良される、固定容量記憶媒体のための可変ビットレート符号化の方法を開示することである。本発明の目的は請求項 1 に記載された方法によって実現される。

本発明の別の目的は、復号化品質が固定ビットレート復号化の場合の品質よりも基本的に改良される、固定容量記憶媒体からの可変ビットレート復号化の方法を開示することであ

50

る。この目的は請求項 2 に記載された方法によって実現される。

本発明の別の目的は、上記の新規の符号化及び復号化方法を夫々に使用する符号化及び復号化装置を提供することである。この目的は請求項 3 及び 4 に記載された装置によって実現される。

本発明の別の目的は、上記本発明の方法に従って符号化又は復号化されるデジタルビデオ信号を開示することである。この目的は請求項 5 に記載された信号によって実現される。

“固定ビットレート”は、各フィールド期間(50Hzテレビジョン信号の場合には20ms)に一定数のビットがチャンネル(例えば、システムマルチプレクサ)に出力されることを意味し、一方、“可変ビットレート”は、他の全てのビット送出スケジュールをカバーする。

固定ビットレート環境において、カーネルMPEGエンコーダ回路の後にある程度のサイズのバッファが接続される。このバッファはカーネルMPEGエンコーダの本質的に可変性のビットレート出力を平滑化する。ビデオバッファリング照合子V BVは、使用可能なエンコーダバッファサイズの上限、並びに、このプロファイル@レベル(P@L)に準拠した全てのビットストリームを復号化するためデコーダによって必要とされる最小バッファサイズを与えるエンコーダコントロールにおいて使用される概念である。全ての定義されたプロファイル@レベル(P@L)に対し、最大V BVバッファサイズ値は、ISO/IEC 13818-2の第8節に指定されている。より小さい値も許容され、V BVバッファサイズエレメントによって各ビットストリーム中でシグナリングされ得る。

可変ビットレートアプリケーションにおいて、ビットレートを平滑化するためエンコーダバッファを使用する切迫した必要性はないが、以下の説明から明らかになるように、バッファを確保することが有用である。変更点は、チャンネルビットレートが可変になることである。

固定ビットレート及び可変ビットレートの両方のアプリケーションにおいて、MPEGエンコーダ及びMPEGデコーダは周知のハイブリッドDPCM/DCT構造を使用することにより動作する。

可変ビットレートデータストリームを処理する能力のある記憶装置を使用するアプリケーションのためのビットレート制御の目的は、固定ビットレートの場合とは相違する。望ましくは、

- ・時間に関して均一な品質を維持し、
- ・所定の容量の媒体上で記憶時間を最大限に延ばすべきである。

従って、エンコーダ制御戦略はこの状況に適合されるべきである。

このエンコーダ制御の理解を助けるため、最初に固定ビットレートに対するエンコーダ制御を説明する。

従来のエンコーダ制御戦略は3段階又は制御レベルを含む。

#### 1) ビット割付(大域的制御)

異なる符号化フレーム対応(I, P, Bフレーム)へのビットの割付は、相対的な複雑さ $X_i$ ,  $X_p$ ,  $X_b$ に従って決められる。その目標は、フレームタイプの主観的な品質を一致させることである。事前解析を行わない以下のフィードバック実装の場合、ビデオシーンはある時間に亘って十分に静止していると考えられるので、過去のフレームからの情報は次のフレーム又はフィールドのビット割付のため使用することができる。以下の説明において、用語“フレーム”は、一般的に、フレーム若しくはフィールド、又は、(以下に定義される)ユニットよりも大きいGOP(グループオブピクチャ)の他のセクションを示す。

複雑さは、所定の画質でI, P及びBフレームを符号化する相対的なコストを反映する規準によって評価され得る。しかし、エンコーダの複雑さに負担を掛けすぎることなく、必要な測定量が得られるかどうかを考慮しなければならない。

最初に“画質”が測定されるべきである。しかし、一般的に合意された十分に簡単な客観的な測定規準は存在しないことは周知である。従って、通常、再構成されたフレームのS

10

20

30

40

50

N比が画質の近似として扱われる。MPEGエンコーダ制御の場合に、この測定規準は、各符号化フレームに対するSN比の計算を含むので幾分コスト高である。従って、その代わりとしてフレームを符号化するため使用された平均量子化ステップサイズ ( $Q_i$ ,  $Q_p$ ,  $Q_b$ ) が品質指標であるとみなされる。“品質”の測定規準が定義されたので、フレームの“複雑さ”は、発生されたビットと使用された量子化ステップの積によって表現される。換言すると、フレームの復号化が複雑化するの、量子化ステップサイズが粗いにもかかわらず、フレームが多数のビットを生成する場合である。この積は各ユニット  $m$  (例えば、マクロブロック、ブロック、スライス、又は、フレームよりも小さい任意の区画) に対し形成され、フレーム全体に関して加算される。

$$x_m = b_m * q_m \quad (1a)$$

$$X_c = \sum x_m \quad (1b)$$

式中、 $b_m$ はユニット  $m$ を符号化するためのビット数であり、 $q_m$ はユニット  $m$ で使用される平均量子化サイズであり、 $c = i, p$ 又は $b$ は符号化タイプである。

