

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5139303号  
(P5139303)

(45) 発行日 平成25年2月6日(2013.2.6)

(24) 登録日 平成24年11月22日(2012.11.22)

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

請求項の数 19 (全 72 頁)

(21) 出願番号	特願2008-533639 (P2008-533639)	(73) 特許権者	595020643
(86) (22) 出願日	平成18年9月27日 (2006. 9. 27)		クアルコム・インコーポレイテッド
(65) 公表番号	特表2009-510936 (P2009-510936A)		QUALCOMM INCORPORATED
(43) 公表日	平成21年3月12日 (2009. 3. 12)		アメリカ合衆国、カリフォルニア州 92
(86) 国際出願番号	PCT/US2006/037993		121-1714、サン・ディエゴ、モア
(87) 国際公開番号	W02007/038724		ハウス・ドライブ 5775
(87) 国際公開日	平成19年4月5日 (2007. 4. 5)	(74) 代理人	100108855
審査請求日	平成20年5月27日 (2008. 5. 27)		弁理士 蔵田 昌俊
(31) 優先権主張番号	60/721, 416	(74) 代理人	100091351
(32) 優先日	平成17年9月27日 (2005. 9. 27)		弁理士 河野 哲
(33) 優先権主張国	米国 (US)	(74) 代理人	100088683
(31) 優先権主張番号	60/727, 643		弁理士 中村 誠
(32) 優先日	平成17年10月17日 (2005.10.17)	(74) 代理人	100109830
(33) 優先権主張国	米国 (US)		弁理士 福原 淑弘

最終頁に続く

(54) 【発明の名称】冗長データ符号化方法及びデバイス

(57) 【特許請求の範囲】

【請求項 1】

マルチメディアデータを処理する方法であって、  
前記マルチメディアデータの複雑さに関する情報を入手することと、  
前記マルチメディアデータを符号化して符号化されたビットストリームを形成すること  
 であって、前記符号化することは、前記マルチメディアデータの複雑さに基づいて、イン  
 トラコーディングされたデータから成る少なくとも1つの冗長ブロックを前記符号化され  
 たビットストリーム内に挿入することであって、前記イントラコーディングされたデータ  
 は、少なくとも1つの冗長データフレームを具備し、前記少なくとも1つの冗長データフ  
 レームは、チャンネルスイッチフレームを具備する、方法。

10

【請求項 2】

前記複雑さは、空間的複雑さと、時間的複雑さと、を具備する請求項 1 に記載の方法。

【請求項 3】

挿入される冗長ブロックの量は、前記マルチメディアデータの前記複雑さに反比例する  
 請求項 1 に記載の方法。

【請求項 4】

前記少なくとも1つの冗長データフレームは、Iフレームを具備する請求項 1 に記載の  
 方法。

【請求項 5】

イントラコーディングされたデータからなる前記少なくとも1つの冗長ブロックは、実

20

質的にIフレームの全体を含む請求項1に記載の方法。

【請求項6】

前記符号化することは、前記マルチメディアデータを基本層及び拡張層へ符号化することを具備し、前記イントラコーディングされたデータは、前記基本層に関する量子化パラメータを用いて符号化される請求項1に記載の方法。

【請求項7】

前記符号化することは、前記マルチメディアデータを基本層及び拡張層へ符号化することを具備し、前記イントラコーディングされたデータは、前記拡張層に関する量子化パラメータを用いて符号化される請求項1に記載の方法。

【請求項8】

マルチメディアデータを処理するための装置であって、  
前記マルチメディアデータの複雑さに関する情報を入手するように構成されたコンテンツ分類器と、

前記マルチメディアデータを符号化して符号化されたビットストリームを形成するように構成された符号器であって、前記符号化することは、前記マルチメディアデータの複雑さに基づいて、イントラコーディングされたデータからなる少なくとも1つの冗長ブロックを前記符号化されたビットストリーム内に挿入することであって、前記イントラコーディングされたデータは、少なくとも1つの冗長データフレームを具備し、前記少なくとも1つの冗長データフレームは、チャンネルスイッチフレームを具備する符号器と、を具備する、装置。

【請求項9】

マルチメディアデータを処理するための装置であって、  
前記マルチメディアデータの複雑さに関する情報を入手するための手段と、  
前記マルチメディアデータを符号化して符号化されたビットストリームを形成するための手段であって、前記符号化することは、前記マルチメディアデータの複雑さに基づいて、イントラコーディングされたデータからなる少なくとも1つの冗長ブロックを前記符号化されたビットストリーム内に挿入することであって、前記イントラコーディングされたデータは、少なくとも1つの冗長データフレームを具備し、前記少なくとも1つの冗長データフレームは、チャンネルスイッチフレームを具備する手段と、を具備する、装置。

【請求項10】

前記複雑さは、空間的複雑さと、時間的複雑さと、を具備する請求項9に記載の装置。

【請求項11】

挿入される前記冗長ブロックの量は、前記マルチメディアデータの複雑さに反比例する請求項9に記載の装置。

【請求項12】

前記少なくとも1つの冗長データフレームは、Iフレームを具備する請求項8に記載の装置。

【請求項13】

イントラコーディングされたデータからなる前記少なくとも1つの冗長ブロックは、実質的にIフレームの全体を具備する請求項8に記載の装置。

【請求項14】

符号化するための前記手段は、前記マルチメディアデータを基本層及び拡張層へ符号化することを具備し、前記イントラコーディングされたデータは、前記基本層に関する量子化パラメータを用いて符号化される請求項9に記載の装置。

【請求項15】

符号化するための前記手段は、前記マルチメディアデータを基本層及び拡張層へ符号化することを具備し、前記イントラコーディングされたデータは、前記拡張層に関する量子化パラメータを用いて符号化される請求項9に記載の装置。

【請求項16】

マルチメディアデータを処理するためのプロセッサであって、

10

20

30

40

50

前記マルチメディアデータの複雑さに関する情報を入手し、  
前記マルチメディアデータを符号化して符号化されたビットストリームを形成するよう  
に構成され、

前記符号化することは、前記マルチメディアデータの複雑さに基づいて、イントラコー  
ディングされたデータからなる少なくとも1つの冗長ブロックを前記符号化されたビット  
ストリーム内に挿入することであって、前記イントラコーディングされたデータは、少な  
くとも1つの冗長データフレームを具備し、前記少なくとも1つの冗長データフレームは  
、チャンネルスイッチフレームを具備する、プロセッサ。

【請求項17】

前記複雑さは、空間的複雑さと、時間的複雑さと、を具備する請求項16に記載のプロ  
セッサ。

10

【請求項18】

マルチメディアデータを処理するための命令を具備する機械によって読み取り可能な媒  
体であって、前記命令は実行時に、機械に、

前記マルチメディアデータの複雑さに関する情報を機械に入手させることと、

前記マルチメディアデータを機械に符号化させて符号化されたビットストリームを形成  
することであって、前記符号化することは、前記マルチメディアデータの複雑さに基づい  
て、イントラコーディングされたデータからなる少なくとも1つの冗長ブロックを前記符  
号化されたビットストリーム内に挿入することを行わせ、

前記イントラコーディングされたデータは、少なくとも1つの冗長データフレームを具備  
し、前記少なくとも1つの冗長データフレームは、チャンネルスイッチフレームを具備する  
機械によって読み取り可能な媒体。

20

【請求項19】

前記複雑さは、空間的複雑さと、時間的複雑さと、を具備する請求項18に記載の機械  
によって読み取り可能な媒体。

【発明の詳細な説明】

【技術分野】

【0001】

35 U.S.C. § 119に基づく優先権の主張

本特許出願は、(a) “A VIDEO TRANSCODER FOR REAL-TIME STREAMING AND MOBILE BR  
OADCAS T APPLICATIONS” (リアルタイムストリーミング及びモバイル放送用途に関する映  
像トランスコーダ) という題名を有する仮特許出願番号60/721,416 (出願日:  
2005年9月27日)、(b) “A VIDEO TRANSCODER FOR REAL-TIME STREAMING AND M  
OBILE BROADCAST APPLICATIONS” (リアルタイムストリーミング及びモバイル放送用途に  
関する映像トランスコーダ) という題名を有する仮特許出願番号60/789,377 (出  
願日: 2006年4月4日)、(c) “METHOD AND APPARATUS FOR SPATIO-TEMPORAL D  
EINTERLACING AIDED BY MOTION COMPENSATION FOR FIELD-BASED VIDEO” (フィールドに  
基づく映像に関する動き補償によって補助された空間-時間的デインターレーシングに  
関する方法及び装置) という題名を有する仮特許出願番号60/727,643 (出願日:  
2005年10月17日)、(d) “METHOD AND APPARATUS FOR SHOT DETECTION IN VID  
EO STREAMING” (映像ストリーミングにおけるショット検出に関する方法及び装置) とい  
う題名を有する仮特許出願番号60/727,644 (出願日: 2005年10月17日)  
)、(e) “A METHOD AND APPARATUS FOR USING AN ADAPTIVE GOP STRUCTURE IN VIDEO  
STREAMING” (映像ストリーミングにおいて適応型GOPを用いるための方法及び装置)  
という題名を有する仮特許出願番号60/727,640 (出願日: 2005年10月17日)  
)、(f) “INVERSE TELECINE ALGORITHM BASED ON STATE MACHINE” (ステートマシ  
ンに基づく逆テレシネアルゴリズム) という題名を有する仮特許出願番号60/730,  
145 (2005年10月24日)、及び(g) “SPATIO-TEMPORAL DEINTERLACING AID  
ED BY MOTION COMPENSATION FOR FIELD-BASED MULTIMEDIA DATA” (フィールドに基づくマ  
ルチメディアデータに関する動き補償によって補助された空間-時間的デインターレーシ

30

40

50

ング)という題名を有する仮特許出願番号60/789,048(出願日:2006年4月3日)に対する優先権を主張するものである。これらの7つのすべての仮特許出願は、本特許出願の譲受人に対して譲渡されておりさらに本明細書において参照されることによって明示で本明細書に組み入れられている。

#### 【0002】

同時係属特許出願の参照

本特許出願は、本特許出願の譲受人に対して譲渡されておりさらに本明細書において参照されることによって明示で本明細書に組み入れられている、“CONTENT CLASSIFICATION FOR MULTIMEDIA PROCESSING”(マルチメディア処理に関するコンテンツ分類)という題名を有する米国特許出願番号11/373,577(出願日:2006年3月10日)に

10

関連するものである。本出願は、リアルタイムストリーミングを目的とする映像データの映像トランスコーディングに関する装置及び方法に関するものである。本発明は、特に、モバイル放送用途におけるリアルタイムストリーミングに関する映像データのトランスコーディングに関するものである。

#### 【背景技術】

#### 【0003】

多くのマルチメディア用途、例えば、無線映像ストリーミング及びビデオテレフォニー、においては、限られた帯域幅資源及び利用可能な帯域幅の可変性に起因して効率的な映像圧縮が有用である。一定の映像コーディング基準、例えば、MPEG-4(ISO/IEC)、H.264(ITU)、又は同様の映像コーディング基準、は、無線放送等の用途に非常に適した効率の高いコーディングを提供する。幾つかのマルチメディアデータ、例えば、デジタルテレビ放送、は、一般的には、その他の基準、例えば、MPEG-2、に従ってコーディングされる。従って、1つの基準(例えば、MPEG-2)に従ってコーディングされたマルチメディアデータを無線放送前に他の基準(例えば、H.264)にトランスコーディング又は変換するためにトランスコーダが用いられる。

20

レートが最適化されたコーデックの向上は、誤り耐性、誤り回復、及びスケーラビリティに関する利点をもたらすことが可能である。さらに、マルチメディアデータ自体から決定された情報の使用は、誤り耐性、誤り回復、及びスケーラビリティを含む、符号化に関するさらなる向上をもたらすことが可能である。従って、ストリーミングマルチメディア情報のモバイル放送を含む多くのマルチメディアデータ用途において使用するためにマルチメディアデータ自体から決定された情報を使用し、スケーラブルであり、誤り耐性を有するマルチメディアデータの非常に効率的な処理及び圧縮を提供するトランスコーダが必要である。

30

#### 【発明の開示】

#### 【発明の概要】

#### 【0004】

説明及び例示されている本発明のコンテンツに基づくトランスコーディング装置及び方法の各々は、幾つかの側面を有しており、いずれの単一の側面も、望ましい属性を確保する役割を単独で果たしているわけではない。以下では、この開示の適用範囲を限定することなしに、本発明のより顕著な特長が簡単に説明される。当業者は、この説明を検討後に、そして特に「発明を実施するための最良の形態」という題名の部分を読んだ後に、このコンテンツ駆動型トランスコーディングの特長がマルチメディアデータ処理装置及び方法に関する改良点をどのようにして提供するかを理解するであろう。

40

#### 【0005】

本明細書において説明される発明上の側面は、様々なマルチメディアデータ符号化方法に関して及び符号器、例えば、トランスコーダにおいて用いられる符号器、の様々なモジュール又は構成要素においてコンテンツ情報を用いることに関するものである。トランスコーダは、コンテンツ情報を用いてマルチメディアデータのトランスコーディングを調整することができる。前記コンテンツ情報は、他のソース、例えば、映像とともに受信され

50

るメタデータ、から受信することができる。前記トランスコーダは、様々な異なる処理動作を通じてコンテンツ情報を生成するように構成することができる。幾つかの側面においては、前記トランスコーダは、前記マルチメディアデータのコンテンツ分類を生成し、前記コンテンツ分類は、1つ以上の符号化プロセスにおいて用いられる。幾つかの側面においては、コンテンツによって駆動されるトランスコーダは、前記マルチメディアデータの空間的及び時間的コンテンツ情報を決定し、チャンネル間でのコンテンツを認識した均一な品質の符号化、及びコンテンツ分類に基づく圧縮/ビット割り当てに関して前記コンテンツ情報を使用することができる。

#### 【0006】

幾つかの側面においては、マルチメディアデータのコンテンツ情報（例えば、メタデータ、コンテンツメトリクス及び/又はコンテンツ分類）が入手又は計算され、前記マルチメディアデータを符号化のために処理する際に用いるために前記トランスコーダの構成要素に提供される。例えば、プリプロセッサは、シーンの変化の検出、逆テレシネ（“IVTC”）、デインターレーシング、動き補償及び雑音抑制（例えば、2Dウェーブレット変換）及び空間-時間的雑音の低減、例えば、アーティファクト除去、デリンギング、デブロッキング、及び/又は雑音除去、に関して一定のコンテンツ情報を用いることができる。幾つかの側面においては、プリプロセッサは、標準品位（SD）からクォーターピデオグラフィックスアレイ（QVGA）へのダウンサンプリング時に空間解像度ダウンサンプリング、例えば、適切な“安全”な及び“アクション処理”エリアの決定、に関して

10

20

#### 【0007】

幾つかの側面においては、符号器は、コンテンツ情報を計算するように構成されるコンテンツ分類モジュールを含む。前記符号器は、各MBに関する量子化パラメータ（QP）を決定する際のビットレート制御（例えば、ビット割り当て）に関して、動き推定、例えば、色動き推定（ME）を行う、動きベクトル（MV）予測を行うことに関して、基本層及び拡張層を提供する際のスケラビリティに関して、及び、例えば適応型イントラリフレッシュ、境界アライメントプロセス、及び拡張層において冗長Iフレームデータを提供することを含む予測階層及び誤り耐性方式に影響を与えるためにコンテンツ分類を用いることによる誤り耐性に関して、コンテンツ分類を用いることができる。幾つかの側面においては、前記トランスコーダは、チャンネル間において最適なマルチメディアデータ品質を維持するためにデータマルチプレクサと調整しながら前記コンテンツ分類を用いる。幾つかの側面においては、前記符号器は、高速チャンネル切り換えを可能にするためにIフレームが符号化されたデータ内において周期的に現れるように強制するために前記コンテンツ分類情報を用いることができる。該実装は、誤り耐性に関して前記符号化されたデータ内において要求されることがあるIブロックを利用することもでき、このため、予測階層を通じて（例えば、コンテンツ分類に基づいて）ランダムアクセス切り換えと誤り耐性を有効な形で結合させることによって誤りに対する頑健性（robustness）を向上させながらコーディング効率を向上させることができる。

30

#### 【0008】

一側面においては、マルチメディアデータを処理する方法は、マルチメディアデータのコンテンツ情報を入手することと、前記マルチメディアデータを符号化して符号化されたビットストリームを形成すること、とを含み、前記符号化することは、前記コンテンツ情報に基づいて、イントラコーディングされたデータから成る少なくとも1つの冗長ブロックを前記符号化されたビットストリーム内に挿入することを具備する。幾つかの側面においては、前記コンテンツ情報は、コンテンツ分類を具備する。前記コンテンツ分類は、前記マルチメディアデータの複雑さに基づくことができる。さらに、前記複雑さは、空間的複雑さ、時間的複雑さ、又は空間的複雑さと時間的複雑さの両方を具備することができる。幾つかの事例においては、挿入される冗長データ量は、前記マルチメディアデータの前記複雑さに反比例する。例示される側面は、前記イントラコーディングされたデータが少なくとも1つの冗長データフレームを具備し、前記少なくとも1つの冗長データフレーム

40

50

がIフレームを具備し、イントラコーディングされたデータから成る前記少なくとも1つの冗長フレームがチャンネルスイッチフレームを具備し、及び/又はイントラコーディングされたデータから成る前記少なくとも1つの冗長ブロックがIフレームの実質的な全体を具備する例を含む。前記符号化することは、基本層及び拡張層への前記マルチメディアデータを符号化することを含むことができ、前記イントラコーディングされたデータは、前記基本層に関する量子化パラメータ、又は前記拡張層に関する量子化パラメータを用いて符号化される。

【0009】

他の側面においては、マルチメディアデータを処理するための装置は、マルチメディアデータのコンテンツ情報を入手するための手段と、前記マルチメディアデータを符号して符号化されたビットストリームを形成するための手段と、を含み、前記符号化することは、前記コンテンツ情報に基づいて、イントラコーディングされたデータから成る少なくとも1つの冗長ブロックを前記符号化されたビットストリーム内に挿入することを具備する。

10

【0010】

さらに他の側面においては、マルチメディアデータを処理するための装置は、マルチメディアデータのコンテンツ情報を入手するように構成されたコンテンツ分類器と、前記マルチメディアデータを符号化して符号化されたビットストリームを形成するように構成された符号器と、を含み、前記符号化することは、前記コンテンツ情報に基づいて、イントラコーディングされたデータから成る少なくとも1つの冗長ブロックを前記符号化されたビットストリーム内に挿入することを具備する。

20

【0011】

他の側面においては、プロセッサは、マルチメディアデータのコンテンツ情報を入手し、前記マルチメディアデータを符号化して符号化されたビットストリームを形成するように構成され、前記符号化することは、前記コンテンツ情報に基づいて、イントラコーディングされたデータから成る少なくとも1つの冗長ブロックを前記符号化されたビットストリーム内に挿入することを具備する。

【0012】

他の側面においては、実行時にマルチメディアデータのコンテンツ情報を機械に入手させる命令と、前記マルチメディアデータを機械に符号化させて符号化されたビットストリームを形成する命令であって、前記符号化することは、前記コンテンツ情報に基づいて、イントラコーディングされたデータから成る少なくとも1つの冗長ブロックを前記符号化されたビットストリーム内に挿入することを具備する命令と、を具備する、機械によって読み取り可能な媒体。

30

【0013】

該当する場合において、同一のものについては図面全体に渡って同一の参照番号を付すことに注意すること。

【発明を実施するための最良の形態】

【0014】

以下の発明を実施するための最良の形態は、本開示において説明される一定の側面を対象にしたものである。しかしながら、本発明は、多数の異なる方法で具体化することが可能である。本明細書において「1つの側面」又は「側面」を参照しているときには、該側面と関連して説明される特定の特長、構造、又は特徴が少なくとも1つの側面の中に含まれていることを意味する。本明細書の様々な箇所における「一側面において」、「一側面により」、又は「幾つかの側面において」という表現の出現は、必ずしもすべてが同じ側面に言及しているわけではなく、さらに、その他の側面と別個であるか又は互いに排他的であるわけでもない。さらに、幾つかの側面によって示すことができその他の側面によって示すことができない様々な特長が説明される。同様に、幾つかの側面に関する要求であることができその他の側面に関する要求であることができない様々な要求が説明される。

40

【0015】

50

以下の説明は、例を徹底的に理解できるようにするために提供される詳細を含む。しかしながら、これらの例は、1つの例又は側面におけるプロセス又はデバイスに関するすべての詳細が本明細書において説明されていないか又は例示されていない場合でも実践可能であることが当業者によって理解される。例えば、電氣的構成要素は、これらの例を不必要に詳細にして不明確にしないようにするために該構成要素のすべての電氣的接続又はすべての電氣的要素を示していないブロック図において示すことができる。その他の場合においては、該構成要素、その他の構造及び技術は、これらの例をさらに説明するための詳細さで示すことができる。

【0016】

本開示は、符号化中のマルチメディアデータのコンテンツ情報を用いた符号化及びトランスコーディング装置及び方法を制御することに関するものである。(マルチメディアデータの)“コンテンツ情報”又は“コンテンツ”は、マルチメディアデータのコンテンツに関連する情報を意味する広義の言葉であり、例えば、メタデータと、マルチメディアデータから計算されたメトリックスと、1つ以上のメトリックスと関連づけられたコンテンツ関連情報、例えばコンテンツ分類と、を含むことができる。コンテンツ情報は、特定の用途に依存して、符号器に提供すること又は符号器によって決定することができる。コンテンツ情報は、シーン変化検出、時間的処理、空間-時間的雑音低減、ダウンサンプリング、量子化のためのビットレートの決定、スケラビリティ、誤り耐性、放送チャンネル間での最適なマルチメディア品質の維持、及び高速チャンネル切り換えを含むマルチメディアデータ符号化の数多くの側面に関して用いることができる。トランスコーダは、これらの側面のうちの1つ以上を用いて、マルチメディアデータの処理を調整し、コンテンツ関連の符号化されたマルチメディアデータを生成することができる。トランスコーディングの側面について説明する本明細書内の記述及び図は、符号化に関する側面及び復号に関する側面にも適用可能である。

【0017】

トランスコーダ装置及び方法は、1つの形式から他の形式へのトランスコーディングに関するものであり、本明細書においては、無線チャンネルを通じて幾つかの側面の例示であるモバイルデバイスに送信するためにMP EG-2映像を拡張されたスケラブルなH.264形式にトランスコーディングすることに関連して具体的に説明される。しかしながら、MP EG-2映像をH.264形式にトランスコーディングすることについての説明は、本発明の適用範囲を限定することは意図されておらず、本発明の幾つかの側面の単なる典型例であるにすぎない。開示される装置及び方法は、ランダムアクセス及び階層化能力を備えた誤り耐性の高い符号化をサポートする非常に効率的なアーキテクチャを提供し、MP EG-2及びH.264以外の映像形式のトランスコーディング及び/又は符号化に対しても適用可能である。

【0018】

本明細書において用いられる“マルチメディアデータ”又は単に“マルチメディア”は、(音声データを含むことができる)映像データ、音声データ、又は映像データと音声データの両方を含む広義の言葉である。本明細書において用いられる“映像データ”又は“映像”は、1つ以上の画像又は関連する一連の画像を含み、テキスト、画像情報及び/又は音声データを内蔵する、フレームに基づく又はフィールドに基づくデータを意味する広義の言葉であり、別記がないかぎりマルチメディアデータを指すために用いることができる(例えば、これらの言葉は、互換可能な形で用いることができる)。

【0019】

以下では、トランスコーダの様々な構成要素の例及びマルチメディアデータを符号化するためにコンテンツ情報を用いることができるプロセス例が説明される。

【0020】

図1Aは、マルチメディアデータ放送システム100の幾つかの側面のデータフローを示すブロック図である。システム100において、マルチメディアデータプロバイダ106は、符号化されたマルチメディアデータ104をトランスコーダ200に通信する。符

10

20

30

40

50

号化されたマルチメディアデータ104は、トランスコーダ200によって受信され、トランスコーダ200は、ブロック110においてマルチメディアデータ104を処理して生のマルチメディアデータにする。ブロック110における処理は、符号化されたマルチメディアデータ104を復号及び構文解析し、他の形式に符号化するための準備としてマルチメディアデータをさらに処理する。復号されたマルチメディアデータは、ブロック112に提供され、ブロック112において、マルチメディアデータは、予め決められたマルチメディア形式又は基準に符号化される。マルチメディアデータが符号化された時点で、ブロック114において、例えば無線放送システム（例えば、携帯電話放送ネットワーク、又は他の通信ネットワーク）を介して送信するための準備が行われる。幾つかの側面においては、受信されたマルチメディアデータ104は、MPEG-2基準に従って符号化されている。トランスコーディングされたマルチメディアデータ104が復号された後は、トランスコーダ200は、H.264基準に従ってマルチメディアデータを符号化する。

10

#### 【0021】

図1Bは、図1Aのブロック110及び112において処理を行うように構成することができるトランスコーダ130のブロック図である。トランスコーダ130は、マルチメディアデータを受信し、マルチメディアデータを復号及び構文解析してパケット化されたエレメンタリストリーム（例えば、サブタイトル、音声、メタデータ、“生”映像、CCデータ、及び表示時刻タイムスタンプ）にし、これらのデータを符号化して希望される形式にし、符号化されたデータをさらなる処理又は送信のために提供するように構成することができる。トランスコーダ130は、符号化されたデータを2つ以上のデータグループ、例えば、符号化された第1のデータグループ及び符号化された第2のデータグループで提供するように構成することができ、階層符号化と呼ばれる。幾つかの側面例においては、階層符号化方式における様々なデータグループ（又は層）は、第1のデータグループにおいて符号化されたデータの品質のほうが第2のデータグループにおいて符号化されたデータよりも低くなる（例えば、表示されたときにより低い視覚品質レベルを提供する）ように異なる品質レベルで符号化、及び形式化することができる。

20

#### 【0022】

図1Cは、マルチメディアデータをトランスコーディングするように構成することができる、及び図1Aのブロック110及び112において描かれている処理の一部又は全部を行うように構成することができるプロセッサ140のブロック図である。プロセッサ140は、本明細書において説明される、復号、構文解析、前処理、及び符号化を含むトランスコーディングプロセスのうちの一つ以上を行い、コンテンツ情報を処理に関して用いるモジュール124a...nを含むことができる。プロセッサ140は、内部メモリ122も含み、直接又は他のデバイスを通じて間接的に外部メモリ120と通信するように構成することができる。プロセッサ140は、プロセッサ140の外部の一つ以上のデバイスと通信するように構成された通信モジュール126も含み、マルチメディアデータを受信すること及び符号化されたデータ、例えば、第1のデータグループにおいて符号化されたデータ及び第2のデータグループにおいて符号化されたデータ、を提供することを含む。幾つかの側面例においては、階層符号化方式における様々なデータグループ（又は層）は、第1のデータグループにおいて符号化されたデータの品質のほうが第2のデータグループにおいて符号化されたデータよりも低くなる（表示されたときにより低い視覚品質レベルを提供する）ように、異なる品質レベルで符号化し、形式化することができる。

30

40

#### 【0023】

トランスコーダ130又は（トランスコーディング用に構成された）プリプロセッサ140、その構成要素、及び内蔵されるプロセスは、ハードウェア、ソフトウェア、ファームウェア、ミドルウェア、マイクロコード、又はその組合せによって実装することができる。例えば、構文解析器、復号器、プリプロセッサ、又は符号器は、独立型構成要素であること、ハードウェア、ファームウェア、ミドルウェアとして他のデバイスの構成要素内に組み込むこと、又はプロセッサにおいて実行されるマイクロコード又はソフトウェア、

50

又はその組合せにおいて実装することができる。ソフトウェア、ファームウェア、ミドルウェア又はマイクロコード内に実装されるときは、動き補償、ショット分類及び符号化プロセスを行うプログラムコード又はコードセグメントは、機械によって読み取り可能な媒体、例えば記憶媒体、に格納することができる。コードセグメントは、手順、関数、サブプログラム、プログラム、ルーチン、サブルーチン、モジュール、ソフトウェアパッケージ、クラス、命令の組合せ、データ構造、又はプログラム文を表すことができる。コードセグメントは、情報、データ、引数、パラメータ、メモリコンテンツを渡す及び/又は受け取ることによって他のコードセグメント又はハードウェア回路に結合させることができる。

#### 【 0 0 2 4 】

##### トランスコーダアーキテクチャ例

図 2 は、図 1 のマルチメディア放送システム 1 0 0 において示されるトランスコーダ 2 0 0 に関して用いることができるトランスコーダ例のブロック図である。トランスコーダ 2 0 0 は、以下においてさらに説明される構文解析器 / 復号器 2 0 2 と、プリプロセッサ 2 2 6 と、符号器 2 2 8 と、同期化層 2 4 0 と、を具備する。トランスコーダ 2 0 0 は、本明細書において説明されるように、トランスコーディングプロセスの 1 つ以上の側面に関するマルチメディアデータ 1 0 4 のコンテンツ情報を用いるように構成される。コンテンツ情報は、マルチメディアメタデータを通じてトランスコーダ 2 0 0 の外部のソースから入手すること、又はトランスコーダによって、例えば、プリプロセッサ 2 2 6 又は符号器 2 2 8 によって計算することができる。図 2 に示される構成要素は、1 つ以上のトランスコーディングプロセスに関してコンテンツ情報を用いるトランスコーダ内に含めることができる構成要素を例示したものである。特定の実装においては、トランスコーダ 2 0 0 の構成要素のうちの 1 つ以上は、除外することができ又は追加の構成要素を含めることができる。さらに、本明細書においてはプロセス又はデバイスのすべての詳細を説明することができないにもかかわらず当業者が本発明を実践するのを可能にするためにトランスコーダ及びトランスコーディングプロセスの一部が説明される。

