

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第5124562号
(P5124562)

(45) 発行日 平成25年1月23日(2013.1.23)

(24) 登録日 平成24年11月2日(2012.11.2)

(51) Int.Cl.

H04N 7/32 (2006.01)

F 1

H04N 7/137

Z

請求項の数 7 (全 55 頁)

(21) 出願番号	特願2009-504355 (P2009-504355)
(86) (22) 出願日	平成19年4月9日(2007.4.9)
(65) 公表番号	特表2009-533900 (P2009-533900A)
(43) 公表日	平成21年9月17日(2009.9.17)
(86) 國際出願番号	PCT/US2007/008937
(87) 國際公開番号	W02007/117707
(87) 國際公開日	平成19年10月18日(2007.10.18)
審査請求日	平成22年3月10日(2010.3.10)
(31) 優先権主張番号	11/400,788
(32) 優先日	平成18年4月7日(2006.4.7)
(33) 優先権主張国	米国(US)

(73) 特許権者	500046438 マイクロソフト コーポレーション アメリカ合衆国 ワシントン州 98052 2-6399 レッドモンド ワン マイ クロソフト ウェイ
(74) 代理人	100077481 弁理士 谷 義一
(74) 代理人	100088915 弁理士 阿部 和夫
(72) 発明者	シーピン ツォ アメリカ合衆国 98052 ワシントン 州 レッドモンド ワン マイクロソフト ウェイ マイクロソフト コーポラシ ョン インターナショナル パテンツ内

最終頁に続く

(54) 【発明の名称】テクスチャレベルに基づく量子化調整

(57) 【特許請求の範囲】

【請求項 1】

ビデオエンコーダで使用する方法であって、

ビデオピクチャシーケンス中の現在のビデオピクチャの現在の平滑領域を識別するステップであって、前記ビデオピクチャシーケンスは、複数の先行ピクチャの表示が表示順序において前記現在のビデオピクチャの表示の前に起こり、複数の将来ピクチャの表示が表示順序において前記現在のビデオピクチャの前記表示の後に起こる前記表示順序を有する、ステップと、

前記現在の平滑領域に対して時間分析を実施するステップであって、前記時間分析は、前記複数の先行及び/または将来ピクチャの少なくとも1つの中の対応する領域が平滑であるかどうか判定することを含み、前記時間分析は、量子化の調整を行うためのより重い重みを、前記現在のビデオピクチャから時間的に離れている先行及び/または将来ピクチャの対応する領域内の平滑性よりも前記現在のビデオピクチャに時間的に近い先行及び/または将来ピクチャの対応する領域内の平滑性におく、ステップと、

前記時間分析に少なくとも部分的に基づいて、前記現在の平滑領域内での量子化を調整するステップと、

前記現在のビデオピクチャに対するエンコードされたデータを出力するステップとを含むことを特徴とする方法。

【請求項 2】

前記現在の平滑領域内での量子化を前記調整するステップは、前記現在の平滑領域用の

量子化ステップサイズを小さくすることをさらに含むことを特徴とする請求項 1 に記載の方法。

【請求項 3】

前記現在の平滑領域用の量子化ステップサイズを小さくすることにより、前記現在のビデオピクチャが復元されるときの前記現在の平滑領域内での輪郭化アーチファクトの発生を防ぐことを特徴とする請求項 2 に記載の方法。

【請求項 4】

前記現在の平滑領域は階調傾斜領域であることを特徴とする請求項 1 に記載の方法。

【請求項 5】

前記ビデオエンコーダは、テクスチャマップを分析して、前記現在の平滑領域を識別し
、前記ビデオエンコーダは、前記時間分析において平均輝度値を分析することを特徴とする請求項 1 に記載の方法。 10

【請求項 6】

請求項 1 に記載の方法をコンピュータに実施させるコンピュータ実行可能命令を格納したことを特徴とする 1 つまたは複数のコンピュータ可読記録媒体。

【請求項 7】

ビデオエンコーダで使用する方法であって、

ビデオピクチャシーケンス中の現在のビデオピクチャの現在の平滑領域を識別するステップであって、前記ビデオピクチャシーケンスは、複数の先行ピクチャの表示が表示順序において前記現在のビデオピクチャの表示の前に起こり、複数の将来ピクチャの表示が表示順序において前記現在のビデオピクチャの前記表示の後に起こる前記表示順序を有する、ステップと、 20

前記現在の平滑領域に対して時間分析を実施するステップであって、前記時間分析は、前記複数の将来ピクチャの少なくとも 1 つの中の対応する領域が平滑であるかどうか判定することを含み、前記時間分析は、前記将来ピクチャの 1 つまたは複数に対して実施される、ステップと、

前記時間分析に少なくとも部分的に基づいて、前記現在の平滑領域内での量子化を調整するステップと、

前記現在のビデオピクチャに対するエンコードされたデータを出力するステップと
を含むことを特徴とする方法。 30

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、映像におけるテクスチャレベルに基づいて量子化を調整する技術及びツールに関する。

【背景技術】

【0002】

DVD、インターネットを介した音楽配信及びデジタルカメラがますます普及するのに伴い、デジタルメディアが一般的になっている。エンジニアは、品質を保ちながら、様々な技術を用いて、デジタル音声、映像及び画像を効率的に処理する。こうした技術を理解するために、音声、映像及び画像情報がコンピュータにおいてどのようにして表され処理されるかを理解することが助けとなる。 40

【0003】

I . コンピュータにおけるメディア情報の表現

コンピュータは、メディア情報を、その情報を表す一連の数字として処理する。例えば、ある 1 つの数字は、ピクチャの各基本小領域に対する明るさの強度または赤、緑もしくは青など、色成分の強度を表すことができ、その結果、ピクチャのデジタル表現は、このような数字の 1 つまたは複数の配列からなる。このような各数字は、サンプルと呼ぶこともできる。カラー画像の場合、従来は、各要素領域の色を表すのに複数のサンプルを使つたが、通常は 3 つのサンプルが使われる。要素領域用のこうしたサンプルの組は、ピクセ 50

ルと呼ぶことができ、「ピクセル」という言葉は、「ピクチャ要素」という概念を指す短縮形である。例えば、1つのピクセルは、要素領域を表すのに必要な赤、緑及び青色光の強度を表す3つのサンプルからなり得る。このようなピクセルタイプは、RGBピクセルと呼ばれる。(映像に対する)サンプルの深さ、解像度及びフレームレートを含むいくつかの要因が、メディア情報の品質に影響する。

【0004】

サンプルの深さは、サンプルを表すのに使うことができる数の範囲を示す、一般にビット単位で測定される属性である。サンプル用により多くの値が可能な場合、強度のより微妙な変化及び/またはより大きな範囲の値を数がとらえ得るので、品質はより高くなり得る。解像度は概して、ある程度の継続期間(音声の場合)または空間(画像もしくは個々のビデオピクチャの場合)に渡るサンプルの数を指す。解像度がより高い画像は、他の画像よりくっきりと見え、より認識可能な有用な細部を含む傾向にある。フレームレートは、映像に対する時間分解能に関する一般的な用語である。フレームレートがより高い映像は、自然物の滑らかな動きを他の映像よりうまくまねる傾向にあり、同様に、時間的次元でのより詳細な細部を含むものと見なすことができる。こうした要因すべてに対して、高品質のための妥協点は、表1が示すような、サンプル深さ、解像度及びフレームレートを表すのに必要なビットレートによって情報を格納し伝送するコストである。

【0005】

【表1】

ピクセルごとのビット (サンプル深さ×ピクセルごとのサンプル)	解像度 (ピクセル数、幅×高さ)	フレームレート (1秒あたりのフレーム数)	ビットレート (1秒あたりのビット数(単位:百万))
8(0~255の値、モノクロ)	160x120	7.5	1.2
24(0~255の値、RGB)	320x240	15	27.6
24(0~255の値、RGB)	640x480	30	221.2
24(0~255の値、RGB)	1280x720	60	1327.1

10

20

30

表1：異なる品質レベルの未加工映像用のビットレート

【0006】

高品質映像(HDTVなど)を格納し送るには高ビットレートが必要なのにも関わらず、企業及び消費者は、高品質コンテンツを作成し、配信し、再生するために、コンピュータにますます依存している。こうした理由により、技術者は、圧縮(ソース符号化及びソース復号化とも呼ばれる)を用いてデジタル映像のビットレートを削減する。圧縮により、情報をより低いビットレートの形にコンバートすることによって情報を格納し伝送するコストが下がる。圧縮において損失をなくすことができる場合は、映像の品質は低下しないが、ビットレートの減少は、映像の複雑さによって制限を受ける。または、圧縮において損失があり得る場合は、映像の品質は低下するが、ビットレートの減少はより目ざまい。圧縮解除(復号化とも呼ばれる)は、圧縮された形から、元の情報のバージョンを復元する。「コーデック」とは、エンコーダ/デコーダシステムである。

【0007】

概して、映像圧縮技術は、「イントラ(intra)」圧縮、及び「インター(inter)」すなわち予測圧縮を含む。映像フレームではイントラ圧縮技術は、一般にIフレームまたは

40

50

キーフレームと呼ばれる個々のフレームを圧縮する。インター圧縮技術は、先行フレーム及び／または後続フレームを参照してフレームを圧縮し、インター圧縮されたフレームは一般に予測フレーム、Pフレーム、またはBフレームと呼ばれる。

【0008】

I I . W i n d o w s (登録商標) M e d i a V i d e o バージョン8及びバージョン9におけるインター圧縮及びイントラ圧縮

本件特許出願人のW i n d o w s (登録商標) M e d i a V i d e o バージョン8 [WMV8]は、ビデオエンコーダ及びビデオデコーダを含む。WMV8エンコーダは、イントラ圧縮及びインター圧縮を使用し、WMV8デコーダは、イントラ圧縮解除及びインター圧縮解除を使用する。W i n d o w s (登録商標) M e d i a V i d e o バージョン9 [WMV9]では、多くの操作に対して同様のアーキテクチャを使用する。10

【0009】

A. イントラ圧縮

図1は、WMV8エンコーダにおける、キーとなるフレーム中のサンプルのブロック105の、ブロックに基づいたイントラ圧縮100を示す。ブロックとは、1組のサンプル、例えば、 8×8 配置のサンプルである。WMV8エンコーダは、キーとなるビデオフレームを 8×8 ブロックに分け、ブロック105などの個々のブロックに 8×8 離散コサイン変換(DCT)110を適用する。DCTとは、 8×8 ブロックのサンプル(空間情報)を、周波数情報である 8×8 ブロックのDCT係数115にコンバートするタイプの周波数変換である。DCT演算自体は、無損失(lossless)またはほぼ無損失である。ただし、元のサンプル値に比較して、DCT係数は、エンコーダが圧縮するのにより効率的である。というのは、重要な情報のほとんどが低周波係数(従来通りには、ブロック115の左上)に集中し、高周波係数の多く(従来通りには、ブロック115の右下)が、ゼロまたはゼロに近い値を有するからである。20

【0010】

エンコーダは次いで、DCT係数を量子化し(120)、結果として、 8×8 ブロックの量子化されたDCT係数125を生じる。量子化は損失性(lossy)である。低周波数DCT係数は、比較的高い値をもつ傾向にあるので、量子化は通常、結果として精度を損なわせるが、係数の情報を完全には損なわせない。一方、高周波数DCT係数は、ゼロまたはゼロに近い値をもつ傾向があるので、高周波係数の量子化は通常、結果として、値がゼロの隣接領域を生じる。さらに、いくつかのケースでは、高周波数DCT係数は、低周波数DCT係数より粗く量子化され、結果として、高周波数DCT係数の精度／情報をより多く損なわせる。30

【0011】

エンコーダは次いで、無損失圧縮の形をしたエントロピーエンコーディング用に 8×8 ブロックの量子化されたDCT係数125を用意する。エントロピーエンコーディングの正確なタイプは、係数がDC係数(最低周波数)、最上位行または左列にあるAC係数(他の周波数)、それとも別のAC係数であるかに応じて変わり得る。

【0012】

エンコーダは、DC係数126を、隣接する 8×8 ブロックのDC係数136との差分としてエンコードする。このブロックは、エンコードされたブロックの以前エンコードされた隣(例えば、最上位または左)にある。(図1は、フレーム中の、エンコードされたブロックの左に位置している隣のブロック135を示す)。エントロピーエンコーダは、その差分をエンコードする(140)。40

【0013】

エントロピーエンコーダは、左列または最上位行のAC係数を、隣接する 8×8 ブロックの対応する左列または最上位行との差分としてエンコードすることができる。これは、AC係数予測の例である。図1は、左列137の(実際には、左に)隣接するブロック135との差分147としてエンコードされた、左列127のAC係数を示す。差分符号化により、差分係数がゼロ値をもつ機会が増える。残っているAC係数は、ブロック12550

の量子化されたDCT係数からのものである。

【0014】

エンコーダは、 8×8 ブロック 145 の量子化されたAC DCT係数を、1次元配列 155 に取り込み(150)、次いで、取り込まれたAC係数を、ランレンジス符号化 160 の変形を用いてエントロピー符号化する。エンコーダは、1つまたは複数のラン/レベル/ラストテーブル 165 からエントロピーコードを選択し、エントロピーコードを出力する。

【0015】

B. インター圧縮

WMV8 エンコーダにおけるインター圧縮は、ブロックベースでの動き補償予測符号化を使用し、この予測符号化の後には、見逃しエラー(residual error)の変換符号化が続く。図2及び図3は、WMV8 エンコーダにおける、予測フレームに対するブロックに基づいたインター圧縮を示す。具体的には、図2は、予測フレーム 210 に対する動き推定を示し、図3は、予測フレームの、動きを補償したブロックに対する予測残差(prediction residual)の圧縮を示す。

10

【0016】

例えば、図2では、WMV8 エンコーダは、予測フレーム 210 中のマクロブロック 215 用の動きベクトルを計算する。動きベクトルを計算するために、エンコーダは、参照フレーム 230 のサーチエリア 235 内をサーチする。サーチエリア 235 内部で、エンコーダは、完全に一致する候補マクロブロックを見つけるために、予測フレーム 210 にあるマクロブロック 215 を、様々な候補マクロブロックと比較する。エンコーダは、一致するマクロブロック用の(エントロピー符号化された)動きベクトルを指定する情報を出力する。動きベクトルは、動きベクトル予測子を参照して差分符号化される。

20

【0017】

動きベクトル予測子に差分を加えることによって動きベクトルを再構成した後、デコーダは、参照フレーム 230 からの情報を使ってマクロブロック 215 に対する予測マクロブロックを計算するために動きベクトルを使う。参照フレーム 230 は、エンコーダ及びデコーダで利用可能な、先行して再構成されたフレームである。予測が完全であることはまれなので、エンコーダは通常、予測マクロブロックとマクロブロック 215 自体の間の、ピクセル差分からなるブロック(エラーブロックまたは残差ブロックとも呼ばれる)をエンコードする。

30

【0018】

図3は、WMV8 エンコーダにおける、エラーブロック 335 の計算及びエンコードの例を示す。エラーブロック 335 は、予測ブロック 315 と元の現在のブロック 325 との間の差分である。エンコーダは、エラーブロック 335 にDCT 340 を適用し、その結果、係数からなる 8×8 ブロック 345 を生じさせる。エンコーダは次いで、DCT係数を量子化(350)し、その結果、量子化されたDCT係数 355 からなる 8×8 ブロックを生じさせる。エンコーダは、 8×8 ブロック 355 を走査(360)して1次元配列 365 にし、そうすることによって、係数が、概して最も低い周波数から最も高い周波数の順に並べられるようとする。エンコーダは、ランレンジス符号化 370 の変形を用いて、走査された係数をエントロピー符号化する。エンコーダは、1つまたは複数のラン/レベル/ラストテーブル 375 からエントロピーコードを選択し、エントロピーコードを出力する。

40

【0019】

図4は、インター符号化ブロック用の対応する復号化処理 400 の例を示す。図4を要約すると、デコーダが、1つまたは複数のラン/レベル/ラストテーブル 415 を用いた可変長復号化 410、及びランレンジス復号化 420 を用いて、予測残差を表すエントロピー符号化情報を復号する(410、420)。デコーダは、エントロピー復号された情報を格納する1次元配列 425 を逆走査(430)して、2次元ブロック 435 にする。デコーダは、データを逆量子化し逆DCT し(共に 440)、その結果、再構成された工

50

ラープロック 445 を生じさせる。別個の動き補償経路において、デコーダは、参照フレームからの置換のための動きベクトル情報 455 を用いて、予測ブロック 465 を計算する。デコーダは、予測ブロック 465 を、再構成されたエラーブロック 445 と組み合わせて(470)、再構成されたブロック 475 を形成する。エンコーダは、逆量子化、逆 DCT、動き補償及び参照フレームとして使用するためにフレームを復元するための合成も実施する。

【0020】

I I I . 損失性圧縮及び量子化

先行セクションでは、量子化、損失性圧縮のための機構、及び無損失圧縮とも呼ばれるエントロピー符号化に言及した。無損失圧縮は、忠実性を全く低下させずに情報から冗長性を除去することによって、情報のビットレートを削減する。例えば、すべてが正確に同じ色合いの赤である、一連の 10 個の連續したピクセルは、特定の色合いの赤に対するコードとして、10 という数は連續したピクセルの「ランレンジス」として表すことができ、この連續は、圧縮解除によって、その色合いの赤に対するコード及びその色合いの赤を有する指示された数(10)の連續したピクセルから完璧に復元される。無損失圧縮技術は、品質を全く犠牲にせずにビットレートを削減するが、ビットレートを一定の点まで削減し得るだけである。ビットレートの低下は、ソースエントロピーと呼ばれる、入力データの統計的特徴抽出における、固有の量の変動性によって制限される。

10

【0021】

対照的に、損失性圧縮を用いると、品質はある程度損なわれるが、到達する可能性のあるビットレートの低下はより劇的である。例えば、それぞれがわずかに異なる色合いの赤である一連の 10 個のピクセルは、正確に同じ特定の近似の赤色の 10 個のピクセルとして近似することができる。損失性圧縮技術を用いると、ビットレートを、無損失圧縮技術より削減することができるが、ビットレート削減の一部は、品質を低下させることによって達成され、損なわれた品質は完全に回復することができない。損失性圧縮はしばしば、損失性圧縮が情報の近似を確立し無損失圧縮技術が近似を表すのに適用されるシステム設計において、無損失圧縮とともに用いられる。例えば、それぞれがわずかに異なる色合いの赤である一連の 10 個のピクセルは、ある特定の色合いの赤に対するコードとして、10 という数は連續したピクセルのランレンジスとして表すことができる。概して、エンコーダは、品質とビットレートとの兼ね合いをはかるために、量子化を変化させる。比較的粗い量子化により、結果として品質がより大きく低下されるが、ビットレートもより大きく削減される。圧縮解除において、元の連續は次いで、近似された同じ赤色の 10 個のピクセルとして復元されることになる。

20

【0022】

可能な一定義によると、量子化は、損失性圧縮に共通して使われる不可逆マップ関数の近似に使われる用語であり、特定された 1 組の可能な出力値があり、その 1 組の可能な出力値の各メンバは、結果としてその特定の出力値を選択させる、関連づけられた 1 組の入力値を有する。スカラーまたはベクトル、一様または非一様、及び適応または非適応量子化を含む、様々な量子化技術が開発されている。

30

【0023】

A . スカラー量子化法

可能な一定義によると、スカラー量子化法とは、入力値 x の、時には復元された値と呼ばれる、被量子化値 $Q[x]$ への近似機能的マッピング $x \rightarrow Q[x]$ である。図 5 は、スカラー量子化法用の「階段」I/O 関数 500 を示す。水平軸は、実数入力変数 x に対する数直線であり、垂直軸は、対応する被量子化値 $Q[x]$ を示す。数直線は、閾値 510 などの閾値によって分割される。1 対の隣接閾値の間の所与の範囲内の x の各値には、同じ被量子化値 $Q[x]$ が割り当てられる。例えば、範囲 520 内の x の各値には、同じ被量子化値 530 が割り当てられる。(ある閾値で、可能な 2 つの被量子化値の一方が、システムに応じて入力 x に割り当てられる)。全体的に、被量子化値 $Q[x]$ は、不連続な階段パターンを呈する。マッピングが数直線に沿って続く距離は、システムに依存し、通

40

50

常、有限個数の閾値の後に終わる。数直線上の閾値の配置は、(図5に示すように)均一に間隔をあけてもよく、不均一に間隔をあけててもよい。

【0024】

スカラー量子化法は、2つの別個の段階に分解することができる。第1の段階は分類法段階であり、ここで分類法関数マッピング $x \rightarrow [x]$ は、入力 x を、整数値であることが多い量子化索引 $[x]$ にマップする。本質的に、分類法は、入力数直線またはデータセットをセグメント化する。図6Aは、汎用化された分類法600及びスカラー量子化法用の閾値を示す。図5のように、実数変数 x に対する数直線は、閾値610などの閾値によってセグメント化される。範囲620など、所与の範囲内の x の各値には、同じ被量子化値 $Q[x]$ が割り当てられる。図6Bは、分類法650及びスカラー量子化法用の閾値の数字の例を示す。10

【0025】

第2段階では、リコンストラクタ機能マッピング $k \rightarrow [k]$ が、各量子化索引 k を復元値 $[k]$ にマップする。本質的に、リコンストラクタは、分類法によって判定される各領域の復元のために、入力数線分に相対したある特定の高さを有するステップを配置する(または、データセット値のサブセットを選択する)。リコンストラクタ機能マッピングは、例えば、ルックアップテーブル使って実装することができる。全体的に、分類法は、以下のようなリコンストラクタに関する。20

$$Q[x] = [x] \quad (1)$$

【0026】

一般的な使用において、「量子化」という用語はしばしば、エンコーディング中に実施される分類法段階を説明するのに使われる。「逆量子化」という用語は、同様に、エンコーディングまたは復号化どちらの間に実施されるのかに関わらず、リコンストラクタ段階を説明するのに使われる。20

【0027】

このような量子化法を使うことによって混入される歪みは、差分に基づく歪み測度 $d(x - Q[x])$ で計算することができる。通常、このような歪み測度は、 $x - Q[x]$ がゼロからそれるのに従って $d(x - Q[x])$ が増大するプロパティを有し、通常は各復元値は対応する分類領域の範囲内にあるので、関数方程式 $Q[x] = x$ によって形成されることになる直線は、階段図(図5に示す)のすべてのステップを通過することになり、したがって $Q[Q[x]]$ は通常、 $Q[x]$ に等しくなる。概して、量子化法は、量子化法が結果として、所与のビットレートの出力に対して他の量子化法より低い平均値の歪みを生じる場合、速度歪み期間中に比較的深く検討される。より形式的には、量子化法は、ソースランダム変数 X に対して、期待される(すなわち、平均または統計的中央)値の歪み測度30

【0028】

【数1】

$$\overline{D} = E_X\{d(X - Q[X])\}$$

【0029】

が $[X]$ の等価な、またはより低いエントロピー H に対してより低い場合、より深く検討される。最も共通して使われる歪み測度は、二乗エラー歪み測度であり、 $d(|x - y|) = |x - y|^2$ である。二乗エラー歪み測度が使われるとき、歪み測度40