各グループオブピクチャGOPの先頭で、現在のGOP  $p$ に関して利用可能なビットバジェット(ビットの蓄え) $B_G(g)$ は、先行のGOPからの繰り上げ $B_G(g-1)$ も考慮して、

$$B_F(n=0) = B_G(g) = (\text{ビットレート}) * (\text{1 GOP当たりのフレーム}) / (\text{フレームレート}) - B_G(g-1) \quad (2)$$

のように表される。現在のGOPの各フレーム  $n$  に対し、符号化タイプ  $c$  に依存したビットの目標数  $T_c$  は、計算された複雑さ及び現在のGOP内の各符号化タイプの残りのフレーム数 ( $N_i, N_p, N_b$ ) に従って決定される。GOPのフレーム  $n-1$  の符号化後に、残りのビット  $B_F(n)$  を考慮することにより、以下の式を使用することができる。

$$T_c = B_F(n) * (X_c * K_c) / (N_i * X_i * K_i + N_p * X_p * K_p + N_b * X_b * K_b) \quad (3)$$

係数  $K_c$  はフレーム符号化タイプの種々の視覚的重要性を反映するため使用され、典型的に  $K_i = K_p = 1$  かつ  $K_b < 1$  である。

式(1a)及び(1b)に従う複雑さの計算は、通常、シーン変化に対する合理的な応答を保証するため、最新のI、P及びBフレームからのデータを用いて、各フレームの後に更新される。GOPの残りのフレームに対するビットバジェット

$$B_F(n+1) = B_F(n) - b_m \quad (4)$$

は、現在のフレームのため使用されたビット数  $b_m$  を減じた後、各対応の残りのフレーム数にセットされた  $N_i, N_p, N_b$  に関して式(3)を用いて再分配される。

GOPの最後のフレーム  $N$  を符号化した後、残りのビット

$$B_G(g) = B_F(N)$$

は、次のGOP  $g+1$  に繰り越される。

また、長時間の平均ビットレートだけを一定に保ちながら1GOPあたりに不定の量のビットを明示的に許容してもよい。實際上、これは、式(2)を用いてGOPの先頭だけでバジェットを再度埋める代わりに、各フレームの後に、

$$B_F(n) = B_F(n-1) - b_m + (\text{ビットレート}) / (\text{フレームレート}) \quad (5)$$

を用いて、ビットバジェット  $B_F(n)$  を再度埋めることにより実現される。このバジェットは、次に、現在のGOP内の残りのフレームだけではなく、GOP内のI、P及びBフレームの数に従って分配される。GOPの途中でシーン変化が生じた場合に、上記の変更を行う効果として、シーン変化のあるPフレームはより多くのビットを獲得し、品質を改良する。しかし、VBVの使用量及び/又はビットバジェットの減少に起因して、殆どの場合に、次の通常のIフレームは必要なビットを獲得しないので、次のGOPの全体的な品質は低下する。従って、好ましくは、“1GOPあたりに一定のビットバジェット”概念が固定ビットレートアプリケーションに使用される。

各フレーム毎に最小限のビットバジェットが少なくともオーバーヘッド情報を符号化するため必要である。これは、複雑さの一番少なく、非常に少ないビットしか割り付けられない傾向があるBフレームの場合に最も重要である。固定ビットレートのレート制御は、固

10

20

30

40

50

定の最小限ビットバジェットを使用してもよいが、必要なオーバーヘッドビットバジェットの評価は過去のフレームに基づいて行われる場合がある。

フレーム（通常はIフレーム）に割付可能な最小限のビット数は、V B Vの状態によって制限される。最初に空のエンコーダバッファが与えられた場合に、V B Vバッファサイズのビットを上回るビットは単一のフレームで消費され得ない。

2) バッファ制御 (局所的制御)

個々のフレーム毎に、ビット生成は、制御理論の比例コントローラに対応したバーチャルバッファ概念を使用して、割り付けられた目標ビット数の範囲内に収まるように制御される。この概念は図5に示されている。バーチャルバッファV B U<sub>c</sub>は、M P E G 2エンコーダM E Nの各ユニット(マクロブロック)によって発生されたビットb<sub>m</sub>と、Mが1フレーム当たりのユニット数を表す場合にこのユニットに対する目標の部分t<sub>m</sub> = T<sub>c</sub> / Mとの間の差e<sub>m</sub>が補充される。

10

一つのフレームタイプについて目標と合致する困難さは他のフレームタイプに伝搬すべきではないので、できる限り異なるサイズV<sub>co</sub>を有する独立したバーチャルバッファV B U<sub>c</sub>がフレーム符号化タイプc = i, p, b毎に使用される。フレームの端におけるバーチャルバッファV B U<sub>c</sub>の状態d<sub>m</sub>は同じ符号化タイプの次のフレームにd<sub>o</sub>として伝搬する。

$$d_m = d_o + (b_m - t_m) \quad (6)$$

各ユニットmに対し、量子化ステップサイズq<sub>m</sub>がこのバーチャルバッファのサイズV<sub>co</sub>に対する使用量d<sub>m</sub>から計算手段Q Cにおいて計算される。