#### 【 0 0 2 5 】

図 5 は、トランスコーダ 2 0 0 の様々な構成要素の動作及び/又はプロセスの時間的關係を図示したタイミング図である。図 5 に示されるように、最初に、符号化されたストリーミング映像 1 0 4 ( 符号化されたマルチメディアデータ )、例えば、M P E G - 2 映像、が、構文解析器 2 0 5 ( 図 2 ) によって任意の時間ゼロ ( 0 ) において受信される。次に、映像ストリームは、例えば復号器 2 1 4 と組み合わせた構文解析器 2 0 5 によって構文解析され 5 0 1、多重分離され 5 0 2、復号される 5 0 3。示されるように、これらのプロセスは、データ処理のストリーム出力をプリプロセッサ 2 2 6 に提供するために、わずかなタイミングのオフセットを伴う形で並行して起きることができる ( 図 2 )。時間  $T_1$  5 0 4 において、プリプロセッサ 2 2 6 が処理結果の出力を開始する上で十分なデータを復号器 2 1 4 から受信した時点で、残りの処理ステップは、性質上順次的になり、前処理後は、時間  $T_f$  5 0 8 における再符号化が完了するまで第 1 パス符号化 5 0 5、第 2 パス符号化 5 0 6、及び再符号化 5 0 7 が順に生じる。

#### 【 0 0 2 6 】

本明細書において説明されるトランスコーダ 2 0 0 は、様々なマルチメディアデータをトランスコーディングするように構成することができ、多くのプロセスは、トランスコーディングされるあらゆる型のマルチメディアデータに適用される。本明細書において提供される例の一部は、M P E G - 2 データを H . 2 6 4 データにトランスコーディングすることに特に関するものであるが、これらの例は、本開示を該データに限定することは意味しない。以下において説明される符号化側面は、あらゆる適切なマルチメディアデータ基準を他の適切なマルチメディアデータ基準にトランスコーディングすることに適用することができる。

#### 【 0 0 2 7 】

##### 構文解析器 / 復号器

再度図2に関して、構文解析器/復号器202は、マルチメディアデータ104を受信する。構文解析器/復号器202は、マルチメディアデータ104を受信し、該データを構文解析して映像エレメンタリストリーム(ES)206、音声ES208、表示時刻タイムスタンプ(PTS)210及びその他のデータ、例えばサブタイトル212、にするトランスポートストリーム構文解析器(“構文解析器”)205を含む。ESは、1つの型のデータ(映像又は音声)を単一の映像又は音声符号器から搬送する。例えば、映像ESは、シーケンスヘッダー及びすべての小部分を含むデータシーケンスに関する映像データを具備する。パケット化されたエレメンタリストリーム、すなわちPES、は、パケット化されている単一のESから成り、典型的には各々が追加されたパケットヘッダーで始まる、PESストリームは、1つのソースからの、例えば1つの映像又は音声符号器からの、1つの型のデータのみを含む。PESパケットは、トランスポートパケットの固定パケット長に対応しない可変長を有し、トランスポートパケットよりもはるかに長くすることができる。トランスポートパケットがPESストリームから形成されるときには、PESヘッダーは、トランスポートパケットペイロードの始めで、トランスポートパケットヘッダーの直後に配置することができる。残りのPESパケットコンテンツは、PESパケットがすべて用いられるまで後続するトランスポートパケットのペイロードを満たす。最後のトランスポートパケットは、例えばバイト、例えば、バイト=0xFF(すべて1)を結めることによって、固定長まで満たすことができる。

10

#### 【0028】

構文解析器205は、ここに示される構文解析器/復号器202の一部である復号器214に映像ES206を通信する。その他の構成においては、構文解析器205及び復号器214は、別個の構成要素である。PTS210は、トランスコーダPTS生成器215に送信され、トランスコーダPTS生成器215は、トランスコーダ200から放送システムに送信されるデータを手配する際に用いるためにトランスコーダ200専用の別個の表示時刻タイムスタンプを生成することができる。トランスコーダPTS生成器215は、データ放送の同期化を調整するためにトランスコーダ200の同期化層240にデータを提供するように構成することができる。

20

#### 【0029】

図3は、構文解析器205が上述される様々なパケット化されたエレメンタリストリームを構文解析するときに従うことができるプロセス300の一例の流れ図である。プロセス300は、ブロック302において開始し、マルチメディアデータ104がコンテンツプロバイダ106から受信される(図1)。プロセス300は、ブロック304に進み、構文解析器205の初期設定が行われる。初期設定は、独立して生成された取得コマンド306によってトリガーすることができる。例えば、構文解析器205から独立して、外部で受信されたTVスケジュール及びチャンネルラインアップ情報に基づくプロセスが、取得コマンド306を生成することができる。さらに、初期設定及び主処理の両方を援助するためにリアルタイムトランスポートストリーム(TS)バッファ記述子308を入力することができる。

30

#### 【0030】

ブロック304において例示されるように、初期設定は、コマンド構文検証を取得することと、第1パスPSI/PSIP/SI(プログラム専用情報/プログラム及びシステム情報プロトコル/システム情報)の処理を行うことと、取得コマンド又はPSI/PSIP/SI一貫性検証のいずれかに特に関連する処理を行うことと、各PESに関するPESバッファを割り当てることと、(例えば、希望される取得開始インスタントとのアライメントに関する)タイミングを設定すること、とを含むことができる。PESバッファは、構文解析されたESデータを保持し、各構文解析されたESデータを対応する音声復号器216、試験符号器220、復号器214、又はトランスコーダPTS生成器215に通信する。

40

#### 【0031】

初期設定後、プロセス300は、受信されたマルチメディアデータ104の主処理のた

50

めにブロック310に進む。ブロック310における処理は、目標パケット識別子(PID)フィルタリングと、連続的PSI/PSIP/SIモニタリング及び処理と、着信したマルチメディアデータが適切なPESバッファ内に渡されるようにするための(例えば、希望される取得継続時間を達成させるための)タイミングプロセスと、を含むことができる。ブロック310におけるマルチメディアデータの処理の結果、PESバッファ‘読み取り’のプログラム記述子及び表示が生成され、後述されるように復号器214とインタフェースする(図2)。

#### 【0032】

ブロック310後、プロセス300は、ブロック314に進み、タイマー割り込みの生成及びPESバッファの消費に後続するPESバッファの解放を含む構文解析動作の終了が起きる。PESバッファは、記述子内において引用されたプログラムのすべての関連するエレメンタリストリーム、例えば、音声、映像、及びサブタイトルストリーム、に関して存在することが注記される。

10

#### 【0033】

再度図2に関して、構文解析器205は、トランスコーダ実装に対応するために音声ES208を音声復号器216に送信し、符号化されたテキスト216を同期化層240及び音声情報の復号に提供する。サブタイトル情報212は、テキスト符号器220に引き渡される。復号器214からのクロズドキャプション(CC)データ218もテキスト符号器220に提供され、テキスト符号器220は、サブタイトル情報212及びCCデータ218をトランスコーダ220によって生成された形式で符号化する。

20

#### 【0034】

構文解析器/復号器202は、映像ES206を受信する復号器214も含む。復号器214は、映像データと関連づけられたメタデータを生成することができ、符号化された映像のパケット化されたエレメンタリストリームを復号して(例えば、標準定義形式の)生映像224にし、映像ESストリーム内の映像クロズドキャプションされたデータを処理する。

#### 【0035】

図4は、復号器214によって行うことができる復号プロセス400の一例を示す流れ図である。プロセス400は、ブロック402における映像エレメンタリストリームデータ206の入力から開始する。プロセス400は、ブロック404に進み、復号器が初期設定される。初期設定は、映像シーケンスヘッダー(VSH)の検出と、第1パスVHS、映像シーケンス(VS)、及び(映像形式、三原色、及びマトリクス係数を含む)VS表示拡張処理を行うことと、復号されたピクチャ、関連づけられたメタデータ及びクロズドキャプション(CC)データを各々バッファリングするためのデータバッファを割り当てること、とを含む幾つかのタスクを含むことができる。さらに、構文解析器205によって提供された映像PESバッファ‘読み取り’情報406が入力される(例えば、図3のブロック310においてプロセス300によって生成することができる)。

30

#### 【0036】

ブロック404において初期設定後は、プロセス400は、ブロック408に進み、映像ESの主処理が復号器214によって行われる。主処理は、新しいデータの利用可能性に関して映像PESバッファ‘読み取り’情報又は“インタフェース”をポーリングすることと、映像ESを復号することと、ピクチャ境界における画素データを再構築及び格納することと、映像及びa/vを同期化することと、ピクチャ境界におけるメタデータを生成及び保存することと、ピクチャ境界におけるCCデータを保存すること、とを含む。主処理408の結果ブロック410は、シーケンス記述子、復号されたピクチャバッファ記述子、メタデータバッファ記述子、及びCCデータバッファ記述子の生成を含む。

40

#### 【0037】

主処理408後、プロセス400は、ブロック412に進み、終了プロセスを行う。終了プロセスは、予め決められたしきい値を超える特定の継続時間中に新しいデータが発生しないこと、シーケンス終了符号の検出、及び/又は明示の終了信号の検出を含む終了条

50

件を決定することを含むことができる。終了プロセスは、復号されたピクチャ、関連づけられたメタデータ、及びCCデータバッファを後述されるプリプロセッサによる消費後に解放することをさらに含むことができる。プロセス400は、ブロック414において終了し、入力として受信される映像ESを待つ状態に入ることができる。

#### 【0038】

##### プリプロセッサ

図2、及びより詳細な図6は、1つ以上の前処理動作に関してコンテンツ情報を用いることができるプリプロセッサ226のサンプル側面を示す。プリプロセッサ226は、メタデータ222及び復号された“生”映像データ224を構文解析器/復号器202から受信する。プリプロセッサ226は、映像データ224及びメタデータ222に関して一定の型の処理を行い、処理されたマルチメディア（例えば、基本層基準フレーム、拡張層基準フレーム、帯域幅情報、コンテンツ情報）及び映像を符号器228に提供するように構成することができる。マルチメディアデータの該前処理は、視覚上の鮮明度、エリアシング対策、及びデータの圧縮効率を向上させることができる。一般的には、プリプロセッサ226は、構文解析器復号器202内の復号器214によって提供される映像シーケンスを受信し、該映像シーケンスを符号器228によるさらなる処理（例えば、符号化）のためにプログレッシブ映像シーケンスに変換する。幾つかの側面においては、プリプロセッサ226は、逆テレシネと、デインターレーシングと、フィルタリング（例えば、アーティファクト除去、デリンギング、デブロッキング、及び雑音除去）と、リサイジング（例えば、標準品位からクォータービデオグラフィックスアレイ（QVGA）への空間解像度ダウンサンプリング）と、GOP構造生成（例えば、複雑さマップ生成の計算、シーン変化検出、及びフェード/フラッシュ検出）と、を含む数多くの動作に関して構成することができる。

#### 【0039】

プリプロセッサ226は、前処理動作のうちの一つ以上に影響を与えるために復号器からのメタデータを用いることができる。メタデータは、マルチメディアデータのコンテンツ（“コンテンツ情報”）に関連する、説明する、又は分類する情報を含むことができ、特に、メタデータは、コンテンツ分類を含むことができる。幾つかの側面においては、メタデータは、符号化動作に関して希望されるコンテンツ情報を含まない。このような場合には、プリプロセッサ226は、コンテンツ情報を決定して該コンテンツ情報を前処理動作に関して用いるように構成することができ及び/又はコンテンツ情報をトランスコーダ200のその他の構成要素、例えば復号器228、に提供する。幾つかの側面においては、プリプロセッサ226は、GOPパーティショニングに影響を与え、適切な型のフィルタリングを決定し、及び/又は符号器に通信される符号化パラメータを決定するために該コンテンツ情報を用いることができる。

#### 【0040】

図6は、プリプロセッサ226に含めることができる様々なプロセスブロックの説明例を示し、プリプロセッサ226によって行うことができる処理を例示する。この例においては、プリプロセッサ226は、メタデータ及び映像222、224を受信し、（処理された）メタデータ及び映像を具備する出力データ614を符号器228に提供する。典型的には、受信することができる3つの型の映像が存在する。第1に、受信された映像は、プログレッシブ映像であることができ、この場合は、デインターレーシングは要求されない。第2に、映像データは、テレシネされた映像、24fpsムービーシーケンス、この場合は映像、から変換された、インターレーシングされた映像であることができる。第3に、映像は、テレシネされない、インターレーシングされた映像であることができる。プリプロセッサ226は、後述されるようにこれらの型の映像を処理することができる。

#### 【0041】

ブロック601において、プリプロセッサ226は、受信された映像データ222、224がプログレッシブ映像であるかどうかを決定する。幾つかの場合においては、このことは、メタデータが該情報を含む場合はメタデータから決定すること、又は映像データ自

10

20

30

40

50

体の処理によって決定することができる。例えば、後述される逆テレシネプロセスは、受信された映像 2 2 2 がプログレッシブ映像であるかどうかを決定することができる。受信された映像 2 2 2 がプログレッシブ映像である場合は、プロセスはブロック 6 0 7 に進み、白色ガウス雑音等の雑音を低減させるためのフィルタリング（又は雑音除去器）動作が映像に関して行われる。ブロック 6 0 1 において、映像データ 2 2 2、2 2 4 がプログレッシブ映像でない場合は、プロセスは、位相検出器 6 0 4 へのブロック 6 0 4 に進む。

【 0 0 4 2 】

位相検出器 6 0 4 は、テレシネにおいて発生した映像と標準放送形式で始まった映像を区別する。映像がテレシネされたとする決定が行われた場合は（位相検出器 6 0 6 から出る「はい」決定経路）、テレシネされた映像は、逆テレシネ 6 0 6 において原形式に戻される。冗長フレームが識別されて除去され、同じ映像フレームから導き出されたフィールドが完全な画像内に再度織り込まれる。再構築されたフィルム画像のシーケンスは、1 秒間の 1 / 2 4 の定期的な間隔で写真記録されたため、GOP パーティション 6 1 2 又は復号器 2 2 8 において行われた動き推定プロセスは、時間の基礎が不正規であるテレシネされたデータではなく逆テレシネされた画像を用いるほうがより正確である。

【 0 0 4 3 】

一側面においては、位相検出器 6 0 4 は、映像フレームを受信後に一定の決定を行う。これらの決定は、( i ) 現在の映像がテレシネ出力からであり 3 : 2 プルダウン位相が図 3 8 に示される 5 つの位相  $P_0$ 、 $P_1$ 、 $P_2$ 、 $P_3$ 、及び  $P_4$  のうちの 1 つであるかどうか、及び ( i i ) 映像は従来の NTSC として生成されたかどうかを含む。その決定は、位相  $P_5$  として表される。これらの決定は、図 2 に示される位相検出器 6 0 4 の出力として現れる。“はい”のラベルが付された位相検出器 6 0 4 からの経路は、逆テレシネ 6 0 6 を作動させ、同じ写真画像から形成されたフィールドを分類して結合させることができるように正確なプルダウン位相が提供されていることを示す。“いいえ”のラベルが付された位相検出器 6 0 4 からの経路は、みかけの NTSC フレームを最適な処理のためにフィールドに分離するためにデインターレーサ 6 0 5 を同様に作動させる。異なる型の映像を何時でも受信することができるため、位相検出器 6 0 4 は、映像フレームを連続的に解析することができる。一例として、NTSC 基準に準拠した映像をコマーシャルとして映像内に挿入することができる。逆テレシネ後、結果的に得られたプログレッシブ映像は、白色ガウス雑音を低減させるために用いることができる雑音除去器（フィルタ）6 0 7 に送られる。

【 0 0 4 4 】

従来の NTSC 映像が認識されたときには（位相検出器 6 0 1 からの「いいえ」経路）、圧縮のためにデインターレーサ 6 0 5 に送信される。デインターレーサ 6 0 5 は、インターレーシングされたフィールドをプログレッシブ映像に変換し、プログレッシブ映像に関する雑音除去動作を行うことができる。1 つのデインターレーシング処理例が以下において説明される。

【 0 0 4 5 】

テレビ等の伝統的なアナログ映像デバイスは、インターレーシングされた方法で映像を提供する。すなわち、該デバイスは、偶数の走査線（偶数フィールド）及び奇数の走査線（奇数フィールド）を提供する。信号サンプリングの観点からは、このことは、以下の式によって表されるパターンでの空間 - 時間的サブサンプリングに相当する。

【 数 1 】

$$F(x, y, n) = \begin{cases} \Theta(x, y, n), & \text{偶数フィールドに対して } y \bmod 2 = 0 \text{ のとき、} \\ \Theta(x, y, n), & \text{奇数フィールドに対して } y \bmod 2 = 1 \text{ のとき、} \\ \text{消去、その他の場合} \end{cases} \quad [1]$$

【 0 0 4 6 】

10

20

30

40

50

ここで、 $\Theta$  は、原フレームピクチャを表し、 $F$  は、インターレーシングされたフィールドを表し、 $(x, y, n)$  は、画素の水平方向の、垂直方向の、及び時間的位置を表す。

【0047】

一般性を失うことなしに、本開示全体を通じて  $n = 0$  は偶数フィールドであると仮定することができる、このため上記の方程式 1 は以下のように単純化される。

【数 2】

$$F(x, y, n) = \begin{cases} \Theta(x, y, n), & y \bmod 2 = n \bmod 2 \text{ のとき、} \\ \text{消去、その他の場合} \end{cases} \quad [2]$$

10

【0048】

デシメーションは水平次元においては行われなため、サブサンプリングパターンは、次の  $n \sim y$  座標で描くことができる。

【0049】

デインターレーサの最終目標は、インターレーシングされた映像（フィールドのシーケンス）をインターレーシングされていないプログレッシブフレーム（フレームのシーケンス）に変換することである。換言すると、偶数及び奇数のフィールドを内挿し、全フレームピクチャを“復元”又は生成することである。この動作は、方程式 3 によって表すことができる。

20

【数 3】

$$F_o(x, y, n) = \begin{cases} F(x, y, n), & y \bmod 2 = n \bmod 2, \\ F_i(x, y, n), & \text{その他の場合} \end{cases} \quad [3]$$

30

【0050】

ここで、 $F_i$  は、消失している画素に関するデインターレーシング結果を表す。

【0051】

図 40 は、インターレーシングされたマルチメディアデータからプログレッシブフレームを生成するために  $Wmed$  フィルタリング及び動き推定を用いるデインターレーサ 605 の側面の一定の側面を示すブロック図である。図 40 の上部分は、現在のフィールド、2つの前フィールド（ $PP$  フィールド及び  $P$  フィールド）、及び2つの後続フィールド（次のフィールド及び次の次のフィールド）からの情報を用いて生成することができる動き強度マップ 4002 を示す。以下においてさらに詳細に説明されるように、動き強度マップ 4002 は、現在のフレームを分類するか又は2つ以上の異なる動きレベルに分割することができる、空間-時間的フィルタリングによって生成することができる。幾つかの側面においては、動き強度マップ 4002 は、以下の方程式 4 乃至 8 を参照しつつ説明されるように、静止エリア、スローモーションエリア、及び高速モーションエリアを識別するために生成される。空間-時間フィルタ、例えば、 $Wmed$  フィルタ 4004 は、インターレーシングされたマルチメディアデータを動き強度マップに基づいて判定基準を用いてフィルタリングし、空間-時間的な仮のデインターレーシングされたフレームを生成する。幾つかの側面においては、 $Wmed$  フィルタリングプロセスは、 $[-1, 1]$  の水平の近隣エリアと、 $[-3, 3]$  の垂直の近隣エリアと、図 40 に示される5つのフィールド（ $PP$  フィールド、 $P$  フィールド、現在のフィールド、次のフィールド、次の次のフィールド）によって表される5つの隣接フィールドから成る時間的近隣エリアと、を含み、 $Z$

40

50

<sup>1</sup> は、1つのフィールドの遅延を表す。現在のフィールドに関して、次のフィールド及びPフィールドは、非パリティフィールドであり、PPフィールド及び次の次のフィールドは、パリティフィールドである。空間-時間的フィルタリングに関して用いられる“近隣エリア”は、フィルタリング動作中に実際に用いられるフィールド及び画素の空間的及び時間的位置を意味し、例えば図6及び7において示されるように「開口」として示すことができる。

#### 【0052】

デインターレーサ605は、Wmedフィルタ4004によって生成された空間-時間的な仮のデインターレーシングされたフレームをフィルタリングするように構成された雑音除去器(雑音除去フィルタ)4006を含むこともできる。空間-時間的な仮のデインターレーシングされたフレームの雑音を除去することは、特にソースインターレーシングされたマルチメディアデータシーケンスが白色雑音によって汚染されている場合に、後続する動き探索プロセスをより正確にする。さらに、Wmedピクチャ内の偶数と奇数の行間におけるエリアスを少なくとも部分的に除去することもできる。雑音除去器4006は、ウェーブレット縮退及びウェーブレットウィナーフィルタに基づく雑音除去器を含む様々なフィルタとして実装することができる。雑音除去器は、候補のWmedフレームが動き補償情報を用いてさらに処理される前に候補のWmedフレームから雑音を除去するために用いることができ、Wmedフレーム内に存在する雑音を除去すること及び信号の周波数コンテンツにかかわらず信号の存在を保持することができる。ウェーブレットフィルタを含む様々な型の雑音除去フィルタを用いることができる。ウェーブレットは、空間領域及びスケーリング領域の両方において所定の信号を突き止めるために用いられる1つのクラスの関数である。ウェーブレットに関する基本的な考え方は、ウェーブレット表示の小さい変化が原信号の対応する小さい変化を生み出すように異なるスケール又は解像度において信号を解析することである。

#### 【0053】

ウェーブレット縮退又はウェーブレットウィナーフィルタは、雑音除去器として利用することもできる。ウェーブレット縮退は、雑音信号のウェーブレット変換を行い、後続して、大きい係数は影響を受けないようにしながら小さいウェーブレット係数を縮退させてゼロ(又はより小さい値)にすることから成る。最後に、逆変換が行われて推定信号が取得される。

#### 【0054】

雑音除去フィルタリングは、雑音環境における動き補償精度を向上させる。ウェーブレット縮退雑音除去は、ウェーブレット変換領域の縮退を含むことができ、典型的には、3つのステップ、すなわち、線形順ウェーブレット変換、非線形縮退雑音除去、及び線形逆ウェーブレット変換、を具備する。ウィナーフィルタは、加法性雑音及びぶれによって劣化された画像を改良するために用いることができるMSE最適線形フィルタである。該フィルタは、当業において一般的に知られており、例えば、上記の“ウェーブレット縮退による理想的な空間適応”において、及び本明細書において参照されることによってその全体が明示で本明細書に組み入れられている、S.P. Ghael, A. M. Sayeed, and R. G. Baraniuk, “Improvement Wavelet denoising via empirical Wiener filtering,” (経験的ウィナーフィルタリングによるウェーブレット雑音除去の向上)、Proceedings of SPIE, vol. 3169, pp. 389 - 399, San Diego, July 1997によって説明されている。

#### 【0055】

幾つかの側面においては、雑音除去フィルタは、(4, 2)双直交立方Bスプラインウェーブレットフィルタの側面に基づく。1つの該フィルタは、以下の順及び逆変換によって定義することができる。

【数4】

$$h(z) = \frac{3}{4} + \frac{1}{2}(z + z^{-1}) + \frac{1}{8}(z + z^{-2}) \quad (\text{順変換}) \quad [4]$$

【0056】

及び

【数5】

$$g(z) = \frac{5}{4}z^{-1} - \frac{5}{32}(1 + z^{-2}) - \frac{3}{8}(z + z^{-3}) - \frac{3}{32}(z^2 + z^{-4}) \quad (\text{逆変換})$$

10

【0057】

雑音除去フィルタの利用は、雑音環境における動き補償精度を向上させることができる。該フィルタの実装は、本明細書において参照されることによってその全体が明示で本明細書に組み入れられている、“Ideal spatial adaptation by wavelet shrinkage,” (ウェーブレット縮退による理想的な空間適応) D.L. Donoho and I.M. Johnstone, Biometrika, vol. 8, pp. 425 - 455, 1994、においてさらに説明されている。

20

【0058】

図40の下部は、インターレーシングされたマルチメディアデータの動き情報(例えば、動きベクトル候補、動き推定、動き補償)を決定するための側面を示す。特に、図40は、選択されたフレームの動き補償された仮のプログレッシブフレームを生成してWmed仮フレームと結合させてその結果得られる、デインターレーシングされた現在のフレーム4014として示される、“最終的な”プログレッシブフレームを形成するために用いられる動き推定及び動き補償方式を示す。幾つかの側面においては、インターレーシングされたマルチメディアデータの動きベクトル(“MV”)候補(又は推定)が、外部の動き推定器からデインターレーサに提供され、双方向動き推定器及び補償器(“ME/MC”)4018に関する出発点を提供するために用いられる。幾つかの側面においては、MV候補選択器4022は、処理中のブロックのMV候補に関する近隣ブロックに関する以前に決定されたMV、例えば、前の処理されたブロック、例えば、デインターレーシングされた前フレーム4020内のブロック例のMV、を使用する。動き補償は、前のデインターレーシングされたフレーム70及び次の(例えば、将来の)Wmedフレーム4008に基づいて双方向で行うことができる。現在のWmedフレーム4010及び動き補償された(“MC”)現在のフレーム4016は、結合器4012によってマージ、すなわち結合される。その結果得られたデインターレーシングされた現在のフレーム4014、現在はプログレッシブフレーム、は、デインターレーシングされた前フレーム4020として用いるためにME/MC4018に戻され、さらに後続処理のためにデインターレーサ605に外部通信される。

30

40

【0059】

インターフィールド補間を具備するデインターレーシング予測方式を、Wmed + MCデインターレーシング方式によるイントラフィールド補間から切り離すことが可能である。換言すると、空間-時間的Wmedフィルタリングは、主にイントラフィールド補間目的で用いることができ、インターフィールド補間は、動き補償中に行うことができる。このことは、Wmed結果のピーク信号・雑音比を小さくするが、不正確なインターフィールド予測モード決定による不良な画素は、Wmedフィルタリングプロセスから除去されるため、動き補償が適用された後の視覚品質のほうが良い。

【0060】

50

適切な逆テレシネ又はデインターレーシング処理後は、ブロック608において、プログレッシブ映像は、エリアス抑止及びリサンプリング（例えば、リサイジング）に関して処理される。幾つかのリサンプリング側面においては、ピクチャサイズのリサイジングのために多相リサンプリングが実装される。ダウンサンプリングの一例においては、原ピクチャとリサイジングされたピクチャの比は、 $p/q$ であることができ、 $p$ 及び $q$ は、相対的に素の整数である。総位相数は、 $p$ である。幾つかの側面における多相フィルタのカットオフ周波数は、約0.5のリサイジング率に関しては0.6である。リサイジングされたシーケンスの高周波数応答を向上させるために、カットオフ周波数は、リサイジング比とは正確には一致しない。このことは、不可避免的に何らかのエリアシングを可能にする。しかしながら、人間の目は、ぶれたエリアスのないピクチャよりも鮮明でエリアスが存在するピクチャを好むことがよく知られている。

10

【0061】

図41は、多相リサンプリング例であり、リサイジング比が $3/4$ である場合の位相を示す。図41に示されるカットオフ周波数も $3/4$ である。上図においては、原画素は縦軸によって示される。フィルタ波形を表すためにこれらの軸を中心にした $\text{sinc}$ 関数も描かれている。選択されたカットオフ周波数は、リサンプリング比とまったく同じであるため、 $\text{sinc}$ 関数のゼロは、図41において交差によって示される、リサイジング後の画素の位置と重なり合う。リサイジング後の画素値を見つけるために、貢献度を以下の方程式において示されるように原画素から合計することができる。

【数6】

20

$$v(x) = \sum_{i=-\infty}^{\infty} u(i) \times \text{sinc}(\pi f_c (i-x)) \quad [6]$$

【0062】

ここで、 $f_c$ は、カットオフ周波数である。上記の1D多相フィルタは、水平次元及び垂直次元の両方に適用することができる。

【0063】

30

他のリサンプリング（リサイジング）側面は、オーバースキャンを説明することである。NTSCテレビ信号においては、画像は、486本の走査線を有しており、デジタル映像においては、各走査線上に720の画素を有することが可能である。しかしながら、テレビ上では、サイズと画面形式との間の不一致に起因して画像全体が見えるわけではない。画像のうちで見えない部分は、オーバースキャンと呼ばれる。

【0064】

米映画テレビ技術者協会（SMPTE）では、放送事業者が可能な限り多くのテレビにおいて有用情報を可視エリア内に入れやすいようにするため、セーフアクションエリア（safe action area）及びセーフタイトルエリア（safe title area）と呼ばれる動作フレームの特定の大きさを定めた。SMPTEによって推奨される方法である「テレビシステムに関するセーフアクションエリア及びセーフタイトルエリアテストパターンに関する仕様」に関するRP27.3-1989を参照すること。セーフアクションエリアは、“すべての意味のある動作が行われなければならない”エリアであるとSMPTEによって定義されている。セーフタイトルエリアは、“大部分の家庭用テレビ受像機における可視性を保証するためにすべての有用な情報を入れることができる”エリアであると定義されている。

40

【0065】

例えば、図25に関して、セーフアクションエリア2510は、画面の中央の90%を占め、5%のボーダーが全周囲に存在する。セーフタイトルエリア2505は、画面の中央の80%を占め、10%のボーダーが存在する。次に図26に関して、セーフタイトル

50

エリアは非常に小さいため、幾つかの放送局では、画像内においてより多くのコンテンツを追加するため、白色の長方形の窓 2 6 1 5 の内部に存在するセーフアクションエリア内にテキストを入れる。

【 0 0 6 6 】

通常は、オーバースキャン内には黒のボーダーを見ることができる。例えば、図 2 6 において、画像の上側 2 6 2 0 及び下側 2 6 2 5 に黒のボーダーが現れる。H . 2 6 4 映像では動き推定の際に境界拡大を用いるため、これらの黒のボーダーはオーバースキャンで除去することができる。拡大された黒のボーダーは、残留物を増大させる可能性がある。控えめなことに、境界は 2 % だけカットし、次にリサイジングを行うことができる。リサイジング用フィルタは、適宜生成することができる。多相ダウンサンプリング前にオーバースキャンを除去するための打ち切りが行われる。

