

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第4955686号
(P4955686)

(45) 発行日 平成24年6月20日 (2012. 6. 20)

(24) 登録日 平成24年3月23日 (2012. 3. 23)

(51) Int. Cl.	F I
HO 4 N 7/32 (2006. 01)	HO 4 N 7/137 Z
HO 4 N 7/30 (2006. 01)	HO 4 N 7/133 Z

請求項の数 14 (全 32 頁)

(21) 出願番号	特願2008-531271 (P2008-531271)	(73) 特許権者	500046438
(86) (22) 出願日	平成18年9月13日 (2006. 9. 13)		マイクロソフト コーポレーション
(65) 公表番号	特表2009-508445 (P2009-508445A)		アメリカ合衆国 ワシントン州 9805
(43) 公表日	平成21年2月26日 (2009. 2. 26)		2-6399 レッドモンド ワン マイ
(86) 国際出願番号	PCT/US2006/035640		クロソフト ウェイ
(87) 国際公開番号	W02007/033242	(74) 代理人	100077481
(87) 国際公開日	平成19年3月22日 (2007. 3. 22)		弁理士 谷 義一
審査請求日	平成21年9月7日 (2009. 9. 7)	(74) 代理人	100088915
(31) 優先権主張番号	11/226, 590		弁理士 阿部 和夫
(32) 優先日	平成17年9月14日 (2005. 9. 14)	(72) 発明者	グオビン シェン
(33) 優先権主張国	米国 (US)		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

最終頁に続く

(54) 【発明の名称】 効率的な統合デジタルビデオトランスコーディング

(57) 【特許請求の範囲】

【請求項 1】

コンピュータ実装方法であって、
 統合トランスコーダにより、符号化されたビットストリームを受信するステップと、
 前記統合トランスコーダにより、前記符号化されたビットストリームをトランスコード
 するステップであって、

前記統合トランスコーダにより、前記符号化されたビットストリームを部分的に復号
 化して中間データストリーム生成し、前記符号化されたビットストリームは前記第 1 のメ
 ディア形式に関連付けられている圧縮技術の第 1 のセットにより符号化され、

前記統合トランスコーダにより、圧縮技術の第 2 のセットを使用して前記中間データ
 ストリームを符号化してトランスコードされたビットストリームを生成し、圧縮技術の前
 記第 2 のセットは前記第 2 のメディア形式に対応するように、トランスコードするステッ
 プと

を有し、前記統合トランスコーダは、

前記部分的な復号化に関連する第 1 の変換と前記符号化に関連する第 2 の変換とを、単
 一のスケーリング行列として組み合わせ、前記単一のスケーリング行列をスカラー乗算に
 よって置き換えることによって、前記部分的な復号化と前記符号化とを統合し、

前記トランスコードするステップは、残誤差補償を実行し、

閾値ベースのドリフト制御メカニズムに応答して、残誤差バッファへのブロック再量
 子化誤差累積を動的にオンまたはオフにするステップと、

10

20

ブロックアクティビティの評価にตอบสนองして、残誤差バッファ内の累積された誤差の動き補償を動的にオンまたはオフにするステップと、

前記部分的復号化からの動き補償累積残誤差と再構成された残との総和に基づく決定にตอบสนองして、早期検出を動的にオンまたはオフにしてブロックが符号化されるかどうかを決定するステップと

のうちの1つまたは複数を有することを特徴とするコンピュータ実装方法。

【請求項2】

前記第1のメディア形式は、MPEG-2であり、前記第2のメディア形式は、WMVであることを特徴とする請求項1に記載のコンピュータ実装方法。

【請求項3】

前記第1のメディア形式は、MPEG-2であり、前記第2のメディア形式は、MPEG-4であることを特徴とする請求項1に記載のコンピュータ実装方法。

【請求項4】

前記統合トランスコーダは、適切な誤差補償を介して誤差伝搬を防ぐ閉ループトランスコーダであることを特徴とする請求項1に記載のコンピュータ実装方法。

【請求項5】

前記統合トランスコーダは、誤差伝搬を防がない開ループトランスコーダであることを特徴とする請求項1に記載のコンピュータ実装方法。

【請求項6】

前記統合トランスコーダは、スケーリングコンポーネントにおいて第1および第2のメディア形式にそれぞれ関連付けられている第1および第2の変換をマージすることを特徴とする請求項1に記載のコンピュータ実装方法。

【請求項7】

トランスコードするステップは、さらに、前記トランスコーディングの品質または速度をそれぞれ高めるために残誤差補償に関連付けられている1つまたは複数のオペレーションを動的にオンまたはオフにするステップを有することを特徴とする請求項1に記載のコンピュータ実装方法。

【請求項8】

トランスコードするステップは、残誤差補償を実行し、トランスコードするステップは、さらに、

早期基準フレームが処理されていると判定したことにตอบสนองして、1つまたは複数のオペレーションを動的に調節して前記早期基準フレームを高い品質でかつ遅い速度でトランスコードするステップを有し、

前記オペレーションは、ブロック再量子化誤差累積、累積誤差の動き補償、および特定のブロックが符号化されているかどうかを判定する早期検出を有すること

を特徴とする請求項1に記載のコンピュータ実装方法。

【請求項9】

部分的に復号化するステップは、さらに、

すべてのフレームについて入力フレームサイズに比例する符号化フレームサイズでレート制御を実行するステップを有し、

前記レート制御では目標フレームサイズと実際の結果として得られるフレームサイズとの間の累積された差を継続的に補償すること

を特徴とする請求項1に記載のコンピュータ実装方法。

【請求項10】

前記方法は、さらに、

高い離散コサイン変換(DCT)係数を破棄して、DCT領域内の前記中間データストリームを縮小するステップを有し、

明示的なピクセル領域縮小は、実行されないこと

を特徴とする請求項1に記載のコンピュータ実装方法。

【請求項11】

トランスコードするステップは、さらに、混合ブロック処理および対応する復号化ループと無関係に前記符号化されたビットストリームから完全な分解能の画像の再構成により前記符号化されたビットストリームのブロックを縮小するステップを有することを特徴とする請求項 1 に記載のコンピュータ実装方法。

【請求項 1 2】

トランスコードするステップは、さらに、任意縮小目標比を得るために 2 段階縮小オペレーションを実行するステップを有することを特徴とする請求項 1 に記載のコンピュータ実装方法。

【請求項 1 3】

前記 2 段階縮小オペレーションは、さらに、
D C T 領域復号化ループ内で第 1 段階縮小オペレーションを実行するステップと、
ピクセル領域内の前記復号化ループの外部で第 2 段階縮小オペレーションを実行するステップと

を有することを特徴とする請求項 1 2 に記載のコンピュータ実装方法。

【請求項 1 4】

前記 2 段階縮小オペレーションは、さらに、
第 1 段階縮小オペレーションを実行して、前記任意縮小目標比の間である第 1 の結果を得るステップと、

前記第 1 の結果を復号化ループ内に入力するステップと、
空間領域内で第 2 段階縮小オペレーションを実行して、前記任意縮小目標比を得るステップと

を有することを特徴とする請求項 1 2 に記載のコンピュータ実装方法。

【発明の詳細な説明】

【技術分野】

【0 0 0 1】

本発明は、効率的な統合デジタルビデオトランスコーディングに関する。

【背景技術】

【0 0 0 2】

デジタルビデオコンテンツは、典型的には、特定のデータ形式を対象として生成される。ビデオデータ形式は、一般に、特定のビットレート、空間分解能、フレームレートなど
有する、特定のビデオ符号化標準または専用符号化アルゴリズムに準拠する。このような
符号化標準としては、M P E G - 2 および W I N D O W S (登録商標) M e d i a V
i d e o (W M V) がある。ほとんどの既存のデジタルビデオコンテンツは、M P E G -
2 データ形式に従って符号化されている。W M V は、ストリーミング技術分野における公
認コーデックとして広く受け入れられ、インターネット上に広く展開され、H D - D V D
コンソーシアムにより採用され、現在は、S M P T E 標準とみなされている。ビデオ符号
化標準が異なれば、圧縮能力および画質も変わる。

【0 0 0 3】

トランスコーディングとは、ある圧縮ビットストリームを他の圧縮ビットストリームに
変換する一般的プロセスのことである。多くの場合、デバイスの能力と配布ネットワーク
とを一致させるために、M P E G - 2 から W M V、H. 2 6 4、さらにはスケーラブル形
式への変換など、ある符号化形式から他の符号化形式への変換をビットストリームに対し
実行することが望ましい。トランスコーディングは、無線チャネル上での伝送用のビット
ストリームの V C R 類似機能、ロゴ挿入、または拡張誤り耐性能力能力などの何らかの特
定の機能を実現するためにも使用できる。

【発明の開示】

【発明が解決しようとする課題】

【0 0 0 4】

図 1 は、従来の C a s c a d e d P i x e l - D o m a i n T r a n s c o d e r
(C P D T) システムを示しており、このシステムは、入力ビットストリームを復号化す

10

20

30

40

50

るフロントエンドデコーダと、異なる符号化パラメータセットまたは新しい形式の新しいビットストリームを生成するエンコーダとをカスケード接続したものである。この従来のトランスコーディングアーキテクチャの欠点の1つは、実用的展開に関して、典型的にはその複雑さが障害となっている点である。その結果、図1のCPDTトランスコーディングアーキテクチャは、典型的には、改善されたスキームの性能ベンチマークとして使用される。

【0005】

図2は、従来のカスケードDCT領域トランスコーダ(CDDT)アーキテクチャを示しており、これは、図1のCPDTアーキテクチャを簡素化したものである。図2のシステムは、機能を空間的／時間的分解能のダウンスケーリングおよび符号化パラメータ変更10に制限する。CDDTを使用すると、図1のCPDTトランスコーダにより実装されるDCT/I DCTプロセスが不要になる。それでも、CDDTは、DCT領域内でMCを実行するが、これは、典型的には、多大な時間を必要とし、また大量の計算を必要とするオペレーションである。これは、DCTブロックがMCブロックと重なり合うことが多いためである。その結果、CDDTアーキテクチャは、典型的には、DCT領域内でMCを実行するために複雑な、大量の計算を必要とする浮動小数点行列オペレーションを適用する必要がある。さらに、動きベクトル(MV)の微調整は、典型的には、CDDTアーキテクチャを使用したのでは実行不可能である。

【課題を解決するための手段】

【0006】

ここでは、後でさらに説明される概念の抜粋を簡素化された形式で紹介する。ここでは、特許請求された主題の重要な特徴または本質的な特徴を識別することを意図されていないし、また特許請求された主題の範囲を決定する補助手段として使用されることも意図されていない。

【0007】

上記に関して、効率的な統合デジタルビデオトランスコーディングについて説明する。一態様では、統合トランスコーダは、符号化されたビットストリームを受信する。統合トランスコーダは、第1のメディアデータ形式に関連付けられている圧縮技術の第1のセットに基づいて符号化されたビットストリームを部分的に復号化することにより符号化されたビットストリームをトランスコードする。この復号化オペレーションは、中間データストリームを生成する。次いで、統合トランスコーダは、第2のメディアデータ形式に関連付けられている圧縮技術の第2のセットを使用して中間データストリームを符号化する。圧縮技術の第1のセットおよび第2のセットは同じでない。

【発明を実施するための最良の形態】

【0008】

図中、コンポーネントの参照番号の一番左の数字は、そのコンポーネントが最初に出現する特定の図面であることを示す。

【0009】

説明および例示のために、図の中で色を使用して、以下の規約を示す。青色の実線矢印は、実際の、または残りの画像データに関するピクセル領域信号を表す。赤色の実線の矢印は、DCT領域内の信号を表す。オレンジ色の破線の矢印は、動き情報を表す。