20

所謂“非線形量子化器”テーブルは(非線形量子化とは称されていないがISO/IEC 13818-2のテーブル7-6を参照のこと)、量子化スケールコード(quantiser\_scale\_code)用の利用可能な31個の符号語の、実際の量子化ステップサイズq<sub>m</sub>への十分に進歩的な非線形マッピングを獲得するため使用される。量子化スケールコードは単純に32 \* d<sub>m</sub> / V<sub>co</sub>に設定してもよい。このテーブルによって利用可能な非常に粗い量子化は、バッファが充満状態に近いときに“パニック”状況进行处理するため使用され得る。

制御理論では比例コントローラの“反応”パラメータと称されるバーチャルバッファのサイズV<sub>co</sub>は、制御のきつさを決定する。そのサイズは、発生されたビットの数が目標からかなり偏る場合に、非常に(場合によってはV B Vよりも)大きくすることが可能である。

30

シーン変化が現在のG O Pで生じたか否かは先験的には分からないので、大きいバーチャルバッファは、例えば、Pフレームのため使用可能である。シーン変化がある場合、ビット生成は予測できない程高い。しかし、これにより画質が低下するのは、量子化ステップが著しく増加される程度にバーチャルバッファが使用された場合に限られる。予測できなかった高いビット生成に対する反応はできるだけ遅延され、一方、予測できない低いビット生成は直ちに量子化ステップサイズを減少させる方が有利である。これは、例えば、ユニットmを符号化しているときのバーチャルバッファ使用量d<sub>m</sub>とフレームの開始時の初期使用量d<sub>o</sub>との関係に依存する異なる反応パラメータV<sub>co</sub>、V d<sub>co</sub>(即ち、バーチャルバッファサイズ)によって以下の式を用いて実現され得る。

もし d<sub>n</sub> > d<sub>o</sub> ならば

$$\left. \begin{aligned}
 & \text{量子化スケールコード} = 32 * d_n / V_{co} \\
 & \text{さもなければ} \\
 & \text{量子化スケールコード} = 32 * \\
 & \quad (d_o / V_{co} - (d_o - d_n) / V d_{co})
 \end{aligned} \right\} (7)$$

40

d<sub>m</sub>が減少するときの反応であるV d<sub>co</sub>はV<sub>co</sub>よりも小さくする必要がある。q<sub>m</sub>は、ISO/IEC 13818-2のテーブル7-6を参照することによって得られる。

V B Vの状態は、バーチャルバッファに基づく制御に加えて、特に、制御が“ゆるい”と

50

き、即ち、バーチャルバッファが大きいとき、監視されるべきである。フレームの目標並びにGOP全体の目標が超えられている間、デコーダ物理バッファを表すV B Vは、決してアンダーフロー若しくはオーバーフローしない。

また、M P E Gは、ある程度のビット数を上回るユニット(マクロブロック)の頻出を許可しない(ISO/IEC 13818-2のテーブル8 - 6を参照のこと)。これも同様に監視、制御されるべきである。

### 3) 内容適応的量子化

画像の異なる部分の局所的な複雑さの変化を反映するため、ユニット毎の局所的な複雑さ測定規準が計算される。一般に、測定規準はユニットの輝度ブロックのACエネルギーから得られる。

獲得された複雑さの係数は、先に計算された量子化ステップサイズをある程度、例えば、0.5乃至2.0の範囲内の倍率で修正するため使用される。この最終的な(量子化スケールコードとして符号化された)量子化ステップサイズ情報は、先のユニット(マクロブロック)に対する量子化ステップサイズ情報と異なるときに限りビットストリームで伝送される。この目的は、フレーム内の主観的な品質が一様であるようにビットを分散させることである。人間の視覚系の特性を活用する別のスキームを適用してもよい。

好ましくは、制御レベル3)は、粗い量子化の場合にはビット生成の厳密な制御が行えなくなるので禁止される。

レベル1)のビット割付が成功した場合、レベル2)は主として、フレーム全体を通じて一定のバーチャルバッファ状態、即ち、一定量子化ステップサイズを維持すべきである。しかし、ビット割付が失敗し、及び/又は、複雑さがフレーム内で非常に不均一に分散される場合があるので、實際上、上記の点は必ずしも完全に真ではない。

制御スキームの準最適性能の理由は、ステップ1)で計算された複雑さの値は所定の画像に対し一定ではなく、容易にモデル化できない形で量子化ステップサイズに依存することである。これは12個の異なるシーケンスのIフレームについて量子化ステップサイズQに対する複雑さXを表す、図6に示されている。自然な画像シーケンスを符号化するとき、たとえ、シーンが完全に時間的に一様であるとしても、後続のPフレームに対するビット割付にある程度の振動が観察される。適用された量子化ステップサイズとは関係がない複雑さの測定規準が使用される場合、この挙動は回避される利点がある。

