

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号
特許第4988567号
(P4988567)

(45) 発行日 平成24年8月1日 (2012. 8. 1)

(24) 登録日 平成24年5月11日 (2012. 5. 11)

(51) Int. Cl.

F I

HO 4 N 7/26 (2006. 01)

HO 4 N 7/13 Z

HO 3 M 7/30 (2006. 01)

HO 3 M 7/30 Z

請求項の数 17 (全 32 頁)

(21) 出願番号	特願2007-518338 (P2007-518338)	(73) 特許権者	503260918
(86) (22) 出願日	平成17年6月24日 (2005. 6. 24)		アップル インコーポレイテッド
(65) 公表番号	特表2008-504750 (P2008-504750A)		アメリカ合衆国 9 5 0 1 4 カリフォル
(43) 公表日	平成20年2月14日 (2008. 2. 14)		ニア州 クパチーノ インフィニット ル
(86) 国際出願番号	PCT/US2005/022616		ープ 1
(87) 国際公開番号	W02006/004605	(74) 代理人	100076428
(87) 国際公開日	平成18年1月12日 (2006. 1. 12)		弁理士 大塚 康德
審査請求日	平成20年5月8日 (2008. 5. 8)	(74) 代理人	100112508
(31) 優先権主張番号	60/583, 418		弁理士 高柳 司郎
(32) 優先日	平成16年6月27日 (2004. 6. 27)	(74) 代理人	100115071
(33) 優先権主張国	米国 (US)		弁理士 大塚 康弘
(31) 優先権主張番号	60/643, 918	(74) 代理人	100116894
(32) 優先日	平成17年1月9日 (2005. 1. 9)		弁理士 木村 秀二
(33) 優先権主張国	米国 (US)	(74) 代理人	100130409
			弁理士 下山 治

最終頁に続く

(54) 【発明の名称】 マルチパスのビデオ符号化

(57) 【特許請求の範囲】

【請求項 1】

複数の符号化パスを用いて複数のイメージを含んだビデオシーケンスを繰り返し符号化する複数段階の方法であって、

ビデオ符号器による第 1 の複数の符号化パスの第 1 の符号化パスのための第 1 の名目量子化パラメータを特定するステップと、

前記第 1 の複数の符号化パスの前記第 1 の符号化パスにおいて、前記第 1 の名目量子化パラメータに基づいて各イメージのためのイメージ固有の量子化パラメータを生成し、各イメージのための前記イメージ固有の量子化パラメータを用いて各イメージを符号化するステップと、

前記第 1 の複数の符号化パスの第 2 の符号化パスのための第 2 の名目量子化パラメータを、前記第 1 の名目量子化パラメータに基づいて特定するステップと、

前記第 1 の複数の符号化パスの前記第 2 の符号化パスにおいて、前記第 2 の名目量子化パラメータに基づいて各イメージのためのイメージ固有の量子化パラメータを生成し、各イメージのための前記イメージ固有の量子化パラメータを用いて各イメージを符号化するステップと、

を備え、

前記第 1 の複数の符号化パスの前記第 2 の符号化パスにおいて生成される前記イメージ固有の量子化パラメータは、前記第 1 の複数の符号化パスの前記第 1 の符号化パスにおいて生成される前記イメージ固有の量子化パラメータとは異なり、

前記方法は、

前記ビデオ符号器による第2の複数の符号化パスの第1の符号化パスのための第1の基準視覚マスキング強度を特定するステップと、

前記第2の複数の符号化パスの前記第1の符号化パスにおいて、前記第1の複数の符号化パスの間に特定された特定の名目量子化パラメータと前記第1の基準視覚マスキング強度とに基づいて各イメージのためのイメージ固有の量子化パラメータを生成し、各イメージのための前記イメージ固有の量子化パラメータを用いて各イメージを符号化するステップと、

前記第2の複数の符号化パスの第2の符号化パスのための第2の基準視覚マスキング強度を、前記第1の基準視覚マスキング強度に基づいて特定するステップと、

前記第2の複数の符号化パスの前記第2の符号化パスにおいて、前記特定の名目量子化パラメータと前記第2の基準視覚マスキング強度とに基づいて各イメージのためのイメージ固有の量子化パラメータを生成し、各イメージのための前記イメージ固有の量子化パラメータを用いて各イメージを符号化するステップと、

をさらに備え、

前記第2の複数の符号化パスの前記第2の符号化パスにおいて生成される前記イメージ固有の量子化パラメータは、前記第2の複数の符号化パスの前記第1の符号化パスにおいて生成される前記イメージ固有の量子化パラメータとは異なることを特徴とする方法。

【請求項2】

前記複数のイメージに関して許容できる符号化が識別されると前記第1の複数の符号化パスを停止するステップをさらに備えることを特徴とする請求項1に記載の方法。

【請求項3】

前記複数のイメージに関して許容できる符号化は、前記ビデオシーケンスの符号化に関する目標ビットレートに関して特定の範囲内に収まる、前記複数のイメージの符号化であることを特徴とする請求項2に記載の方法。

【請求項4】

複数の符号化パスを用いて複数のイメージを含んだビデオシーケンスを繰り返し符号化するためのコンピュータプログラムを記憶しているコンピュータ読み取り可能な記憶媒体であって、前記コンピュータプログラムは、以下に列挙する命令セットを含むことを特徴とするコンピュータ読み取り可能な記憶媒体。

ビデオ符号器による第1の複数の符号化パスの第1の符号化パスのための第1の名目量子化パラメータを特定すること。

前記第1の複数の符号化パスの前記第1の符号化パスにおいて、前記第1の名目量子化パラメータに基づいて各イメージのためのイメージ固有の量子化パラメータを生成し、各イメージのための前記イメージ固有の量子化パラメータを用いて各イメージを符号化すること。

前記第1の複数の符号化パスの第2の符号化パスのための第2の名目量子化パラメータを、前記第1の名目量子化パラメータに基づいて特定すること。

前記第1の複数の符号化パスの前記第2の符号化パスにおいて、前記第2の名目量子化パラメータに基づいて各イメージのためのイメージ固有の量子化パラメータを生成し、各イメージのための前記イメージ固有の量子化パラメータを用いて各イメージを符号化すること。

ここで、前記第1の複数の符号化パスの前記第2の符号化パスにおいて生成される前記イメージ固有の量子化パラメータは、前記第1の複数の符号化パスの前記第1の符号化パスにおいて生成される前記イメージ固有の量子化パラメータとは異なる。

前記ビデオ符号器による第2の複数の符号化パスの第1の符号化パスのための第1の基準視覚マスキング強度を特定すること。

前記第2の複数の符号化パスの前記第1の符号化パスにおいて、前記第1の複数の符号化パスの間に特定された特定の名目量子化パラメータと前記第1の基準視覚マスキング強度とに基づいて各イメージのためのイメージ固有の量子化パラメータを生成し、各イメー

10

20

30

40

50

ジのための前記イメージ固有の量子化パラメータを用いて各イメージを符号化すること。

前記第 2 の複数の符号化パスの第 2 の符号化パスのための第 2 の基準視覚マスキング強度を、前記第 1 の基準視覚マスキング強度に基づいて特定すること。

前記第 2 の複数の符号化パスの前記第 2 の符号化パスにおいて、前記特定の名目量子化パラメータと前記第 2 の基準視覚マスキング強度とに基づいて各イメージのためのイメージ固有の量子化パラメータを生成し、各イメージのための前記イメージ固有の量子化パラメータを用いて各イメージを符号化すること。

ここで、前記第 2 の複数の符号化パスの前記第 2 の符号化パスにおいて生成される前記イメージ固有の量子化パラメータは、前記第 2 の複数の符号化パスの前記第 1 の符号化パスにおいて生成される前記イメージ固有の量子化パラメータとは異なる。

10

【請求項 5】

前記第 2 の名目量子化パラメータは、前記第 1 の複数の符号化パスの前記第 1 の符号化パスの符号化結果にさらに基づくことを特徴とする請求項 4 に記載のコンピュータ読み取り可能な記憶媒体。

【請求項 6】

前記複数のイメージに関して許容できる符号化が識別されると前記第 2 の複数の符号化パスを停止することをさらに含むことを特徴とする請求項 4 に記載のコンピュータ読み取り可能な記憶媒体。

【請求項 7】

前記複数のイメージに関して許容できる符号化は、前記ビデオシーケンスの符号化に関する目標ビットレートに関して特定の範囲内に収まる、前記複数のイメージの符号化であることを特徴とする請求項 6 に記載のコンピュータ読み取り可能な記憶媒体。

20

【請求項 8】

各イメージのための前記イメージ固有の量子化パラメータは、前記イメージ内の平均ピクセル輝度に基づいて導出されることを特徴とする請求項 1 に記載の方法。

【請求項 9】

各イメージのための前記イメージ固有の量子化パラメータは、前記イメージ内の平均ピクセル輝度のパワー関数に基づいて導出されることを特徴とする請求項 1 に記載の方法。

【請求項 10】

イメージは複数のピクセルセットを含み、

30

各イメージのための前記イメージ固有の量子化パラメータは、イメージ内の全ピクセルセットに亘るピクセル値の絶対差分の総和 (S A D) の平均のパワー関数に基づいて導出される

ことを特徴とする請求項 1 に記載の方法。

【請求項 11】

各イメージのための前記イメージ固有の量子化パラメータは、イメージ間の動きに起因する耐えられ得る歪みの量を示す時間的屬性に基づいて導出されることを特徴とする請求項 1 に記載の方法。

【請求項 12】

各イメージのための前記イメージ固有の量子化パラメータは、前記イメージ内に定義された複数のピクセル領域に関する動き補償された誤差信号の絶対値の総和に基づいて導出されることを特徴とする請求項 1 に記載の方法。

40

【請求項 13】

各イメージのための前記イメージ固有の量子化パラメータは、前記イメージのピクセル値の絶対差分の総和 (S A D) の平均と、前記イメージよりも時間的に前のイメージセットのピクセル値の絶対差分の総和 (S A D) の平均と、前記イメージよりも時間的に後のイメージセットのピクセル値の絶対差分の総和 (S A D) の平均と、に基づいて導出されることを特徴とする請求項 1 に記載の方法。

【請求項 14】

各イメージのための前記イメージ固有の量子化パラメータは、前記イメージ内の平均ピ

50

クセル輝度に基づいて導出されることを特徴とする請求項 4 に記載のコンピュータ読み取り可能な記憶媒体。

【請求項 15】

各イメージのための前記イメージ固有の量子化パラメータは、前記イメージ内の平均ピクセル輝度のパワー関数に基づいて導出されることを特徴とする請求項 4 に記載のコンピュータ読み取り可能な記憶媒体。

【請求項 16】

各イメージに関する前記イメージ固有の量子化パラメータは、当該イメージの全体よりも小さい、当該イメージの一部分のためのものであることを特徴とする請求項 1 に記載の方法。

10

【請求項 17】

各イメージに関する前記イメージ固有の量子化パラメータは、当該イメージの全体のためのものであることを特徴とする請求項 1 に記載の方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、マルチパスのビデオ符号化に関する。

【背景技術】

【0002】

ビデオ符号器は、様々な符号化方式を使用することにより、ビデオイメージ（動画像）（例えば、ビデオフレーム）のシーケンスを符号化する。ビデオ符号化方式は、通常、ビデオフレーム、またはビデオフレームの諸部分（例えば、ビデオフレーム中のピクセルの集合（ピクセルセット））をフレーム内、またはフレーム間に関して符号化する。フレーム内符号化されたフレームまたはピクセルセットとは、他のフレーム、または他のフレームの中のピクセルセットとは独立に符号化されたフレームまたはピクセルセットである。フレーム間符号化されたフレームまたはピクセルセットとは、1 以上の他のフレーム、または 1 以上の他のフレームの中の 1 以上のピクセルセットを参照して符号化されたフレームまたはピクセルセットである。

20

【0003】

ビデオフレームを圧縮する際、一部の符号器は、符号化されるべきビデオフレームまたはビデオフレームの集合（ビデオフレームセット）のための「ビット予算(bit budget)」を提供する、「レートコントローラ」を実施する。ビット予算は、ビデオフレームまたはビデオフレームセットを符号化するのに割り当てられたビットの数を指定する。ビット予算を効率的に割り当てることにより、レートコントローラは、ある制約（例えば、目標ビットレート、その他）に鑑みて、最高品質の圧縮されたビデオストリームを生成しようと試みる。

30

【0004】

今日まで、様々な単一パスのレートコントローラおよびマルチパスのレートコントローラが、提案されてきた。単一パスのレートコントローラは、一連のビデオイメージを 1 回のパスで符号化する符号化方式のためのビット予算を提供するのに対して、マルチパスのレートコントローラは、一連のビデオイメージを複数回のパスで符号化する符号化方式のためのビット予算を提供する。

40

【0005】

単一パスのレートコントローラは、リアルタイムの符号化状況において役立つ。他方、マルチパスのレートコントローラは、一連の制約に基づく特定のビットレートに関する符号化を最適化する。符号化におけるビットレートを制御する際に、フレームまたはフレーム内のピクセルセットの空間的複雑性、または時間的複雑性を考慮に入れる、レートコントローラは、現在、多くはない。また、ほとんどのマルチパスのレートコントローラは、所望されるビットレートに鑑みて、フレームおよびフレーム内のピクセルセットのうちの少なくとも一方に関して、最適な量子化パラメータを使用する符号化ソリューションを求