【0010】

[概要]

効率的なデジタルビデオトランスコーディングのシステムおよび方法を図4から14を参照して以下に説明する。これらのシステムおよび方法では、入力ビットストリーム内の情報を使用することで、アプリケーションに、誤差伝搬の動的な制御、およびそれによる、ビデオビットストリームトランスコーディングの速度と品質の選択的な制御を実行させることができる。この選択的制御により、アプリケーションは、閉ループトランスコーディング(高速トランスコーディングプロファイル)から開ループ(高品質トランスコーディングプロファイル)へのトランスコーディング方式の継ぎ目のないスケーリングが可能50

になる。従来のトランスコーディングアーキテクチャ（例えば、図 1 の C P D T および図 2 の C D D T）とは対照的に、効率的なデジタルビデオトランスコーディングのアーキテクチャは、統合され、これにより異なる種類の離散コサイン変換（D C T）または D C T 類似変換を組み合わせることで 1 つのトランスコーディングモジュールが形成される。効率的なビデオトランスコーディングを行うシステムおよび方法は、高速ルックアップテーブルで再量子化処理を実装し、三重閾値アルゴリズムを使用して精細ドリフト制御メカニズムを実現する。

【0011】

一実装では、効率的なデジタルビデオトランスコーディングで、ビットストリームデータ形式（例えば、M P E G - 2 など）を W M V にトランスコードする場合、高品質プロファイルトランスコーディングオペレーションは、W M V の高度な符号化機能をサポートする。一実装では、高速プロファイルトランスコーディングオペレーションは、任意分解能 2 段階的縮小（例えば、高精細度（H D）から標準精細度（S D）にトランスコードする場合）を実行する。このような 2 段階的縮小オペレーションでは、縮小比の一部は、D C T 領域内で効率よく得られるが、縮小比オペレーションは、実質的に分解能を低くして空間領域内において実行される。

【0012】

[例示的な概念的基礎]

図 3 は、M P E G - 2 を W M V に変換する例示的な非統合カスケードピクセル領域トランスコーディング分割アーキテクチャ 300 を示す。この分割アーキテクチャは、復号化および符号化のオペレーションを別々のモジュールがそれぞれ実行するため、統合されない。図 3 の分割アーキテクチャは、効率的なデジタルビデオトランスコーディングのための統合されたシステムおよび方法について続いて説明するための概念的基礎をなす。表 1 は、図 3 の説明のため、シンボルおよびそれぞれの意味をまとめたものである。

【0013】

10

20

【表 1】

表 1

e_{i+1}	トランスコーダの一部を符号化することにより符号化されるフレーム (i+1) の誤差。
$\hat{B}_i$	元の分解能でMPEG-2デコーダにより再構成されたフレーム i。
$\tilde{B}_i$	元の分解能でエンコーダにより再構成されたフレーム i。
$\hat{b}_i$	低減された分解能でMPEG-2デコーダにより再構成されたフレーム i。
$\tilde{b}_i$	低減された分解能でエンコーダにより再構成されたフレーム i。
$\hat{r}_{i+1}$	MPEG-2デコーダにより再構成されたフレーム (i+1) の残余。
$\tilde{r}_{i+1}$	エンコーダにより再構成されたフレーム (i+1) の残余。
$MC_{mp2}(B, mv)$	16×16ブロックベースの、MPEG-2デコーダによる参照画像Bおよび動きベクトルmvを使用する動き補償予測。
$MC_{vc1}(B, mv)$	16×16または8×8ブロックベースのいずれかの、トランスコーダ308 (エンコーダ) による参照画像Bおよび動きベクトルmvを使用する動き補償予測。
$MC'_{mp2}(b, mv)$	8×8以下のブロックベースの、MPEG-2フィルタリングを使用する、低減された分解能基準bおよび動きベクトルmvによる動き補償予測。
$MC'_{vc1}(b, mv)$	8×8以下のブロックベースの、トランスコーダ308フィルタリングを使用する、低減された分解能基準Bおよび動きベクトルmvによる動き補償予測。
MV	元のフレーム分解能内の動きベクトル。
mv	低減されたフレーム分解能内の動きベクトル。

【0014】

説明および例示のために、システム300は、ビットレートの低減、空間分解能の低減、およびその組合せとともにMPEG-2からWMVへのトランスコーディングに関して説明される。多くの既存のデジタルビデオコンテンツは、MPEG-2データ形式により符号化される。WMVは、ストリーミング技術分野における公認コーデックとして広く受け入れられ、インターネット上に広く展開され、HD-DVDコンソーシアムにより採用され、現在は、SMPTE標準とみなされている。

【0015】

MPEG-2およびWMVは、圧縮および画質に関する様々な能力を備える。例えば、MPEG-2およびWMVによりそれぞれ使用される圧縮技術は、非常に異なる。例えば、動きベクトル(MV)精度および動き補償(MC)フィルタリング技術は、異なる。MPEG-2では、動き精度は、最大でも1/2ピクセル精度でしかなく、また補間法は、双一次フィルタリングである。対照的に、WMVでは、動き精度は、最大1/4ピクセル精度まで高められ、2つの補間法、つまり、双一次フィルタリングおよび双三次フィルタリングがサポートされる。さらに、フィルタリングプロセスに関わる丸め制御パラメータがある。WMVを使用すると、MPEG-2ビットレートに比べて、ビデオビットレートは最大50%まで低減され、しかも画質低下は無視できるくらい小さい。

【0016】

他の実施例では、MPEG-2およびWMVにより使用される変換は、異なる。例えば、MPEG-2は、標準DCT/I DCTを使用し、変換サイズは、 8×8 に固定される。対照的に、WMVでは、変換カーネル行列の成分がすべての小さな整数である整数変換（VC1-T）を使用する。さらに、変換サイズは、 8×8 、 8×4 、 4×8 、および 4×4 のいずれかを使用してブロックからブロックへのWMVを使用することで変更することができる。MPEG-2は、フレームレベルの最適化をサポートしない。その一方で、WMVは、性能最適化のため様々なフレームレベル構文をサポートする。WMVは、強度補正、範囲低減、および動的分解能変更などの他の多くの高度な符号化機能をサポートする。

【0017】

10

上記を考慮し、分解能の変更なしでビットレート低減を行うために、図3に示されているMPEG-2デコーダとWMVエンコーダのブリッジとなるフィルタリングプロセスは、全域通過フィルタである（つまり、効果がない）。したがって、フレーム（ $i+1$ ）に対するエンコーダへの入力は、

【0018】

【数1】

$$e_{i+1} = \hat{f}_{i+1} + MC_{mp2}(\hat{B}_i, MV_{mp2}) - MC_{vc1}(\tilde{B}_i, MV_{vc1}) \quad (1)$$

【0019】

で表される。

20

【0020】

この実装では、図3のWMV符号化効率は、より精細な動き精度から得られる。WMVでは、MPEG-2の場合のように共通の $1/2$ ピクセル精度のほかに $1/4$ ピクセル動き精度が許される。さらに、WMVでは、MCフィルタリングに対する双三次補間と呼ばれるより適切な、より複雑な補間が可能である。双一次補間は、 $1/2$ ピクセルMCに対するMCモジュール（ MC_{mp2} ）においてMPEG-2に使用される。双一次補間法は、MPEG-2双一次補間が丸め制御を行わないことを除き、WMVで使用されるのと似た方法である。高速化するために、 $1/2$ ピクセルの動き精度をエンコーダ部分に実現することができる。これに対する理由の1つは、絶対的なオリジナルフレームの欠如である（つまり、ビットストリーム入力データ（BS-IN）は、すでに圧縮されている）。したがって、この実施例では、より正確な、しかも意味のある動きベクトルを得ることは難しい。その一方で、MPEG-2デコーダから得られる動き情報（つまり、 $MV_{vc1} = MV_{mp2}$ ）は、直接再利用できる。分解能変更はないため、この仮定ではMV精度の低下はない。エンコーダが、さらに、双一次補間を使用し、丸め制御パラメータを常にオフに強制するように制約されている場合、動き補償が線形オペレーションであるとの妥当な仮定の下で、丸め誤差を無視すると（つまり、 $MC_{vc1} = MC_{mp2}$ ）、式1は、

30

【0021】

【数2】

$$e_{i+1} = \hat{f}_{i+1} + MC_{mp2}(\hat{B}_i - \tilde{B}_i, MV_{mp2}) \quad (2)$$

40

【0022】

のように簡略化される。式2により、図3の基準C P D Tトランスコーダを簡素化できる。このような簡素化されたアーキテクチャは、図5を参照しつつ以下で説明される。簡素化されたアーキテクチャについて説明する前に、効率的なデジタルビデオトランスコーディングを行う例示的なシステムについて、最初に説明する。

【0023】

〔例示的なシステム〕

必要というわけではないが、パーソナルコンピュータなどのコンピューティングデバイスによって実行されるコンピュータープログラム命令の一般的背景状況において効率的なデジタルビデオトランスコーディングについて説明する。一般に、プログラムモジュール

50

は、特定のタスクを実行する、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、オブジェクト、コンポーネント、データ構造などを含む。これらのシステムおよび方法は、前記の文脈において説明されているが、以下で説明される活動およびオペレーションは、ハードウェアでも実装することができる。

【0024】

図4は、効率的なデジタルビデオトランスコーディングを行う例示的なシステム400を示している。この実装では、システム400のオペレーションは、ハイブリッドDCTおよびブロックベース動き補償(MC)ビデオ符号化方式に関して説明され、多くのビデオ符号化標準および専用形式がこの方式に基づく。より具体的には、システム400は、MPEG-2からWMVにトランスコードするために使用されるアーキテクチャ、コンポーネント、およびオペレーションと併せて説明される。しかし、MPEG-2をWMVにトランスコードするためのシステム400により具現化されるスケーラブルな複雑度および効率のトランスコーディングについて説明されるアーキテクチャ、コンポーネント、およびオペレーションは、さらに、MPEG-2およびWMVに加えて他のビットストリームデータ形式変換に適用することもできることは理解されうる。例えば、一実装では、システム400を使用して、MPEG-2ビットストリームをMPEG-4ビットストリームにトランスコードし、MPEG-4ビットストリームデータをWMVビットストリームデータにトランスコードするといったことを行う。このような代替実施形態では、システム400の以下で説明されるトランスコーディングアーキテクチャ(コンポーネントおよびそれに関連するオペレーションを含む)では、復号化され、符号化され、それぞれのデータ形式であるビットストリームデータの種別を考慮する。

【0025】

この実装では、システム400は、汎用コンピューティングデバイス402を含む。コンピューティングデバイス402は、パーソナルコンピュータ、ラップトップ、サーバ、ハンドヘルドまたはモバイルコンピューティングデバイスなどのどれかのタイプのコンピューティングデバイスを表す。コンピューティングデバイス402は、第1のデータ形式(例えば、MPEG-2)で符号化されたビットストリームを異なるデータ形式(例えば、WMV)で符号化されたビットストリームにトランスコードするプログラムモジュール404およびプログラムデータ406を備える。プログラムモジュール404は、例えば、効率的なデジタルビデオトランスコーディングモジュール408(「トランスコーディングモジュール408」)および他のプログラムモジュール410を含む。トランスコーディングモジュール408は、符号化されたメディア412(例えば、MPEG-2メディア)を、トランスコードされたメディア414(例えば、WMVメディア)にトランスコードする。他のプログラムモジュール410としては、例えば、オペレーティングシステムおよびトランスコーディングモジュール408のビデオビットストリームトランスコーディング機能を利用するアプリケーションなどがある。一実装では、このアプリケーションは、オペレーティングシステムの一部である。一実装では、トランスコーディングモジュール408は、そのトランスコーディング機能を、アプリケーションプログラミングインターフェース(API)416を介してアプリケーションに公開する。