10

【 0 0 6 7 】

再度図 6 に関して、プログレッシブ映像は、ブロック 6 1 0 に進み、デブロッカー及びデリンギング動作が行われる。映像圧縮用途においては、2 つの型のアーティファクト、“ブロッキング”及び“リンギング”が共通して発生する。ブロッキングアーティファクトは、圧縮アルゴリズムが各フレームをブロック（例えば、8 × 8 のブロック）に分割することが原因で発生する。各ブロックは、幾つかの小さい誤りが存在する状態で再構築され、ブロックのエッジにおける誤りは、近隣ブロックのエッジにおける誤りとコントラストを成すことがしばしばあり、ブロック境界を可視にする。対照的に、リンギングアーティファクトは、画像特徴のエッジの周囲における歪みとして現れる。リンギングアーティファクトは、符号器が高周波 D C T 係数を量子化する際に多くの情報を捨てすぎることが原因で発生する。幾つかの説明例においては、デブロッキング及びデリンギングは両方とも、低域 F I R（有限インパルス応答）を用いてこれらの可視のアーティファクトを隠すことができる。

20

【 0 0 6 8 】

デブロッキング処理の一例においては、フレームの境界におけるエッジ及びデブロッキングフィルタプロセスが使用不能になるエッジを除くフレームのすべての 4 × 4 ブロックエッジにデブロッキングフィルタを適用することができる。このフィルタリングプロセスは、マクロブロックアドレスの昇順で処理されたフレーム内の全マクロブロックによるフレーム構築プロセスが完了後にマクロブロックに基づいて行わなければならない。各マクロブロックに関して、最初に垂直エッジが左から右にフィルタリングされ、その後に水平エッジが上から下にフィルタリングされる。図 3 9 に示されるように、水平方向及び垂直方向に関して、4 つの 1 6 サンプルエッジに対してルマデブロッキングフィルタプロセスが行われ、2 つの 8 サンプルエッジに対して各クロマ成分に関するデブロッキングフィルタプロセスが行われる。以前のマクロブロックに対するデブロッキングプロセス作業によって既に修正されている現在のマクロブロックの左上のサンプル値は、現在のマクロブロックに対するデブロッキングフィルタプロセスへの入力として用いなければならない。現在のマクロブロックのフィルタリング中にさらに修正することができる。垂直エッジのフィルタリング中に修正されたサンプル値は、同じブロックに関する水平エッジのフィルタリング用入力として用いることができる。デブロッキングプロセスは、ルマ成分及びクロマ成分に関して別々に呼び出すことができる。

30

40

【 0 0 6 9 】

デリンギング処理の一例においては、エッジ付近のエリアを平滑化するために 2 D フィルタを適応的に用いることができる。エッジ画素は、ぶれを回避するためのフィルタリングがほとんど又はまったく行われたい。

【 0 0 7 0 】

G O P パーティショナ

デブロッキング及びデリンギング後は、プログレッシブ映像は、G O P パーティショナ 6 1 2 によって処理される。G O P ポジショニングは、ショット変化を検出することと、複雑さマップ（例えば、時間的、空間的帯域幅マップ）を生成することと、適応型 G O P

50

パーティショニングと、を含むことができる。これらの各々が以下において説明される。

【 0 0 7 1 】

A . シーン変化検出

ショット検出は、ピクチャグループ ( GOP ) 内のフレームがシーンの変化が生じていることを示すデータを提示したときを決定することに関するものである。一般的には、GOP 内においては、フレームは、2 , 3 ( 又はそれ以上の ) 隣接フレーム内では有意な変化はしていない場合があり、又は、低速の変化又は高速の変化が生じている場合もある。当然のことであるが、これらのシーン変化分類は、必要な場合は、特定の用途に依存してより大きな変化レベルに細分することができる。

【 0 0 7 2 】

ショット又はシーンの変化の検出は、映像の効率的な符号化にとって重要である。典型的には、GOP が有意な変化をしていないときには、GOP の最初の I フレームは、幾つかの予測フレームによって後続され、映像の後続する復号及び表示が視覚的に受け入れ可能になるように十分に映像を符号化することができる。しかしながら、突然又はゆっくりにかかわらずシーンが変化中のときには、後続して復号された視覚的に受け入れ可能な結果を生み出すために追加の I フレーム及びそれよりも予測性が低い符号化 ( P フレーム及び B フレーム ) が必要になることがある。

【 0 0 7 3 】

既存の符号化システムの性能を向上させるショット検出及び符号化システム及び方法が以下において説明される。該側面は、プリプロセッサ 2 2 6 の GOP パーティショナ 6 1 2 に実装すること ( 図 7 )、又はプリプロセッサとともに又はプリプロセッサなしで動作することができる符号器デバイス内に含めることができる。該側面は、突然のシーン変化が生じているかどうか、シーンがゆっくりと変化しているかどうか、又は映像符号化を特に複雑化するカメラのフラッシュがシーン内に存在するかどうかを決定するために隣接する映像データフレーム間の統計的比較を含む統計値 ( 又はメトリックス ) を利用する。これらの統計値は、プリプロセッサから入手して符号化デバイスに送ることができ、又は ( 例えば、動き補償を行うように構成されたプロセッサによって ) 符号化デバイス内において生成することができる。結果的に得られた統計値は、シーン変化検出決定に役立つ。トランスコーディングを行うシステムにおいては、適切なプリプロセッサ又は構成可能なプロセッサがしばしば存在する。プリプロセッサがデインターレーシングによって補助された動き補償を行う場合は、動き補償統計値を利用可能であり、直ちに用いることができる。該システムにおいては、ショット検出アルゴリズムは、システムの複雑さをわずかに増大させる可能性がある。

【 0 0 7 4 】

本明細書において説明されるショット検出器例は、前フレーム、現在のフレーム、及び次のフレームからの統計値を利用する必要があるだけであり、従って、非常に低いレテンシーを有する。ショット検出器は、突然のシーンの変化、クロスフェーディングとその他の低速のシーン変化、及びカメラのフラッシュを含む幾つかの異なる型のショットイベントを区別する。符号器内の異なる戦略を用いて異なる型のショットイベントを決定することによって、符号化効率及び視覚品質が向上される。

【 0 0 7 5 】

シーン変化検出は、映像コーディングシステムが I フレームを固定された間隔で挿入することによってビットを知的に保存するようにするために用いることができる。幾つかの側面においては、( 例えば、メタデータ内に組み込まれた又はプリプロセッサ 2 2 6 によって計算された ) プリプロセッサによって入手されたコンテンツ情報は、シーン変化検出のために用いることができる。例えば、コンテンツ情報に依存し、後述されるしきい値及びその他の判定基準は、異なる型の映像コンテンツに関して動的に調整することができる。

【 0 0 7 6 】

映像符号化は、通常は、構造化されたピクチャグループ ( GOP ) に対して行われる。

10

20

30

40

50

GOPは、通常はイントラコーディングされたフレーム（Iフレーム）から開始し、一連のP（予測）又はB（双方向）フレームが後続する。典型的には、Iフレームは、フレームを表示するために要求される全データを保存することができ、Bフレームは、（前フレームから変更されたデータを含むだけであるか又は次のフレーム内のデータとは異なる）前フレーム及び後続フレーム内のデータに依存し、Pフレームは、前フレームから変化しているデータを含む。共通の使用法においては、符号化された映像においては、Iフレームは、Pフレーム及びBフレームとともに散在する。サイズ（例えば、フレームの符号化に用いられるビット数）の点では、Iフレームは、典型的には、Pフレームよりもはるかに大きく、Pフレームは、Bフレームよりも大きい。効率的な符号化、送信及び復号処理を行うために、GOPの長さは、大きなIフレームからの効率の損失を低減させる上で十分に長く、符号器と復号器との間の mismatch、すなわちチャネル損傷、に対処する上で十分に短くすべきである。さらに、Pフレーム内のマクロブロック（MB）は、同じ理由でイントラコーディングすることができる。

10

**【0077】**

シーン変化検出は、映像符号器が適切なGOP長を決定し、しばしば不要なIフレームを固定された間隔で挿入する代わりにGOP長に基づいてIフレームを挿入するために用いることができる。実際のストリーミング映像システムにおいては、通常は、ビットエラー又はパケット損失によって通信チャネルが損傷される。Iフレーム又はIMBをどこに配置するかは、復号された映像の品質及び視覚上の経験に有意な影響を与える可能性がある。1つの符号化方式は、共配置されている前ピクチャ又はピクチャ部分からの有意な変化を有するピクチャ又はピクチャ部分に関してイントラコーディングされたフレームを用いることである。通常は、これらの領域は、動き推定では有効に及び効率的に予測することはできず、符号化は、該領域がインターフレームコーディング技術（例えば、Bフレーム及びPフレームを用いた符号化）から免除されている場合により効率的に行うことができる。チャネル損傷の状況においては、これらの領域は、誤り伝播を被る可能性があり、該誤り伝播は、イントラフレーム符号化によって低減又は排除（又はほぼ排除）することができる。

20

**【0078】**

GOP映像の一部分は、2つ以上のカテゴリに分類することができ、各領域は、特定の実装に依存する異なるイントラフレーム符号化基準を有することができる。一例として、映像は、3つのカテゴリ、すなわち、突然のシーン変化、クロスフェーディングとその他の低速のシーン変化、及びカメラのフラッシュライト、に分類することができる。

30

**【0079】**

突然のシーンの変化は、通常はカメラの操作に起因する、前フレームとは有意な量だけ異なるフレームを含む。これらのフレームのコンテンツは、前フレームのコンテンツと異なるため、突然のシーン変化のフレームは、Iフレームとして符号化すべきである。

**【0080】**

クロスフェーディング及びその他の低速のシーンの変化は、通常はコンピュータによるカメラ撮影の処理に起因する、シーンの低速切り換えを含む。2つの異なるシーンを漸次混合することは、人間の目には見やすいように思えるが、映像コーディング上は難題である。動き補償は、これらのフレームのビットレートを有効に低下させることができず、これらのフレームに関してより多くのイントラMBを更新することができる。

40

**【0081】**

カメラのフラッシュライト、又はカメラのフラッシュイベントは、フレームのコンテンツがカメラのフラッシュを含むときに発生する。該フラッシュは、継続時間が相対的に短く（例えば、1フレーム）、極端に明るいため、フラッシュを描くフレーム内の画素は、隣接フレーム上の対応するエリアと比較して異常に高い輝度を呈する。カメラのフラッシュライトは、ピクチャの輝度を突然かつ素早く変化させる。通常は、カメラのフラッシュライトの継続時間は、人間の視覚系（HVS）の時間的マスキング継続時間よりも短く、典型的には44msであると定義されている。人間の目は、これらの短時間の突然の明る

50

さの品質は感じず、このため粗く符号化することができる。フラッシュライトフレームは、動き補償では有効に処理することができず、将来のフレームに関する不良な予測候補であるため、これらのフレームの粗い符号化は、将来のフレームの符号化効率を低下させない。フラッシュライトとして分類されたシーンは、「人工的な」高輝度に起因してその他のフレームを予測するために用いるべきではなく、さらに、その他のフレームも、同じ理由でこれらのフレームを予測するために有効に用いることができない。これらのフレームは、いったん識別された時点で、相対的に大量の処理を要求する可能性があるため取り出すことができる。1つの選択肢は、カメラのフラッシュライトフレームを除去してその代わりにDC係数を符号化することである。該解決方法は単純で、高速計算が可能であり、多くのビットを節約する。

10

## 【0082】

上記のフレームカテゴリのいずれかが検出されたときに、ショットイベントが宣言される。ショット検出は、符号化品質を向上させる上で有用であるだけでなく、映像コンテンツの探索及びインデキシングにおいても役立つことができる。シーン検出プロセスの1つの例示的側面が以下において説明される。この例においては、ショット検出プロセスは、最初に、ショット検出のための処理対象となっている選択されたフレームに関する情報、又はメトリックスを計算する。メトリックスは、映像の双方向動き推定及び補償処理からの情報と、その他の輝度に基づくメトリックスと、を含むことができる。

## 【0083】

双方向動き推定/補償を行うためには、現在のフレームのすべての $8 \times 8$ ブロックを、最も近くに隣接するフレームのうちの2つ、過去の1つ、及び将来の1つ、の中のブロックとマッチングさせる双方向動き補償器で映像シーケンスを前処理することができる。動き補償器は、すべてのブロックに関する動きベクトル及び差分メトリックスを生成する。図29は、現在のフレームCの画素を過去のフレームP及び将来の(又は次の)フレームNとマッチングさせる例を示し、マッチングされた画素への動きベクトル(過去の動きベクトル $MV_P$ 及び将来の動きベクトル $MV_N$ )を描く図である。以下では、双方向動きベクトル生成及び関連する符号化についての一般的説明が図32を参照しながら一般的に示される。

20

## 【0084】

双方向動き情報、例えば、対応する隣接フレーム内の(最良のマッチングであった)MBを識別する動き情報、を決定後、現在のフレームと次のフレーム及び前フレームとの様々な比較によって(例えば、GOPパーティション612内の動き補償器又は他の適切な構成要素によって)追加のメトリックスを生成することができる。動き補償器は、すべてのブロックに関する差分メトリックスを生成することができる。差分メトリックスは、差分二乗和(SSD)又は差分絶対値和(SAD)であることができる。本明細書においては、一般性を失うことなしに、SADが一例として用いられる。

30

## 【0085】

すべてのフレームに関して、SAD比、“コントラスト比”とも呼ばれる、は以下のよう

## 【数7】

40

$$\gamma = \frac{\varepsilon + SAD_P}{\varepsilon + SAD_N} \quad [6]$$

## 【0086】

ここで、 $SAD_P$ 及び $SAD_N$ は、前進及び後退差分メトリックスのそれぞれの差分絶対値和である。分母は、「ゼロによる割り算」の誤りを防止するために小さい正数を含むことに注目すべきである。分子も、分母内の1の影響を均衡化させるためのを含む。例えば、前フレーム、現在のフレーム、及び次のフレームが同一である場合は、動き探索は

50

、 $SAD_P = SAD_N = 0$ を生み出すはずである。この場合は、上記計算は、0又は無限大ではなく  $= 1$ を生成する。

【0087】

すべてのフレームに関して輝度ヒストグラムを計算することができる。典型的には、マルチメディア画像は、8ビットの輝度深度（例えば、“ビン”数）を有する。幾つかの側面により輝度ヒストグラムを計算するために用いられる輝度深度は、ヒストグラムを入手するために16に設定することができる。その他の側面においては、輝度深度は、処理中のデータの型、利用可能な計算能力、又はその他の予め決められた判定基準に依存する適切な数字に設定することができる。幾つかの側面においては、輝度深度は、計算された又は受信されたメトリックス、例えばデータのコンテンツ、に基づいて動的に設定することができる。

10

【0088】

以下の方程式は、輝度ヒストグラム差分（ラムダ）の一計算例を示す。

【数8】

$$\lambda = \frac{\sum_{i=1}^{16} |N_{Pi} - N_{Ci}|}{N} \quad [7]$$

20

【0089】

ここで、 $N_{Pi}$ は、前フレームに関する*i*番目のビン内のブロック数であり、 $N_{Ci}$ は、現在のフレームに関する*i*番目のビン内のブロック数であり、 $N$ は、フレーム内の総ブロック数である。前フレーム及び現在のフレームの輝度ヒストグラム差分が完全に異なる（又はばらばらである）場合は、 $= 2$ である。

【0090】

この情報を用いて、フレーム差分メトリックス（ $D$ ）は、以下のように計算される。

【数9】

$$D = \frac{\gamma_C}{\gamma_P} + A\lambda(2\lambda + 1) \quad [8]$$

30

【0091】

ここで、 $A$ は、用途別に選択される定数であり、

【数10】

$$\gamma_C = \frac{\varepsilon + SAD_P}{\varepsilon + SAD_N}, \quad \text{及び} \quad \gamma_P = \frac{\varepsilon + SAD_{PP}}{\varepsilon + SAD_C}$$

40

【0092】

である。

【0093】

選択された（現在の）フレームは、フレーム差分メトリックスが方程式9に示される判定基準を満たしている場合は突然のシーン変化として分類される。

【数 1 1】

$$D = \frac{\gamma_c}{\gamma_p} + A\lambda(2\lambda + 1) \geq T_1 \quad [9]$$

【0094】

ここで、Aは、用途別に選択される定数であり、 $T_1$ は、しきい値である。

【0095】

一例において、シミュレーションにより、 $A = 1$ 及び $T_1 = 5$ に設定することは、優れた検出性能を達成させることが示される。現在のフレームが突然のシーン変化フレームである場合は、 $\gamma_c$ は大きくすべきであり、 $\gamma_p$ は小さくすべきである。比 $\gamma_c / \gamma_p$ は、メトリックスがコンテキストの活動レベルに合わせて正規化されるように $\gamma_c$ 単独の代わりに用いることができる。

【0096】

上記判定基準は、輝度ヒストグラム差分(ラムダ)を非線形で用いることに注目すべきである。図16は、 $\lambda^*(2\lambda + 1)$ は、凸関数であることを示す。 $\lambda$ が小さい(例えば、ゼロに近い)ときは、ほとんどプリエンファシスである。 $\lambda$ が大きいほど、より多くのエンファシスが関数によって行われる。このプリエンファシスにより、1.4よりも大きいあらゆる $\lambda$ に関して、しきい値 $T_1$ が5に設定されている場合に突然のシーン変化が検出される。

【0097】

現在のフレームは、シーン強度メトリックスDが一定数の連続フレームに関して方程式5に示される判定基準を満たしている場合にクロスフェーディング又は低速のシーン変化であると決定される。

【数 1 2】

$$T_2 \leq D < T_1 \quad [10]$$

【0098】

ここで、 $T_1$ は、上記の同じしきい値であり、 $T_2$ は、他のしきい値である。

【0099】

フラッシュライトイベントは、通常は、輝度ヒストグラムをより明るい側に移動させる。この例示的側面のカメラにおいては、輝度ヒストグラム統計値は、現在のフレームがカメラのフラッシュライトを具備するかどうかを決定するために用いられる。方程式11及び12に示されるように、ショット検出プロセスは、現在のフレームマイナスの輝度が前フレームの輝度よりも一定のしきい値 $T_3$ だけ大きいかどうか、及び現在のフレームの輝度が次のフレームの輝度よりもしきい値 $T_3$ だけ大きいかどうかを決定することができる。

【数 1 3】

$$\bar{Y}_C - \bar{Y}_P \geq T_3 \quad [11]$$

$$\bar{Y}_C - \bar{Y}_N \geq T_3 \quad [12]$$

10

20

30

40

50

## 【 0 1 0 0 】

上記の判定基準が満たされない場合は、現在のフレームは、カメラのフラッシュライトを具備するとして分類されない。上記の判定基準が満たされている場合は、以下の方程式において示されるように、ショット検出プロセスは、後退差分メトリックス  $SAD_p$  及び前進差分メトリックス  $SAD_N$  が一定のしきい値  $T_4$  よりも大きいかどうかを決定する。

【数 1 4】

$$SAD_p \geq T_4 \quad [13]$$

$$SAD_N \geq T_4 \quad [14]$$

10

## 【 0 1 0 1 】

ここで、

【数 1 5】

$$\bar{Y}_C$$

## 【 0 1 0 2 】

は、現在のフレームの平均輝度であり、

【数 1 6】

$$\bar{Y}_p$$

20

## 【 0 1 0 3 】

は、前フレームの平均輝度であり、

【数 1 7】

$$\bar{Y}_N$$

30

## 【 0 1 0 4 】

は、次のフレームの平均輝度であり、 $SAD_p$  及び  $SAD_N$  は、現在のフレームと関連づけられた前進及び後退差分メトリックスである。

【 0 1 0 5 】

ショット検出プロセスは、現在のフレームの輝度が前フレームの輝度及び次のフレームの輝度よりも大きいかどうかを最初に決定することによってカメラのフラッシュイベントを決定する。より大きくない場合は、フレームはカメラのフラッシュイベントではない。しかしながら、より大きい場合は、フレームはカメラのフラッシュイベントである可能性がある。次に、ショット検出プロセスは、後退差分メトリックスがしきい値  $T_3$  よりも大きいかどうか、及び前進差分メトリックスがしきい値  $T_4$  よりも大きいかどうかを評価することができる。これらの両方の条件が満たされている場合は、ショット検出プロセスは、現在のフレームがカメラのフラッシュライトを有するとして分類する。判定基準が満たされていない場合は、フレームは、いずれの型のショットイベントとしても分類されず、又は、フレームに関して行うべき符号化を識別するデフォルト分類を与えることができる（例えば、フレームをドロップする、Iフレームとして符号化する）。

40

【 0 1 0 6 】

上記においては、 $T_1$ 、 $T_2$ 、 $T_3$ 、及び  $T_4$  に関する幾つかの典型値が示されている。典型的には、これらのしきい値は、ショット検出の特定の実装の試験を通じて選択され

50

る。幾つかの側面においては、しきい値  $T_1$ 、 $T_2$ 、 $T_3$ 、及び  $T_4$  のうちの1つ以上が予め決定され、該値は、符号化デバイス内のショット分類器内に組み入れられる。幾つかの側面においては、しきい値  $T_1$ 、 $T_2$ 、 $T_3$ 、及び  $T_4$  のうちの1つ以上は、ショット分類器に供給された情報（例えば、メタデータ）に基づいて又はショット分類器自体によって計算された情報に基づいて処理中に（例えば、動的に）設定することができる。

【0107】

ショット検出情報を用いて映像を符号化することは、典型的には符号器内において行われるが、本明細書においては、ショット検出開示を完全にすることを目的として説明される。図30に関して、符号化プロセス301は、ショット検出情報を用いて、フレームシーケンス内の検出されたショットに基づいて映像を符号化することができる。プロセス301は、ブロック303に進み、現在のフレームが突然のシーン変化として分類されているかどうかを確認する。そのように分類されている場合は、ブロック305において、現在のフレームは、Iフレームとして符号化することができ、GOP境界を決定することができる。そのように分類されていない場合は、プロセス301は、ブロック307に進む。現在のフレームがブロック309においてゆっくりと変化しているシーンの一部として分類されている場合は、現在のフレーム、及びゆっくりと変化しているシーン内のその他のフレームは、予測フレーム（例えば、Pフレーム又はBフレーム）として符号化することができる。次に、プロセス301は、ブロック311に進み、現在のフレームがカメラのフラッシュを具備するフラッシュライトシーンとして分類されているかどうかを確認する。そのように分類されている場合は、ブロック313において、フレームは、特別な処理、例えば、除去又はフレームに関するDC係数の符号化、を目的として識別することができる。そのように分類されていない場合は、現在のフレームの分類は行われておらず、現在のフレームは、その他の判定基準に従って符号化すること、Iフレームとして符号化すること、又はドロップすることができる。

10

20

【0108】

上述される側面においては、圧縮すべきフレームと隣接する2つのフレームとの間の差分量は、フレーム差分メトリックスDによって示される。有意な量の一方向の輝度変化が検出された場合は、フレーム内におけるクロスフェード効果であることを意味する。クロスフェードがより顕著になるほど、Bフレームを用いることによってより多くの利得を達成することができる。幾つかの側面においては、以下の方程式において示されるように修正されたフレーム差分メトリックスが用いられる。

30

【数18】

$$D_1 = \begin{cases} \left(1 - \alpha + 2\alpha \frac{|d_p - d_N|}{d_p + d_N}\right) \times D, & Y_p - \Delta \geq Y_C \geq Y_N + \Delta \text{ or } Y_p + \Delta \leq Y_C \leq Y_N - \Delta \text{ のとき,} \\ D, & \text{その他の場合} \end{cases} \quad [15]$$

【0109】

ここで、 $d_p = |Y_C - Y_p|$  及び  $d_N = |Y_C - Y_N|$  は、それぞれ、現在のフレームと前フレームとの間のルマ差分、及び現在のフレームと次のフレームとの間のルマ差分であり、 $\alpha$  は、実装に依存することができるため通常の実験において決定することができる定数であり、 $\Delta$  は、0乃至1の間の値を有する重み付け変数である。

40

【0110】

B. 帯域幅マップ生成

プリプロセッサ226（図6）は、マルチメディアデータを符号化するために用いることができる帯域幅マップを生成するように構成することもできる。幾つかの側面においては、符号器228（図7）内のコンテンツ分類モジュール712が代わりに帯域幅マップを生成する。

【0111】

50

人間の視覚品質  $V$  は、符号化の複雑さ  $C$  及び割り当てられたビット  $B$  (帯域幅とも呼ばれる) の両方の関数であることができる。図 15 は、この関係を示すグラフである。符号化の複雑さメトリックス  $C$  は、人間の視覚の観点から空間周波数及び時間周波数を考慮することが注目されるべきである。人間の目がより感じやすい歪みに関しては、複雑さ値はそれに対応して高くなる。典型的には、 $V$  は、 $C$  においては単調減少しており、 $B$  においては単調増加していると仮定することができる。

【0112】

一定の目視品質を達成するために、符号化対象であって以下の 2 つの方程式において表される判定基準を満たしている  $i$  番目のオブジェクト (フレーム又は  $MB$ ) に帯域幅 ( $B_i$ ) が割り当てられる。

【数 19】

$$B_i = B(C_i, V) \quad [16]$$

$$B = \sum_i B_i \quad [17]$$

【0113】

上記の 2 つの方程式において、 $C_i$  は、 $i$  番目のオブジェクトの符号化の複雑さであり、 $B$  は、利用可能な総帯域幅であり、 $V$  は、オブジェクトに関する達成された視覚品質である。人間の視覚品質は、方程式として公式化するのは困難である。従って、上記の方程式は、正確には定義されていない。しかしながら、3D モデルがすべての変数において連続的であると仮定した場合、帯域幅比 ( $B_i / B$ ) は、( $C, V$ ) 対の近隣においては不変であるとして処理することができる。帯域幅比  $\beta_i$  は、以下の方程式において定義される。

【数 20】

$$\beta_i = B_i / B \quad [18]$$

【0114】

これで、ビット割り当ては、以下の方程式において表されるように定義することができる。

【数 21】

$$\beta_i = \beta(C_i)$$

$$1 = \sum_i \beta_i \quad \text{但し } (C_i, V) \in \delta(C_0, V_0) \quad [19]$$

【0115】

ここで、 $\delta$  は、“近隣”を示す。

【0116】

符号化の複雑さは、空間及び時間の両方において、人間の視覚感度によって影響される

10

20

30

40

50

。ジロッドの人間の視力モデルは、空間上の複雑さを定義するために用いることができるモデルの一例である。このモデルは、局部空間周波数及び周囲照明を検討する。その結果得られたメトリックスは、 $D_{csat}$  と呼ばれる。プロセス内の前処理時点において、ピクチャがイントラコーディングされるか又はインターコーディングされるかは不明であり、両方に関する帯域幅比が生成される。ビットは、異なる映像オブジェクトの  $I_{INTRA}$  間の比に従って割り当てられる。イントラコーディングされたピクチャに関しては、帯域幅比は、以下の方程式で表される。

【数 2 2】

$$\beta_{INTRA} = \beta_{0INTRA} \log_{10} (1 + \alpha_{INTRA} Y^2 D_{csat}) \quad [20]$$

10

【0 1 1 7】

上記方程式において、 $Y$  は、マクロブロックの平均輝度成分であり、 $I_{INTRA}$  は、二乗輝度に関する重み付け係数であり、 $D_{csat}$  項が後続し、 $\beta_{0INTRA}$  は、

【数 2 3】

$$1 = \sum_i \beta_i$$

20

【0 1 1 8】

を保証するための正規化係数である。例えば、 $I_{INTRA}$  に関する値 = 4 は、良い視覚品質を達成させる。コンテンツ情報（例えば、コンテンツ分類）は、映像の特定のコンテンツに関する希望される良い視覚品質に対応する値に  $I_{INTRA}$  を設定するために用いることができる。一例においては、映像コンテンツが“画面上にアナウンサーが映し出される”ニュース放送を具備する場合は、情報画像及び映像の表示可能な部分は、音声部分よりも重要性が低いとみなすことができるため、視覚品質レベルをより低く設定することができ、データを符号化するためにより少ないビットを割り当てることができる。他の例においては、映像コンテンツがスポーツ行事を具備する場合は、コンテンツ情報は、表示された画像のほうが視聴者にとってより重要である可能性があるためより高い視覚品質に対応する値に  $I_{INTRA}$  を設定するために用いることができ、従ってデータを符号化するためにより多くのビットを割り当てることができる。

30

【0 1 1 9】

この関係について理解するために、帯域幅は符号化の複雑さと対数的に割り当てられることが注目されるべきである。二乗輝度項  $Y^2$  は、規模がより大きい係数はより多くのビットを符号化のために用いるということを反映している。対数が負の値になるのを防止するため、括弧内の項に 1 が加えられる。その他の基を有する対数も使用可能である。

【0 1 2 0】

時間的複雑さは、フレーム差分メトリックス、例えば絶対値差分和 (SAD)、とともに動き量（例えば、動きベクトル）を考慮に入れて 2 つの連続するフレーム間の差分を測定するフレーム差分メトリックス評価基準によって決定される。

40

【0 1 2 1】

インターコーディングされたピクチャに関するビット割り当ては、空間的及び時間的複雑さを考慮することができる。このことは、以下のように表される。

【数 2 4】

$$\beta_{INTER} = \beta_{0INTER} \log_{10} \left( 1 + \alpha_{INTER} \cdot SSD \cdot D_{cstat} \exp(-\gamma \|MV_P + MV_N\|^2) \right) \quad [21]$$