50

めて、解空間を十分に探索することもしない。

【発明の開示】

【発明が解決しようとする課題】

【0006】

したがって、当技術分野において、新規な技術を使用して、ビデオイメージの集合（ビデオイメージセット）を符号化するためのビットレートを制御しながら、ビデオイメージおよびビデオイメージの諸部分のうちの少なくとも一方の空間的複雑性、または時間的複雑性を考慮する、レートコントローラの必要性が存在する。また、当技術分野において、符号化ソリューションを十分に調べて、ビデオイメージおよびビデオイメージの諸部分のうちの少なくとも一方に関して、最適な量子化パラメータセットを使用する符号化ソリューションを特定する、マルチパスのレートコントローラの必要性も存在する。

10

【課題を解決するための手段】

【0007】

本発明の一部の実施形態は、いくつかのイメージ（例えば、ビデオシーケンスのいくつかのフレーム）を符号化するマルチパス符号化方法を提供する。本方法は、それらのイメージを符号化する符号化操作を繰り返し実行する。符号化操作は、本方法がイメージに関する量子化パラメータを計算するために使用する、名目量子化パラメータ(nominal QP)に基づく。符号化操作の数回の異なる繰り返し中、本方法は、いくつかの異なる名目量子化パラメータを使用する。本方法は、終了の基準に達すると（例えば、本方法が、許容できるイメージの符号化を特定すると）、繰り返しを停止する。

20

【0008】

本発明の一部の実施形態は、ビデオシーケンスを符号化するための方法を提供する。本方法は、ビデオの中の第1のイメージの複雑性を定量化する第1の属性を特定する。また、本方法は、特定された第1の属性に基づき、第1のイメージを符号化するための量子化パラメータも特定する。次に、本方法は、特定された量子化パラメータに基づき、第1のイメージを符号化する。一部の実施形態では、この方法は、ビデオの中のいくつかのイメージに関して、以上3つの操作を実行する。

【0009】

本発明の一部の実施形態は、ビデオイメージ、およびビデオイメージの諸部分のうちの少なくとも一方の「視覚マスキング」属性に基づき、ビデオイメージのシーケンスを符号化する。イメージ、またはイメージの一部の視覚マスキングは、そのイメージまたはイメージ部分の中で、どれだけの符号化アーチファクトが許容されることが可能であるかの指標である。イメージまたはイメージ部分の視覚マスキング属性を表現するのに、一部の実施形態は、そのイメージまたはイメージ部分の輝度エネルギーを定量化する視覚マスキング強度を計算する。一部の実施形態では、輝度エネルギーは、イメージまたはイメージ部分の平均ルーマ（Luma, 輝度）エネルギーまたは平均ピクセルエネルギーの関数として測定される。

30

【0010】

輝度エネルギーの代わりに、または輝度エネルギーと併せて、イメージまたはイメージ部分の視覚マスキング強度は、イメージまたはイメージ部分の活動エネルギーも定量化することが可能である。活動エネルギーは、イメージまたはイメージ部分の複雑性を表現する。一部の実施形態では、活動エネルギーは、イメージまたはイメージ部分の空間的複雑性を定量化する空間的成分、および、イメージ間の動きに起因する、許容されマスキングされることが可能な歪みの量を定量化する動き成分のうちの少なくとも一方を含む。

40

【0011】

本発明の一部の実施形態は、ビデオシーケンスを符号化するための方法を提供する。本方法は、ビデオの中の第1のイメージの視覚マスキング属性を特定する。また、本方法は、特定された視覚マスキング属性に基づき、第1のイメージを符号化するための量子化パラメータも特定する。次に、本方法は、特定された量子化パラメータに基づき、第1のイメージを符号化する。

50

【 0 0 1 2 】

本発明の新規な諸特徴は、添付の特許請求の範囲で示される。しかし、説明のため、本発明のいくつかの実施形態が、以下の図で示される。

【発明を実施するための最良の形態】

【 0 0 1 3 】

本発明の以下の詳細な説明では、本発明の多数の詳細、実施例、および実施形態が示され、説明される。しかし、本発明は、示される諸実施形態に限定されないこと、ならびに本発明は、説明される特定の詳細および実施例の一部を伴うことなしに実施されてもよいことが、当業者には明瞭かつ明白であろう。

【 0 0 1 4 】

10

[I . 定義]

このセクションは、本明細書で使用されるいくつかの記号に関する定義を与える。

【 0 0 1 5 】

R_T は、フレームシーケンスの符号化に対して所望されるビットレートである、目標ビットレートを表す。通常、このビットレートは、毎秒のビット数単位 (ビット / 秒) で表現され、所望される最終ファイルサイズ、シーケンスの中のフレーム数、およびフレームレートから計算される。

【 0 0 1 6 】

R_p は、パス p の終了時における、符号化されたビットストリームのビットレートを表す。

20

【 0 0 1 7 】

E_p は、パス p の終了時における、ビットレートの誤差のパーセンテージを表す。一部のケースでは、そのパーセンテージは、

【 0 0 1 8 】

【数 1】

$$100 \times \frac{(R_T - R_p)}{R_T}.$$

【 0 0 1 9 】

30

として計算される。

【 0 0 2 0 】

は、最終ビットレートの許容誤差を表す。

【 0 0 2 1 】

c は、第 1 の Q P 探索段階に関するビットレートの許容誤差を表す。

【 0 0 2 2 】

Q P は、量子化パラメータを表す。

【 0 0 2 3 】

$Q P_{N o m (p)}$ は、フレームシーケンスに関する、パス p での符号化において使用される名目量子化パラメータ (nominal QP) を表す。 $Q P_{N o m (p)}$ の値は、目標ビットレートに達するように、第 1 の Q P 調整段階において本発明のマルチパス符号器によって調整される。

40

【 0 0 2 4 】

$M Q P_p (k)$ は、パス p におけるフレーム k に関する量子化パラメータ (Q P) である、マスキングされたフレーム Q P を表す。一部の実施形態は、名目 Q P、およびフレームレベルでの視覚マスキングを使用することにより、この値を計算する。

【 0 0 2 5 】

$M Q P_{M B (p)} (k, m)$ は、フレーム k およびパス p における個別のマクロブロック (マクロブロックインデックス m を有する) に関する量子化パラメータ (Q P) である、マスキングされたマクロブロック Q P を表す。一部の実施形態は、 $M Q P_p (k)$ およ

50

びマクロブロックレベルでの視覚マスキングを使用することにより、 $MQP_{MB}(p)(k, m)$ を計算する。

【0026】

$F(k)$ は、フレーム k に関するマスキング強度と呼ばれる値を表す。マスキング強度 $F(k)$ は、フレームに関する複雑性の尺度であり、一部の実施形態では、この値は、符号化アーチファクト／雑音が、どれだけ目に見えるように現れるかを判定するのに使用され、フレーム k の $MQP_p(k)$ を計算するのに使用される。

【0027】

$R(p)$ は、パス p における基準マスキング強度を表す。基準マスキング強度は、フレーム k の $MQP_p(k)$ を計算するのに使用され、目標ビットレートを達成するために、第2の段階において本発明のマルチパス符号器によって調整される。

10

【0028】

$MB(k, m)$ は、フレーム k の中のインデックス m を有するマクロブロックに関するマスキング強度を表す。マスキング強度 $MB(k, m)$ は、マクロブロックに関する複雑性の尺度であり、一部の実施形態では、符号化アーチファクト／雑音が、どれだけ目に見えるように現れるかを判定するのに使用され、 $MQP_{MB}(p)(k, m)$ を計算するのに使用される。

【0029】

$AMQP_p$ は、パス p におけるフレーム群にわたる平均のマスキングされた QP を表す。一部の実施形態では、この値は、パス p におけるすべてのフレームにわたる平均の $MQP_p(k)$ として計算される。

20

【0030】

[II. 概要]

本発明の一部の実施形態は、所与のビットレートでフレームシーケンスを符号化することに関して、最良の視覚的品質を実現する符号化方法を提供する。一部の実施形態では、この方法は、量子化パラメータ QP をすべてのマクロブロックに割り当てる視覚マスキングプロセスを使用する。この割り当ては、イメージまたはビデオフレームの中のより明るい領域内、または空間的に複雑な領域内の符号化アーチファクト／雑音が、より暗い領域内、または均一の領域内におけるほどは、目に見えないという認識に基づく。

【0031】

30

一部の実施形態では、この視覚マスキングプロセスは、本発明のマルチパス符号化プロセス（マルチパスの符号化プロセス）の一環として実行される。この符号化プロセスは、最終的な符号化されたビットストリームが、目標ビットレートに達するようにするために、名目量子化パラメータを調整し、基準マスキング強度パラメータ R を介して、視覚マスキングプロセスを制御する。後段でさらに説明されるとおり、名目量子化パラメータを調整すること、およびマスキングアルゴリズムを制御することにより、各ピクチャ（すなわち、通常のビデオ符号化方式においては、各フレーム）に関する QP 値、および各ピクチャ内の各マクロブロックが調整される。

【0032】

一部の実施形態では、マルチパス符号化プロセスは、シーケンス全体に関する名目 QP および R を全体的に調整する。他の諸実施形態では、このプロセスは、ビデオシーケンスをセグメントに分割し、各セグメントに関して、名目 QP および R が調整される。後段の説明は、マルチパス符号化プロセスが使用されるフレームシーケンスについて述べる。そのシーケンスには、一部の実施形態では、シーケンス全体が含まれるのに対して、他の諸実施形態では、あるシーケンスのあるセグメントだけが含まれることが、当業者には認識されよう。

40

【0033】

一部の実施形態では、本方法は、3つの符号化段階を有する。これら3つの段階は、(1)パス0において実行される初期分析段階、(2)パス1乃至パス N_1 において実行される第1の探索段階、および(3)パス $N_1 + 1$ 乃至 $N_1 + N_2$ において実行される第2

50

の探索段階である。

【 0 0 3 4 】

初期分析段階において（すなわち、パス 0 中に）、本方法は、名目 $Q P$ （符号化のパス 1 において使用されるべき $Q P_{N o m (1)}$ ）の初期値を特定する。また、初期分析段階中、本方法は、第 1 の探索段階におけるすべてのパスにおいて使用される、基準マスキング強度 R の値も特定する。

【 0 0 3 5 】

第 1 の探索段階で、本方法は、符号化プロセスの N_1 回の繰り返し（すなわち、 N_1 回のパス）を実行する。各パス p 中に各フレーム k に関して、プロセスは、特定の量子化パラメータ $M Q P_p (k)$ 、ならびにフレーム k 内の個々のマクロブロック m に関する特定の量子化パラメータ $M Q P_{M B (p)} (k , m)$ を使用することによってフレームを符号化する。ただし、 $M Q P_{M B (p)} (k , m)$ は、 $M Q P_p (k)$ を使用して計算される。

10

【 0 0 3 6 】

第 1 の探索段階において、量子化パラメータ $M Q P_p (k)$ は、パスとパスの間で変化する名目量子化パラメータ $Q P_{N o m (p)}$ から導出されるので、パスとパスの間で変化する。つまり、第 1 の探索段階中、各パス p の終りに、プロセスは、パス $p + 1$ に関する名目 $Q P_{N o m (p + 1)}$ を計算する。一部の実施形態では、名目 $Q P_{N o m (p + 1)}$ は、先行するパスからの名目 $Q P$ 値およびビットレート誤差に基づく。他の諸実施形態では、名目 $Q P_{N o m (p + 1)}$ 値は、第 2 の探索段階において各パスの終りに、異なる形で計算される。

20

【 0 0 3 7 】

第 2 の探索段階で、本方法は、符号化プロセスの N_2 回の繰り返し（すなわち、 N_2 回のパス）を実行する。第 1 の探索段階の場合と同様に、プロセスは、特定の量子化パラメータ $M Q P_p (k)$ 、ならびにフレーム k 内の個々のマクロブロック m に関する特定の量子化パラメータ $M Q P_{M B (p)} (k , m)$ を使用することにより、各パス p 中に各フレーム k を符号化する。ただし、 $M Q P_{M B (p)} (k , m)$ は、 $M Q P_p (k)$ から導出される。

【 0 0 3 8 】

やはり、第 1 の探索段階の場合と同様に、量子化パラメータ $M Q P_p (k)$ は、パスとパスの間に変化する。しかし、第 2 の探索段階中、このパラメータは、パスとパスの間に変化する基準マスキング強度 $R (p)$ を使用して計算されるという理由により変化する。一部の実施形態では、基準マスキング強度 $R (p)$ は、先行するパスからのビットレートの誤差、および R の値に基づいて計算される。他の諸実施形態では、この基準マスキング強度は、第 2 の探索段階において各プロセスの終りに、異なる値となるように計算される。

30

【 0 0 3 9 】

マルチパス符号化プロセスは、視覚マスキングプロセスに関連して説明されるが、符号器は、これら両方のプロセスを一緒に使用しなくてもよいことが、当業者には認識されよう。例えば、一部の実施形態では、マルチパス符号化プロセスは、 R を無視し、前述した第 2 の探索段階を省くことにより、視覚マスキングなしに、所与の目標ビットレート近傍のビットストリームを符号化するのに使用される。

40

【 0 0 4 0 】

視覚マスキングおよびマルチパス符号化プロセスを、本出願のセクション I I I およびセクション I V においてさらに説明する。

【 0 0 4 1 】

[I I I . 視覚マスキング]

名目量子化パラメータが与えられると、視覚マスキングプロセスはまず、基準マスキング強度 (R) およびフレームのマスキング強度 (F) を使用して、各フレームに関するマスキングされたフレーム量子化パラメータ ($M Q P$) を計算する。次に、このプロセ

