

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第5149188号
(P5149188)

(45) 発行日 平成25年2月20日 (2013. 2. 20)

(24) 登録日 平成24年12月7日 (2012. 12. 7)

(51) Int. Cl.

H04N 7/32 (2006.01)

F I

H04N 7/137

Z

請求項の数 29 (全 78 頁)

(21) 出願番号	特願2008-533623 (P2008-533623)	(73) 特許権者	595020643
(86) (22) 出願日	平成18年9月27日 (2006. 9. 27)		クァアルコム・インコーポレイテッド
(65) 公表番号	特表2009-510933 (P2009-510933A)		QUALCOMM INCORPORATED
(43) 公表日	平成21年3月12日 (2009. 3. 12)		ED
(86) 国際出願番号	PCT/US2006/037949		アメリカ合衆国、カリフォルニア州 92
(87) 国際公開番号	W02007/038702		121-1714、サン・ディエゴ、モア
(87) 国際公開日	平成19年4月5日 (2007. 4. 5)		ハウス・ドライブ 5775
審査請求日	平成20年5月26日 (2008. 5. 26)	(74) 代理人	100108855
(31) 優先権主張番号	60/721, 416		弁理士 蔵田 昌俊
(32) 優先日	平成17年9月27日 (2005. 9. 27)	(74) 代理人	100091351
(33) 優先権主張国	米国 (US)		弁理士 河野 哲
(31) 優先権主張番号	60/727, 643	(74) 代理人	100088683
(32) 優先日	平成17年10月17日 (2005. 10. 17)		弁理士 中村 誠
(33) 優先権主張国	米国 (US)	(74) 代理人	100109830
			弁理士 福原 淑弘

最終頁に続く

(54) 【発明の名称】 コンテンツ情報を使用してマルチメディアアトランスコーディングを調整するコンテンツ駆動トランスコーダ

(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、該方法は下記を備える：

マルチメディアデータを受け取ること；および、

前記マルチメディアデータのコンテンツを分類すること、ここにおいて、前記分類することは、下記を備える：

前記マルチメディアデータの空間情報および時間情報を判定すること、

前記マルチメディアデータの空間的な複雑度および時間的な複雑度を判定すること、および

前記マルチメディアデータの前記判定された空間情報および時間情報と前記判定された空間的な複雑度および時間的な複雑度に基づいて前記マルチメディアデータの前記コンテンツを分類すること、

前記マルチメディアデータの前記コンテンツ分類に基づいて前記マルチメディアデータを第1データレイヤおよび第2データレイヤに符号化すること、ここで、前記第1データレイヤが、前記第2データレイヤと独立に復号化可能であるように構成され、前記第1データレイヤおよび前記第2データレイヤが、異なる品質レベルで符号化され、

ここにおいて、前記第1のデータレイヤは、

前記第1のデータレイヤの残差を**残差係数** C_{base} に変換することであって、前記残差が符号化されるマルチメディアデータと時間基準フレームの最良一致マクロブロックとの差である、変換すること、

10

20

変換された前記残差係数 C_{base} から第 1 のデータレイヤ係数 C'_{base} を計算すること、および

第 1 の量子化パラメータ Q_b によって前記第 1 のデータレイヤ係数 C'_{base} を量子化すること

のステップによって符号化され、

ここにおいて、前記第 2 のデータレイヤは、

前記第 2 のデータレイヤの残差を残差係数 C_{enh} に変換することであって、前記残差が符号化されるマルチメディアデータと時間基準フレームの最良一致マクロブロックとの差である、変換すること、

変換された前記残差係数 C_{enh} から第 2 のデータレイヤ係数 C'_{enh} を計算すること、および

10

第 2 の量子化パラメータによって前記第 2 のデータレイヤ係数 C'_{enh} を量子化すること

のステップによって符号化され、

ここにおいて、

【数 1】

$$C'_{base} = \begin{cases} 0, & \text{if } C_{base} \text{ および } C_{enh} \text{ の符号が反対である場合} \\ \min(C_{base}, C_{enh}), & \text{それ以外場合} \end{cases}$$

$$C'_{enh} = C_{enh} - Q_b^{-1}(Q_b(C'_{base}))$$

20

または

【数 2】

$$C'_{base} = \begin{cases} C_{base}, & |Q_b^{-1}Q_b(C_{base}) - C_{enh}| < |C_{enh}| \text{ の場合} \\ 0, & \text{それ以外の場合} \end{cases}$$

30

および、

【数 3】

$$C'_{enh} = C_{enh} - Q_b^{-1}(Q_b(C'_{base}))$$

40

ここにおいて、 Q_b^{-1} は前記第 1 のデータレイヤの逆量子化パラメータであり、第 1 の量子化パラメータは、前記第 2 の量子化パラメータよりも大きくセットされ、符号化された第 2 のデータレイヤが使用可能でない場合に、表示可能なマルチメディアデータを形成するために、符号化された前記第 1 データレイヤが復号化されることができ、符号化された前記第 1 のデータレイヤと符号化された前記第 2 のデータレイヤとの両方が使用可能である場合に、表示可能なマルチメディアデータを形成するために符号化された前記第 1 データレイヤおよび符号化された前記第 2 データレイヤが組み合わされて復号化されることができる。

【請求項 2】

請求項 1 に記載の方法、ここにおいて、

50

前記第 1 データレイヤが、I フレームおよび P フレームを備え、前記第 2 データレイヤが、I フレーム、P フレーム、および B フレームを備える。

【請求項 3】

請求項 1 に記載の方法、ここにおいて、

前記第 1 データレイヤが、ベースレイヤを備え、前記第 2 データレイヤが、エンハンスメントレイヤを備える。

【請求項 4】

請求項 1 に記載の方法、ここにおいて、

前記符号化することは、前記コンテンツ分類に基づいて前記マルチメディアデータの少なくとも一部のビットレートを割り当てることを備える。

10

【請求項 5】

請求項 1 に記載の方法、ここにおいて、

前記符号化することは、さらに下記を備える：

シーン変化を検出するのに前記コンテンツ分類を使用すること；および、

前記検出されたシーン変化に基づいて前記第 1 データレイヤおよび前記第 2 データレイヤ内に I フレームを含めるかどうかを決定すること。

【請求項 6】

請求項 1 に記載の方法、ここにおいて、

前記符号化することは、さらに、前記コンテンツ分類に基づいて前記マルチメディアデータを符号化するフレームレートを決定することを備える。

20

【請求項 7】

請求項 1 に記載の方法、ここにおいて、

前記符号化することは、前記コンテンツ分類に基づいて前記マルチメディアデータの動き推定を実行することを備える。

【請求項 8】

さらに下記を備える請求項 1 に記載の方法：

前記第 1 データレイヤの符号化の第 1 フレームレートを決定すること；

前記第 2 データレイヤの符号化の第 2 フレームレートを決定すること、

ここにおいて、前記第 1 フレームレートが、前記第 2 フレームレートより低い。

【請求項 9】

30

請求項 1 に記載の方法、ここにおいて、

前記符号化することは、前記コンテンツ分類に基づいて前記マルチメディアデータに対して誤り弾力性処理を実行することを備える。

【請求項 10】

請求項 1 に記載の方法、

ここにおいて、前記第 1 量子化パラメータが、データを符号化するための第 1 ステップサイズを備え、

さらにここにおいて、前記第 2 量子化パラメータが、データを符号化するための第 2 ステップサイズを備え、前記第 1 ステップサイズが、前記第 2 ステップサイズより大きい。