【0030】

【数2】

$$(\overline{D})$$

【0031】

の期待値は、平均二乗エラーと呼ばれる。

【0032】

10

20

30

40

50

B . デッドゾーン + 均一閾値量子化法

不均一量子化法は、すべての分類法領域に対して、均一に間隔をあけられていない閾値を有する。可能な一定義によると、デッドゾーンプラス均一閾値量子化法（「D Z + U T Q」）は、ゼロ入力値（デッドゾーン（「D Z」）と呼ばれる）を含むものを除くすべての分類法領域に対して均一に間隔をあけられた閾値をもつ量子化法である。一般的な意味では、D Z サイズが他の分類法領域とは異なるので、D Z + U T Q は、不均一量子化法である。

【 0 0 3 3 】

D Z + U T Q は、2つのパラメータに基づいて表現することができる分類法索引マッピング規則 $x \rightarrow [x]$ を有する。図7は、D Z + U T Q に対する階段I/O関数 7 0 0 を示し、図8 A は、汎用化された分類法 8 0 0 及び D Z + U T Q に対する閾値を示す。0より大きいパラメータ s は、D Z 以外のすべてのステップに対するステップサイズを示す。数学的に、 $i = 0$ に関して、すべての s_i は s に等しい。 0 以上のパラメータ z は、D Z サイズと、他のステップのサイズとの比を示す。数学的に、 $s_0 = z \cdot s$ である。図8 Aにおいて、z は 2 なので、D Z は、他の分類ゾーンの 2 倍の広さである。

【 0 0 3 4 】

D Z + U T Q に対する索引マッピング規則 $x \rightarrow [x]$ は、

【 0 0 3 5 】

【 数 3 】

$$A[x] = \text{sign}(x) * \max\left(0, \left\lfloor \frac{|x|}{s} - \frac{z}{2} + 1 \right\rfloor\right) \quad (2),$$

20

【 0 0 3 6 】

と表現することができ、上式 (2) で、

【 0 0 3 7 】

【 数 4 】

$\lfloor \cdot \rfloor$

【 0 0 3 8 】

は、引数以下の最小の整数を示し、sign(x) は、

30

【 0 0 3 9 】

【 数 5 】

$$\text{sign}(x) = \begin{cases} +1, & x \geq 0 \text{ の場合} \\ -1, & x < 0 \text{ の場合} \end{cases} \quad (3).$$

【 0 0 4 0 】

として定義される関数である。

【 0 0 4 1 】

図8 B は、分類法 8 5 0 及び $s = 1$ 、 $z = 2$ の D Z + U T Q に対する閾値の数字の例を示す。図5、図6 A、図6 B は、 $z = 1$ の特殊なケースの D Z + U T Q を示す。U T Q 形の量子化法は、様々な統計ソースに対して良好な性能をもつ。具体的には、D Z + U T Q 形は、ラブリッシュソースとして知られる統計的ランダム変数ソースに対して最適である。

40

【 0 0 4 2 】

一部のシステム設計（図示せず）では、D Z + U T Q 分類規則を十分に特徴づけるために、追加検討する必要があり得る。実用的な理由により、分類関数 $[x]$ から結果として生じ得る値の範囲を、何らかの合理的な有限範囲に限定する必要があり得る。この限定は、クリッピングと呼ばれる。例えば、いくつかのこののようなシステムでは、分類規則は、より厳密に、

【 0 0 4 3 】

50

【数6】

$$A[x] = \text{sign}(x) * \min \left[g, \max \left(0, \left\lfloor \frac{|x|}{s} - \frac{z}{2} + 1 \right\rfloor \right) \right] \quad (4),$$

【0044】

と定義することができ、上式(4)で、gは [x] の絶対値に対する極限値である。

【0045】

各量子化索引に対する復元値を判定するのに、異なる復元規則を用いることができる。相互運用性の達成にのみ焦点を当てている標準及び製品仕様ではしばしば、必ずしも分類規則を特定せずに復元値を指定する。言い換えると、一部の仕様では、機能的マッピング $x \rightarrow [x]$ を定義せずに機能的マッピング $k \rightarrow [k]$ を定義している場合がある。これにより、標準 / 仕様に準拠するように構築されたデコーダが、情報を正しく復元することができる。対照的に、エンコーダにはしばしば、依然として標準 / 仕様に準拠したまま 10 で、望み通りに分類法を変更する自由が与えられる。

【0046】

量子化閾値を調整する多数のシステムが開発されている。多くの標準及び製品は、単純化するために、(例えば、典型的な単純分類規則のための)典型的な中央値復元規則に対応する復元値を指定する。しかし、分類のために、閾値は実際には、一定の入力値がより一般的な(したがって、より低いビットレートの)索引にマップされるように調整される場合があり、復元値は最適により近づく。 20

【0047】

多くのシステムでは、量子化の程度は、量子化ステップサイズによって測定される。比較的粗い量子化では、より広範囲の入力値に対応する、より大きい量子化ステップサイズを用いる。比較的細かい量子化では、比較的小さい量子化ステップサイズを用いる。しばしば、シグナリング及び復元のために、量子化ステップサイズは、最小の量子化ステップサイズの倍数としてパラメータ化される。

【0048】

C. 量子化アーチファクト

上述したように、損失性圧縮は、品質の低下を引き起こす傾向にある。例えば、わずかに異なる値の一連の10個のサンプルは、正確に同じ特定の近似値をもつ10個のサンプルとして、量子化を用いて近似することができる。この種類の量子化は、一連の10個のサンプルをエンコードするためのビットレートを削減し得るが、元の10個のサンプルの細部が損なわれる。 30

【0049】

いくつかのケースでは、量子化は、詳細部の単純損失より人工的に見え視覚的に混乱させる傾向にある、可視的なアーチファクトを生じる。例えば、平滑な、テクスチャ加工されていないコンテンツは、輪郭化アーチファクト(contouring artifacts)、すなわち異なる2つの量子化出力値の領域の間に現れるアーチファクトの影響を受けやすい。というのは、人間の視覚システムは、平滑コンテンツにおける微妙な変化(特に輝度の違い)に敏感だからである。上記の例を用いて、一連の10個のサンプルの輝度値が第1のサンプルから第10のサンプルまで徐々に、継続的に変わるケースを検討する。量子化は、最初の5個のサンプル値を1つの値として、最後の5個のサンプル値を別の値として近似することができる。この種類の量子化は、マスキング効果によりテクスチャ加工されたエリア内で可視的なアーチファクトを作成し得るが、平滑領域では、2組の5個のサンプルの間の復元された画像中に可視線またはステップを作成し得る。 40

【0050】

IV. VC-1における差分量子化

差分量子化において、エンコーダは、ピクチャの異なる部分に対して、量子化ステップサイズ(本明細書では、量子化パラメータまたは一部の実装形態ではQ Pとも呼ぶ)を変化させる。通常、マクロブロックレベルまたは他のサブピクチャレベルでのQ Pを変化さ 50

することを伴う。エンコーダは、Q Pをどのようにして変化させるかの決定を行い、そうした決定を、必要に応じてデコーダに信号通知する。

【0051】

例えば、VC - 1 エンコーダは、任意選択で、圧縮のための差分量子化を選ぶ。エンコーダは、Q Pが個々のピクチャ中のマクロブロックの間で変わり得るかどうかを示すために、ピクチャレベルの上のシンタクスレベルでビットストリーム要素(D QUANT)を送る。エンコーダは、ピクチャ Q Pを示すために、ピクチャレベルのビットストリーム要素、すなわち PQ INDEXを送る。D QUANT = 0 の場合、PQ INDEXによって指し示される Q Pは、ピクチャ中のすべてのマクロブロックに対して使われる。D QUANT = 1 または 2 の場合、同じピクチャ中の異なるマクロブロックを、異なる Q Pに対しても使うことができる。10

【0052】

VC - 1 エンコーダは、差分量子化に対して複数の手法を用いることができる。1つの手法では、ただ2つの異なる Q Pがピクチャに対して使われる。これは、二値差分量子化と呼ばれる。例えば、ピクチャエッジにあるマクロブロックに対して1つの Q Pが使われ、ピクチャのそれ以外の部分にあるマクロブロックに対して別の Q Pが使われる。これは、全体的な表示品質を保つのに詳細部が比較的重要でない、ピクチャエッジでビットを節約するのに有用であり得る。または、マクロブロックごとに信号通知される1ビット値が、使用可能な2つの Q P 値のどちらをマクロブロックに対して使うべきか示す。多値差分量子化と呼ばれる別の手法では、より多数の異なる Q Pを、ピクチャ中の個々のマクロブロックに対して使うことができる。20

【0053】

エンコーダは、D QUANTが非ゼロのとき、ピクチャレベルのビットストリーム要素、すなわち VOP DQUANT を送る。VOP DQUANT は、可能性としては、ピクチャのどの部分が、ピクチャ Q P以外の Q Pを使うことができるかを示す DQP ROLF ILE を含む他の要素からなる。異なる任意のマクロブロックがピクチャ Q P以外の Q Pを使うことができることを DQP ROLF ILE が示す場合、ビットストリーム要素 DQB ILE VEL が存在する。DQB ILE VEL = 1 の場合、各マクロブロックは、2つの Q Pの一方を使う(二値量子化)。DQB ILE VEL = 0 の場合、各マクロブロックは、どの Q Pを使ってもよい(多値量子化)。30

【0054】

ビットストリーム要素 MQDFFF は、二値量子化用のマクロブロックに対する1ビットセレクタを信号通知するために、マクロブロックレベルで送られる。多値量子化の場合、MQDFFF は、ピクチャ Q Pとマクロブロック Q Pまたはマクロブロックに対するエスケープ符号化された絶対 Q Pとの間の差分を示す。

【0055】

V. 他の標準及び製品

多数の国際標準では、ビデオデコーダの侧面及び圧縮された映像情報用の形式を指定している。直接的に、または暗黙により、こうした標準では、一定のエンコーダの詳細も指定しているが、他のエンコーダ詳細は特定されていない。一部の標準では、静止画像の圧縮/圧縮解除に対処し、他の標準では、音声の圧縮/圧縮解除に対処している。多数の企業では、音声、静止画像及び映像用のエンコーダ及びデコーダを生産している。他の様々な種類の信号(例えば、ハイパースペクトル画像、グラフィックス、テキスト、財務情報など)も、圧縮技術を用いて、普通に表され、格納され、または伝送される。40

【0056】

様々な映像標準では、異なるピクチャタイプに対して異なる量子化ステップサイズの使用を認め、レート及び品質管理のための量子化ステップサイズの変化を認めている。

【0057】

標準では通常、量子化法の設計を十分には指定していない。ほとんどが、エンコーダ分類規則 x [x] 及び / またはデコーダ復元規則 k [k]において、何らかの変形50

を認めている。D Z 比 $z = 2$ 以上の使用は、いくつかのエンコーディング設計において暗黙であった。例えば、一部の標準における、予測される領域に対する復元値の間隔をあけることは、 $z = 2$ の使用を含意している。標準にあるこうした例における復元値は、 $z = 2$ である D Z + U T Q 分類の使用のために、適切に間隔をあけられる。 $z = 1$ （または少なくとも $z < 2$ ）に基づく設計が、いくつかの標準では量子化に使われてきた。こうしたケースにおいて、復元値は、ゼロの辺りで、かつゼロから離れて等しく間隔をあけられる。

【 0 0 5 8 】

デジタル映像に対する映像圧縮の決定的な重要性が明らかになっているので、映像圧縮が十分に発展した分野であることは意外ではない。しかし、従来の映像圧縮技術の利益がどのようなものであっても、以下に述べる技術及びツールの利点は有していない。

10

【発明の開示】

【発明が解決しようとする課題】

【 0 0 5 9 】

本出願は、映像におけるテクスチャレベルに基づいて量子化を調整する技術及びツールを記載する。

【課題を解決するための手段】

【 0 0 6 0 】

例えば、ビデオエンコーダが、エンコーディング中にビットを割り振るときの平滑領域用の適応平滑化閾値及び平滑領域の時間分析を用いて、映像の知覚品質を向上させる。

20

【 0 0 6 1 】

本要約は、詳細な説明においてさらに後で説明する一そろいの概念を、簡素化された形で導入するためにある。本要約は、本特許請求対象の主要な特徴または本質的な特徴を識別することは意図しておらず、本特許請求対象の範囲を限定するのに使われることも意図していない。

【 0 0 6 2 】

一態様において、ビデオエンコーダは、ビデオピクチャシーケンス中の現在のビデオピクチャの現在の平滑領域（例えば、階調傾斜領域）を識別する。シーケンスは、先行ピクチャの表示が現在のピクチャの表示の前に起こり、将来ピクチャの表示が現在のピクチャの表示の後に起こる表示順序を有する。エンコーダは、現在の平滑領域に対して時間分析を実施する。時間分析は、先行及び／または将来ピクチャの少なくとも 1 つの中にある対応する領域が平滑であるかどうか判定することを含む。時間分析に少なくとも部分的に基づいて、エンコーダは、現在の平滑領域内で、量子化を調整する（例えば、輪郭化アーチファクトの混入を避けるように、量子化ステップサイズを低下させる）。時間分析は、調整可能な数の将来ピクチャに対して実施することができる。時間分析は、現在のピクチャに時間的により近い将来ピクチャの対応する領域における平滑性に、より大きな重みがあるとし得る。エンコーダは、テクスチャマップを分析して、現在の平滑領域を識別することができ、時間分析における平均輝度値を分析することができる。

30

【 0 0 6 3 】

別の態様では、ビデオエンコーダは、ビデオピクチャシーケンスに対する差分量子化間隔（例えば、一定または適応的に調整可能な間隔）を判定し、該間隔は間隔番号を含む。エンコーダは、シーケンス中の予測された差分量子化ピクチャに対する差分量子化を実施するときに間隔を用いる。間隔に従い、エンコーダは、予測された差分量子化ピクチャの 1 つの後の、少なくともその間隔番号の予測ピクチャに対する差分量子化をスキップする。例えば、差分量子化は、第 1 の予測ピクチャ用の 1 つまたは複数の差分量子化ステップサイズを選択すること、及び第 2 の予測ピクチャ用の 1 つまたは複数の差分量子化ステップサイズを変更することを含み、1 つまたは複数の差分量子化ステップサイズは、第 1 の予測ピクチャ用のピクチャ量子化ステップサイズとは異なり、第 2 の予測ピクチャは、第 1 の予測ピクチャからの間隔の外にある。

40

【 0 0 6 4 】

50

別の態様では、ビデオエンコーダは、現在のピクチャにおけるテクスチャを（例えば、テクスチャマップを分析することによって）分析し、分析されたテクスチャに少なくとも部分的に基づいて、現在のピクチャ用の平滑化閾値を設定する。エンコーダは、現在のピクチャのテクスチャデータを平滑化閾値と比較し、平滑化閾値に従い、現在のピクチャ中の少なくとも1つの平滑領域の発見に基づいて、現在のピクチャの少なくとも一部に対する差分量子化を調整する。エンコーダは、ロック階調値の階調値ヒストグラムにスライディングウィンドウを適用することによって、テクスチャを分析することができる。差分量子化を調整するために、エンコーダは、現在のピクチャ中の平坦なロックのパーセンテージを判定し、このパーセンテージを1つまたは複数のパーセント閾値と比較することができる。あるいは、エンコーダは、現在のピクチャ中のテクスチャ領域内の分離された平坦なロックを識別し、分離された平坦なロック向けの差分量子化をスキップすることができる。

【0065】

上記及び他の目的、特徴及び利点は、添付の図面を参照して進められる以下の詳細な説明から、より明らかになるであろう。

【発明を実施するための最良の形態】

【0066】

本出願は、映像の効率的な圧縮のための技術及びツールに関する。記載する様々な実施形態において、ビデオエンコーダは、ビデオをエンコードする技術、及びビットストリーム形式または異なる層もしくはレベルを含むシンタクスとともに使用するための、対応するシグナリング技術を組み込んでいる。記載する技術及びツールのいくつかは、インターレースまたはプログレッシブフレームに適用することができる。

【0067】

本明細書において記載する実装形態に対する様々な代替形態が可能である。例えば、フローチャート図を参照して記載する技術は、フローチャートに示す段階の順序を変更することによって、あるいは特定の段階を繰り返しまたは省くことなどによって、変えることができる。例えば、分析における初期段階（例えば、ピクチャに対するテクスチャ情報を取得することや、平滑領域を検出する際にテクスチャ分析を実施すること）は、以降の段階（例えば、ピクチャ用のエンコーディングの決定を行うことや、平滑領域を検出する際に時間分析を実施すること）が始まる前に完了してもよく、異なる段階に対する動作は、ロック単位、マクロロック単位、または他の領域単位で交互に行ってもよい。別の例として、いくつかの実装形態は具体的なマクロロック形式を参照して記載してあるが、他の形式を用いてもよい。

【0068】

様々な技術及びツールは、組み合わせても、それぞれ独立しても用いることができる。異なる実施形態が記載した技術及びツールの1つまたは複数を実装する。本明細書において記載するいくつかの技術及びツールは、ビデオエンコーダにおいても、ビデオエンコーディングに具体的に限定されない他の何らかのシステムにおいても用いることができる。

【0069】

I . コンピューティング環境

図9は、説明する実施形態のいくつかを実装することができる適切なコンピューティング環境900の一般化した例を示す。コンピューティング環境900は、使用または機能の範囲に対するどのような限定を示すことも意図していない。というのは、こうした技術及びツールは、多様な汎用または専用コンピューティング環境において実装することができるからである。

【0070】

図9を参照すると、コンピューティング環境900は、少なくとも1つの処理ユニット910及びメモリ920を含む。図9において、この最も基本的な構成930は、破線内に含まれている。処理ユニット910は、コンピュータ実行可能命令を実行し、実プロセッサでも仮想プロセッサでもよい。多重処理システムでは、処理能力を向上するために、

10

20

30

40

50

複数の処理ユニットがコンピュータ実行可能命令を実行する。メモリ920は、揮発性メモリ（例えば、レジスタ、キャッシュ、RAM）、不揮発性メモリ（例えば、ROM、EEPROM、フラッシュメモリなど）、またはこの2つの何らかの組合せでよい。メモリ920は、記載した技術及びツールの1つまたは複数を有するビデオエンコーダを実装するソフトウェア980を格納する。

【0071】

コンピューティング環境は、追加の特徴を有することもできる。例えば、コンピューティング環境900は、記憶装置940、1つまたは複数の入力デバイス950、1つまたは複数の出力デバイス960、及び1つまたは複数の通信接続970を含む。バス、コントローラ、またはネットワークなどの相互接続機構（図示せず）が、コンピューティング環境900のコンポーネントを相互接続する。一般に、オペレーティングシステムソフトウェア（図示せず）が、コンピューティング環境900において実行される他のソフトウェアのための動作環境を提供し、コンピューティング環境900のコンポーネントの活動を調整する。10

【0072】

記憶装置940は、取外し可能でも固定型でもよく、磁気ディスク、磁気テープまたはカセット、CD-ROM、DVD、あるいは情報を格納するのに使うことができるとともにコンピューティング環境900においてアクセスすることができる他のどの媒体も含む。記憶装置940は、ビデオエンコーダを実装するソフトウェア980用の命令を格納する。20

【0073】

入力デバイス（群）950は、キーボード、マウス、ペン、またはトラックボールなどのタッチ式入力デバイス、音声入力デバイス、走査デバイス、あるいはコンピューティング環境900への入力を実現する別のデバイスでよい。音声またはビデオエンコードのためにには、入力デバイス（群）950は、サウンドカード、ビデオカード、TVチューナカードでも、アナログまたはデジタルの形で、あるいはコンピューティング環境900に音声またはビデオサンプルを読み込ませるCD-ROMまたはCD-RWで音声またはビデオ入力を受け入れる類似のデバイスでもよい。出力デバイス（群）960は、ディスプレイ、プリンタ、スピーカ、CDライタ、またはコンピューティング環境900からの出力を実現する別のデバイスでよい。30

【0074】

通信接続（群）970は、別の計算エンティティへの、通信媒体を介した通信を可能にする。こうした通信媒体は、コンピュータ実行可能命令、音声またはビデオ入出力、あるいは他のデータなどの情報を、変調データ信号の形で伝える。変調データ信号とは、情報を信号にエンコードするような方式で設定または変更される信号特性の1つまたは複数を有する信号である。限定ではなく例として、通信媒体は、電気、光学、RF、赤外線、音響、または他のキャリアを有して実現される有線または無線技術を含む。

【0075】

こうした技術及びツールは、コンピュータ可読媒体という一般的な状況において説明することができる。コンピュータ可読媒体は、コンピューティング環境においてアクセスすることができる、利用可能などの媒体でもよい。限定ではなく例として、コンピューティング環境900の場合、コンピュータ可読媒体は、メモリ920、記憶装置940、通信媒体、及び上記のどの組合せも含む。40

【0076】

こうした技術及びツールは、コンピュータ実行可能命令、例えばコンピューティング環境において、ターゲットである実プロセッサまたは仮想プロセッサ上で実行される、プログラムモジュールに含まれるコンピュータ実行可能命令という一般的な状況で説明することができる。概して、プログラムモジュールは、特定のタスクを実施し、または特定の抽象データ型を実装する、ルーチン、プログラム、ライブラリ、オブジェクト、クラス、コンポーネント、データ構造などを含む。プログラムモジュールの機能は、様々な実施形態50

での要望に応じて、組み合わせることも、プログラムモジュールの間に振り分けることができる。プログラムモジュール用のコンピュータ実行可能命令は、ローカルまたは分散コンピューティング環境において実行することができる。

【0077】

提示として、詳細な説明では、コンピューティング環境におけるコンピュータ動作を記述するのに、「決定する」及び「分析する」のような用語を用いる。こうした用語は、コンピュータによって実施される動作を高度に抽象化したものであり、人間によって実施される行為と混同すべきではない。こうした用語に対応する実際のコンピュータ動作は、実装形態に応じて変わる。

【0078】

I I . 一般化したビデオエンコーダ

図10は、説明するいくつかの実施形態を実装することができる、一般化したビデオエンコーダ1000のブロック図である。エンコーダ1000は、現在のピクチャ1005を含む一連のビデオピクチャを受け取り、圧縮された映像情報1095を、記憶装置、バッファ、または通信接続への出力として生じる。出力ビットストリームの形式は、Windows(登録商標) Media VideoまたはVC-1形式、MPEG-x形式(例えば、MPEG-1、MPEG-2、もしくはMPEG-4)、H.26x形式(例えば、H.261、H.262、H.263、もしくはH.264)、あるいは他の形式でよい。

【0079】

エンコーダ1000は、ビデオピクチャを処理する。ピクチャという用語は概して、ソースの、符号化されたまたは復元された画像データを指す。プログレッシブ映像の場合、ピクチャは、プログレッシブ映像フレームである。インターレース映像の場合、ピクチャは、コンテキストに応じて、インターレース映像フレーム、フレームの上位フィールド、またはフレームの下位フィールドを指し得る。エンコーダ1000はブロックに基づき、フレームに対して4:2:0のマクロブロック形式を使用する。図11に示すように、マクロブロック1100は、4つの8×8輝度(またはluma)ブロック(Y1からY4)及び2つの8×8彩度(またはchroma)ブロック(U及びV)を含み、こうした彩度ブロックは、4つの輝度ブロックと同じ位置に配置されるが、従来の4:2:0マクロブロック形式に従って水平及び垂直方向の解像度が半分である。フィールドに対しては、同じまたは異なるマクロブロック組織及び形式を使うことができる。こうした8×8のブロックは、異なる段階、例えば、周波数変換段階及びエントロピー符号化段階でさらに下位分割することができる。エンコーダ1000は、8×8ブロック及び16×16マクロブロックとは異なるサイズまたは構成のサンプルの組に対して、操作を実施することができる。あるいは、エンコーダ1000は、オブジェクトに基づき、または異なるマクロブロックまたはブロック形式を使用する。