50

スは、フレームレベルおよびマクロブロックレベルのマスキング強度 (F および M_B) に基づき、各マクロブロックに関するマスキングされたマクロブロック量子化パラメータ (MQP_{M_B}) を計算する。視覚マスキングプロセスが、マルチパス符号化プロセスにおいて使用される場合、一部の実施形態における基準マスキング強度 (R) は、前述し、後段でさらに説明するとおり、第1の符号化パス中に特定される。

【0042】

< A . フレームレベルのマスキング強度を計算すること >

1 . 第1のアプローチ

フレームレベルのマスキング強度 $F(k)$ を計算するのに、一部の実施形態は、以下の数式 (A) を使用する。すなわち、

$$F(k) = C * \text{power}(E * \text{avgFrameLuma}(k), \quad) * \text{power}(D * \text{avgFrameSAD}(k), \quad F), (A)$$

ただし、

- $\text{avgFrameLuma}(k)$ は、 b が、1以上の整数である (例えば、 $b = 1$ または $b = 4$)、 $b \times b$ の領域を使用して計算される、フレーム k 内の平均ピクセル輝度であり、

- $\text{avgFrameSAD}(k)$ は、フレーム k 内のすべてのマクロブロックにわたる $MbSAD(k, m)$ の平均であり、

- $MbSAD(k, m)$ は、インデックス m を有するマクロブロック内のすべての 4×4 ブロックに関する関数、 $\text{Calc4x4MeanRemovedSAD}(4 \times 4_block_pixel_values)$ によって与えられる値の合計であり、

- F 、 C 、 D 、および E は、定数であり、かつ (或いは)、局所的な統計に適合されており、

かつ、

- $\text{power}(a, b)$ は、 a^b を意味する。

【0043】

関数 $\text{Calc4x4MeanRemovedSAD}$ に関する擬似コードは、以下のとおりである。すなわち、

```
Calc4x4MeanRemovedSAD(4x4_block_pixel_values)
```

```
{
```

```
    所与の  $4 \times 4$  ブロック内のピクセル値の平均値を計算する；
```

```
    ピクセル値から平均値を引き、絶対値を計算する；
```

```
    1つ前の工程で得られた絶対値を合計する；
```

```
    合計を返す(return the sum;)；
```

```
}
```

【0044】

2 . 第2のアプローチ

他の諸実施形態は、フレームレベルのマスキング強度を異なる形で計算する。例えば、前述した数式 (A) は、フレームのマスキング強度を基本的に以下のように計算する。すなわち、

$$F(k) = C * \text{power}(E * \text{Brightness_Attribute}, \text{exponent0}) * \text{power}(\text{scalar} * \text{Spatial_Activity_Attribute}, \text{exponent1})$$

である。

【0045】

数式 (A) では、フレームの $\text{Brightness_Attribute}$ は、 $\text{avgFrameLuma}(k)$ と等しくなり、 $\text{Spatial_Activity_Attribute}$ は、フレーム内のすべてのマクロブロックにわたる平均マクロブロック $SAD(MbSAD(k, m))$ 値である $\text{avgFrameSAD}(k)$ と等しくなる。ただし、平均マクロブロック SAD は、マクロブロック内のすべての 4×4 ブロックに関する ($\text{Calc4x4MeanRemovedSAD}$ によって与えられる) 平均隔たり 4×4 ピクセル偏差 (mean removed 4×4 pixel variation) の絶対値の合計に等しい。この Spat i

10

20

30

40

50

`al_Activity_Attribute`は、符号化されているフレーム内のピクセル領域内における空間的イノベーション（空間的变化）の量を測定する。

【0046】

他の諸実施形態は、いくつかの連続するフレームにわたるピクセル領域内の時間的イノベーション（時間的变化）の量を含むように活動性(activity)の尺度を拡張する。具体的には、それらの実施形態は、以下のとおりフレームのマスキング強度を計算する。すなわち、

$$F(k) = C \cdot \text{power}(E \cdot \text{Brightness_Attribute}, \text{exponent0}) \cdot \text{power}(\text{scalar} \cdot \text{Activity_Attribute}, \text{exponent1}) \quad (B)$$

である。

10

【0047】

この数式では、`Activity_Attribute`が、以下の数式(C)によって与えられる。すなわち、

$$\text{Activity_Attribute} = G \cdot \text{power}(D \cdot \text{Spatial_Activity_Attribute}, \text{exponent_beta}) + E \cdot \text{power}(F \cdot \text{Temporal_Activity_Attribute}, \text{exponent_delta}) \quad (C)$$

である。

【0048】

一部の実施形態では、`Temporal_Activity_Attribute`は、フレーム間の動きに起因する、許容される（すなわち、マスキングされる）ことが可能な歪みの量を定量化する。それらの実施形態の一部では、フレームの`Temporal_Activity_Attribute`は、ある定数に、フレーム内の定義されたピクセル領域の動き補償された誤差信号の絶対値の合計を掛けた値に等しい。他の諸実施形態では、`Temporal_Activity_Attribute`は、以下の数式(D)によって与えられる。すなわち、

20

【0049】

【数D】

`Temporal_Activity_Attribute` =

$$\sum_{j=-1}^{-N} (W_j \cdot \text{avgFrameSAD}(j)) + \sum_{j=1}^M (W_j \cdot \text{avgFrameSAD}(j)) + W_0 \cdot \text{avgFrameSAD}(0) \quad (D)$$

30

【0050】

である。

【0051】

数式(D)では、「`avgFrameSAD`」は、(前述したとおり、)フレーム内の平均マクロブロックSAD(`MbSAD(k, m)`)値を表し、`avgFrameSAD(0)`は、現在のフレームに関する`avgFrameSAD`であり、負のjは、現在のフレームより前の時間インスタンスを指し示し、正のjは、現在のフレームより後の時間インスタンスを指し示す。このため、`avgFrameSAD(j = -2)`は、現在のフレームより前の2つのフレームの平均フレームSADを表し、`avgFrameSAD(j = 3)`は、現在のフレームより後の3つのフレームの平均フレームSADを表す。

40

【0052】

また、数式(D)において、変数Nおよび変数Mは、現在のフレームより前のフレームの数、および現在のフレームより後のフレームの数をそれぞれ示す。特定のフレーム数に基づいて値Nおよび値Mを単に選択する代わりに、一部の実施形態は、現在のフレームの時間より前の、特定の時間の長さ、およびその時間より後の、特定の時間の長さに基づき、値Nおよび値Mを計算する。動きマスキングを時間的長さと互いに関係付けることは、動きマスキングを所定のフレーム数と互いに関係付けることよりも有利である。これは、動きマスキングを時間的長さと互いに関係付けることが、見る人の時間ベースの視覚的認識にまさに一致しているからである。他方、そのようなマスキングをフレーム数と互いに

50

関係付けることには、異なるディスプレイが、異なるフレームレートでビデオを提示するので、表示時間が定まらないという難点がある。

【 0 0 5 3 】

数式 (D) において、「 W 」は、一部の実施形態では、フレーム j が、現在のフレームから離れるにつれ、減少する重み係数を指す。やはり、この数式において、第 1 の合計は、現在のフレームより前にマスキングされることが可能な動きの量を表し、第 2 の合計は、現在のフレームより後にマスキングされることが可能な動きの量を表し、最後の項 ($avgFrameSAD(0)$) は、現在のフレームのフレーム SAD を表す。

【 0 0 5 4 】

一部の実施形態では、重み係数は、シーン変化を考慮に入れるように調整される。例えば、一部の実施形態は、ルックアヘッド範囲内 (すなわち、 M 個のフレームの範囲内) の来たるべきシーン変化を考慮に入れるが、シーン変化後のいずれのフレームも考慮に入れない。例えば、それらの実施形態は、シーン変化後のルックアヘッド範囲内のフレームに関しては、重み係数を 0 に設定することが可能である。また、一部の実施形態は、ルックビハインド範囲内 (すなわち、 N 個のフレームの範囲内) で、シーン変化に先立つフレーム、またはシーン変化時のフレームを考慮に入れない。例えば、それらの実施形態は、前のシーンに関係する、または前のシーン変化の前に来るルックビハインド範囲内のフレームに関しては、重み係数を 0 に設定することが可能である。

【 0 0 5 5 】

3 . 第 2 のアプローチの変形例

a) `Temporal_Activity_Attribute` に対する過去のフレーム、および将来のフレームの影響を制限すること

前述の数式 (D) は、`Temporal_Activity_Attribute` を基本的に以下の関係で表現する。すなわち、

$$\text{Temporal_Activity_Attribute} = \text{Past_Frame_Activity} + \text{Future_Frame_Activity} + \text{Current_Frame_Activity}$$

ただし、`Past_Frame_Activity (PFA)` は、

【 0 0 5 6 】

【 数 2 】

$$\sum_{i=1}^N (W_i \cdot avgFrameSAD(i))$$

【 0 0 5 7 】

に等しく、`Future_Frame_Activity (FFA)` は、

【 0 0 5 8 】

【 数 3 】

$$\sum_{j=1}^M (W_j \cdot avgFrameSAD(j))$$

【 0 0 5 9 】

に等しく、`Current_Frame_Activity (CFA)` は、`avgFrameSAD (current)` に等しい。

【 0 0 6 0 】

一部の実施形態は、`Temporal_Activity_Attribute` の計算を変更して、`Past_Frame_Activity` も、`Future_Frame_Activity` も、`Temporal_Activity_Attribute` の値を過度にコントロールしないようにする。例えば、一部の実施形態は、最初、`PFA` が、

【 0 0 6 1 】

【数 4】

$$\sum_{i=1}^N (W_i \cdot \text{avgFrameSAD}(i))$$

【 0 0 6 2 】

と等しくなり、F F A が、

【 0 0 6 3 】

【数 5】

$$\sum_{j=1}^M (W_j \cdot \text{avgFrameSAD}(j))$$

10

【 0 0 6 4 】

と等しくなるように定義する。

【 0 0 6 5 】

これらの実施形態は、次に、P F A が、スカラー掛ける F F A より大きいかどうかを判定する。大きい場合、それらの実施形態は、次に、P F A を、P F A 上限値（例えば、スカラー掛ける F F A）と等しくなるように設定する。P F A 上限値と等しくなるように P F A を設定することに加え、一部の実施形態は、F F A を 0 に設定することと、C F A を 0 に設定することの組み合わせも実行することができる。他の諸実施形態は、P F A と C F A のいずれか、または両方を、P F A、C F A、および F F A の重み付き組み合わせに設定することが可能である。

20

【 0 0 6 6 】

同様に、P F A 値および F F A 値を重み付き合計に基づいて最初に定義した後、一部の実施形態は、F F A 値が、スカラー掛ける P F A より大きいのかも判定する。大きい場合、それらの実施形態は、次に、F F A を、F F A 上限値（例えば、スカラー掛ける P F A）と等しくなるように設定する。F F A 上限値と等しくなるように F F A を設定することに加え、一部の実施形態は、P F A を 0 に設定することと、C F A を 0 に設定することの組み合わせも実行することができる。他の諸実施形態は、F F A と C F A のいずれか、または両方を、F F A、C F A、および P F A の重み付き組み合わせに設定することが可能である。

30

【 0 0 6 7 】

（重み付き合計に基づく、P F A 値および F F A 値の初期計算後の）続いて行われる P F A 値および F F A 値の可能な調整により、これらの値のいずれかが、Temporal_Activity_Attribute を過度にコントロールすることも防止される。

【 0 0 6 8 】

b) Spatial_Activity_Attribute および Temporal_Activity_Attribute の Activity_Attribute に対する影響を制限すること

前述の数式 (C) は、基本的に、以下の関係で Activity_Attribute を表す。すなわち、

40

$$\text{Activity_Attribute} = \text{Spatial_Activity} + \text{Temporal_Activity}$$

ただし、Spatial_Activity は、 $\text{scalar} * (\text{scalar} * \text{Spatial_Activity_Attribute})$ に等しく、Temporal_Activity は、 $\text{scalar} * (\text{scalar} * \text{Temporal_Activity_Attribute})$ に等しい。

【 0 0 6 9 】

一部の実施形態は、Activity_Attribute の計算を変更して、Spatial_Activity も、Temporal_Activity も、Activity_Attribute の値を過度にコントロールしないようにする。例えば、一部の実施形態は、最初、Spatial_Activity (S A) が、 $\text{scalar} * (s$

50

$scalar * Spatial_Activity_Attribute$) と等しくなるように定義し、 $Temporal_Activity(TA)$ が、 $scalar * (scalar * Temporal_Activity_Attribute)$ と等しくなるように定義する。

【0070】

それらの実施形態は、次に、 SA が、スカラー掛ける TA よりも大きいかどうかを判定する。大きい場合、それらの実施形態は、次に、 SA を、 SA 上限値（例えば、スカラー掛ける TA ）と等しくなるように設定する。そのようなケースにおいて SA 上限と等しくなるように SA を設定することに加え、一部の実施形態は、 TA 値を、0 に、または TA と SA の重み付き組み合わせに設定することも可能である。

10

【0071】

同様に、指数方程式に基づいて SA 値および TA 値を最初に定義した後、一部の実施形態は、 TA 値が、スカラー掛ける SA よりも大きいのかも判定する。大きい場合、それらの実施形態は、次に、 TA を、 TA 上限値（例えば、スカラー掛ける SA ）と等しくなるように設定する。そのようなケースにおいて TA 上限と等しくなるように TA を設定することに加え、一部の実施形態は、 SA 値を 0 に、または SA と TA の重み付き組み合わせに設定することも可能である。

【0072】

（指数方程式に基づく、 SA 値および TA 値の初期計算後の）続いて行われる SA 値および TA 値の可能な調整により、これらの値のいずれかが、 $Activity_Attribute$ を過度にコントロールすることも防止される。