【0122】

上記方程式において、 $MV_P$  及び  $MV_N$  は、現在の MB に関する順方向及び逆方向動きベクトルである（図 29 参照）。イントラコーディングされた帯域幅式における  $Y^2$  は、二乗差分和（SSD）によって代えられることを注記することができる。上記方程式における  $MV_P + MV_N^2$  の役割について理解するために、人間の視覚系の次の特徴に注目すること。すなわち、スムーズで予測可能な動き（小さい  $MV_P + MV_N^2$ ）が生じるエリアは、目を引き、目によって追跡することができ、典型的には、静止領域よりも多くの歪みを許容することができない。しかしながら、高速の又は予測不能な動き（大きい  $MV_P + MV_N^2$ ）が生じるエリアは、追跡することができず、有意な量子化を許容することができる。実験では、 $INTER = 1$ 、 $\beta = 0.001$  が良い視覚品質を達成させることを示している。

10

【0123】

C. 適応型 GOP パーティショニング

プリプロセッサ 226 によって行うことができる他の処理例においては、図 6 の GOP パーティショナ 612 は、いっしょにコーディングされたピクチャグループの構成を適応的に変更することができ、MPEG2 を用いる例を参照しながら説明される。幾つかのこれまでの映像圧縮基準（例えば、MPEG2）は、正規の構造を課すことはできるが GOP が正規の構造を有することを要求しない。MPEG2 シーケンスは、常に I フレーム、すなわち、以前のピクチャを基準にせず符号化されているフレーム、から始まる。MPEG2 GOP 形式は、通常は、I フレームに後続する P 又は予測ピクチャの GOP 内のスペースを固定することによって符号器において予め手配されている。P フレームは、一部分が以前の I 又は P フレームから予測されているピクチャである。開始 I フレームと後続する P フレームとの間のフレームは、B フレームとして符号化される。“B” フレーム（B は双方向を表す）は、前の及び次の I 又は P ピクチャを基準として個々に又は同時に用いることができる。I フレームを符号化するために必要なビット数は、平均すると、P フレームを符号化するために必要なビット数を上回り、同様に、P フレームを符号化するために必要なビット数は、平均すると、B フレームに関して要求されるビット数を上回る。スキップされたフレームは、使用される場合は、表示するためのビットは必要ない。

20

30

【0124】

P 及び B フレームを使用する上での基本概念は、さらにはより最近の圧縮アルゴリズムにおいては、映像を表現するために必要なデータのレートを低くするためにフレームをスキップすることは、時間的冗長性を排除することである。時間的冗長性が高い、すなわちピクチャ間での変化がほとんどない、ときには、以前に復号された I 又は P ピクチャはのちにその他の P 又は B ピクチャを復号するための基準として用いられるため、P、B、又はスキップされたピクチャの使用は、映像ストリームを効率的に表現する。

40

【0125】

適応型 GOP パーティショニングは、この概念を適応的に用いることに基づく。フレーム間の差分は定量化され、ピクチャを I、P、B、又はスキップされたフレームによって表す決定は、定量化された差分に関する適切な試験が行われた後に自動的に実施される。適応型構造は、固定型 GOP 構造では入手できない利点を有する。固定された構造は、コンテンツの変化がほとんど起こっていない可能性を無視し、適応型手順は、それよりもはるかに多くの B フレームを各 I と P との間、又は 2 つの P の間に挿入するのを可能にし、それによって、フレームシーケンスを適切に表現するために必要なビット数を少なくする。逆に、映像コンテンツの変化が有意であるときには、予測されたフレームと基準フレー

50

ムとの間の差分が大きすぎるため、Pフレームの効率は大幅に低下される。これらの状態  
下においては、マッチングオブジェクト ( m a t c h i n g o b j e c t ) が動き探索  
領域から外れる可能性があり、又はカメラの角度の変化による歪みに起因してマッチング  
オブジェクト間の類似性が低下する。その時点において、Pフレーム又はI及びその隣接  
するPフレームは、互いにより接近しているものを選択すべきであり、より少ない数のB  
フレームを挿入すべきである。固定されたGOPは、その調整を行うことができない。

【 0 1 2 6 】

本明細書において開示されるシステムにおいては、これらの状態は自動的に検知される  
。GOP構造は柔軟であり、これらのコンテンツの変化に適応するように構築される。シ  
ステムは、距離に関して同じ追加性質を有する、フレーム間の距離の評価基準であると考  
えることができるフレーム差分メトリックスを評価する。概念上は、フレーム間の距離  $d_{12}$  及び  $d_{23}$  を有するフレーム  $F_1$ 、 $F_2$ 、及び  $F_3$  が与えられている場合、 $F_1$  と  $F_3$  との間の距離は、少なくとも  $d_{12} + d_{23}$  であるとされる。この距離に関するメトリ  
ックスに基づいてフレーム割り当てが行われる。

10

【 0 1 2 7 】

GOPパーティショナは、受信されたフレームにピクチャ型を割り当てることによって  
動作する。ピクチャ型は、各ブロックをコーディングする際に要求される予測方法を示す  
。

【 0 1 2 8 】

Iピクチャは、その他のピクチャを基準にせずにコーディングされる。これらのピクチャ  
は独立型であるため、復号が開始することができるデータストリーム内においてアクセ  
スポイントを提供する。前フレームまでの“距離”がシーン変化しきい値を超える場合は  
、I符号化型がフレームに割り当てられる。

20

【 0 1 2 9 】

Pピクチャは、動き補償予測に関して前のI又はPピクチャを用いることができる。こ  
れらのピクチャは、予測対象となっているブロックから変位する ( d i s p l a c e ) こ  
とができる以前のフィールド又はフレーム内のブロックを符号化のための基礎として用い  
る。基準ブロックが考慮対象ブロックから減じられた後に、残留ブロックは、典型的には  
空間的冗長性の除去に関する離散コサイン変換を用いて符号化される。フレームとPフレ  
ームになるために割り当てられた最後のフレームとの間の“距離”が、典型的には第1の  
しきい値よりも小さい第2のしきい値を超える場合は、P符号化型がフレームに割り当て  
られる。

30

【 0 1 3 0 】

Bフレームピクチャは、上述されるように、前の及び次のP又はIピクチャを動き補償  
に関して用いることができる。Bピクチャ内のブロックは、前方、後方又は双方向に予測  
することができ、又は、その他のフレームを基準にせずにイントラコーディングすること  
が可能である。H. 264においては、基準ブロックは、32のフレームからの32のブ  
ロックの線形的組合せであることができる。フレームをI又はP型に割り当てることがで  
きない場合は、該フレームとその直前のフレームとの間の“距離”が、典型的には第2の  
しきい値よりも小さい第3のしきい値よりも大きい場合に、B型として割り当てられる。

40

【 0 1 3 1 】

フレームが符号化されたBフレームになるために割り当てることができない場合は、“  
スキップフレーム”状態に割り当てられる。このフレームは、実質的には前フレームのコ  
ピーであるためスキップすることができる。

【 0 1 3 2 】

表示順の点で隣接するフレーム間の差分を定量化するメトリックスを評価することは、  
行われるこの処理の第1の部分である。このメトリックスは、上述される距離であり、こ  
のメトリックスを用いてすべてのフレームの適切な型を評価することができる。従って、  
Iと隣接するPとの間、又は2つの連続するPフレーム間のスペースは、可変であること  
ができる。メトリックスの計算は、ブロックに基づく動き補償器によって映像フレームを

50

処理することから開始し、ブロックは、映像圧縮の基本単位であり、通常は  $16 \times 16$  画素から成るが、 $8 \times 8$ 、 $4 \times 4$  及び  $8 \times 16$  等の他のブロックサイズも可能である。2つのデインターレーシングされたフィールドから成るフレームに関しては、動き補償は、フィールドごとに行うことができ、基準ブロックの探索は、フレーム内ではなくフィールド内において行われる。現在のフレームの第1のフィールド内のブロックに関しては、後続するフレームのフィールドに順方向基準ブロックが存在し、同様に、現在のフレームの直前のフレームのフィールドにおいて逆方向基準ブロックが存在する。現在のブロックは、組み立てられて補償されたフィールドに入れられる。プロセスは、フレームの第2のフィールドが続く。2つの補償されたフィールドが結合されて、順方向及び逆方向補償フレームを形成する。

10

## 【0133】

逆テレシネにおいて生成されたフレーム606に関しては、基準ブロックの探索は、再構築されたフィルムフレームだけであるためフレームのみに基づく。2つの基準ブロック及び2つの差分、前進及び後退、が見つげ出され、さらに順方向及び逆方向補償フレームにも至る。要約すると、動き補償器は、すべてのブロックに関する動きベクトル及び差分メトリックスを生成するが、ブロックは、デインターレーサ605の出力が処理される場合はNTSCフィールドの一部であり、逆テレシネの出力が処理される場合はフィルムフレームの一部である。メトリックスの差分は、考慮中のフィールド又はフレーム内のブロックと、前進差分又は後退差分のいずれを評価中であるかに依存して、前のフィールド又はフレーム又は直後のフィールド又はフレーム内における前記ブロックと最も良くマッチするブロックとの間におけるメトリックスの差分が評価されることに注目すること。この計算には、輝度値のみが含まれる。

20

## 【0134】

動き補償ステップは、2組の差分を生成する。これらは、現在の輝度値のブロックと、時間的に現在のフレームの直前及び直後のフレームから取り出された基準ブロックにおける輝度値との間の差分である。各前進差分及び各後退差分の絶対値は、各画素に関して決定され、各々がフレーム全体に関して別々に合計される。フレームを具備するデインターレーシングされたNTSCフィールドが処理されるときには両フィールドが2つの合計計算の中に入れられる。この方法により、前進差分及び後退差分の合計された絶対値  $SAD_p$  及び  $SAD_N$  が見つげ出される。

30

## 【0135】

すべてのフレームに関して、この関係を用いて  $SAD$  比が計算される。

## 【数25】

$$\gamma = \frac{\varepsilon + SAD_p}{\varepsilon + SAD_N} \quad [22]$$

## 【0136】

ここで、 $SAD_p$  及び  $SAD_N$  は、前進差分及び後退差分のそれぞれの合計された絶対値である。“ゼロによる割り算”の誤りを防止するために小さい正数が分子に加えられる。同様の項が分母に加えられ、 $SAD_p$  又は  $SAD_N$  のいずれかがゼロに近いときにの感度をさらに低下させる。

40

## 【0137】

代替側面においては、差分は、SSD、二乗差分の和、及び  $SAD$ 、絶対値差分和、又は  $SATD$  であることができ、ブロック要素の差分が取られる前に二次元離散コサイン変換を適用することによって画素値ブロックが変換される。これらの和は、アクティブ映像のエリア全体にわたって評価されるが、その他の側面においてはより小さいエリアを用いることができる。

50

## 【 0 1 3 8 】

受信された状態の（動き補償されない）すべてのフレームの輝度ヒストグラムも計算される。ヒストグラムは、輝度値ブロックが利用可能な場合に二次元離散コサイン変換を輝度値ブロックに適用した結果である  $16 \times 16$  の係数配列内の DC 係数、すなわち  $(0, 0)$  係数、に基づいて動作する。同様に、ヒストグラムでは、 $16 \times 16$  ブロック内の 256 の輝度値の平均値を用いることができる。輝度深度が 8 ビットである画像に関しては、ビン数は 16 に設定される。次のメトリックスは、ヒストグラム差分を評価する。

## 【 数 2 6 】

$$\lambda = \frac{1}{N} \sum_{i=1}^{16} |N_{Pi} - N_{Ci}| \quad [23]$$

10

## 【 0 1 3 9 】

上記方程式において、 $N_{Pi}$  は、 $i$  番目のビン内の前フレームからのブロック数であり、 $N_{Ci}$  は、 $i$  番目のビン内に属する現在のフレームからのブロック数であり、 $N$  は、フレーム内の総ブロック数である。

## 【 0 1 4 0 】

これらの中間結果が組み立てられて現在のフレームの差分メトリックスが以下のように形成される。

20

## 【 数 2 7 】

$$D = \frac{\gamma_C}{\gamma_P} + \lambda(2\lambda + 1) \quad [24]$$

## 【 0 1 4 1 】

ここで、 $\gamma_C$  は、現在のフレームに基づく SAD 比であり、 $\gamma_P$  は、前フレームに基づく SAD 比である。シーンがスムーズな動きを有し、そのルマヒストグラムがほんのわずかに変化する場合は、 $D$  はほぼ 1 に等しい。現在のフレームが突然のシーン変化を表示する場合は、 $\gamma_C$  は大きくなり、 $\gamma_P$  は小さくすべきである。メトリックスがコンテキストの活動レベルに合わせて正規化されるようにするため  $\gamma_C$  単独の代わりに比

30

## 【 数 2 8 】

$$\frac{\gamma_C}{\gamma_P}$$

## 【 0 1 4 2 】

が用いられる。

40

## 【 0 1 4 3 】

図 4 2 は、圧縮型をフレームに割り当てるプロセスを示す。方程式 1 9 において定義される現在のフレームの差分  $D$  が、フレーム割り当てに関して行われる決定のための基礎である。決定ブロック 4 2 0 2 が示すように、考慮中のフレームがシーケンス内の第 1 のフレームである場合は、はいと記された決定経路をたどってブロック 4 2 0 6 に進み、それによってフレームは I フレームである宣言する。ブロック 4 2 0 8 において、累積されたフレーム差分がゼロに設定され、プロセスは（ブロック 4 2 1 0 において）開始ブロックに戻る。考慮中のフレームがシーケンス内の第 1 のフレームでない場合は、決定が行われたブロック 4 2 0 2 からいいえと記された経路をたどり、テストブロック 4 2 0 4 におい

50

て、現在のフレーム差分がシーン変化しきい値に照らして試験される。現在のフレームの差分がそのしきい値よりも大きい場合は、はいと記された決定経路をたどってブロック 4 2 0 6 に進み、I フレームの割り当てが再度行われる。

#### 【 0 1 4 4 】

現在のフレームの差分が、シーン変化しきい値よりも小さい場合は、いいえ経路をたどってブロック 4 2 1 2 に進み、現在のフレームの差分が累積フレーム差分に加えられる。流れ図をさらにたどり決定ブロック 4 2 1 4 において、累積されたフレーム差分が、一般的にはシーン変化しきい値よりも小さいしきい値  $t$  と比較される。累積されたフレーム差分が  $t$  よりも大きい場合は、制御はブロック 4 1 2 6 に移行し、フレームは P フレームになるために割り当てられ、ステップ 4 2 1 8 において累積されたフレーム差分がゼロに設定される。累積されたフレーム差分が  $t$  よりも小さい場合は、制御は、ブロック 4 2 1 4 からブロック 4 2 2 0 に移行する。ブロック 4 2 2 0 において、現在のフレームの差分が  $t$  よりも小さい と比較される。現在のフレームの差分が よりも小さい場合は、該フレームは、ブロック 4 2 2 2 においてスキップすることが割り当てられ、プロセスが戻る。現在のフレームの差分が よりも大きい場合は、該フレームは、ブロック 4 2 2 6 において B フレームとなることが割り当てられる。

#### 【 0 1 4 5 】

##### 符号器

再度図 2 に関して、トランスコーダ 2 0 0 は、処理されたメタデータ及び生映像をプリプロセッサ 2 2 6 から受信する符号器 2 2 8 を含む。メタデータは、ソース映像 1 0 4 において最初に受け取られた情報と、プリプロセッサ 2 2 6 によって計算された情報と、を含むことができる。符号器 2 2 8 は、第 1 パス符号器 2 3 0 と、第 2 パス符号器 2 3 2 と、再符号器 2 3 4 と、を含む。符号器 2 2 8 は、第 2 パス符号器 2 3 2 から、プリプロセッサ 2 2 6 と同様に第 1 パス符号器 2 3 0、及び再符号器 2 3 4 に情報（例えば、メタデータ、誤り耐性情報、コンテンツ情報、符号化されたビットレート情報、基本層及び拡張層のバランス情報、及び量子化情報）を提供することができるトランスコーダ制御 2 3 1 からの入力を受信する。符号器 2 2 8 は、プリプロセッサ 2 2 6 から受信されたコンテンツ情報及び / 又は符号器 2 2 8 自体によって、例えば、コンテンツ分類モジュール 7 1 2（図 7）によって生成されるコンテンツ情報を用いて、受信された映像を符号化する。

#### 【 0 1 4 6 】

図 7 は、図 2 に示される符号器 2 2 8 に関して用いることができる典型的な 2 パス符号器に含めることができる機能モジュールのブロック図である。図 7 には機能モジュールの様々な側面が示されるが、図 7 及び本明細書における説明は、符号器内に組み込むことができる全機能に必ずしも対処しているわけではない。従って、以下においては、機能モジュールの一定の側面は、後述される基本層及び拡張層の符号化に関する解説に沿って説明される。

#### 【 0 1 4 7 】

##### 基本層及び拡張層の符号化

符号器 2 2 8 は、SNR スケラブル符号器であることができ、SNR スケラブル符号器は、プリプロセッサ 2 2 6 からの生映像及びメタデータを符号化し、本明細書においては基本層とも呼ばれる第 1 の符号化データグループ、及び本明細書においては拡張層とも呼ばれる 1 つ以上の追加の符号化データグループに分けることができる。符号化アルゴリズムは、基本層係数及び拡張層係数を生成し、これらの係数は、復号時に、両層が復号に関して利用可能なときに復号器において結合することができる。両層が利用可能でないときには、基本層の符号化は、基本層を単一層として復号することを可能にする。

#### 【 0 1 4 8 】

該多層符号化プロセスの一側面が、図 3 1 を参照しながら説明される。ブロック 3 2 1 において、完全にイントラコーディングされたマクロブロック（イントラコーディングされた MB）によって I フレームが符号化される。H . 2 6 4 においては、I フレーム内のイントラコーディングされた MB は、有意な量のコーディング利得を提供する完全に利用

10

20

30

40

50

された空間予測によって符号化される。2つのサブモード、すなわち、イントラ4×4及びイントラ16×16が存在する。基本層が空間予測によって提供されたコーディング利得を利用する場合は、基本層は、拡張層を符号化及び復号する前に符号化及び復号する必要がある。Iフレームの2パス符号化及び復号が用いられる。基本層において、基本層量子化パラメータ $QP_b$ は、粗い量子化ステップサイズを変換係数に提供する。拡張層では、原フレームと再構築された基本層フレームとの間の画素に関する差分が符号化される。拡張層は、より微細な量子化ステップサイズを提供する量子化パラメータ $QP_e$ を用いる。符号化手段、例えば図2の符号器228、は、ブロック321において符号化を行う。

【0149】

ブロック323において、符号器は、処理中のGOP内のP及び/又はBフレームに関する基本層データ及び拡張層データを符号化する。符号化手段、例えば符号器228は、ブロック323において符号化を行うことができる。ブロック325において、符号化プロセスは、符号化すべきさらなるP又はBフレームが存在するかどうか確認する。符号化手段、例えばSNRスケラブル符号器228は、325において動作することができる。より多くのP又はBフレームが残っている場合は、GOP内の全フレームの符号化が終了するまでステップ323が繰り返される。P及びBフレームは、インターコーディングされたマクロブロック(インターコーディングされたMB)から成るが、P及びBフレーム内には後述されるようにイントラコーディングされたMBが存在することができる。

【0150】

復号器が基本層データと拡張層データを区別するようにするために、符号器228は、ブロック327においてオーバーヘッド情報を符号化する。オーバーヘッド情報の型は、例えば、層数を識別するデータ、層を基本層として識別するデータ、層を拡張層として識別するデータ、層間の関係を識別するデータ(例えば、層2は、基本層1に関する拡張層であり、層3は、拡張層2に関する拡張層である)、又は層を一連の拡張層における最後の拡張層として識別するデータを含む。オーバーヘッド情報は、関係する基本層及び/又は拡張層データと結びつけられたヘッダー内に含めること、又は別個のデータメッセージ内に含めることができる。符号化手段、例えば図2の符号器228、は、ブロック327においてプロセスを実行することができる。

【0151】

単層復号を有するようにするためには、2つの層の係数を逆量子化前に結合させなければならない。従って、これらの2つの層の係数は、対話形式で生成しなければならない。その他の場合は、有意な量のオーバーヘッドを導入する可能性がある。オーバーヘッドが増大する1つの理由は、基本層符号化及び拡張層符号化が異なる時間的基準を用いることが可能なためである。基本層及び拡張層の係数を生成するためにはアルゴリズムが必要であり、これらの係数は、両層が利用可能なときに逆量子化前に復号器において結合させることができる。同時に、アルゴリズムは、拡張層が利用可能でないときに又は復号器が例えば電力節約等の理由で拡張層を復号しないことに決定したときに受け入れ可能な基本層映像に対応すべきである。以下では、該プロセスの一例の詳細が、直下における標準的な予測コーディングに関する概要を示しつつ説明される。

【0152】

Pフレーム(又はインターコーディングされたセクション)は、現在のピクチャ内の領域と基準ピクチャ内の最もマッチングしている予測領域との間の時間的冗長性を利用することができる。基準フレーム内の最もマッチングする予測領域の位置は、動きベクトルで符号化することができる。現在の領域フレームと最もマッチングしている基準予測領域との間の差は、残留誤り(又は予測誤り)と呼ばれる。

【0153】

図32は、例えばMPEG-4でのPフレーム構築プロセス例を示す。プロセス331は、図31のブロック323において生じる可能性があるプロセス例をより詳細に示す。プロセス331は、5×5マクロブロックから成る現在のピクチャ333を含み、ここで、この例におけるマクロブロック数は任意である。マクロブロックは、16×16画素か

10

20

30

40

50

ら成る。画素は、8ビット輝度値(Y)及び2つの8ビット色差値(Cr及びCb)によって定義することができる。MPEGにおいては、Y、Cr及びCb成分は、4:2:0形式で格納することができ、Cr及びCb成分は、X方向及びY方向に2だけダウンサンプリングされる。従って、各マクロブロックは、256のY成分、64のCr成分及び64のCb成分から成ることになる。現在のピクチャ333のマクロブロック335は、現在のピクチャ333とは異なる時点において基準ピクチャ337から予測される。Y、Cr及びCb値の点で符号化中の現在のマクロブロック335に最も近い最もマッチングしているマクロブロック339の位置を突き止めるための探索が基準ピクチャ337において行われる。基準ピクチャ337内の最もマッチングしているマクロブロック339の位置は、動きベクトル341において符号化される。基準ピクチャ337は、復号器が現在のピクチャ333の構築前に再構築していることになるIフレーム又はPフレームであることができる。最もマッチングしているマクロブロック339が現在のマクロブロック335から減じられ(Y、Cr及びCb成分の各々に関する差分が計算され)その結果残留誤り343が発生する。残留誤り343は、2D離散コサイン変換(DCT)345によって符号化され、次に量子化される347。量子化347は、例えば、より少ないビットを高周波数係数に割り当てより多くのビットを低周波数係数に割り当てることによって空間圧縮を提供するために行うことができる。残留誤り343の量子化された係数は、動きベクトル341及び基準ピクチャ333識別情報とともに、現在のマクロブロック335を表す符号化された情報である。符号化された情報は、将来用いるためにメモリに格納するか又は例えば誤り訂正又は画像拡張を目的として作業を行うこと、又はネットワーク349を通じて送信することができる。

#### 【0154】

残留誤り343の符号化された量子化された係数は、符号化された動きベクトル341とともに、後続する動き推定及び補償に関する基準フレームの一部として用いるために符号器内において現在のマクロブロック335を再構築するために用いることができる。符号器は、このPフレーム再構築に関して復号器の手順をエミュレーションすることができる。復号器のエミュレーションは、符号器及び復号器の両方が同じ基準ピクチャに関して作業をすることになる。本明細書においては、再構築プロセスは、符号器内において行われるか、さらなるインターコーディングのために行われるか、又は復号器において行われるかに関わらず提示される。Pフレームの再構築は、基準フレーム(又は参照中のピクチャ又はフレームの一部分)が再構築された後に開始することができる。符号化された量子化された係数は、逆量子化351され、次に、2D逆DCT、又はIDCT 353が行われ、その結果復号された又は再構築された残留誤り355が生じる。符号化された動きベクトル341は、復号され、既に再構築されている基準ピクチャ337内の既に再構築されている最もマッチングしているマクロブロック357を突き止めるために用いられる。次に、再構築された残留誤り355が再構築された最もマッチングするマクロブロック357に加えられて再構築されたマクロブロック359が形成される。再構築されたマクロブロック359は、メモリに格納すること、独立して又はその他の再構築されたマクロブロックとともにピクチャ内に表示すること、又は画像エンハンスメントのためにさらに処理することができる。

#### 【0155】

Bフレーム(双方向予測でコーディングされたセクション)は、現在のピクチャ内の領域、前フレーム内の最もマッチングしている予測領域及び後続ピクチャ内の最もマッチングしている予測領域の間における時間的冗長性を利用することができる。後続する最もマッチングしている予測領域及び前の最もマッチングしている予測領域が結合されて、結合された双方向の予測された領域が形成される。現在のピクチャ領域と最もマッチングしている結合された双方向予測領域との間の差分が残留誤り(又は予測誤り)である。後続する基準ピクチャ内の最もマッチングしている予測領域及び前基準ピクチャ内の最もマッチングしている予測領域の位置は、2つの動きベクトルで符号化することができる。

#### 【0156】

10

20

30

40

50

図33は、符号器228によって行うことができる基本層及び拡張層の係数の符号化に関する符号器プロセス例を示す。基本層及び拡張層が符号化されてSNRスケラブルビットストリームが提供される。図33は、図31のステップ323において行われるようなインターMB残留誤り係数符号化例を示す。しかしながら、イントラMB係数を符号化するために同様の方法を用いることも可能である。符号化手段、例えば図2の符号器構成要素228、は、図33に示されるプロセス及び図32のステップ323を実行することができる。(符号化される)原映像データ406(映像データは、この例においてはルマ及びクロマ情報を具備する)は、基本層の最もマッチングしているマクロブロック302及び拡張層の最もマッチングしているマクロブロック365に入力される。両ループ363及び365の目的は、加算器367及び369においてそれぞれ計算される残留誤りを最小にすることである。ループ363及び365は、示されるように並行して、又は順次で実行することができる。ループ363及び365は、最もマッチングしているマクロブロックと原データ361との間の残留誤りを最小にする最もマッチングしているマクロブロックを識別するために、基準フレームを含むバッファ371及び373をそれぞれ探索するための論理を含む(バッファ371及び373は、同じバッファであることができる)。基本層ループ363は一般的には拡張層ループ365よりも粗い量子化ステップサイズ(より高いQP値)を利用するため、ループ363及びループ365の残留誤りは異なることになる。変換ブロック375及び377は、各ループの残留誤りを変換する。

10

【0157】

20

次に、変換された係数は、選択器379において構文解析されて基本層係数および拡張層係数になる。選択器379の構文解析は、後述されるように幾つかの形態をとることができる。構文解析技術の1つの共通する特長は、拡張層係数 $C'_{enh}$ は、基本層係数 $C'_{base}$ のディファレンシャルリファインメントになるような形で計算されることである。基本層のリファインメントになるような形で拡張層を計算することは、復号器がそれ自体によって基本層係数を復号して画像の妥当な表現を有すること、又は基本層係数と拡張層係数を結合して画像のリファインメントされた表現を有することを可能にする。次に、選択器379によって選択された係数が量子器381および383によって量子化される。(量子化器381及び383によってそれぞれ計算された)量子化された係数

【数29】

30

$$\tilde{C}'_{base} \text{ 及び } \tilde{C}'_{enh}$$

【0158】

は、メモリに格納するか又はネットワークを通じて復号器に送信することができる。

【0159】

復号器内のマクロブロックの再構築とマッチングさせるために、逆量子化器385は、基本層残留誤り係数を逆量子化する。逆量子化された残留誤り係数は、逆変換され387、バッファ371内で見つけ出された最もマッチングするマクロブロックに追加され389、その結果復号器内において再構築されるものとマッチングする再構築されたマクロブロックが得られる。量子化器383、逆量子化器391、逆変換器393、加算器397及びバッファ373は、基本層ループ363において行われたのと同様の計算を拡張ループ365において行う。さらに、加算器393は、拡張層の再構築において用いられた逆量子化された拡張層係数及び基本層係数を結合させるために用いられる。拡張層量子化器及び逆量子化器は、一般的には、基本層よりも微細な量子化器ステップサイズ(より低いQP)を利用する。

40

【0160】

図34、35及び36は、図33の選択器379において採用することができる基本層

50

係数及び拡張層係数の選択器プロセスの例を示す。図2の符号器228等の選択する手段は、図34、35及び35において描かれるプロセスを実行することができる。図34を一例として用いて、変換された係数は、構文解析され、以下の方程式において示されるように基本層係数及び拡張層係数に分けられる。

【数30】

$$C'_{base} = \begin{cases} 0, & C_{base} \text{ 及び } C_{enh} \text{ が反対の符号のとき} \\ \min(C_{base}, C_{enh}), & \text{その他の場合} \end{cases} \quad [25]$$

$$C'_{enh} = C_{enh} - Q_b^{-1}(Q_b(C'_{base})) \quad [26]$$

10

【0161】

ここで、“min”関数は、2つの引数の数学上の最小値又は最小の大きさであることができる。方程式25は、ブロック401として描かれ、方程式26は、図34の加算器510として描かれている。方程式26において、 $Q_b$ は、基本層量子化器381を表し、 $Q_b^{-1}$ は、基本層の逆量子化器385を表す。方程式26は、拡張層係数を、方程式25を用いて計算された基本層係数のディファレンシャルリファインメントに変換する。

【0162】

図35は、基本層及び拡張層係数選択器379の他の例を示す。この例においては、ブロック405に含まれている方程式( )は、以下を表す。

【数31】

$$C'_{base} = \begin{cases} C_{base}, & |Q_b^{-1}Q_b(C_{base}) - C_{enh}| < |C_{enh}| \text{ のとき} \\ 0, & \text{その他の場合} \end{cases} \quad [27]$$

20

30

【0163】

加算器407は、以下の2つの方程式において示されるように拡張層係数を計算する。

【数32】

$$C'_{enh} = C_{enh} - Q_b^{-1}(Q_b(C'_{base})) \quad [28]$$

【0164】

ここで、 $C'_{base}$ は、方程式27によって与えられる。

40

【0165】

図36は、基本層及び拡張層選択器379の他の例を示す。この例においては、基本層係数は不変であり、拡張層は、量子化/逆量子化された基本層係数と原拡張層係数との間の差に等しい。