【請求項 11】

40

下記を備える請求項 1 に記載の方法：

前記符号化することは、前記コンテンツ分類に基づいて前記マルチメディアデータ内の雑音を減らすことを備える。

【請求項 12】

雑音を減らすことが、人為的結果の除去を実行することを備える、請求項 11 に記載の方法。

【請求項 13】

請求項 11 に記載の方法、ここにおいて、

雑音を減らすことが、デリンギングフィルタを用いて前記マルチメディアデータの少なくとも一部を処理することを備え、ここにおいて、

50

前記デリリングフィルタの強さが、前記マルチメディアデータの前記コンテンツ分類に基づく。

【請求項 14】

請求項 11 に記載の方法、ここにおいて、

雑音を減らすことが、デブロッキングフィルタを用いて前記マルチメディアデータの少なくとも一部を処理することを備え、ここにおいて、

前記デブロッキングフィルタの強さが、前記マルチメディアデータの前記コンテンツ分類に基づく。

【請求項 15】

雑音を減らすことが、前記マルチメディアデータの選択された周波数をフィルタリングすることを備える、請求項 11 に記載の方法。

10

【請求項 16】

前記デリリングフィルタの前記強さが、前記マルチメディアデータのコンテンツ分類に基づく、請求項 13 に記載の方法。

【請求項 17】

前記デブロッキングフィルタの前記強さが、前記マルチメディアデータのコンテンツ分類に基づく、請求項 14 に記載の方法。

【請求項 18】

符号化することが、前記マルチメディアデータをダウンサンプリングすることを備える、請求項 1 に記載の方法。

20

【請求項 19】

請求項 1 に記載の方法、ここにおいて、

符号化することは、前記マルチメディアデータの符号化のビットレートを決定するために、下記を備える：

前記マルチメディアデータに品質レベルを関連付けること；および、

前記マルチメディアデータの前記品質レベルおよび前記コンテンツ分類を使用すること。

【請求項 20】

マルチメディアデータを処理する装置、該装置は下記を備える：

マルチメディアデータを受け取り、前記マルチメディアデータのコンテンツを分類し、前記マルチメディアデータの前記コンテンツ分類に基づいて前記マルチメディアデータを第 1 データレイヤおよび第 2 データレイヤに符号化するように構成された符号化器、

30

ここにおいて、前記第 1 データレイヤが、前記第 2 データレイヤと独立に復号化可能であるように構成され、前記第 1 データレイヤおよび前記第 2 データレイヤが、異なる品質レベルで符号化され、前記分類することは、前記マルチメディアデータの空間情報および時間情報を判定すること、前記マルチメディアデータの空間的な複雑度および時間的な複雑度を判定すること、および前記マルチメディアデータの前記判定された空間情報および時間情報と前記判定された空間的な複雑度および時間的な複雑度に基づいて前記マルチメディアデータの前記コンテンツを分類することを備え、

ここにおいて、前記第 1 のデータレイヤは、

40

前記第 1 のデータレイヤの残差を残差係数 C_{base} に変換することであって、前記残差が符号化されるマルチメディアデータと時間基準フレームの最良一致マクロブロックとの差である、変換すること、

変換された前記残差係数 C_{base} から第 1 のデータレイヤ係数 C'_{base} を計算すること、および

第 1 の量子化パラメータ Q_b によって前記第 1 のデータレイヤ係数 C'_{base} を量子化すること

のステップによって符号化され、

ここにおいて、前記第 2 のデータレイヤは、

前記第 2 のデータレイヤの残差を残差係数 C_{enh} に変換することであって、前記残差

50

が符号化されるマルチメディアデータと時間基準フレームの最良一致マクロブロックとの差である、変換すること、

変形された前記残差係数 C_{enh} から第 2 のデータレイヤ係数 C'_{enh} を計算すること、および

第 2 の量子化パラメータによって前記第 2 のデータレイヤ係数 C'_{enh} を量子化すること

のステップによって符号化され、

ここにおいて、

【数 4】

$$C'_{base} = \begin{cases} 0, & \text{if } C_{base} \text{ および } C_{enh} \text{ の符号が反対である場合} \\ \min(C_{base}, C_{enh}), & \text{それ以外場合} \end{cases}$$

$$C'_{enh} = C_{enh} - Q_b^{-1}(Q_b(C'_{base}))$$

10

または

【数 5】

$$C'_{base} = \begin{cases} C_{base}, & |Q_b^{-1}Q_b(C_{base}) - C_{enh}| < |C_{enh}| \text{ の場合} \\ 0, & \text{それ以外の場合} \end{cases}$$

20

および、

【数 6】

$$C'_{enh} = C_{enh} - Q_b^{-1}(Q_b(C'_{base}))$$

30

ここにおいて、 Q_b^{-1} は前記第 1 のデータレイヤの逆量子化パラメータであり、第 1 の量子化パラメータは、前記第 2 の量子化パラメータよりも大きくセットされ、符号化された第 2 のデータレイヤが使用可能でない場合に、表示可能なマルチメディアデータを形成するために、符号化された前記第 1 データレイヤが復号化されることができ、符号化された前記第 1 のデータレイヤと符号化された前記第 2 のデータレイヤとの両方が使用可能である場合に、表示可能なマルチメディアデータを形成するために符号化された前記第 1 データレイヤおよび符号化された前記第 2 データレイヤが組み合わされて復号化されることができ

40

【請求項 2 1】

請求項 2 0 に記載の装置、ここにおいて、

前記符号化器が、前記コンテンツ分類に基づいて前記マルチメディアデータの動き推定を実行し、データ動き補償情報を生成するように構成された動き推定モジュールを備え、

さらにここにおいて、前記符号化モジュールが、さらに、前記マルチメディアデータを符号化するのに前記動き補償情報を使用するように構成される。

【請求項 2 2】

請求項 2 0 に記載の装置、ここにおいて、

前記符号化器が、前記第 1 および第 2 の量子化パラメータを決定する量子化モジュール

50

をさらに備える。

【請求項 2 3】

請求項 2 0 に記載の装置、ここにおいて、

前記符号化器が、前記コンテンツ分類に基づいて前記マルチメディアデータの少なくとも一部のビットレートを提供するように構成されたビット割当モジュールをさらに備える。

【請求項 2 4】

請求項 2 0 に記載の装置、ここにおいて、

前記符号化器が、シーン変化を検出するように構成されたシーン変化検出モジュールをさらに備え、前記符号化モジュールが、さらに、前記検出されたシーン変化に基づいて、符号化されたマルチメディアデータに I フレームを含めるように構成される、請求項 2 4 に記載の装置。

10

【請求項 2 5】

請求項 2 0 に記載の装置、ここにおいて、

前記符号化器が、前記コンテンツ分類に基づいて前記マルチメディアデータのフレームレートを決定するように構成されたフレームレートモジュールをさらに備え、

ここにおいて、前記符号化モジュールが、前記フレームレートに基づいて前記マルチメディアデータを符号化する。

【請求項 2 6】

前記符号化器が、さらに、前記コンテンツ分類に基づいて前記マルチメディアデータに対して誤り処理を実行するように構成される、請求項 2 0 に記載の装置。

20

【請求項 2 7】

実行時に、機械に下記を行わせる命令を備えるコンピュータ可読媒体：

マルチメディアデータを受け取ること；

前記マルチメディアデータのコンテンツを分類すること、ここにおいて、前記分類することは、下記を備える：

前記マルチメディアデータの空間情報および時間情報を判定すること、

前記マルチメディアデータの空間的な複雑度および時間的な複雑度を判定すること、および

前記マルチメディアデータの前記判定された空間情報および時間情報と前記判定された空間的な複雑度および時間的な複雑度に基づいて前記マルチメディアデータの前記コンテンツを分類すること、および、

30

前記マルチメディアデータの前記コンテンツ分類に基づいて前記マルチメディアデータを符号化された第 1 データレイヤおよび符号化された第 2 データレイヤに符号化すること、ここにおいて、前記第 1 データレイヤが、前記第 2 データレイヤと独立に復号化可能であるように構成され、前記第 1 データレイヤおよび前記第 2 データレイヤが、異なる品質レベルで符号化され、

ここにおいて、前記第 1 のデータレイヤは、

前記第 1 のデータレイヤの残差を残差係数 $C_{b a s e}$ に変換することであって、前記残差が符号化されるマルチメディアデータと時間基準フレームのベストマッチングマクロブロックとの差である、変換すること、

40

変形された前記残差係数 $C_{b a s e}$ から第 1 のデータレイヤ係数 $C'_{b a s e}$ を計算すること、および

第 1 の量子化パラメータ Q_b によって前記第 1 のデータレイヤ係数 $C'_{b a s e}$ を量子化すること

のステップによって符号化され、および

ここにおいて、前記第 2 のデータレイヤは、

前記第 2 のデータレイヤの残差を残差係数 $C_{e n h}$ に変換することであって、前記残差が符号化されるマルチメディアデータと時間基準フレームの最良一致マクロブロックとの差である、変換すること、