可変ビットレートアプリケーションの場合、エンコーダ及びデコーダのバッファから始めて以下の点を更に考慮する必要がある。

上記の如く、物理バッファは可変ビットレートの場合でも有利である。その理由は、他のフレームタイプよりも膨大な量のビットがIフレームによって生成されるからである。例えば、5 M b i t / sの固定ビットレートを用いて合理的な品質で標準的なTV画像(C C I R 601解像度)を符号化するとき、“瞬時データレート”が(フレームによって発生されたビット)/(1フレーム期間)により定義される場合に、Iフレーム用の典型的な瞬時データレートは、15 M b i t / sよりも高くなり得る。

記憶装置のピークデータレートがこの値よりも低い限り、バッファは、固定ビットレート符号化よりも高い品質を得ることができなければならない。メインプロファイル@メインプロファイル(Main Profile @ Main Profile)に対する1.75 M b i t / sのV B Vバッファサイズは、例えば、初期に空のバッファを想定するのに充分である45 M b i t / sにピーク瞬時データレートを制限する。

エンコーダバッファの可変ビットレート出力及びデコーダバッファの可変ビットレート入力、漏出しやすいバケツとしてモデル化することができる。データは、常にチャンネルのピークレートで漏出する。データはデータがエンコーダバッファ内にある限りエンコーダバッファから漏出し、そのデータはデコーダバッファが一杯ではない限りデコーダバッファに漏入する。

制御目的のため、物理的バッファ状態は1フレーム(フィールド)期間毎に1回だけ監視することができる。しかし、これは、サイズのバッファ部分[(ピークデータレート)/(フレーム期間)]が、フレーム期間中に起こり得る一様ではないビット生成に対処する

10

20

30

40

50

必要があることを意味する。ユニット毎（マクロブロック毎）に基づくバッファ制御はこの状況を回避するため使用され得る点が有利である。

可変ビットレート制御は以下の目標に到達する必要がある。

始めに仮定されたように、可変ビットレートエンコーダ制御は、記憶媒体上の記録の全区間に亘って品質を一致させることを目標とすべきである。特に、ビットは、付加的（視覚的）な品質の利益が期待できない限り使われるべきではない。レート制御の点から、2通りのシナリオを区別することができる。

シナリオ1) ビデオの最大区間を所定の容量の記憶媒体に合わせる：

十分な品質は制御ステップ2) から量子化ステップサイズを適当な値に設定することによりアプリケーション設計者によって選択され得る。ステップ3) の内容適応的量子化は、各フレーム内で主観的品質の調和を扱うので、依然として行われるべきである。實際上、この領域における改良は、人の観察者によるアーティファクト可視性をよりよく反映する視覚モデルが使用される場合に可能である。

レート制御はエンコーダバッファのオーバーフローを防止するためにのみ必要である。可変チャンネルデータレートは一時的に零にすることが許可されるので、オーバーフローは可変ビットレートアプリケーションでは発生しない。

シナリオ1 a) 保証されたビデオの最小区間：

以下のシナリオ2) に記載されるようなレート制御を用いて、媒体上の最小記録区間を保証することが可能になる。ビデオの複雑さに依存して、記録区間はこの最小値よりも大きくしてもよい。

シナリオ2) 所定の容量の媒体上に所定区間のビデオを合わせる：

ビデオ題材を通じて複雑さの分布が先験的に解析されていない場合を想定する。量子化ステップサイズとは無関係であり、かつ、合理的な低コストのエンコーダから入手可能な変数から導き出すことができる利用可能な複雑さの測定規準が存在しない限り、その区間を通じてビデオ品質を一致させる制御解決法はある種の発見的であって最適ではない方法を使用する必要がある。

可変ビットレート制御アプローチの基本的な考え方は、始めからビットレートが変化する非常にゆるい制御を有し、一方、記憶媒体の容量の端に達したとき、又は、予め選択された最小ビットが符号化されるべき画像題材の残りの区間を通じて保証できないようなとき、制御がきつくなることである。

しかし、このアプローチでは、制御の始めを特別に扱うことにより一定の品質を示すことができるが、制御の終わりに向かって品質の低下する可能性がある固定レートが実施されることに注意する必要がある。

この欠点は、記録の終わりに向けて使用される固定ビットレート  $R_0$  が記録全体の平均ビットレート  $R_{avg}$  よりも高く選択される場合、ある程度緩和することができる点が有利である。平均ビットレートは、

$$R_{avg} = B / T \quad (8)$$

によって計算することができ、式中、 $B$  は総記憶媒体容量であり、 $T$  はビデオ題材区間である。

このビットレート  $R_0$  は、記録の最初の部分の品質が固定ビットレート符号化によって全体を記録した場合よりも優れている状態を維持されつつ、終わりの方の殆どのビデオ内容が許容可能な品質を示すように選択される方が有利である。