20

【0073】

< B . マクロブロックレベルのマスキング強度を計算すること >

1 . 第 1 のアプローチ

一部の実施形態では、マクロブロックレベルのマスキング強度 $M_B(k, m)$ は、以下のとおり計算される。すなわち、

$$M_B(k, m) = A * \text{power}(C * \text{avgMbLuma}(k, m), \quad) * \text{power}(B * \text{MbSAD}(k, m), \quad M_B), (F)$$

ただし、

・ $\text{avgMbLuma}(k, m)$ は、フレーム k 、マクロブロック m における平均ピクセル輝度であり、

30

・ M_B 、 A 、 B 、および C は、定数であり、かつ（或いは）、局所的な統計に適合されている。

【0074】

2 . 第 2 のアプローチ

前述した数式（ F ）は、マクロブロックのマスキング強度を基本的に以下のとおり計算する。すなわち、

$$M_B(k, m) = D * \text{power}(E * \text{Mb_Brightness_Attribute}, \text{exponent0}) * \text{power}(scalar * \text{Mb_Spatial_Activity_Attribute}, \text{exponent1})$$

である。

【0075】

40

数式（ F ）において、マクロブロックの $Mb_Brightness_Attribute$ は、 $\text{avgMbLuma}(k, m)$ に等しく、 $Mb_Spatial_Activity_Attribute$ は、 $\text{avgMbSAD}(k)$ に等しい。この $Mb_Spatial_Activity_Attribute$ は、符号化中のマクロブロック内のピクセル領域内における空間的イノベーションの量を測定する。

【0076】

フレームのマスキング強度の場合と全く同様に、一部の実施形態は、いくつかの連続するフレームにわたるピクセル領域内の時間的イノベーションの量を含むように、マクロブロックのマスキング強度における活動性の尺度を拡張することが可能である。具体的には、それらの実施形態は、マクロブロックのマスキング強度を以下のとおり計算する。すな

50

わち、

$$M_B(k,m) = D \cdot \text{power}(E \cdot \text{Mb_Brightness_Attribute}, \text{exponent0}) \cdot \text{power}(\text{scalar} \cdot \text{Mb_Activity_Attribute}, \text{exponent1})$$

ただし、Mb_Activity_Attributeは、以下の数式(H)によって与えられる。すなわち、

$$\text{Mb_Activity_Attribute} = F \cdot \text{power}(D \cdot \text{Mb_Spatial_Activity_Attribute}, \text{exponent_beta}) + G \cdot \text{power}(F \cdot \text{Mb_Temporal_Activity_Attribute}, \text{exponent_delta}) \quad (H)$$

である。

【0077】

マクロブロックに関するMb_Temporal_Activity_Attributeの計算は、フレームに関するMb_Temporal_Activity_Attributeの前述した計算と同様であることが可能である。例えば、それらの実施形態の一部では、Mb_Temporal_Activity_Attributeは、以下の数式(I)によって与えられる。すなわち、

【0078】

【数I】

$$\text{Mb_Temporal_Activity_Attribute} =$$

$$\sum_{i=1}^N (W_i \cdot \text{MbSAD}(i,m)) + \sum_{j=1}^M (W_j \cdot \text{MbSAD}(j,m)) + \text{MbSAD}(m) \quad (I)$$

【0079】

である。

【0080】

数式(I)の中の変数は、セクションIII.Aにおいて定義された。数式(F)において、フレームi内、またはフレームj内のマクロブロックmは、現在のフレーム内のマクロブロックmと同一の位置におけるマクロブロックであることが可能である。或いは、フレームi内、またはフレームj内のマクロブロックmは、現在のフレーム内のマクロブロックmと一致すると最初に予測されたフレームi内、またはフレームj内のマクロブロックであることが可能である。

【0081】

数式(I)によって与えられるMb_Temporal_Activity_Attributeは、数式(D)によって与えられる、フレームのTemporal_Activity_Attributeの変更(前述のセクションIII.A.3で説明された)と同様の形で変更されることが可能である。具体的には、数式(I)によって与えられるMb_Temporal_Activity_Attributeは、過去のフレーム内、および将来のフレーム内におけるマクロブロックの過度の影響を制限するように変更されることが可能である。

【0082】

同様に、数式(H)によって与えられるMb_Activity_Attributeも、数式(C)によって与えられる、フレームのActivity_Attributeの変更(セクションIII.A.3で前述した)と同様の形で変更されることが可能である。具体的には、数式(H)によって与えられるMb_Activity_Attributeは、Mb_Spatial_Activity_AttributeおよびMb_Temporal_Activity_Attributeの過度の影響を制限するように変更されることが可能である。

【0083】

< C . マスキングされたQP値を計算すること >

マスキング強度の値(F および M_B)、および基準マスキング強度の値(R)に基づき、視覚マスキングプロセスは、2つの関数CalcMQPおよびCalcMQPf

10

20

30

40

50

or MBを使用することにより、フレームレベルおよびマクロブロックレベルにおけるマスキングされたQP値を計算することができる。これら2つの関数に関する擬似コードは、以下のとおりである。すなわち、

```
CalcMQP(nominalQP,  $R$ ,  $F(k)$ , maxQPFrameAdjustment)
{
    QPFrameAdjustment =  $F * (F(k) - R) / R$ ;
    [minQPFrameAdjustment, maxQPFrameAdjustment]の範囲内に入るようにQPFrameAdjustmentをクリッピングする;
    maskedQPofFrame = nominalQP + QPFrameAdjustment;
    許容範囲内に入るようにmaskedQPofFrameをクリッピングする;
    (フレームkに関する)maskedQPofFrameを返す;
}
```

```
CalcMQPforMB(maskedQPofFrame,  $F(k)$ ,  $M_B(k, m)$ , maxQPMacroblockAdjustment)
{
    if(  $F(k) > T$ ) ただし、Tは、適切に選択された閾値
        QPMacroblockAdjustment =  $M_B * (M_B(k, m) - F(k)) / F(k)$ ;
    else
        QPMacroblockAdjustment = 0;
    [minQPMacroblockAdjustment, maxQPMacroblockAdjustment]の範囲内に入るようにQP
    MacroblockAdjustmentをクリッピングする;
    maskedQPofMacroblock = maskedQPofFrame + QPMacroblockAdjustment;
    有効なQP値範囲内に入るようにmaskedQPofMacroblockをクリッピングする;
    maskedQPofMacroblockを返す;
}
```

である。

【0084】

前述の関数において、 F および M_B は、所定の定数であること、または局所的な統計に適合させられることが可能である。

【0085】

[IV. マルチパス符号化]

図1は、本発明の一部の実施形態に係るマルチパス符号化方法を概念的に示すプロセス100を提示する。この図に示されるとおり、プロセス100は、以下の3つのサブセクションで説明される3つの段階を有する。

【0086】

< A . 分析および初期(最初の)QP選択 >

図1に示されるとおり、プロセス100は、最初、マルチパス符号化プロセスの初期分析段階中に(すなわち、パス0中に)、基準マスキング強度($R_{(1)}$)の初期値、および名目量子化パラメータ($QP_{nom(1)}$)の初期値を計算する(105で)。初期基準マスキング強度($R_{(1)}$)は、第1の探索段階中に使用されるのに対して、初期名目量子化パラメータ($QP_{nom(1)}$)は、第1の探索段階の第1のパス中に(すなわち、マルチパス符号化プロセスのパス1中に)使用される。

【0087】

パス0の始めに、 $R_{(0)}$ は、何らかの任意の値、または実験的結果に基づいて選択された値(例えば、 R 値の通常範囲の中央値)であることが可能である。シーケンスの分析中、各フレームに関してマスキング強度 $F(k)$ が計算され、次に、基準マスキング強度 $R_{(1)}$ が、パス0の終りにおいて $avg(F(k))$ と等しくなるように設定される。また、基準マスキング強度 R に関する他の決定も可能である。例えば、基準マスキング強度 R は、値 $F(k)$ の中央値として、または例えば、値 $F(k)$ の重み付き平均値などの、値 $F(k)$ に関する他の算術関数として計算されてもよい。

【 0 0 8 8 】

異なる複雑性を有する、初期 Q P 選択のいくつかのアプローチが存在する。例えば、初期名目 Q P は、任意の値（例えば、26）として選択されることが可能である。代わりに、符号化実験に基づき、目標ビットレートに対して許容できる品質をもたらすことが知られている値が、選択されることが可能である。

【 0 0 8 9 】

また、初期名目 Q P 値は、空間分解能、フレームレート、空間的 / 時間的複雑性、および目標ビットレートに基づき、ルックアップテーブルから選択されることが可能である。一部の実施形態では、この初期名目 Q P 値は、以上のパラメータの各々に基づく距離の尺度を使用して、テーブルから選択されてもよい。あるいは、以上のパラメータの重み付き距離の尺度を使用して選択されてもよい。

10

【 0 0 9 0 】

また、この初期名目 Q P 値は、レートコントローラを使用する（マスキングなしの）高速符号化中に、フレーム Q P 値が選択されるにつれ、フレーム Q P 値の調整された平均に設定されることも可能である。ただし、その平均は、パス 0 に関するビットレートパーセンテージのレート誤差 E_0 に基づいて調整されている。また、同様に、初期名目 Q P は、フレーム Q P 値の重み付きの調整された平均に設定されることも可能である。ただし、各フレームに関する重みは、飛ばされるマクロブロック（スキップドマクロブロック）として符号化されない、そのフレーム内のマクロブロックのパーセンテージによって決まる。代わりに、初期名目 Q P は、基準マスキング強度を $R(0)$ から $R(1)$ に変更することの効果は考慮に入れられる限り、レートコントローラを使用する（マスキングなしの）高速符号化中に、フレーム Q P 値が選択されるにつれ、フレーム Q P 値の調整された平均、または調整された重み付き平均に設定されることも可能である。

20

【 0 0 9 1 】

< B . 第 1 の探索段階：名目 Q P 調整 >

105 の後、マルチパス符号化プロセス 100 は、第 1 の探索段階に入る。第 1 の探索段階では、プロセス 100 は、シーケンスに対する N_1 回の符号化を実行する。ただし、 N_1 は、第 1 の探索段階中のパスの回数を表す。第 1 の段階の各パス中、プロセスは、変化する名目量子化パラメータを、一定の基準マスキング強度とともに使用する。

【 0 0 9 2 】

具体的には、第 1 の探索段階における各パス p 中、プロセス 100 は、各フレーム k に関する特定の量子化パラメータ $MQP_p(k)$ 、およびフレーム k 内のそれぞれの個別のマクロブロックに関する特定の量子化パラメータ $MQP_{MB(p)}(k, m)$ を計算する（107 で）。所与の名目量子化パラメータ $QP_{Nom(p)}$ および基準マスキング強度 $R(p)$ に関するパラメータ $MQP_p(k)$ および $MQP_{MB(p)}(k, m)$ の計算は、セクション III で説明されている（ただし、 $MQP_p(k)$ および $MQP_{MB(p)}(k, m)$ は、セクション III で前述した関数、 $CalcMQP$ および $CalcMQP_{forMB}$ を使用して計算される）。107 中の第 1 のパス（すなわち、パス 1）において、名目量子化パラメータおよび第 1 段階の基準マスキング強度は、初期分析段階 105 中に計算されたパラメータ $QP_{Nom(1)}$ および基準マスキング強度 $R(1)$ である。

30

40

【 0 0 9 3 】

107 の後、プロセスは、107 で計算された量子化パラメータ値に基づき、シーケンスを符号化する（110 で）。次に、符号化プロセス 100 は、終了すべきかどうかを判定する（115 で）。異なる実施形態は、全体的な符号化プロセスを終了させることに關して、異なる基準を有する。マルチパス符号化プロセスを完全に終了させる終了条件の実施例には、以下が含まれる。すなわち、

- ・ $|E_p| < \epsilon$ である場合。ただし、 ϵ は、最終ビットレートにおける許容誤差である。
- ・ $QP_{Nom(p)}$ が、Q P 値の有効範囲の上限または下限にある場合。
- ・ パスの回数が、容認できるパスの最大回数 P_{MAX} を超えた場合。

50

【 0 0 9 4 】

一部の実施形態は、以上の終了条件のすべてを使用する可能性があるのに対して、他の諸実施形態は、これらの条件の一部だけを使用する可能性がある。さらに別の諸実施形態は、符号化プロセスを終了させることにに関して、他の終了条件を使用することが可能である。

【 0 0 9 5 】

マルチパス符号化プロセスが、終了することを決めた場合（ 1 1 5 で）、プロセス 1 0 0 は、第 2 の探索段階を省き、 1 4 5 に進む。 1 4 5 で、プロセスは、最後のパス p からのビットストリームを最終結果として保存し、その後、終了する。

【 0 0 9 6 】

他方、プロセスが、終了すべきではないと判定した場合（ 1 1 5 で）、プロセスは、次に、第 1 の探索段階を終了させるべきかどうかを判定する（ 1 2 0 で）。やはり、異なる諸実施形態は、第 1 の探索段階を終了させることにに関して、異なる基準を有する。マルチパス符号化プロセスの第 1 の探索段階を終了させる終了条件の実施例には、以下が含まれる。すなわち、

- ・ $Q P_{N o m (p+1)}$ が、 $Q P_{N o m (q)}$ と同一であり、かつ $q = p$ である場合（その場合、ビットレートの誤差は、名目 $Q P$ を変更しても、それよりも下げることができない）。