【0166】

基本層及び拡張層の残留誤り係数に加えて、復号器は、MBがどのように符号化されているかを識別する情報が必要である。符号化手段、例えば図2の符号器構成要素228は、イントラコーディング及びインターコーディングされた部分のマップ、例えば、マクロブロック(又はサブマクロブロック)がイントラコーディング又はインターコーディングされているとして識別され(さらに、例えば、順方向、逆方向又は双方向

50

を含むいずれの型のインターコーディングであるか及びインターコーディングされた部分がいずれのフレームを基準にしているかも識別する) MB マップ、を含むことができるオーバーヘッド情報を符号化することができる。一側面例においては、MB マップ及び基本層係数は、基本層において符号化され、拡張層係数は、拡張層において符号化される。

【0167】

P フレーム及び B フレームは、インター MB と同様にイントラコーディングされた MB も含むことができる。ハイブリッド映像符号器は、P 又は B フレーム内の一定のマクロブロックをイントラコーディングされた MB として符号化することを決定するためにレート歪み (RD) 最適化を用いるのが一般的である。イントラコーディングされた MB が拡張層インター MB に依存しない単層復号を有するために、近隣のインター MB は、基本層イントラコーディングされた MB の空間予測には関して用いられない。計算の複雑さが拡張層復号に関して常に不変であるようにするために、基本層の P 又は B フレーム内のイントラコーディングされた MB に関して、拡張層におけるリファインメントをスキップをことが可能である。

10

【0168】

P 又は B フレーム内のイントラコーディングされた MB は、インター MB よりも多くのビットを要求する。この理由により、P 又は B フレーム内のイントラコーディングされた MB は、より高い QP で基本層品質のみで符号化可能である。このことは、映像品質のある程度の劣化をもたらすことになるが、この劣化は、上述されるように基本層及び拡張層においてインター MB 係数を用いてのちのフレーム内においてリファインメントされる場合は目立たないはずである。2つの理由がこの劣化を目立たなくする。第1の理由は、人間の視覚系 (HVS) の特徴であり、他方の理由は、インター MB がイントラ MB をリファインメントすることである。第1のフレームから第2のフレームに位置を変化させるオブジェクトの場合は、第1のフレーム内の幾つかの画素は、第2のフレームでは見ることができず (カバーされる情報)、第2のフレーム内の幾つかの画素は、初めて見ることができる (カバーされない情報)。人間の目は、カバーされない及びカバーされる視覚情報に対して敏感でない。このため、カバーされていない情報に関しては、より低い品質で符号化されるにもかかわらず、目は違いを見分けることができない。同じ情報が後続する P フレーム内にとどまっている場合は、拡張層はより低い QP を有するため拡張層における後続 P フレームがその情報をリファインメントできる確率が高くなる。

20

30

【0169】

P 又は B フレームにおいてイントラコーディングされた MB を導入する他の共通技術は、イントラリフレッシュと呼ばれる。この場合は、幾つかの MB は、標準 R-D 最適化ではインターコーディングされた MB であることが要求されるが、イントラコーディングされた MB としてコーディングされる。基本層に含まれているこれらのイントラコーディングされた MB は、QP<sub>b</sub> 又は QP<sub>e</sub> のいずれかによって符号化することができる。QP<sub>e</sub> が基本層に関して用いられる場合は、QP<sub>b</sub> 又は QP<sub>e</sub> のいずれかで符号化することができる。QP<sub>e</sub> が基本層に関して用いられる場合は、拡張層においてはリファインメントは不要である。QP<sub>b</sub> が基本層に関して用いられる場合は、リファインメントが必要な場合がある。その他の場合は、拡張層において品質低下に気づくことが可能になる。インターコーディングは、コーディング効率という意味ではイントラコーディングよりも効率的であるため、拡張層におけるこれらのリファインメントは、インターコーディングされる。この方法により、基本層係数は、拡張層に関して用いられない。従って、品質は、拡張層において新たな動作を導入せずに向上される。

40

【0170】

拡張層では、提供する高い圧縮品質に起因して B フレームが一般的に用いられる。しかしながら、B フレームは、P フレームのイントラコーディングされた MB を基準にしなければならないことがある。B フレームの画素が拡張層品質で符号化される場合は、上述されるように、P フレームのイントラコーディングされた MB の品質がより低いことに起因してあまりにも多いビットを要求する可能性がある。上述されるように、HVS の品質を

50

利用することによって、Pフレームのより低い品質のイントラコーディングされたMBを基準にするときにはBフレームMBをより低い品質で符号化することができる。

【0171】

P又はBフレーム内のイントラコーディングされたMBの1つの極端な事例は、符号化中の映像内におけるシーンの変化の存在に起因してP又はBフレーム内の全MBがイントラモードで符号化されることである。この場合は、フレーム全体を基本層品質でコーディングすることができ、拡張層におけるリファインメントは行われず、Bフレームにおいてシーンの変化が生じた場合で、Bフレームは拡張層でしか符号化されないと仮定した場合は、Bフレームは、基本層品質で符号化するか又は単にドロップすることが可能である。Pフレームにおいてシーンの変化が生じた場合は、変更は必要でなく、Pフレームをドロップするか又は基本層品質で符号化することができる。“SCALABLE VIDEO CODING WITH TWO LAYER ENCODING AND SINGLE LAYER DECODING”(2層符号化及び単層復号を有するスケーラブルな映像コーディング)という題名を有し、本特許出願の譲受人によって所有され、本明細書において参照されることによってその全体が本明細書に組み入れられている、同時係属米国特許出願番号[Attorney docket/ref. no. 050078]において、スケーラブル層符号化がさらに説明されている。

10

【0172】

符号器第1パス部分

図7は、図2の符号器228の説明例を示す。示されているブロックは、符号器228内に含めることができる様々な符号器処理を示す。この例においては、符号器228は、区分線704上方の第1パス部分706と、線704下方の第2パス部分706(図2の第2パス符号器232及び再符号器234の機能を含む)と、を含む。

20

【0173】

符号器228は、メタデータ及び生映像をプリプロセッサ226から受信する。メタデータは、プリプロセッサ226によって受信された又は計算されたメタデータを含むことができ、映像のコンテンツ情報に関連するメタデータを含む。符号器228の第1パス部分702は、機能に関して後述される第1パス符号化702に含めることができる典型的プロセスを示す。当業者が知ることになるように、該機能は様々な形態で(例えば、ハードウェア、ソフトウェア、ファームウェア、又はその組合せ内に)実装することができる。

30

【0174】

図7は、適応型イントラリフレッシュ(AIR)モジュールを示す。AIRモジュール710は、メタデータに基づいてIフレームをインスタンス化するIフレームインスタンス化モジュール708に入力を提供する。第1パス部分702は、メタデータ及び映像を受信し、映像に関連するコンテンツ情報を決定するように構成されたコンテンツ分類モジュール712を含むことも可能である。コンテンツ情報は、メタデータ及び映像も受信するレート制御ビット割り当てモジュール714に提供することができる。制御ビット割り当てモジュール714は、レートビット制御情報を決定し、該情報をモード決定モジュール715に提供する。コンテンツ情報及び映像は、イントラモデル(歪み)モジュール716に提供することができ、イントラモデル(歪み)モジュール716は、イントラコーディング歪み情報をモード決定モジュール715及び基本層及び拡張層に関するスケーラビリティレート-歪みモジュール718に提供する。基本層及び拡張層に関するスケーラビリティレート-歪みモジュール718にインターコーディング歪み情報を提供する動き推定(歪み)モジュール720に映像及びメタデータが提供される。基本層及び拡張層に関するスケーラビリティレート-歪みモジュール718は、モード決定モジュール715に提供される動き推定モジュール720及びイントラモデル歪みモジュール716からの歪み推定を用いてスケーラビリティレート-歪み情報を決定する。モード決定モジュール715は、スライス/MB順序設定(ordering)モジュール722からの入力も受け取る。スライス/MB順序設定モジュール722は、(第2パス部分706に示される)誤り耐性モジュール740から入力を受け取り、映像の独立して符号化可能な部分(

40

50

スライス)を誤り耐性に関するアクセスユニット境界とアライメントさせることに関する情報をモード決定モジュール715に提供する。モード決定モジュール715は、入力に基づいて符号化モード情報を決定し、“最良の”コーディングモードを第2パス部分706に提供する。該第1パス部分702符号化の幾つかの例に関するさらなる例示的説明が以下に示される。

**【0175】**

上述されるように、コンテンツ分類モジュール712は、プリプロセッサ226によって供給されたメタデータ及び生映像を受信する。幾つかの例においては、プリプロセッサ226は、マルチメディアデータからコンテンツ情報を計算し、該コンテンツ情報を(例えば、メタデータ内の)コンテンツ分類モジュール712に提供し、コンテンツ分類モジュール712は、コンテンツ情報を用いてマルチメディアデータに関するコンテンツ分類を決定することができる。幾つかのその他の側面においては、コンテンツ分類モジュール712は、マルチメディアデータから様々なコンテンツ情報を決定するように構成され、さらにコンテンツ情報を決定するように構成することもできる。

**【0176】**

コンテンツ分類モジュール712は、異なる型のコンテンツを有する映像に関する異なるコンテンツ分類を決定するように構成することができる。異なるコンテンツ分類は、マルチメディアデータを符号化する側面、例えば、量子化パラメータを決定するためのビットレート(例えば、ビット割り当て)を決定する、動き推定、スケラビリティ、誤り耐性を決定する、チャンネル間での最適なマルチメディアデータ品質を維持する、及び高速チャンネル切り換え方式(例えば、Iフレームが高速チャンネル切り換えを許容するように定期的に強制する)に関して用いられる異なるパラメータを得ることが可能になる。一例により、符号器228は、コンテンツ分類に基づいてレート-歪み(R-D)最適化及びビットレート割り当てを決定するように構成される。コンテンツ分類を決定することは、コンテンツ分類に基づいて希望されるビットレートに対応する所定の品質レベルにマルチメディアデータを圧縮することを可能にする。さらに、マルチメディアデータのコンテンツを分類する(例えば、人間の視覚系に基づいてコンテンツ分類を決定する)ことによって、その結果受信デバイスのディスプレイ上において得られる通信されたマルチメディアデータの知覚品質(perceptive quality)は、映像コンテンツに依存して作られる。

**【0177】**

コンテンツ分類モジュール712がコンテンツを分類するために行う手順の一例として、図9は、コンテンツ分類モジュール712が動作する際に従うことができる典型的プロセスを例示するプロセス900を示す。示されるように、プロセス900は、入力ブロック902において開始し、コンテンツ分類モジュール712は、生のマルチメディアデータ及びメタデータを受信する。次に、プロセス900は、ブロック904に進み、コンテンツ分類モジュール712は、マルチメディアデータの空間情報及び時間情報を決定する。幾つかの側面においては、空間情報及び時間情報は、空間的マスキング及び時間的マスキング(例えば、フィルタリング)によって決定される。空間情報及び時間情報は、シーン変化データ及び動きベクトル(MV)平滑化を含むメタデータに基づいて決定することができる。次に、プロセス900は、ブロック912に進み、空間的複雑さ、時間的複雑さ、及び感度の推定を行う。次に、プロセス900はブロック916に進み、ブロック904及び912において、決定された空間データ、時間データ及び感度データの結果に基づいてマルチメディアデータのコンテンツが分類される。さらにブロック916において、特定のレート-歪み(R-D)曲線を選択することができ及び/又はR-D曲線データを更新することができる。次に、プロセス900は、出力ブロック918に進み、出力は、複雑さ-歪みマップ又は空間的活動及び時間的活動(例えば、コンテンツ分類)を示す値及び/又は選択されたR-D曲線を含むことができる。再度図7に関して、コンテンツ分類モジュール712は、上述される、レート制御ビット割り当てモジュール714、イントラモデル(歪み)モジュール716、及びIフレームインスタンス化モジュール70

10

20

30

40

50

8 に出力を提供する。

【0178】

コンテンツ情報

コンテンツ分類モジュール712は、以下においてさらに説明される、空間的複雑さ、時間的複雑さ、コントラスト比値、標準偏差、及びフレーム差分メトリックスを含む様々なコンテンツ関連メトリックスを含む様々なコンテンツ情報をマルチメディアデータから計算するように構成することができる。

【0179】

コンテンツ分類モジュール712は、マルチメディアデータの空間的複雑さ及び時間的複雑さを決定するように、及びテクスチャ値を空間的複雑さと、動き値を時間的複雑さと関連づけるように構成することができる。コンテンツ分類モジュール712は、符号化中のマルチメディアデータのコンテンツに関連する前処理されたコンテンツ情報をプリプロセッサ226から受信し、又は代替として、プリプロセッサ226は、コンテンツ情報を計算するように構成することができる。上述されるように、コンテンツ情報は、例えば、1つ以上の $D_{c a s t}$ 値と、コントラスト比値と、動きベクトル(MV)と、絶対値差分和(SAD)と、を含むことができる。

【0180】

一般的には、マルチメディアデータは、1つ以上の画像シーケンス、又はフレームシーケンスを含む。各フレームは、処理のために画素ブロックに分割することができる。空間的複雑さは、フレーム内における空間的詳細レベルの評価基準について一般的に説明する広義の用語である。主に無地(plain)の又は変化しない又は変化が小さい輝度及び色差エリアを有するシーンは、小さい空間的複雑さを有することになる。空間的複雑さは、映像データのテクスチャと関連づけられる。空間的複雑さは、この側面においては、 $D_{c s a t}$ と呼ばれる人間の視覚感度メトリックスに基づき、前記人間の視覚感度メトリックスは、局所空間周波数及び周囲照明の関数として各ブロックに関して計算される。当業者は、ビジュアル画像の空間周波数パターン及び照明とコントラストに関する特徴を用いて人間の視覚系を利用するための技術を認識している。人間の視覚系の遠近法上の制限を利用するための幾つかの感度メトリックスが知られており、本明細書において説明される方法とともに用いることが可能である。

【0181】

時間的複雑さは、フレームシーケンス内のフレーム間で基準とされるマルチメディアデータにおける動きレベルの評価基準を一般的に説明するために用いられる広義の用語である。動きがほとんど又はまったくないシーン(例えば、映像データのフレームシーケンス)は、時間的複雑さが小さい。時間的複雑さは、各マクロブロックに関して計算することができ、 $D_{c s a t}$ 値、動きベクトル及び1つのフレームと他のフレーム(例えば、基準フレーム)との間の絶対画素差分和に基づくことができる。

【0182】

フレーム差分メトリックスは、動き量(例えば、動きベクトル又はMV)及び予測子と現在のマクロブロックとの間の絶対値差分和(SAD)として表される残留エネルギーを考慮した2つの連続フレーム間の差分の評価基準となる。フレーム差分は、双方向又は単一方向の予測効率の評価基準も提供する。

【0183】

動きが補償されたデインターレーシングを行うことが可能なプリプロセッサから受信された動き情報に基づいたフレーム差分メトリックスの一例は以下のとおりである。デインターレーサは、双方向動き推定を行い、従って、双方向動きベクトル及びSAD情報を利用可能である。各マクロブロックに関してSAD\_MVによって表されるフレーム差分は、以下のように導き出すことができる。

10

20

30

40

【数 3 3】

$$\text{SAD\_MV} = \log_{10} [\text{SAD} * \exp(-\min(1, \text{MV}))] \quad [29]$$

【0 1 8 4】

ここで、

【数 3 4】

$$\text{MV} = \text{Square\_root}(\text{MV}_x^2 + \text{MV}_y^2), \text{ SAD} = \min(\text{SAD}_N, \text{SAD}_P)$$

10

【0 1 8 5】

であり、 $\text{SAD}_N$  は、逆方向基準フレームから計算された  $\text{SAD}$  であり、 $\text{SAD}_P$  は、順方向基準フレームから計算された  $\text{SAD}$  である。

【0 1 8 6】

フレーム差分を推定する他の手法が方程式 6 乃至 8 を参照しながら上述されている。 $\text{SAD}$  比（又はコントラスト比）は、方程式 6 において上述のように計算することができる。すべてのフレームの輝度ヒストグラムも決定することができ、ヒストグラム差分は、方程式 7 を用いて計算される。フレーム差分メトリックス  $D$  は、方程式 8 に示されるように計算することができる。

20

【0 1 8 7】

一例においては、コントラスト比及びフレーム差分メトリックスは、映像コンテンツ分類を入手するために次の方法で利用され、所定の映像シーケンス内の特徴を信頼できる形で予測可能である。本明細書においては符号器 2 2 8 において生じる状況が説明されているが、プリプロセッサ 2 2 6 は、コンテンツ分類（又はその他のコンテンツ情報）を決定し、メタデータを介してコンテンツ分類を符号器 2 2 8 に渡すように構成することもできる。以下の例において説明されるプロセスは、 $R-D$  曲線に基づく解析から得られた分類と同様に、コンテンツを 8 つの可能なクラスに分類する。分類プロセスは、シーンの複雑さ及び各スーパーフレーム内でのシーン変化発生数に依存して各スーパーフレームに関して 0 乃至 1 の範囲内の値を出力する。プリプロセッサ内のコンテンツ分類モジュールは、フレームコントラスト及びフレーム差分値からコンテンツ分類メトリックスを入手するために各スーパーフレームに関して以下のステップ（1）乃至（5）を実行することができる。

30

【0 1 8 8】

1．平均フレームコントラスト及びフレームコントラスト偏差をマクロブロックコントラスト値から計算する。

【0 1 8 9】

2．シミュレーションから得られた値を用いてフレームコントラスト及びフレーム差分値、それぞれ 4 0 及び 5、を正規化する。

40

【0 1 9 0】

3．例えば一般化された方程式を用いてコンテンツ分類メトリックスを計算する。

【数 3 5】

$$\begin{aligned} \text{CCMetric} = & \text{CCW1} * \text{I\_Frame\_Contrast\_Mean} + \text{CCW2} \\ & * \text{Frame\_Difference\_Mean} - \text{CCW3} * \text{I\_Contrast\_Deviation}^2 * \exp \\ & (\text{CCW4} * \text{Frame\_Difference\_Deviation}^2) \end{aligned} \quad [30]$$

【0191】

ここで、CCW1、CCW2、CCW3及びCCW4は、重み付け係数である。この例においては、値は、CCW1に関しては0.2、CCW2に関しては0.9、CCW3に関しては0.1、CCW4に関しては-0.00009が選択される。

10

【0192】

4. スーパーフレーム内でのシーン変化数を決定する。一般的には、スーパーフレームは、特定の期間において表示することができるピクチャ又はフレームのグループを意味する。典型的には、期間は1秒である。幾つかの側面においては、スーパーフレームは、(30/fps映像の場合は)30フレームを具備する。その他の側面においては、スーパーフレームは、24フレーム(24/fps映像)を具備する。シーン変化数に依存して、以下の事例のうちの1つが実行される。

20

【0193】

(a) シーン変化なし：スーパーフレーム内においてシーン変化がないときには、メトリックスは、以下の方程式において示されるようにフレーム差分値のみに完全に依存する。

【数 3 6】

$$\begin{aligned} \text{CCMetric} = & (\text{CCW2} + (\text{CCW1}/2)) * \text{Frame\_Difference\_Mean} - (\text{CCW3} - \\ & (\text{CCW1}/2)) * 1 * \exp(-\text{CCW4} * \text{Frame\_Difference\_Deviation}^2) \end{aligned} \quad [31]$$

30

【0194】

(b) 単一のシーン変化：スーパーフレーム内において単一のシーン変化フレームが観察されたときには、以下のようにデフォルト方程式を用いてメトリックスが計算される。

【数 3 7】

$$\begin{aligned} \text{CCMetric} = & \text{CCW1} * \text{I\_Frame\_Contrast\_Mean} + \text{CCW2} \\ & * \text{Frame\_Difference\_Mean} - \text{CCW3} * \text{I\_Contrast\_Deviation}^2 * \exp \\ & (\text{CCW4} * \text{Frame\_Difference\_Deviation}^2) \end{aligned} \quad [32]$$

40

【0195】

(c) 2つのシーン変化：所定のスーパーフレーム内において多くて2つのシーン変化が存在することが観察されたときには、以下の方程式において示されるように、最初のスーパーフレームは後のスーパーフレームによって素早くリフレッシュされるため、最後のスーパーフレームは、最初のスーパーフレームよりも大きい重みが与えられる。

【数 3 8】

$$\begin{aligned} \text{CCMetric} = & 0.1 * \text{I\_Frame\_Contrast\_Mean1} + \\ & \text{CCW1} * \text{I\_Frame\_Contrast\_Mean2} + (\text{CCW2} - 0.1) * \text{Frame\_Difference\_Mean} - \\ & \text{CCW3} * \text{I\_Contrast\_Deviation1}^2 * \text{I\_Contrast\_Deviation2}^2 * \exp \\ & (\text{CCW4} * \text{Frame\_Difference\_Deviation}^2) \end{aligned} \quad [33]$$

【0196】

(d) 3つ以上のシーン変化：所定のスーパーフレームが3つ以上のIフレーム（例えばN）を有することが観察される場合は、以下の方程式において示されるように、最後のIフレームはより大きい重みが与えられ、その他のすべてのIフレームは0.05の重みが与えられる。

10

【数 3 9】

$$\begin{aligned} \text{CCMetric} = & 0.05 * \text{I\_Frame\_Contrast\_Mean}_{(1...N-1)} + \text{CCW1} * \\ & \text{I\_Frame\_Contrast\_Mean}_{(N)} + (\text{CCW2} - (0.05 * (N-1))) * \text{Frame\_Difference\_Mean} - \\ & \text{CCW3} * \text{I\_Contrast\_Deviation}_{(N)}^2 * \text{I\_Contrast\_Deviation}_{(1...N-1)}^2 \\ & * \exp(\text{CCW4} * \text{Frame\_Difference\_Deviation}^2) \end{aligned} \quad [34]$$

20

【0197】

5. フレーム差分平均が0.05よりも小さいときに動きが少ないシーンの場合はメトリックスに関して補正を用いることができる。(CCOFFSET)0.33のオフセットがCCMetricに加えられる。

【0198】

コンテンツ分類モジュール712は、D<sub>c s a t</sub>値、動きベクトル及び/又は絶対値差分和を用いて、マクロブロックに関する空間的複雑さ（又は指定された映像データ量）を決定する。時間的複雑さは、フレーム差分メトリックスの評価基準（動きベクトルを有する動き量、及びフレーム間の絶対値差分和を考慮した2つの連続フレーム間の差分）によって決定される。

30

【0199】

幾つかの側面においては、コンテンツ分類モジュール712は、帯域幅マップを生成するように構成することができる。例えば、帯域幅マップ生成は、プリプロセッサ226が帯域幅マップを生成しない場合にコンテンツ分類モジュール712によって行うことができる。

【0200】

テクスチャ値及び動き値の決定

40

マルチメディアデータ内の各マクロブロックに関して、コンテンツ分類モジュール712は、テクスチャ値を空間的複雑さと、動き値を時間的複雑さと関連づける。テクスチャ値は、マルチメディアデータの輝度値と関連し、低いテクスチャ値は、データの近隣画素の輝度値の変化が小さいことを示し、高いテクスチャ値は、データの近隣画素の輝度値の変化が大きいことを示す。

【0201】

テクスチャ値及び動き値が計算された時点で、コンテンツ分類モジュール712は、動き情報及びテクスチャ情報の両方を考慮することによってコンテンツ分類を決定する。コンテンツ分類モジュール712は、分類中である映像データに関するテクスチャを、マクロブロックの輝度値の複雑さを一般的に示す相対的テクスチャ値、例えば、“低”テクスチ

50

ヤ、“中”テクスチャ、又は“高”テクスチャ、と関連づける。さらに、コンテンツ分類モジュール712は、分類中の映像データに関して計算された動き値を、マクロブロックの動き量を一般的に示す相対的動き値、例えば、“少ない”動き、“中度の”動き、又は“多い”動き、と関連づける。代替側面においては、動き及びテクスチャに関してより少ない又はより多いカテゴリを用いることができる。次に、関連づけられたテクスチャ値及び動き値を考慮することによってコンテンツ分類メトリックスが決定される。

#### 【0202】

図8は、テクスチャ値及び動き値がコンテンツ分類とどのように関連づけられているかを示す分類チャート例を示す。当業者は、該分類チャートを例えばルックアップテーブル又はデータベース内に実装するための数多く方法に精通する。分類チャートは、映像データコンテンツに関する予め決められた評価に基づいて生成される。映像データ分類を決定するために、(“x軸”上の)“低”、“中”、又は“高”のテクスチャ値が、(“y軸”上の)“少ない”、“中度の”、又は“多い”の動き値と相互参照される。交差ブロックにおいて示されるコンテンツ分類が映像データに割り当てられる。例えば、“高”のテクスチャ値及び“中度の”動き値は、7の分類になる。図8は、この例においては、8つの異なるコンテンツ分類と関連づけられている相対的なテクスチャ値及び動き値の様々な組合せを示す。幾つかのその他の側面においては、より多い又はより少ない分類を用いることができる。“CONTENT CLASSIFICATION FOR MULTIMEDIA PROCESSING”(マルチメディアデータ処理に関するコンテンツ分類)という題名を有し、本特許出願の譲受人に譲渡され、本明細書において参照されることによって明示で本明細書に組み入れられている、同時係属米国特許出願番号11/373,577(出願日:2006年3月10日)において、コンテンツ分類の例示的側面のさらなる説明が開示されている。

#### 【0203】

レート制御ビット割り当て

本明細書において説明されるように、マルチメディアデータコンテンツ分類は、映像の知覚品質を一定に維持しながらビット管理を有効に向上させるために符号化アルゴリズムにおいて用いることができる。例えば、分類メトリックスは、シーン変化検出、ビットレート割り当て制御、及びフレームレートアップコンバージョン(FRUC)に関するアルゴリズムにおいて用いることができる。映像データ通信においては圧縮器/伸長器(コーデック)システム及びデジタル信号処理アルゴリズムが共通して用いられ、帯域幅を保全するように構成することができるが、品質と帯域幅保全との間には相反する関係がある。最良のコーデックは、映像の品質低下を最小にしながら最大の帯域幅保全を提供する。

#### 【0204】

一説明例においては、レート制御ビット割り当てモジュール714は、コンテンツ分類を用いてビットレート(例えば、マルチメディアデータを符号化するために割り当てられるビット数)を決定し、符号器228のその他のプロセス及び構成要素によって用いるためにビットレートをメモリに格納する。映像データの分類から決定されたビットレートは、一貫した品質レベルでマルチメディアデータを提供する一方で帯域幅の保全を援助することができる。一側面においては、異なるビットレートを8つの異なるコンテンツ分類の各々と関連づけることができ、そのビットレートがマルチメディアデータを符号化するために用いられる。その結果得られる効果は、マルチメディアデータの異なるコンテンツ分類に異なるビット数が符号化のために割り当てられるが、知覚品質は、ディスプレイ上で見たときには類似している又は一貫していることである。

#### 【0205】

一般的には、より高いコンテンツ分類を有するマルチメディアデータは、より多いレベルの動き及び/又はより高いレベルのテクスチャを示しており、符号化時により多くのビットが割り当てられる。より低い分類(テクスチャがより低く動きがより小さいことを示す)を有するマルチメディアデータは、より少ないビットが割り当てられる。特定のコンテンツ分類のマルチメディアデータに関しては、ビットレートは、マルチメディアデータを見るための選択された目標知覚品質レベルに基づいて決定することができる。マルチメ

10

20

30

40

50

ディアデータ品質を決定することは、人間がマルチメディアデータを見て等級を付けることによって決定することができる。幾つかの代替側面においては、マルチメディアデータ品質の推定は、例えば信号・雑音比アルゴリズムを用いた自動試験システムによって行うことができる。一側面においては、各特定の品質レベルを達成するために必要な一組の標準品質レベル（例えば、5つ）及び対応するビットレートが、各コンテンツ分類のマルチメディアデータに関して予め決定される。一組の品質レベルを決定するために、特定のコンテンツ分類のマルチメディアデータは、マルチメディアデータが一定のビットレートをを用いて符号化されるときにマルチメディアデータの目で知覚された品質の数値による表示を提供する平均意見点数（MOS）を生成することによって評価することができる。MOSは、1乃至5の範囲内の単一の数として表すことができ、1は、最低の知覚品質であり、5は、最高の知覚品質である。その他の側面においては、MOSは、5つよりも多い又は少ない品質レベルを有することができ、各品質レベルに関して異なる記述を用いることができる。

10

#### 【0206】

マルチメディアデータ品質を決定することは、人間がマルチメディアデータを見て等級を付けることによって決定することができる。幾つかの代替側面においては、マルチメディアデータ品質の推定は、例えば信号・雑音比アルゴリズムを用いた自動試験システムによって行うことができる。一側面においては、各特定の品質レベルを達成するために必要な一組の標準品質レベル（例えば、5つ）及び対応するビットレートが、各コンテンツ分類のマルチメディアデータに関して予め決定される。

20

#### 【0207】

一定のコンテンツ分類のマルチメディアデータに関する目で知覚された品質レベルとビットレートとの間の関係を知ることは、目標の（例えば、希望される）品質レベルを選択することによって決定することができる。ビットレートを決定するために用いられる目標品質レベルは、予め選択する、ユーザーが選択する、ユーザーからの又は他のプロセスからの入力を要する自動プロセス又は半自動プロセスを通じて選択する、又は予め決められた判定基準に基づいて符号化デバイス又はシステムによって動的に選択することができる。目標品質レベルは、例えば、符号化アプリケーションの型、又はマルチメディアデータを受信するクライアントデバイスの型に基づいて選択することができる。