50

変換された前記残差係数 C_{enh} から第 2 のデータレイヤ係数 C'_{enh} を計算すること、および

第 2 の量子化パラメータによって前記第 2 のデータレイヤ係数 C'_{enh} を量子化すること

のステップによって符号化され、

ここにおいて、

【数 7】

$$C'_{base} = \begin{cases} 0, & \text{if } C_{base} \text{ および } C_{enh} \text{ の符号が反対である場合} \\ \min(C_{base}, C_{enh}), & \text{それ以外場合} \end{cases} \quad 10$$

$$C'_{enh} = C_{enh} - Q_b^{-1}(Q_b(C'_{base}))$$

または

【数 8】

$$C'_{base} = \begin{cases} C_{base}, & |Q_b^{-1}Q_b(C_{base}) - C_{enh}| < |C_{enh}| \text{ の場合} \\ 0, & \text{それ以外の場合} \end{cases} \quad 20$$

および、

【数 9】

$$C'_{enh} = C_{enh} - Q_b^{-1}(Q_b(C'_{base})) \quad 30$$

ここにおいて、 Q_b^{-1} は前記第 1 のデータレイヤの逆量子化パラメータであり、第 1 の量子化パラメータは、前記第 2 の量子化パラメータよりも大きくセットされ、符号化された第 2 のデータレイヤが使用可能でない場合に、表示可能なマルチメディアデータを形成するために、符号化された前記第 1 データレイヤが復号化されることができ、符号化された前記第 1 のデータレイヤと符号化された前記第 2 のデータレイヤとの両方が使用可能である場合に、表示可能なマルチメディアデータを形成するために符号化された前記第 1 データレイヤおよび符号化された前記第 2 データレイヤが組み合わされて復号化されることができる。

【請求項 28】

請求項 27 に記載のコンピュータ可読媒体、ここにおいて、

符号化する前記命令が、下記を行う命令を備える：

前記マルチメディアデータの前記コンテンツ分類に基づいて前記マルチメディアデータの少なくとも一部のビットレートを割り当てること。

【請求項 29】

下記を行う構成を備えるプロセッサ：

マルチメディアデータを受け取ること；

前記マルチメディアデータのコンテンツを分類すること、ここにおいて、前記分類することは、下記を備える：

前記マルチメディアデータの空間情報および時間情報を判定すること、

前記マルチメディアデータの空間的な複雑度および時間的な複雑度を判定すること、および

前記マルチメディアデータの前記判定された空間情報および時間情報と前記判定された空間的な複雑度および時間的な複雑度に基づいて前記マルチメディアデータの前記コンテンツを分類すること、および、

前記マルチメディアデータの前記コンテンツ分類に基づいて前記マルチメディアデータを符号化された第1データレイヤおよび符号化された第2データレイヤに符号化すること、
 ここにおいて、前記第1データレイヤが、前記第2データレイヤと独立に復号化可能であるように構成され、前記第1データレイヤおよび前記第2データレイヤが、異なる品質レベルで符号化され、

10

ここにおいて、前記第1のデータレイヤは、

前記第1のデータレイヤの残差を残差係数 C_{base} に変換することであって、前記残差が符号化されるマルチメディアデータと時間基準フレームの最良一致マクロブロックとの差である、変換すること、

変換された前記残差係数 C_{base} から第1のデータレイヤ係数 C'_{base} を計算すること、および

第1の量子化パラメータ Q_b によって前記第1のデータレイヤ係数 C'_{base} を量子化すること

のステップによって符号化され、

ここにおいて、前記第2のデータレイヤは、

20

前記第2のデータレイヤの残差を残差係数 C_{enh} に変換することであって、前記残差が符号化されるマルチメディアデータと時間基準フレームの最良一致マクロブロックとの差である、変換すること、

変換された前記残差係数 C_{enh} から第2のデータレイヤ係数 C'_{enh} を計算すること、および

第2の量子化パラメータによって前記第2のデータレイヤ係数 C'_{enh} を量子化すること

のステップによって符号化され、

ここにおいて、

【数10】

30

$$C'_{base} = \begin{cases} 0, & \text{if } C_{base} \text{ および } C_{enh} \text{ の符号が反対である場合} \\ \min(C_{base}, C_{enh}), & \text{それ以外場合} \end{cases}$$

$$C'_{enh} = C_{enh} - Q_b^{-1}(Q_b(C'_{base}))$$

または

【数11】

40

$$C'_{base} = \begin{cases} C_{base}, & |Q_b^{-1}Q_b(C_{base}) - C_{enh}| < |C_{enh}| \text{ の場合} \\ 0, & \text{それ以外の場合} \end{cases}$$

および、

【数 1 2】

$$C'_{enh} = C_{enh} - Q_b^{-1}(Q_b(C'_{base}))$$

ここにおいて、 Q_b^{-1} は前記第 1 のデータレイヤの逆量子化パラメータであり、第 1 の量子化パラメータは、前記第 2 の量子化パラメータよりも大きくセットされ、符号化された第 2 のデータレイヤが使用可能でない場合に、表示可能なマルチメディアデータを形成するために、符号化された前記第 1 データレイヤが復号化されることができ、符号化された前記第 1 のデータレイヤと符号化された前記第 2 のデータレイヤとの両方が使用可能である場合に、表示可能なマルチメディアデータを形成するために符号化された前記第 1 データレイヤおよび符号化された前記第 2 データレイヤが組み合わされて復号化されることができる。

10

【発明の詳細な説明】

【技術分野】

【0001】

米国特許法第 119 条の下での優先権の主張

本願は、(a) 2005 年 9 月 27 日出願の米国仮出願第 60/721,416 号、名称「A VIDEO TRANSCODER FOR REAL-TIME STREAMING AND MOBILE BROADCAST APPLICATIONS」、
(b) 2006 年 4 月 4 日出願の米国仮出願第 60/789,377 号、名称「A VIDEO TRANSCODER FOR REAL-TIME STREAMING AND MOBILE BROADCAST APPLICATIONS」、
(c) 2005 年 10 月 17 日出願の米国仮出願第 60/727,643 号、名称「METHOD AND APPARATUS FOR SPATIO-TEMPORAL DEINTERLACING AIDED BY MOTION COMPENSATION FOR FIELD-BASED VIDEO」、
(d) 2005 年 10 月 17 日出願の米国仮出願第 60/727,644 号、名称「METHOD AND APPARATUS FOR SHOT DETECTION IN VIDEO STREAMING」、
(e) 2005 年 10 月 17 日出願の米国仮出願第 60/727,640 号、名称「A METHOD AND APPARATUS FOR USING AN ADAPTIVE GOP STRUCTURE IN VIDEO STREAMING」、
(f) 2005 年 10 月 24 日出願の米国仮出願第 60/730,145 号、名称「INVERSE TELECINE ALGORITHM BASED ON STATE MACHINE」、および
(g) 2006 年 4 月 3 日出願の米国仮出願第 60/789,048 号、名称「SPATIO-TEMPORAL DEINTERLACING AIDED BY MOTION COMPENSATION FOR FIELD-BASED MULTIMEDIA DATA」の優先権を主張するものである。この 7 つの米国仮出願のすべてが、本願の譲受人に譲渡され、参照によって本明細書に明示的に組み込まれている。

20

30

【0002】

同時係属特許出願の参照

本願は、本願の譲受人に譲渡され、参照によって本明細書に明白に組み込まれている、2006 年 3 月 10 日出願の米国仮出願第 11/373,577 号、名称「CONTENT CLASSIFICATION FOR MULTIMEDIA PROCESSING」に関連する。

40

【0003】

本願は、リアルタイムストリーミング用のビデオデータのビデオトランスコーディングの装置および方法を対象とし、より具体的には、移動体ブロードキャストアプリケーションでのリアルタイムストリーミング用のビデオデータのトランスコーディングを対象とする。

50

【背景技術】

【0004】

効率的なビデオ圧縮は、制限された帯域幅リソースおよび使用可能帯域幅の変動性に起因して、無線ビデオストリーミングおよびビデオテレフォニなどの多数のマルチメディアアプリケーションで有用である。MPEG-4 (ISO/IEC)、H.264 (国際電気通信連合)、または類似するビデオ符号化など、ある種のビデオ符号化標準規格は、無線ブロードキャストなどの応用によく適する高効率符号化を提供する。一部のマルチメディアデータ、たとえばディジタルテレビジョンプレゼンテーションは、一般に、MPEG-2などの他の標準規格に従って符号化される。したがって、トランスコーダが、無線ブロードキャストの前に、ある標準規格（たとえば、MPEG-2）に従って符号化されたマルチメディアデータを別の標準規格（たとえば、H.264）にトランスコーディングまたは変換するのに使用される。