上記の説明から分かるように、各種の記録に対し、未知変数は符号化の開始時の量子化ステップサイズと、終わりの方の固定ビットレート  $R_0$  である。残念ながら、ビデオの複雑さに関する知識がない場合、所定の量子化ステップサイズでのビット生成を評価することは困難であるので、最初に使用する適切な量子化ステップサイズを見つけることは容易に実現できない。

また、全ての時間に物理的バッファ使用量を監視し、一定の量子化から固定レート動作へ移行させる制御スキームを構築することは難しい。この移行は、時点  $t$  までに生成されたビットの量と、未だ符号化されるべき区間とに依存する。初期に著しく多量のビットが生

10

20

30

40

50

成され、残りの区間が短い場合、きついレート制御が必要であることを意味する。

従って、以下の修正レート制御スキームが提案される。

全記録に亘ってビットレートを制御し続けることが好ましい。非常にゆるい制御 (= 非常に大きいバッチャルバッファ) を有し、かつ、始めの1GOP当たりの一定ビットバジェット  $B_G(g)$  を除去することにより、瞬時ビットレート (即ち、準固定量子化ステップサイズ) のかなり大きい変動を許容することが可能になるので、少なくともシーン変化及び数秒の区間の異なるシーンは、このゆるい制御による恩恵をうける。

第1に、記録の最初及び最後における平均ビットレートの合理的な値を定義する必要がある。この記録の平均ビットレートが式(8)に従って計算された  $R_{avg}$  であるとする、画像シーケンス全体が時間的に一定の複雑さを有する場合、平均ビットレートの0.95倍のビットレートは最悪の場合に無視可能な損失を生じ、一方、最後で平均ビットレートの1.5倍のビットレートは、偶然最後に発生した略最悪の場合の画像内容を十分に適切に符号化し得ることを考慮して、例えば、

$$R_{start} = 0.95 * R_{avg} \quad (9a)$$

$$R_{end} = 1.5 * R_{avg} \quad (9b)$$

とするのが有利である。

全時間に亘って ( $R_{start}$  から  $R_{end}$  まで直線的に移行する場合)、 $R_{avg}$  は、

$$R_{avg} = R_{start} * \tau_1 + 0.5 * (R_{start} + R_{end}) * \tau_2 + R_{end} * \tau_3 \quad (10a)$$

により構成されると考えられ、但し、時間間隔  $\tau_1$ 、 $\tau_2$  及び  $\tau_3$  は、

$$\tau_1 + \tau_2 + \tau_3 = 1 \text{ (記憶媒体の総時間)} \quad (10b)$$

として正規化される。

上記の仮定の場合、もう一つのパラメータを選択する必要がある。 $\tau_3$  を0.01に設定することは、ビット生成全体の精密な制御が非常に短い時間に再確立され得ることを表す。例えば、1時間のビデオ区間Tの場合、 $\tau_3 * T = 36$ 秒に対応する。 $\tau_3$  は全体的なビット数適応期間であるとみなし得る。式(9a)乃至(10b)を解くことにより、

$$\tau_1 = 0.79 \text{ かつ } \tau_2 = 0.20$$

が得られる。

上記の数値を用いる場合、ゆるい制御は記録時間の約79%の区間で行われ、一方、移行フェーズは総時間の約20%で生じる。ゆるい制御の時間の割合が増加した別の移行特性を実現することが可能である。

レート制御の時間的変動パラメータは以下の式を用いて記述される。

$$\alpha = \begin{cases} 0 & , \tau < \tau_1 \\ (\tau - \tau_1) / (\tau_2 - \tau_1), & \tau_1 \leq \tau \leq \tau_2 \\ 1 & , \tau > \tau_2 \end{cases} \quad (11)$$

時点  $\tau_1$  と  $\tau_2$  の間に に関する別の移行特性を実現することが可能である。

レート制御のステップ1)で使用されるビットレートは、

$$R(\tau) = (1 - \alpha) * R_{start} + \alpha * R_{end} \quad (12)$$

である。

バッチャルバッファサイズ  $V_{co}$ 、 $c = i, p, b$  は、以下の通り修正する方が有利である。

$$V_c(\tau) = ((1 - \alpha) * V_{huge} + \alpha) * V_{co} \quad (13)$$

例えば、 $V_{huge} = 100$  のサイズを増加させてもよい。非常に遅い反応に対応するこの巨大なバッチャルバッファは、複雑な画像題材の短い時間間隔を意図している。しかし、長い複雑な部分に続くビデオの複雑ではない部分は、非常に粗い量子化ステップサイズで符号化できることは望ましくない。従って、ビット生成 (= バッチャルバッファ状態) が予測以上に減少する場合、即ち、目標よりも低下する場合に、上記のより迅速な反応パラメータの考え方を必要とする必要がある。このパラメータ  $V_{dc}$  は、

$$V_{dc}(\tau) = ((1 - \alpha) * V_{dc\_huge} + \alpha) * V_{dc\_co} \quad (14)$$