- ・ $|E_p| < \epsilon_c$ であり、 $\epsilon_c > 0$ である場合。ただし、 ϵ_c は、第 1 の探索段階に関するビットレートの許容誤差である。

- ・ パスの回数が、 P_1 を超えている場合。ただし、 P_1 は、 $P_{M A X}$ 未満である。

- ・ パスの回数が、 P_1 未満である P_2 を超えており、かつ $|E_p| < \epsilon_2$ であり、 $\epsilon_2 > 0$ である場合。

【 0 0 9 7 】

一部の実施形態は、以上の終了条件のすべてを使用する可能性があるのに対して、他の諸実施形態は、これらの条件の一部だけを使用する可能性がある。さらに別の諸実施形態は、第 1 の探索段階を終了させることにに関して、他の終了条件を使用することが可能である。

【 0 0 9 8 】

マルチパス符号化プロセスが、第 1 の探索段階を終了させることを決めた場合（ 1 2 0 で）、プロセス 1 0 0 は、次のサブセクションで説明される、第 2 の探索段階に進む。他方、プロセスが、第 1 の探索段階を終了させるべきではないと判定した場合（ 1 2 0 で）、プロセスは、第 1 の探索段階における次のパスに関する名目 $Q P$ を更新する（ 1 2 5 で）（すなわち、 $Q P_{N o m (p+1)}$ を定義する）。一部の実施形態では、名目 $Q P_{N o m (p+1)}$ は、以下のとおり更新される。パス 1 の終りで、それらの実施形態は、

$Q P_{N o m (p+1)} = Q P_{N o m (p)} + E_p$ と定義する。ただし、 E_p は、定数である。パス 2 からパス N_1 までの各パスの終りで、それらの実施形態は、次に、

$Q P_{N o m (p+1)} = \text{InterpExtrap}(0, E_{q1}, E_{q2}, Q P_{N o m (q1)}, Q P_{N o m (q2)})$

と定義する。ただし、 InterpExtrap は、以下にさらに説明される関数である。また、上の数式では、 $q1$ および $q2$ は、パス p までのすべてのパスの中で最低である、対応するビットレート誤差を有するパス番号であり、 $q1$ 、 $q2$ 、および p は、以下の関係を有する。すなわち、

$1 \leq q1 \leq q2 \leq p$

である。

【 0 0 9 9 】

以下は、 InterpExtrap 関数に関する擬似コードである。x が、 $x1$ から $x2$ までの間にない場合、この関数は、補外（外挿）関数であることに留意されたい。それ以外の場合、この関数は、補間（内挿）関数である。

InterpExtrap(x,x1,x2,y1,y2)

```
{
    if(x2 != x1) y = y1 + (x - x1) * (y2 - y1) / (x2 - x1);
    else y = y1;
    return y;
}
```

【0100】

名目QP値は、通常、整数値に丸められ、QP値の有効範囲内に入るようにクリッピングされる。他の諸実施形態は、前述したアプローチとは異なる形で、名目 $QP_{Nom(p+1)}$ 値を計算することも可能であることが、当業者には認識されよう。

10

【0101】

125の後、プロセスは、107に戻り、次のパス（すなわち、 $p := p + 1$ ）を開始し、このパスに関して、各フレームkに関する特定の量子化パラメータ $MQP_p(k)$ 、ならびに現在のパスpに関するフレームk内のそれぞれの個別のマクロブロックmに関する特定の量子化パラメータ $MQP_{MB(p)}(k, m)$ を計算する（107で）。次に、プロセスは、それらの新たに計算された量子化パラメータに基づき、フレームシーケンスを符号化する（110で）。そして、110から、プロセスは、前述した115に進む。

【0102】

< C . 第2の探索段階：基準マスキング強度調整 >

プロセス100は、第1の探索段階を終了すべきであると判定した場合（120で）、130に進む。第2の探索段階で、プロセス100は、シーケンスの N_2 回の符号化を実行する。ただし、 N_2 は、第2の探索段階中のパスの回数を表す。各パス中、プロセスは、同一の名目量子化パラメータ、および変化する基準マスキング強度を使用する。

20

【0103】

130で、プロセス100は、パス $N_1 + 1$ である次のパス、すなわちパス $p + 1$ に関する基準マスキング強度 $R(p + 1)$ を計算する。パス $N_1 + 1$ において、プロセス100は、135でフレームシーケンスを符号化する。異なる諸実施形態は、パスpの終りに、異なる形で基準マスキング強度 $R(p + 1)$ を計算する（130で）。2つの代替のアプローチを以下に説明する。

【0104】

一部の実施形態は、先行するパスからのビットレートの誤差、および R の値に基づき、基準マスキング強度 $R(p)$ を計算する。例えば、パス N_1 の終りに、一部の実施形態は、

30

$$R(N_1 + 1) = R(N_1) + R(N_1) \times \text{Konst} \times E_{N_1}$$
であると定義する。

【0105】

mが、1より大きい整数であるパス $N_1 + m$ の終りに、一部の実施形態は、

$$R(N_1 + m) = \text{InterpExtrap}(0, E_{N_1 + m - 2}, E_{N_1 + m - 1}, R(N_1 + m - 2), R(N_1 + m - 1))$$
であると定義する。

40

【0106】

代わりに、一部の実施形態は、

$$R(N_1 + m) = \text{InterpExtrap}(0, E_{N_1 + m - q_2}, E_{N_1 + m - q_1}, R(N_1 + m - q_2), R(N_1 + m - q_1))$$
であると定義する。ただし、 q_1 および q_2 は、最良の誤差を与えた先行するパスである。

【0107】

他の諸実施形態は、セクションIで定義されたAMQPを使用することにより、第2の探索段階における各パスの終りに、基準マスキング強度を計算する。所与の名目QP、および R の何らかの値に関してAMQPを計算するための1つのやり方を、関数GetA

50

v g M a s k e d Q P の擬似コードに関連して以下に説明する。

GetAvgMaskedQP(nominalQP, R)

```
{
    sum=0;
    for(k=0;k < numframes;k++){
        MQP(k)=CalcMQP(nominalQP,  $R$ ,  $F(k)$ ,maxQPFrameAdjustment)を使用して
        計算された、フレーム k に関するmaskedQP; //前段を参照
        sum+=MQP(k);
    }
    return sum/numframes;
}
```

10

【0108】

AMQPを使用する一部の実施形態は、先行するパスからのビットレートの誤差、およびAMQPの値に基づき、パス $p+1$ に関する所望されるAMQPを計算する。そのAMQPに対応する $R(p+1)$ が、次に、関数 $Search(AMQP(p+1), R(p))$ によって与えられる探索手続きを介して求められる。この関数の擬似コードは、このサブセクションの終りにおいて与えられる。

【0109】

例えば、パス N_1 の終りにおいて一部の実施形態は、 $AMQP_{N_1+1}$ を計算する。ただし、

20

$N_1 > 1$ である場合、 $AMQP_{N_1+1} = InterpExtrap(0, E_{N_1-1}, E_{N_1}, AMQP_{N_1-1}, AMQP_{N_1})$ であり、かつ

$N_1 = 1$ である場合、 $AMQP_{N_1+1} = AMQP_{N_1}$ である。

【0110】

次に、以上の実施形態は、

$R(N_1+1) = Search(AMQP_{N_1+1}, R(N_1))$
であると定義する。

【0111】

パス N_1+m の終りに（ただし、 m は、1より大きい整数）、一部の実施形態は、

$AMQP_{N_1+m} = InterpExtrap(0, E_{N_1+m-2}, E_{N_1+m-1}, AMQP_{N_1+m-2}, AMQP_{N_1+m-1})$

30

であり、かつ

$R(N_1+m) = Search(AMQP_{N_1+m}, R(N_1+m-1))$
であると定義する。

【0112】

所望されるAMQP、および R の何らかの既定値が与えられると、所望されるAMQPに対応する R は、一部の実施形態では、以下の擬似コードを有する $Search$ 関数を使用して求められることが可能である。すなわち、

Search(AMQP, R)

```
{
    interpolateSuccess=True; //別の設定が行われるまで

    reLumaSad0=refLumaSad1=refLumaSadx= $R$ ;
    errorInAvgMaskedQp=GetAvgMaskedQp(nominalQp,refLumaSadx)-AMQP;
    if(errorInAvgMaskedQp > 0){
        ntimes=0;
        do{
            ntimes++;
            refLumaSad0=(refLumaSad0*1.1);
            errorInAvgMaskedQp=GetAvgMaskedQp(nominalQp,refLumaSad0)-amqp;
```

40

50

```

        }while(errorInAvgMaskedQp > 0 && ntimes < 10);
        if(ntimes >=10) interpolateSuccess=False;
    }
    else{ //errorInAvgMaskedQp < 0
        ntimes=0;
        do{
            ntimes++;
            refLumaSad1=(refLumaSad1*0.9);
            errorInAvgMaskedQp=GetAvgMaskedQp(nominalQp, refLumaSad1)-amqp;
        }while(errorInAvgMaskedQp < 0 && ntimes < 10);
        if(ntimes >=10) interpolateSuccess=False;
    }
    ntimes=0;
    do{
        ntimes++;
        refLumaSadx=(refLumaSad0+refLumaSad1)/2; //単純な連続近似
        errorInAvgMaskedQp=GetAvgMaskedQp(nominalQp, refLumaSadx)-AMQP;
        if(errorInAvgMaskedQp > 0) refLumaSad1=refLumaSadx;
        else refLumaSad0=refLumaSadx;
    }while(ABS(errorInAvgMaskedQp) > 0.05 && ntimes < 12);
    if(ntimes >=12) interpolateSuccess=False;
    }
    if(interpolateSuccess) return refLumaSadx;
    else return R
}

```

である。

【 0 1 1 3 】

以上の擬似コードにおいて、数値 1 0、1 2、および 0 . 0 5 は、適切に選択された閾値で置き換えられてもよい。

【 0 1 1 4 】

フレームシーケンスの符号化を介して、次のパス（パス $p + 1$ ）に関する基準マスキング強度を計算した後、プロセス 1 0 0 は、1 3 2 に進み、次のパス（すなわち、 $p := p + 1$ ）を開始する。各符号化パス p 中の各フレーム k 、および各マクロブロック m に関して、プロセスは、各フレーム k に関する特定の量子化パラメータ $MQP_p(k)$ 、ならびにフレーム k 内の個々のマクロブロック m に関する特定の量子化パラメータ $MQP_{MB(p)}(k, m)$ を計算する（1 3 2 で）。所与の名目量子化パラメータ $QP_{Nom(p)}$ 、および基準マスキング強度 $R(p)$ に関するパラメータ $MQP_p(k)$ および $MQP_{MB(p)}(k, m)$ の計算は、セクション I I I において説明されている（ただし、 $MQP_p(k)$ および $MQP_{MB(p)}(k, m)$ は、セクション I I I において前述した関数、 $CalcMQP$ および $CalcMQP_{forMB}$ を使用することによって計算される）。1 3 2 の間の第 1 のパス中、基準マスキング強度は、1 3 0 において計算されたばかりの基準マスキング強度である。また、第 2 の探索段階中、名目 QP も、第 2 の探索段階全体を通して、一定のままである。一部の実施形態では、第 2 の探索段階中の名目 QP は、第 1 の探索段階中に最良の符号化ソリューションをもたらした（すなわち、最低のビットレート誤差の符号化ソリューションをもたらした）名目 QP である。

【 0 1 1 5 】

1 3 2 の後、プロセスは、1 3 0 で計算された量子化パラメータを使用してフレームシーケンスを符号化する（1 3 5 で）。1 3 5 の後、プロセスは、第 2 の探索段階を終了させるべきかどうかを判定する（1 4 0 で）。異なる諸実施形態は、パス p の終りに第 2 の探索段階を終了させることに関して、異なる基準を使用する。そのような基準の実施例は

、以下のとおりである。すなわち、

- ・ $|E_p| < \quad$ の場合。ただし、 \quad は、最終ビットレートの許容誤差である。
- ・ パスの回数が、許容されるパスの最大回数を超えている場合。

【 0 1 1 6 】

一部の実施形態は、以上の終了条件のすべてを使用する可能性があるのに対して、他の諸実施形態は、これらの条件の一部だけを使用する可能性がある。さらに別の諸実施形態は、第 1 の探索段階を終了させることにに関して、他の終了条件を使用することが可能である。

【 0 1 1 7 】

プロセス 1 0 0 は、第 2 の探索段階を終了させるべきでないと判定すると (1 4 0 において)、1 3 0 に戻り、符号化の次のパスに関する基準マスキング強度を再計算する。1 3 0 から、プロセスは、1 3 2 に進んで、量子化パラメータを計算し、次に、1 3 5 に進んで、その新たに計算された量子化パラメータを使用することにより、ビデオシーケンスを符号化する。

【 0 1 1 8 】

他方、プロセスは、第 2 の探索段階を終了させることを決めた場合 (1 4 0 で)、1 4 5 に進む。1 4 5 で、プロセス 1 0 0 は、最後のパス p からのビットストリームを最終結果として保存し、その後、終了する。

【 0 1 1 9 】

[V . 復号器入力バッファアンダフロー制御]

本発明の一部の実施形態は、復号器によって使用される入力バッファの使用に関して最適な符号化ソリューションを特定するため、ビデオシーケンスの目標ビットレートに対する様々な符号化を調べるマルチパス符号化プロセスを提供する。一部の実施形態は、このマルチパスプロセスは、図 1 のマルチパス符号化プロセス 1 0 0 に従う。