【0080】

図10に戻ると、エンコーダシステム1000は、予測ピクチャ及びイントラ符号化されたキーピクチャを圧縮する。提示として、図10は、エンコーダシステム1000を通過するキーピクチャ用の経路、及び予測ピクチャ用の経路を示す。エンコーダシステム1000のコンポーネントの多くは、キーピクチャ及び予測ピクチャ両方の圧縮に使われる。こうしたコンポーネントによって実施される正確な動作は、圧縮される情報のタイプに応じて変わり得る。

【0081】

予測ピクチャ(例えば、プログレッシブPフレームまたはBフレーム、インターレースPフィールドまたはBフィールド、あるいはインターレースPフレームまたはBフレーム)は、1つまたは複数の他のピクチャ(一般に、参照ピクチャまたはアンカと呼ばれる)からの予測(または差)によって表される。予測残差とは、予測されたものと元のピクチャの間の差である。対照的に、キーピクチャ(例えば、プログレッシブIフレーム、インターレースIフィールド、またはインターレースIフレーム)は、他のピクチャを参照せ

10

20

30

40

50

ずに圧縮される。

【0082】

現在のピクチャ 1005 が予測ピクチャである場合、動き推定装置 1010 は、1つまたは複数の参照ピクチャ、例えばピクチャストア 1020 にバッファリングされている、圧縮解除された先行ピクチャ 1025 を参照して、マクロブロックまたは現在のピクチャ 1005 の他のサンプルの組の動きを推定する。現在のピクチャ 1005 が双方向予測ピクチャである場合、動き推定装置 1010 は、現在のピクチャ 1005 における動きを、圧縮解除された最大 4 つの参照ピクチャ（例えば、インターレース B フィールド用）を参考して推定する。一般に、動き推定装置は、時間的に先行する 1 つまたは複数の参照ピクチャ、及び 1 つまたは複数の将来の参照ピクチャを参照して、B ピクチャにおける動きを推定するが、B ピクチャは、異なる時間方向から予測される必要はない。エンコーダシステム 1000 は、多様な参照ピクチャ用に別々のストア 1020 及び 1022 を使うことができる。
10

【0083】

動き推定装置 1010 は、1 サンプル、1 / 2 サンプル、1 / 4 サンプル、または他のインクリメントによって動きを推定することができ、動き推定の精度を、ピクチャ単位または他の単位で切り換えることができる。動き推定装置 1010（及び補償装置 1030）も、（例えば、双三次と双線形の間の）参照ピクチャサンプル補間タイプを、フレームごとに、または他の単位で切り換えることができる。動き推定の精度は、水平方向及び垂直方向において同じでも異なってもよい。動き推定装置 1010 は、差分動きベクトル情報などの動き情報 1015 をサイド情報として出力する。エンコーダ 1000 は、例えば、動きベクトル用に 1 つまたは複数の予測子を計算すること、動きベクトルと予測子との間の差分を計算すること、及び差分をエントロピー符号化することによって、動き情報 1015 をエンコードする。動きベクトルを再構成するために、動き補償装置 1030 は、予測子を差分動きベクトル情報と合成する。
20

【0084】

動き補償装置 1030 は、再構成されたピクチャ（群）1025 に再構成された動きベクトルを適用して、動きが補償された現在のピクチャ 1035 を形成する。ただし、予測が完璧であることはまれであり、動きが補償された現在のピクチャ 1035 と元の現在のピクチャ 1005 との間の差が予測残差 1045 である。以降のピクチャ復元の間、予測残差 1045 は、動きが補償された現在のピクチャ 1035 に加えられて、元の現在のピクチャ 1005 により近い復元ピクチャを取得する。ただし、損失性圧縮では、何らかの情報が依然として元の現在のピクチャ 1005 から失われる。あるいは、動き推定装置及び動き補償装置は、別のタイプの動き推定 / 補償を適用する。
30

【0085】

周波数変換装置 1060 は、空間ドメイン映像情報を周波数ドメイン（すなわち、スペクトル）データに変換する。ブロックベースのビデオピクチャに対して、周波数変換装置 1060 は、サンプルデータまたは予測残差データのブロックに DCT、DCT の変形、または他のブロック変換を適用し、周波数変換係数のブロックを生じる。代替的には、周波数変換装置 1060 は、フーリエ変換など別の従来の周波数変換を適用し、あるいはウェーブレットまたはサブバンド分析を用いる。周波数変換装置 1060 は、8 × 8、8 × 4、4 × 8、4 × 4 または他のサイズの周波数変換を適用することができる。
40

【0086】

量子化装置 1070 が次いで、スペクトルデータ係数のブロックを量子化する。量子化装置 1070 は、ピクチャ単位または他の単位（例えば、マクロブロック単位）によって変わるステップサイズで、スペクトルデータに均一なスカラー量子化を適用する。あるいは、量子化装置 1070 は、スペクトルデータ係数に別のタイプの量子化、例えば、均一でない量子化、ベクトル量子化、または非適応量子化を適用し、あるいは周波数変換を使わないエンコーダシステムにおいて、空間ドメインデータを直接量子化する。ある実装における量子化に関連する技術及びツールは以下で説明する。
50

【 0 0 8 7 】

適応量子化に加え、エンコーダ 1000 は、フレームドロップ、適応フィルタリング、またはレート制御のための他の技術を使うことができる。

【 0 0 8 8 】

エンコーダ 1000 は、特定のタイプの情報をもたない（例えば、マクロブロックに対する差分動きベクトルも残差情報ももたない）マクロブロックである、スキップされたマクロブロックに対して特殊なシグナリングを用いることができる。

【 0 0 8 9 】

再構成された現在のピクチャが、後続の動き推定／補償用に必要とされる場合、逆量子化装置 1076 が、量子化されたスペクトルデータ係数に対して逆量子化を実施する。逆周波数変換装置 1066 が次いで、周波数変換装置 1060 の動作の逆を実施し、（予測ピクチャに対する）再構成された予測残差、または再構成されたキーピクチャを生じる。現在のピクチャ 1005 がキーピクチャだった場合、再構成されたキーピクチャは、再構成された現在のピクチャ（図示せず）と見なされる。現在のピクチャ 1005 が予測ピクチャだった場合、再構成された予測残差が、動き補償された現在のピクチャ 1035 に追加されて、再構成された現在のピクチャを形成する。ピクチャストア 1020、1022 の一方または両方が、動きが補償された予測に使うために、再構成された現在のピクチャをバッファリングする。いくつかの実施形態では、エンコーダは、再構成されたフレームに非プロック化フィルタを適用して、ピクチャにおける不連続及び他のアーチファクトを適合的に平滑化する。

10

20

【 0 0 9 0 】

エントロピーコーダ 1080 が、量子化装置 1070 の出力、ならびに特定のサイド情報（例えば、動き情報 1015、量子化ステップサイズ（QP））を圧縮する。一般的なエントロピー符号化技術は、算術符号化、差分符号化、ハフマン符号化、ランレンジス符号化、LZ 符号化、ディクショナリ符号化及び上記の組合せを含む。エントロピーコーダ 1080 は一般に、異なる種類の情報（例えば、DC 係数、AC 係数、異なる種類のサイド情報）に対して異なる符号化技術を使用し、ある特定の符号化技術において、複数のコードテーブルの中から選択を行うことができる。

【 0 0 9 1 】

エントロピーコーダ 1080 は、多重化装置（MUX）1090 に、圧縮された映像情報 1095 を提供する。MUX 1090 は、バッファを含むことができ、バッファレベル指示子（indicator）は、コントローラにフィードバックすることができる。MUX 1090 の前または後で、圧縮された映像情報 1095 は、ネットワークを介した伝送用にチャネル符号化することができる。チャネル符号化は、圧縮された映像情報 1095 に、誤差検出及び修正データを適用することができる。

30

【 0 0 9 2 】

コントローラ（図示せず）は、動き推定装置 1010、周波数変換器 1060、量子化装置 1070、逆量子化装置 1076、エントロピーコーダ 1080 及びバッファ 1090 など、様々なモジュールから入力を受け取る。コントローラは、エンコーディング中に中間結果を評価し、例えば、歪みを推定し、他の速度歪み分析を実施する。コントローラは、動き推定装置 1010、周波数変換器 1060、量子化装置 1070 及びエントロピーコーダ 1080 などのモジュールと連携して、エンコーディング中に符号化パラメータを設定し変更する。エンコーダは、エンコーディング中に異なる符号化パラメータ選択肢を評価するとき、異なるパラメータ設定を評価するために、特定の段階（例えば、量子化及び逆量子化）を繰り返し実施することができる。エンコーダは、ある段階で、次の段階に進む前に、パラメータを設定することができる。あるいは、エンコーダは、異なる符号化パラメータを連帯して評価してもよい。評価されるべき符号化パラメータ決定のツリー及び対応するエンコーディングのタイミングは、実装に依存する。

40

【 0 0 9 3 】

エンコーダ 1000 内部のモジュールの間に示される関係は、エンコーダ内の情報の全

50

体的な流れを示す。他の関係は、単純化するために図示していない。具体的には、図10は通常は、ビデオシーケンス、ピクチャ、マクロブロック、ブロックなどに対して使われるエンコーダ設定、モード、テーブルなどを示すサイド情報を見せない。このようなサイド情報は、完成されると、出力ビットストリームに入れて、通常はサイド情報のエントロピーエンコーディングの後に送られる。

【0094】

ビデオエンコーダの特定の実施形態は通常、汎用化されたエンコーダ1000の変形または補完バージョンを用いる。実装及び所望される圧縮のタイプに応じて、エンコーダのモジュールは、追加し、省略し、多様なモジュールに分け、他のモジュールと結合し、かつ／または同様のモジュールと置き換えることができる。例えば、コントローラは、エンコーダの異なるモジュールに関連づけられた多様なコントローラモジュールに分けることができる。代替実施形態では、異なるモジュール及び／または他の構成のモジュールを有するエンコーダが、記載した技術の1つまたは複数を実施する。10

【0095】

I I I . 知覚モデルを用いた映像コンテンツの特徴抽出

映像コンテンツは、知覚モデルを用いて特徴づけることができる。このことは、エンコーダが異なる種類の映像コンテンツに対して適切なエンコーディング決定を行うのを助け得る。エンコーダは、エンコーディング前にピクチャを分析して、ピクチャの異なる部分（例えば、ブロック、マクロブロックなど）におけるコンテンツの特徴抽出（characterization）を行うことができる。例えば、青い空など、ビデオピクチャの比較的平滑な部分は、歪みを混入させるのには比較的受容可能でないと特徴づけることができる。というのは、一定の種類の品質劣化（例えば、量子化アーチファクト）が、平滑領域内で人間によってより簡単に知覚されるからである。対照的に、歪みは概して、テクスチャ領域内で比較的目立ちにくい（したがって、より受容可能）。20

【0096】

図12を参照すると、図10を参照して上述したようなビデオエンコーダが、ビデオピクチャの部分（例えば、マクロブロックのブロック、マクロブロック、または他の領域）を（例えば、平滑領域、エッジ領域、テクスチャ領域などとして）特徴づけ、特徴抽出に従って1つまたは複数のエンコーディング技術を適応させるプロセス1200において、適応エンコーディング技術を実装する。記載した技術の多くは、ブロック、マクロブロックまたは他の領域上など、ピクチャ内部の適応エンコーディングを提供する。こうした技術は、画像の異なる部分を分類しそれに従ってエンコードするのに、情報を用いる。より具体的には、ビデオエンコーダが、コンテンツをその知覚特性に基づいて分類するために、ピクチャの部分を特徴づける。30

【0097】

段階1210で、ビデオエンコーダは、ビデオピクチャの1つまたは複数の部分を特徴づける。例えば、エンコーダは、ビデオピクチャのブロックを、ブロック内部の強度分散に基づいて特徴づける。一実装形態では、エンコーダは、ブロックまたはブロックのダウンサンプリングされたバージョン中のピクセルに対して、ピクセルとその隣接ピクセルとの間の差分の和を計算する。この差分値の和は、ピクセルとその周辺ピクセルとの間の強度分散を測定する。例えば、周辺ピクセルは、ピクセルの上または下、左または右、あるいは対角線など、ピクセルに隣接する、またはほぼ隣接する他の2つ以上のピクセルを含む。ピクセルの強度とその周辺ピクセルの強度との間の差分は、輝度及び／または彩度データにおける差分に基づいて計算される。言い換えると、差分は、輝度サンプル及び／または彩度サンプルを用いて計算される。計算された平均差分値は、ピクセル（例えば、ピクセル階調）に割り当てられる。差分値は、ブロック中のピクセルに対して（例えば、ブロック階調）、またはそのサブサンプリングされた何らかのセットに対して、このようにして計算される。ブロック中のピクセルに割り当てられた差分値は、ブロック値を表現することができる、ブロックに対する特徴抽出または分類（例えば、平滑、エッジ、またはテクスチャ。テクスチャまたは非テクスチャ。平滑または非平滑、など）を判定するため4050

に評価される。一例では、ブロック中のピクセルに対するピクセル階調は、ブロックの階調に対する中央差分値（例えば、ブロック中央値）を判定するために評価される。したがって、ブロック内部のピクセルの間の強度差により、ブロック、マクロブロック、または他のビデオピクチャ領域に対する強度分散を測定する。

【0098】

ブロック中央値は、ブロック値を判定するのに必要とされない。強度分散またはブロック特徴抽出は、ブロック中のピクセルに割り当てられた差分値に対する平均値（例えば、ブロック平均）にも基づき得る。ブロック中央値または平均は、ブロックを分類するのに使うことができ、かつ／または、より微細化された制御機能への入力として使うことができる。特徴抽出または制御機能は、エンコーディングの1つまたは複数の側面を適応的に変化させる。10

【0099】

あるいは、ブロックを特徴づけるための強度分散を計算する代わりに、エンコーダは、別の基準値を用いる。

【0100】

段階1220で、エンコーダは、特徴抽出に基づいて、ビデオピクチャを適応的にエンコードする。一実装形態では、輪郭化アーチファクトの除去または削減のためのエンコーディング技術が、ブロックの特徴抽出に基づいて実施される。例えば、階調傾斜検出、DCシフト検出、AC係数保存及び適応差分量子化が、一定の平滑領域に対して実施され、テクスチャ加工された領域が、ビットレートを保護するためにより強く量子化される。20

【0101】

図12は、ピクチャの多様な部分に対する適応エンコーディング段階1220に先行する特徴抽出段階1210を示すが、こうした段階は、ピクチャにおいてブロック単位で繰り返しきこっても、他の何らかの単位で並べられてもよい。

【0102】

段階1230で、エンコーダは、適応エンコードされたビットストリームを信号通知する。ブロック特徴抽出に基づいてエンコードするためにエンコーダによって差分量子化が使われる場合、例えば、ビデオエンコーダは、ビデオデコーダに差分量子化を信号通知するシグナリング方式を用いて情報を圧縮されたビットストリームの中にエンコードする。

【0103】

段階1240で、対応するビデオデコーダが、ビデオピクチャに関するエンコードされたデータを含む、適応エンコードされたビットストリームを読む。例えば、ビデオデコーダは、信号通知された差分量子化情報を読む。段階1250で、デコーダは、圧縮されたビットストリームを復号化し、例えば、信号通知された差分量子化情報に従ってブロックを逆量子化する。30

【0104】

A. ブロックに基づいた例示的特徴抽出

図13は、輝度及び／または彩度データを用いてブロックを特徴づける、ブロックに基づいた操作を示す図である。輝度ブロック「Y」(1302)は、4:2:0のマクロブロック形式の、8×8ブロックのマクロブロックである。そうすることが必要なわけではないが、この例では、ピクセルブロック向けの対応する彩度ブロック1304、1306も、階調ブロック1308を計算する際に使われる。そうすることが必要なわけではないが、この例に示すように、輝度ブロック1302は、因子2で水平に、かつ垂直に（例えば、サンプルのペアの単純平均化によって）ダウンサンプリングされて(1312)、4×4次元の彩度ブロックに合致する輝度ブロック1310を作成する。40

【0105】

ダウンサンプリングされた輝度ブロック1310に示すように、ピクセル1314に対する輝度サンプルの強度値は、ダウンサンプリングされた輝度ブロック1310中のピクセル1314に近い4ピクセルに対するサンプルと比較され、ピクセル1314に対するサンプルとその周辺の垂直及び水平ピクセルに対するサンプルとの間の差分の平均和が計50

算される。この例では、ピクセル 1314 は、位置 $Y' (r, c)$ に位置する。その周辺ピクセルと比較した、このピクセル 1314 向けの輝度強度値に対する差分の平均和は、以下のようになる。

$$L_1(r, c) = [| Y'(r, c) - Y'(r, c-1) | + | Y'(r, c) - Y'(r-1, c) | + | Y'(r, c) - Y'(r, c+1) | + | Y'(r, c) - Y'(r+1, c) |] / 4 \quad (5)$$

【0106】

上に示すように、 $Y'(r, c)$ は、ダウンサンプリングされたブロック Y' 中の行 r 及び列 c にあるピクセル 1314 の輝度成分である。 $L_1(r, c)$ は、ピクセル 1314 が、ブロック Y' 中のその隣のピクセルとは輝度強度がどの程度異なるかを指示する。この輝度強度差の測定値は、ピクセル階調の例である。

【0107】

任意選択で、彩度データ 1304、1306 は、強度差を判定するために、輝度データの代わりに独自に検討することもでき、輝度データと一緒に検討することもできる。ピクセル 1314 の輝度強度値及び彩度強度値に対する差の平均和は、以下の式に示すように、周辺ピクセルに対するサンプルの強度値における差の平均として表すことができる。

$$G_1(r, c) = \{ [| Y'(r, c) - Y'(r, c-1) | + | Y'(r, c) - Y'(r-1, c) | + | Y'(r, c) - Y'(r, c+1) | + | Y'(r, c) - Y'(r+1, c) |] + [| U(r, c) - U(r, c-1) | + | U(r, c) - U(r-1, c) | + | U(r, c) - U(r, c+1) | + | U(r, c) - U(r+1, c) |] + [| V(r, c) - V(r, c-1) | + | V(r, c) - V(r-1, c) | + | V(r, c) - V(r, c+1) | + | V(r, c) - V(r+1, c) |] \} / 12 \quad (6)$$

【0108】

$G_1(r, c)$ は、ダウンサンプリングされたブロック中の (r, c) に位置するピクセルに対するピクセル階調の例であり、ピクセル階調は、ピクセル 1314 が、その周辺にある隣のピクセルとは輝度及び彩度強度がどの程度異なるかを指示する。この例では、ピクセル階調値 $G_1(r, c)$ は、直接的に垂直または水平なピクセルに基づくが、近隣にある他のピクセルは検討しない。他のピクセルデータは、他の変形におけるピクセル階調の作成において検討できることも企図される。例えば、対角線ピクセルは、与えられた配置の一部、またはその代わりと見なすことができよう。あるいは、より長い範囲（例えば、2 または 3 ピクセル）に渡る強度差を検討することもできよう。

【0109】

$G_1(r, c)$ は、単一のピクセルが、その隣とは輝度及び彩度強度がどの程度異なるかを指示する。ブロック全体に対する強度分散を特徴づけるために、ブロック内部の複数のまたはすべてのピクセルに対して同じ分析が実施される。このような一例において、複数のピクセル階調のブロック 1308 が作成され、そこからブロック階調が導出される。上述したように、ピクセル階調またはブロック階調の計算は、輝度比較のみ、彩度比較のみ、または輝度及び彩度比較の両方を合わせて含み得る。

【0110】

そうすることが望ましい場合、 $G_1(r, c)$ を見つけるための上記の式は、欠けていいるブロック境界値を明らかにするように変形させることができる。例えば、ブロック外のサンプルは、境界値を明らかにするように式 $G_1(r, c)$ を適応させる場合、ブロック内部の他の隣接サンプルと同じになるように外挿することも、仮定することもできる。あるいは、例えば、式の分母を通分することができ、例えば特定の方向の周辺サンプルがブロック外であるときは、そうした周辺サンプルは、比較において無視してよい。図示したように、複数のピクセル階調のブロック 1308 は、そのブロック中のすべてのピクセルに対するピクセル階調データを提供し得る。あるいは、複数のピクセル階調のブロック 1308 は、そのブロック中のすべてのピクセルに対するピクセル階調データを含み得るわけではない。

10

20

30

40

50

【0111】

図14は、図13のブロック1308中の複数のピクセル階調のヒストグラムである。より具体的には、ヒストグラム1400は、ブロックがどのように特徴づけられ、または値をつけられるかを視覚化する。この例では、30より下の8個のピクセル階調値、及び30より上の8個のピクセル階調値がある。したがって、このブロック階調に対する中央値は30である。（偶数個の候補の場合、中央値は、2つの中間候補値の平均として、または2つの中間候補値の一方もしくは他方として計算することができる）。中央値は、ブロックを平滑、テクスチャまたはエッジとして特徴づけるのに使うことができる。当然ながら、ピクセル階調または複数のピクセル階調のブロックが取得されると、ブロックを特徴づけるのに、他の基準を用いることもできる。例えば、ブロックは、ピクセル階調値の平均によって特徴づけることができる。ブロック値は、割り当てられると、特徴抽出方式において（例えば、平滑または非平滑。平滑、テクスチャ、エッジなど）、またはより微細化された制御機能において使うことができる。ブロック値は、適応エンコーディング戦略においてブロックがどのように扱われるか判定するのに使うことができる。10

【0112】

ブロック値は、複数のピクセル階調を並べ替え、並べられた値から中央階調値を選択することによって選択することができる。例えば、ブロック内部の1組のピクセル階調、例えば{10、14、28、36、38}は、その組の中の中央値ピクセル階調、すなわち28に等しい、割り当てられたブロック値を有する。別の例では、ブロック値は、数値例に先行するその組の中の平均階調、すなわち25.2に基づいて判定される。当然ながら、こうした組は、完全なブロック階調またはそのサブセットから取得することができる。20

【0113】**C．特徴抽出情報の例示的な使用**

図15は、図13、14の例に続く、例示的ブロック特徴抽出フレームワークのグラフである。図示したように、0から最大30までを含む範囲のブロック値をもつブロックが、平滑ブロックとして特徴づけられることになる。30より大きいが60以下の範囲のブロック値をもつブロックは、テクスチャブロックとして特徴づけられることになり、60より大きいブロック値をもつブロックは、エッジブロックとして特徴づけられることになる。

【0114】

あるいは、エンコーダは、別の特徴抽出フレームワーク、例えば、ブロックまたはビデオピクチャの他の部分に対する他の及び／または追加の特徴抽出を含むものを用いる。異なる階調及び基準に対して、フレームワークは、スケール及び／または次元数が変わり得る。

【0115】

エンコーダは、エンコーディング決定を行うとき、ブロックまたはビデオピクチャの他の部分の特徴抽出を用いることができる。表2は、例示的適応符号化方式の特徴を、図15を参照して記載したブロック特徴抽出に関係づける。図示したように、異なるように特徴づけられたブロックは、1つまたは複数の適応特徴によって異なるように扱われる。