10

【0005】

レートが最適化されたコーデックの改善は、誤り弾力性 (error resiliency)、誤り回復、およびスケーラビリティにおける利益を提供することができる。さらに、マルチメディアデータ自体から判定される情報の使用は、誤り弾力性、誤り回復、およびスケーラビリティを含む符号化に関する追加の改善をも提供することができる。したがって、マルチメディアデータ自体から決定される情報を使用するマルチメディアデータの高効率の処理および圧縮を提供し、スケーラブルであり、ストリーミングマルチメディア情報の移動体ブロードキャストを含む多数のマルチメディアデータ応用での使用に関して誤りに柔軟であるトランスコーダの必要が存在する。

20

【特許文献1】米国仮出願第60/721,416号

【特許文献2】米国仮出願第60/789,377号

【特許文献3】米国仮出願第60/727,643号

【特許文献4】米国仮出願第60/727,644号

【特許文献5】米国仮出願第60/727,640号

【特許文献6】米国仮出願第60/730,145号

【特許文献7】米国仮出願第60/789,048号

【特許文献8】米国仮出願第11/373,577号

【特許文献9】米国特許出願[弁理士整理番号第050078号]、名称「SCALABLE VIDEO CODING WITH TWO LAYER ENCODING AND SINGLE LAYER DECODING」

30

【非特許文献1】「Ideal spatial adaptation by wavelet shrinkage」、D.L. DonohoおよびI.M. Johnstone、Biometrika、vol. 8、425~455頁、1994年

【非特許文献2】S.P. Ghael、A.M. Sayeed、およびR.G. Baraniuk、「Improvement Wavelet denoising via empirical Wiener filtering」、Proceedings of SPIE、vol. 3169、389~399頁、San Diego、1997年7月

40

【非特許文献3】RP 27.3-1989 (Specifications for Safe Action and Safe Title Areas Test Pattern for Television Systems)

【発明の開示】

【発明の概要】

【0006】

説明され図示される発明的なコンテンツベースのトランスコーディングの装置および方法のそれぞれは、複数の態様を有し、どの単一の態様も、その望ましい属性の責任を単独で負うものではない。本開示の範囲を限定せずに、そのより顕著な特徴を、これから短く説明する。この議論を考慮した後に、特に「詳細な説明」と題するセクションを読んだ後

50

に、このコンテンツ駆動トランスコーディングの特徴が、マルチメディアデータ処理の装置および方法に関する改善をどのように提供するかが理解されるであろう。

【 0 0 0 7 】

本明細書で説明する発明的態様は、マルチメディアデータを符号化するさまざまな方法のために、および符号化器、たとえばトランスコード内で使用される符号化器のさまざまなモジュールまたは構成要素内で、コンテンツ情報を使用することに関する。トランスコードは、コンテンツ情報を使用し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。コンテンツ情報は、別のソース、たとえばビデオと共に受け取られるメタデータから受け取ることができる。トランスコードを、さまざまな異なる処理動作を介し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。いくつかの態様で、トランスコードは、マルチメディアデータのコンテンツ分類を生成し、このコンテンツ分類が、その後、1つまたは複数の符号化処理に使用される。いくつかの態様で、コンテンツ駆動トランスコードは、マルチメディアデータの空間的および時間的なコンテンツ情報を決定し、チャンネルにまたがるコンテンツを意識した均一な品質の符号化およびコンテンツ分類ベースの圧縮/ビット割当にそのコンテンツ情報を使用することができる。

10

【 0 0 0 8 】

いくつかの態様において、マルチメディアデータのコンテンツ情報（たとえば、メタデータ、コンテンツメトリックス、および/またはコンテンツ分類）が、取得されまたは計算され、その後、符号化に関するマルチメディアデータの処理に使用するためにトランスコードの構成要素に供給される。たとえば、プリプロセッサが、シーン変化検出、逆テレシネ（「I V T C」）の実行、デインターレーシング、動き補償および雑音抑圧（たとえば、2 D ウェーブレット変換）、ならびに時空間雑音低減、たとえば人為的な結果の除去、デリング（de-ringing）、デブロッキング、および/または雑音除去に、ある種のコンテンツ情報を使用することができる。いくつかの態様で、プリプロセッサは、空間分解能ダウンスampling、たとえば、標準品位（SD）からQuarter Video Graphics Array（QVGA）にダウンスamplingする時の適当な「セーフ」エリア（safe area）および「アクションハンドリング」エリア（action handling area）の判定に、コンテンツ情報を使用することもできる。

20

【 0 0 0 9 】

いくつかの態様において、符号化器は、コンテンツ情報を計算するように構成されたコンテンツ分類モジュールを含む。符号化器は、動き推定、たとえばカラー動き推定（ME）の実行、動きベクトル（MV）予測の実行のためにMBごとに量子化パラメータ（QP）を決定する際のビットレート制御（たとえば、ビット割当）、ベースレイヤおよびエンハンスメントレイヤを提供する際のスケラビリティ、ならびにたとえば適応イントラリフレッシュ、境界整列処理、およびエンハンスメントレイヤでの冗長Iフレームデータの提供を含む予測階層方式および誤り弾力性方式に影響を与えるためにコンテンツ分類を使用することによる誤り弾力性にコンテンツ分類を使用することができる。いくつかの態様で、トランスコードは、チャンネルにまたがる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多重化装置と協調してコンテンツ分類を使用する。いくつかの態様で、符号化器は、高速なチャンネルの切替を可能にするために、符号化されたデータ内にIフレームが周期的に現れることを強制するのにコンテンツ分類情報を使用することができる。そのような実施態様は、誤り弾力性のために符号化されたデータ内で要求される可能性があるIブロックを利用することもでき、ランダムアクセスの切替および誤り弾力性（たとえば、コンテンツ分類に基づく）を、予測階層を介して効果的に組み合わせて、符号化効率を改善すると同時に誤りに対する頑健性を高めることができる。

30

40

【 0 0 1 0 】

一態様において、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を受け取ることと、マルチメディアデータのコンテンツに基づいてマルチメディアデータを第1データグループおよび第2データグループに符号化することであって、第1データグループが、第2データグループと独立に復号化可能であるように構成され、第1データグル

50

ープおよび第2データグループが、異なる品質レベルで符号化される、符号化することを含む。第1の態様の一態様で、第1データグループは、IフレームおよびPフレームを備え、第2データグループは、Iフレーム、Pフレーム、およびBフレームを備える。もう1つの態様において、第1データグループは、ベースレイヤを備え、第2データグループは、エンハンスメントレイヤを備える。第3の態様で、この方法は、マルチメディアデータのコンテンツを分類することをさらに備え、符号化は、コンテンツ分類に基づく。第4の態様において、符号化は、第1データグループを符号化するためのマルチメディアデータの第1量子化パラメータを決定することと、第2データグループを符号化するための第2量子化パラメータを決定することとを備え、第1量子化パラメータおよび第2量子化パラメータの決定は、コンテンツ分類に基づく。第5の態様において、符号化は、コンテンツ分類に基づいてマルチメディアデータの少なくとも一部のビットレートを割り当てることを備える。第6の態様において、符号化は、さらに、シーン変化を検出するのにコンテンツ分類を使用することと、検出されたシーン変化に基づいて第1データグループおよび第2データグループ内にIフレームを含めるかどうかを決定することとを含む。第7の態様において、符号化は、コンテンツ分類に基づいてマルチメディアデータを符号化するフレームレートを決定することを含む。第8の態様において、符号化は、コンテンツ分類に基づいてマルチメディアデータの動き推定を実行することを含む。第9の態様において、この方法は、第1データグループの符号化の第1フレームレートを決定することと、第2データグループの符号化の第2フレームレートを決定することとをも含み、第1フレームレートは、第2フレームレートより低い。第10の態様において、符号化は、コンテンツ分類に基づいてマルチメディアデータに対して誤り弾力性処理を実行することを含む。第11の態様において、符号化は、第2データグループが使用可能でない場合に表示可能なマルチメディアデータを形成するために第1データグループを復号化することができ、第1データグループと第2データグループとの両方が使用可能である場合に表示可能なマルチメディアデータを形成するために第1データグループおよび第2データグループを組み合わせて復号化することができるように、第1データグループおよび第2データグループを符号化することを含む。第12の態様において、第1量子化パラメータは、データを符号化するための第1ステップサイズを含み、第2量子化パラメータは、データを符号化するための第2ステップサイズを含み、第1ステップサイズは、第2ステップサイズより大きい。第13の態様において、この方法は、マルチメディアデータのコンテンツを分類することをさらに備え、符号化は、コンテンツ分類に基づき、符号化は、コンテンツ分類に基づいてマルチメディアデータ内の雑音を減らすことを備える。第14の態様において、雑音を減らすことは、人為的な結果の除去を実行することを含む。第15の態様において、雑音を減らすことは、デリンギングフィルタを用いてマルチメディアデータの少なくとも一部を処理することを備え、デリンギングフィルタの強さは、マルチメディアデータのコンテンツに基づく。第16の態様において、雑音を減らすことは、デブロッキングフィルタを用いてマルチメディアデータの少なくとも一部を処理することを含み、デブロッキングフィルタの強さは、マルチメディアデータのコンテンツに基づく。第17の態様において、雑音を減らすことは、マルチメディアデータの選択された周波数をフィルタリングすることを含む。第18の態様において、デリンギングフィルタの強さは、マルチメディアデータのコンテンツ分類に基づく。第19の態様において、デブロッキングフィルタの強さは、マルチメディアデータのコンテンツ分類に基づく。第20の態様において、符号化は、マルチメディアデータをダウンサンプリングすることを備える。最後に、第21の態様において、符号化は、マルチメディアデータの符号化のビットレートを決定するために、マルチメディアデータに品質レベルを関連付けることと、マルチメディアデータの品質レベルおよびコンテンツ情報を使用することとを含む。