【 0 1 2 0 】

復号器入力バッファ (「復号器バッファ」) の使用率は、符号化されたイメージシーケンス (例えば、フレーム) の復号化中、ある程度変動する。これは、符号化されたイメージのサイズの変動、復号器が符号化されたデータを受け取る速度、復号器バッファのサイズ、復号化プロセスの速度、その他の、様々な要因のためである。

【 0 1 2 1 】

復号器バッファアンダフローとは、イメージが、復号器側に完全に到着する前に、復号器が、次のイメージを復号化する準備ができていない状況を意味する。一部の実施形態のマルチパス符号器は、復号器バッファをシミュレートし、シーケンス内の選択されたセグメントを再符号化して、復号器バッファアンダフローを防止する。

【 0 1 2 2 】

図 2 は、本発明の一部の実施形態のコーデックシステム 2 0 0 を概念的に示す。このシステムは、復号器 2 0 5 および符号器 2 1 0 を含む。この図では、符号器 2 1 0 は、符号器 2 1 0 が、復号器 2 0 5 の同様のコンポーネントの動作をシミュレートすることを可能にする、いくつかのコンポーネントを有する。

【 0 1 2 3 】

具体的には、復号器 2 0 5 は、入力バッファ 2 1 5 と、復号化プロセス 2 2 0 と、出力バッファ 2 2 5 とを有する。符号器 2 1 0 は、シミュレートされた復号器入力バッファ 2 3 0、シミュレートされた復号化プロセス 2 3 5、およびシミュレートされた復号器出力バッファ 2 4 0 を保持することにより、以上のモジュールをシミュレートする。本発明の説明を妨げないように、図 2 は、復号化プロセス 2 2 0 および符号化プロセス 2 4 5 を単一のブロックとして示すように単純化されている。また、一部の実施形態では、シミュレートされた復号化プロセス 2 3 5、およびシミュレートされた復号器出力バッファ 2 4 0 は、バッファアンダフロー管理のために利用されず、したがって、この図では、単に例示のために示されている。

【 0 1 2 4 】

復号器は、入力バッファ 215 を保持して、入ってくる符号化イメージの速度および到着時間の変動を平滑化する。復号器に、データがなくなった（アンダフロー）場合、または入力バッファがいっぱいになった場合（オーバフロー）、ピクチャ復号化が止まるので、または入ってくるデータが破棄されるので、目に見える復号化の不連続が存在する。これらのケースのいずれも、望ましくない。

【0125】

アンダフロー条件を解消するために、符号器 210 は、一部の実施形態では、イメージシーケンスをまず符号化し、それらをストレージ 255 の中に格納する。例えば、符号器 210 は、マルチパス符号化プロセス 100 を使用して、イメージシーケンスの第 1 の符号化を獲得する。次に、符号器 210 は、復号器入力バッファ 215 をシミュレートし、バッファアンダフローを生じさせるイメージを再符号化する。すべてのバッファアンダフロー条件が取り除かれた後、再符号化されたイメージが、ネットワーク接続（インターネット、ケーブル、PSTN 線、その他）、非ネットワークの直接接続、媒体（DVD、その他）、その他であることが可能な接続 260 を介して、復号器 205 に供給される。

【0126】

図 3 は、一部の実施形態の符号器の符号化プロセス 300 を示す。このプロセスは、復号器バッファがアンダフローになることを生じさせない最適な符号化ソリューションを見出そうと試みる。図 3 に示されるとおり、プロセス 300 は、所望される目標ビットレートを満たすイメージシーケンスの第 1 の符号化（例えば、シーケンス内の各イメージに関する平均ビットレートが、所望される平均目標ビットレートを満たす）を特定する（302 で）。例えば、プロセス 300 は、マルチパス符号化プロセス 100 を使用して（302 で）、イメージシーケンスの第 1 の符号化を得ることが可能である。

【0127】

302 の後、符号化プロセス 300 は、接続速度（すなわち、復号器が、符号化されたデータを受け取る速度）、復号器入力バッファのサイズ、符号化されたイメージのサイズ、復号化プロセス速度、その他などの、様々な要因を考慮することにより、復号器入力バッファ 215 をシミュレートする（305 で）。310 で、プロセス 300 は、符号化されたイメージのセグメントが復号器入力バッファをアンダフローさせるかどうかを判定する。アンダフロー条件を判定する（その後、解消させる）のに符号器が使用する技術は、後段でさらに説明する。

【0128】

符号化されたイメージが、アンダフロー条件を生じさせないとプロセス 300 が判定した場合（310 で）、プロセスは、終了する。他方、符号化されたイメージのいずれかのセグメント内にバッファアンダフロー条件が存在するとプロセス 300 が判定した場合（310 で）、プロセス 300 は、符号化パラメータを、先行する符号化プロセスからのこれらのパラメータの値に基づき、改良する（315 で）。次に、プロセスは、アンダフローを伴うセグメントを再符号化して（320 で）、セグメントのビットサイズを小さくする。セグメントを再符号化した後、プロセス 300 は、そのセグメントを調べて（325 で）、アンダフロー条件が解消されたかどうかを判定する。

【0129】

セグメントが、依然として、アンダフローを生じさせるとプロセスが判定した場合（325 で）、プロセス 300 は、315 に進んで、アンダフローを解消するように符号化パラメータをさらに改良する。一方、セグメントが、アンダフローを全く生じさせないとプロセスが判定した場合（325 で）、プロセスは、ビデオシーケンスを再検査して再符号化するための開始点を、320 における前回の繰り返しにおいて再符号化されたセグメントの終りの後のフレームとして指定する（330 で）。次に、335 で、プロセスは、315 および 320 で指定されたアンダフローセグメントの後に続く最初の IDR フレームまで（かつ、そのフレームを除外して）、330 で指定されたビデオシーケンスの部分を再符号化する。335 の後、プロセスは、305 に戻り、復号器バッファをシミュレートして、ビデオシーケンスの残りの部分が、再符号化の後、依然として、バッファアンダフ

10

20

30

40

50

ローを生じさせるかどうかを判定する。305からのプロセス300の流れを、以上に説明した。

【0130】

< A . 符号化されたイメージのシーケンス内のアンダフローセグメントを特定すること >

前述したとおり、符号器は、復号器バッファ条件をシミュレートして、符号化された、または再符号化されたイメージシーケンス内のいずれかのセグメントが、復号器バッファの中でアンダフローを生じさせるかどうかを判定する。一部の実施形態では、符号器は、符号化されたイメージのサイズ、帯域幅などのネットワーク条件、復号器要因（例えば、入力バッファサイズ、イメージを除去するのにかかる初期時間および名目時間、復号化プロセス時間、各イメージの表示時間、その他）を考慮するシミュレーションモデルを使用する。

10

【0131】

一部の実施形態では、MPEG-4 AVC符号化ピクチャバッファ(CPB)モデルが、復号器入力バッファの状態をシミュレートするのに使用される。CPBは、MPEG-4 H.264標準において、仮想参照デコーダ(Hypothetical Reference Decoder)(HRD)のシミュレートされた入力バッファを指すのに使用される用語である。HRDは、符号化プロセスが、生成することができる適合するストリームの変動性に対する制約を指定する仮想の復号器モデルである。CPBモデルは、周知であるが、便宜上、以下のセクション1で説明する。CPBおよびHRDのより詳細な説明は、Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification (ITU-T Rec. H.264/ISO/IEC 14496-10 AVC)で見ることができる。

20

【0132】

1. CPBモデルを使用して、復号器バッファをシミュレートすること

以下の段落は、一部の実施形態において、CPBモデルを使用して、復号器入力バッファがどのようにシミュレートされるかを説明する。イメージnの最初のビットが、CPBに入り始める時間は、初期到着時間 $t_{ai}(n)$ と呼ばれ、以下のとおり導出される。すなわち、

30

・イメージが、最初のイメージ（すなわち、イメージ0）である場合、 $t_{ai}(0) = 0$ であり、

・イメージが、符号化されている、または再符号化されているシーケンス内の最初のイメージではない場合（すなわち、 $n > 0$ である場合）、 $t_{ai}(n) = \text{Max}(t_{af}(n-1), t_{ai,earliest}(n))$

である。

【0133】

上の数式において、

・ $t_{ai,earliest}(n) = t_{r,n}(n) - \text{initial_cpb_removal_delay}$

40

である。ただし、 $t_{r,n}(n)$ は、以下に指定されるCPBからのイメージnの名目除去時間であり、 $\text{initial_cpb_removal_delay}$ は、初期バッファリング期間である。

【0134】

イメージnに関する最終到着時間は、

$$t_{af}(n) = t_{ai}(n) + b(n) / \text{BitRate}$$

によって導出される。ただし、 $b(n)$ は、イメージnのサイズ（単位はビット）である。

【0135】

一部の実施形態では、符号器は、H.264規格におけるように、ビットストリームの

50

オプションの部分（オプションナルパート）から名目除去時間を読み取る代わりに、以下に説明されるとおり、名目除去時間の独自の計算を行う。イメージ 0 に関して、C P B からのイメージの名目除去時間は、

$$t_{r,n}(0) = \text{initial_cpb_removal_delay}$$

によって指定される。

【0136】

イメージ n ($n > 0$) に関して、C P B からのイメージの名目除去時間は、

$$t_{r,n}(n) = t_{r,n}(0) + \sum_{i=0}^{n-1} (t_i)$$

によって指定される。ただし、 $t_{r,n}(n)$ は、イメージ n の名目除去時間であり、 t_i は、ピクチャ i に関する表示時間である。

10

【0137】

イメージ n の除去時間は、以下のとおり指定される。

- ・ $t_{r,n}(n) \geq t_{af}(n)$ である場合、 $t_r(n) = t_{r,n}(n)$ であり、
- ・ $t_{r,n}(n) < t_{af}(n)$ である場合、 $t_r(n) = t_{af}(n)$ である。

【0138】

イメージ n のサイズ、 $b(n)$ が、余りにも大きいため、名目除去時間における除去が妨げられることを示すのは、この後者 ($t_r(n) = t_{af}(n)$) のケースである。

【0139】

2. アンダフローセグメントの検出

前のセクションで説明されるとおり、符号器は、復号器入力バッファの状態をシミュレートし、所与の時点におけるバッファ内のビット数を獲得することができる。或いは、符号器は、その名目除去時間と最終到着時間の差（すなわち、 $t_b(n) = t_{r,n}(n) - t_{af}(n)$ ）を介して、それぞれの個別のイメージが、復号器入力バッファの状態をどのように変化させるかを追跡することができる。 $t_b(n)$ が、0 未満である場合、バッファは、時点 $t_{r,n}(n)$ と時点 $t_{af}(n)$ の間にアンダフローを来たしており、場合によっては、 $t_{r,n}(n)$ より前、および $t_{af}(n)$ の後にもアンダフローを来たしている。

20

【0140】

アンダフローに直接に関わっているイメージは、 $t_b(0)$ が、0 未満であるかどうかを試験することにより、容易に見出すことができる。しかし、0 未満の $t_b(n)$ を有するイメージは、必ずしもアンダフローを生じさせるわけではなく、逆に、アンダフローを生じさせるイメージは、0 未満の $t_b(n)$ を有さない可能性もある。一部の実施形態は、アンダフローセグメントを次のように定義する。すなわち、アンダフローがその最悪の点に達するまで、復号器入力バッファを絶えず空にすることによってアンダフローを生じさせる、一続きの連続するイメージ（復号化順の）として定義する。

30

【0141】

図 4 は、一部の実施形態における、イメージ数に対する、名目除去時間と最終イメージ到着の差 $t_b(n)$ のプロットである。このプロットは、1500 個の符号化されたイメージのシーケンスに関して描かれている。図 4 a は、アンダフローセグメントを示し、矢印が、そのセグメントの始まりと終りを示している。簡明にするため、矢印によって明示されていない、第 1 のアンダフローセグメントの後に出現する別のアンダフローセグメントが、図 4 a に存在することに留意されたい。

40

【0142】

図 5 は、305 におけるアンダフロー検出動作を実行するのにエンコーダが使用するプロセス 500 を示す。プロセス 500 は、前述したとおり、復号器入力バッファの状態をシミュレートすることにより、各イメージの最終到着時間、 t_{af} 、および名目除去時間、 $t_{r,n}$ をまず判定する（505 で）。このプロセスは、バッファアンダフロー管理の繰り返しプロセス中に数回、呼び出されることが可能であるので、あるイメージ番号を開始点として受け取り、その与えられた開始イメージからイメージシーケンスを調べることにより留意されたい。明らかに、最初の繰り返しに関して、開始点は、シーケンス内の最初の

50

イメージである。

【 0 1 4 3 】

5 1 0 で、プロセス 5 0 0 は、復号器入力バッファにおける各イメージの最終到着時間を、復号器によるそのイメージの名目除去時間と比べる。名目除去時間より後の最終到着時間を有するイメージが全く存在しないとプロセスが判定した（すなわち、アンダフロー条件は、全く存在しない）場合、プロセスは、終了する。他方、最終到着時間が、名目除去時間より後であるイメージが見つかった場合、プロセスは、アンダフローが存在すると判定し、5 1 5 に進んで、アンダフローセグメントを識別する。

【 0 1 4 4 】

5 1 5 で、プロセス 5 0 0 は、アンダフロー条件が改善し始める（すなわち、 $t_b(n)$ が、一続きのイメージにわたってさらに負にならない）次の大域最小値まで、復号器バッファが、絶えず空にされることが始まるイメージのセグメントとして、アンダフローセグメントを識別する。次いで、プロセス 5 0 0 は、終了する。一部の実施形態では、アンダフローセグメントの始まりは、関連するフレーム間符号化されたイメージセットの開始を示すフレーム内符号化されたイメージである I フレームで始まるように、さらに調整される。アンダフローを生じさせる 1 つまたは複数のセグメントが特定されると、符号器は、そのアンダフローを解消することに取り組む。以下のセクション B が、単一セグメントのケース（すなわち、符号化されたイメージのシーケンス全体が、単一のアンダフローセグメントだけを含む）におけるアンダフローの解消を説明する。その後、セクション C が、マルチセグメントのアンダフローのケースに関するアンダフローの解消を説明する。