【0116】**【表2】**

特徴抽出	DCシフト検出	階調傾斜検出	量子化
平滑	はい	はい	より低いQP
エッジ	いいえ	いいえ	より高いQP
テクスチャ	いいえ	いいえ	より高いQP

表2：適応エンコーディング特徴

【0117】

表 2 に示す様々な適応特徴については、本文書全体を通して論じ、後でさらに論じる。あるいは、エンコーダは、ブロック特徴抽出への、適応特徴決定の別のマッピングを用いる。さらに、本明細書において記載するいくつかの特徴は、映像コンテンツの特徴抽出を検討する必要はない。

【 0 1 1 8 】

I V . テクスチャレベルに基づく差分量子化

差分量子化において、エンコーダは、ピクチャの異なる部分に対して、量子化ステップサイズ（本明細書では、量子化パラメータまたは一部の実装形態では Q P とも呼ぶ）を変化させる。通常、マクロブロックまたは他のサブピクチャレベルでの Q P を変化させることを伴う。エンコーダは、Q P をどのようにして変化させるかの決定を行い、こうした決定を、必要に応じてデコーダに信号通知することができる。10

【 0 1 1 9 】

以前のエンコーダでは、二値差分量子化（2つの Q P の間で変化させる）及び多値差分量子化（3つ以上の Q P の間で変化させる）を用いていた。例えば、ある二値差分量子化手法では、ピクチャエッジにあるマクロブロックに対して1つの Q P が使われ、ピクチャのそれ以外の部分にあるマクロブロックに対して別の Q P が使われる。これは、全体的な表示品質を保つのに詳細部が比較的重要でない、ピクチャエッジでビットを節約するのに有用であり得る。多値差分量子化手法では、より多数の異なる Q P を、ピクチャ中の個々のマクロブロックに対して使うことができる。例えば、エンコーダは、マクロブロックに対して Q P を選び、現在のピクチャに対する Q P とマクロブロックに対する Q P との間の差分を信号通知することができる。20

【 0 1 2 0 】

量子化アーチファクトに対する知覚感度は、空間及び時間ドメイン両方におけるビデオのテクスチャレベルと大いに関連する。高テクスチャレベルはしばしば、品質劣化及び量子化アーチファクトを隠し得るマスキング効果を結果として生じる。ただし、テクスチャレベルがより低い領域（例えば、平滑領域）では、劣化及び量子化アーチファクトは、より可視的である。以前のエンコーダでは、ビデオピクチャの一部の部分（例えば、ピクチャエッジ）に対して量子化調整を行っていたが、本明細書において記載するより包括的なコンテンツに基づく差分量子化戦略により、多くのシナリオにおいて速度歪み性能が向上する。30

【 0 1 2 1 】

したがって、記載する技術及びツールの多くでは、テクスチャに基づく差分量子化（本明細書では、テクスチャ D Q と呼ばれる）を用いて、様々なテクスチャレベルに基づいてビットを割り振り、より優れた知覚品質を達成する。テクスチャ D Q から、テクスチャ情報に基づいて、かついくつかのケースでは、時間分析情報など他の情報に基づいて映像を符号化するための異なる Q P が選ばれる。エンコーダは、テクスチャ情報（及び可能であれば他の情報）を分析し、ピクチャ中の 8 × 8 ブロックやまたはマクロブロックなど、適切な領域（テクスチャ D Q 領域）にテクスチャ D Q を適用する。記載する技術及びツールの多くでは、潜在テクスチャ D Q 領域として平滑領域に焦点を当てる。平滑領域は、平坦な領域（不变またはほぼ不变の色のエリア）及び階調傾斜領域（領域に渡る、不变またはほぼ不变の速度で変化する色のエリア）を含む。平滑領域は、ノイズ、フィルム粒子、または他の色変化の小さいエリアに割り込まれても、平滑と見なすことができる。40

【 0 1 2 2 】

図 16 は、テクスチャ情報に基づいて差分量子化を適用する、汎用化された技術 1600 を示すフローチャートである。図 10 のエンコーダ 1000 などのエンコーダまたは他のツールが、技術 1600 を実施する。

【 0 1 2 3 】

段階 1610 で、エンコーダは、現在のピクチャのテクスチャ情報（例えば、異なる領域が平滑、エッジ、またはテクスチャ領域のいずれであるかを示す特徴抽出やブロック値）を取得する。段階 1620 で、エンコーダは、現在のピクチャ中のテクスチャ D Q 領域50

(例えば、輪郭化アーチファクトが存在し得る平滑領域)またはテクスチャDQ領域群を見つける。段階1630で、エンコーダは、テクスチャDQ領域(群)にテクスチャDQを適用し、ピクチャをエンコードする。例えば、平滑領域が、高テクスチャ領域より小さいQPで符号化される。エンコードするべきより多くのピクチャがある場合、エンコーダは、段階1640で次のピクチャ入手し、必要に応じて次のピクチャにテクスチャDQを選択的に適用する。エンコーダは、ビデオピクチャに対するエンコードされたデータを、例えば記憶装置、通信接続、またはバッファに出力する。

【0124】

領域が平滑領域として扱われるべきかどうか判定するのに、異なるテクスチャDQ領域検出技術を用いることができる。例えば、エンコーダは、ある特定の領域がテクスチャDQ領域と見なされるべきかどうか判定するのに、異なるテクスチャ基準及び/または異なるテクスチャ閾値を用いることができる(かつ、閾値を適応的に調整することができる)。より優れた知覚ビデオ品質のためのビットを割り振るのに、適応量子化値マッピングを用いることができる。差分量子化決定は、時間分析にも基づき得る(すなわち、将来ピクチャを考察して、長期にわたる領域の特性に基づく決定を行う)。

10

【0125】

差分量子化決定は、イントラピクチャ及び予測ピクチャ両方に対して行うことができる。予測ピクチャの場合、差分量子化ピクチャの間のPピクチャ及びBピクチャ差分量子化間隔を制御することができる。さらに、優的な(dominant)高テクスチャエリアが存在する場合のピクチャのテクスチャを観察することによって、平滑領域テクスチャ閾値を、比較的小さいQPで(優的な高テクスチャエリアに比較して)比較的平滑な領域を符号化するように緩和することができる。

20

【0126】

上のセクションIIIで図12~15を参照して記載したのと同様の技術は、現在のピクチャに対するテクスチャマップを生成するのに用いることができる。例えば、エンコーダは、セクションIIIで記載したように、ピクチャに対するテクスチャレベル用の階調を、ピクチャに対するY、U、Vチャネル中の第1の導出物(差)として算出する。マクロブロック形式が4:2:0の場合、算出プロセスを高速化するために、エンコーダは、Yチャネルを2:1の因子で水平及び垂直にダウンサンプリングすることができる。エンコーダは、各ピクセルに対するY、U、Vの階調を水平及び垂直方向両方に合計する。最大解像度の8×8ブロックの場合、エンコーダは、ダウンサンプリングされたピクチャ中の対応する4×4ブロックにおける階調の和の平均を計算して、ブロック階調値として使う。階調の平均の計算は、計算上、セクションIIIで記載した中央値の計算より複雑でない。

30

【0127】

あるいは、エンコーダは、ピクチャのテクスチャ情報を、他の何らかのやり方で取得する。例えば、エンコーダは、階調を算出する異なる階調方向を選び、輝度チャネルに対してのみ階調を算出する、などを行う。テクスチャ情報は、どのようにして取得されまたは算出されようとも、次いで、テクスチャDQ決定を行うのに使うことができる。

【0128】

40

テクスチャマップは、ピクチャの相異なる部分のテクスチャレベルを示す。例えば、テクスチャマップは、ピクチャ中の平滑領域(例えば、ブロック、マクロブロック、エッジ、または他のエリア)及びテクスチャ加工された領域を識別するのに使うことができる。記載した差分量子化技術は、テクスチャマップ中の情報に基づいて、ピクチャの適切な部分に対して実施することができる。あるいは、エンコーダは、最初にテクスチャマップを作成することなく、テクスチャ情報を用いる。

【0129】

A. 時間分析 (temporal analysis)

現在のビデオピクチャからのテクスチャ情報に加えて、正確な差分量子化決定を行うのに、時間分析を用いることができる。時間分析を用いる理由の1つは、平滑領域がいくつ

50

かのピクチャに渡って平滑なままである場合、特に他のピクチャが動き補償において平滑領域を参照するとき、平滑領域に対して比較的小さいQPを使う際の影響力が、より大きくなるからである。逆に、平滑ブロックが将来ピクチャ中の高テクスチャまたはエッジブロックで置き換えられた場合、比較的小さいQPを使うことの利益が1つ失われる。したがって、エンコーダは、現在のピクチャ中に平滑領域を見つけた後で将来ピクチャを考察し、領域の平滑性が将来ピクチャ中でどのように変わるかに基づいて差分量子化決定を行う。エンコーダは、先行ピクチャ、例えば、表示順序において現在のビデオピクチャに先行するBピクチャを考察するが、動き補償において現在のビデオピクチャを参照することもできる。

【0130】

10

図17は、時間分析を用いてテクスチャDQ決定を行う例示的技術1700を示す。図10のエンコーダ1000などのエンコーダまたは他のツールが、技術1700を実施する。

【0131】

段階1710で、エンコーダは、ビデオシーケンスにある現在のピクチャ中の現在のブロックに対してテクスチャ分析を実施する。例えば、エンコーダは、ブロックの階調情報を考察する。エンコーダは、階調情報を、ブロックに対する階調閾値と比較し、ブロックを平滑または非平滑（例えば、テクスチャ、エッジ）と分類することができ、この場合、階調閾値は、固定され、または現在のピクチャまたはビデオシーケンスの他の部分に対して動的に設定される。あるいは、エンコーダは、現在のピクチャ中の他の何らかの部分向けにテクスチャ分析を実施する。

20

【0132】

段階1720で、エンコーダは時間分析を実施する。エンコーダは、自動的に、または現在のブロックが平滑ブロックと分類された場合のみ、時間分析を実施することができる。例えば、エンコーダは、現在のピクチャ中の平滑ブロックが、将来ピクチャ中で平滑なままであるかどうか判定する。平滑な場合、現在のピクチャ中の平滑領域は、後で比較的小さいQPで符号化される。あるいは、エンコーダは、現在のピクチャ中の平滑ブロックが先行ピクチャ中、または先行及び将来のピクチャ両方の中でも平滑だったかどうか判定する。

【0133】

30

エンコーダが分析する先行及び／または将来のピクチャの数は、実装に応じて変わり得る。平滑領域が将来ピクチャ（例えば、次のピクチャまたは他の何らかの時間的に近いピクチャ）中で、テクスチャ加工された領域で置き換えられた場合、現在のピクチャ中の平滑領域は、より大きいQPで符号化することができよう。というのは、比較的小さいQPを使うことの利点は、おそらく持続性がないからである。一実装形態では、差分量子化決定を行う際、時間的により近いピクチャが、より離れたピクチャより重く重みづけされる。エンコーダが考察する先行及び／または将来ピクチャの重みづけ及び数は、実装に応じて変わり得る。

【0134】

40

算出を簡単にするために、エンコーダは、将来ピクチャ中の現在のブロックと、対応するブロックを比較するための単一の値を見つけることができる。例えば、輝度値は平滑ブロック中でかなり一貫性があるので、将来ピクチャ中の対応するブロックの類似を測定するために、ブロックに対する輝度値の平均値が算出される。以下の例式(7)で、将来ピクチャ中の対応するブロックの将来の平滑性の「強度」S(t)は、現在のブロック中の平均輝度値と将来ピクチャ中の対応するブロックとの間の重みづけされた絶対差、2つの将来ピクチャ中の対応するブロックの平均輝度値などの和によって算出される。

【0135】

【数7】

$$S(t) = C(n) * \sum_{i=1}^n (n-i+1) * |M(t+i) - M(t+i-1)| \quad (7)$$

【0136】

上式(7)で、nは時間的な「先見」ピクチャの総数であり、C(n)は、2/(n*(n+1))となるように定義される規格化因子であり、M(t)は、時間tでのピクチャ中のブロック(または対応するブロック)に対する輝度値の平均値である。エンコーダは、将来の平滑性の代わりに、またはそれに加えて、過去の平滑性を測定することもできる。あるいは、エンコーダは、平滑性の時間分析において別の重みづけシステム及び/または平滑性基準値を用いる。10

【0137】

図17を再度参照すると、段階1730で、エンコーダは、テクスチャ分析及び時間分析の結果を用いて、ブロックをテクスチャDQブロックと分類するかどうか判定する。例えば、エンコーダは、平滑ブロック(ただし他のブロックではなく)に対する平滑性強度S(t)を計算し、平滑性強度S(t)を時間平滑化閾値と比較する。時間平滑化閾値は、固定することもでき、動的に設定することもできる。

【0138】

図17において、エンコーダが、現在のブロックが平滑ブロックであること、及び先行及び/または将来ピクチャ中の対応するブロックも平滑であることを発見した場合、エンコーダは、段階1740で、テクスチャDQブロックのカウントに現在のブロックを加える。エンコーダは、テクスチャDQブロックのカウントを使って、ピクチャに対してテクスチャDQを実施するかどうか判定することができる。あるいは、エンコーダは、時間分析を他の何らかのやり方で用いて、テクスチャDQ決定を行う。20

【0139】

分析するべきブロックがそれ以上ある場合、エンコーダは、段階1750で次のブロック入手し、図17に示すプロセスを繰り返す。これは、エンコーダが現在のビデオピクチャのブロックを評価するまで続く。その時点で、エンコーダは、エンコーディング決定において平滑ブロックのカウントまたは時間分析の他の結果を用いる。

【0140】

図17は、ブロック単位での時間分析を実施するエンコーダを示すが、代替的には、エンコーダは、時間分析をマクロブロック単位で、または他の何らかの領域単位で実施する。30

【0141】

B. テクスチャDQ閾値及び分離平滑ブロックフィルタリング

エンコーダが時間分析を用いても用いなくても、エンコーダは、テクスチャDQをいつ適用するか決定する際、他のいくつかの機構を用いることができる。エンコーダは、DQを実施するかどうかについての決定を行い、行う場合、テクスチャDQ領域に対してQPがどの程度細ければよいか決定するのに、1つまたは複数の普及閾値(例えば、ピクチャ中の平滑ブロックのパーセンテージ)を使うことができる。例えば、ピクチャ中の平滑ブロックの数またはパーセンテージが閾値より上の場合、エンコーダは、平滑コンテンツを小さいQPでエンコードするあまりに多くのビットを費やすことを避けるために、比較的粗いステップサイズを選ぶことができる。エンコーダは、平滑ブロックの数またはパーセンテージが、ピクチャ中で少しでもテクスチャDQを使うのに十分かどうか判定するためのより低い閾値も有し得る。40

【0142】

別のビットレート削減法は、特定の平滑ブロックが、大部分がテクスチャ加工された領域内にあるときは、そうした平滑ブロックをテクスチャブロックとして扱うことである。これは分離平滑ブロックフィルタリングと呼ぶことができる(平滑ブロックは、このようにしてフィルタリングされるように完全に「分離される」必要があるわけではないが)。50

例えば、テクスチャ加工されたブロックで囲まれる平滑ブロックは、テクスチャ加工されたブロックより小さいQ Pで符号化される必要はない。というのは、平滑ブロック中の量子化アーチファクトは、テクスチャ加工された周辺コンテンツによっておそらくマスクされるからである。その結果、エンコーダは、分離平滑ブロックに対してテクスチャD Qを実施しないことを選ぶことができる。エンコーダは、ピクチャ中の平滑ブロックの数またはパーセンテージを算出するときに、分離平滑ブロックを軽視してもよい。

【0143】

図18は、閾値及び分離平滑ブロックフィルタリングを用いてテクスチャD Q決定を行う例示的技術1800を示す。図10のエンコーダ1000などのエンコーダまたは他のツールが、技術1800を実施する。

10

【0144】

段階1810で、エンコーダは、現在のピクチャ中に平滑ブロックを見つける。あるいは、エンコーダは、図17を参照して記載したテクスチャ分析及び時間分析を実施する。あるいは、エンコーダは、他の何らかのやり方で、現在のピクチャ中に平滑ブロックを見つける。

【0145】

段階1820で、エンコーダは、分離平滑ブロックフィルタリングを実施する。あるいは、エンコーダは、現在のピクチャ中で非平滑ブロックによって囲まれる単平滑ブロックを削除する。エンコーダは、異なる多くの決定モデルを使って、分離平滑ブロックフィルタリングを実施することができる。例えば、エンコーダは、平滑ブロックを、そのすべての隣接ブロックがテクスチャ加工されたブロックであるときのみ、テクスチャ加工されたブロックとして扱うことができる。あるいは、エンコーダは、平滑ブロックを、その一定数の隣接ブロックがテクスチャ加工されている場合に、テクスチャ加工されたブロックとして扱うことができる。あるいは、エンコーダは、分離平滑ブロックを、より大きい（例えば、2もしくは3）グループごとにまとめて、かつ／またはブロック（群）が分離されるかどうかについての他の何らかのテストを用いて削除する。

20

【0146】

段階1830で、エンコーダは、ピクチャ中の平滑ブロックのパーセンテージを、低閾値（例えば、ピクチャ中の総ブロックの1～2%）と突き合わせる。平滑ブロックのパーセンテージが低閾値を下回る場合、エンコーダは、このピクチャに対してテクスチャD Qが使われないと判定する（段階1840）。平滑ブロックのパーセンテージが低閾値を上回る場合、エンコーダは、段階1850で、パーセンテージを、高閾値と突き合わせる。この高い方の閾値が、平滑ブロックに対するQ Pを選ぶのに使われる。パーセンテージが高閾値より高い場合、エンコーダは、テクスチャD Qを実施するが、ビットレートを削減するために、平滑ブロックに対して、比較的粗いQ Pを選ぶ（段階1860）。パーセンテージが高閾値より低い場合、エンコーダは、平滑ブロックに対して比較的細かいQ Pを選ぶ（段階1870）。分析するべきピクチャがそれ以上ある場合（段階1880）、エンコーダは、他のピクチャに対してプロセスを繰り返すことができる。閾値の数及び閾値パーセンテージ値は、実装に応じて変わり得る。

30

【0147】

代替的には、エンコーダは、テクスチャD Q閾値を用いずに分離平滑ブロックフィルタリングを実施し、または分離平滑ブロックフィルタリングなしでテクスチャD Q閾値を使う。あるいは、エンコーダは、分離平滑ブロックフィルタリングなしで、またはD Q閾値を使わずにテクスチャD Qを実施する。

40

【0148】

C. 適応テクスチャレベル閾値

エンコーダは、固定されたテクスチャレベルまたは平滑化閾値を使って、所与のブロックがテクスチャD Qブロック（例えば、平滑ブロック）と見なされるべきかどうか判定することができる。D Qシグナリングのビットレートコスト（例えば、「すべてのマクロブロック」二値D Qシグナリングシナリオにおいて、マクロブロックごとに1ビット）及び

50

ピクチャのいくつかの部分を比較的小さいQ Pで量子化するビットレートコストを考慮に入れると、閾値は、テクスチャD Qのコストに対するチェックとして作用する。例えば、エンコーダは、(図13、14を参照して記載した技術または他の何らかの技術を用いて) ブロックに対するブロック値を取得し、ブロック値を固定されたテクスチャレベル / 平滑化閾値(例えば、図15を参照して記載した閾値)と比較する。

【0149】

エンコーダは、テクスチャレベル / 平滑化閾値を適応的に変更することもできる。例えば、平滑ブロックの知覚性はたくさんの高テクスチャコンテンツを有するピクチャ中で変わり得るので、ブロックを平滑ブロックと分類するテクスチャレベル閾値は、中間テクスチャまたは高テクスチャピクチャ中で緩和してよい。これは、適応テクスチャレベル閾値の例である。エンコーダは、ある閾値範囲内の、いくつかの異なる閾値を選択させる。一実装形態では、平滑ブロックに対する適応テクスチャレベル閾値は、ブロック値14とブロック値30の間で変化させることができる。異なるテクスチャレベル閾値に対して、異なる差分量子化マッピングを用いることができる。適応テクスチャレベル閾値は、より平滑な領域内で品質を向上させるために、より高いテクスチャフレーム中より平滑な領域にビットを割り振るのに有用であり得る。

10

【0150】

図19は、高テクスチャピクチャに対するテクスチャレベル閾値を選択的に調整する技術1900を示す。図10のエンコーダ1000などのエンコーダまたは他のツールが、技術1900を実施する。エンコーダは、ピクチャ中の優的な高テクスチャコンテンツの存在を検出することによって、テクスチャレベル閾値を調整するかどうか判定する。一実装形態では、高テクスチャコンテンツの検出は、テクスチャヒストグラムにおけるサイズが10のスライディングウィンドウ中のテクスチャ「エネルギー」の算出によって実装される。

20

【0151】

図19を参照すると、エンコーダは、適応テクスチャレベル閾値技術1900における段階1910で、ピクチャのテクスチャ情報(例えば、テクスチャレベルヒストグラム)を取得する。例えば、エンコーダは、上述したようにテクスチャマップを取得し、情報からテクスチャレベルヒストグラムを作成する。

【0152】

30

段階1920で、エンコーダは、ピクチャが高テクスチャピクチャであるかどうかチェックする。ピクチャが高テクスチャピクチャの場合、エンコーダは、段階1930で、ピクチャに対するテクスチャレベル閾値を調整する。ピクチャが高テクスチャピクチャでない場合、エンコーダは、テクスチャレベル閾値を調整せずにピクチャを処理する(段階1940)。エンコーダは次いで、他のピクチャに対するテクスチャレベル閾値を分析し選ぶことができる(段階1950)。あるいは、エンコーダは、ピクチャ中の異なるレベルの高テクスチャコンテンツに対して、スライド式の異なるテクスチャレベル閾値を適用する。

【0153】

40

例えば、ピクチャ中の優的な高テクスチャコンテンツの程度をチェックするために、エンコーダは、ピクチャに対するテクスチャヒストグラムを計算する。エンコーダは、テクスチャヒストグラムにおけるスライディングウィンドウを適用して、テクスチャエネルギーを算出し、ピークまたは顕著な高テクスチャ帯域を判定する。式(8)は、エンコーダがウィンドウにおけるテクスチャエネルギーを算出するための一方法を示す。スライディングウィンドウは、最小テクスチャレベル閾値 g_0 (デフォルトで30)からスライドし始め、エンコーダは、 g_0 でのウィンドウ値 $W(g)$ を計算する。スライディングウィンドウは、そのウィンドウに対するテクスチャエネルギーの算出の後に右へ1シフトし、エンコーダは、新規値 g_0 で始まる次のウィンドウ値 $W(g)$ を計算する。これは、エンコーダがヒストグラムにおいて表されるテクスチャレベルの最大に達するまで続く。

【0154】

50

$F(g)$ をピクセルごとのテクスチャレベルのヒストグラムとする。 $E(g)$ を、そのテクスチャレベルに対するテクスチャエネルギーとし、 $E(g) = F(g) * g$ である。エンコーダは、スライディングウィンドウのテクスチャエネルギー $W(g)$ を以下のように算出する。

【0155】

【数8】

$$W(g) = \sum_{g=g_0}^{g_0+10} (F(g) * g) \quad (8).$$

【0156】

10

最大スライディングウィンドウエネルギー $W(g)$ が、全体的なピクチャエネルギーのある特定のパーセント閾値を超えた場合、その最大スライディングウィンドウエネルギー $W(g)$ に対する g_0 が、平滑領域に対する閾値を調整するのに使われる。