【0026】

[高速プロファイルトランスコーディング]

図5は、誤差伝搬のない例示的な簡素化された統合閉ループカスケードピクセル領域トランスコーダを示す。説明および例示のために、図5のコンポーネントは、図4のコンポーネント参照しつつ説明される。例えば、図5のアーキテクチャは、図4のトランスコーディングモジュール408の例示的なアーキテクチャの一実装を表す。図3のアーキテクチャと比較して図5のアーキテクチャを参照し、これが、独立したエンコーダおよびデコーダコンポーネントを持たない統合アーキテクチャであることに留意されたい。さらに、MV微調整動き推定モジュールが、MPEG-2デコーダ内のMCから取り除かれることに留意されたい。さらに、WMVエンコーダ内のMCは、累積再量子化誤差に作用するMCにマージされる。このようにして、図5のトランスコーディングアーキテクチャは、プ

ログレッシブおよびインターレースのビデオデータ形式の高速トランスコーディングに関わる計算の複雑さを大幅に低減する。

【0027】

WMV変換は、MPEG-2で使用されていると異なることに留意されたい。MPEG-2では、標準浮動小数点DCT／IDCTが使用されるが、WMVでは、エネルギーパッキング特性がDCTに類似している整数変換が採用される。その結果、MPEG-2デコーダのIDCTとWMVエンコーダのVC1-Tは、互いに相殺することはない。WMVの整数変換は、DCT／IDCTの整数実装と異なる。WMV整数変換は、すべての変換係数が小さな整数であるように慎重に設計される。従来のトランスコーダは、第1の変換に関して符号化されたビットストリームを第1の変換と同じではない第2の変換にトランスコードするように統合されていない。

10

【0028】

式3は、 8×8 VC1-Tの例示的な変換行列を示している。

【0029】

【数3】

$$T_8 = \begin{bmatrix} 12 & 12 & 12 & 12 & 12 & 12 & 12 & 12 \\ 16 & 15 & 9 & 4 & -4 & -9 & -15 & -16 \\ 16 & 6 & -6 & -16 & -16 & -6 & 6 & 16 \\ 15 & -4 & -16 & -9 & 9 & 16 & 4 & -15 \\ 12 & -12 & -12 & 12 & 12 & -12 & -12 & 12 \\ 9 & -12 & 4 & 15 & -15 & -4 & 16 & -9 \\ 6 & -16 & 16 & -6 & -6 & 16 & -16 & 6 \\ 4 & -9 & 15 & -16 & 16 & -15 & 9 & -4 \end{bmatrix} \quad (3)$$

20

【0030】

式3と後述の式4および5との組合せは、2つの異なる変換が、トランスコーディングモジュール408のスケーリングコンポーネント内にどのように実装されるかを示している（図4）。一実装では、VC1-Tの精度は、16ビット精度であり、これは、MMX実装に非常に適している。その結果、コーデックの複雑さを大幅に低減できる。

【0031】

図6は、例示的な簡素化された閉ループDCT領域トランスコーダを示している。図6のアーキテクチャは、トランスコーディングモジュール408の例示的なアーキテクチャの一実装を表す（図4）。図6のアーキテクチャ600は、図5のアーキテクチャ500と比較して簡素化されているアーキテクチャである。図6を参照し、 C_8 を標準DCT変換行列、 B を逆量子化MPEG-2 DCTブロック、 b を B のDCTとすると、MPEG-2 IDCTは、

30

【0032】

【数4】

$$b = C'_8 B C_8$$

【0033】

のように計算される。そこで、

40

【0034】

【数5】

$\hat{B}$

【0035】

を b のVC1-Tとすると、

【0036】

【数6】

$\hat{B}$

【0037】

50

は

【0038】

【数7】

$$\hat{B} = T_8 b T_8' \circ N_{88}$$

【0039】

として計算され、ただし、 $\circ$ は、2つの行列の成分毎の乗算を表し、 N_{88} は、VC1-T変換の正規化行列であり、これは、

$$c_8 = \begin{bmatrix} 8/288 & 8/289 & 8/292 & 8/298 & 8/288 & 8/28 \\ 9 & 8/292 & 8/298 \end{bmatrix}$$

10

として、

【0040】

【数8】

$$N_{88} = c_8 \cdot c_8'$$

【0041】

のように計算される。

【0042】

【数9】

$\hat{B}$

【0043】

20

は、Bから直接計算され、その際に、式

【0044】

【数10】

$$\hat{B} = T_8 (C_8' B C_8) T_8' \circ N_{88} \quad (4)$$

【0045】

を使用する。

【0046】

【数11】

$$T_8 C_8'$$

30

【0047】

および

【0048】

【数12】

$$C_8 T_8'$$

【0049】

が対角行列に非常に近いことを検証するために、近似を適用する場合、式4は、行列Bの要素毎のスケーリングとなる。つまり、

【0050】

【数13】

40

$$S_{88} = \text{diag}(T_8 C_8') \cdot \text{diag}(C_8 T_8') \circ N_{88}$$

【0051】

として、

【0052】

【数14】

$$\hat{B} = B \circ S_{88} \quad (5)$$

【0053】

となる。

50

【0054】

式5は、WMVエンコーダのVC1-TとMPEG-2デコーダのIDCTをマージできることを示している。その結果、図5のアーキテクチャは、図6に示されているアーキテクチャにさらに簡素化されうる。詳細な比較を行うと、2つのDCT/I DCTモジュールは、2つのCV1-Tおよび逆VC1-Tモジュールで置き換えられることがわかる。一実装では、単純なスケーリングモジュールも加えられる。このアーキテクチャでは、2つのスイッチが、アクティビティマスクとともに埋め込まれる。これらの埋め込まれたコンポーネントは、後述のように、トランスコーダ408のトランスコーディングコーディングオペレーションの複雑さを動的に制御するために使用される（図4）。この時点で、これらのコンポーネントは、接続されている。WMV変換の16ビット算術オペレーション特性は、PCおよびDSPの並列処理に役立つ。したがって、計算の複雑さは、大幅に低減される。さらに、スケーリング行列 S_{88} のすべての成分は、互いに関して実質的に近接しているため、この計算、および一実装は、スカラー乗算で置き換えられる。

10

【0055】

図5および6は、フィードバックループが関与するそれぞれの例示的な閉ループトランスコーディングアーキテクチャを示している。この実装では、VC-1逆量子化、VC-1逆変換、残誤差累積、および累積誤差に対するMCを含む、フィードバックループは、VC-1再量子化プロセスによりもたらされる誤差を補償する。再量子化誤差は、図1に示されているような、ビットレート削減トランスコーダに対するドリフト誤差の主要原因である。図5および6のトランスコーディングアーキテクチャは、誤差補償があっても完全にドリフト誤差をなくすわけでないが、ドリフト誤差は非常に小さい。これは、動き補償フィルタリングの際のドリフト誤差の残りの原因が丸め誤差であるからである。残誤差補償のメリットの1つは、図5および6のアーキテクチャは、表2に関して後述のように、補償プロセスを動的にオンまたはオフにする機能を実現することである。図6のトランスコーディングアーキテクチャでは、実質的に最適な方法で、SDからSDへの、またはHDからHDへの変換など、MPEG-2からWMVへの純粋なビットレート削減トランスコーディングを実行する。

20

【0056】

より具体的には、従来のカスケードトランスコーダアーキテクチャ（例えば、図1および2のアーキテクチャ）は、複雑度柔軟性を欠いている。計算量削減に関して、そのような従来のアーキテクチャが達成できる大半のことは、MV再利用およびモードマッピングを介してのものである。その一方で、累積残誤差補償アーキテクチャ、例えば、図6のアーキテクチャ（および後述の図8と10のアーキテクチャ）は、複雑度に関するスケラビリティを組み込んでいる。表2は、図6のスイッチの例示的な意味を示している。

30

【0057】

【表2】

表2

トランスコーディング速度および品質の動的制御のための例示的なスイッチ

S_0	ブロックレベル	誤差累積スイッチ
S_1	ブロックレベル	誤差更新スイッチ
S_2	ブロックレベル	早期スキップブロック決定スイッチ

40

【0058】

図4のトランスコーディングモジュール408がドリフトのない簡素化を実装した後、アプリケーションは、複雑度とトランスコーディング速度を加速する品質との間のトレードオフの関係を動的に設定することができる。この実装では、品質は、速度を引き換えに、速度は、品質を引き換えに設定することができる。言い換えると、ある種のドリフト誤差は、他の簡素化されたトランスコーダにおいて許容されることがあるということである

50

。この戦略では、高速な方法で持ち込まれるドリフト誤差は、制限され、完全に制御可能である。この考慮事項に基づき、3つのスイッチ（ S_0 、 S_1 、および S_2 ）は、図6、8、および10のアーキテクチャで実現される。これらのスイッチは、残誤差補償ベースのアーキテクチャにのみ使用される。これらのスイッチは、いくつかの時間のかかるオペレーションを選択的にスキップし、複雑度を実質的に低減するが、誤差はごくわずかしかり込まない。様々なスイッチの意味は、表2にまとめられている。これらのスイッチに関連付けられている計算決定は、それぞれのスイッチに関して後述の基準により効率的に得られる。

【0059】

スイッチ S_0 は、いつブロックの再量子化誤差を累積して残誤差バッファに入れるべきかを制御する。標準再構成セレクトと比較して、スイッチ S_0 の役割は、高速ルックアップテーブルベースの再量子化プロセスを採用し、三重閾値アルゴリズムを介してより精密なドリフト制御メカニズムを実現することにより改善される。その結果、スイッチ S_0 に関して得られる観察結果はすべて、考慮される。例えば、一実装では、DCT領域エネルギー差をインジケータとして使用することができる。

【0060】

スイッチ S_1 は、最も時間のかかるモジュールの実行時期、累積残誤差のMCを制御する。一実装では、スイッチ S_1 はオンである。基準フレームに対し、2値アクティビティマスクが作成される。アクティビティマスクのそれぞれの要素は、

【0061】

【数15】

$$Activity(Block_i) = \begin{cases} 1, & Energy(block_i) > Th \\ 0, & Energy(block_i) \leq Th \end{cases}$$

【0062】

により決定されるように、 8×8 ブロックのアクティブ性に対応するが、ただし、 $Energy(block_i)$ は、累積残誤差バッファ内のブロックのエネルギーである。一実装では、 $Energy(block_i)$ は、計算された空間領域またはDCT領域である。 $Energy(block_i)$ は、絶対値の総和により近似することができる。MVが、低アクティビティの領域に属しているブロックを指している場合、その特定のブロックに対する累積残誤差のMCは、スキップされる。

【0063】

スイッチ S_2 は、早期検出を実行して、ブロック誤差を符号化すべきかどうかを決定する。これは、エンコーダが粗い量子化ステップサイズを適用する場合にアプリケーションをトランスレートするのに特に有用である。この実装では、入力信号（累積残誤差のMCとMPEG-2デコーダから再構成された残差の総和）が、閾値よりも弱い場合、誤差が符号化されないようにスイッチ S_2 はオフにされる。