【 0 1 4 5 】

< B . 単一セグメントのアンダフローの解消 >

図 4 (a) を参照すると、 $t_b(n)$ 対 n の曲線が、下降する傾きで n 軸と 1 回だけ交差する場合、シーケンス全体の中に 1 つだけのアンダフローセグメントが存在する。アンダフローセグメントは、ゼロ交差点より前の最も近い局所最大値で始まり、ゼロ交差点とそのシーケンスの終りの間の、次の大域最小値で終わる。バッファが、アンダフローから回復する場合、セグメントの終点の後には、上昇する傾きを持つ曲線による別のゼロ交差点が続き得る。

【 0 1 4 6 】

図 6 は、一部の実施形態において、単一のイメージセグメント内でアンダフロー条件を解消するのに符号器が利用する（3 1 5、3 2 0、および 3 2 5 で）プロセス 6 0 0 を示す。6 0 5 で、プロセス 6 0 0 は、バッファに入る入力ビットレートと、セグメントの終りで見られる最長の遅延（例えば、最小 $t_b(n)$ ）との積を計算することにより、ビットの総数を推定して、アンダフローセグメント内の（ B ）を小さくする。

【 0 1 4 7 】

次に、6 1 0 で、プロセス 6 0 0 は、前回の符号化パス（または直近の複数回のパス）からの、現在のセグメント内の平均のマスキングされたフレーム $Q P (A M Q P)$ 、および総ビット数を使用して、そのセグメントに関する所望されるビット数、 $B_T = B - B_p$ を得るための所望される $A M Q P$ を推定する。ただし、 p は、そのセグメントに関するプロセス 6 0 0 の現在の繰り返し回数である。その繰り返しが、特定のセグメントに関するプロセス 6 0 0 の最初の繰り返しである場合、 $A M Q P$ および総ビット数は、3 0 2 で特定された初期（最初の）符号化ソリューションから導出された、そのセグメントに関する $A M Q P$ および総ビット数である。他方、その繰り返しが、プロセス 6 0 0 の最初の繰り返しではない場合、それらのパラメータは、プロセス 6 0 0 の前回のパス、または前の数回のパスにおいて得られた符号化ソリューション、または符号化ソリューション群から導出されることが可能である。

【 0 1 4 8 】

次に、6 1 5 で、プロセス 6 0 0 は、所望される $A M Q P$ を使用して、マスキング強度 $F(n)$ に基づく、平均のマスキングされたフレーム $Q P$ 、 $M Q P(n)$ を変更して、より多くのマスキングを許容することができるイメージが、より多くのビット削減を受け

10

20

30

40

50

るようにする。次に、プロセスは、315で定義されたパラメータに基づき、ビデオセグメントを再符号化する(620で)。次に、プロセスは、セグメントを調べて(625で)、アンダフロー条件が解消されたかどうかを判定する。図4(b)は、プロセス600が、アンダフローセグメントに適用されて、そのセグメントを再符号化した後の、図4(a)のアンダフロー条件の解消を示す。アンダフロー条件が解消されると、プロセスは、終了する。それ以外の場合、プロセスは、605に戻り、総ビットサイズを小さくするように符号化パラメータをさらに調整する。

【0149】

< C . 複数のアンダフローセグメントに対するアンダフロー解消 >

シーケンス内に複数のアンダフローセグメントが存在する場合、セグメントの再符号化により、すべての後続のフレームに関するバッファ充满時間、 $t_b(n)$ が変わる。変更されたバッファ条件を考慮に入れるのに、符号器は、下降する傾きを有する最初のゼロ交差点から(すなわち、最低の n で)始めて、一度に1つのアンダフローセグメントを探索する。

10

【0150】

アンダフローセグメントは、そのゼロ交差点より前の最も近い局所最大値で始まり、そのゼロ交差点と次のゼロ交差点(あるいは、ゼロ交差がもはや存在しない場合、シーケンスの終り)の間の、次の大域最小値で終わる。1つのセグメントを見出した後、符号器は、そのセグメントの終りにおいて $t_b(n)$ を0に設定し、すべての後続のフレームに関してバッファシミュレーションを再び行うことにより、そのセグメントの中のアンダフローを仮想的に除去し、更新されたバッファ充满度を推定する。

20

【0151】

次に、符号器は、変更されたバッファ充满度を使用して、次のセグメントを探索すること続ける。すべてのアンダフローセグメントが、前述したとおり、特定されると、符号器は、単一セグメントのケースと全く同じように、他のセグメントとは独立に、各セグメントに関して、AMQPを導出し、マスキングされたフレームQPを変更する。

【0152】

他の諸実施形態は、異なる形で実施されることも可能であることが、当業者には認識されよう。例えば、一部の実施形態は、復号器の入力バッファのアンダフローを生じさせる複数のセグメントを特定しない。代わりに、一部の実施形態は、前述したバッファシミュレーションを実行して、アンダフローを生じさせる第1のセグメントを特定する。そのようなセグメントを特定した後、それらの実施形態は、そのセグメントを訂正して、そのセグメントの中のアンダフロー条件を正し、その後、訂正された部分の後から符号化を再開する。シーケンスの残りの部分の符号化の後、それらの実施形態は、次のアンダフローセグメントに関して、そのプロセスを繰り返す。

30

【0153】

< D . バッファアンダフロー管理の応用例 >

前述した復号器バッファアンダフローに関する技術は、多数の符号化システムおよび復号化システムに適用される。そのようなシステムのいくつかの実施例を以下に説明する。

【0154】

図7は、ビデオストリーミングサーバ710といくつかのクライアント復号器715~725を接続するネットワーク705を示す。クライアントは、每秒300kbや每秒3Mbなどの、異なる帯域幅を有するリンクを介してネットワーク705に接続される。ビデオストリーミングサーバ710は、符号器730からクライアント復号器715~725への符号化されたビデオイメージのストリーミングを制御している。

40

【0155】

ストリーミングビデオサーバは、ネットワークにおける最も遅い帯域幅(すなわち、每秒300Kb)、および最小のクライアントバッファサイズを使用して、符号化されたビデオイメージをストリーミングすることを決めることが可能である。そのケースでは、ストリーミングサーバ710は、每秒300Kbの目標ビットレートに対して最適化された

50

1つだけの符号化されたイメージセットを必要とする。他方、サーバは、異なる帯域幅、および異なるクライアントバッファ条件に対して最適化された、異なる符号化を生成し、格納することができる。

【0156】

図8は、復号器のアンダフロー管理のための応用例の別の実施例を示す。この実施例では、HD-DVDプレーヤ805が、ビデオ符号器810からの符号化されたビデオデータを格納しているHD-DVD840から、符号化されたビデオイメージを受け取っている。HD-DVDプレーヤ805は、入力バッファ815と、簡明にするために1つのブロック820として示された復号化モジュールセットと、出力バッファ825とを有する。

10

【0157】

プレーヤ805の出力は、TV830またはコンピュータディスプレイ端末装置835などのディスプレイデバイスに送られる。HD-DVDプレーヤは、非常に高い帯域幅、例えば、毎秒29.4Mbを有することが可能である。ディスプレイデバイス上で高品質のイメージを維持するため、符号器は、ビデオイメージが次のように符号化されることを確実にする。すなわち、イメージシーケンスにおいて、復号器入力バッファに時間どおり送り届けられることが不可能なほど大きなセグメントが存在しないようにする。

【0158】

[VI. コンピュータシステム]

図9は、本発明の一実施形態が実施されるコンピュータシステムを提示する。コンピュータシステム900は、バス905と、プロセッサ910と、システムメモリ915と、読み取り専用メモリ920と、永久記憶デバイス925と、入力デバイス群930と、出力デバイス群935とを含む。バス905は、コンピュータシステム900の多数の内部デバイスを通信するように接続するすべてのシステムバス、周辺バス、およびチップセットバスをひとまとめにして表す。例えば、バス905は、プロセッサ910を、読み取り専用メモリ920、システムメモリ915、および永久記憶デバイス925と通信するように接続する。

20

【0159】

以上の様々なメモリユニットから、プロセッサ910は、本発明のプロセスを実行するために、実行されるべき命令、および処理されるべきデータを取り出す。読み取り専用メモリ(ROM)920が、プロセッサ910、ならびにコンピュータシステムの他のモジュール群によって必要とされている静的データおよび命令を格納する。

30

【0160】

他方、永久記憶デバイス925は、読み取り-書き込みメモリデバイスである。このデバイスは、コンピュータシステム900がオフである場合でも、命令およびデータを格納する不揮発性メモリユニットである。本発明の一部の実施形態は、大容量記憶デバイス(磁気ディスクまたは光ディスク、ならびにその対応するディスクドライブなどの)を永久記憶デバイス925として使用する。

【0161】

他の諸実施形態は、取り外し可能な記憶デバイス(フロッピー(登録商標)ディスクまたはzip(登録商標)ディスク、ならびにその対応するディスクドライブなど)を永久記憶デバイスとして使用する。永久記憶デバイス925と同様に、システムメモリ915は、読み取り書き込みメモリデバイスである。しかし、記憶デバイス925とは異なり、システムメモリは、ランダムアクセスメモリなどの、揮発性読み取り-書き込みメモリである。システムメモリは、プロセッサが、ランタイムに必要とする命令およびデータの一部を格納する。一部の実施形態では、本発明のプロセスは、システムメモリ915、永久記憶デバイス925、および、読み取り専用メモリ920のうちの少なくとも1つの中に格納される。

40

【0162】

また、バス905は、入力デバイス群930および出力デバイス群935にも接続する

50

。入力デバイス群は、ユーザが、コンピュータシステムに対して情報を通信し、コマンドを選択することを可能にする。入力デバイス群 930 には、英数字キーボードおよびカーソルコントローラが含まれる。出力デバイス群 935 は、コンピュータシステムによって生成されたイメージを表示する。出力デバイス群には、プリンタ、ならびに陰極線管 (CRT) または液晶ディスプレイ (LCD) などのディスプレイデバイスが含まれる。

【0163】

最後に、図 9 に示されるとおり、バス 905 は、コンピュータ 900 を、ネットワークアダプタ (図示せず) を介してネットワーク 965 にも接続する。このようにして、コンピュータは、コンピュータのネットワーク (ローカルエリアネットワーク (「LAN」)、ワイドエリアネットワーク (「WAN」)、またはイントラネットなど)、またはネットワークのネットワーク (インターネットなど) の一部であることが可能である。コンピュータシステム 900 のコンポーネントのいずれか、またはすべてが、本発明に関連して使用されることが可能である。しかし、他の任意のシステム構成が、本発明に関連して使用されることも可能であることが当業者には認識されよう。

【0164】

本発明を多数の特定の詳細に関連して説明してきたが、本発明は、本発明の趣旨を逸脱することなく、他の特定の形態で実施されることも可能であることが当業者には認識されよう。例えば、復号器入力バッファをシミュレートする H264 の方法を使用することの代わりに、バッファサイズ、バッファの中のイメージの到着時間および除去時間、ならびにイメージの復号化時間および表示時間を考慮する他のシミュレーション方法を使用して

【0165】

前述したいくつかの実施形態は、平均の除去された SAD を計算して、マクロブロック内のイメージ変化の指示を得た。しかし、他の諸実施形態は、異なる形でイメージ変化を識別することが可能である。例えば、一部の実施形態は、マクロブロックのピクセルの予測されるイメージ値を予測することが可能である。それらの実施形態は、次に、その予測値をマクロブロックのピクセルの輝度値から引き、それらの減算の絶対値を合計することにより、マクロブロック SAD を生成する。一部の実施形態では、予測値は、そのマクロブロック内のピクセルの値だけでなく、近隣のマクロブロックの 1 つまたは複数のマクロブロックの中のピクセルの値にも基づく。

【0166】

また、前述した諸実施形態は、導出された空間的マスキング値および時間的マスキング値を直接使用する。他の諸実施形態は、連続する空間的マスキング値および連続する時間的マスキング値のうちの少なくとも一方に平滑化フィルタリングを適用してから、それらの値を使用することを、ビデオイメージを介してそれらの値の一般的な傾向を選び出すために行う。このため、本発明は、以上の例示的な詳細によって限定されないことが、当業者には理解されよう。

【図面の簡単な説明】

【0167】

【図 1】本発明の一部の実施形態の符号化方法を概念的に示すプロセスを示す図である。

【図 2】一部の実施形態のコーデックシステムを概念的に示す図である。

【図 3】一部の実施形態の符号化プロセスを示す流れ図である。

【図 4】(a) は、一部の実施形態におけるアンダフロー条件を示す、イメージ番号に対するイメージの名目除去時間と最終到着時間の差のプロットである。(b) は、アンダフロー条件が解消された後の、図 4 a に示されたのと同じイメージに関する、イメージ番号に対するイメージの名目除去時間と最終到着時間の差のプロットである。