【0011】

第2の態様において、マルチメディアデータを処理する装置は、マルチメディアデータを受け取り、マルチメディアデータのコンテンツに基づいてマルチメディアデータを第1データグループおよび第2データグループに符号化するように構成された符号化器であっ

10

20

30

40

50

て、第1データグループが、第2データグループと独立に復号化可能であるように構成され、第1データグループおよび第2データグループが、異なる品質レベルで符号化される、符号化器を含む。第1の態様の一態様において、符号化器は、マルチメディアデータのコンテンツ分類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を含み、符号化モジュールは、さらに、コンテンツ分類に基づいてマルチメディアデータを符号化するように構成される。第2の態様において、符号化器は、さらに、第1データグループを符号化するためのマルチメディアデータの第1量子化パラメータおよび第2データグループを符号化するための第2量子化パラメータを決定するように構成され、第1量子化パラメータおよび第2量子化パラメータの決定は、マルチメディアデータのコンテンツ分類に基づく。第3の態様において、符号化器は、コンテンツ分類に基づいてマルチメディアデータの動き推定を実行し、データ動き補償情報を生成するように構成された動き推定モジュールを備え、符号化モジュールは、さらに、マルチメディアデータを符号化するのに動き補償情報を使用するように構成される。第4の態様において、符号化器は、コンテンツ分類に基づいてマルチメディアデータの量子化パラメータを決定する量子化モジュールをも含み、符号化器は、さらに、マルチメディアデータを符号化するのに量子化パラメータを使用するように構成される。第5の態様において、符号化器は、さらに、コンテンツ分類に基づいてマルチメディアデータの少なくとも一部のビットレートを提供するように構成されたビット割当モジュールをも含む。第6の態様において、符号化器は、シーン変化を検出するように構成されたシーン変化検出モジュールをも含み、符号化モジュールは、さらに、検出されたシーン変化に基づいて、符号化されたマルチメディアデータにIフレームを含めるように構成される。第7の態様において、符号化器は、コンテンツ分類に基づいてマルチメディアデータのフレームレートを決定するように構成されたフレームレートモジュールをも含み、符号化モジュールは、フレームレートに基づいてマルチメディアデータを符号化する。第8の態様において、符号化器は、コンテンツ分類に基づいて第1データグループおよび第2データグループを符号化するようにも構成される。第9の態様において、符号化器は、さらに、コンテンツ分類に基づいてマルチメディアデータに対して誤り処理を実行するようにも構成される。

【0012】

第3の態様において、マルチメディアデータを処理する装置は、マルチメディアデータを受け取るための手段と、マルチメディアデータのコンテンツに基づいてマルチメディアデータを符号化された第1データグループおよび符号化された第2データグループに符号化するための手段であって、第1データグループが、第2データグループと独立に復号化可能であるように構成され、第1データグループおよび第2データグループが、異なる品質レベルで符号化される、符号化するための手段とを含む。第1の態様の一態様において、受け取るための手段は、符号化器を備える。もう1つの態様において、符号化するための手段は、符号化器を備える。第3の態様において、符号化するための手段は、マルチメディアデータのコンテンツ分類を決定するための手段を備え、符号化するための手段は、コンテンツ分類に基づいてマルチメディアデータを符号化する。第4の態様において、符号化するための手段は、符号化器を備えるトランスコーダを備える。

【0013】

第4の態様において、機械可読媒体は、実行時に、機械に、マルチメディアデータを受け取ることと、マルチメディアデータのコンテンツに基づいてマルチメディアデータを符号化された第1データグループおよび符号化された第2データグループに符号化することであって、第1データグループが、第2データグループと独立に復号化可能であるように構成され、第1データグループおよび第2データグループが、異なる品質レベルで符号化されるような、符号化することを行わせる命令を備える。第4の態様の一態様において、コンピュータ可読媒体は、マルチメディアデータのコンテンツを示すコンテンツ分類を生成する命令をさらに備え、マルチメディアデータの符号化された第1データグループおよび符号化された第2データグループへの符号化は、コンテンツ分類に基づいてマルチメディアデータを符号化することを含む。もう1つの態様において、符号化は、第1データ

グループを符号化するためのマルチメディアデータの第1量子化パラメータを決定することとおよび第2データグループを符号化するための第2量子化パラメータを決定することとを備え、第1量子化パラメータおよび第2量子化パラメータを決定することは、コンテンツ分類に基づく。第3の態様で、マルチメディアデータを符号化する命令は、マルチメディアデータのコンテンツに基づいてマルチメディアデータの少なくとも一部のビットレートを割り当てる命令を含む。

【0014】

第5の態様において、プロセッサは、マルチメディアデータを受け取り、マルチメディアデータのコンテンツに基づいてマルチメディアデータを符号化された第1データグループおよび符号化された第2データグループに符号化する構成であって、第1データグループが、第2データグループと独立に復号化可能であるように構成され、第1データグループおよび第2データグループが、異なる品質レベルで符号化される、構成を備える。第5の態様の一態様において、プロセッサは、さらに、マルチメディアデータのコンテンツを示すコンテンツ分類を生成する構成をさらに備え、符号化は、コンテンツ分類に基づいてマルチメディアデータを符号化することを備える。もう1つの態様において、プロセッサは、さらに、第1データグループを符号化するためのマルチメディアデータの第1量子化パラメータおよび第2データグループを符号化するための第2量子化パラメータを決定する構成を備え、第1量子化パラメータおよび第2量子化パラメータは、コンテンツ分類に基づく。

【0015】

適当な場合に、図面の複数の図を通じて同様の符号が同様の部分を指すことに留意されたい。

【詳細な説明】

【0016】

次の詳細な説明は、この開示で論じられるある種の態様を対象とする。しかし、本発明を、複数の異なる形で実施することができる。本明細書での「一態様」または「態様」への言及は、その態様に関連して説明される特定の特徴、構造、または特性が、少なくとも1つの態様に含まれることを意味する。本明細書のさまざまな場所での句「一態様において」、「一態様による」、または「いくつかの態様において」の出現は、必ずしもすべてが同一の態様に言及しているのではなく、他の態様と相互に排他的な別々の態様または代替の態様でもない。さらに、一部の態様によって示されるが他の態様によっては示されない場合があるさまざまな特徴を説明する。同様に、一部の態様の要件であるが他の態様の要件ではない場合があるさまざまな要件を説明する。