【0064】

一実装では、スイッチ S_0 、 S_1 、および S_2 に対する閾値は、前の方の基準フレームが高い品質、遅い速度で処理されるように調節される。これは、スイッチの目的が、品質と速度との間のトレードオフの関係を高めることであり、また予測符号化の特性があるからである。

【0065】

[高品質プロファイルトランスコーダ]

ビットレート変更が大きな変更でないか、または入力ソース品質があまり高くない場合、図6のアーキテクチャでは、MPEG-2ビットストリームをWMVビットストリームに変換する際にビットレート削減を実質的に最適化する。他方、入力ソースが高品質であり、高品質出力が望ましく、さらにトランスコーディングの速度は、中程度の要求条件（例えば、リアルタイム）としてよい。MV微調整が行われる図3のカスケードピクセル領域トランスコーダ（CDPT）などの高品質プロファイルトランスコーダは、これらの基

10

20

30

40

50

準を満たす。このアーキテクチャを使用すると、最高の符号化効率が必ず達成されるように、WMVエンコーダのすべての拡張符号化機能をオンにすることができる。

【0066】

[分解能変更]

従来のメディアトランスコーディングシステムでは、一般に、空間分解能縮小機能でトランスコードする際に3つの誤差源がある。これらの誤差は、以下のとおりである。

- ・ 縮小：縮小されたビデオを得る際に生じる誤差。典型的には、画質と複雑度との間のトレードオフの関係を考慮するように縮小フィルタのオペレーションを設計する場合、特に空間領域において縮小する場合に、ハード配線を選択する。

- ・ 再量子化誤差：純粋なビットレート削減トランスコーディングプロセスの場合と同様に、これは、再量子化ステップサイズが粗い再量子化による誤差である。

- ・ MV誤差：MVが不正であると、動き補償予測は誤ることになる。その結果、再量子化誤差がどれだけ補償されようと、またビットレートがどの程度高かろうと、新しいMVおよびモードに基づいて動き補償を再計算しなければ完全な結果を得ることは困難である。WMVがBフレームに対するMVモードを1つかサポートしないので、これは、Bフレームをトランスコードする従来のシステムの問題である。これは、最適化を実行する必要がある場合に、符号化モードが変更される、例えば、4MVモードから1MVモードに変わるため、問題となる可能性もある。さらに、一般的に、クロミナンス成分に対する問題が存在するが、それは、典型的には単一のMVで補償されるからである。（これは、Pフレームに適用される場合には説明されている効率的なデジタルビデオトランスコーディングアーキテクチャの問題ではない。これに対する理由の1つは、WMVがPフレームに対する4MV符号化モードをサポートすることである。）

【0067】

トランスコーディングモジュール408（図4）のオペレーションでは、ここで説明されるように、最後の2つの誤差発生源を解消する。

【0068】

[再量子化誤差補償]

Dはダウンサンプリングフィルタリングを表すものとする。図3のアーキテクチャを参照すると、フレーム（i+1）に対するVC-1エンコーダへの入力は、

【0069】

【数16】

$$e_{i+1} = D(\hat{r}_{i+1}) + D(MC_{mp2}(\hat{b}_i, MV_{mp2})) - MC_{vc1}(\tilde{b}_i, mv_{vc1}) \quad (6)$$

【0070】

に従って導かれることがわかる。 $MC_{vc1} = MC_{mp2}$ 、 $mv_{mp2} = mv_{vc1} = MV_{mp2} / 2$ であると仮定する。近似は、

【0071】

【数17】

$$D(MC_{mp2}(\hat{b}_i, MV_{mp2})) = MC'_{mp2}(D(\hat{b}_i), D(MV_{mp2})) = MC'_{mp2}(\hat{b}_i, mv_{mp2}) \quad (7),$$

【0072】

である。式6は、

【0073】

【数18】

$$e_{i+1} = D(\hat{r}_{i+1}) + MC'_{mp2}(\hat{b}_i - \tilde{b}_i, mv_{mp2}) \quad (8)$$

【0074】

に簡素化される。

【0075】

式8の第1の項

10

20

30

40

50

【0076】

【数19】

$$D(\hat{f}_{i+1})$$

【0077】

は、復号化されたMPEG-2残留信号の縮小プロセスを指している。この第1の項は、空間領域ローパスフィルタリングおよびデシメーションを使用して決定することができる。しかし、DCT領域縮小機能を使用してこの項を求めることで、複雑度が低減され、PSNRおよび画質が向上する。DCT領域縮小結果は、係数 $(-1, 0, 9, 16, 9, 0, -1) / 32$ を使用する空間領域双一次フィルタリングまたは空間領域7タップフィルタリングを通じて得られる結果よりも実質的によい。この実装では、DCT領域縮小機能は、左上 4×4 低周波DCT係数のみを保持する。つまり、保持されているDCT係数上で標準 4×4 IDCTを適用すると、空間的2:1縮小画像（つまり、図4のトランスコードされたメディア414）が得られる。

10

【0078】

式8の第2の項

【0079】

【数20】

$$MC'_{mp2}(\hat{b}_i - \bar{b}_i, mv_{mp2})$$

20

【0080】

は、縮小分解能に対する再量子化誤差補償を意味する。この実装では、MPEG-2デコーダのMCおよびWMVエンコーダのMCは、低減された分解能で累積再量子化誤差に作用する単一MCプロセスにマージされる。

【0081】

図7は、4つの 4×4 DCTブロックのオペレーションを1つの 8×8 DCTブロックにマージする例示的なマージオペレーションを示している。実用上の問題が1つ残っている。DCT領域縮小では、4つの 8×8 DCT（元の分解能のMPEG-2マクロブロック（MB）内のブロック B_1 から B_4 ）は、低減された分解能で、そのままDCT領域において、新しいMBの 8×8 ブロックの4つの 4×4 サブブロックにマッピングされる（例えば、図7を参照）。WMVでは、PフレームおよびBフレームに対し、 4×4 変換タイプが許される。その結果、上述のスケーリングを除き、ほかにする必要はない。しかし、Iフレームについては、 8×8 変換タイプのみが許される。そのため、Iフレームを扱うときには、トランスコーディングモジュール408（図4）は、4つの 4×4 低周波DCTサブブロックを 8×8 DCTブロック1個

30

【0082】

【数21】

$$\hat{B}$$

【0083】

に変換する。一実装では、これは、4つの 4×4 DCTサブブロックをピクセル領域に逆変換し、次いで、新しい 8×8 VC1-Tを適用することにより行われる。一実装では、計算の複雑度を低減するために、これはDCT領域内で行われる。

40

【0084】

例えば、

【0085】

【数22】

$$\hat{B}_1, \hat{B}_2, \hat{B}_3$$

【0086】

および

50

【0087】

【数23】

 $\hat{B}_4$

【0088】

は、それぞれ B_1 、 B_2 、 B_3 、および B_4 の4つの 4×4 低周波サブブロックを表すものとし、 C_4 は、 4×4 標準IDCT変換行列であり、 T_8 は、整数WMV変換行列であり、さらに、 $T_8 [T_L, T_R]$ とし、 T_L および T_R は 8×4 行列である。このシナリオでは、

【0089】

【数24】

10

 $\hat{B}$

【0090】

は、

【0091】

【数25】

 $\hat{B}_1, \hat{B}_2, \hat{B}_3$

【0092】

および

【0093】

20

【数26】

 $\hat{B}_4$

【0094】

から、式

【0095】

【数27】

$$\hat{B} = (T_L C_4') \hat{B}_1 (T_L C_4')' + (T_L C_4') \hat{B}_2 (T_R C_4')' + (T_R C_4') \hat{B}_3 (T_L C_4')' + (T_R C_4') \hat{B}_4 (T_R C_4')'$$

【0096】

30

を使用して直接計算される。何らかの操作をした後、

【0097】

【数28】

 $\hat{B}$

【0098】

は、

【0099】

【数29】

$$\hat{B} = (X + Y)C' + (X - Y)D'$$

40

【0100】

のように効率的に計算されるが、ただし、

【0101】

【数 3 0】

$$C = (T_L C'_4 + T_R C'_4) / 2$$

$$D = (T_L C'_4 - T_R C'_4) / 2$$

$$X = C(\hat{B}_1 + \hat{B}_3) + D(\hat{B}_1 - \hat{B}_3)$$

$$Y = C(\hat{B}_2 + \hat{B}_4) + D(\hat{B}_2 - \hat{B}_4)$$

【0 1 0 2】

である。一実装では、上記の式のCおよびDは両方とも、事前に計算される。最終結果は、 N_{88} で正規化される。

10

【0 1 0 3】

図8は、簡素化されたDCT領域数値2：1分解能縮小トランスコーダに対する例示的なアーキテクチャ800を示している。一実装では、図4のトランスコーディングモジュール408は、例示的なアーキテクチャ800を実装するものである。このアーキテクチャのスイッチは、表2を参照しつつ上で説明されているように、図6のと同じ機能を有する。図8、および一実装を参照すると、第1の2つのモジュール（MPEG-2 VLDおよび逆量子化）は、図6に示されているのとは比べて簡素化されていることがわかる。これは、トランスコーディングモジュール408は、 8×8 ブロックから左上 4×4 部分のみを取り出す。

20

【0 1 0 4】

ドリフト誤差補償が低減された分解能で行われる従来の低ドリフトのトランスコーダと比べて、図6および8のトランスコーダは、混合ブロック処理モジュールを含まない。これは、WMVが、インター符号化されたマクロブロック内の 8×8 ブロックに対するイントラ符号化モードをサポートしているからである。言い換えると、元の分解能のイントラMBは、低減された分解能のインターMBのイントラ 8×8 ブロックにマッピングされるということである。したがって、MBモードマッピング規則は、

【0 1 0 5】

【数 3 1】

$$\text{mode_new} = \begin{cases} \text{INTRA} & \text{すべての mode_orig} = \text{INTRA} \text{ の場合} \\ \text{SKIP} & \text{すべての mode_orig} = \text{SKIP} \text{ の場合} \\ \text{INTER} & \text{その他の場合} \end{cases}$$

30

【0 1 0 6】

に示されているように、非常に単純な規則になる。既存の混合ブロック処理オペレーションは、典型的には、完全分解能の画像を再構成するために復号化ループを必要とする。したがって、混合ブロック処理を取り除くと、従来のシステムに比べて計算量を実質的に削減できる。

【0 1 0 7】

簡素化されたDCT領域2：1分解能縮小トランスコーディングアーキテクチャ800は、Pフレームについては実質的にドリフトを含まない。これは、4MV符号化モードの結果である。ドリフト誤差の唯一の原因は、縮小フィルタリングを使用するCPDTアーキテクチャと比べると、MVを $1/4$ 分解能から $1/2$ 分解能に丸めること（ $m_{v_{mp2}} = m_{v_{vc1}}$ を保証する）とMCおよび縮小の非可換的特性である。そのような残りの誤差は、ローパス縮小フィルタリングのせいで無視できるくらい小さい（例えば、DCT領域またはピクセル領域内で達成される）。

40

【0 1 0 8】

図9は、一実施形態による、2：1空間分解能縮小トランスコーディングオペレーションのインターレースメディアに対し4つの 4×4 DCTブロックのオペレーションを1つの 8×8 DCTブロックにマージする実施例を示す。2：1縮小は、元のフレームの分解