【0157】

図20は、 g_0 からの新しい適応平滑領域閾値を判定するのに使われる疑似コード200を示す。 g_0 が100を超える場合、適応閾値は30に設定される。エンコーダは、 g_0 が30より小さいかどうかをチェックし、小さい場合、適応閾値を14に設定する。 $30 < g_0 < 100$ の場合、適応閾値は、テーブル $g_i\text{F}\text{l}\text{a}\text{t}\text{T}\text{h}\text{T}\text{a}\text{b}\text{l}\text{e}$ にある値に設定される。ビデオ品質を保つために、最後の適応閾値との新しい適応閾値の最大差は、シーン変更キーピクチャを除くすべてのピクチャに対して + / - 4 を上限とする。適応平滑閾値は、テクスチャ加工されたブロックを識別するのに使われる閾値を超えるべきでない。例えば、図20において、最高適応閾値は30である。

20

【0158】

あるいは、エンコーダは、他の何らかのやり方で（例えば、異なるテクスチャ強度またはエネルギー基準値で、スライディングウィンドウなしで、異なるように構成されたスライディングウィンドウで、テーブルまたは他のデータ構造中の異なる閾値で、適応閾値の間の差に上限を設けずに、他の何らかのやり方で適応閾値における差に上限を設けて、など）、テクスチャレベル閾値を適応的に調整する。

【0159】

30

D. I ピクチャ及びPピクチャの差分量子化

記載した差分量子化技術及びツールは、イントラピクチャ及び予測ピクチャ上で単独でも、組み合わせても用いることができる。Iピクチャ差分量子化（IピクチャDQ）という用語は、Iピクチャへの差分量子化の適用を指し、Pピクチャ差分量子化（PピクチャDQ）という用語は、Pピクチャへの差分量子化の適用を指す。IピクチャDQを使用すると、結果としてより高い品質Iピクチャを生じ、品質の向上は、これらIピクチャに従属する予測ピクチャに対して、より長く保つことができる。PピクチャDQは、イントラ及びインターブロック両方においてPピクチャの品質をさらに向上させ得るが、こうしたPピクチャの品質は、予測の元になるピクチャの品質にも依存することになる。同様に、以後の予測ピクチャの品質に与えるPピクチャDQの影響力は、以後の予測ピクチャの、予測の元になるピクチャとの類似性に依存することになる。

40

【0160】

E. 差分量子化間隔

IピクチャDQ及びPピクチャDQは両方とも、本明細書において記載する技術の1つまたは複数を用いて、異なるテクスチャレベルブロックに対して異なるQPを適用するかどうか決定する。品質とビット使用の釣合いをとるために、PピクチャDQに対して費やされるビットの量を制御するのに、PピクチャDQ間隔を用いることができる。例えば、エンコーダは、n-1であるn個のPピクチャごとに、1つに対してPピクチャDQを使うことを選ぶが、差分量子化Pピクチャの間の間隔中のピクチャに対してはPピクチャDQをスキップする。エンコーダは、いくつかのPピクチャの知覚品質を向上させるように、差分量子化用にビットを費やし、こうした品質向上は、他の予測ピクチャに持ち越され

50

る。同時に、D Q 間隔は、エンコーダが予測ピクチャの差分量子化用に費やすビットの全体数を制限するのを助ける。

【0161】

代替的には、エンコーダは別の間隔を選択する。例えば、エンコーダは、I ピクチャごとにただ 1 つの P ピクチャに対して P ピクチャ D Q を使うことを選んでも、他の何らかの間隔を選んでもよい。間隔は、固定でも適応でもよい。例えば、エンコーダは、エンコードされるコンテンツのタイプに基づいて、P ピクチャ D Q 間隔を適応的に調整してよい。

【0162】

V . 階調傾斜検出

映像圧縮において混入される可視的な様々なアーチファクトのうち、輪郭化が、量子化によって引き起こされ得る特定の 1 つのアーチファクトである。輪郭化アーチファクトは、そうでなければ連続する、非常に平滑な領域であるもの、例えば空、水などにおける、構造化された階調の不連続として、人間の目によって知覚される。このような不連続は、非常に混乱させるものとなる場合があり、ピクチャの他の部分が可視的な歪みがほとんどなく符号化されていても、人間観察者に、ピクチャ全体がひどく歪んでいると結論づけさせる可能性がある。

【0163】

階調傾斜領域は、輪郭化アーチファクトを生じさせ得る。一定義によると、領域が平滑または比較的平滑であるが領域内部でピクセル値が徐々に変わること、領域は、階調傾斜領域であると見なされる。したがって、階調傾斜領域及び平面領域が両方とも平滑領域であると見なされるが、階調傾斜領域は、平面領域とは異なる。一定義によると、平面領域は、平面領域全体を通して不变または比較的不变なピクセル値によって特徴づけられる。階調傾斜領域は通常、はっきりとしたエッジ及び広範なテクスチャ細部を欠いている。

【0164】

図 21 は、階調傾斜領域の 2 つの例を示す。各領域内の階調傾斜方向は、矢印で表される。階調傾斜領域 2100 内で、輝度値は、領域の上から下に向かって徐々に増大する。階調傾斜領域 2100 内の傾斜の方向は、領域の各部分において同じである。階調傾斜領域 2110 内で、輝度値は、領域の中心からエッジに向かって徐々に増大する。階調傾斜の方向は、階調傾斜領域 2110 内部で変化する。ただし、狭い近隣にある各点での階調傾斜方向は、中心点を含む近隣を除いた、近隣にある他の点での階調傾斜方向の小さい角内である。図 21 に示すように、階調傾斜領域は、階調傾斜方向が領域全体において不变の領域、及び階調傾斜方向が近隣において小さく変化する領域を含む。

【0165】

図 22A は、階調傾斜領域 2210 、テクスチャ加工された領域 2220 、シャープエッジ領域 2230 及び平坦領域 2240 を有する例示的ピクチャ 2200 を示す図である。図 22B は、階調傾斜領域 2210 内での量子化の結果を示す図である。現在見えている（例えば、マクロブロック 2250 内部の）バンド効果は、輪郭アーチファクトである。図 22C は、マクロブロック 2250 の詳細を示す。マクロブロック 2250 中の輝度サンプルの上半分に対する変換係数の量子化は、DC 値 68 に起因する均一な値を結果として生じる。マクロブロック 2250 中の輝度サンプルの下半分に対する変換係数の量子化は、DC 値 70 に起因する均一な値を結果として生じる。したがって、輝度サンプルに対する変換係数の量子化により、マクロブロック 2250 中の上半分の 8 × 8 ブロックと下半分の 8 × 8 ブロックとの間の可視的な輪郭アーチファクトが作成されている。

【0166】

多くの既存ビデオエンコーダでは、ピクチャ中の輪郭化アーチファクトを削減しようとする試みにおいてビデオピクチャ全体に適用される技術を用いている。このような技術は、特に輪郭化アーチファクトをほとんどまたは全く含まない領域内で、ビットの浪費を結果として生じ得る。したがって、記載するいくつかの技術及びツールにより、エンコーダは、輪郭化アーチファクトが起こりそうな階調傾斜領域を検出することができる。階調傾斜領域が検出されると、エンコーダは、階調傾斜領域内での輪郭化アーチファクトの混入

10

20

30

40

50

を削減し、または避ける符号化決定（例えば、Q P の調整）を行うことができる。そうすることによって、エンコーダは、ビットをより効果的に割り振り、より優れた表示品質を達成することができる。

【0167】

階調傾斜領域を検出するために、エンコーダは、以下の技術の1つまたは複数を実装することができる。

1. 検出された領域（群）内での輪郭化アーチファクトの混入の削減または除去に注目した符号化決定を用いた階調傾斜領域検出。

2. 計算コストを削減し、かつ／またはフィルム粒子などの異常があるにも関わらず正確な階調傾斜領域検出を可能にするための、領域に基づく階調推定及びダウンサンプリング。10

3. 局所的な近隣における漸進階調変化を検出するための階調一貫性チェック。

4. ピクチャ中の階調傾斜領域（群）の存在に関するバイナリ決定を行うためのバケット投票（bucket voting）。

5. エンコーダが適切な符号化決定を行うのを助けるための（例えば、マクロブロックレベルでの）階調傾斜マスク及び階調方向マップの生成。

【0168】

図23は、汎用化された領域に基づく階調傾斜検出技術2300を示す。図10のエンコーダ1000などのエンコーダまたは他のツールが、技術2300を実施する。いくつかのケースでは、領域に基づく階調傾斜検出技術2300は、ピクチャ中の各ピクセルに対する階調傾斜方向を見つける必要をなくすことによって、階調傾斜コンテンツのより速い検出を可能にする。例えば、ピクチャは、同じサイズの非重複矩形領域に分けられる。領域のサイズは、実装形態に応じて変わり得る。一実装形態では、領域は、 16×16 マクロブロック（4つの 8×8 ブロック）である。好ましくは、領域は、マクロブロック整列を可能にするサイズである。20

【0169】

段階2310で、エンコーダは、現在の領域が平滑領域であるかどうかチェックする。例えば、それに割り当てられたブロック階調値が30より小さい場合、エンコーダは、 8×8 ブロックが平滑として特徴づけられるピクチャのテクスチャマップを用い、またはエンコーダは、セクションIIまたはIVで記載した別の技術を用いて、現在の領域が平滑であるかどうかチェックする。領域が多様なブロックを含むとき、領域は、領域内に含まれるすべてのブロックが平滑である場合は（または代替的には、ほぼ最小数のブロックが平滑である場合は）平滑領域であると見なされる。異なる実装形態では、ある特定の領域またはブロックが平滑であるかどうか判定するための異なる基準を用いてよい。例えば、ピクチャがダウンサンプリングされる場合、領域が平滑であるかどうか判定するための基準は、異なり得る。30

【0170】

領域が平滑でない場合、次の領域が処理される（段階2320）。平滑領域に対して、エンコーダは、段階2330で階調方向を見つける。例えば、エンコーダは、図26、27を参照して記載したような技術を用いて階調方向を見つける。あるいは、エンコーダは、他の何らかの技術を用いて階調方向を見つける。40

【0171】

段階2340で、エンコーダは、領域に対する階調方向を見つけるのに用いられる技術及び基準に依存する閾値及び／または決定実行論理を使って、領域に対して階調傾斜決定を行う。処理されるべき領域がそれ以上ある場合、エンコーダは、次の領域を処理する（段階2320）。一実装形態では、ピクチャ中の異なる領域に対する初期階調方向を計算した後、エンコーダは、ピクチャ内でスライディングウィンドウを適用することによって、異なる領域内に階調傾斜が存在するかどうかを示すバイナリマスクを生成する。バイナリマスクにおける情報により、エンコーダは、正確な階調傾斜決定を行うことができる。

【0172】

図24は、図10に示したようなビデオエンコーダ内の例示的な階調傾斜領域（gradient slope region）検出装置（GSR検出装置）2400のブロック図である。GSR検出装置2400は、現在のピクチャ2405から入力としてピクセルデータを入手する。

【0173】

ピクチャサイズ及び可能性としては他の因子に応じて、GSR検出装置2400は、ダウンサンプリングモジュール2410においてダウンサンプリングを実施するかどうか判定する。例示的ダウンサンプリング技術については、後で説明する。

【0174】

階調算出器2420は、（おそらくはダウンサンプリングされた）ピクセルデータ及びテクスチャマップ2425を入力として入手し、平滑領域に対する階調を算出する。例えば、階調算出器2420は、図26、27を参照して記載したような技術を用い、または他の何らかの技術を用いる。階調算出における例示的領域サイズは 16×16 であるが、領域のサイズは、実装に応じて変わり得る。ダウンサンプリングが適用されるかどうか、及びどの程度適用されるかに応じて、階調が算出される対象の領域は、元のピクチャ2405中で異なる量のエリアを表し得る。階調算出器2420は、平滑領域に対する階調方向を示すマップまたは他のデータ構造を出力する。

10

【0175】

一貫性チェック器2430は、平滑領域に対する算出された階調を入手し、例えば、後で説明するように、そうした階調の角の一貫性をチェックする。一貫性チェック器2430は、一貫性マップまたは算出された階調に対する一貫性情報を示す他のデータ構造を生じる。

20

【0176】

決定モジュール2440は、（一貫性チェックの後で）追加の決定規則を使って、平滑領域が階調傾斜領域と見なされるべきかどうか判定する。例示的決定規則及び基準については、後で説明する。決定モジュール2440は、一貫性マップまたは一貫性情報を示す他のデータ構造を検討し、算出された階調方向または他の情報を検討してもよい。決定モジュール2440は、階調算出において使われる領域サイズと同じまたは異なるサイズの領域に対する、マップまたは他のデータ構造における決定情報を出力する。

【0177】

各領域に対する決定は、マスクジェネレータ2450に与えられ、マスクジェネレータ2450は、ピクチャ中の領域に対する階調傾斜決定を示す階調傾斜マスク及び／またはバイナリ階調傾斜決定マスク2495を生じる。例えば、マスク2495は、各階調傾斜領域に対して「1」に等しいビット及び他の領域に対して「0」に等しいビットを含む。算出された階調を入力として受諾し、マスクジェネレータ2450は、ダウンサンプリング及びマスク決定を明らかにする、元のピクチャの異なる領域に対する最終的な階調傾斜を示す別のマスク2495を生じ得る。GSR検出装置2400が、階調算出の前にダウンサンプリングを実施すると、マスクジェネレータ2450は、ダウンサンプリングされた領域に対する階調傾斜を、元のピクチャの対応する領域に割り当てることができる。

30

【0178】

GSR検出装置2400のコンポーネントは、図24において別々のモジュールとして示してあるが、こうしたコンポーネントの機能は、実装形態に応じて、異なるモジュールに再構成し、結合し、または分けることができる。さらに、GSR検出装置2400のコンポーネントは、他の実装形態では省略することができる。例えば、ダウンサンプリングは必要とされない。GSR検出装置は、テクスチャマップを入力としてとる必要がなく、その代わりに、領域が平滑であるかどうかの指示を他の何らかのソースから受け取ることができる。GSR検出装置は、一貫性チェック器を使う必要がない。GSR検出装置は、領域が階調傾斜領域であるかどうかについて何らかの種類の決定を行うことになるが、決定がどのようにして行われるかの詳細（決定モジュールにおける決定規則を含む）は、実装に応じて変わり得る。階調傾斜決定は、バイナリマスクに含まれる必要はなく、他の何らかのやり方で、エンコーダの他の部分に伝達することができる。

40

50

【0179】

A. ダウンサンプリングを用いた領域に基づく階調方向推定

計算コストを削減するために、領域に対する階調方向を見つけるのに先立って、ダウンサンプリングを用いることができる。一実装形態では、元のピクチャ幅が1280より大きく高さが720より大きい場合、元のピクチャは、4:1ダウンサンプリングされる。例えば、ピクチャ幅が1920ピクセルかつピクチャ高さが1080ピクセルの1080p配置において、デコーダは、幅が480ピクセルかつ高さが270ピクセルのダウンサンプリングされたピクチャを生じる。

【0180】

通常、ダウンサンプリングされたピクチャは、同じサイズの複数の非重複矩形領域に分けられる。例えば、ダウンサンプリングの後、各 16×16 領域は、元の最大解像度ピクチャの4つのマクロブロック(16ブロック)に対応する。ダウンサンプリングされたピクチャ中の領域は、領域が対応する少なくとも12個のブロックが平滑である場合、平滑領域であると見なされる。領域サイズは、実装に依存し、階調推定における領域と元のピクチャ中の領域との間の関係は、ダウンサンプリング比に応じて変化する。

10

【0181】

ダウンサンプリングは、フィルム粒子などの異常が存在するのにも関わらず、階調傾斜領域検出の正確さを向上させるのにも有用である。例えば、図25に示すようなDC値のブロックをもつ、ピクチャの一的部分2500を検討する。ピクチャ部分2500の大部分は、ピクチャ部分の上から下に向かって徐々に増大するDC値で示したように、一貫した階調傾斜方向を有する。ただし、白いサンプル値は、最大解像度で異常な階調傾斜方向を生じるフィルム粒子によって影響を受けるDC値を表す。単純2:1ダウンサンプリングを水平かつ垂直に行うことにより、暗色の縁のあるサンプル値が、階調傾斜方向を算出するのに使われる。ダウンサンプリングされた値は一貫した階調傾斜を保つので、フィルム粒子は、階調傾斜の検出に影響を与えない。

20

【0182】

ダウンサンプリングは、他のピクチャ解像度に対して用いることができ、他のダウンサンプリング比も用いることができる。

【0183】

B. 階調傾斜方向の算出

30

一実装形態では、平滑領域に対する階調傾斜方向を算出するために、(図26において定義される)2つの 16×16 コンパスオペレータ K_H 及び K_V が領域に適用される。こうすることにより、領域に対して2つの階調 g_x 、 g_y が生じ、一方は水平方向用で一方は垂直方向用である。 16×16 領域に対して、コンパスオペレータは、領域の一部の値に正の重み、及び領域の他の値に負の重みを与える。あるいは、コンパスオペレータは、階調を他の何らかのやり方で計算する。

【0184】

で示される、階調方向の角度表示は、2つの階調から導出され、[0、255]内の整数にマップされる。図27の疑似コード2700は、図26のコンパスオペレータを使って、領域に対する階調方向を計算する例示的手順(

40

【0185】

【数9】



【0186】

は、要素ごとの積を示す)を示す。領域がテクスチャ加工された領域またはエッジ領域の場合、ルーチンは-2を返す。領域が平滑だが平坦である(領域に対する階調 g_x 及び g_y への低絶対値によって指示される)場合、ルーチンは-1を返す。それ以外の場合、ルーチンは、階調傾斜を、同じ逆正接値に対する傾斜方向の間を区別するためのオフセット(例えば、正の逆正接値が、右上傾斜、それとも左下傾斜を示すのか)を用いて、水

50

平階調 g_x に対する垂直階調 g_y の逆正接として計算し、傾斜値の範囲を正の数として表す。

【 0 1 8 7 】

あるいは、階調方向は、他の何らかのやり方で計算される。例えば、エンコーダは、異なるコンパスオペレータ、傾斜領域に対する異なる閾値、傾斜を計算するための異なる論理、及び／または傾斜情報の異なる表現を用いる。

【 0 1 8 8 】

C . 近隣階調一貫性チェック

エンコーダは、領域が階調傾斜領域と見なされるべきかどうかについて正確な決定を行うのを助けるために、領域に対する階調一貫性チェックを実施することができる。階調一貫性チェックは、階調傾斜コンテンツ検出における「虚報」を避けるのを助ける。一実装形態では、階調傾斜一貫性チェックは、階調傾斜の一貫性を判定するための、 3×3 スライディングウィンドウ（3領域かける3領域）の使用を伴う。10

【 0 1 8 9 】

図28は、階調傾斜領域に対する一貫性チェックを実施する技術を示す。図10のエンコーダ1000などのエンコーダまたは他のツールが、技術2800を実施する。

【 0 1 9 0 】

段階2810で、エンコーダは、スライディングウィンドウをピクチャ中の現在の領域に位置決めする。段階2820で、エンコーダは、スライディングウィンドウ中の領域の階調方向をチェックする。次いで、段階2830で、エンコーダは、現在の領域に対する一貫性決定を行う。例えば、（可能性としてはダウンサンプリングされた）ピクチャ中の検出された平滑領域の階調方向を与えられて、階調一貫性チェックが、 3×3 の隣接領域を含むスライディングウィンドウを用いて実施される。ウィンドウは、ラスタ走査順に動かされることにより、ウィンドウは、（例えば、ウィンドウを領域上でセンタリングし、一貫性チェックを実施し、次いで、ウィンドウをピクチャを横断するように左から右に動かすことによって）ピクチャ中の領域上に位置決めされる。所与のウィンドウ値に対して、一貫性チェックは、ウィンドウ内部の9個すべての領域の最大階調方向と最小階調方向（例えば、図27を参照）との間の差が、32より小さくなる（傾斜が0から255の数で表されるとき、45度に等しい）よう要求する。この条件が満たされる場合、 3×3 組の領域に対する移動するウィンドウの値は1となり、満たされない場合は0となる。あるいは、エンコーダは、異なる機構を用いて、例えば、異なるサイズのスライディングウィンドウ、最大傾斜～最小傾斜に対する異なる傾斜範囲閾値、傾斜一貫性に対する分散などの異なる測度、及び／または異なるチェックパターンを用いて、傾斜方向の一貫性をチェックし、あるいは領域の組とは反対の各領域に対するスライディングウィンドウ値を計算する。一貫性チェックは、異なる表現の傾斜情報に対しては変化する。2030

【 0 1 9 1 】

エンコーダは次いで、次の領域の組を処理することができる（段階2840）。出力として、エンコーダは、マスクまたは決定情報を示す他のデータ構造を生じる。例えば、エンコーダは、スライディングウィンドウを位置決めし、ピクチャ中の領域の組に対して一貫性チェックを実施し、領域の各組に1（一貫した傾斜）または0という決定を割り当てることによって取得されたバイナリー貫性マスク（本明細書では、consistency Maskと呼ばれる）を生じる。40

【 0 1 9 2 】

任意選択で、エンコーダは、決定情報をに対してさらに処理を実施する。一部の実装形態では、エンコーダは、ピクチャに対する階調一貫性決定を改善するのを助けるために、一貫性マスクに対するモーフオロジー演算（morphological operation）を実施する。2つの可能なモーフオロジー演算は、Erode及びDilateである。

【 0 1 9 3 】

例えば、Erode演算は、consistency Mask中のすべてのビットに対して実施され、その後にDilate演算が続く。Erode演算において、最も近い450

つのピクセル（ここでは、consistency Mask中の値）中で、複数が最初に0の印をつけられた場合、最初に1の印をつけられたビットには、0の印がつけられる。Dilate演算において、最も近い4つのピクセル中で、複数が最初に1の印をつけられた場合、最初に0の印をつけられたビットには、1の印がつけられる。

【0194】

あるいは、エンコーダは、モーフォロジー演算も決定情報の他の後処理も使わずにマスクを生成する。

【0195】

D. 決定規則及びバケット投票

一貫性チェックを実施した後でも、平滑領域の発生率は相当低い場合もあり、平滑領域が相当分離される場合もあるので、平滑領域を特別にエンコードするのは非効率的であろう。例えば、モーフォロジー演算を適用した後でも、依然として、差分量子化を必要としないように十分に分離される、consistency Maskで表される階調傾斜領域が存在する場合がある。一部の実装形態において、エンコーダは、決定規則（例えば、バケット投票を含む）を使って、ピクチャ中の階調傾斜領域にDQが適用されるべきかどうか決定するのを助ける。図24のGSR検出装置2400では、決定モジュール2440がこのような決定を行う。

【0196】

一実装形態では、エンコーダは、現在のピクチャが、consistency Maskに基づく重要な階調傾斜を含むかどうかに関する1つまたは複数のバイナリ決定を行う。マスクconsistency Maskは、同じサイズの25個の矩形領域（バケットと呼ばれる）に分けられ、各行に5バケット、各列に5バケットがある。（したがって、「バケット」領域は、決定に使われる領域及び階調算出に使われる領域より大きい）。各バケット内部の1がカウントされる。Buckets[i][j]を、場所(i, j)でバケットに含まれる1の数とし、0 i, j 4である。水平及び垂直バケットの見積もり、すなわち、それぞれバケットの各列中の1の数及びバケットの各行中の1の数も、以下の関係によって算出される。