に従って時間的に変動するが、この場合、全区間に亘って拘束反応が望ましいので、初期推測値として  $V d_{\text{huge}} = 1$  が推奨される。

ビットバジェット  $B_F(n)$  の再補充は 1 GOP につき 1 回以上行われる。フレーム  $n$  についての再補充は、

$$B_F(n) = B_F(n-1) + (1 - \quad) * R(\quad) / (\text{フレームレート})$$

GOP  $g$  についての再補充は、

$$B_F(n=0) = B_G(g) = \quad * R(\quad) * (1 \text{ GOP 当たりのフレーム}) / (\text{フレームレート}) + B_G(g-1)$$

である。

これらは、以下に説明するようにエンコーダバッファ、即ち、漏れやすいバケツとしての VBV を空にすることは別に、本明細書の最初のパラグラフで説明したように固定ビットレート制御に対してなされるべき唯一の変更であり得る。

原則として、本発明のデジタルビデオデータを符号化する方法は、所定のデータ記憶容量を有し、可変的な瞬時記憶ビットレートを許容するメモリ又は記録手段上の記憶に相当であり、

- 上記記憶容量の前半部分に対し、データは、平均ビットレートよりも僅かに低いビットレートと、平均ビットレート制御よりも著しくゆるい制御とを用いて符号化され、

- 上記記憶容量の後半部分に対し、データは、平均ビットレートよりも高いビットレートと、きついビットレート制御とを用いて符号化される。

原則として、本発明の方法は、所定データ記憶容量を有し、復号化バッファ手段への可変的な瞬時読み出しビットレートを許容するメモリ又は記録手段に記憶されたデジタルビデオデータの復号化に相当であり、

- 上記記憶容量の前半部分の間に、データは、平均デコーダバッファ手段入力ビットレートよりも僅かに低いビットレートと、平均許容可能バッファ使用量レベル変動よりも著しく高いレベル変動とを用いて復号化され、

- 上記記憶容量の後半部分の間に、データは、平均デコーダバッファ手段入力ビットレートよりも高いビットレートと、平均許容可能バッファ使用量レベル変動よりも低いレベル変動とを用いて復号化される。

本発明の方法の有利な更なる実施例は夫々の従属請求項に記載されている。

原則として、本発明の装置は、所定のデータ記憶容量を有し、可変的な瞬時記憶ビットレートを許容するメモリ又は記録手段上の記憶用のデジタルビデオデータの符号化に相当であり、

離散コサイン変換手段と、量子化手段と、逆量子化手段と、逆離散コサイン変換手段と、画像記憶手段と、符号化された出力信号を送出し上記量子化手段の出力に接続されたエントロピー符号化手段とを含む符号化ループを使用し、

上記量子化手段、上記逆量子化手段及び上記エントロピー符号化手段は、

- 上記記憶容量の前半部分に対し、データは、平均ビットレートよりも僅かに低いビットレートと、平均ビットレート制御よりも著しくゆるい制御とを用いて符号化され、

- 上記記憶容量の後半部分に対し、データは平均ビットレートよりも高いビットレートと、きついビットレート制御とを用いて符号化されるように、制御される。

原則として本発明の装置は、所定のデータ記憶容量を有し、復号化バッファ手段への可変的な瞬時読み出しビットレートを許容するメモリ又は記録手段に記憶されたデジタルビデオデータの復号化に相当であり、

エントロピー復号化手段と、逆量子化手段と、逆離散コサイン変換手段と、画像記憶手段に供給される復号化された出力信号を送出する加算手段とを含み、上記画像記憶手段の出力信号は上記加算手段の第 2 の入力に供給される回路を使用し、

上記復号化バッファ手段は、

- 上記記憶容量の前半部分の間に、データは、平均デコーダバッファ手段入力ビットレートよりも僅かに低いビットレートと、平均許容可能バッファ使用量レベル変動よりも著しく高いレベル変動とを用いて復号化され、

10

20

30

40

50

- 上記記憶容量の後半部分の間に、データは、平均デコーダバッファ手段入力ビットレートよりも高いビットレートと、平均許容可能バッファ使用量レベル変動よりも低いレベル変動とを用いて復号化されるように制御される。

本発明の装置の有利な更なる実施例は夫々の従属請求項に記載されている。

原則として、本発明のデジタルビデオ信号は、所定のデータ記憶容量を有し、可変的な瞬時記憶ビットレートを許容するメモリ及び記録手段上の記憶に適当であり、

- 上記記憶容量の前半部分に対し、上記信号は、平均ビットレートよりも僅かに低いビットレートと、平均ビットレート制御よりも著しくゆるい制御とを用いて符号化され、
- 上記記憶容量の後半部分に対し、上記信号は、平均ビットレートよりも高いビットレートと、きついビットレート制御とを用いて符号化される。