50

能を、水平方向と垂直方向の両方において2だけ変更する。一実装では、このインターレースプロセスは、図4のトランスコーディングモジュール408により実装される。より具体的には、インターレース符号化コンテンツでは、すべてのMB内の左上 8×4 サブブロックは、ショートカットMP EG-2デコーダにより再構成され、両方のフィールドは、垂直方向でローパスフィルタにより平滑化され、次いで1つのフィールドが、WMV符号化プロセスの前に削除される。

【0109】

[MV誤差補償]

WMVは、4MV符号化モードをサポートするが、典型的には、Pフレームを符号化することのみ対象とする。その結果、システム400（図4）は、入力MP EG-2ストリーム内にBフレームがないか、またはより低い時間的分解能に向けてトランスコーダ実行時にBフレームが破棄される場合に、図6のアーキテクチャを実行する。これに対する理由の1つは、WMVがBフレームについてMB毎にMVを1つしか許さないという点である。このようなシナリオでは、トランスコーディング408（図4）は、元の分解能でMBに関連付けられている4つのMVから新しい動きベクトルを合成する。前述のMV合成方法のそれぞれは、互換性を有する。一実装では、トランスコーディングモジュール408は、中央値フィルタリングを実行する。説明されているように、MVが不正であると、動き補償予測は誤ることになる。さらに悪いことに、再量子化誤差がどれだけ補償されようと、またビットレートがどの程度高かろうと、新しいMVに基づいて動き補償を再実行しない場合に完全な結果を得ることは困難である。そこで、このような動き誤差を補償できるアーキテクチャを実現する。

【0110】

再び、図3のアーキテクチャを参照すると、Bフレームであると仮定される、フレーム $(i+1)$ に対するVC-1への入力は、

【0111】

【数32】

$$e_{i+1} = D(\hat{r}_{i+1}) + D(MC_{mp2}(\hat{E}_i, MV_{mp2})) - MC_{vc1}(\tilde{b}_i, mv_{vc1}) \quad (9);$$

【0112】

のように導かれ、近似は

【0113】

【数33】

$$D(MC_{mp2}(\hat{E}_i, MV_{mp2})) = MC'_{mp2}(D(\hat{E}_i), D(MV_{mp2})) = MC'_{mp2}(\hat{b}_i, mv_{mp2}) \quad (10)$$

【0114】

である。

【0115】

式9は、

【0116】

【数34】

$$e_{i+1} = D(\hat{r}_{i+1}) + MC'_{mp2}(\hat{b}_i, mv_{mp2}) - MC'_{vc1}(\tilde{b}_i, mv_{vc1}) \quad (11)$$

【0117】

に簡素化される。式11に関して、

【0118】

10

20

30

40

【数 3 5】

$$\begin{aligned}
e_{i+1} &= D(\hat{f}_{i+1}) + MC'_{mp2}(\hat{b}_i, mv_{mp2}) - MC'_{vc1}(\tilde{b}_i, mv_{vc1}) \\
&= D(\hat{f}_{i+1}) + [MC'_{mp2}(\hat{b}_i, mv_{mp2}) - MC'_{vc1}(\hat{b}_i, mv_{vc1})] + MC'_{vc1}(\hat{b}_i, mv_{vc1}) - \\
&\quad MC'_{vc1}(\tilde{b}_i, mv_{vc1}) \\
&= D(\hat{f}_{i+1}) + [MC'_{mp2}(\hat{b}_i, mv_{mp2}) - MC'_{vc1}(\hat{b}_i, mv_{vc1})] + MC'_{vc1}(\hat{b}_i - \tilde{b}_i, mv_{vc1}) \quad (12)
\end{aligned}$$

【0 1 1 9】

が得られる。

【0 1 2 0】

10

式 1 2 の中の角かっこ内の 2 つの項は、矛盾した MV（つまり、 mv_{mp2} は、 mv_{vc1} と異なる）により引き起こされるか、または MPEG-2 と WMV との間の異なる MC フィルタリング法により引き起こされる動き誤差を補償する。この目的に対する対応するモジュールは、ハイライトされ、図 1 0 では、淡黄色ブロックにまとめられる。

【0 1 2 1】

図 1 0 は、一実施形態による、完全なドリフト補償が行われる例示的な簡素化された 2 : 1 縮小トランスコーダアーキテクチャを示している。一実装では、図 4 のトランスコーディングモジュール 4 0 8 は、図 1 0 の例示的なアーキテクチャを実行する。式 1 2 を参照する際に、

【0 1 2 2】

20

【数 3 6】

$$MC'_{mp2}(\hat{b}_i, mv_{mp2})$$

【0 1 2 3】

は、元のインター-MB に対応するすべての 8×8 ブロック、および $1/4$ ピクセル精度を有する $mv_{mp2} = MV_{mp2} / 2$ について実行されることに留意されたい。VC-1 エンコーダで使用される MV は、単一の MV : $mv_{vc1} = \text{median}(MV_{mp2}) / 2$ である。動き誤差補償モジュールに関して、 mv_{vc1} の精度は、 $1/4$ ピクセルレベルになりうることに留意されたい。式 1 2 の最後の項は、基準フレームの再量子化誤差を補償する。B フレームは、他のフレームの基準ではないため、誤差に強い。その結果、アプリケーション側で、高速化を達成するため誤差補償機能を安全にオフにすることができる。ここでもまた、このような近似は、B フレームのみを対象としたものである。動き誤差補償に対する MC は、再構成されたピクセルバッファに作用するが、再量子化誤差補償に対する MC は、累積残誤差バッファに作用することに留意されたい。

30

【0 1 2 4】

MC に関して、イントラ-インター、またはインター-イントラの変換を適用することができる。これは、MPEG-2 デコーダが B フレームおよび基準フレームを再構成したからである。この実装では、この変換は、図 1 0 の混合ブロック処理モジュール内で実行される。2 つのモード合成法が可能である。一実装では、基本モードが合成モードとして選択される。例えば、元の分解能の 4 つの MB のモードが、2 つの双方向予測モード、1 つの逆方向予測モードと 1 つの順方向予測モードである場合、双方向予測モードが、低減された分解能の MB に対するモードとして選択される。他の実装では、最大の誤差をもたらすモードが選択される。この実施例に関して、逆方向モードを使用すると、最大の誤差が生じる。このシナリオでは、誤差を補償できるように逆方向モードが選択される。結果から、後者の技術は、前者のモード選択技術に比べてわずかに優れた品質を示すことがわかる。

40

【0 1 2 5】

式 1 2 による例示的なアーキテクチャが、図 1 0 に示されている。表 3 に示されているように、特にこのアーキテクチャに対する 4 つのフレームレベルのスイッチがある。

【0 1 2 6】

50

【表 3】

表 3
例示的なフレームレベルのスイッチ

S_{IP}	フレームレベル	IおよびPフレームのみについてスイッチを閉じる。
S_P	フレームレベル	Pフレームのみについてスイッチを閉じる。
S_B	フレームレベル	Bフレームのみについてスイッチを閉じる ($= !S_{IP}$)。
$S_{IP/B}$	フレームレベル	Bフレームがある場合、IおよびPフレームのみについてスイッチを閉じる。

10

【0127】

4つのフレームレベルのスイッチにより、異なるフレームタイプに異なる符号化経路が確実なものとされる。特に、アーキテクチャは、Bフレーム (S_{IP}) に対する残誤差累積を実行せず、IおよびPフレーム (S_B) に対するMV誤差補償を実行せず、生成すべきBフレーム ($S_{IP/B}$) がない場合には基準フレームを再構成しない。対応する4つの元のMVが著しく矛盾する場合のみMV誤差が補償される必要があるため、フレームレベルスイッチ S_B はブロックレベルスイッチに変えることができることに留意されたい。

20

【0128】

より具体的には、スイッチ S_{IP} は、IフレームまたはPフレームのみについて閉じられ、スイッチ S_P は、Pフレームのみについて閉じられ、スイッチ S_B は、Bフレームについてのみ閉じられる。結果として得られるアーキテクチャは、図3の基準カスケードピクセル領域トランスコーダほど複雑ではない。これに対する理由の1つは、明示的なピクセル領域縮小プロセスが回避されることである。その代わりに、ピクセル領域縮小は、高いDCT係数を単純に破棄することによりDCT領域において暗黙のうちに実行される。このアーキテクチャでは、表2に関して上で説明されているように、様々なスイッチを使用することにより優れた複雑度スケーラビリティが実現される。

30

【0129】

超高速のトランスコーディング速度を必要とするアプリケーションでは、図10のアーキテクチャは、すべてのスイッチをオフにすることにより開ループ内に構成されうる。この開ループアーキテクチャは、さらに、MPEG-2の逆量子化プロセスとWMVの再量子化プロセスをマージすることにより最適化されうる。また、MPEG-2の逆ジグザグスキャンモジュール (VLD内) をWMVエンコーダ内のモジュールと組み合わせることもできる。

【0130】

[クロミナンス成分]

MPEG-2およびWMVのクロミナンス成分に関して、クロミナンス成分 (UV) のMVおよび符号化モードは、ルミナンス成分 (Y) から導かれる。低減された分解能のMBに対応する元の分解能の4つのMBすべてが、矛盾しない符号化モードを持つ場合 (つまり、すべてのインター符号化またはすべてのイントラ符号化)、問題はない。しかし、その場合でなければ、MPEG-2およびWMVの異なる派生規則により問題が生じる。MPEG-2では、MBがインターモードで符号化される場合に、UVブロックはインター符号化される。しかし、WMVでは、MBがインターモードで符号化され、イントラ符号化された 8×8 Yブロックが3個未満である場合にのみ、UVブロックはインター符号化される。この問題は、PフレームとBフレームの両方に存在する。図4のトランスコーディングモジュール408は、以下のようにこれらの問題を解消する。

40

- ・ インター-イントラ変換：インター符号化されたMBが3つのイントラ符号化され

50

た 8×8 Y ブロックを持つ場合（インター符号化された MB で 4 つすべての 8×8 Y ブロックをイントラ符号化することは不可能である）、UV ブロックはイントラ符号化される。この場合、元の分解能の 1 つの MB は、対応する UV ブロックとともにインター符号化される。これらの UV ブロックは、インターモードからイントラモードに変換される。人視覚系（HVS）は、クロミナンス信号に対する感度が弱いので、トランスコーディングモジュール 408 では、空間的隠し技術を使用して、 8×8 UV ブロックをインターモードからイントラモードに変換する。一実装では、DC 距離は、隠し方向を決定するためのインジケータとして使用される。隠しは、単純なコピーまたは他の補間法を介して行われる。

・ イントラ-インター変換：インター符号化された MB が 1 つまたは 2 つのイントラ符号化された 8×8 Y ブロックを有している場合、トランスコーディングモジュール 408 は、UV ブロックをインター符号化する。このシナリオでは、元の分解能の 4 つの対応する MB のうちに 1 つまたは 2 つのイントラ符号化された MB がある。これらの UV ブロックは、イントラモードからインターモードに変換される。この実装では、トランスコーディングモジュール 408 は、ゼロ設定法と呼ばれる時間的隠し技術を使用して、これらのブロックを処理し、これにより復号化ループを回避する。

【0131】

誤差隠しオペレーションを使用してクロミナンス成分のモード変換を処理すると、現在のフレームに持ち込まれた誤差は無視できるくらい小さく、したがって無視できるが、ただし、その後のフレームで色ずれを生じることがある。クロミナンス成分に対するドリフトは、典型的には、不正な動きにより生じる。これに対処し、品質を向上させるために、一実装では、トランスコーディングモジュール 408 は、クロミナンス成分に再構成ベースの補償を使用する（つまり、クロミナンス成分に対し淡黄色モジュールを常に適用する）。