【0197】

【数10】

$$\begin{aligned} \text{BucketProjection_H}[j] &= \sum_{0 \leq i \leq 4} \text{Buckets}[i][j] \\ \text{BucketProjection_V}[j] &= \sum_{0 \leq i \leq 4} \text{Buckets}[i][j] \end{aligned} \quad (9)$$

【0198】

本実装形態において、ピクチャは、以下の条件のどれかが満たされる場合、重要な階調傾斜を含むものと見なされる。

1.（バケット分散に関係なく）consistency Maskにおけるピクセルの少なくとも6%が1の印をつけられている、あるいは

2. バケットの1つまたは複数において、ピクセルの少なくとも75%が1の印をつけられている、あるいは

3. バケット見積もりの1つまたは複数において、ピクセルの少なくとも20%が1の印をつけられている。

【0199】

例えば、サイズが960×1440のダウンサンプリングされたピクチャに対する16×16領域は、サイズが20×30のマスクで表され（各値は、ダウンサンプリングされたピクチャの3×3組の領域用）、マスクは、25バケットに分けられ、各バケットは、一貫性マスクの24領域に対応する。各バケットは、総計が25×24 = 600ビットであるconsistency Maskのうち24ビットを含む。エンコーダは、図29に示すように分散した、各バケット中の1の数をカウントする。エンコーダは、1の総数がすべてのビットの6%より大きいかどうかチェックする。この場合、（図29に示す）1

10

20

30

40

50

の総数は 83 であり、これはすべてのビットの 6 %より大きい。したがって、この場合のエンコーダは、上記条件 1 を満たすことにより、バケット見積もりをスキップするはずである。1 の総数が条件 1 に対する閾値より下だった場合、エンコーダは、いずれかのバケット中のビットの 75 %が 1 である（条件 2）かどうかチェックし、必要な場合は、水平及び垂直バケット見積もり（条件 3）をチェックして、階調傾斜領域であると指示される領域が、後で説明するマクロブロックレベル階調傾斜マスクなど、階調傾斜マスク及び決定マスクが生成されるべきようなものであるかどうか判定する。

【0200】

あるいは、エンコーダは、マスク `consistencyMask` または他の表現における一貫性情報を処理する、他の決定規則を使う。例えば、条件 1、2、3 に示すパーセント閾値は、実装に応じて変わり得る。あるいは、条件の 1つまたは複数が省略され、または条件が、他の条件（例えば、バケット見積もりに対する異なる方向など）で順序付けられ、置き換えられ、もしくは補完される。一貫性情報のチェックとは別に、エンコーダは、DQ がピクチャ中の階調傾斜領域に適用されるべきかどうか、またはどの程度の DQ が適用されるべきか決定する際、階調値及び / または他の情報を検討することもできる。別の代替形態として、エンコーダは、こうした決定規則を完全に省いても、単に階調傾斜マスクを生成する際に `consistencyMask` を使ってもよい。

10

【0201】

E. マクロブロックレベル階調傾斜マスク生成

階調傾斜情報を、以後のエンコーダ決定実行に有用な形で提供するために、エンコーダは、情報をマップ、マスク、または他のデータ構造に入れる。情報は、階調傾斜領域存在 / 不在情報ならびに階調傾斜領域に対する実際の階調方向値を含み得る。

20

【0202】

階調傾斜存在 / 不在情報に対しては、階調傾斜領域が検出された場合、エンコーダは、階調傾斜マスクを生じる。例えば、エンコーダは、領域レベルマスク（`consistencyMask` など）を元のピクチャに対するマクロブロックレベルに戻すようにコンパートすることによって、マクロブロックレベル階調傾斜マスク（本明細書では、MBS1opeMask と呼ばれる）を生じることにより、可能なダウンサンプリングを検討する。`consistencyMask` 中の各値は、元のピクチャ中の 9つのマクロブロック、またはピクチャが 4 : 1 ダウンサンプリングされている場合は 36 のマクロブロックに対応することに留意されたい。`consistencyMask` 中の、値が 1 の各ビットに対しては、エンコーダは、平滑でないマクロブロックを除く、MBS1opeMask 中の対応するマクロブロックに 1 の印をつける。平滑性チェックはやはり、階調傾斜検出における虚報を避けるのを助ける。例えば、一実装形態では、エンコーダは、テクスチャマップを使ってマクロブロック中のブロックに関するテクスチャ情報を取得し、マクロブロックは、マクロブロック内部の 4つすべてのブロックが平滑である場合のみ、平滑と見なされる。

30

【0203】

あるいは、エンコーダは、階調決定情報を他の何らかの形で提供し、かつ / またはマクロブロック平滑性に対して他の何らかの決定を用いる。

40

【0204】

階調方向情報に対して、各領域の階調方向を、それに対応するすべての平滑なマクロブロックを割り当てることによって、階調方向マップが生成される。そうする際、エンコーダは、階調算出前のダウンサンプリングによる、元のピクチャのマクロブロックと階調領域との間の可能なサイズ差を明らかにする。

【0205】

生成された階調傾斜マスク及び階調方向マップは次いで、エンコーダにおいて、より優れた符号化決定を行うのに使われる。概して、階調傾斜領域検出装置によって生成された結果は、エンコーダによって、他の符号化決定を行うのに用いることができる。例えば、エンコーダは、生成された階調傾斜マスク及び / または階調方向マップに基づいて量子化

50

決定を行うことができる。可能なエンコーダ決定のいくつかについては、後で説明する。

【0206】

V I . 非ゼロ A C 係数を維持するための量子化調整

通常、ピクチャは、エンコーダ内のレート制御ユニットによってピクチャレベル量子化パラメータを割り当てられる。同じピクチャレベル Q P を使うと、高度にテクスチャ加工されたマクロブロックを表すのに使われるビットの量は通常、低テクスチャ加工されたマクロブロックを表すのに使われるビットの量よりはるかに大きい（10 ~ 50 倍も大きい）。しかし、人間の視覚システムは、高度にテクスチャ加工された密なエリア内では、歪みに対して、低くテクスチャ加工された平滑なエリア内よりも鈍感なので、低テクスチャ加工されたマクロブロックに対しては比較的小さい Q P 、及び高度にテクスチャ加工されたマクロブロックに対してはより大きい Q P を使うことは意味のあることである。10

【0207】

このことは、マクロブロックを人間の視覚の重要性によって分類し（通常、ブロックの分散またはブロック内部の階調の強度を用いる）、目標数のビットを、何らかの知覚重みづけに比例して割り当てる、よく用いられる戦略につながる。修正されるべき、各マクロブロックに対する量子化パラメータは、ピクチャレベル量子化法を重みづけによって修正することによって選択される。

【0208】

変化が非常に少ない平滑領域内では、ブロックがしばしば、妥当な低さの Q P でも、 D C 係数のみの（非ゼロ A C 係数は残らない）エネルギーをもつように量子化されることが、実験により示されている。驚くべきことに、極度に平滑な領域内の隣接ブロック中の D C 値は、ブロックごとに1だけ変化し、知覚されるブロック状の、輪郭化アーチファクトは、絶対項におけるこのような小さい差で予期するよりずっと激しい。そうでなければ満足に符号化されていたピクチャ中の比較的小さい領域内でのこのタイプのアーチファクトの出現により、ピクチャ全体に対する知覚される全体的品質が低下させられる。20

【0209】

従来の速度歪みに基づくとともに知覚に基づくマクロブロック Q P 選択技術は、この状況をうまく扱わない。速度歪みの最適化により、平滑ブロックは、絶対項における歪みが小さいので満足に符号化されたと見なされることになり、したがって、こうしたブロックに対してそれ以上ビットが費やされることはない。一方、典型的な知覚に基づく方法では、マクロブロックを知覚クラスに分類し、マクロブロックの知覚クラスに従って、事前定義されたオフセットをピクチャレベル量子化パラメータに加え、または引くことによって、量子化パラメータを各マクロブロックに割り当てる。事前定義されたオフセットが非常に強度（例えば、平滑領域に対する Q P を1に削減する）でない限り、このような方法は、変化の小さい平滑ブロックが単一の非ゼロ D C 係数に量子化されず、すべての A C 係数がゼロに量子化されることを保証できない。しかし、非常に強度なオフセットの設定により、知覚品質を向上させるのにビットを必要としなくてもよいマクロブロック中で費やされるビットが増加する場合があり、ビットレートを非効率的に上昇させ、エンコーダによってレート制御用に選択されたピクチャレベル量子化パラメータと衝突する。30

【0210】

したがって、後で説明するいくつかの技術及びツールは、ブロッキングまたは輪郭アーチファクトの混入を削減し、または取り除くのに十分なビットが平滑領域に割り振られるように、ピクチャ中でビットを選択的に、思慮深く割り振る。

【0211】

例えば、エンコーダは、I ピクチャ中の各マクロブロックに対して Q P を算出し、量子化パラメータを選択して、平滑ブロックに十分なビットを割り振り、そうすることによって、I ピクチャにおける、知覚されるブロッキングアーチファクトを削減する。1つまたは複数の平滑ブロックを有する各マクロブロックに対して、マクロブロックのブロックごとに量子化された少なくとも N 個の非ゼロ A C 係数があるように、 Q P が選択され、ここで N は 1 以上の整数である。しばしば、維持される A C 係数は、変換の最低周波数 A C 基50

底関数に対する係数であり、この関数は、漸進値の変化をブロックに渡り水平に及び／または垂直に特徴づける。これは、特に変化が少ない平滑領域に対して、各ブロックの、知覚される表示品質をもち直す傾向にある。一実装形態では、エンコーダは、依然として A C 係数を所望通りに維持する、ピクチャ Q P を超えない最大 Q P を選択する。非ゼロ A C 係数が維持されない状況（例えば、非常に平坦なブロック）があり得る。ただし、概して、このように、エンコーダは、比較的小さい Q P でビットを費やすことに過度に積極的ではなく、ピクチャ Q P との衝突を削減し、または避ける。

【 0 2 1 2 】

妥当な N 値により、選択された Q P は、ほとんどのマクロブロックに対して変わらない。すなわち、ほとんどのマクロブロックに対してはピクチャ Q P と同じであり、ごくわずかの平滑ブロックが影響を受ける。N の妥当な値は 1、2、3 または 4 である。選択された Q P は、低テクスチャのマクロブロックに対して比較的変わりやすい。一実装形態では、N = 1 または 2 であると、ピクチャのビットレートが増大し過ぎずに、知覚される品質が向上する。10

【 0 2 1 3 】

図 3 0 は、1 つまたは複数の非ゼロ A C 係数を維持するのを助けるためのマクロブロック Q P を選択する例示的技術 3 0 0 0 を示す。図 1 0 のエンコーダ 1 0 0 0 などのエンコーダまたは他のツールが、技術 3 0 0 0 を実施する。

【 0 2 1 4 】

段階 3 0 1 0 で、エンコーダは、マクロブロックの各輝度ブロックの N 番目に大きい A C 係数を見つける。例えば、N = 2 の場合、エンコーダは、 16×16 マクロブロックの 4 つの 8×8 ブロックそれぞれの 2 番目に大きい A C 係数を見つける。A C (0)、A C (1)、A C (2)、A C (3) を、それぞれ、4 つの輝度ブロック 0、1、2、3 に対する N 番目に大きい係数とする。マクロブロック中の異なるブロック組織の場合、N 番目の係数は、マクロブロック中のより多くの、またはより少ないブロックに由来する。20

【 0 2 1 5 】

段階 3 0 2 0 で、エンコーダは、こうした N 番目の係数値のうち、最小のものを見つける。4 つのブロックのうち N 番目の係数に対して、 $A C_{min} = \min(A C(0), A C(1), A C(2), A C(3))$ である。他のブロック番号に対しては、 $A C_{min}$ は異なるやり方で計算される。30

【 0 2 1 6 】

段階 3 0 3 0 で、エンコーダは、マクロブロックに対する Q P を、 $A C_{min}$ がその Q P に対するデッドゾーン閾値外となるように設定する。デッドゾーン閾値は、Q P の値が量子化に使われるときに A C 係数をゼロに量子化する「カットオフ」閾値である。デッドゾーン閾値は通常、所与の Q P に対して所定であり、それに比例する。デッドゾーン閾値は、0 と第 1 の復元ポイントとの間のいずれかのポイントで選択される。エンコーダが、均一量子化法または不均一量子化法のどちらかを使うとき、第 1 の復元ポイントは、Q P 値、及び均一それとも不均一量子化が用いられるかに依存する。一実装形態では、第 1 の復元ポイントは、量子化された係数レベル = 1 の復元された値であり、レベルは、均一量子化用には $2 * Q P$ であり、不均一量子化用には $3 * Q P$ である。均一量子化の場合、カットオフ閾値は 0 と $2 * Q P$ の間となる。不均一量子化の場合、カットオフ閾値は 0 と $3 * Q P$ の間となる。例えば、デッドゾーン閾値 Z (Q P) は、均一量子化用には $Z(QP) = 6 * QP / 5$ 、不均一量子化用には $Z(QP) = 2 * QP$ として選択される。あるいは、他のカットオフ閾値を使うこともできる。40

【 0 2 1 7 】

A C 係数 A C は、 $A b s(A C) < Z(Q P)$ の場合、ゼロに量子化される。マクロブロック向けに Q P を設定する（段階 3 0 3 0 ）ために、エンコーダは、 $A C_{min}$ を Q P の候補値に対する Z (Q P) と比較し、ピクチャ Q P で開始し、量子化法に対する最小 Q P に達し（例えば、Q P = 1）または不等式 $A b s(A C_{min}) >= Z(Q P)$ が満たされるまで Q P を減少させることによって、少なくとも N 個の A C 係数を維持することになる50

、マクロブロック向けQ P ($Q P_m$)を見つけることができる。不等式 $A b s (A C_{min}) >= Z (Q P)$ が満たされる場合、エンコーダは、マクロブロックに対する閾値 Q P を、不等式を満たす第 1 の Q P (すなわち、最高適格 Q P) となるように設定する。あるいは、エンコーダは、他の論理を用いて、マクロブロックに対する Q P を計算し、例えば、最低 Q P から始め、または Q P 値のバイナリサーチを用いる。

【0218】

マクロブロック中のすべてのブロックを量子化するための $Q P_m$ を用いるプロセスは、無制約ビットレート量子化と呼ぶことができる。制約ビットレート量子化技術において、エンコーダは、上述したように、単独でマクロブロックの輝度ブロックそれぞれに対して所望の数の非ゼロ A C 係数を生じるのに必要とされる最大 Q P (ピクチャ Q P 以下) を判定する (例えば、それぞれ、ブロック 0、1、2、3 に対して $Q P_0$ 、 $Q P_1$ 、 $Q P_2$ 、 $Q P_3$)。その結果、 $Q P_m$ は、 $Q P_0$ 、 $Q P_1$ 、 $Q P_2$ 、 $Q P_3$ の最小に等しくなる。¹⁰ ビット使用を削減するために、エンコーダは、 $Q P_m$ の代わりにブロック i ($i = 0, 1, 2, 3$ 、など) を量子化するのに $Q P_i$ を使うことができる。マクロブロック全体に対して单一の Q P を指定するエンコーダにおいて、エンコーダは代わりに、 $Q P_m$ を使ってブロックを量子化するとき、各ブロック i 用の $Q P_i$ を使って量子化される際に非ゼロである A C 係数だけを保持すればよく、それにより、ブロック中の他の A C 係数が $Q P_m$ による量子化で維持されることになっても、所与のブロック中の最上位の N 個の非ゼロ A C 係数のみを保存する。図 30 に示す量子化プロセスに対して、各輝度ブロックに対する量子化プロセスは、2 パスプロセスとして実施することができる。最初のパスにおいて、エンコーダは、係数が $Z (Q P_i)$ より小さい場合は D C T 係数をゼロに「閾値化」し、それ以外の場合は、同じ D C T 係数を保持する。次いで、「閾値化」された D C T 係数は、 $Q P_m$ を使う同じやり方で量子化される。²⁰

【0219】

あるいは、エンコーダは、非ゼロ A C 係数を他の何らかのやり方で維持する。例えば、エンコーダは、マクロブロック単位以外の単位 (例えば、ブロック単位) で Q P を選択することができる。エンコーダは、I ピクチャ、P ピクチャ、または B ピクチャ、あるいはその組合せに対する A C 係数を維持することができる。

【0220】

可能な最小 Q P で、量子化された非ゼロ係数の数が N より小さい場合、N は、それにしたがって調整してよい。³⁰

【0221】

V III . D C シフトに対する差分量子化

典型的な損失性エンコーディングシナリオでは、量子化されたすべての D C 及び A C 係数を、逆量子化後に正確に回復することができるわけではない。例えば、一部のビデオコーデックでは、D C 係数値は、いくつかの Q P 及び D C 係数値に対して 1 だけシフトする (すなわち、事前量子化値に対して相対して 1 だけ増大または低下する)。この現象は D C シフトの例である。いくつかの D C 係数値の表現は、1 つまたは複数のより低い Q P での量子化及び逆量子化を通して無損失であるが、他の、より高い Q P では損失性である。⁴⁰

【0222】

A C 係数すべてが 0 に量子化され、D C 係数を正確に回復することができない、いくつかのブロックを有する領域は、D C シフトエリア内で可視的な輪郭化アーチファクトを呈する。輪郭化アーチファクトのあるこのようないくつかの領域はしばしば、空、水または光線など、平滑な階調傾斜領域である。図 31 は、量子化及び逆量子化後の階調傾斜領域内の 3 つの隣接し合うブロックにおける D C シフトを示す図である。階調傾斜領域内の 3 つの隣接し合うブロック 3102、3104、3106 の D C 値は、それぞれ、量子化より前には 68、69、70 である。量子化及び逆量子化の後、ブロック 3104 の D C 値は 70 にシフトされる。図 31 に示すように、3 つの隣接し合うブロックの D C 値は現在、68、70、70 である。このようなブロックが階調傾斜領域内にあるとき、量子化された D C 値は、知覚可能な輪郭化アーチファクトを引き起こし得る。例えば、図 22 A ~ C を再度参⁵⁰

照すると、階調傾斜領域 2210 が量子化されており、結果として、図 22B において可視的な輪郭化アーチファクトを生じる。図 22C に示すように、マクロブロック 2250 の上半分のブロックに対する DC 係数の量子化は、結果として、DC 値 68 から復元された均一な値を生じ、下半分のブロックに対する DC 係数の量子化は、結果として、DC 値 70 から復元された均一な値を生じる。

【0223】

したがって、後で説明するいくつかの技術及びツールは、ビデオエンコーダによって、DC シフトエリアを検出し、DC シフトエリアへの輪郭化アーチファクトの混入を削減し、または避けるように量子化を調整するのに用いられる。

【0224】

図 32 は、DC シフトエリアへの輪郭化アーチファクトの混入を削減し、または避けるように量子化を調整する、汎用化された技術 3200 を示すフローチャートである。図 10 のエンコーダ 1000 などのエンコーダまたは他のツールが技術 3200 を実施する。

【0225】

段階 3210 で、エンコーダは、シフトエリアを検出する。DC シフトエリアの検索は、前の階調傾斜検出で手伝うことができる。例えば、エンコーダは、後で説明するように、1つまたは複数の階調傾斜領域を検出し（または以前計算された階調傾斜検出情報を用い）、次いで、階調傾斜領域（群）内の DC シフトブロックを識別することによって、DC シフトエリアを検出する。

【0226】

段階 3220 で、エンコーダは、DC シフトエリア内の量子化を調整する。例えば、エンコーダは、DC シフトによって引き起こされる輪郭化アーチファクトの混入を削減し、または避けるために、差分量子化（DQ）を用いて、DC シフトブロックを符号化することができます。エンコーダは、いくつかのマクロブロック（DC シフトブロックを有するもの）に対する QP を削減するが、他のブロック用の QP は変更しない。DC シフトブロックを有するマクロブロック用の QP を削減することにより、マクロブロック用の DC 値を無損失に保つのが助けられ、そうすることによって、輪郭化アーチファクトの混入を削減しまたは避けることができる。エンコーダは、二値 DQ または多値 DQ を使って、DC シフト問題を解決することができ、そうすることによって、ピット使用を制御しながら表示品質を向上させる。分析するべきピクチャがそれ以上ある場合、エンコーダは、次のピクチャを処理する（段階 3230）。

【0227】

あるいは、エンコーダは、DC シフトエリア向けの量子化を、マクロブロック単位で、または他の何らかの単位で調整する。

【0228】

A. 階調傾斜検出

階調傾斜検出は、ピクチャ中の1つまたは複数の階調傾斜領域を識別するのに用いることができる。階調傾斜領域（群）は、特に量子化後に領域（群）内のブロックが非ゼロ DC 係数値及びゼロだけの AC 係数をもつ時、輪郭化アーチファクトを呈する傾向にある。見つけられると、このような領域（群）は、輪郭化アーチファクトの原因となり得る DC シフトブロックに対してチェックしてよい。

【0229】

例えば、エンコーダは、本明細書において記載した技術（セクション V）または他の何らかの技術を用いて階調傾斜領域を見つける。ブロック中の非ゼロのみの係数が量子化後に DC 係数である場合、エンコーダは、そのブロックを、DC シフトエリア調整用の候補として扱う。あるいは、エンコーダは、追加ブロックを、DC シフトエリア調整用の候補として検討する。

【0230】

B. DC シフトブロックの識別

エンコーダは、特定の候補ブロックを、DC シフトブロックとして識別する。DC シフ

10

20

30

40

50

トブロックの識別は、ブロックを圧縮するのに使われる量子化法及びQ Pの詳細に依存する。例えば、いくつかの復元されたD C係数が、あるQ Pでの元の値からシフトするが、比較的粗いQ Pでシフトする。

【0231】

あるエンコーダにおける異なるQ P用のD Cシフト係数の例を、以下のテーブルに挙げる。テーブルは、Q Pの異なる値に対するD Cシフトを呈するD C係数値を示し、Q Pは、パラメータP Q I n d e x（及び、可能性としては、ハーフステップパラメータ）から明示的に、またはパラメータP Q I n d e x（及び、可能性としては、ハーフステップパラメータ）から暗黙的に導出される。テーブルに列挙されていないD C値は、例示的エンコーダにおける指示されたQ Pに対して無損失であり、3を下回るQ Pに対するD C値（テーブルには示していない）はすべて無損失である。例示的エンコーダは、テーブルに示すものより高いQ Pに対して、D Cシフト調整を実施しない。例示的エンコーダにおいて、D C係数の量子化は、異なる量子化法（例えば、均一、不均一）に対して同じである。どのD C係数値がD Cシフト係数であるかは、異なるビデオコーデックにおいて異なる。D Z係数の量子化が異なる量子化法において異なる場合、異なる量子化法（例えば、均一、不均一）が、結果として、異なるシフトパターンを生じ得る。

【0232】

【表3】

表3：例示的DCシフト係数

PQIndex (暗黙的OP)	PQIndex (明示的OP)	元のDC	シフト されたDC	元のDC	シフト されたDC	元のDC	シフト されたDC
3-5.5	3-5.5	6	7	96	97	186	187
		15	16	105	106	195	196
		24	25	114	115	204	205
		33	34	123	124	213	214
		42	43	132	133	222	223
		51	52	141	142	231	232
		60	61	150	151	240	241
		69	70	159	160	249	250
		78	79	168	169		
		87	88	177	178		
6- 7.5,9- 10.5	6-7.5	2	1	92	93	178	177
		6	7	97	96	183	182
		11	12	102	101	187	188
		16	15	106	107	192	193
		21	20	111	112	197	196
		25	26	116	115	202	201
		30	31	121	120	207	206
		35	34	126	125	211	212
		40	39	158	157	216	217
		45	44	130	131	221	220
		49	50	135	136	225	226
		54	53	140	139	230	231
		59	58	144	145	235	236
		63	64	149	150	240	239
		68	69	154	155	245	244
30	40	73	74	159	158	249	250
		78	77	164	163	254	255
		83	82	168	169		
		87	88	173	174		