#### 【0208】

図7の説明例においては、レート制御ビット割り当てモジュール714は、コンテンツ分類モジュール712から両方のデータを受信し、プリプロセッサ226から直接メタデータを受信する。レート制御ビット割り当てモジュール714は、符号器228の第1パス部分に常駐し、レート制御微調整738は、第2パス部分706に常駐する。この2パスレート制御側面は、第1パス（レート制御ビット割り当てモジュール714）が（例えば、256 kbps）の長期平均ビットレートを目標とする）1つのスーパーフレームルックアヘッドによるコンテキスト適応型ビット割り当てを行い、ピークレートを制限し、第2パス（レート制御微調整モジュール738）は、2層スケラビリティに関して第1パスの結果をリファインメントし、レート適応化を行う。レート制御は、4つのレベル、すなわち、（1）GOPレベル - I、P、B、及びFフレームのビット分布がGOP内部において一様にならないように制御する、（2）スーパーフレームレベル - 最大のスーパーフレームサイズに関する厳格な限度を制御する、（3）フレームレベル - コンテンツ情報（例えば、コンテンツ分類）に基づく、マルチメディアデータフレームの空間的及び時間的複雑さに従ってビット要求を制御する、及び（4）マクロブロックレベル - コンテンツ情報（例えば、コンテンツ分類）に基づく、空間的及び時間的複雑さマップに基づいてマクロブロックのビット割り当てを制御する、に対して作用する。

30

40

#### 【0209】

レート制御モジュール714の動作の典型的流れ図が図10に示される。図10に示されるように、プロセス1000は、入力ブロック1002において開始する。レート制御モジュール714は、様々な入力を受信し、図7にはこれらの様々な入力がすべて示され

50

ているわけではない。例えば、入力情報は、プリプロセッサ 226 からのメタデータと、目標ビットレートと、符号器バッファサイズ（又は、同等として、レート制御に関する最大遅延時間）と、初期レート制御遅延と、フレームレート情報と、を含むことができる。さらなる入力情報は、ピクチャグループ（GOP）レベルにおける入力を含むことができ、例えば、最大スーパーフレームサイズ、GOP の長さ及び P/B フレーム分布（シーン変化情報を含む）、希望される基本層及び拡張層の手配、将来の 30 フレームに関する GOP 内のピクチャに関する複雑さ - 歪みメトリックスを含む。その他の入力情報は、ピクチャレベルにおける入力を含み、（コンテンツ分類モジュール 712 から受信された）現在のピクチャに関する複雑さ - 歪みマップと、量子化パラメータ（QP）と、（スライディングウィンドーにわたって当てはめられた）過去の 30 フレームのビット細分化と、を含む。最後に、マクロブロック（MB）レベルにおける入力情報は、例えば、基準ピクチャ内の共配置されたマクロブロック（MB）の平均絶対値差分（MAD）と、（スキップされたかどうかにかかわらず）量子化後のマクロブロックのコーディングされたブロックパターン（CBP）と、を含む。

#### 【0210】

ブロック 1002 における入力後、プロセス 1000 は、ビットストリームを符号化するための初期設定に関するブロック 1004 に進む。同時並行して、バッファ初期設定 1006 が行われる。次に、ブロック 1008 において示されるように GOP が初期設定され、GOP ビット割り当て 1010 が初期設定の一部として受信される。GOP 初期設定後は、流れはブロック 1012 に進み、スライスが初期設定される。この初期設定は、ブロック 1014 によって示されるようにヘッダービットの更新を含む。ブロック 1004、1008 及び 1012 の初期設定が行われた後は、ブロック 1016 によって示されるように基本単位又はマクロブロック（MB）に関するレート制御（RC）が行われる。ブロック 1016 におけるマクロブロックのレート制御決定の一環として、符号器 228 内のインタフェースを介して入力が受信される。これらの入力は、マクロブロック（MB）ビット割り当て 1018 と、二次モデルパラメータの更新 1020 と、メジアンからのメジアン絶対偏差（“MAD”、頑健な分散推定値）パラメータの更新 1022 と、を含むことができる。次に、プロセス 1000 は、1つのピクチャ 1024 を符号化後の動作の実行のためにブロック 1024 に進む。この手順は、ブロック 1026 によって示されるようにバッファパラメータの更新を受信することを含む。次に、プロセス 1000 は、出力ブロック 1028 に進み、レート制御モジュール 714 は、図 7 に示されるようにモード決定モジュール 715 によって用いられる各マクロブロック MB に関する量子化パラメータ QP を出力する。

#### 【0211】

##### 動き推定

動き推定モジュール 720 は、メタデータ及び生映像の入力をプリプロセッサ 226 から受信し、ブロックサイズ、動きベクトル歪みメトリックス、及び基準フレーム識別子を含むことができる出力をモード決定モジュール 715 に提供する。図 11 は、動き推定モジュール 720 の典型的動作を示す。示されるように、プロセス 1100 は、入力 1102 から開始する。フレームレベルにおいて、モジュール 720 は、基準フレーム ID 及び動きベクトルの入力を受信する。マクロブロックレベルにおいて、入力 1102 は、入力画素及び基準フレーム画素を含む。プロセス 1100 は、ステップ 1104 に続き、色動き推定（ME）及び動きベクトル予測が行われる。このプロセスを実行するために、MPEG-2 動きベクトル、ルマ動きベクトル MV 1106、動きベクトル平滑化 1108、及び非因果動きベクトル 1110 を含む様々な入力が受信される。次に、プロセス 1100 は、ブロック 1112 に進み、動きベクトル探索アルゴリズム又は方法、例えばヘキサゴナル又はダイヤモンド探索方法、が実行される。ブロック 1112 におけるプロセスへの入力は、ブロック 1114 によって示されるように絶対値差分和（SAD）、二乗差分和（SSD）、及び/又はその他のメトリックスを含むことができる。動きベクトル探索が行われた時点で、プロセス 1100 は、終了ブロック 1116 に進み、終了処理が行

われる。次に、プロセス100は、出力ブロック1118において終了し、ブロックサイズ、動きベクトル(MV)、歪みメトリックス、及び基準フレーム識別子の出力を生み出す。

#### 【0212】

基本層及び拡張層に関するスケーラビリティR-D

図13は、スケーラビリティR-Dモジュール718によって実行することができるスケーラビリティプロセス1300の典型的流れ図を示す。プロセス1300は、開始ブロック1302において開始し、ブロック1304に進み、スケーラビリティR-Dモジュール718は、動き推定モジュール720から入力を受信し、動き推定を行う。動き推定は、ブロック1306によって示されるように基本層基準フレーム、拡張層基準フレーム、及びコーディング対象となる原フレームの入力に依存する。該情報は、GOPパーティション612によって計算し、例えばメタデータを介してスケーラビリティR-Dモジュール718に通信することができる。プロセス1300は、ブロック1308に進み、データ基本層及び拡張層データのスケラビリティ情報を決定する。次に、ブロック1310において示されるように基本層符号化が行われ、その後ブロック1312における拡張層符号化が続く。拡張層の符号化は、ブロック1314によって示されるように、層間予測に関する基本層コーディング結果を入力して用いることができ、従って、基本層符号化後に一時的に実行される。この件は、“SCALABLE VIDEO CODING WITH TWO LAYER ENCODING AND SINGLE LAYER DECODING”(2層符号化及び単層復号を有するスケーラブルな映像コーディング)という題名を有する同時係属米国特許出願番号[Attorney docket/ref. no. 050078]においてさらに説明されている。符号化が完了後は、プロセス1300は、ブロック1316において終了する。

#### 【0213】

スライス/マクロブロック順序設定

第1パス部分702は、スライス/マクロブロック順序設定モジュール722も含み、スライス/マクロブロック順序設定モジュール722は、第2パス部分において誤り耐性モジュール740から入力を受信し、スライスアライメント情報をモード決定モジュール715に提供する。スライスは、独立して復号可能な(エントロピー復号)コーディングされた映像データの塊である。アクセスユニット(AU)は、正確に1つの主のコーディングされたピクチャを常に含む一組のNALユニットを各々具備するコーディングされた映像フレームである。主のコーディングされたピクチャに加えて、アクセスユニットは、1つ以上の冗長なコーディングされたピクチャ又コーディングされたピクチャのスライス又はスライスデータパーティションを含まないその他のNALユニットを内蔵することもできる。アクセスユニットの復号の結果、復号されたピクチャが常に得られる。

#### 【0214】

フレームは、最高の時間ダイバーシティを提供する物理層パケットの時分割多重化ブロック(TDMカプセルと呼ばれる)であることができる。スーパーフレームは、1つの時間単位(例えば、1sec)に対応し、4つのフレームを含む。時間領域においてスライス及びAU境界をフレーム境界にアライメントさせることは、壊れたデータの最も効率的な分離及び局在化が得られる。深いフェード中には、TDMカプセル内の隣接データのほとんどが誤りによる影響を受ける。時間的ダイバーシティに起因して、残りのTDMカプセルは、無傷の状態になる確率が非常に高い。壊れていないデータは、影響を受けたTDMカプセルから失われたデータを復元する及び隠すために利用することができる。データシンボルが変調する周波数副搬送波における分離を通じて周波数ダイバーシティが達成される周波数領域多重化(FDM)に対しても同様の論理が適用される。さらに、無線ネットワークにおいてしばしば適用される(送信機アンテナ及び受信機アンテナにおける分離を通じての)空間ダイバーシティ及びその他の形態のダイバーシティに対しても同様の論理が適用される。

#### 【0215】

スライス及びAUをフレームにアライメントさせるために、アウターコード(FEC)

コードブロック生成及びM A C層カプセル化も同様にアライメントすべきである。図20は、スライス及びA Uにおけるコーディングされた映像データ又は映像ビットストリームの構成を示す。コーディングされた映像は、階層化された映像コーディングが適用される1つ以上のビットストリーム、例えば、基本層ビットストリーム及び拡張層ビットストリーム、において構築することができる。

#### 【0216】

映像ビットストリームは、図20においてフレームI' 2005、フレーム3' 2010及びフレームM' 2015によって示されるA Uを具備する。A Uは、スライス1' 2020、スライス2' 2025、及びスライスN' 2030によって示されるように、データスライスを具備する。スライスの各始まりは、開始符号によって識別され、ネットワーク適応化に提供される。一般的には、Iフレーム又はイントラコーディングされたA Uは大きく、Pフレーム又は順方向の予測されたフレームが後続し、Bフレームが後続する。A Uを複数のスライスにコーディングすることは、スライス間での空間予測が制限されていること及びスライスヘッダーもオーバーヘッドの一因であることに起因して、コーディングされたビットレートの点で有意なオーバーヘッドコストを発生させる。スライス境界は再同期化地点であり、P L Pが破壊されているときはP L P内のスライスに誤りが限定され、他方、P L Pが複数のスライス又は複数のスライスの一部を含む場合は、誤りはP L P内の全スライス又はスライスの一部に対して影響を及ぼすことになるため、隣接する物理層パケットをスライスに制限することは誤りを制御する。

#### 【0217】

Iフレームは典型的に大きく、例えば数十キロビット程度であるため、複数のスライスに起因するオーバーヘッドは、Iフレームのサイズ全体又は総ビットレートの大きな割合ではない。さらに、イントラコーディングされたA U内により多くのスライスを有することは、より良い及びより頻繁な再同期化及びより効率的な空間的誤りの隠蔽を可能にする。さらに、P及びBフレームは、Iフレームから予測されるため、Iフレームは、映像ビットストリームにおいて最も重要な情報を搬送する。Iフレームは、チャンネル取得に関するランダムアクセスポイントとしての働きもする。

#### 【0218】

次に図21に関して、Iフレームをフレーム境界に慎重にアライメントさせ、I A Uを有するスライスを同じくフレーム境界に慎重にアライメントさせることは、最も効率的な誤り制御、誤り保護を可能にする(フレーム1' 2105に属する1つのスライスが失われた場合、フレーム2' 2110はフレーム1' 2105から有意な時間だけ分離されているので無傷の状態である確率が高いため、再同期化及び誤り隠蔽を通じて誤りからの回復を行うことができる)。

#### 【0219】

Pフレームは、典型的には、数キロビット程度の大きさが設定されているため、Pフレームのスライス及び整数のPフレーム数をフレーム境界にアライメントさせることは、(Iフレームに関する理由と同様の理由で)悪影響を及ぼすような効率損失のない誤り耐性を可能にする。該側面においては時間的誤り隠蔽を採用することができる。代替として、連続するPフレームが異なるフレームで到着するように分散させることは、Pフレーム間における追加の時間的ダイバーシティを提供し、このことは、時間的隠蔽が以前に再構築されたI又はPフレームからの動きベクトル及びデータに基づくことに起因して可能になる。Bフレームは、極端に小さい(数百ビット)から中程度(数千ビット)であることができる。従って、悪影響を及ぼす効率損失なしで誤り耐性を達成させるためには整数のBフレームをフレーム境界にアライメントさせることが望ましい。

#### 【0220】

##### モード決定モジュール

図12は、モード決定モジュール715の幾つかの動作例を示す。示されるように、プロセス1200は、入力ブロック1202において開始する。一説明例においては、モード決定モジュール715に入力される様々な情報は、スライス型と、I n t r a 4 x 4 c

10

20

30

40

50

ostと、Intra16x16costと、IntraUV8x8costと、IntraY16x16Modeと、IntraUVModeと、動きベクトルデータ(MVD)と、量子化パラメータ(QP)と、SpPredMB4x4Yと、SpPredMB16x16Yと、SpPredMB8x8Uと、SpPredMB8x8Vと、レート歪みフラッグと、生YMB画素と、生UMB画素と、生VMB画素と、を含む。次に、プロセス1200は、ブロック1204符号化初期設定に進み、ブロック1204符号化初期設定は、ブロック1206によって示されるように符号器初期設定を指示する入力信号又はインタフェースによって開始させることができる。初期設定は、許容されたモード(スキップ、指示を含む)を設定することと、モード重みを設定することと(要求される場合であり、デフォルトは、全モードに関して等しい重みになる)、バッファを設定すること、とを含むことができる。初期設定後は、プロセス1200は、ブロック1208に進み、モード決定に関する主処理が行われ、各々の許容されたモードに関するマクロブロック(MB)モードコストの計算、重み付け係数による各MBモードコストの重み付け、及び最低MBモードコストモードを選択することを含む。これらの動作に関わる入力は、ブロック1210及び1212によって示されるように、動き推定(例えば、MVD及び予測)と、空間的予測(例えば、すべてのイントラコスト及び予測)と、を含む。モード決定モジュール715には、特に圧縮率を向上させるブロック1214のエントロピー符号化がインタフェースされている。次に、プロセス1200は、ブロック1216に進み、符号器第2パス部分706に情報を渡すようにバッファが更新される。最後に、プロセス1200は、ブロック1218に進み、“最良の”符号化モードを符号器第2パス部分706に通信することができる。

10

20

#### 【0221】

##### 符号器第2パス部分

再度図7に関して、符号器228の第2パス部分706は、符号化の第2パスを行うための第2パス符号器モジュール232を含む。第2パス符号器232は、モード決定モジュール715から出力を受信する。第2パス符号器232は、MC/変換量子化モジュール726と、ジグザグ(ZZ)/エントロピー符号器728と、を含む。第2パス符号器232の結果は、スケーラビリティモジュール730及びビットストリームパッキングモジュール731に出力され、該モジュールは、符号化された基本層及び拡張層を(図2に示される)同期化層240を介してトランスコーダ200による送信のために出力する。図2に示されるように、第2パス符号器232及び再符号器234からの基本層及び拡張層が同期化層240によってアセンブルされ、基本層と、拡張層と、データPES244(例えば、CC及びその他のテキストデータ)と、音声PES246と、を含むパケット化されたPES242になることが注記される。音声符号器236は、復号された音声情報218を受信し、代わりに、情報を符号化して符号化された情報238を同期化層240に出力する。

30

#### 【0222】

##### 再符号器

再度図7に関して、符号器第2パス部分706は、図2の再符号器234に対応する再符号器234も含む。再符号器234は、第1パス部分702の出力も受信し、MC/変換量子化726及びZZ/エントロピーコーディング728部分を含む。さらに、スケーラビリティモジュール730は、再符号器234に出力する。再符号器234は、再符号化の結果得られた基本層及び拡張層を、シンクロナイザ(例えば、図2に示される同期化層240)への送信のためにビットストリームパッキングモジュール731に出力する。図7の符号器228例は、(例えば、圧縮効率を向上させるために)第2パス符号化を微調整するのを援助するために第2パス符号器232内のMC/変換量子化モジュール234及び再符号器234内のZZ/エントロピーモジュール736の両方にビットストリームパッキングフィードバックを提供するレート制御微調整モジュール738も含む。

40

#### 【0223】

##### 誤り耐性モジュール

50

図7に示される符号器228例は、第2パス部分706内の誤り耐性モジュール740も含む。誤り耐性モジュール740は、ビットストリームパッキングモジュール731及びスライス/MB順序設定モジュール722と通信する。誤り耐性モジュール740は、プリプロセッサ228からメタデータを受信し、誤り耐性方式、例えば、スライス及びアクセスユニットをフレーム境界にアライメントさせること、予測階層及び選択的イントラリフレッシュ、を選択する。誤り耐性方式の選択は、メタデータ内で受信された情報、又はビットストリームパッキングモジュール731及びスライス/MB順序設定モジュール722から誤り耐性モジュールに通信された情報に基づくことができる。誤り耐性モジュール740は、選択された誤り耐性プロセスを実装するために第1パス部分702内のスライス/マクロブロック(MB)順序設定モジュールに情報を提供する。誤りが発生しやすい環境を通じての映像送信は、より明瞭でより少ない誤りを含むデータを視聴者に提示することができる誤り耐性戦略及びアルゴリズムを採用することができる。以下の誤り耐性に関する説明は、個々の既存の又は将来の用途、トランスポート及び物理層又はその他の技術、又はその組合せに対して適用することができる。有効な誤り頑健性アルゴリズムは、OSI層間での誤りが発生しやすい性質及び誤り保護能力に関する理解を、通信システムの望ましい性質、例えば低レイテンシー及び高スループットと関連して統合したものである。誤り耐性処理は、マルチメディアデータのコンテンツ情報、例えば、マルチメディアデータのコンテンツ分類に基づくことができる。主な利点の1つは、フェーディング及びマルチパスチャネルに関する誤りからの回復力である。後述される誤り耐性手法は、符号器228内に(特に、誤り耐性モジュール740及びスライス/MB順序設定モジュール722内に)組み込むことができるプロセスに特に関連するものであり、誤りが発生しやすい環境におけるデータ通信に一般的に拡大することができる。

#### 【0224】

##### 誤り耐性

予測に基づくハイブリッド圧縮システムに関して、イントラコーディングされたフレームは、時間的な予測なしで独立してコーディングされる。インターコーディングされたフレームは、過去のフレーム(Pフレーム)及び将来のフレーム(Bフレーム)から時間的に予測することができる。最良の予測子(predictor)は、基準フレーム(1つ以上)内での探索プロセスを通じて識別することができ、最良のマッチを識別するためにSAD等の歪み評価基準が用いられる。現在のフレームの予測コーディングされた領域は、様々なサイズ及び形状のブロック(16×16、32×32、8×4等)又は例えばセグメンテーションを通じてオブジェクトとして識別された画素グループであることができる。

#### 【0225】

時間的予測は、典型的には数多くのフレーム(例えば、10乃至数10のフレーム)にわたって拡大し、フレームがIフレームとしてコーディングされたときに終了され、GOPは、典型的にはIフレーム周波数によって定義される。最高のコーディング効率を達成させるためには、例えばGOPはシーンであり、GOP境界がシーンの境界とアライメントされ、シーン変化フレームがIフレームとしてコーディングされる。動きが少ないシーケンスは、相対的に静止した背景を具備し、動きは、一般的には前景オブジェクトに制限される。該少ない動きのシーケンスのコンテンツ例は、最も視聴されているコンテンツの30%超がこの性質を有するニュース番組及び天気予報番組を含む。動きが少ないシーケンスにおいては、領域のほとんどがインターコーディングされ、予測されたフレームは、中間の予測されたフレームを通じてIフレームを再参照する。

#### 【0226】

図22に関して、Iフレーム内のイントラコーディングされたブロック2205は、コーディングされたフレーム(又はAU)P1内のインターコーディングされたブロック2210に関する予測子である。この例においては、これらのブロックの領域は、背景の静止部分である。イントラコーディングされたブロック2205の誤りに対する感度は、その“重要度”がより高いことも意味する“優れた”予測子であるため、連続的な時間的予

10

20

30

40

50

測を通じて上昇する。さらに、イントラコーディングされたブロック 2205 は、予測チェーンと呼ばれるこの時間的予測チェーンのおかげで、ディスプレイ内により長い時間（図内の例におけるシーンの継続時間）持続する。

【0227】

予測階層は、この重要度レベル又は持続性評価基準に基づいて生成され、親（イントラコーディングされたブロック 2205）が最上部、子供が最下部に位置するブロックツリーであると定義される。P1 内のインターコーディングされたブロック 215 は、予測階層の第 2 のレベル上にあり、以下同様であることに注目すること。葉は、予測チェーンを終了させるブロックである。

【0228】

予測階層は、コンテンツの種類（ニュースだけでなく、例えば音楽及びスポーツ）に関わりなく映像シーケンスに関して生成することができ、予測に基づく映像（及びデータ）圧縮全般に適用可能である（このことは、本特許出願において説明される全発明に適用される）。予測階層が構築された時点で、後述される誤り耐性アルゴリズム、例えば適応型イントラリフレッシュ、は、より効果的に適用することができる。重要度評価基準は、例えば隠蔽動作及び適応型イントラリフレッシュを適用してコーディングされたビットストリームの誤り耐性を向上させることを通じての所定のブロックの誤りからの回復能力に基づくことができる。重要度評価基準に関する推定は、ブロックが持続性メトリックスとも呼ばれる予測子として用いられる回数に基づくことができる。持続性メトリックスは、予測誤りの伝播を抑止することによってコーディング効率を向上させるためにも用いられる。持続性メトリックスは、より高い重要度を有するブロックに関するビット割り当ても増大させる。

【0229】

適応型イントラリフレッシュ

適応型イントラリフレッシュは、マルチメディアデータのコンテンツ情報に基づくことができる誤り耐性技術である。イントラリフレッシュプロセスにおいては、幾つかの MB は、標準的な R-D 最適化がインターコーディングされた MB であることを要求するにもかかわらずイントラコーディングされる。AIR は、イントラコーディングされた MB を P 又は B フレーム内に導入するために動き重み付きイントラリフレッシュを採用する。これらのイントラコーディングされた MB は、基本層内に含まれ、QP<sub>0</sub> 又は QP<sub>1</sub> のいずれかによって符号化することができる。QP<sub>0</sub> が基本層に関して用いられる場合は、拡張層ではリファインメントは行われるべきではない。QP<sub>1</sub> が基本層に関して用いられる場合は、リファインメントが適しており、リファインメントが行われない場合は、拡張層において品質低下が顕著になる。インターコーディングは、コーディング効率の点でイントラコーディングよりも効率的であるため、拡張層におけるこれらのリファインメントは、インターコーディングされる。この方法により、基本層係数は拡張層に関しては用いられず、拡張層においては新たな動作を導入せずに品質が向上される。

【0230】

幾つかの側面においては、適応型イントラリフレッシュは、動き重み付きの代わりに又は動き重み付きに加えてマルチメディアデータのコンテンツ情報（例えば、コンテンツ分類）に基づくことができる。例えば、コンテンツ分類が相対的に高い（例えば、シーンが大きい空間的及び時間的複雑さを有する）場合は、適応型イントラリフレッシュは、相対的により多くのイントラコーディングされた MB を P 又は B フレーム内に導入することができる。代替として、コンテンツ分類が相対的に低い（動きがより少なく、小さい空間的及び/又は時間的複雑さを有するシーンであることを示す）場合は、適応型イントラリフレッシュは、より少ないイントラコーディングされた MB を P 又は B フレーム内に導入することができる。誤り耐性を向上させるための該メトリックス及び方法は、無線マルチメディア通信の状況のみではなく、（例えば、グラフィックスレンダリングにおける）データ圧縮及びマルチメディア処理全般に関しても適用することができる。

10

20

30

40

50

## 【 0 2 3 1 】

## チャンネルスイッチフレーム

本明細書において定義されるチャンネルスイッチフレーム（CSF）は、高速チャンネル取得及び放送マルチプレックス内のストリーム間での高速チャンネル変更を目的として放送ストリーム内の該当位置に挿入されたランダムアクセスフレームを表す広義の用語である。チャンネルスイッチフレームは、主フレームが誤りを有する状態で送信された場合に用いることができる冗長データを提供するため、誤り頑健性も向上させる。典型的には、Iフレーム又はプログレッシブIフレーム、例えば、H.264内のプログレッシブ復号器リフレッシュフレームがランダムアクセスポイントとして機能する。しかしながら、頻繁なIフレーム（又はシーン継続時間よりも短い、短いGOP）は、圧縮効率の有意な低下に至る。イントラコーディングされたブロックは誤り耐性に関して用いることができるため、ランダムアクセス及び誤り耐性を予測階層を通じて実効的に結合させることで誤り頑健性を向上させる一方でコーディング効率を向上させることができる。

10

## 【 0 2 3 2 】

ランダムアクセス切り換え及び誤り頑健性の向上は、協調して達成させることができ、コンテンツ分類等のコンテンツ情報に基づくことができる。少ない動きのシーケンスに関しては、予測チェーンは長く、スーパーフレーム又はシーンを再構築するのに必要な情報の有意な部分は、そのシーンの開始時に生じたIフレームに内蔵されている。チャンネルの誤りはバースト性になる傾向があり、フェードが著しくFEC及びチャンネルコーディングが失敗したときには、隠蔽することができない大きな残留誤りが存在する。この状態は少ない動き（従って低ビットレート）のシーケンスの場合に特に激しく、その理由は、コーディングされたデータ量が映像ビットストリーム内において良好な時間ダイバーシティを提供する上で十分でないこと及びこれらのシーケンスはすべてのビットが再構築に役立つようにする圧縮可能性が高いシーケンスであることである。多い動きのシーケンスは、コンテンツの性質に起因して誤りに対する頑健性がより高く、すべてのフレーム内のより多くの新しい情報が、独立して復号可能であって誤り耐性が本質的により高いコーディングされたイントラブロック数を増加させる。予測階層に基づく適応型イントラリフレッシュは、多い動きのシーケンスに関して高い性能を達成させ、少ない動きのシーケンスに関しては性能の向上はそれほどではない。従って、少ない動きのシーケンスに関しては、Iフレームのうちのほとんどを含むチャンネルスイッチフレームが優れたダイバーシティ源である。誤りがスーパーフレームに悪影響を与えた場合は、後続フレーム内での復号は、予測に起因する失われた情報を復元させるCSFから開始し、誤り耐性が達成される。

20

30

## 【 0 2 3 3 】

多い動きのシーケンス、例えば相対的に高いコンテンツ分類（例えば、6乃至8）を有するシーケンス、の場合は、CSFは、SF内において持続するブロック - 優れた予測子であるブロックから成ることができる。CSFのその他のすべての領域は、イントラブロックで終了されることを意味する短い予測チェーンを有するブロックであるためコーディングする必要がない。従って、CSFは、誤り発生時に予測に起因する失われた情報から回復する働きを依然として行う。少ない動きのシーケンスに関するCSFは、Iフレームの大きさと同等であり、より大きな量子化を通じてより低いビットレートでコーディングすることができ、多い動きのシーケンスに関するCSFは、対応するIフレームよりもはるかに小さい。

40

## 【 0 2 3 4 】

予測階層に基づく誤り耐性は、スケーラビリティに関して適切に機能し、効率の高い階層化コーディングを達成することができる。物理層技術における階層的変調をサポートするスケーラビリティは、特定の帯域幅比による映像ビットストリームのデータパーティショニングを要求することがある。これらの帯域幅比は、（例えば、最低のオーバーヘッドを有する）最適なスケーラビリティに関する常に理想的な比であるわけではない。幾つかの側面においては、1:1の帯域幅比を有する2層スケーラビリティが用いられる。映像ビットストリームを等しい大きさの2層にパーティショニングすることは、少ない動きの

50

シーケンスに関しては効率的でない場合がある。少ない動きのシーケンスの場合は、すべてのヘッダー情報及びメタデータ情報を含む基本層は、拡張層よりも大きい。しかしながら、少ない動きのシーケンスに関するCSFのほうが大きいいため、拡張層内の残りの帯域幅内に適切に納まる。

#### 【0235】

多い動きのシーケンスは、1:1までのデータパーティショニングを最低のオーバーヘッドで達成することができる十分な残存情報を有する。さらに、該シーケンスに関するチャンネル切り換えフレームは、多い動きのシーケンスに関するほうがはるかに小さい。従って、予測階層に基づく誤り耐性は、多い動きのシーケンスに関するスケラビリティに関しても同様に適切に機能する。これらのアルゴリズムに関する説明に基づき、中度動きクリップに関して上述される概念を拡大させることが可能であり、提案される概念は、映像コーディング全般に関して適用される。

10

#### 【0236】

##### マルチプレクサ

幾つかの符号器側面においては、符号器によって生成されて符号化されたビットを放送用に準備するために用いられる複数のマルチメディアストリームを符号化するためにマルチプレクサを用いることができる。例えば、図2に示される符号器228の例示的側面においては、同期化層240は、マルチプレクサを具備する。マルチプレクサは、ビットレート割り当て制御を提供するために実装することができる。推定される複雑さをマルチプレクサに提供することができ、マルチプレクサは、多重化された映像チャンネルの集合に

20

#### 【0237】

図18は、複数のマルチメディアストリーム又はチャンネル1802を符号化するシステムを示すブロック図である。マルチメディアストリーム1802は、マルチプレクサ(MUX)1806と通信状態にある各々の符号器1804によって符号化され、マルチプレクサ(MUX)1806は、送信媒体1808と通信状態にある。例えば、マルチメディアストリーム1802は、様々なコンテンツチャンネル、例えば、ニュースチャンネル、スポーツチャンネル、映画チャンネル等に対応することができる。符号器1804は、システムに関して規定された符号化形式に従ってマルチメディアストリーム1802を符号化する。開示される技術の原理及び利点は、映像ストリームの符号化に関して説明されている一方で、例えば音声ストリームを含むマルチメディアストリームに対しても一般的に適用可能である。符号化されたマルチメディアストリームは、マルチプレクサ1806に提供され、マルチプレクサ1806は、様々な符号化されたマルチメディアストリームを結合させて結合されたストリームを送信のために送信媒体1808に送信する。