【0132】

[レート制御]

図 11 は、デコーダに対する例示的な仮想バッファベリファイババッファ（VBV）を示す。図 11 の VBV モデルに基づくデコーダは、典型的には、既存の MPEG-2 ビットストリームを検証する。この実装では、ビデオレートが入力レートに比例して減少する場合、トランスコードされた WMV ビットストリームは、自動的に、VBV の要求条件を満たすことになる。この点で、本明細書の効率的なデジタルビデオトランスコーディングアーキテクチャでは、すべてのフレームに対し、符号化されたフレームサイズが入力フレームサイズに比例する。これらの新規性のあるアーキテクチャでは、目標フレームサイズと実際の結果として得られるフレームサイズとの累積された差を継続的に補償し、学習を介して、異なるビットレート範囲について直線量子化ステップ（QP）マッピング規則を構成する。

【0133】

高いビットレートでは、符号化ビット（B）と MPEG-2 TM-5 レート制御法でも使用される量子化ステップ（QP）との間に近似式がある。

【0134】

【数 37】

$$B = S \cdot \frac{X}{QP} \quad (13)$$

【0135】

ただし、S は、フレームの複雑度であり、X は、モデルパラメータである。フレームの複雑度は異なるコーデックでも同じままであると仮定すると、

【0136】

【数 3 8】

$$QP_{vcl} = \left(\frac{X_{vcl}}{X_{mp2}} \right) \cdot \left(\frac{B_{mp2}}{B_{vcl}} \right) \cdot QP_{mp2} = k \cdot \left(\frac{B_{mp2}}{B_{vcl}} \right) \cdot QP_{mp2}$$

【0 1 3 7】

となるが、ただし、 QP_{vcl} は、WMV再量子化で使用されるQP値であり、 QP_{mp2} は、MP EG-2量子化のQP値であり、 k は、目標ビットレートに関するモデルパラメータである。一実装では、線形モデル

$$QP_{vcl} / QP_{mp2} = k \cdot (B_{mp2} / B_{vcl}) + t \quad (14)$$

が使用される。低、中、および高ビットレートの場合のパラメータ k および t の値は、直線回帰法を使用して表4にまとめられている。 10

【0 1 3 8】

【表 4】

表 4

直線回帰法の例示的なパラメータ値

フレーム タイプ	Iフレーム		Pフレーム		Bフレーム	
	パラメータ		パラメータ		パラメータ	
	k	t	k	t	k	t
低 (<1 Mbps)	0.612861	-0.194954	0.016081	3.128561	0.076037	2.264825
中 (<3 Mbps)	0.314311	0.070494	0.041140	1.400647	0.207292	0.545977
高	0.682409	-0.248120	0.057869	1.115930	0.199024	0.441518

20

【0 1 3 9】

式14に基づく例示的な詳細レート制御アルゴリズムが、表5に示されており、表5に示されているアルゴリズム内の様々なシンボルの意味は、以下の表6において定義されている。 30

【0 1 4 0】

【表 5】

表 5
例示的なレート制御アルゴリズム

```

Initialize  $SumD = 0$ ;
While (MPEG-2 ストリームが終わりでない)
{
    ステップ 1: 1つのMPEG2フレームを復号化し、 $B_{mp2}$ および $QP_{mp2}$ を取得する。
    ステップ 2:  $B_{pred\_vcl} = B_{mp2} \cdot \frac{R_{vcl}}{R_{mp2}}$ 
                 $B_{vcl} = B_{pred\_vcl} + SumD$ 
                If ( $B_{vcl} < 0$ ) then  $B_{vcl} = 1$ ;
                 $QP_{vcl} = (k \cdot \frac{B_{mp2}}{B_{vcl}} + t) \cdot QP_{mp2}$ ;
                 $QP_{vcl}$ を丸めて[1, 31]にクリップする。
    ステップ 3:  $QP_{vcl}$ を使用してこのフレームをWMVフレームに符号化する。
    ステップ 4: 実際の符号化されたWMVフレームサイズ $B_{actual\_vcl}$ を取得する。
                Update  $SumD$ :  $SumD = SumD + B_{pred\_vcl} - B_{actual\_vcl}$ ;
}

```

10

20

【0 1 4 1】

【表 6】

表 6
表 5 のアルゴリズムで使用されるシンボルの定義

B_{mp2}	MPEG-2フレームサイズ
R_{mp2}	MPEG-2ストリームビットレート
R_{vcl}	目標WMVストリームビットレート
B_{pred_vcl}	ビットレートの比により予測されるWMVフレームサイズ
B_{vcl}	符号化する予想WMVフレームサイズ (新しいビットレート)
B_{actual_vcl}	実際の符号化されたWMVフレームサイズ
$SumD$	予測されるWMVフレームサイズと初めからの実際のWMVフレームサイズとの累積する差

30

【0 1 4 2】

[任意分解能変更]

40

例えば、レガシーSD受信機／プレーヤをサポートするため行われるHD分解能からSD分解能へのコンテンツの変換は有用である。HD形式の典型的な分解能は、 $1920 \times 1080i$ および $1280 \times 720p$ であるが、SDに対しては、 $720 \times 480i$ 、NTSCに対しては $720 \times 480p$ である。 $1920 \times 1080i$ から $720 \times 480i$ までの水平および垂直の縮小比は、それぞれ、 $8/3$ および $9/4$ である。アスペクト比を保つために、最終縮小比は、 $8/3$ となるように選択され、その結果得られる画像サイズは、 720×404 となる。同様に、 $1280 \times 720p$ から $720 \times 480p$ では、縮小比は、 $16/9$ となるように選択され、その結果得られる画像サイズは、 720×404 となる。デコーダ／プレーヤにより、 720×480 の完全画像となるように黒色バーが挿入される (ビットストリームにパディングされる代わりに)。

50

【0143】

デジタル信号処理理論によれば、縮小比 m/n に対する実質的に最適な縮小方法は、最初に、信号を n 倍でアップサンプリングし（つまり、元のサンプルとサンプルとの間に $n-1$ 個のゼロを挿入し）、ローパスフィルタ（例えば、多数のタップを有するサイン関数）を適用し、次いでその結果得られた信号を m 倍でデシメートすることである。このようなオペレーションを実行することで、縮小により入り込むスペクトルエイリアシングは、最大限抑制される。しかし、このプロセスは、非常に多量の計算を必要とし、入力信号が高品位であるためリアルタイムで実行することが困難である。この計算複雑度を低減するために、新規性のある２段階縮小戦略がとられる。

【0144】

10

図12は、一実施形態による、任意空間分解能縮小機能を持つトランスコーダを示している。一実装では、図4のトランスコーディングモジュール408は、図12のアーキテクチャを実行する。一実装では、任意縮小トランスコーダは、図12などの非統合トランスコーダである。他の実装では、図12に関して後述される、以下の任意縮小トランスコーディングオペレーションは、図5、6、8、および／または10に示されているような統合トランスコーダで実行される。

【0145】

図12を参照すると、システム1200は、任意縮小目標を得るために２段階縮小オペレーションを実行する。第１段階縮小の結果は、復号化ループ内に埋め込まれる。これにより、復号化オペレーションの複雑度が低減される。例えば、 $8/3$ の縮小比を得るために、縮小オペレーションが最初に実行され、 $2/1$ に縮小する。この第１段階縮小の結果は、復号化ループ内に入力され、そこで、第２段階縮小が、空間領域内で実行される。この実施例では、第２段階縮小オペレーションは、 $4/3$ の縮小を行い、 $8/3$ 縮小比を得る。他の実施例では、システム1200により、 $4/3$ 縮小を２回適用して（２段階で）、縮小比 $16/9$ が得られる。この２段階縮小方法では、すでに説明されているDCT領域縮小戦略を使用し、第１段階縮小結果を復号化ループ内に完全に埋め込む。分解能は、第１段階縮小後に著しく低減されるため、ピクセル領域上で最適な縮小方法を適用し続けることができる。

20

【0146】

図12を参照する際に、複数のMV

30

【0147】

【数39】

$\left(\left\lfloor \frac{m}{n} \right\rfloor \times \left\lfloor \frac{m}{n} \right\rfloor \text{ および } \left\lceil \frac{m}{n} \right\rceil \times \left\lceil \frac{m}{n} \right\rceil\right)$ の間

【0148】

が新しいMB（MVスケーリングおよびフィルタリングモジュール）に関連付けられていることに留意されたい。

【0149】

[例示的な手順]

図13は、一実施形態による、効率的なデジタルビデオトランスコーディングを行う手順1300を例示する。一実装では、図4のトランスコーディングモジュール408は、手順1300のオペレーションを実行する。図13を参照すると、ブロック1302において、この手順は符号化されたビットストリーム（例えば、図4の符号化されたメディア412）を受け取る。ブロック1304で、この手順は、第１のメディアデータ形式（例えば、MPEG-2、MPEG-4など）に関連する圧縮技術の第１のセットに従って符号化されたビットストリームを部分的に復号化する。この部分的復号化オペレーションで、中間データストリームを生成する。統合トランスコーダは、完全な復号化を実行しない。例えば、「概念的な」MPEG-2デコーダのMCが、WMVエンコーダのMCとマージされる場合、復号化オペレーションを、MPEG-2復号化を実行するものとして記述することは難しい。ブロック1306で、中間データストリームの縮小が望ましい場合、

40

50

この手順は、第1の縮小段階において符号化されたビットストリームに関連付けられているデータを縮小する。第1の縮小段階は、復号化ループのDCT領域内で実行される。ブロック1308で、2段階縮小が望ましい場合、この手順は、さらに、空間領域において、DCT領域内で縮小されたデータを縮小する（ブロック1306を参照）。

【0150】

ブロック1310で、圧縮技術の第1のセットに従って復号化されたデータは、圧縮技術の第2のセットにより符号化される。一実装では、手順1300は、図に示され、また図12および14に関して説明されているような非統合トランスコーディングアーキテクチャ内に実装される。この実装では、圧縮技術の第2のセットは、圧縮技術の第1のセットと同じである。他の実装では、手順1300は、図に示され、また図5～11および14に関して説明されているような統合トランスコーディングアーキテクチャ内に実装される。この実装では、圧縮技術の第2のセットは、圧縮技術の第1のセットと同じでない。例えば、一実装では、圧縮技術の第1のセットは、MPEG-2に関連付けられ、圧縮技術の第2のセットは、WMVに関連付けられる。

【0151】

[例示的な動作環境]

図14は、効率的なデジタルビデオトランスコーディングを完全にまたは部分的に実装できる好適なコンピューティング環境の一実施例を示している。例示的なコンピューティング環境1400は、図4の例示的なシステム400の好適なコンピューティング環境の一例にすぎず、本明細書で説明されているシステムおよび方法の使用または機能性の範囲に関する制限を示唆する意図はない。コンピューティング環境1400は、コンピューティング環境1400に示されている1つのコンポーネントまたはその組合せに関係する何らかの依存関係または要求条件がその環境にあるものと解釈すべきでない。

【0152】

本明細書で説明されている方法およびシステムは、他の数多くの汎用または専用コンピューティングシステム、環境、または構成で動作する。使用に適していると思われるよく知られているコンピューティングシステム、環境、および／または構成の実施例として、限定はしないが、パーソナルコンピュータ、サーバコンピュータ、マルチプロセッサシステム、マイクロプロセッサベースのシステム、ネットワークPC、ミニコンピュータ、メインフレームコンピュータ、上記システムまたはデバイスを含む分散コンピューティング環境などがある。フレームワークのコンパクト版またはサブセット版も、ハンドヘルドコンピュータまたは他のコンピューティングデバイスなどの資源の限られているクライアント内に実装することができる。本発明は、通信ネットワークを通じてリンクされているリモート処理デバイスによりタスクが実行されるネットワーク接続コンピューティング環境内で実施される。