【0017】

次の説明は、例の完全な理解を提供するために詳細を含む。しかし、当業者は、例または態様に含まれる処理またはデバイスのすべての詳細が本明細書で説明されないか図示されない場合であっても、これらの例を実践できることを理解する。たとえば、電気的な構成要素が、不必要な詳細において例を不明瞭にしないようにするために、その構成要素のすべての電気的な接続またはすべての電気的な要素を示してはいないブロック図で示される場合がある。他の場合に、そのような構成要素、他の構造、および技法が、例をさらに説明するために詳細に示される場合がある。

【0018】

本開示は、符号化されるマルチメディアデータのコンテンツ情報を使用して符号化およびトランスコーディングの装置および方法を制御することに関する。「コンテンツ情報」または「コンテンツ」(マルチメディアデータの)は、マルチメディアデータのコンテンツに関する情報を意味する広義の用語であり、たとえば、メタデータ、マルチメディアデータから計算されるメトリックス、およびたとえばコンテンツ分類など、1つまたは複数のメトリックスに関連するコンテンツ関連情報を含めることができる。コンテンツ情報は、特定の応用例に応じて、符号化器に供給するか、符号化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、時空間雑音低減、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、スケーラビリティ、誤り弾力性、ブロードキャストチャンネルにまたがる最適なマルチメディア品質の維持、および高速なチャンネルの切替を含む、マルチメディアデータ符号化の多数の態様に使用することができる。これらの態様のうちの1つまたは複数を使用して、トランスコーダは、マルチメディアデータの処理を調整し、コンテンツに関連して符号化されたマルチメディアデータを作ることができる。トランスコーディングの態様を説明する本明細書の説明および図面は、符号化の態様および復号化の態様に適用可能とすることもできる。

【0019】

トランスコーダの装置および方法は、あるフォーマットから別のフォーマットへのトランスコーディングに関し、本明細書では、いくつかの態様の事例となる移動体デバイスへの無線チャンネルを介する伝送のための、質を高められたスケーラブルなH.264フォーマットへのMPEG-2ビデオのトランスコーディングに関して具体的に説明される。しかし、H.264フォーマットへのMPEG-2ビデオのトランスコーディングの説明は、本発明の範囲を限定するものとして意図されたものではなく、本発明のいくつかの態様の例示にすぎない。開示される装置および方法は、ランダムアクセス機能およびレイヤリング機能を有する誤りに対して弾力的な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、MPEG-2およびH.264以外のビデオフォーマットのトランスコーディングおよび/または符号化にも適用可能とすることができる。

【0020】

本明細書で使用される「マルチメディアデータ」または単に「マルチメディア」は、ビデオデータ（オーディオデータをも含むことができる）、オーディオデータ、またはビデオデータとオーディオデータとの両方を含む広義の用語である。本明細書で使用される「ビデオデータ」または「ビデオ」は、テキスト、画像情報、および/またはオーディオデータを含む1つまたは複数の画像または画像の関連するシーケンスを含むフレームベースのデータまたはフィールドベースのデータを指す広義の用語であり、そうではないと指定されない限り、マルチメディアデータを指すのに使用することができる（たとえば、これらの用語を交換可能に使用することができる）。

【0021】

下の説明は、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使用することができる、トランスコーダのさまざまな構成要素の例および処理の例である。

【0022】

マルチメディアブロードキャストシステム

図1Aは、マルチメディアデータブロードキャストシステム100のいくつかの態様のデータの流れを示すブロック図である。システム100内で、マルチメディアデータ供給器106が、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によって受け取られ、トランスコーダ200は、ブロック110内でマルチメディアデータ104を生元のマルチメディアデータへと処理する。ブロック110での処理は、符号化されたマルチメディアデータ104を復号化し、解析し、さらに、別のフォーマットに符号化するためにマルチメディアデータを準備するためにマルチメディアデータを処理する。復号化されたマルチメディアデータは、ブロック112に供給され、ここで、マルチメディアデータが、所定のマルチメディアフォーマットまたはマルチメディア標準規格に符号化される。マルチメディアデータが一旦符号化されたならば、ブロック114で、そのマルチメディアデータを、たとえば無線ブロードキャストシステム（たとえば、セルラ電話ブロードキャスト網または別の通信網を介して）を介した伝送のために準備する。いくつかの態様において、受け取られたマルチメディアデータ104は、MPEG-2標準規格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が復号化された後に、トランスコーダ200は、そのマルチメディアデータをH.264標準規格に符号化する。

【0023】

図1Bは、図1Aのブロック110および112での処理を実行するように構成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を、マルチメディアデータを受け取り、そのマルチメディアデータを復号化し、パケット化された基本のストリーム（たとえば、サブタイトル、オーディオ、メタデータ、「生」ビデオ、CCデータ、およびプレゼンテーションタイムスタンプ）に解析し、これを所望のフォーマットに符号化し、さらなる処理または伝送のために符号化されたデータを供給するように構成することができる。データを複数のデータグループ、たとえば符号化された第1データグループおよび符号化された第2データグループに供給するように、トランスコーダ130を構成することができ、これを、階層化された符号化（layered encoding）と称する。複数の態様のいくつかの例で、階層化された符号化方式でのさまざまなデータグループ（または層）を、異なるレベルの品質で符号化し、第1データグループ内で符号化されるデータが第2データグループ内で符号化されるデータより低い品質を有する（たとえば、表示された時により低い視覚的品質レベルをもたらす）ようにフォーマットすることができる。

10

【0024】

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でき、図1Aのブロック110および112に示された処理の一部またはすべてを実行するように構成できるプロセッサ140のブロック図である。プロセッサ140は、復号化、解析、前処理、および符号化を含む本明細書で説明するトランスコーディング処理のうちの1つまたは複数を実行し、処理にコンテンツ情報を使用する、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をも含み、プロセッサ140を、「直接に」または「別のデバイスを介して間接に」のいずれかで外部メモリ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1つまたは複数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をも含み、それはマルチメディアデータを受け取るため、および第1データグループ内で符号化されたデータおよび第2データグループ内で符号化されたデータなどのような符号化されたデータを供給するためを含む。複数の態様のいくつかの例で、階層化された符号化方式でのさまざまなデータグループ（または層）を、異なるレベルの品質で符号化し、第1データグループ内で符号化されるデータが第2データグループ内で符号化されるデータより低い品質を有する（たとえば、表示された時により低い視覚的品質レベルをもたらす）ようにフォーマットすることができる。

20

30

【0025】

トランスコーダ130またはプリプロセッサ140（トランスコーディング用に構成された）の構成要素とそれに含まれる処理とを、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、またはその任意の組合せで実施することができる。たとえば、パーサ、復号化器、プリプロセッサ、または符号化器を、独立型の構成要素とするか、別のデバイスの構成要素にハードウェア、ファームウェア、ミドルウェアとして組み込むか、プロセッサ上で実行されるマイクロコードまたはソフトウェアで実施するか、これらの組合せとすることができる。ソフトウェア、ファームウェア、ミドルウェア、またはマイクロコードで実施される時に、動き補償処理、ショット分類処理、および符号化処理を実行するプログラムコードまたはコードセグメントを、記憶媒体などの機械可読媒体に格納することができる。コードセグメントは、プロシージャ、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、あるいは命令、データ構造、またはプログラムステートメントの任意の組合せを表すことができる。情報、データ、引数、パラメータ、またはメモリ内容を渡し、かつ/または受け取ることによって、一のコードセグメントを別のコードセグメントまたはハードウェア回路に結合することができる。

40

【0026】

トランスコーダアーキテクチャの例示的な例

図2に、図1のマルチメディアブロードキャストシステム100に示されたトランスコーダ200に使用できるトランスコーダの例のブロック図を示す。トランスコーダ200

50

は、下でさらに説明する、解析器／復号化器 202、プリプロセッサ 226、符号化器 228、および同期化層 240を備える。本明細書で説明するように、トランスコーディング処理の1つまたは複数の態様に関してマルチメディアデータ 104のコンテンツ情報を使用するように、トランスコード 200は構成される。コンテンツ情報は、トランスコード 200の外部ソースから、マルチメディアメタデータを介して取得可能であるか、またはトランスコードによって、たとえばプリプロセッサ 226または符号化器 228によって計算されることができる。図2に示された構成要素は、1つまたは複数のトランスコーディング処理にコンテンツ情報を使用する、トランスコードに含めることができる構成要素の例示である。特定の実装において、トランスコード 200の構成要素のうちの1つまたは複数の除外することができ、あるいは、追加の構成要素を含めることができる。さらに、トランスコードおよびトランスコーディング処理の諸部分は、処理またはデバイスのあらゆる詳細が本明細書で説明されない場合であっても当業者が本発明を実施することを可能にするために説明される。