【0233】

【表4】

PQIndex (暗黙的OP)	PQIndex (明示的OP)	元のDG	シフト されたDG	元のDG	シフト されたDG	元のDG	シフト されたDG
8,11- 12	8-9	2	1	88	89	171	172
		5	6	92	91	175	174
		9	8	95	96	178	179
		12	11	99	98	182	181
		15	16	102	101	185	186
		19	18	105	106	189	188
		22	23	109	108	192	191
		26	25	112	113	195	196
		29	30	116	115	199	198
		33	32	119	120	202	203
		36	37	123	122	206	205
		40	39	126	127	209	210
		43	44	158	156	213	212
		47	46	130	129	216	217
		50	51	133	134	220	219
		54	53	137	136	223	224
		57	56	140	141	227	226
		60	61	144	143	230	231
		64	63	147	146	234	233
		67	68	150	151	237	236
		71	70	154	153	240	241
		74	75	157	158	244	243
		78	77	161	160	247	248
		81	82	164	165	251	250
		85	84	168	167	254	255

10

20

【0 2 3 4】

【表5】

PQ Index (暗黙的OP)	PQ Index (明示的OP)	元のDG	シフト されたDG	元のDG	シフト されたDG	元のDG	シフト されたDG
13-14	10-11	2	3	90	89	175	174
		5	4	93	92	177	178
		8	7	95	96	180	181
		11	10	98	99	183	184
		13	14	101	102	186	185
		16	17	104	103	189	188
		19	20	107	106	192	191
		22	21	110	109	194	195
		25	24	112	113	197	198
		27	28	115	116	200	201
		30	31	118	119	203	202
		33	34	121	120	206	205
		36	35	124	123	209	208
		39	38	126	127	211	212
		42	41	158	157	214	215
		45	44	129	130	217	218
		47	48	132	133	220	219
		50	51	135	136	223	222
		53	52	138	137	225	226
		56	55	141	140	228	229
		59	58	144	143	231	232
		61	62	146	147	234	235
		64	65	149	150	237	236
		67	68	152	151	240	239
		70	69	155	154	243	242
		73	72	158	157	245	246
		76	75	160	161	248	249
		78	79	163	164	251	250
		81	82	166	167	254	253
		84	85	169	168		
		87	86	172	171		

【0 2 3 5】

10

20

30

【表6】

PQIndex (暗黙的OP)	PQIndex (明示的OP)	元のDG	シフト されたDG	元のDG	シフト されたDG	元のDG	シフト されたDG
15-16	12-13	2	1	87	88	171	170
		4	3	90	89	173	174
		6	7	92	93	176	175
		9	8	95	94	178	179
		11	12	97	98	181	180
		14	13	100	99	183	184
		16	17	102	103	186	185
		19	18	105	104	188	189
		21	22	107	108	191	190
		24	23	110	109	193	194
		26	27	112	111	195	196
		29	28	114	115	198	197
		31	30	117	116	200	201
		33	34	119	120	203	202
		36	35	122	121	205	206
		38	39	124	125	208	207
		41	40	127	126	210	211
		43	44	158	157	213	212
		46	45	129	130	215	216
		48	49	132	131	218	217
		51	50	134	135	220	221
		53	54	137	136	222	223
		56	55	139	140	225	224
		58	57	141	142	227	228
		60	61	144	143	230	229
		63	62	146	147	232	233
		65	66	149	148	235	234
		68	67	151	152	237	238
		70	71	154	153	240	239
		73	72	156	157	242	243
		75	76	159	158	245	244
		78	77	161	162	247	248
		80	81	164	163	249	250
		83	82	166	167	252	251
		85	84	168	169	254	255

10

20

30

【0 2 3 6】

【表7】

PQindex (暗黙的QP)	PQindex (明示的QP)	元のDC	シフト されたDC	元のDC	シフト されたDC	元のDC	シフト されたDC
17-18	14-15	1	2	87	88	171	172
		3	4	89	90	173	174
		5	6	92	91	175	176
		8	7	94	93	178	177
		10	9	96	95	180	179
		12	13	98	99	182	183
		14	15	100	101	184	185
		16	17	103	102	186	187
		19	18	105	104	189	188
		21	20	107	106	191	190
		23	24	109	110	193	194
		25	26	111	112	195	196
		27	28	114	113	198	197
		30	29	116	115	200	199
		32	31	118	119	202	201
		34	35	120	121	204	205
		36	37	122	123	206	207
		39	38	125	124	209	208
		41	40	127	126	211	210
		43	42	158	157	213	212
		45	46	129	130	215	216
		47	48	131	132	217	218
		50	49	133	134	220	219
		52	51	136	135	222	221
		54	53	138	137	224	223
		56	57	140	141	226	227
		58	59	142	143	228	229
		61	60	144	145	231	230
		63	62	147	146	233	232
		65	66	149	148	235	236
		67	68	151	152	237	238
		69	70	153	154	239	240
		72	71	156	155	242	241
		74	73	158	157	244	243
		76	77	160	159	246	247
		78	79	162	163	248	249
		81	80	164	165	250	251
		83	82	167	166	253	252
		85	84	169	168	255	254

10

20

30

【0237】

表3に示すDCシフト係数を有する例示的エンコーダは概して、テクスチャ加工された領域に対して、平滑領域とは異なるQPを使う。例示的エンコーダは通常、3~5の範囲のQPを使って、平滑領域をエンコードする。上の表3に示すように、QP3~5に対して、シフトされたDC値はすべて、元のDC値より1多い。他のエンコーダは、テクスチャ領域と比較して、平滑領域用に異なるQPを使い得る。

【0238】

DCシフトブロックが検出される際の輪郭化アーチファクトの混入を削減し、または避けるのを助けるために、例示的エンコーダは、DC値をマクロブロック中で無損失に保つためのDCシフトブロックを含むそうしたマクロブロック用のQPを変更する。具体的には、例示的エンコーダは、DCシフトブロックを含むマクロブロック用QPをQP=2に削減する。（他のエンコーダは、DCシフトエリア向けに他の何らかのQPを使い得る）。概して、エンコーダは結果として、マクロブロックのブロックのDC係数を無損失に取り扱うことになる、使用可能な最大QPを選択することができる。

【0239】

エンコーダは、階調傾斜領域（群）内のDCシフトブロックを判定するために、プロックごとの平均輝度値を算出する。というのは、平均輝度値は、（変換における拡張を補償した後の）DCシフト値に対応するからである。平均輝度値により、エンコーダは、どのプロックがDCシフトを有するか推定し、または予測する。あるいは、エンコーダは、現実のDC値を算出し、それらの値をDCシフトテーブル中でルックアップして、どのプロックがシフトを有するかを正確に識別する。

【0240】

エンコーダは、階調傾斜領域（群）内で、特定の分離DCシフトブロックを排除するための追加処理を実施することができる。例示的エンコーダにおいて、現在のプロックが、階調傾斜領域内に配置されたDCシフトブロックと識別されると、周辺にある4つの隣接

40

50

ブロックがチェックされる。周辺にある4つの隣接ブロックのいずれかが平滑ブロックであり、現在のブロックのシフトされたDC値より低いDC値をもつ場合、エンコーダは、DC値を無損失に保つために、現在のブロックを含むマクロブロックに対してQP=2を使う。代替的には、エンコーダは、隣接ブロックのチェックを行わず、または、隣接ブロックの他の何らかの配置をチェックして、DCシフトエリア用のQPを変化させるかどうか判定する。

【0241】

C. 多値差分量子化コストモデル

二値DQ及び多値DQは通常、異なるビットレートコストを有する。一実装形態では、「すべてのマクロブロック」二値DQにおけるピクチャQPまたは代替QPを信号通知するのに、マクロブロックごとに1ビットが使われ、多値DQにおけるピクチャQPまたは代替QPを信号通知するのに、マクロブロックごとに少なくとも3ビットが使われる。

【0242】

エンコーダは、多値DQを使ってDCシフトブロックを含む平滑領域内のQPを削減させることができるが、エンコーダはその代わりに、二値DQシナリオにおいて、すべての平滑領域用のQPを（例えば、QP=2に）調整し、ピクチャの残りには比較的粗いピクチャQPを使うことを選ぶことができる。例えば、エンコーダは、多値DQのシグナリングコストが二値DQのコストに相対して高すぎることがわかった場合、上記のことを行うことができる。

【0243】

一実装形態では、それぞれQP=3、4、5、6からQP=2に進む平滑ブロックのコストを算出するのに、以下のテーブルが使われる。

```
g_iSmoothBlockDiffQPCost[4] = {18, 22, 28, 36}
```

このテーブルは、二値DQコストB(QP)のコスト算出の以下の例において使われる。

$$B(QP) = \text{counts_of_total_MBs} + (\text{counts_of_biLevel_Dquant_MBs} - \text{counts_of_DC_Shift_Blocks}) * g_iSmoothBlockDiffQPCost[QP - 3]$$

【0244】

コストB(QP)は、マクロブロックごとの二値コストシグナリングのコストを明らかにし、ブロックに対して3、4、5、または6ではなくQP=2を使う際の増大したビットコストを推定する。多値DQコストM(QP)は、以下のように算出される。

$$M(QP) = (\text{counts_of_frameQP_MBs} * 3) + (\text{counts_of_biLevel_Dquant_MBs} - \text{counts_of_DC_Shift_Blocks}) * 8 + (\text{counts_of_DC_Shift_Blocks} * 3)$$

【0245】

コストM(QP)は、多値DQのシグナリングコストを明らかにし、それにより、いくつかのマクロブロック量子化パラメータに対するエスケープ符号化を仮定する。B(qp) < M(qp)の場合、二値DQが使われることになり、すべての平滑ブロックに対してQP=2が使われることになる。それ以外の場合、多値DQが使われることになる。

【0246】

あるいは、エンコーダは、異なるタイプまたは構成のDQに対して、他のコストモデルを使う。あるいは、エンコーダは、DCシフトブロックが検出されると、ピクチャ全体に対するQPを削減し、または他の何らかの技術を用いて、DCシフトブロックが検出されるときに輪郭化アーチファクトの混入を削減し、または避けるように量子化を変更する。

【0247】

D. ピクチャQP切換え

例示的なエンコーダでは、多値DQは、ピクチャQPからピクチャQP+6までのどのQPを信号通知するのにも3ビットを必要とする。この範囲外のどのQPも、エスケー

10

20

30

40

50

プロセス化により 8 ビットで信号通知されることになる。平滑領域用に使われる代替 Q P は一般に、ピクチャ Q P より小さく、したがってエスケープ符号化を必要とする。

【 0 2 4 8 】

したがって、ピクチャ Q P を切り換えることにより、多値 D Q に対する符号化オーバーヘッドをセーブすることができる。例えば、エンコーダは、上述した多値 D Q コストモデルを使ってピクチャ Q P を選ぶことができる。例えば、3 値のシナリオ（例えば、ピクチャ Q P 、平滑領域 Q P 及び D C シフト Q P ）の場合、ピクチャ Q P 用の異なる候補値に対して多値 D Q コストが計算される。エンコーダは、オーバーヘッドコストが最小のピクチャ Q P を選択することができる。

【 0 2 4 9 】

あるいは、エンコーダは、他の基準を使ってピクチャ Q P を切り換え、またはピクチャ Q P 切換えを実施しない。

【 0 2 5 0 】

E . 高テクスチャマクロブロック用の粗い量子化

多値 D Q を使うという決定が行われた場合、ピクチャ Q P より高い（比較的粗い）Q P を高テクスチャマクロブロックに割り当てるによって、高テクスチャマクロブロック用に粗い量子化を用いることができる。ピクチャに対して多値 D Q を（例えば、D C シフトマクロブロックに対して比較的小さい Q P を使うために）使うという決定は、ピクチャ Q P より高いマクロブロックごとの粗い Q P を信号通知するための追加オーバーヘッドコストがないことを意味する。例えば、ピクチャ Q P + 1 は、量子化レベルでの目立ちやすい差を避けるために、粗い Q P として使うことができ、または他の何らかの Q P を使うこともできる。テクスチャ閾値は、現在のピクチャ用に多値 D Q を使うことをエンコーダが決定した後で、どのマクロブロックが粗い Q P で量子化されることになるか判定するのに使うことができる。

【 0 2 5 1 】

あるいは、エンコーダは、他の基準を使って、特定の領域（例えば、マクロブロック）が粗い Q P を使うべきかどうか判定し、または粗い Q P を使わない。

【 0 2 5 2 】

F . D C シフト量子化のための例示的技術

図 3 3 は、D C シフトエリア内での量子化を、量子化アーチファクトの混入を削減し、または避けるように適合させる複合技術 3 3 0 0 を示すフローチャートである。図 1 0 のエンコーダ 1 0 0 0 などのエンコーダまたは他のツールが、技術 3 3 0 0 を実施する。この複合技術は、上述した様のいくつかを組み合わせた例である。他の技術では、この例を参照して記載した様すべては用いず、または様を異なる順序もしくは代替的なやり方で実施する。

【 0 2 5 3 】

段階 3 3 1 0 で、エンコーダは、例えば、セクション V で記載したように、現在のピクチャ中の 1 つまたは複数の階調傾斜領域を検出する。段階 3 3 2 0 で、エンコーダは、階調傾斜領域（群）内の D C シフトブロックを、例えば D C シフトテーブルを使って検出する。

【 0 2 5 4 】

エンコーダは次いで、ピクチャをどのようにして量子化するか決定する。段階 3 3 3 0 で、エンコーダは、ピクチャに対して二値 D Q を使うかどうか決定する。二値 D Q を使う場合、エンコーダは、D C シフトエリア（段階 3 3 4 0 ）及び他の平滑エリア用のピクチャ Q P より小さい Q P を使う。二値 D Q を使わない場合、段階 3 3 5 0 で、エンコーダは、ピクチャに対して多値 D Q を使うかどうか決定する。多値 D Q を使う場合、エンコーダは、D C シフトエリア用のピクチャ Q P より小さい Q P を使い（段階 3 3 6 0 ）、他の平滑エリアに対しては異なる Q P を使うことができ、高テクスチャマクロブロックが存在する場合は、高テクスチャマクロブロックに対して粗い Q P （ピクチャ Q P より大きいようなものなど）を使う（段階 3 3 7 0 ）。エンコーダが、二値 D Q も多値 D Q も選ばない場

10

20

30

40

50

合、エンコーダは、ピクチャ QP を削減し、削減されたピクチャ QP を DC シフトエリア（段階 3380）ならびに他のエリアに対して使う。あるいは、二値 DQ 及び多値 DQ のコストが両方とも高すぎる場合、エンコーダは、DC シフトエリアに対する QP 削減をスキップする。エンコーダが、量子化方式を選んだら、エンコーダは、段階 3390 でピクチャを圧縮し、処理されるべきピクチャがどれか残っている場合は、次のピクチャを処理する（段階 3395）。

【0255】

以上、様々な実施形態を参照して本発明の原理を説明し示したが、こうした様々な実施形態は、このような原理から逸脱することなく、構成及び細部において変更できることが理解されよう。本明細書において説明したプログラム、処理、または方法は、特にそう示さない限り、特定のどのタイプのコンピューティング環境にも関連せず、限定されないことを理解されたい。様々なタイプの汎用または専用コンピューティング環境が、本明細書において説明した教示とともに使うことができ、こうした教示に従って動作を実施することができる。ソフトウェアで示した実施形態の要素は、ハードウェアでも実装することができ、その逆も可能である。10

【0256】

本発明の原理を適用することができる可能な多くの実施形態を鑑み、本発明として、このようなすべての実施形態が添付の特許請求の範囲及び等価物の範囲及び精神内であると主張する。20

【図面の簡単な説明】

【0257】

【図 1】 8×8 ブロックのサンプルの、ブロックに基づいたイントラフレーム圧縮を示す図である。

【図 2】ビデオエンコーダにおける動き推定を示す図である。

【図 3】ビデオエンコーダにおける 8×8 ブロックの予測残差に対するブロックに基づいた圧縮を示す図である。

【図 4】ビデオエンコーダにおける 8×8 ブロックの予測残差に対するブロックに基づいた圧縮解除を示す図である。

【図 5】スカラー量子化法用の階段 I/O 関数を示すチャートである。

【図 6 A】スカラー量子化法用の分類法及び閾値を示すチャートである。30

【図 6 B】スカラー量子化法用の分類法及び閾値を示すチャートである。

【図 7】DZ + UTQ に対する階段 I/O 関数を示すチャートである。

【図 8 A】DZ + UTQ に対する分類法及び閾値を示すチャートである。

【図 8 B】DZ + UTQ に対する分類法及び閾値を示すチャートである。

【図 9】説明したいいくつかの実施形態を実装することができる適切なコンピューティング環境を示すブロック図である。

【図 10】説明したいいくつかの実施形態を実装することができる、一般化したビデオエンコーダシステムを示すブロック図である。

【図 11】説明したいいくつかの実施形態において用いられるマクロブロック形式を示す図である。40

【図 12】適応ビデオエンコーディング方法を示すフローチャートである。

【図 13】ブロックに対する輝度及び彩度データを用いたピクセル階調の計算を示す図である。

【図 14】図 13 のブロックに対する複数のピクセル階調を示すヒストグラムグラフである。

【図 15】例示的ブロック値特徴抽出フレームワークを示すグラフである。

【図 16】テクスチャ情報に基づく差分量子化を適用する汎用化された技術を示すフローチャートである。

【図 17】時間分析を用いてテクスチャ DQ 決定を行う技術を示すフローチャートである。50

【図18】パーセント閾値及び分離平滑ブロックフィルタリングを使ってテクスチャDQ決定を行う技術を示すフローチャートである。

【図19】高テクスチャピクチャに対するテクスチャレベル閾値を選択的に調整する技術を示すフローチャートである。

【図20】適応テクスチャレベル閾値を判定する例示的疑似コードを示すコード図である。

【図21】階調傾斜領域の2つの例を示す図である。

【図22A】階調傾斜領域、テクスチャ加工された領域、シャープエッジ領域及び平坦領域を有する例示的フレームを示す図である。

【図22B】図22Aの階調傾斜領域内の輪郭化アーチファクトを示す図である。 10

【図22C】図22Bの輪郭化アーチファクトのマクロブロックレベルの細部を示す図である。

【図23】汎用化された領域に基づく階調傾斜検出技術を示すフローチャートである。

【図24】一実装形態による例示的階調傾斜検出装置を示すブロック図である。

【図25】異常な階調傾斜方向を引き起こす可能性があるフィルム粒子を有する階調傾斜領域の4:1ダウンサンプリングを示す図である。

【図26】 16×16 コンパスオペレータ K_H 、 K_V を示す式図である。

【図27】図26のコンパスオペレータを使って領域に対する階調方向を計算する例示的疑似コードを示すコード図である。

【図28】階調傾斜領域に対する一貫性チェックを実施する技術を示すフローチャートである。 20

【図29】バケット投票技術におけるバケットを示す図である。

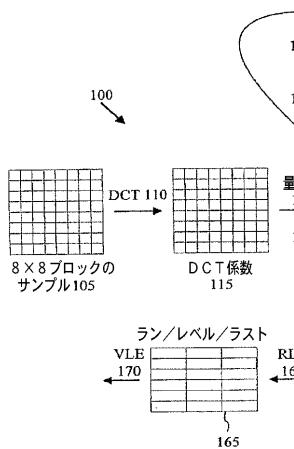
【図30】1つまたは複数の非ゼロAC係数を維持するのを助けるためのマクロブロックQPを選択する例示的技術を示すフローチャートである。

【図31】量子化及び逆量子化後の、階調傾斜領域内の3つの隣接ブロックにおけるDCシフトを示す図である。

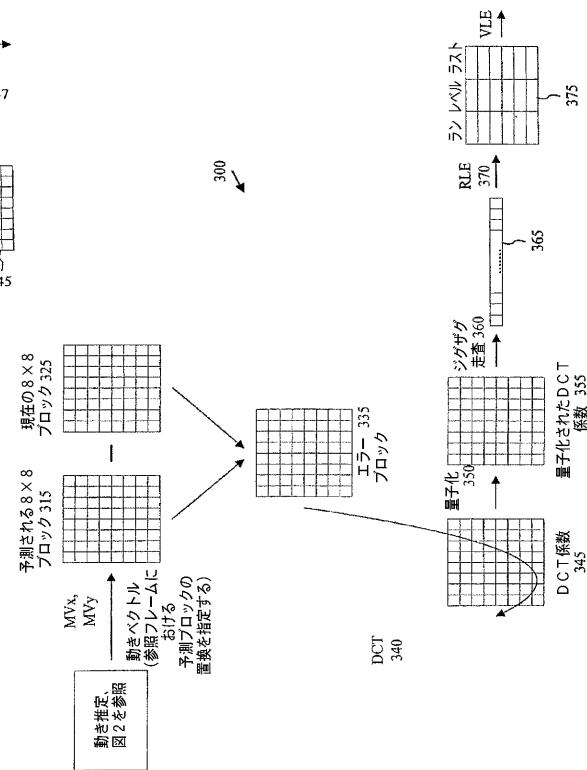
【図32】DCシフトエリア内での輪郭化アーチファクトの混入を削減し、または避けるように量子化を調整する汎用化された技術を示すフローチャートである。