【0153】

図14を参照すると、効率的なデジタルビデオトランスコーディングアーキテクチャを備える例示的なシステムは、例えば、図1のコンピューティングデバイス102に関連付けられているイニシエータオペレーションを実行するコンピュータ1410の形態の汎用コンピューティングデバイスを備える。コンピュータ1410が備えるコンポーネントとしては、限定はしないが、（複数の）演算処理装置1418、システムメモリ1430、およびシステムメモリを備える様々なシステムコンポーネントを演算処理装置1418に結合するシステムバス1421などがある。システムバス1421は、メモリバスまたはメモリコントローラ、周辺機器バス、および様々なバスアーキテクチャを使用するローカルバスを含む数種類のバス構造のうちのいずれでもよい。例えば、限定はしないが、このようなアーキテクチャとしては、Industry Standard Architecture (ISA) バス、Micro Channel Architecture (MCA) バス、Enhanced ISA (EISA) バス、Video Electronics Standards Association (VESA) ローカルバス、およびMezzanineバスとも呼ばれるPeripheral Component

Interconnect (PCI) バスがある。

【0154】

コンピュータ1410は、通常、様々なコンピュータ可読媒体を含む。コンピュータ可読媒体は、揮発性および不揮発性媒体、取り外し可能および取り外し不可能媒体を含む、コンピュータ1410によってアクセスされることができる媒体であればどのような媒体でも使用可能である。例えば、限定はしないが、コンピュータ可読媒体は、コンピュータ記憶媒体および通信媒体を含むことができる。コンピュータ記憶媒体は、コンピュータ可読命令、データ構造体、プログラムモジュール、またはその他のデータなどの情報を格納する方法または技術で実装される揮発性および不揮発性、取り外し可能および取り外し不可能媒体を含む。コンピュータ記憶媒体としては、限定はしないが、RAM、ROM、EEPROM、フラッシュメモリまたはその他のメモリ技術、CD-ROM、デジタル多目的ディスク(DVD)またはその他の光ディスク記憶装置、磁気カセット、磁気テープ、磁気ディスク記憶装置またはその他の磁気記憶デバイス、または所望の情報を格納するために使用することができ、しかもコンピュータ1410によりアクセスできるその他の媒体がある。

10

【0155】

通信媒体は、通常、コンピュータ可読命令、データ構造体、プログラムモジュール、または搬送波もしくはその他のトランスポートメカニズムなどの変調データ信号によるその他のデータを具現するものであり、任意の情報配信媒体を含む。「変調データ信号」という用語は、信号内に情報を符号化するような方法で特性のうちの1つまたは複数が設定または変更された信号を意味する。例えば、限定はしないが、通信媒体としては、有線ネットワークまたは直接配線接続などの有線媒体、および、音響、RF、赤外線、およびその他の無線媒体などの無線媒体がある。上記のいずれの組合せもコンピュータ可読媒体の範囲に収まらなければならない。

20

【0156】

システムメモリ1430は、読み取り専用メモリ(ROM)1431およびランダムアクセスメモリ(RAM)1432などの揮発性および／または不揮発性メモリの形態のコンピュータ記憶媒体を含む。起動時などにコンピュータ1410内の要素間の情報伝送を助ける基本ルーチンを含む基本入出力システム1433(BIOS)は、通常、ROM1431に格納される。通常、RAM1432は、演算処理装置1418に直接アクセス可能な、および／または演算処理装置1418によって現在操作されているデータおよび／またはプログラムモジュールを格納する。例えば、限定はしないが、図14は、オペレーティングシステム1434、アプリケーションプログラム1435、その他のプログラムモジュール1436、およびプログラムデータ1437を例示している。

30

【0157】

コンピュータ1410はさらに、その他の取り外し可能／取り外し不可能な揮発性／不揮発性コンピュータ記憶媒体を備えることもできる。例にすぎないが、図14は、取り外し不可能な不揮発性磁気媒体の読み書きを行うハードディスクドライブ1441、取り外し可能な不揮発性磁気ディスク1452の読み書きを行う磁気ディスクドライブ1451、およびCDROMまたはその他の光媒体などの取り外し可能な不揮発性光ディスク1456の読み書きを行う光ディスクドライブ1455を例示している。例示的な動作環境において使用できる他の取り外し可能／取り外し不可能な揮発性／不揮発性コンピュータ記憶媒体としては、限定はしないが、磁気テープカセット、フラッシュメモリカード、デジタル多目的ディスク、デジタルビデオテープ、ソリッドステートRAM、ソリッドステートROMなどがある。ハードディスクドライブ1441は、典型的には、インターフェース1440などの取り外し不可能メモリインターフェースを介してシステムバス1421に接続され、磁気ディスクドライブ1451および光ディスクドライブ1455は、典型的には、インターフェース1450などの取り外し可能メモリインターフェースによりシステムバス1421に接続される。

40

【0158】

50

図14に例示されている上記のドライブおよび関連コンピュータ記憶媒体は、コンピュータ1410用のコンピュータ可読命令、データ構造体、プログラムモジュール、およびその他のデータを格納する機能を備える。例えば、図14では、ハードディスクドライブ1441は、オペレーティングシステム1444、アプリケーションプログラム1445、その他のプログラムモジュール1446、およびプログラムデータ1447を格納するものとして例示されている。これらのコンポーネントは、オペレーティングシステム1434、アプリケーションプログラム1435、その他のプログラムモジュール1436、およびプログラムデータ1437と同じである場合もあれば異なる場合もあることに留意されたい。オペレーティングシステム1444、アプリケーションプログラム1445、その他のプログラムモジュール1446、およびプログラムデータ1447に対しては、ここで、異なる番号を割り当てて、それらが少なくとも異なるコピーであることを示している。

10

【0159】

ユーザは、キーボード1462、およびマウス、トラックボール、またはタッチパッドと一般に呼ばれるポインティングデバイス1461などの入力デバイスを介してコンピュータ1410にコマンドおよび情報を入力できる。他の入力デバイス（図に示されていない）としては、マイク、ジョイスティック、ペンタブレット、衛星放送受信アンテナ、スキャナなどがある。これらの入力デバイスおよびその他の入力デバイスは、システムバス1421に結合されているユーザ入力インターフェース1460を通じて演算処理装置1418に接続されることが多いが、パラレルポート、ゲームポート、またはユニバーサルシリアルバス（USB）などの他のインターフェースおよびバス構造により接続されることも可能である。この実装では、モニタ1491またはその他の種類のユーザインターフェースデバイスは、さらに、例えばビデオインターフェース1490などのインターフェースを介してシステムバス1421に接続される。

20

【0160】

コンピュータ1410は、リモートコンピュータ1480などの1つまたは複数のリモートコンピュータへの論理接続を使用してネットワーク接続環境で動作する。一実装では、リモートコンピュータ1480は、図1に示されているように、応答側のコンピューティングデバイス106を表している。リモートコンピュータ1480は、パーソナルコンピュータ、サーバ、ルータ、ネットワークPC、ピアデバイス、または他の共通ネットワークノードでもよく、特定の実装に応じて、コンピュータ1410に関係する上述の要素の多くまたはすべてを含むが、メモリ記憶デバイス1481だけが図14に例示されている。図14に示されている論理接続は、ローカルエリアネットワーク（LAN）1481およびワイドエリアネットワーク（WAN）1473を含むが、他のネットワークを含むこともできる。このようなネットワーキング環境は、オフィス、企業全体にわたるコンピュータネットワーク、イントラネット、およびインターネットでは一般的である。

30

【0161】

LANネットワーキング環境で使用される場合、コンピュータ1410は、ネットワークインターフェースまたはアダプタ1470を介してLAN 1471に接続される。WANネットワーキング環境で使用される場合、コンピュータ1410は、典型的には、インターネットなどのWAN 1473上で通信を確立するためモデム1472またはその他の手段を備える。モデム1472は、内蔵でも外付けでもよいが、ユーザ入力インターフェース1460またはその他の適切なメカニズムを介してシステムバス1421に接続されうる。ネットワーク接続環境では、コンピュータ1410またはその一部に関して示されているプログラムモジュールは、リモートメモリ記憶装置デバイスに格納されうる。例えば、限定はしないが、図14はリモートアプリケーションプログラム1485をメモリデバイス1481に置かれているものとして例示している。図に示されているネットワーク接続は例示的であり、コンピュータ間の通信リンクを確立するのに他の手段も使用可能である。

40

【0162】

50

[結び]

上の節では構造的特徴および／または方法論的なオペレーションまたはアクションに固有の言語で効率的なデジタルビデオトランスコーディングアーキテクチャについて説明しているが、付属の特許請求の範囲で定められている実装は、説明された特定の特徴またはアクションに必ずしも限られない。むしろ、説明されている効率的な統合デジタルビデオトランスコーディングアーキテクチャの特定の特徴およびオペレーションは、請求されている主題を実施するための複数の例示的な実装形態として開示されている。

【0163】

例えば、一実装では、説明されている高速および高品質トランスコーディングシステムおよび方法は、トランスコーディング、任意サイズ縮小、レート削減を含めて、MPEG-2からMPEG-4へのトランスコーディング、およびMPEG-4からWMVへのトランスコーディングに使用される。例えば、図6の簡素化された閉ループDCT領域トランスコーダは、MPEG-4をWMVにトランスコードするために使用することができる。MPEG-2 (IS-13818 Part. 2) との違いは、MPEG-2では、MCにおいて1/2ピクセル要素 (pel) MV精度および双一次補間のみを使用することであり、WMVには、そのような同じモード (1/2 pel 双一次) がある。しかし、MPEG-4では、1/2 pel と1/4 pel の両方のMV精度とともに、1/4 pel 位置 (WMVのとは異なる) に対する補間もサポートする。この違いに対処するために、1/2 pel MVがMPEG-4ビデオで使用される場合、トランスコーディングプロセスは、上述のようにMPEG-2からWMVトランスコーディングと同じである。さらに、1/4 pel MVがMPEG-4ビデオに含まれる場合、図6に関して上で説明されているようにMCにおける補間法が異なることで誤差が入り込む。さらに、図10に関して上で説明されている完全ドリフト補償がある簡素化された2:1縮小トランスコーダは、変更と無関係にMPEG-4からWMVへの2:1サイズ縮小トランスコーディングに適用可能である。さらに、図12の上で説明されているレート削減および任意縮小トランスコーディングオペレーションを含む、高品質トランスコーディングは、MPEG-4からWMVトランスコーディングに効果的である。

【図面の簡単な説明】

【0164】

【図1】入力ビットストリームを復号化するフロントエンドデコーダと、異なる符号化パラメータセットまたは新しい形式の新しいビットストリームを生成するエンコーダとをカスケード接続する、従来のカスケードピクセル領域トランスコーダ (CPDT) システムを示す図である。

【図2】図1のCPDTアーキテクチャを簡素化した、従来のカスケードDCT領域トランスコーダ (CDDT) アーキテクチャを示す図である。

【図3】一実施形態により、MPEG-2をWMVにトランスコードする例示的な非統合ピクセル領域トランスコーディング分割アーキテクチャを示す図である。より具体的には、この分割アーキテクチャは、効率的な統合デジタルビデオトランスコーディングの概念的基礎を形成するものである。