10

【0027】

図5に、トランスコード 200のさまざまな構成要素および／または処理の動作の時間的関係を示すグラフとしてタイミング図を示す。図5に示されているように、MPEG-2ビデオなどの符号化されたストリーミングビデオ 104（符号化されたマルチメディアデータ）が、まず、解析器 205（図2）によって任意の時刻0に受け取られる。次に、ビデオストリームが、例えば復号化器 214と組み合わせられた解析器 205により、解析され501、逆多重化され502、復号化される503。図示されているように、これらの処理は、データ処理のストリーム出力をプリプロセッサ 226（図2）に供給するために、わずかなタイミングオフセットを伴って並列に発生することができる。時刻 T_1 504に、プリプロセッサ 226が、処理結果の出力を開始するのに十分なデータを復号化器 214から一旦受け取ったならば、残りの処理ステップは、性質において逐次的になり、第1パス符号化 505、第2パス符号化 506、および再符号化 507は、時刻 T_f 508の再符号化の完了まで、前処理の後に順番に発生する。

20

【0028】

本明細書で説明するトランスコード 200を、さまざまなマルチメディアデータをトランスコードするように構成することができ、処理の多くが、トランスコーディングされるマルチメディアデータのどのタイプにもあてはまる。本明細書で提供する例の一部は、特にH.264データへのMPEG-2データのトランスコーディングに関するが、これらの例は、本開示をそのようなデータに限定することを意味するものではない。下で説明する符号化の諸態様は、任意の適切なマルチメディアデータ標準規格の別の適切なマルチメディアデータ標準規格へのトランスコーディングに適用することができる。

30

【0029】

解析器／復号化器

もう一度図2を参照すると、解析器／復号化器 202は、マルチメディアデータ 104を受け取る。解析器／復号化器 202は、マルチメディアデータ 104を受け取り、そのデータをビデオの基本のストリーム（ES）206、オーディオES 208、プレゼンテーションタイムスタンプ（PTS）210、およびサブタイトル 212などの他のデータに解析するトランスポートストリーム解析器（「解析器」）205を含む。ESは、単一のビデオ符号化器またはオーディオ符号化器からの1つのタイプのデータ（ビデオまたはオーディオ）を運ぶ。たとえば、ビデオESは、シーケンスヘッダおよびそのシーケンスのすべてのサブパートを含む、データの1シーケンスに関するビデオデータを備える。パケット化された基本のストリームすなわち PESは、それぞれが典型的には追加されたパケットヘッダから始まる複数のパケットにされた単一のESからなる。PESストリームは、1つのソースから、たとえば1つのビデオ符号化器またはオーディオ符号化器からの1タイプのみのデータを含む。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時に、PESヘッ

40

50

ダを、トランスポートパケットペイロードの先頭でトランスポートパケットヘッダの直後に置くことができる。残りのPESパケットコンテンツは、PESパケットがすべて使用されるまで、連続するトランスポートパケットのペイロードを充てんする。最終的なトランスポートパケットは、固定長まで、たとえば、バイト、たとえばバイト = $0 \times FF$ (すべて1) を詰め込むことによって、充てんすることができる。

【0030】

解析器205は、ここで示す解析器/復号化器202の一部である復号化器214にビデオES 206を通信する。他の構成では、解析器205および復号化器214が、別々の構成要素である。PTS 210が、トランスコーダPTS生成器215に送られ、トランスコーダPTS生成器215は、トランスコーダ200からブロードキャストシステムに送られるデータを配列するのに使用するために、トランスコーダ200に特有の別々のプレゼンテーション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を、データブロードキャストの同期を調整するためにトランスコーダ200の同期化層240にデータを供給するように構成することができる。

10

【0031】

図3に、上で説明したさまざまなパケット化された基本のストリームを解析する時に解析器205が従うことのできる処理300の1つの例の流れ図を示す。処理300は、マルチメディアデータ104がコンテンツ供給器106(図1)から受け取られる時に、ブロック302で開始される。処理300は、ブロック304に進み、ここで、解析器205の初期化が実行される。初期化は、独立に生成される獲得コマンド306によってトリガすることができる。たとえば、解析器205とは独立であり、外部で受け取られるTVスケジュール情報およびチャンネルラインアップ情報に基づく処理が、獲得コマンド306を生成することができる。さらに、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、初期化処理と主処理との両方で支援するために入力することができる。

20

【0032】

ブロック304に示されているように、初期化は、コマンド構文検証の獲得、第1パスPSI/PSIP/SI(プログラム固有情報/プログラムおよびシステム情報プロトコル/システム情報)処理の実行、獲得コマンドまたはPSI/PSIP/SI整合性検証のいずれかに特に関連する処理の実行、PESごとのPESバッファの割当、およびタイミングをセットすること(たとえば、所望の獲得開始の瞬間と整列させるため)を含むことができる。PESバッファは、解析されたESデータを保持し、各解析されたESデータに対応するオーディオ復号化器216、テスト符号化器220、復号化器214、またはトランスコーダPTS生成器215に通信する。

30

【0033】

初期化の後に、処理300は、受け取られたマルチメディアデータ104の主処理のブロック310に進む。ブロック310での処理は、ターゲットパケット識別子(PID)フィルタリング、連続的なPSI/PSIP/SIの監視および処理、ならびに、入ってくるマルチメディアデータが適切なPESバッファに渡されるようにするためのタイミング処理(たとえば、所望の獲得持続時間を達成するための)を含むことができる。ブロック310でのマルチメディアデータの処理の結果として、PESバッファ「読み取り」のプログラム記述子および表示が生成され、これは、下で説明するように、復号化器214(図2)とインターフェースする。

40

【0034】

ブロック310の後に、処理300はブロック314に進み、ここで、タイマ割込みの生成およびPSEバッファの消費の結果としてのPSEバッファの解放を含めて、解析動作の終了が発生する。PESバッファが、オーディオストリーム、ビデオストリーム、およびサブタイトルストリームなど、記述子に列挙されたプログラムのすべての関連する基本のストリームについて存在することに留意されたい。

【0035】

50

もう一度図2を参照すると、パーサ205は、トランスコーダの実装に対応するためにオーディオ復号化器216にオーディオES 208を送り、同期化層240への符号化されたテキスト216およびオーディオ情報の復号化を供給する。サブタイトル情報212は、テキスト符号化器220に配送される。復号化器214からのクローズドキャプション(CC)データ218も、テキスト符号化器220に供給され、テキスト符号化器220は、トランスコーダ200によって作用されるフォーマットでサブタイトル情報212およびCCデータ218を符号化する。

【0036】

解析器/復号化器202は、復号化器214をも含み、復号化器214は、ビデオES 206を受け取る。復号化器214は、ビデオデータに関連するメタデータを生成し、符号化されたビデオパケット化された基本のストリームを生じるビデオ224(たとえば、標準品位フォーマットの)に復号化し、ビデオクローズドキャプション付きデータをビデオESストリームへと処理することができる。

【0037】

図4に、復号化器214によって実行できる復号化処理400の1つの例を示す流れ図を示す。処理400は、ブロック402でのビデオの基本のストリームのデータ206の入力から始まる。処理400は、ブロック404に進み、ここで、復号化器が初期化される。初期化は、ビデオシーケンスヘッダ(VSH)の検出、第1パスVSH、ビデオシーケンス(VS)、およびVSディスプレイエクステンション(Display Extension)処理(ビデオフォーマット、原色、および行列係数を含む)の実行、ならびに復号化されたピクチャ、関連するメタデータ、およびクローズドキャプション(CC)データをそれぞれバッファリングするためのデータバッファの割当を含む複数のタスクを含むことができる。さらに、解析器205によって供給されるビデオPESバッファ「読み取り」情報406(たとえば、図3のブロック310で処理300によって生成することができる)が入力される。

【0038】

ブロック404での初期化の後に、処理400は、ブロック408に進み、ここで、ビデオESの主処理が、復号化器214によって実行される。主処理は、新しいデータの可用性に関するビデオPSEバッファ「読み取り」情報または「インターフェース」のポーリングと、ビデオESの復号化と、同期しているピクチャ境界での画素データの再構成および格納と、ビデオおよびa/vがピクチャ境界でメタデータを生成し格納することと、ピクチャ境界でのCCデータ格納とを含む。主処理408の結果ブロック410は、シーケンス記述子、復号化済みピクチャバッファ記述子、メタデータバッファ記述子、およびCCデータバッファ記述子の生成を含む。