10

好ましくは、本発明のデジタルビデオ信号は、MPEG2若しくはMPEG1、又は、H261に準拠して符号化されたビデオ信号である。

本発明の記憶媒体、特に、光ディスクは、本発明のデジタルビデオ信号を格納するか、又は、本発明のデジタルビデオ信号が記録されている。

図面の説明

以下、添付図面を参照して本発明の好ましい実施例を説明する。

図面中、

図1は固定ビットレートシステムを表し、

図2は可変ビットレートシステムを表し、

図3はMPEGエンコーダの基本的なブロック図であり、

20

図4はMPEGデコーダの基本的なブロック図であり、

図5は制御ループの構造を示す図であり、

図6は量子化ステップサイズに対する複雑さの関数の例を表わす図である。

好ましい実施例の説明

固定ビットレート又は可変ビットレートアプリケーションにおいて、MPEG2エンコーダMENCは、エンコーダバッファENCBに中間的に記憶される可変ビットレートデータVBRDを生成する。このバッファの使用量レベルはMPEG2エンコーダMENCの符号化処理を制御するため使用され得る。受信側で、デコーダバッファDECBに中間的に記憶された可変ビットレートデータVBRDは、MPEG2デコーダMDECに供給される。デコーダバッファDECBからの可変ビットレートデータVBRDの読み出しはMPEG2デコーダMDECによって制御される。

30

図1において、エンコーダバッファENCBは固定ビットレートデータストリームCBRをデコーダバッファDECBに送出する。しかし、図2では、エンコーダバッファENCBは可変ビットレートデータストリームVBRをデコーダバッファDECBに送出する。図3において、エンコーダビデオデータ入力信号IEは、符号化されるべきマクロブロックデータを含む。イントラフレームデータの場合、減算器SUBは、離散コサイン変換計算手段DCTと量子化手段Qを介して、エンコーダバッファを含むエントロピー符号化手段ECODに上記イントラフレームデータをそのまま伝達する。エントロピー符号化手段ECODはエンコーダビデオデータ出力信号OEを送出する。

インターフレームデータの場合、減算器SUBは、入力信号から予測マクロブロックデータPMDを減算し、離散コサイン変換計算手段DECと量子化手段Qを介して、エントロピー符号化手段ECODに差データを伝達する。量子化手段Qの出力信号は、逆量子化手段 $Q^{-1}_E$ にも供給され、逆量子化手段 $Q^{-1}_E$ の出力信号は、逆離散コサイン変換計算手段 $DCT^{-1}_E$ を介して、再生マクロブロック差データRMDDの形で加算器ADDEに伝達される。加算器ADDEの出力信号はフレームストア及び動き補償手段FS\_\_MC\_\_Eに中間的に記憶され、動き補償手段FS\_\_MC\_\_Eは再生マクロブロックデータRMDDに基づいて動き補償を行い、予測マクロブロックデータPMDを減算器SUB及び加算器ADDEの別の入力に送出する。

40

量子化手段Q、逆量子化手段 $Q^{-1}_E$ 及びエントロピー符号化手段ECODは、上記の本発明の方法に従ってエンコーダバッファの使用量レベルによって制御される。また、動き補

50

償手段  $FS\_MC\_E$  は動き補償を行わない場合もあり得る。

記憶媒体への可変ビットレート出力は、エンコーダバッファから利用可能なデータが存在する場合に、最大可能ビットレートを用いて行うことが可能であり、さもなければ、停止することが可能である。

図4において、デコーダビデオデータ入力信号  $ID$  は、デコーダバッファを含むエントローピー復号化手段  $EDEC$  と、逆量子化手段  $Q^{-1}_D$  と、逆離散コサイン変換計算手段  $DCIT^{-1}_D$  とを介して、デコーダビデオデータ出力信号  $OD$  を送出する加算器  $ADDD$  に伝達される。加算器  $ADDD$  の出力信号は、フレームストアに供給され、動き補償手段  $FS\_MC\_D$  は再生マクロブロックデータに基づいて動き補償を行い、インターフレームマクロブロックデータの場合に、予測マクロブロックデータ  $PMD$  を加算器  $ADDD$  の他方の入力に送出する。イントラフレームマクロブロックデータの場合、加算器  $ADDD$  は、逆離散コサイン変換計算手段  $DCIT^{-1}_D$  の出力信号をそのまま伝達する。

デコーダ側の逆量子化手段  $Q^{-1}_D$  及びエントローピー復号化手段  $EDEC$  は、エンコーダ側の逆量子化手段  $Q^{-1}_E$  及びエントローピー符号化手段  $ECOD$  に対する制御と同様の制御に従って制御される。通常、デコーダ側の動き補償手段  $FS\_MC\_D$  において、エンコーダ側の動き補償手段  $FS\_MC\_E$  で使用された動きベクトル情報と同じものが使用される。また、デコーダ側の動き補償手段  $FS\_MC\_D$  は動き補償を行わなくてもよい。