【図4】一実施形態による、効率的な統合デジタルビデオトランスコーディングを行う例示的なシステムを示す図である。

【図5】一実施形態による、例示的な簡素化された閉ループカスケードピクセル領域トランスコーダを示す図である。

【図6】一実施形態による、例示的な簡素化された閉ループDCT領域トランスコーダを示す図である。

【図7】一実施形態による、4つの4x4 DCTブロックの1つの8x8 DCTブロックへの例示的なマージオペレーションを示す図である。このマージオペレーションは、効率的ビデオコンテンツトランスコーディングの際に実行される。

【図8】一実施形態による、簡素化されたDCT領域数値2:1分解能縮小トランスコーダに対する例示的なアーキテクチャを示す図である。

【図 9】一実施形態による、2 : 1 空間分解能縮小トランスコーディングオペレーションのインターレースメディアに対する 4 つの 4×4 DCT ブロックのオペレーションを 1 つの 8×8 DCT ブロックにマージする実施例を示す図である。

【図 10】一実施形態による、ドリフト補償が十分な例示的な簡素化された 2 : 1 縮小トランスコーダアーキテクチャを示す図である。

【図 11】デコーダに対する例示的な標準仮想バッファベリファイヤバッファ (VBV) モデルを示す図である。

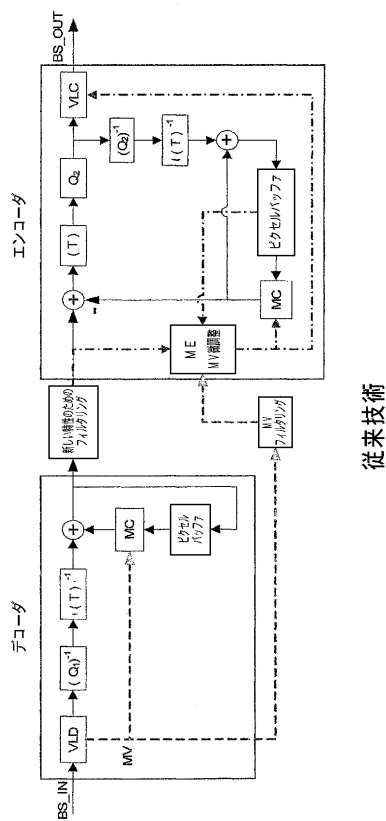
【図 12】一実施形態による、任意空間分解能縮小機能を持つトランスコーダを示す図である。

【図 13】一実施形態による、効率的な統合デジタルビデオトランスコーディングオペレーションの例示的な手順を示す図である。

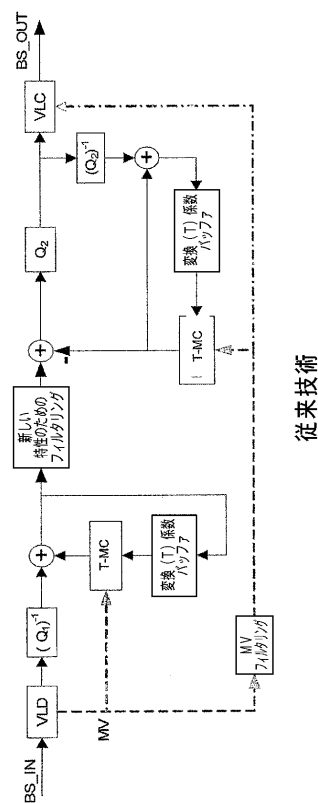
10

【図 14】一実施形態により、効率的な統合デジタルビデオトランスコーディングを部分的にまたは完全に実装できる例示的な環境を示す図である。

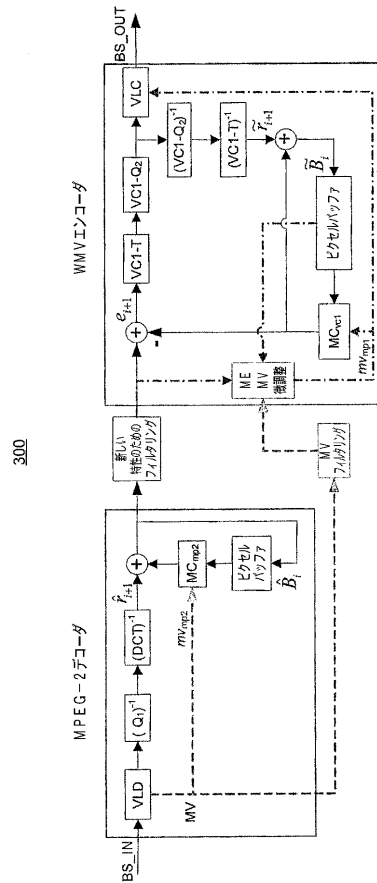
【図 1】



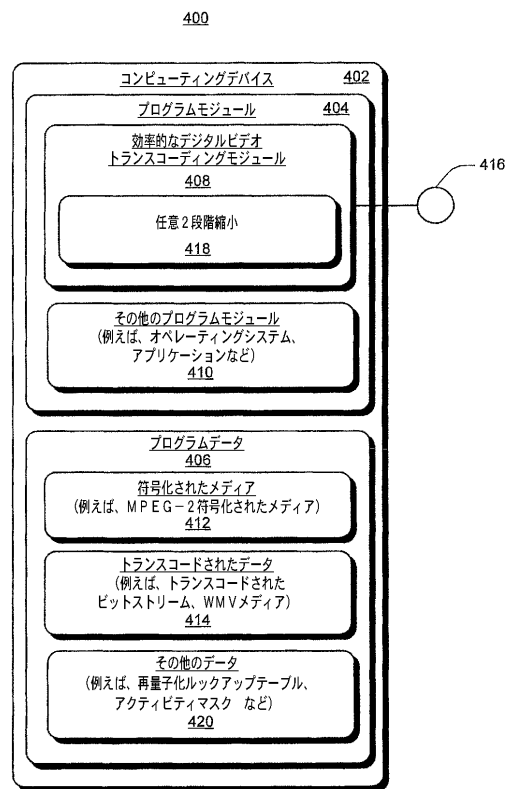
【図 2】



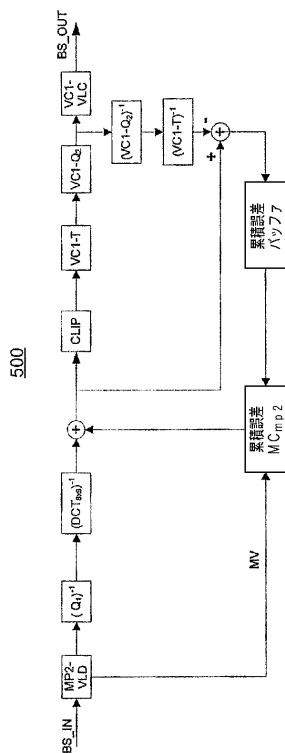
【図 3】



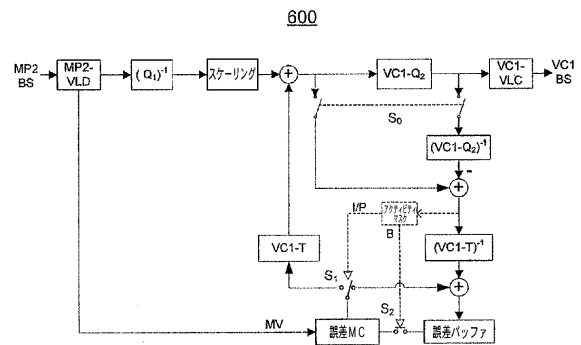
【図 4】



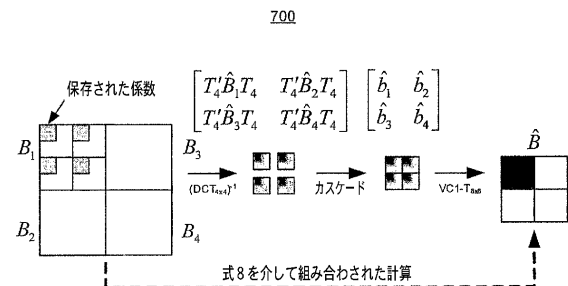
【図 5】



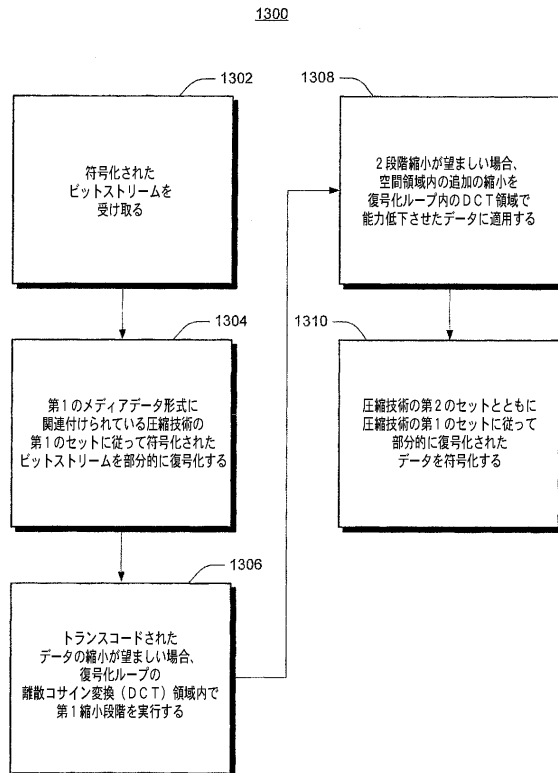
【図 6】



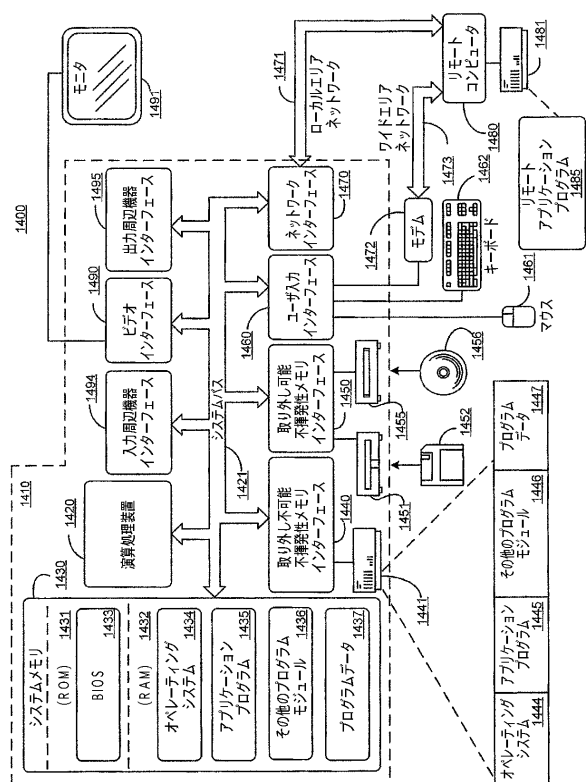
【図 7】



【図 1 3】



【図 14】



フロントページの続き

- (72)発明者 シーペン リー
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 パテント内
- (72)発明者 ワンヨン ツァオ
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 パテント内
- (72)発明者 ユーウェン ホー
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 パテント内

審査官 岩井 健二

- (56)参考文献 特開2004-015744 (JP, A)
特表2004-504739 (JP, A)
特開2003-189309 (JP, A)
特開2001-145113 (JP, A)
特開平11-073410 (JP, A)

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