30

#### 【0238】

送信媒体1808は、様々な媒体、例えば、限定することなしに、デジタル衛星通信、例えば、Direct TV(登録商標)、デジタルケーブル、有線及び無線のインターネット通信、光ネットワーク、携帯電話ネットワーク等に対応することができる。送信媒体1808は、例えば、無線周波数(RF)への変調を含むことができる。典型的には、スペクトル上の制約等に起因して、送信媒体は、制限された帯域幅を有し、マルチプレクサ1806から送信媒体へのデータは、相対的に一定のビットレート(CBR)で維持される。

40

#### 【0239】

従来のシステムにおいては、マルチプレクサ1806の出力部において一定のビットレート(CBR)を使用することは、マルチプレクサ1806に入力される符号化されたマ

50

マルチメディア又は映像ストリームも C B R であることを要求することがある。背景技術において説明されるように、映像コンテンツを符号化時における C B R の使用は、結果的に可変の視覚品質が得られる可能性があり、典型的には望ましくない。

【 0 2 4 0 】

例示されるシステムにおいては、符号器 1 8 0 4 のうちの 2 つ以上は、入力データの予想される符号化上の複雑さを通信する。符号器 1 8 0 4 のうちの 1 つ以上は、それに応答して、適応されたビットレート制御をマルチプレクサ 1 8 0 6 から受信する。このことは、相対的に複雑な映像を符号化することを予想する符号器 1 8 0 4 がこれらの映像フレームに関するより高いビットレート又はより高い帯域幅（1 フレーム当たりより多くのビット）を疑似可変ビットレートの形で受信することを可能にする。このことは、マルチメディアストリーム 1 8 0 2 を一貫した視覚品質で符号化するのを可能にする。相対的に複雑な映像を符号化する特定の符号器 1 8 0 4 によって用いられる余分の帯域幅は、本来であれば符号器が一定のビットレートで動作するために実装された場合にその他の映像ストリーム 1 8 0 4 を符号化するために用いられていたことになるビットから得られる。このことは、マルチプレクサ 1 8 0 6 の出力を一定のビットレート（C B R）に維持する。

10

【 0 2 4 1 】

個々のマルチメディアストリーム 1 8 0 2 は、相対的に“バースト性”を有する可能性がある、すなわち使用される帯域幅が変動する可能性がある一方で、複数の映像ストリームの累積和は、それよりもバースト性が低くなることができる。複雑さがより小さい映像を符号化中のチャンネルからのビットレートは、例えばマルチプレクサ 1 8 0 6 によって、相対的に複雑な映像を符号化中のチャンネルに再割り当てすることができ、このことは、結合された映像ストリーム全体の視覚品質を向上させることができる。

20

【 0 2 4 2 】

符号器 1 8 0 4 は、符号化してひとつに多重化すべき一組の映像フレームの複雑さの表示をマルチプレクサ 1 8 0 6 に提供する。マルチプレクサ 1 8 0 6 の出力は、送信媒体 1 8 0 8 に関して規定されるビットレートよりも高くない出力を提供すべきである。複雑さの表示は、上述されるコンテンツ分類に基づき、選択されたレベルの品質を提供することができる。マルチプレクサ 1 0 0 6 は、複雑さの表示を解析し、割り当てられたビット数又は帯域幅を様々な符号器 1 0 0 4 に提供し、符号器 1 8 0 4 は、この情報を用いて組内の映像フレームを符号化する。このことは、一組の映像フレームを個々に可変ビットレートにし、その一方でグループとして一定のビットレートを達成するのを可能にする。

30

【 0 2 4 3 】

コンテンツ分類は、品質に基づくマルチメディアの圧縮全般をあらゆる一般的圧縮器に関して可能にする際に用いることができる。本明細書において説明されるコンテンツ分類及び方法及び装置は、品質に基づく及び/又はコンテンツに基づくマルチメディアデータ処理の際に用いることができる。一例は、一般的圧縮器に関するマルチメディア全般の圧縮において使用することである。他の例は、伸長器又は復号器又はポストプロセッサにおける伸長又は復号、例えば、補間、リサンプリング、拡張、回復及び提示動作、における使用である。

【 0 2 4 4 】

次に図 1 9 に関して、典型的映像通信システムは、通信ネットワークによって接続された映像符号器及び映像復号器からなる映像圧縮システムを含む。無線ネットワークは、通信チャンネルが経路損失に加えて移動シナリオにおける対数正規フェーディング又はシャドーイング及び多経路フェーディングを呈する 1 つのクラスの誤りが発生しやすいネットワークである。チャンネル上の誤りに対処してアプリケーション層データに関して信頼できる通信を提供するため、R F 変調器は、インターリーバを含む順方向誤り訂正と、畳み込み又はターボ符号化等のチャンネルコーディングと、を含む。

40

【 0 2 4 5 】

映像圧縮は、ソース映像における冗長性を低下させ、コーディングされた映像データの各ビットにおいて搬送される情報量を増大させる。このことは、コーディングされた映像

50

のうちの小さい部分さえも失われたときに品質上の影響を増大させる。映像圧縮システムに内在する空間的及び時間的予測は、損失を悪化させて誤りを伝搬させ、その結果、再構築された映像内において可視のアーティファクトを発生させる。映像符号器における誤り耐性アルゴリズム及び映像復号器における誤り回復アルゴリズムは、映像圧縮システムの誤り頑健性を向上させる。

【 0 2 4 6 】

典型的には、映像圧縮システムは、基礎となるネットワークにとらわれない。しかしながら、誤りが発生しやすいネットワークにおいては、アプリケーション層における誤り保護アルゴリズムをリンク/物理層における F E C 及びチャネルコーディングと統合するか又は整合させることが非常に望ましく、システムの誤り性能を向上させる際に最も高い効率を提供する。

10

【 0 2 4 7 】

図 1 4 は、フレームを符号化するために符号器 2 2 8 において発生する可能性があるレート-歪みデータフローの一例を示す。プロセス 1 4 0 0 は、開始 1 4 0 2 において開始し、決定ブロック 1 4 0 4 に進み、決定ブロック 1 4 0 4 において、シーン変化検出器入力 1 4 1 0 を (例えば、メタデータを介して) プリプロセッサ 2 2 6 から受信し、誤り耐性入力 1 4 0 6 が取得される。選択されたフレームが I フレームであることを情報が示す場合は、プロセスは、フレームをイントラ符号化する。選択されたフレームが P 又は B フレームであることを情報が示す場合は、プロセスは、イントラコーディング及び動き推定 (インター) コーディングを用いてフレームを符号化する。

20

【 0 2 4 8 】

ブロック 1 4 0 4 の条件に関して肯定的な状態が生じた後に、プロセス 1 4 0 0 は、準備ブロック 1 4 1 4 に進み、レート R は、R - D 曲線に基づく希望される目標品質である値  $R = R_{q u a l}$  に設定される。この設定は、R - D 曲線を具備するデータブロック 1 4 1 6 から受信される。次に、プロセス 1 4 0 0 は、ブロック 1 4 1 8 に進み、ブロック 1 4 2 0 のコンテンツ分類プロセスからの画像/映像活動情報 (例えば、コンテンツ分類) に基づいてレート制御ビット割り当て {  $Q_{p i}$  } が行われる。

【 0 2 4 9 】

レート制御ビット割り当てブロック 1 4 1 8 は、ブロック 1 4 2 2 において動き推定のために用いられる。動き推定 1 4 2 2 は、プリプロセッサ 1 4 1 2 からメタデータの入力を、ブロック 1 4 2 4 から動きベクトル平滑化 (M P E G - 2 + 履歴) を、ブロック 1 4 2 6 から複数の基準フレーム (因果 + 非因果マクロブロック M B) を受信することもできる。次に、プロセス 1 4 0 0 は、ブロック 1 4 2 8 に進み、レート制御ビット割り当て (  $Q_{p i}$  ) に関してイントラコーディングされたモードに関するレート計算が決定される。次に、プロセス 1 4 0 0 はブロック 1 4 3 0 に進み、モード及び量子化パラメータが決定される。ブロック 1 4 3 0 のモード決定は、ブロック 1 4 2 2 の入力、誤り耐性 1 4 0 6 入力、及びブロック 1 4 3 2 において決定されるスケラビリティ R - D に基づいて行われる。モードが決定された時点で、フローはブロック 1 4 3 2 に進む。ブロック 1 4 3 0 から 1 4 3 2 への流れは、データが符号器の第 1 パス部分から第 2 パス部分に渡されるときに発生することが注記される。

30

40

【 0 2 5 0 】

ブロック 1 4 3 2 において、変換及び量子化が符号器 2 2 8 の第 2 のパスによって行われる。変換/量子化プロセスは、ブロック 1 4 4 4 によって示されるように調整又は微調整される。この変換/量子化プロセスは、レート制御微調整モジュール (図 7) による影響を受ける可能性がある。次に、プロセス 1 4 0 0 は、ジグザグソーティング及びエントロピーコーディングを行って符号化された基本層を生成するためにブロック 1 4 3 4 に進む。ジグザグソーティングは、量子化されたデータを符号化のために効率的な形式で準備する。エントロピーコーディングは、一連のビット符号を用いて一組の可能なシンボルを表す圧縮技術である。図 3 1 乃至 3 6 を参照して前述されるように、変換/量子化ブロック 1 4 3 2 の拡張層の結果は加算器 1 4 3 6 にも送信され、加算器 1 4 3 6 は、基本層を

50

減じてその結果を拡張層に関するZZ/エントロピーコーダ1438に送信する。追記すると、拡張層は、真のレートのコンテンツ分類1420を更新するために及びレート制御によって用いられるビットレートの長期及び短期の履歴を決定する動作のためにフィードバックされる(ライン1440の真のレートの更新を参照)。

#### 【0251】

図17Aは、誤り耐性に関して冗長データを符号化されたデータ内に挿入するためにマルチメディアデータを符号化するプロセスの流れ図であり、符号化することは、コンテンツ情報に基づく。プロセス1700は、図17Bに示される装置又は本明細書において開示されるトランスコーダ200を含むその他の構成要素によって実行することができる。プロセス1700が開始し、ブロック1702においてマルチメディアデータのコンテンツ情報  
10  
を入手する。この入手は、コンテンツ分類器手段、マルチメディアデータのコンテンツ情報を入手するためのモジュール1712(図17B)又は他の構成要素、例えば図7のコンテンツ分類器712、によって行うことができる。次に、プロセス1700は、ブロック1704に進み、マルチメディアデータを符号化して符号化されたビットストリームを形成し、符号化することは、コンテンツ情報に基づいて、イントラコーディングされたデータから成る少なくとも1つの冗長ブロックを符号化されたビットストリーム内に挿入することを含む。このことは、符号化手段、マルチメディアデータを符号化するためのモジュール1714(図17B)、又は他の例においてはトランスコーダ200(図7)によって行うことができる。データ境界を時間境界とアライメントさせることは、誤り耐性を得ることができる。コンテンツ分類は、マルチメディアデータの複雑さに基づくこと  
20  
ことができ、複雑さは、空間的複雑さと、時間的複雑さと、を具備する。挿入される冗長データ量は、マルチメディアデータの複雑さに反比例することができる。イントラコーディングされたデータは、少なくとも1つの冗長データフレームを含むことができ、少なくとも1つの冗長データフレームは、Iフレーム、又はIフレームの実質的な全体を含む。イントラコーディングされたデータから成る少なくとも1つの冗長フレームは、チャンネルスイッチフレームを含むことも可能である。

#### 【0252】

図17Bは、図17Aに示されるプロセス1700を実行することができる符号化装置1710のブロック図である。装置1710は、コンテンツ分類手段、例えばマルチメディアデータのコンテンツ情報を入手するためのモジュール1712、を含むことができる  
30  
。コンテンツ分類手段の例も図7を参照して例示及び説明されている(例えば、コンテンツ分類モジュール714)。幾つかの側面においては、コンテンツ分類手段は、プリプロセッサ、符号器、又はトランスコーダ内に組み入れることができる。装置1710は、時間領域においてデータ境界をフレーム境界とアライメントさせるためにマルチメディアデータを符号化するためのモジュールであって、図17Aのブロック1704において説明される符号化を行うことができる符号化手段1714も含む。該機能は、符号器又はトランスコーダ内に含めることができる。例えば、トランスコーダ200は、図17Aのブロック1704において説明される符号化を行うことができる符号化手段も含む。

#### 【0253】

図23、24、27及び28は、本明細書において説明される側面を具体化するマルチメディアデータ符号化方法の典型例を示すプロセス流れ図である。図23は、コンテンツ情報に基づいてマルチメディアデータを符号化するプロセス2300を示すプロセス流れ図である。ブロック2305において、プロセス2300は、符号化されたマルチメディアデータを受信し、ブロック2310において、プロセス2300は、マルチメディアデータを復号する。ブロック2315において、プロセス2300は、復号されたマルチメディアデータと関連づけられたコンテンツ情報を決定する。ブロック2320において、プロセス2300は、コンテンツ情報に基づいてマルチメディアデータを符号化する。  
40

#### 【0254】

図24は、コンテンツ情報レベルに基づいてデータ境界をアライメントさせるためにマルチメディアデータを符号化するプロセス2400を示すプロセス流れ図である。プロッ  
50

ク 2 4 0 5 において、プロセス 2 4 0 0 は、マルチメディアデータと関連づけられたコンテンツ情報を入手し、このことは、例えばプリプロセッサ 2 2 6 又は図 7 に示されるコンテンツ分類モジュール 7 1 2 によって行うことができる。ブロック 2 4 1 0 において、プロセス 2 4 0 0 は、コンテンツ情報に基づいてデータ境界をアライメントさせるためにマルチメディアデータを符号化する。例えば、スライス境界及びアクセスユニット境界は、符号化中のマルチメディアデータのコンテンツ分類に基づいてフレーム境界とアライメントされる。符号化されたデータは、後続する処理及びノ又はモバイルデバイスへの送信のために利用可能であり、プロセス 2 4 0 0 が終了する。

【 0 2 5 5 】

図 2 7 は、コンテンツ情報に基づいて適応型イントラリフレッシュ方式を用いてデータを符号化するプロセス 2 7 0 0 を示すプロセス流れ図である。プロセス 2 7 0 0 が開始するときには、マルチメディアデータが入手されている。ブロック 2 7 0 5 において、プロセス 2 7 0 0 は、マルチメディアデータのコンテンツ情報を入手する。コンテンツ情報を入手することは、例えば、プリプロセッサ 2 2 6 又は上述されるコンテンツ分類モジュール 7 1 2 によって行うことができる。プロセス 2 7 0 0 は、ブロック 2 7 1 0 に進み、適応型イントラリフレッシュ誤り耐性方式を用いてマルチメディアデータを符号化し、適応型イントラリフレッシュ誤り耐性方式は、コンテンツ情報に基づく。ブロック 2 7 1 0 の機能は、符号器 2 2 8 によって行うことができる。符号化されたデータは、後続する処理及び送信に関して利用可能になり、プロセス 2 7 0 0 が終了する。

【 0 2 5 6 】

図 2 8 は、マルチメディアコンテンツ情報に基づいて冗長 I フレームを用いてマルチメディアデータを符号化するプロセスを示すプロセス流れ図である。プロセス 2 8 0 0 が開始する時点で、マルチメディアデータを処理に関して利用可能である。ブロック 2 8 0 5 において、プロセス 2 8 0 0 は、マルチメディアデータのコンテンツ情報を入手する。上述されるように、このことは、例えば、プリプロセッサ 2 2 6 及びノ又は符号器 2 2 8 によって行うことができる。ブロック 2 8 1 0 において、プロセス 2 8 0 0 は、コンテンツ情報に基づいて 1 つ以上の追加の I フレームを符号化されたデータ内に挿入するためにマルチメディアデータを符号化する。この符号化は、誤り耐性方式と関連して上述されるように符号器 2 2 8 によって行うことができ、採用された誤り耐性方式に依存して基本層又は拡張層内に I フレームを挿入する。ブロック 2 8 1 0 後は、符号化されたデータは、後続する処理及びノ又はモバイルデバイスへの送信に関して利用可能である。

【 0 2 5 7 】

本明細書において説明される方法は、当業者によって知られる様々な通信ハードウェア、プロセッサ及びシステムにおいて実装できることが注目されるべきである。例えば、クライアントが本明細書において説明されるように動作するための一般要件は、クライアントが、コンテンツ及び情報を表示するためのディスプレイと、クライアントの動作を制御するためのプロセッサと、クライアントの動作に関連するデータ及びプログラムを格納するためのメモリと、を有することである。一側面においては、クライアントは、携帯電話である。他の側面においては、クライアントは、通信能力を有するハンドヘルドコンピュータである。さらに他の側面においては、クライアントは、通信能力を有するパソコンである。さらに、様々な側面を実装するために GPS 受信機等のハードウェアをクライアント内に組み入れることができる。本明細書において開示される側面に関連して説明される様々な例示的な論理、論理ブロック、モジュール、及び回路は、本明細書において説明されている機能を果たすように設計された汎用プロセッサ、デジタル信号プロセッサ (DSP)、特定用途向け集積回路 (ASIC)、フィールドプログラマブルゲートアレイ (FPGA)、その他のプログラミング可能な論理デバイス、ディスクリートゲートロジック、ディスクリートトランジスタロジック、ディスクリートハードウェア構成部品、又はそのあらゆる組合せ、とともに実装又は実行することができる。汎用プロセッサはマイクロプロセッサであることができるが、代替として、従来のものでよいプロセッサ、コントローラ、マイクロコントローラ、又はステートマシンであってもよい。さらに、プロセッサは

10

20

30

40

50

、計算装置の組合せ、例えば、DSPと、1つのマイクロプロセッサ、複数のマイクロプロセッサ、DSPコアと関連する1つ以上のマイクロプロセッサ、又はその他のあらゆる該コンフィギュレーションとの組合せ、として実装することもできる。

【0258】

本明細書において開示される側面に関連して説明される様々な例示的な論理、論理ブロック、モジュール、及び回路は、本明細書において説明される機能を果たすように設計された汎用プロセッサ、デジタル信号プロセッサ(DSP)、特定用途向け集積回路(ASIC)、フィールドプログラマブルゲートアレイ(FPGA)、その他のプログラミング可能な論理デバイス、ディスクリートゲートロジック、ディスクリートトランジスタロジック、ディスクリートハードウェア構成要素、又はその他のあらゆる組合せ、とともに実装又は実行することができる。汎用プロセッサはマイクロプロセッサであることができるが、代替として、従来のどのようなプロセッサ、コントローラ、マイクロコントローラ、又はステートマシンであってもよい。さらに、プロセッサは、計算装置の組合せ、例えば、DSPと、1つのマイクロプロセッサ、複数のマイクロプロセッサ、DSPコアと関連する1つ以上のマイクロプロセッサ、又はその他のあらゆる該コンフィギュレーションとの組合せ、として実装することもできる。

10

【0259】

開示される方法及び装置は、1つの形式で符号化された映像データを、符号化が映像データのコンテンツに基づいており符号化が誤りに対する耐性を有する他の形式に符号化された映像データにトランスコーディングすることを提供する。本明細書において開示される例に関連して説明される方法又はアルゴリズムは、ハードウェア内において直接具体化させること、プロセッサによって実行されるソフトウェアモジュール内において具体化させること、又はこれらの2つの組合せにおいて具体化させることができる。ソフトウェアモジュールは、RAMメモリ、フラッシュメモリ、ROMメモリ、EPROMメモリ、EEPROMメモリ、レジスタ、ハードディスク、取り外し可能なディスク、CD-ROM、又は当業において既知であるその他のあらゆる形態の記憶媒体に常駐することができる。1つの典型的な記憶媒体をプロセッサに結合させ、プロセッサが記憶媒体から情報を読み出すようにすること及び記憶媒体に情報を書き込むようにすることができる。代替として、記憶媒体は、プロセッサと一体化させることができる。プロセッサ及び記憶媒体は、ASIC内に常駐することができる。ASICは、ユーザー端末内に常駐することができる。代替として、プロセッサ及び記憶媒体は、ユーザー端末内において個別構成要素として常駐することができる。

20

30

【0260】

上述される例は、単なる典型例であるにすぎず、当業者は、本明細書において開示される発明概念から逸脱せずに上述される例を数多く利用すること及び逸脱することができる。これらの例に対する様々な修正は、当業者にとっては容易に明確になるであろう。本明細書において定められている一般原理は、本発明において説明される斬新な側面の精神及び適用範囲を逸脱しない形で例えば瞬時メッセージ送信サービス又は一般的な無線データ通信用途におけるその他の例に対しても適用可能である。以上のように、本開示の適用範囲は、本明細書において示される例に限定されることを意図するものではなく、本明細書において開示される原理及び斬新な特長に一致する限りにおいて最も広範な適用範囲が認められるべきである。本明細書において用いられている“典型的な”という表現は、“1つの例、事例、又は実例”を示すことのみを目的とするものである。このため、本明細書において“典型的な”例として説明されているいずれの例も、その他の例よりも好ましい又は有利であるとは必ずしも解釈すべきではない。従って、本明細書において説明される斬新な側面は、以下の請求項の適用範囲のみによって定義されるべきである。

40

【図面の簡単な説明】

【0261】

【図1A】異なる映像形式間のトランスコーディングのためのトランスコーダを含むメディア放送システムのブロック図である。

50

【図 1 B】マルチメディアデータを符号化し、符号化された第 1 のデータグループ及び符号化された第 2 の符号化グループを提供するように構成された符号器のブロック図である。

【図 1 C】マルチメディアデータを符号化するように構成されたプロセッサのブロック図である。

【図 2】図 1 のシステムのトランスコーダの例のブロック図である。

【図 3】図 2 のトランスコーダ内において用いられる構文解析器の動作を示す流れ図である。

【図 4】図 2 のトランスコーダ内において用いられる復号器の動作を示す流れ図である。

【図 5】図 2 のトランスコーダによって行われる一連の動作を示す流れ図である。

【図 6】図 2 のトランスコーダにおいて用いることができるプリプロセッサの一連の動作及び機能を示す流れ図である。

【図 7】図 2 のトランスコーダにおいて用いることができる典型的 2 パス符号器のブロック図である。

【図 8】テクスチャ値及び動き値をコンテンツ分類とどのように関連づけるかの一側面を示す分類図例を示した図である。

【図 9】例えば図 7 の符号器内において用いるためのコンテンツ分類に関する典型的動作を示す流れ図である。

【図 10】例えば図 7 の符号器とともに用いるためのレート制御の動作を示した流れ図である。

【図 11】例えば図 7 の符号器とともに用いるための典型的動き推定器の動作を示す流れ図である。

【図 12】例えば図 7 の符号器とともに用いるための典型的モード決定符号器機能の動作を示す流れ図である。

【図 13】図 7 の符号器において用いるためのスケラビリティに影響を与える典型的動作を示す流れ図である。

【図 14】例えば図 7 の符号器において発生するレート歪みデータフローに影響を与える典型的動作を示す流れ図である。

【図 15】符号化の複雑さと、割り当てられたビットと、人間の視覚品質との間の関係を示すグラフである。

【図 16】非線形シーン検出公式を示すグラフである。

【図 17 A】誤り耐性に関して冗長データを符号化されたデータ内に挿入するためにマルチメディアデータを符号化するためのプロセスの流れ図であり、符号化することは、前記コンテンツ情報に基づく。

【図 17 B】図 17 A に示されるプロセス 1700 を実行することができる符号化装置 1710 のブロック図である。

【図 18】動き推定 / 補償を用いたデインターレーシングプロセスを示した図である。

【図 19】マルチメディア通信システムのブロック図である。

【図 20】拡張層及び基本層における映像ビットストリームの構成を示した図である。

【図 21】映像フレーム境界へのスライスのアライメントを示した図である。

【図 22】予測階層を示すブロック図である。

【図 23】コンテンツ情報に基づいてマルチメディアデータを符号化する方法を示すプロセス流れ図である。

【図 24】コンテンツ情報レベルに基づいてデータ境界をアライメントさせるためにマルチメディアデータを符号化する方法を示すプロセス流れ図である。

【図 25】データフレームのセーフアクションエリア及びセーフタイトルエリアを示した図である。

【図 26】データフレームのセーフアクションエリアを示した図である。

【図 27】マルチメディアコンテンツ情報に基づいて適応型イントラリフレッシュを用いてマルチメディアデータを符号化する方法を示す流れ図である。

10

20

30

40

50

【図28】マルチメディアコンテンツ情報に基づいて冗長なIフレームを用いてマルチメディアデータを符号化するプロセスを示す流れ図である。

【図29】現在のフレームと前フレームMV<sub>p</sub>及び現在のフレームと次のフレームMV<sub>N</sub>との間における動き補償ベクトルを示した図である。

【図30】ショット検出を示す流れ図である。

【図31】基本層及び拡張層を示す流れ図である。

【図32】マクロブロックの符号化を示す概略図である。

【図33】基本層及び拡張層を符号化するためのモジュールを示す概略図である。

【図34】基本層及び拡張層の係数選択器プロセスの例を示した図である。

【図35】基本層及び拡張層の係数選択器プロセスの他の例を示した図である。

【図36】基本層及び拡張層の係数選択器プロセスの他の例を示した図である。

【図37】コンテンツ情報に基づくマルチメディアデータの符号化を示す流れ図である。

【図38】逆テレシネプロセスにおける可能なシステム決定を示した図である。

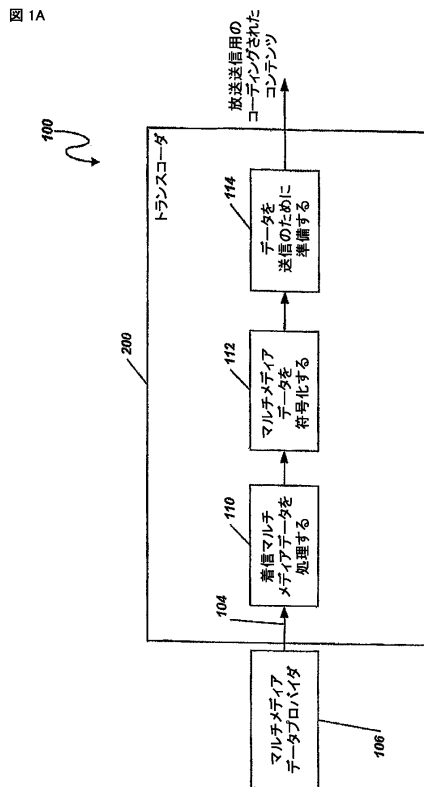
【図39】デブロッキングプロセスによってフィルタリングされるマクロブロック内の境界を示した図である。