【図33】量子化アーチファクトの混入を削減し、または避けるようにDCシフトエリア内の量子化を適応させる複合技術を示すフローチャートである。 30

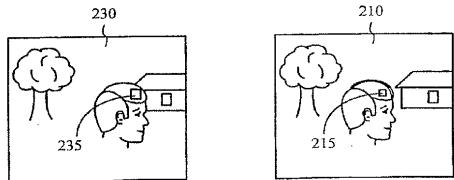
【図1】



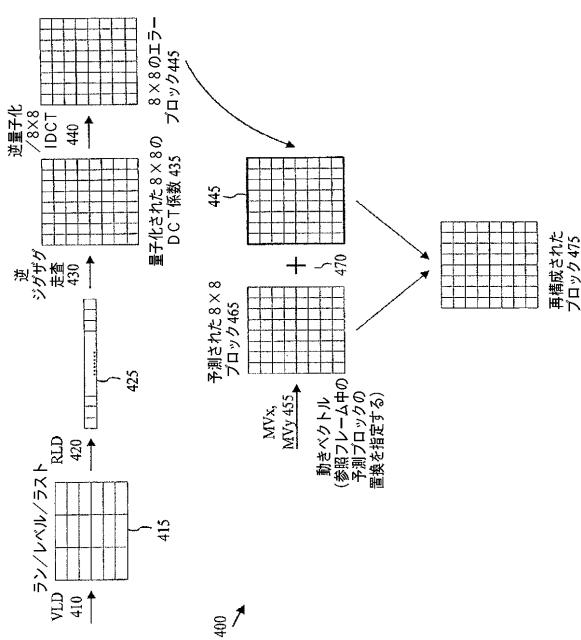
【図3】



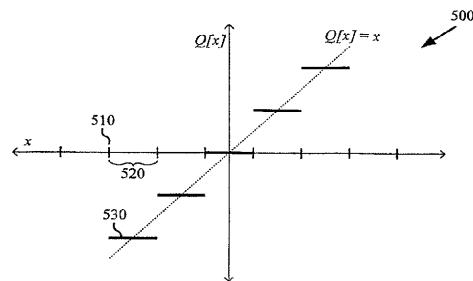
【図2】



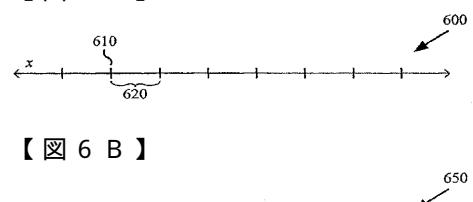
【図4】



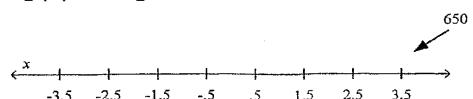
【図5】



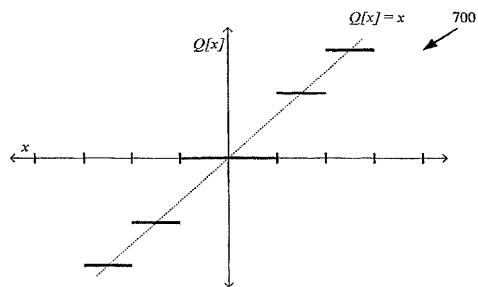
【図6A】



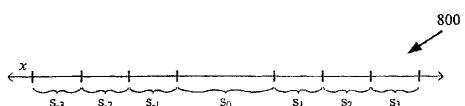
【図6B】



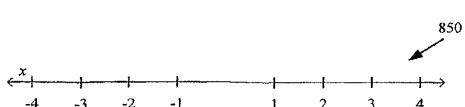
【図7】



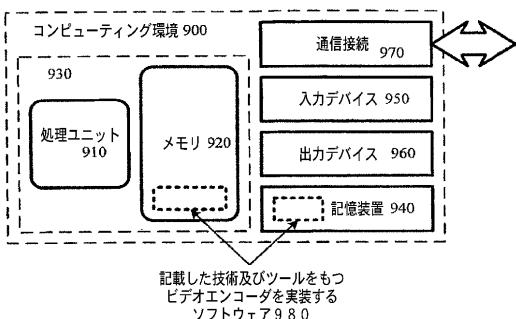
【図8A】



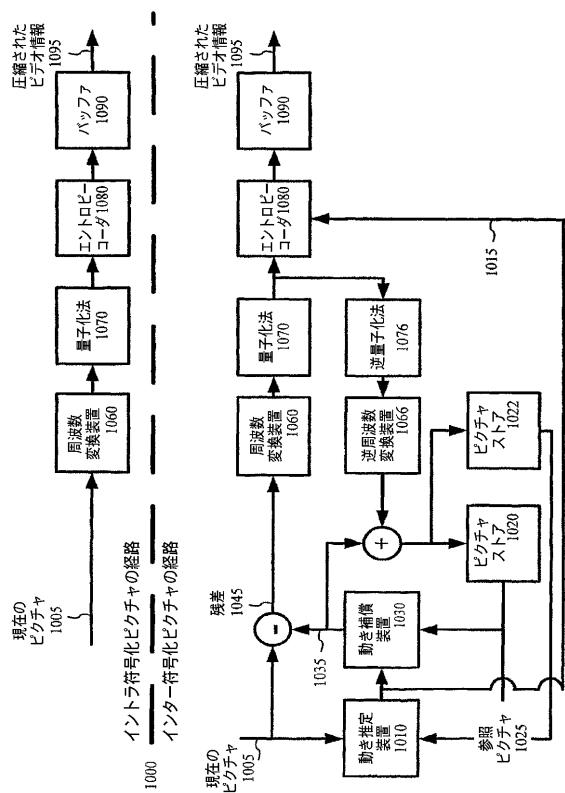
【図8B】



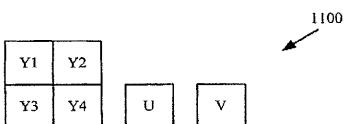
【図9】



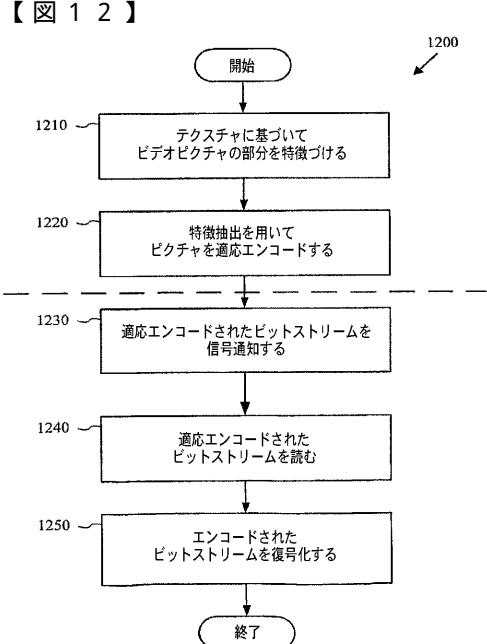
【図10】



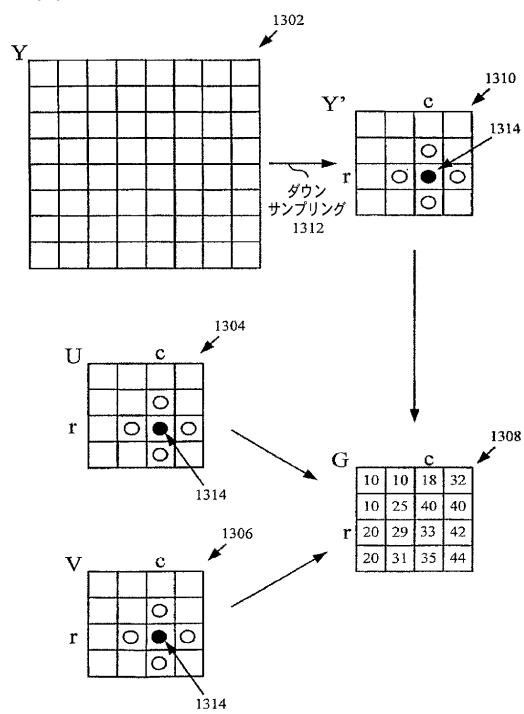
【図11】



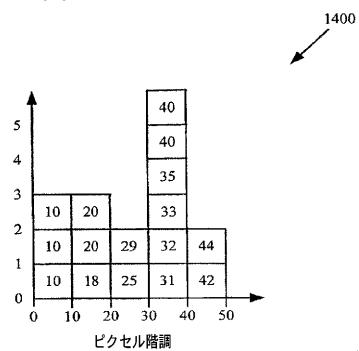
【図12】



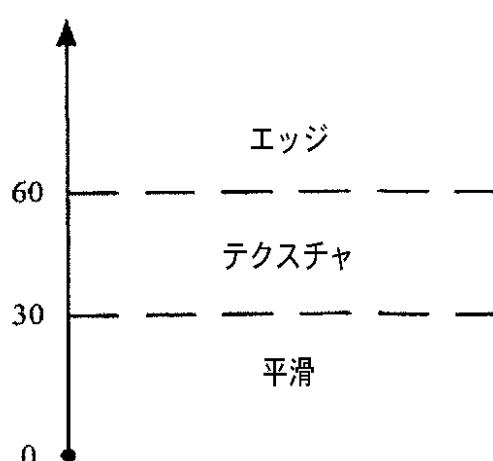
【図13】



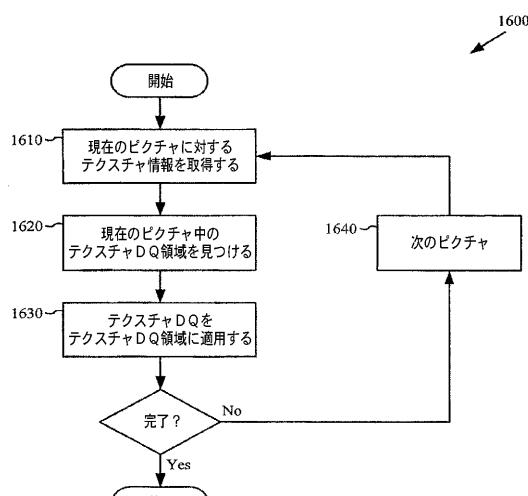
【図14】



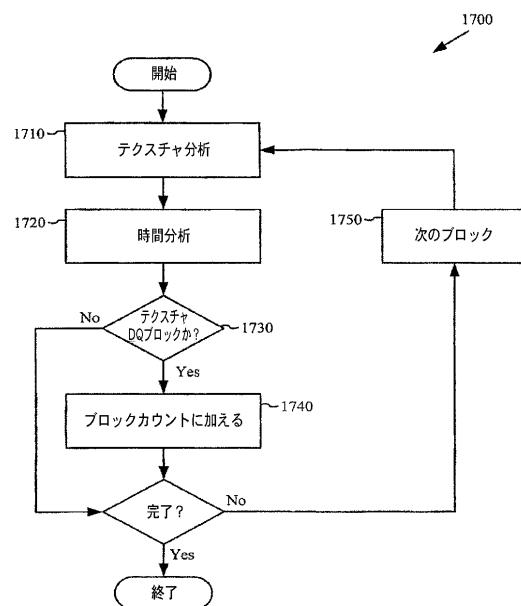
【図15】



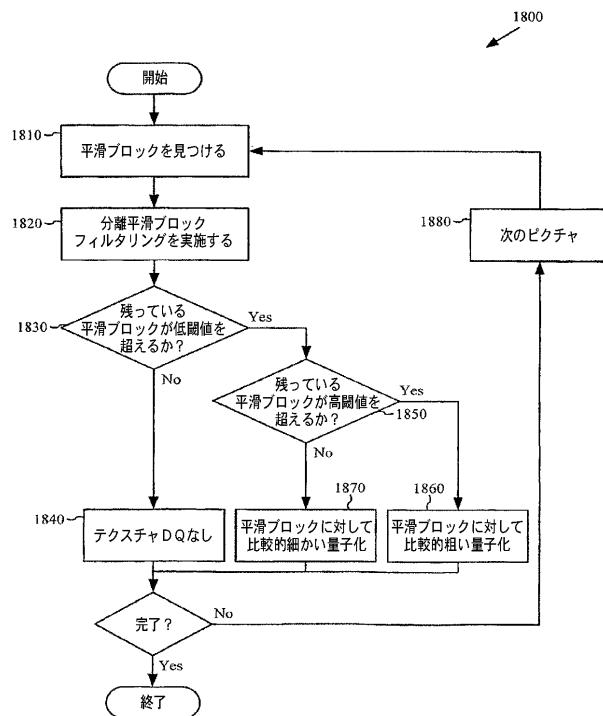
【図16】



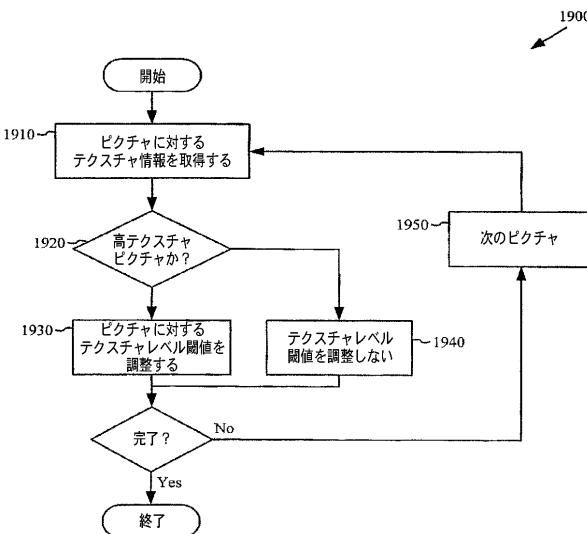
【図17】



【図18】



【図19】



【図20】

```

#define TEXTURELOWTH 30
#define TEXTUREMAX 100
#define MINFLATTH 14
#define MAXFLATTH 28

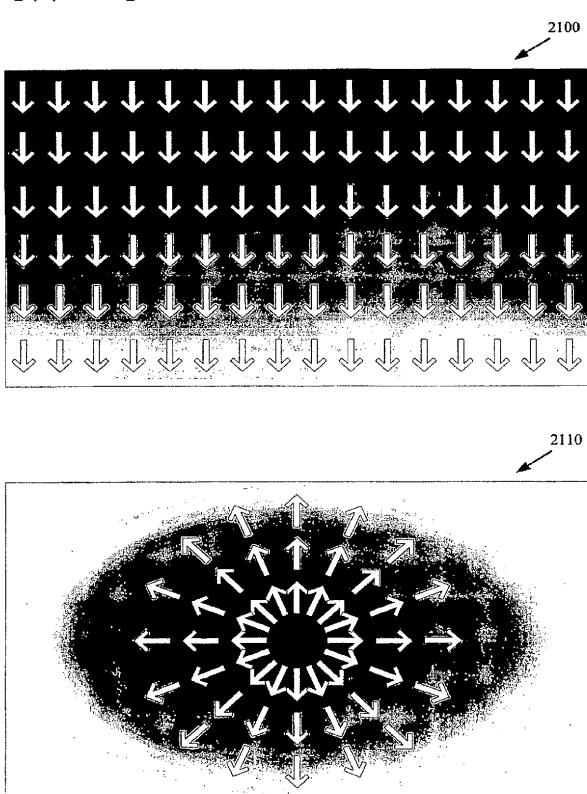
const UInt g_iFlatThTable[TEXTUREMAX-TEXTURELOWTH] = {
    16, 16, 16, 16, 18, 18, 18, 18, 18, 18,
    20, 20, 20, 20, 20, 20, 20, 20, 20,
    22, 22, 22, 22, 22, 22, 22, 22, 22,
    24, 24, 24, 24, 24, 24, 24, 24, 24,
    24, 24, 24, 24, 24, 24, 24, 24, 24,
    26, 26, 26, 26, 26, 26, 26, 26, 26,
    26, 26, 26, 26, 26, 26, 26, 26, 26
};

unsigned integer decideAdaptiveFlatThreshold(unsigned integer g0)
{
    Integer new_adaptive_threshold;
    if(g0 < TEXTURELOWTH) new_adaptive_threshold=MINFLATTH;
    else
        if(g0 >= TEXTUREMAX)
            new_adaptive_threshold=MAXFLATTH;
        else
            new_adaptive_threshold=g_iFlatThTable[g0-TEXTURELOWTH];

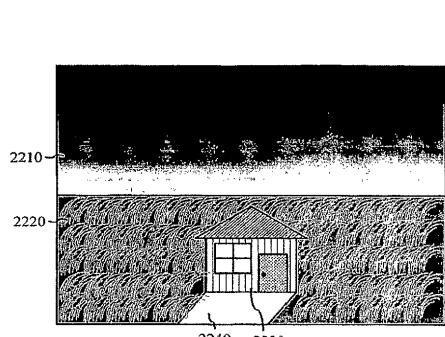
    if(not scene change key frames)
        new_adaptive_threshold = last_adaptive_threshold +
        max(-4, min(4, new_adaptive_threshold - last_adaptive_threshold));

    return new_adaptive_threshold;
}
  
```

【図21】



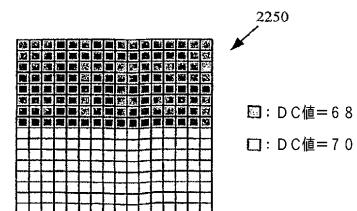
【図 2 2 A】



【図 2 2 B】



【図 2 2 C】



【図 2 5】

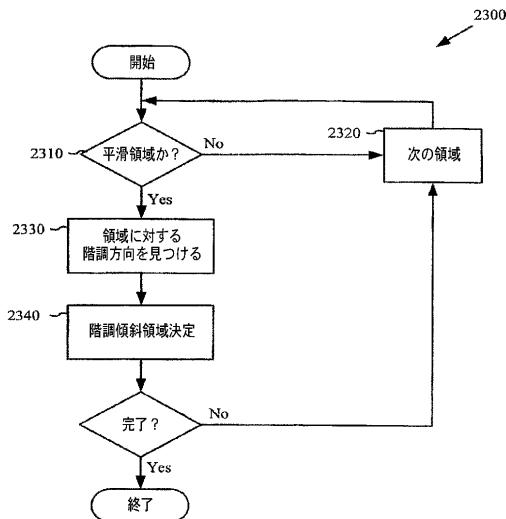
15	15	15	15	15	15	15	15	15
15	15	15	15	15	15	15	15	15
16	16	16	16	16	20	16	16	16
16	16	16	16	21	19	21	16	16
17	17	17	17	17	17	17	17	17
20	19	17	17	17	17	17	17	17
18	21	17	17	17	17	17	17	17
18	18	18	18	18	18	18	18	18

【図 2 6】

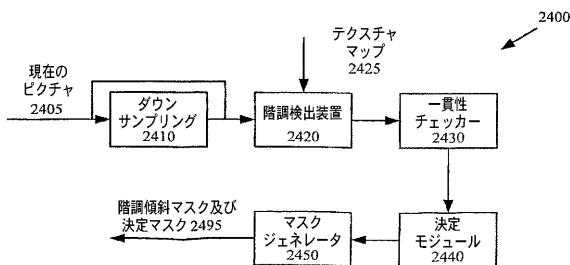
$$K_H = [k_{ij}] \text{ where } 0 \leq i, j \leq 15, k_{ij} = \begin{cases} 1 & \text{if } i < 8 \\ -1 & \text{if } i \geq 8 \end{cases}$$

$$K_V = [k_{ij}] \text{ where } 0 \leq i, j \leq 15, k_{ij} = \begin{cases} 1 & \text{if } j < 8 \\ -1 & \text{if } j \geq 8 \end{cases}$$

【図 2 3】



【図 2 4】



【図 2 7】

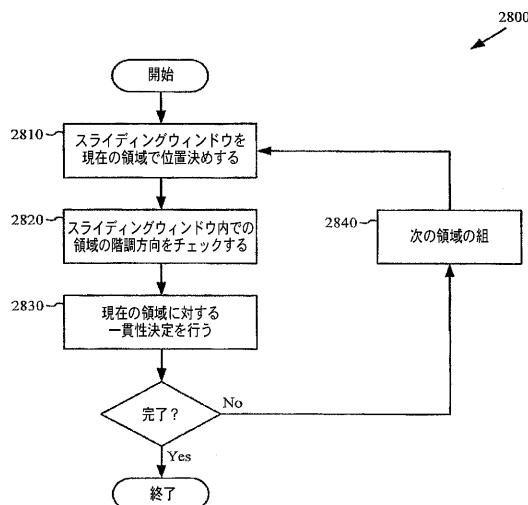
```

ComputeGradientDirection (IN region, OUT gradientDirection)
if (region is NOT smooth) {
    gradientDirection = -2; //texture or edge region
    return;
}

gx = sum of all elements of K_H ⊗ block
gy = sum of all elements of K_V ⊗ block
if (abs(gx) < 32 && abs(GradY) < 32 ) {
    gradientDirection = -1; //flat region
    return;
}
θ = tan⁻¹(gy / gx);
Offset = 0;
if(gx < 0) {
    Offset = 3.14159;
}
Else if(gy < 0) {
    Offset = 2 * 3.14159;
}
θ = θ + Offset;
gradientDirection = ((int)(θ * 2598)) >> 6;
End_ComputeGradientDirection

```

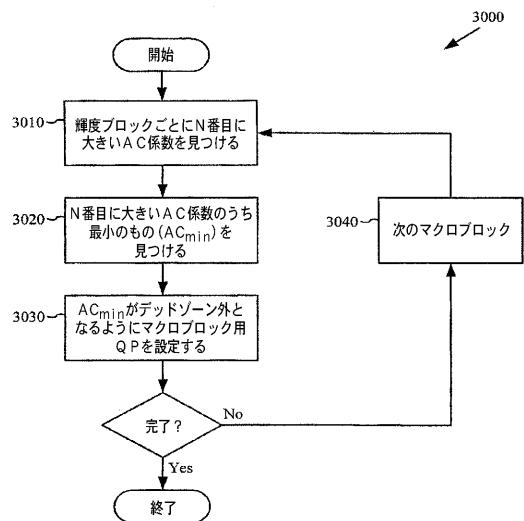
【図28】



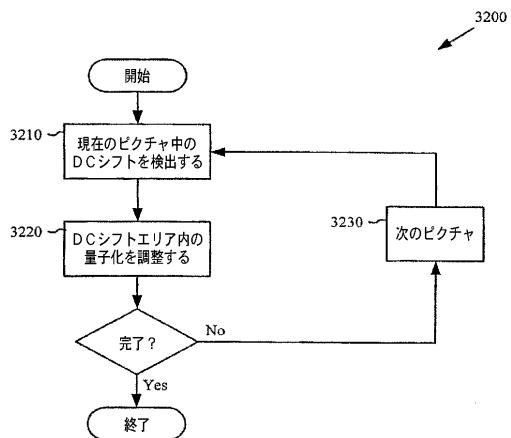
【図29】

	0	1	2	3	4
0	0	5	10	5	0
1	0	6	24	10	0
2	0	5	10	5	0
3	0	0	0	0	0
4	0	2	0	1	0

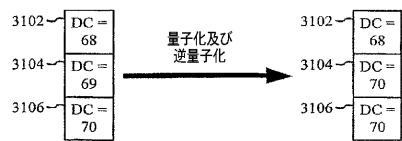
【図30】



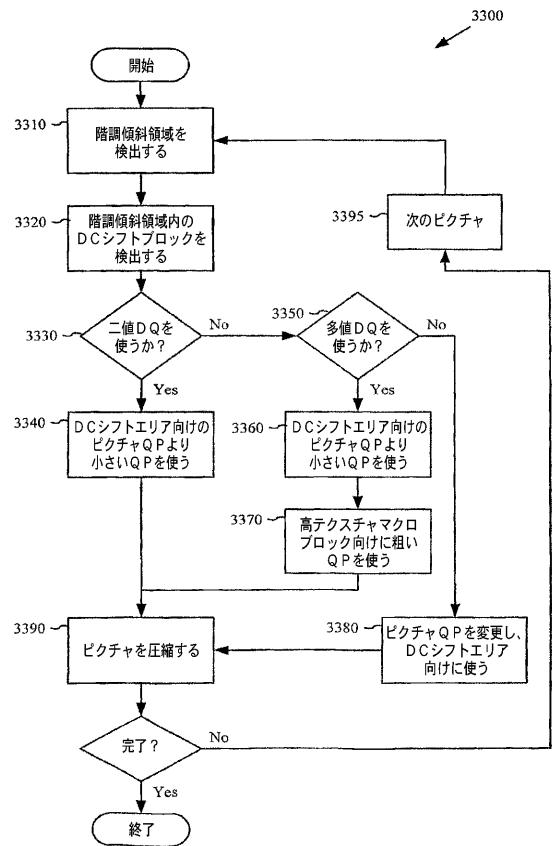
【図32】



【図31】



【図3-3】



フロントページの続き

(72)発明者 チー・ルン・リン

アメリカ合衆国 98052 ワシントン州 レッドモンド ワン マイクロソフト ウェイ マ
イクロソフト コーポレーション インターナショナル パテンツ内

審査官 岩井 健二

(56)参考文献 国際公開第02/080575 (WO, A1)

特開2005-260467 (JP, A)

特開2004-247889 (JP, A)

特開2004-023288 (JP, A)

特開平07-222145 (JP, A)

特開平05-227525 (JP, A)

(58)調査した分野(Int.Cl., DB名)

H04N 7/24 - 7/68