例えば、シーン変化及び複雑なシーン内容の場合のように、高ビットレート要求の状況に対処するため、デコーダバッファは、好ましくは最大可能ビットレートを用いて記憶媒体から補充される。これは、一般的に、デコーダバッファが一杯ではない場合に行われ、デコーダバッファが一杯である場合には、記憶媒体からのデータはデコーダバッファに記憶されない。代替として、デコーダバッファは、データをメモリ又は記録手段に記憶する際の夫々のビットレートに対応したビットレートで補充してもよい。何れの場合も、デコーダバッファからの読み出しビットレートは符号化する際のビットレートと対応し、即ち、許容され、かつ、場合によっては実際に生ずる平均バッファ使用量レベル変動は、総容量の前半部分（80%）の方で大きく、総容量の後半部分（20%）の方で小さい。“平均許容可能バッファ使用量レベル変動”は、一部のGOPの長さに関する変動を意味する。

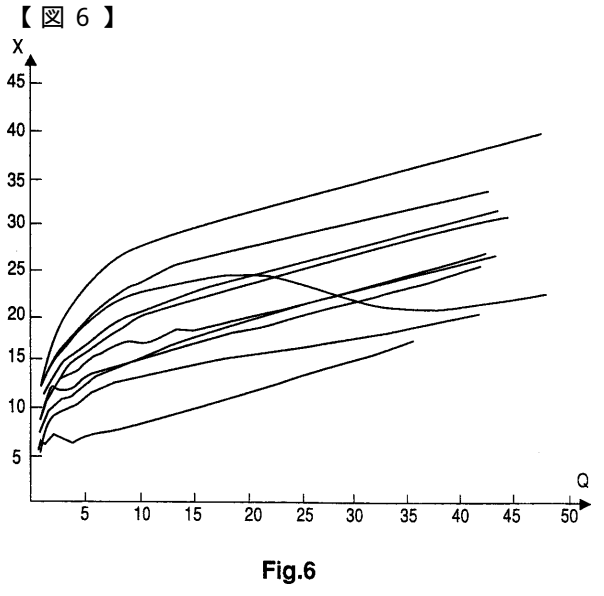
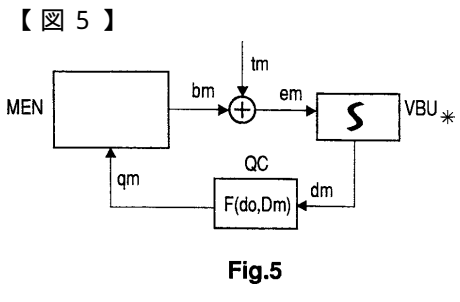
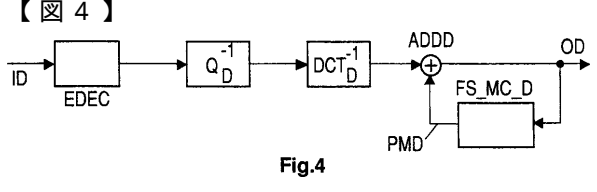
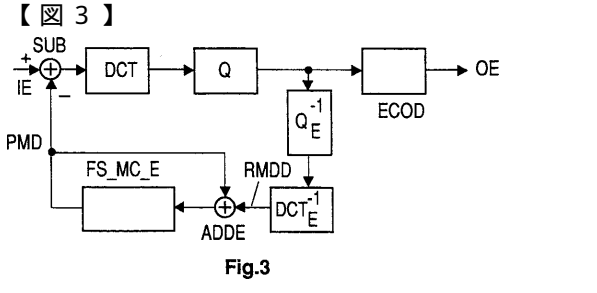
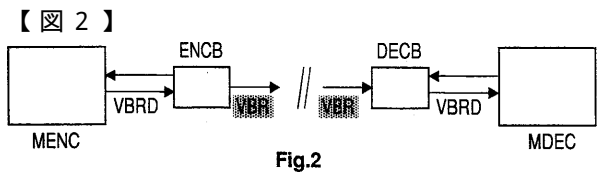
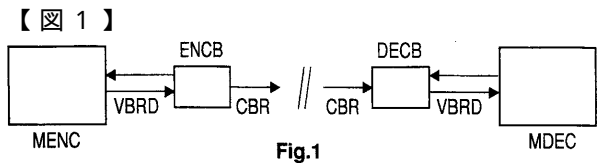
本発明は、ビデオデータを記録するため所定の容量を使用するDVD、他の光学式、光磁気式若しくは光磁性式ディスク、電子式若しくは光学式RAM又はROM、及び、磁気テープシステムのようなあらゆる手段で使用することができる。

本発明は、MPEG2エンコーダ及びデコーダに限定されることはなく、MPEG1 (ISO/IEC 11172-2)、H261 (ISO/IEC)、並びに、例えば、標準TV、HDTV及びテレビ電話アプリケーションにおける他の符号化及び復号化方式と共に使用することが可能である。

10

20

30



---

フロントページの続き

(72)発明者 ヘルペル, カルシュテン  
ドイツ連邦共和国, 3 0 1 7 1 ハノーヴァー, グローセ バルリンゲ 6 1

審査官 矢野 光治

(56)参考文献 特開平08 - 017137 (JP, A)  
特開平11 - 103446 (JP, A)  
特開2000 - 008529 (JP, A)

(58)調査した分野(Int.Cl., DB名)  
H04N 7/24 - 7/68