【図40】空間-時間的デインターレーシングプロセスを示した図である。

【図41】1D多相リサンプリングの例を示した図である。

【図42】映像ストリーミングにおける適応型GOP構造の例を示す流れ図である。

【図1A】



【図1B】

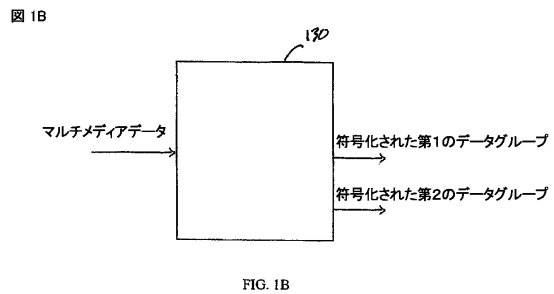


FIG. 1A

FIG. 1B

【図1C】

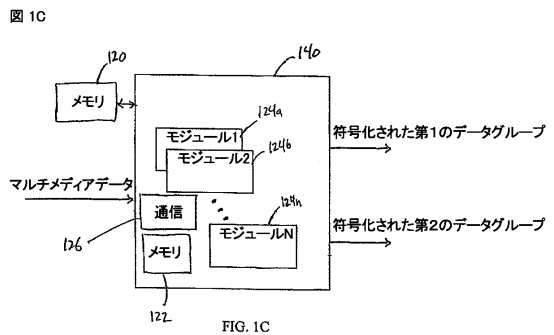


FIG. 1C

【 図 2 】

図 2

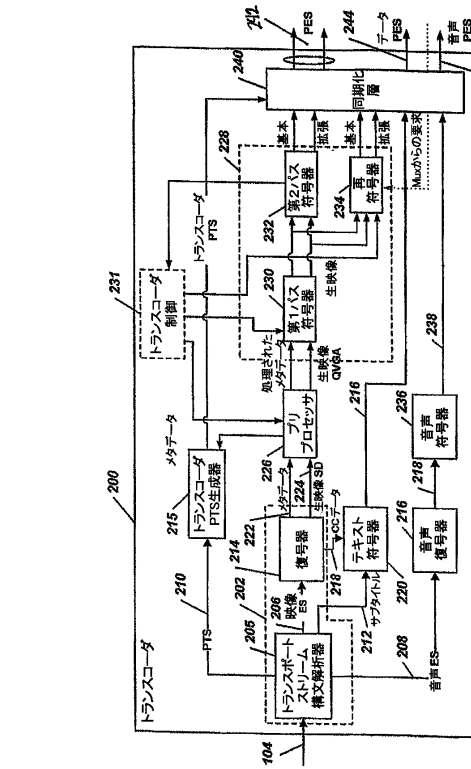


FIG. 2

【 図 4 】

図 4

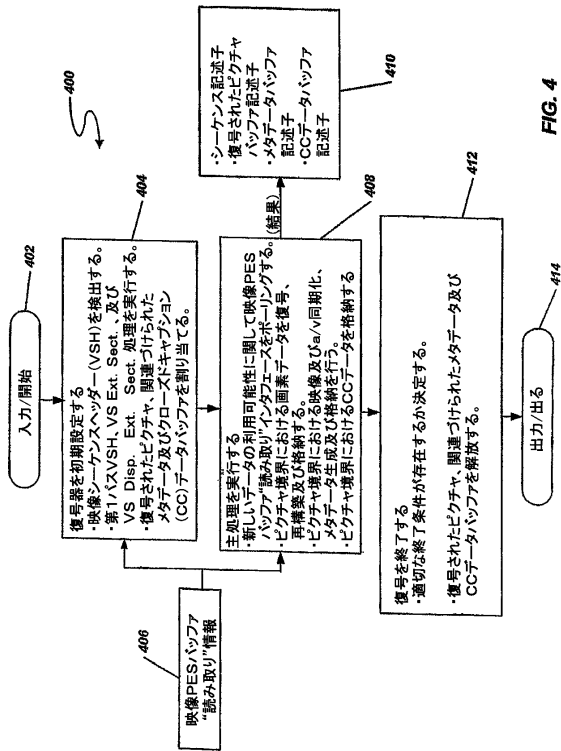


FIG. 4

【 図 3 】

図 3

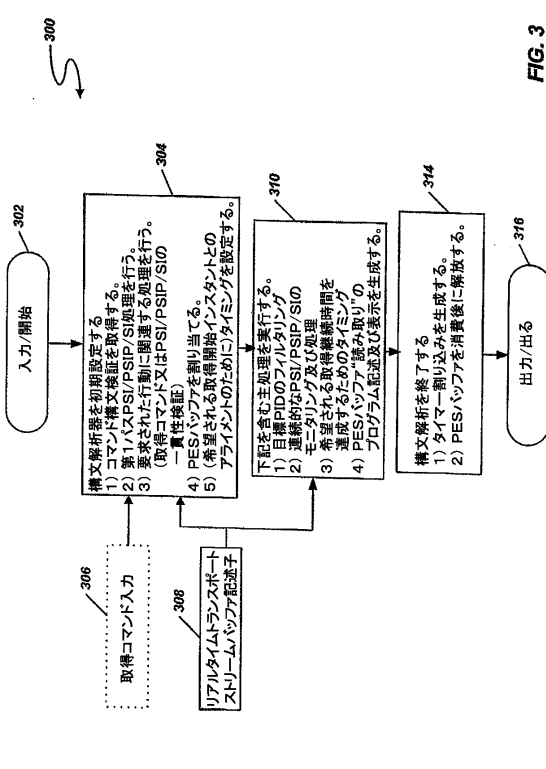


FIG. 3

【 図 5 】

図 5

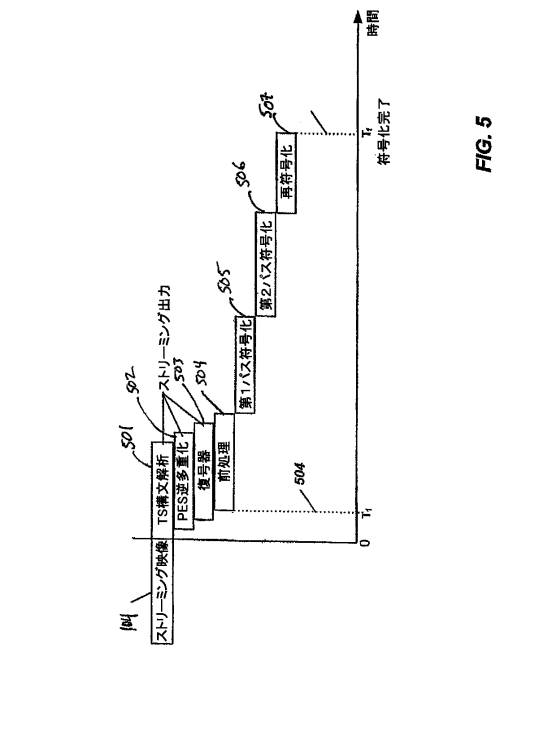


FIG. 5



【 図 10 】

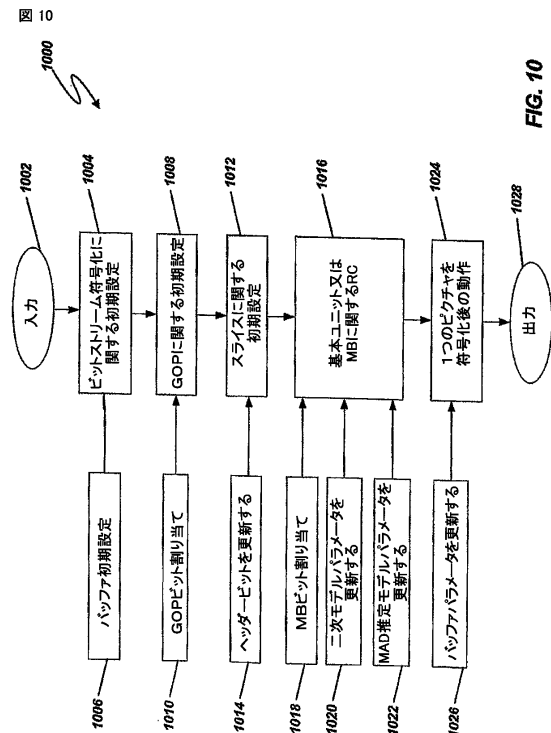


FIG. 10

【 図 11 】

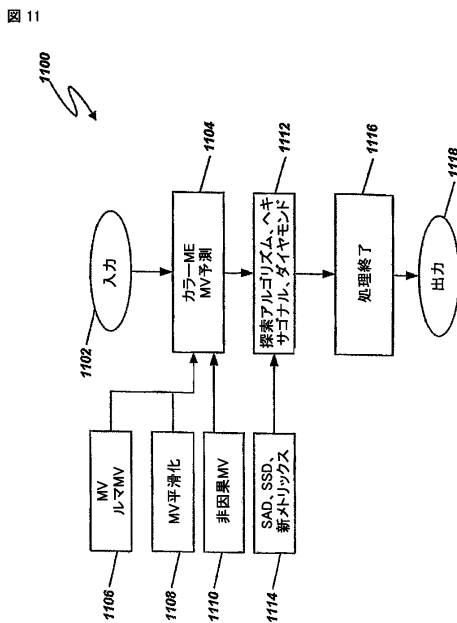


FIG. 11

【 図 12 】

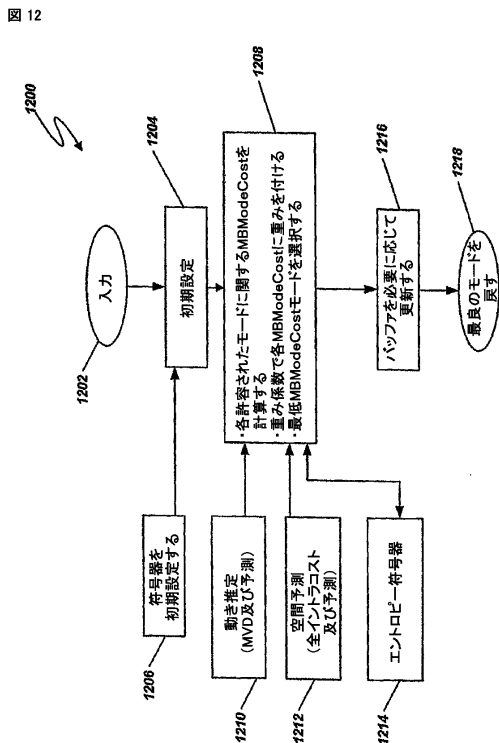


FIG. 12

【 図 13 】

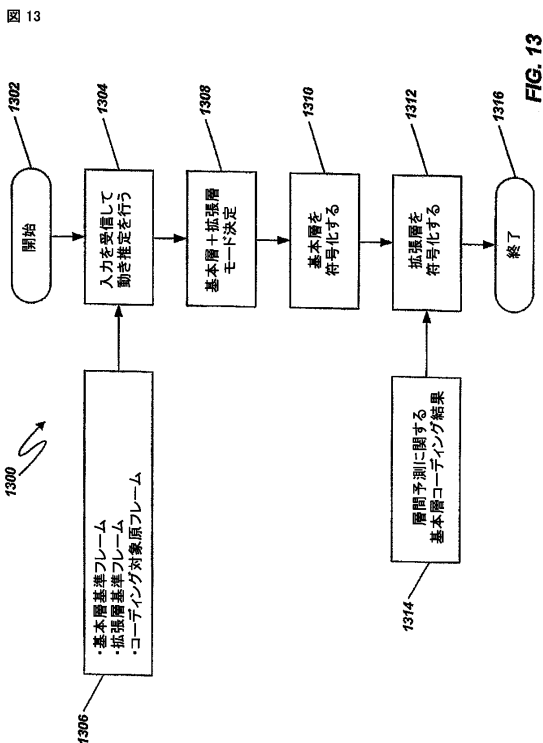


FIG. 13



【 図 18 】

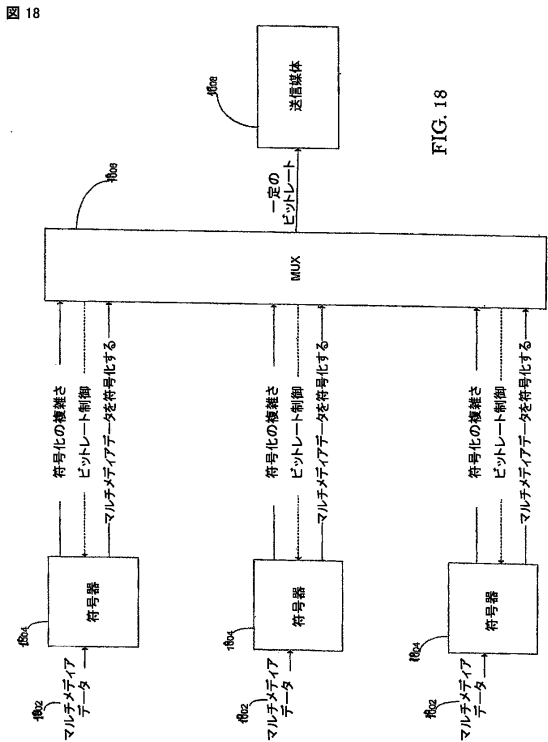


FIG. 18

【 図 19 】

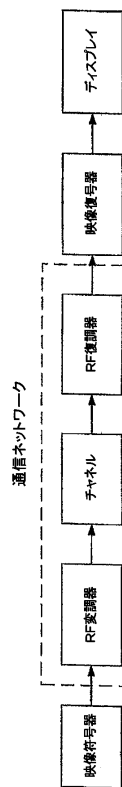


FIG. 19

【 図 20 】

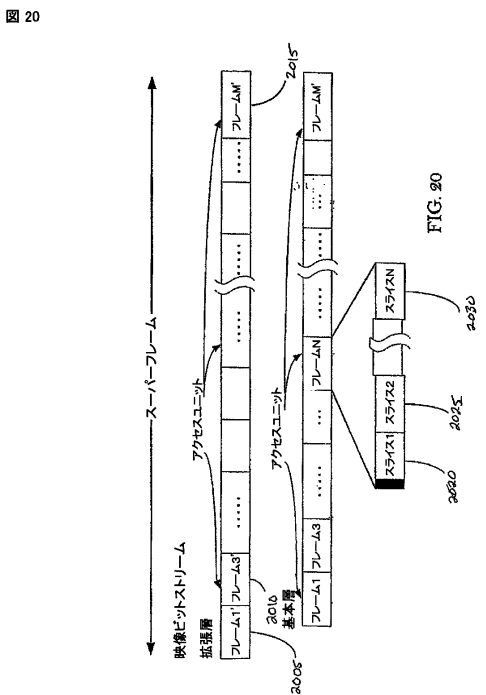


FIG. 20

【 図 21 】

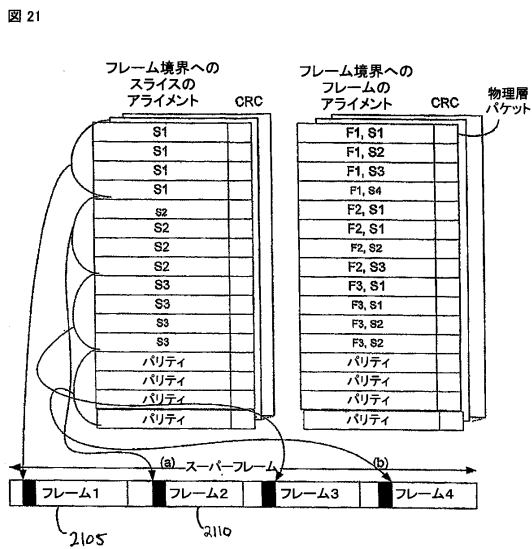
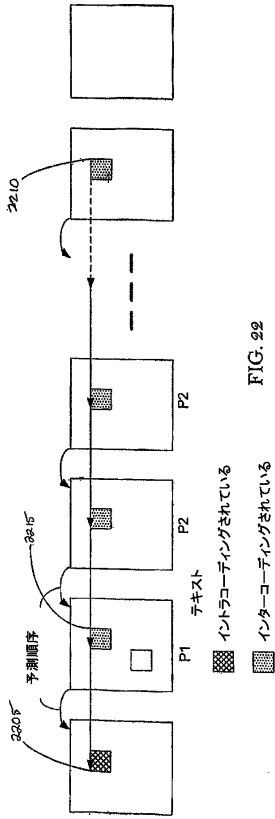


FIG. 21

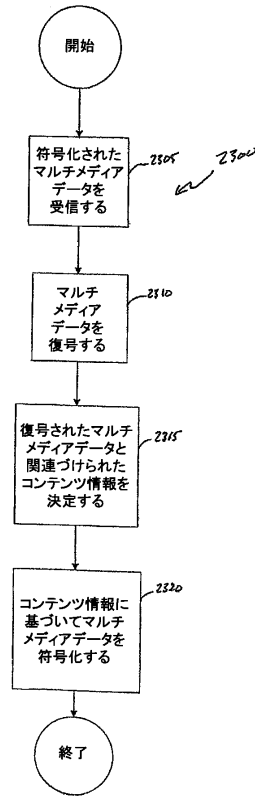
【 図 2 2 】

図 22



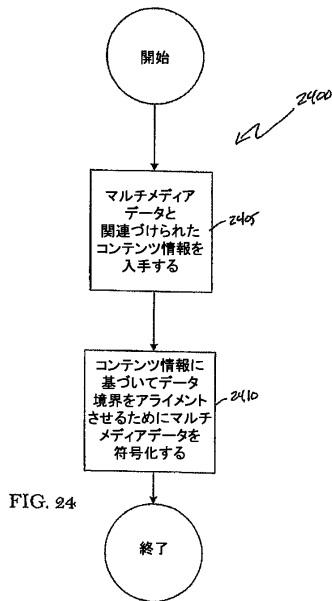
【 図 2 3 】

図 23



【 図 2 4 】

図 24



【 図 2 5 】

図 25

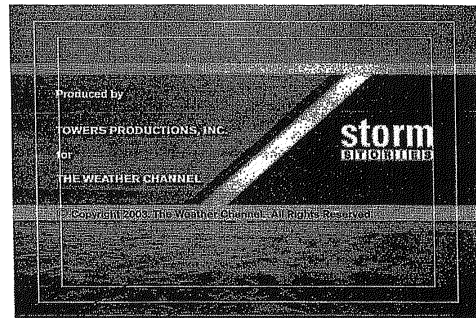


FIG. 25

【図 26】

図 26

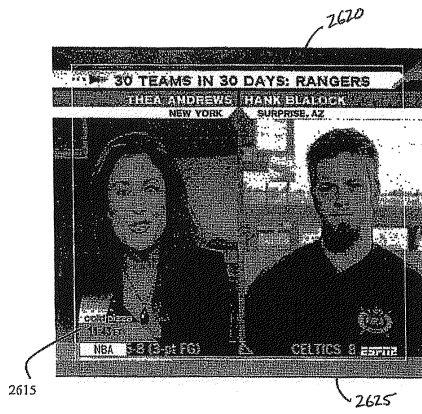


FIG. 26

【図 27】

図 27

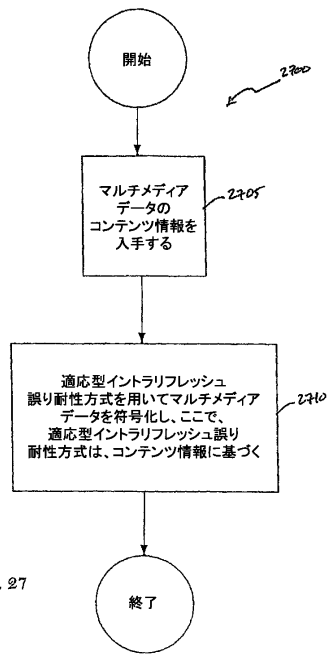


FIG. 27

【図 28】

図 28

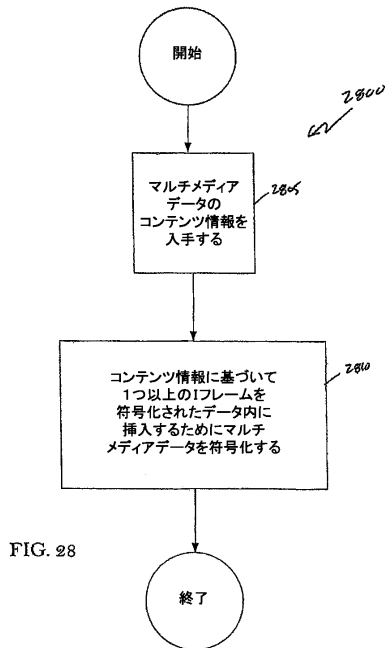


FIG. 28

【図 29】

図 29

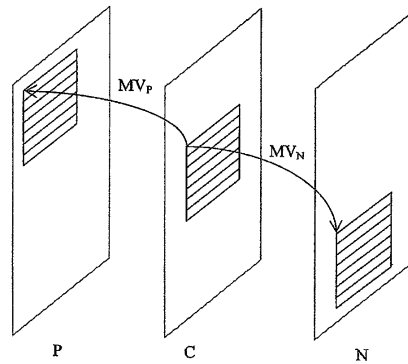


FIG. 29

【図30】

図30

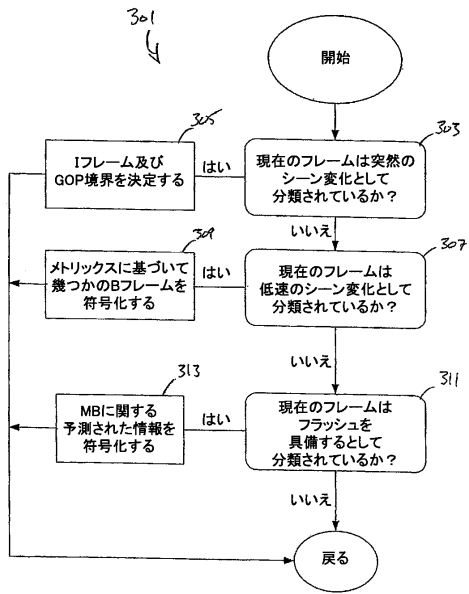


FIG.30

【図31】

図31

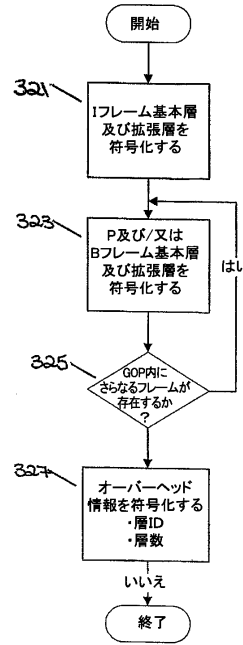


Figure31

【図32】

図32

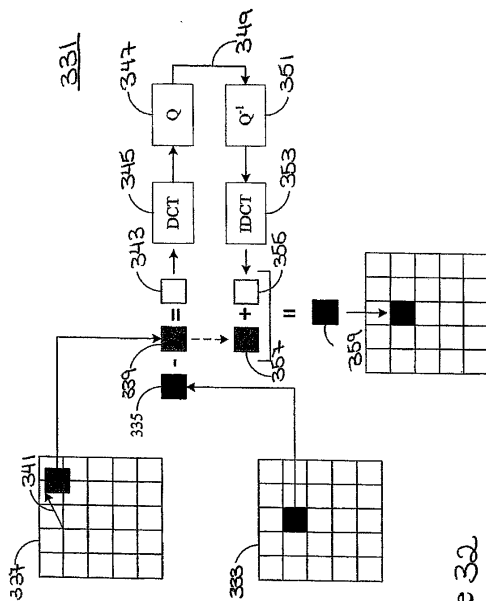


Figure 32

【図33】

図33

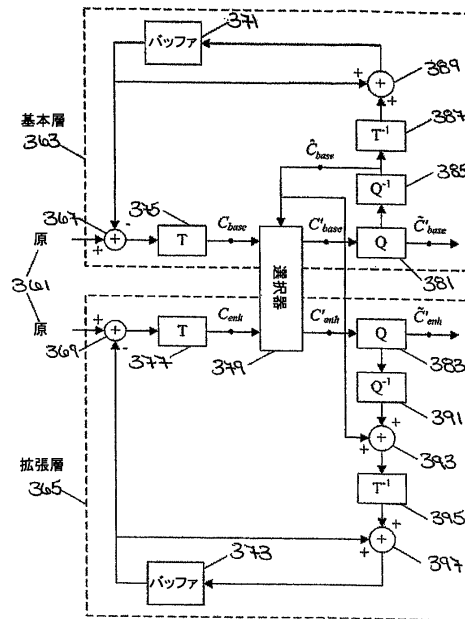


Figure 33

【図34】

図34

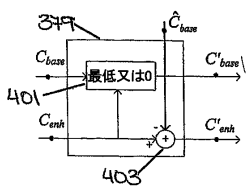


Figure 34

【図35】

図35

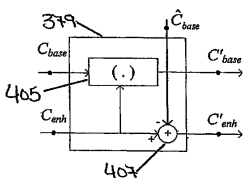


Figure 35

【図36】

図36

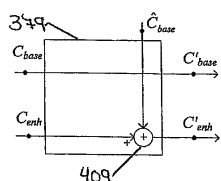


Figure 36

【図39】

図39

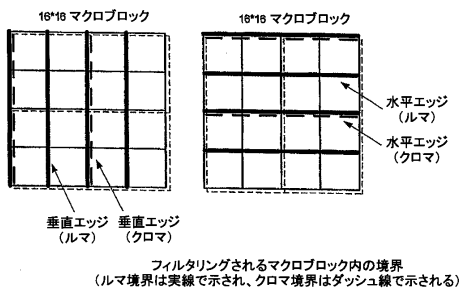


FIG. 39

【図37】

図37

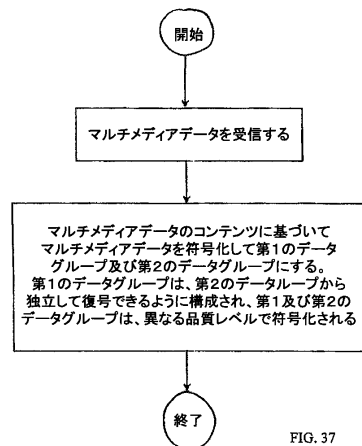


FIG. 37

【図38】

図38

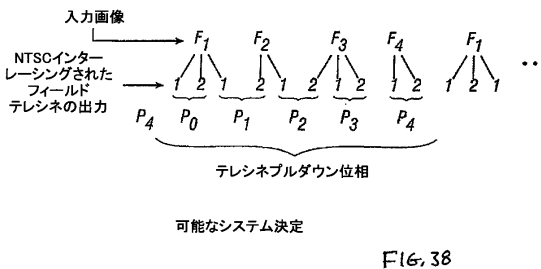


FIG. 38

【図40】

図40

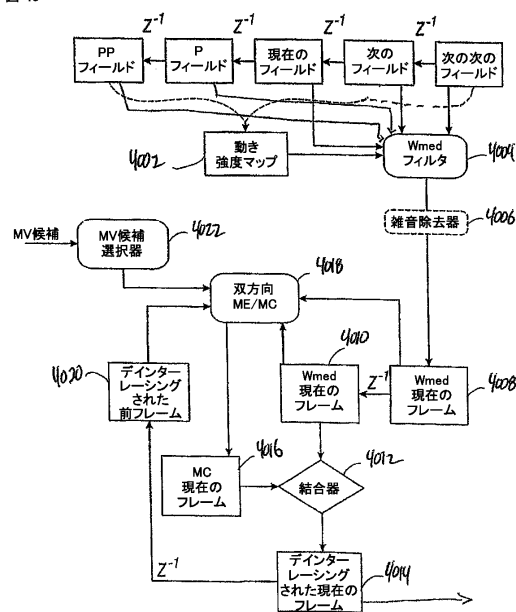


FIG. 40

【図41】

図41

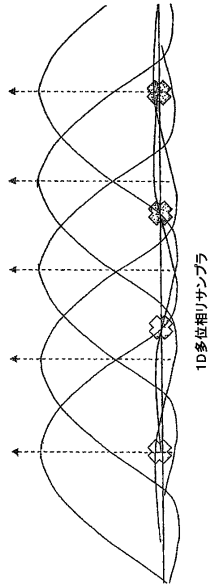


FIG. 41

【図42】

図42

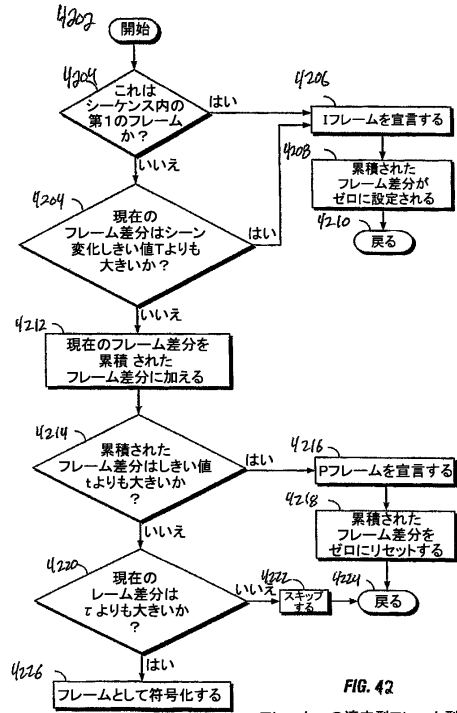


FIG. 42

フレームへの適応型フレーム型割り当て流れ図

## フロントページの続き

- (31)優先権主張番号 60/727,644  
(32)優先日 平成17年10月17日(2005.10.17)  
(33)優先権主張国 米国(US)
- (31)優先権主張番号 60/727,640  
(32)優先日 平成17年10月17日(2005.10.17)  
(33)優先権主張国 米国(US)
- (31)優先権主張番号 60/730,145  
(32)優先日 平成17年10月24日(2005.10.24)  
(33)優先権主張国 米国(US)
- (31)優先権主張番号 60/789,048  
(32)優先日 平成18年4月3日(2006.4.3)  
(33)優先権主張国 米国(US)
- (31)優先権主張番号 60/789,377  
(32)優先日 平成18年4月4日(2006.4.4)  
(33)優先権主張国 米国(US)
- (74)代理人 100075672  
弁理士 峰 隆司
- (74)代理人 100095441  
弁理士 白根 俊郎
- (74)代理人 100084618  
弁理士 村松 貞男
- (74)代理人 100103034  
弁理士 野河 信久
- (74)代理人 100119976  
弁理士 幸長 保次郎
- (74)代理人 100153051  
弁理士 河野 直樹
- (74)代理人 100140176  
弁理士 砂川 克
- (74)代理人 100101812  
弁理士 勝村 紘
- (74)代理人 100092196  
弁理士 橋本 良郎
- (74)代理人 100100952  
弁理士 風間 鉄也
- (74)代理人 100070437  
弁理士 河井 将次
- (74)代理人 100124394  
弁理士 佐藤 立志
- (74)代理人 100112807  
弁理士 岡田 貴志
- (74)代理人 100111073  
弁理士 堀内 美保子
- (74)代理人 100134290  
弁理士 竹内 将訓
- (74)代理人 100127144  
弁理士 市原 卓三

- (74)代理人 100141933  
弁理士 山下 元
- (72)発明者 ラビーン・ドラン、ビジャヤラクシュミ・アール、  
アメリカ合衆国、カリフォルニア州 9 2 1 3 0、サン・ディエゴ、カレ・マー・デ・バレナス  
4 2 7 2
- (72)発明者 ウォーカー、ゴードン・ケント  
アメリカ合衆国、カリフォルニア州 9 2 0 6 4、ボウエイ、ハンティントン・ゲート・ドライブ  
1 4 4 8 4
- (72)発明者 ティアン、タオ  
アメリカ合衆国、カリフォルニア州 9 2 1 2 6、サン・ディエゴ、ニュー・サレム・ストリート  
8 6 8 0、ナンバー145
- (72)発明者 パーミディパティ、パニクマー  
アメリカ合衆国、カリフォルニア州 9 2 1 2 1、サン・ディエゴ、ジェネシー・アベニュー 9  
6 4 5、ナンバーアイ1
- (72)発明者 シ、ファン  
アメリカ合衆国、カリフォルニア州 9 2 1 3 0、サン・ディエゴ、カレ・マー・デ・アーモニア  
4 4 6 0
- (72)発明者 チェン、ペイソン  
アメリカ合衆国、カリフォルニア州 9 2 1 2 6、サン・ディエゴ、カミント・アルバレス 1 0  
8 5 9
- (72)発明者 サブラマニア、シタラマン・ガナパシー  
アメリカ合衆国、カリフォルニア州 9 2 1 2 6、サン・ディエゴ、ヒレリー・ドライブ 9 3 6  
4、ナンバー2204
- (72)発明者 オグズ、セイフラー・ハリト  
アメリカ合衆国、カリフォルニア州 9 2 1 2 2、サン・ディエゴ、フィオア・テラス 5 2 4 0  
、ナンバー417

審査官 上嶋 裕樹

- (56)参考文献 国際公開第2004/114667(WO, A1)  
特開2003-319341(JP, A)  
特開2003-037844(JP, A)  
国際公開第2004/070953(WO, A1)  
特開平07-298272(JP, A)  
特開2000-350217(JP, A)

(58)調査した分野(Int.Cl., DB名)

H04N 7/12  
H04N 7/26  
H04N 7/30  
H04N 7/32