【図 5】一部の実施形態において、アンダフロー検出を実行するのに符号器が使用するプロセスを示す図である。

【図 6】一部の実施形態において、イメージの単一のセグメントの中でアンダフロー条件を解消するのに符号器が利用するプロセスを示す図である。

【図 7】ビデオストリーミングアプリケーションにおけるバッファアンダフロー管理の応用例を示す図である。

【図 8】HD-DVDシステムにおけるバッファアンダフロー管理の応用例を示す図である。

【図 9】本発明の一実施形態が実施されるコンピュータシステムを示す図である。

【図 1】

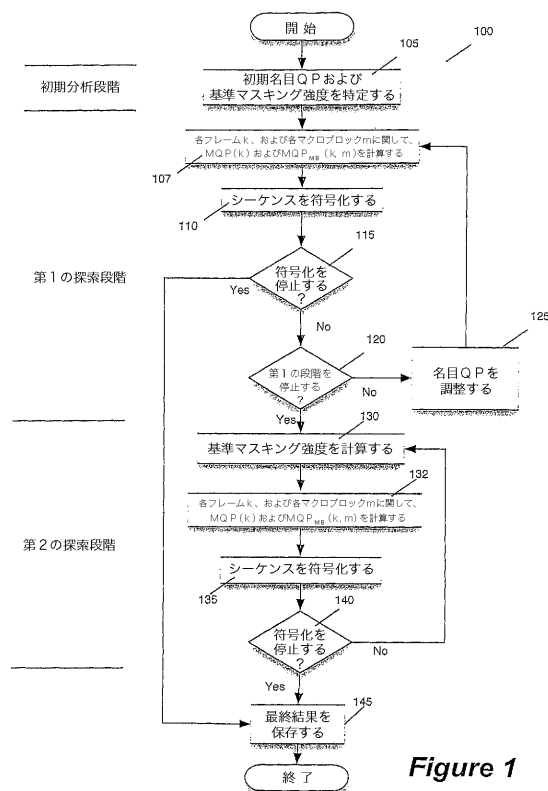


Figure 1

【図 2】

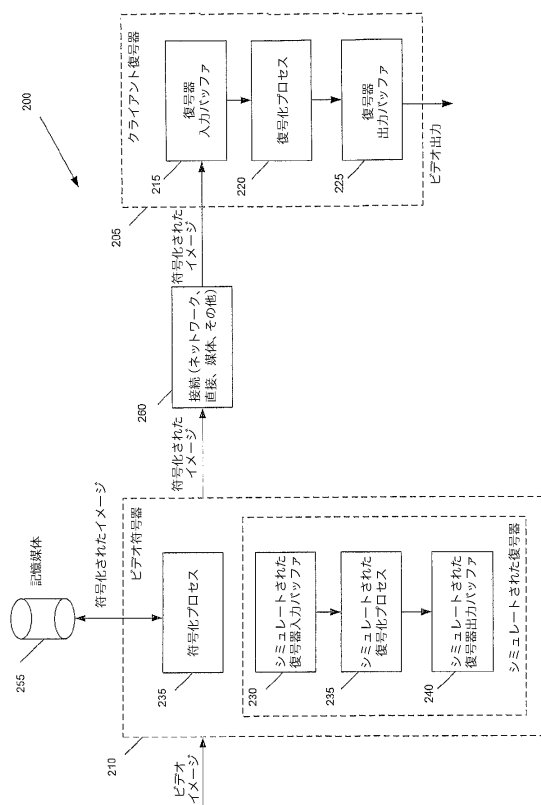


Figure 2

【図 3】

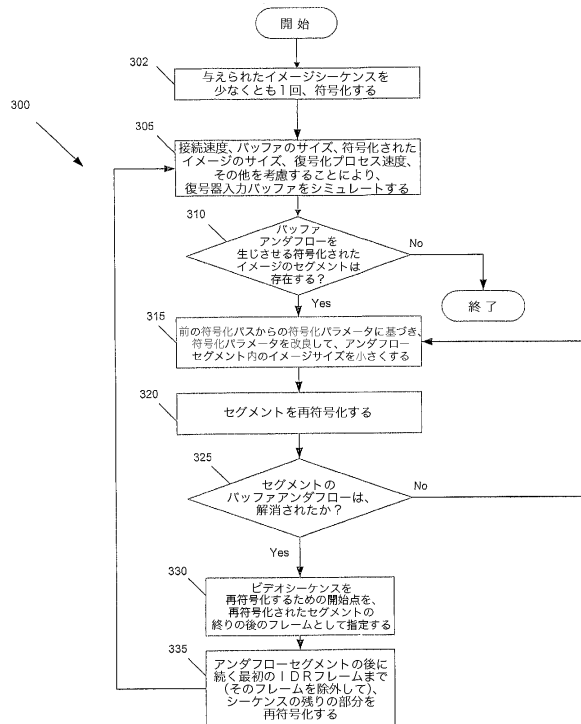


Figure 3

【図 4】

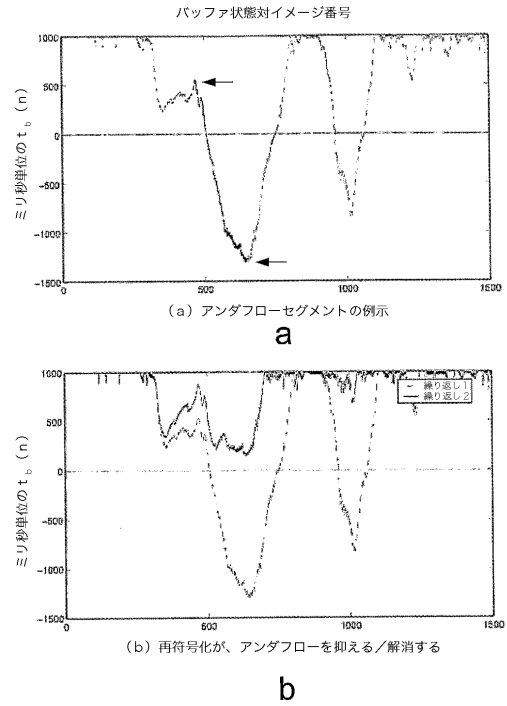


Figure 4

【図 5】

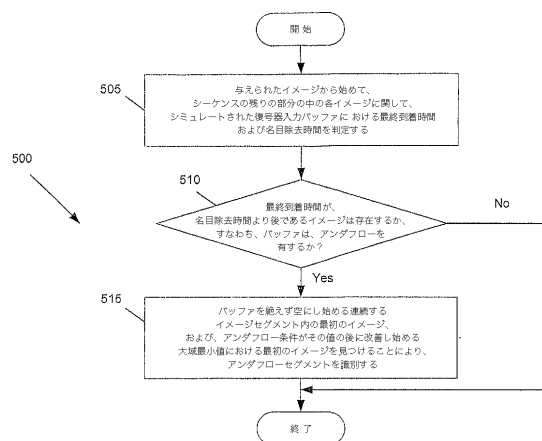


Figure 5

【図 6】

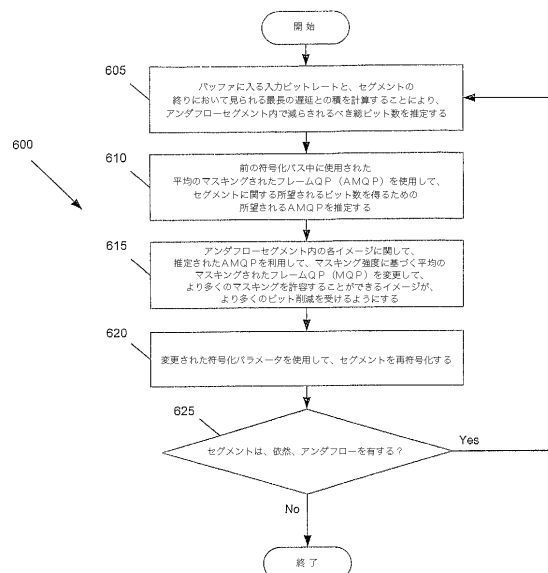


Figure 6

【図 7】

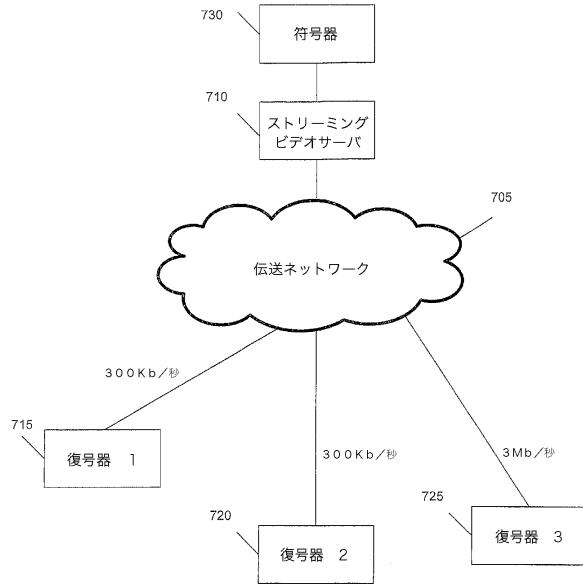


Figure 7

【図 8】

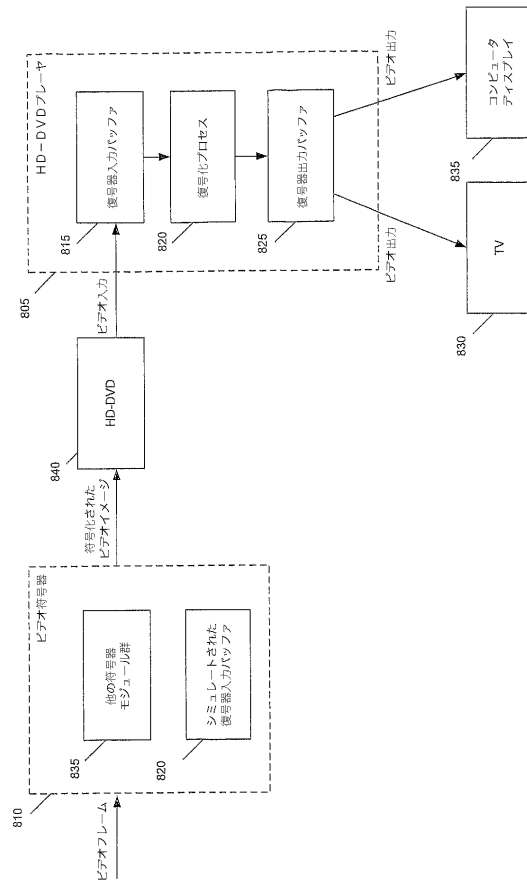


Figure 8

【図 9】

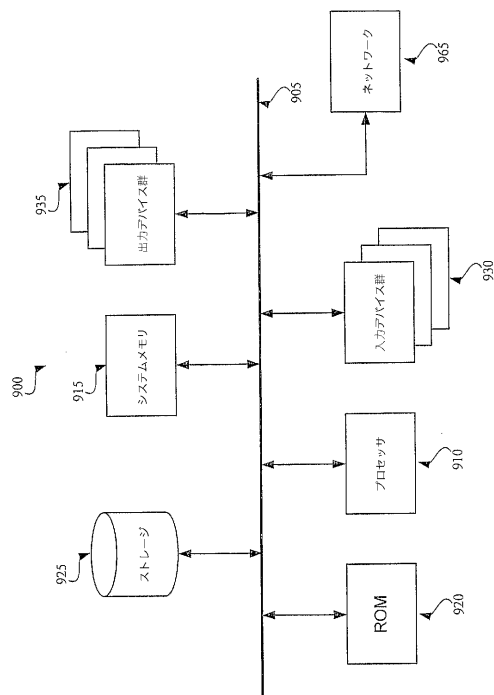


Figure 9

フロントページの続き

(31)優先権主張番号 11/118,616

(32)優先日 平成17年4月28日(2005.4.28)

(33)優先権主張国 米国(US)

(31)優先権主張番号 11/118,604

(32)優先日 平成17年4月28日(2005.4.28)

(33)優先権主張国 米国(US)

(74)代理人 100134474

弁理士 坂田 恭弘

(72)発明者 トン, シン

アメリカ合衆国 カリフォルニア州 9 4 0 4 3 , マウンテン ビュー, キャンベル ドライブ 1 6 9

(72)発明者 ウー, シー - ジュン

アメリカ合衆国 カリフォルニア州 9 4 0 8 6 , サニーベール, コルテス ドライブ 2 0 0 3 , 6

(72)発明者 プン, トーマス

アメリカ合衆国 カリフォルニア州 9 5 1 3 4 , サンノゼ, デスカンソ ドライブ 8 0 , 1 3 2 2

(72)発明者 ドゥミトラ, アドリアナ

アメリカ合衆国 カリフォルニア州 9 4 0 8 7 , サニーベール, ダブリュ・エル カミノ レアル 2 5 0 , アpartment 2 4 0 3

(72)発明者 ハスケル, バリン

アメリカ合衆国 カリフォルニア州 9 4 0 4 0 , マウンテン ビュー, フェアブルック ドライブ 1 1 9 0

(72)発明者 ノーマイル, ジム

アメリカ合衆国 カリフォルニア州 9 4 0 2 4 , ロスアルトス, エスパーグ ロード 1 7 2 5

審査官 金田 孝之

(56)参考文献 特表2 0 0 1 - 5 0 1 4 2 9 (J P , A)

特表2 0 0 1 - 5 2 0 8 5 4 (J P , A)

特開平0 8 - 2 7 5 1 7 6 (J P , A)

特開2 0 0 4 - 1 6 6 1 2 8 (J P , A)

米国特許第0 6 1 9 2 0 7 5 (U S , B 1)

(58)調査した分野(Int.Cl., D B 名)

H04N 7/12

H04N 7/26

H04N 7/30- 7/32

H03M 3/00-11/00