【0039】

主処理408の後に、処理400は、ブロック412に進み、ここで、終了処理を実行する。終了処理は、所定のしきい値を超える特定の持続時間の間に新しいデータが発生しないこと、シーケンス終了コードの検出、および/または明示的終了信号の検出を含む終了条件の判定を含むことができる。終了処理は、さらに、下で説明する、復号化済みピクチャバッファ、関連するメタデータバッファ、およびCCデータバッファのプリプロセッサによる消費の結果としてこれらのバッファを解放することを含むことができる。処理400は、ブロック414で終了し、ここで、処理400は、ビデオESが入力として受け取られるのを待つ状態に入ることができる。

【0040】

プリプロセッサ

図2およびより詳細には図6に、1つまたは複数の前処理動作にコンテンツ情報を使用できるプリプロセッサ226のサンプル態様を示す。プリプロセッサ226は、パーサ/復号化器202からメタデータ222および復号化された「生」ビデオデータ224を受け取る。プリプロセッサ226は、ビデオデータ224およびメタデータ222に対してあるタイプの処理を実行し、処理されたマルチメディア(たとえば、ベースレイヤ基準フ

10

20

30

40

50

レーム、エンハンスメントレイヤ基準フレーム、帯域幅情報、コンテンツ情報）およびビデオを符号化器228に供給するように構成される。マルチメディアデータのそのような前処理は、ビデオの明瞭さ、アンチエイリアシング、およびデータの圧縮効率を改善することができる。一般に、プリプロセッサ226は、解析器復号化器202内の復号化器214によって供給されるビデオシーケンスを受け取り、符号化器228によるさらなる処理（たとえば、符号化）のためにそのビデオシーケンスをプログレッシブビデオシーケンスに変換する。いくつかの態様で、プリプロセッサ226を、逆テレシネ、デインターレーシング、フィルタリング（たとえば、人為的な結果の除去、デリンギング、デブロッキング、および雑音除去）、サイズ変更（たとえば、標準品位からQuarter Video Graphics Array（QVGA）への空間分解能ダウンサンプリング）、およびGOP構造生成（たとえば、計算複雑度マップ生成、シーン変化検出、およびフェード/フラッシュ検出）を含む多数の動作のために構成することができる。

10

【0041】

プリプロセッサ226は、前処理動作のうちの1つまたは複数に影響を与えるために復号化器からのメタデータを使用することができる。メタデータには、マルチメディアデータのコンテンツの記述または分類に関する情報（「コンテンツ情報」）を含めることができ、具体的には、メタデータにコンテンツ分類を含めることができる。いくつかの態様において、メタデータは、符号化動作に望まれるコンテンツ情報を含めない。そのような場合に、コンテンツ情報を決定し、そのコンテンツ情報を前処理動作に使用し、かつ/またはトランスコード200の他の構成要素、たとえば復号化器228にコンテンツ情報を供給するようにプリプロセッサ226を構成することができる。いくつかの態様において、プリプロセッサ226は、GOPパーティショニングに影響し、フィルタリングの適当なタイプを決定し、かつ/または符号化器に通信される符号化パラメータを決定するためにそのようなコンテンツ情報を使用することができる。

20

【0042】

図6に、プリプロセッサ226に含めることができるさまざまな処理ブロックの例示的な例を示し、プリプロセッサ226によって実行できる処理を示す。この例では、プリプロセッサ226は、メタデータおよびビデオ222、224を受け取り、（処理された）メタデータおよびビデオを備える出力データ614を符号化器228に供給する。通常、受け取られる可能性がある3タイプのビデオがある。第1に、受け取られるビデオが、プログレッシブビデオである可能性があり、この場合にはデインターレーシングが不要である場合がある。第2に、ビデオデータが、テレシネされたビデオすなわち24fpsムービーシーケンスから変換されたインターレース式ビデオである可能性があり、この場合には、逆テレシネ動作が必要になる場合がある。第3に、ビデオが、テレシネされていないインターレース式ビデオである可能性がある。プリプロセッサ226は、下で説明するようにこれらのタイプのビデオを処理することができる。

30

【0043】

ブロック601で、プリプロセッサ226は、受け取ったビデオデータ222、224がプログレッシブビデオであるかどうかを判定する。いくつかの場合に、これは、メタデータがそのような情報を含む場合にメタデータから判定することができ、あるいは、ビデオデータ自体の処理によって判定することができる。たとえば、下で説明する逆テレシネ処理は、受け取られたビデオ222がプログレッシブビデオであるかどうかを判定することができる。そうである場合には、この処理はブロック607に進み、ここで、フィルタリング（たとえば、デノイザ（denoiser））動作が、白色ガウス雑音などの雑音を減らすためにビデオに対して実行される。ビデオデータ222、224がプログレッシブビデオではない場合には、ブロック601で、この処理は、ブロック604へ、フェーズ検出器604へ進む。

40

【0044】

フェーズ検出器604は、テレシネで発したビデオと標準ブロードキャストフォーマットで始まったビデオとを区別する。ビデオがテレシネされたとの判断が行われる場合に（

50

フェーズ検出器 604 から出る YES 判断経路)、テレシネされたビデオが、逆テレシネ 606 でそのオリジナルフォーマットに戻される。冗長なフレームが、識別され、除去され、同一ビデオフレームから導出されたフィールドが、完全な画像にもう一度織り込まれる。再構成されたフィルム画像のシーケンスは、1 秒の 1/24 の規則的な間隔で写真のように記録されたので、GOP パーティショナ (partitioner) 612 または復号化器 228 内で実行される動き推定処理は、不規則なタイムベースを有す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使用することによって、より正確である。

【0045】

一態様で、フェーズ検出器 604 は、ビデオフレームの受け取りの後にある種の判断を行う。これらの判断は、(i) テレシネ出力および 3:2 プルダウン・フェーズからの現行ビデオが、図 38 に示された 5 つのフェーズ P_0 、 P_1 、 P_2 、 P_3 、および P_4 のうちの 1 つであるかどうか、および (ii) ビデオが従来の NTSC として生成されたかどうかを含む。その判断を、フェーズ P_5 と表す。これらの判断は、図 2 に示されたフェーズ検出器 604 の出力として現れる。フェーズ検出器 604 からの「YES」と示された経路は、逆テレシネ 606 を作動させ、正しいプルダウン・フェーズが与えられたことを示し、その結果、逆テレシネ 606 が、同一の写真画像から形成されたフィールドを選別でき、これらを組み合わせられるようにする。フェーズ検出器 604 からの「NO」と示された経路は、同様に、最適処理のために明白な NTSC フレームをフィールドに分離するためにデインターレース 605 を作動させる。フェーズ検出器 604 は、異なるタイプのビデオが任意の時に受け取られる可能性があるので、ビデオフレームを連続的に分析することができる。例示として、NTSC 標準規格に準拠するビデオが、コマーシャルとしてビデオに挿入される場合がある。逆テレシネの後に、結果のプロGRESSIVE ビデオが、デノイザ (フィルタ) 607 に送られ、デノイザ 607 は、白色ガウス雑音を減らすのに使用することができる。

【0046】

通常の NTSC ビデオが認識される場合に (フェーズ検出器 601 からの NO 経路)、そのビデオは、圧縮のためにデインターレース 605 に送られる。デインターレース 605 は、インターレース式フィールドを PROGRESSIVE ビデオに変換し、その後、雑音除去動作が、PROGRESSIVE ビデオに対して実行され得る。デインターレース処理の 1 つの例示的な例を、下で説明する。

【0047】

テレビジョンなどの伝統的なアナログビデオデバイスは、インターレースされた形でビデオをレンダリングする、すなわち、そのようなデバイスは、偶数番号の走査線 (偶数フィールド) を送り、次に奇数番号の走査線 (奇数フィールド) を送る。信号サンプリングの観点から、これは、次式によって記述されるパターンでの時空間サブサンプリングと同等である。

【数 1】

$$F(x, y, n) = \begin{cases} \Theta(x, y, n), & y \bmod 2 = 0 \text{ (偶数フィールド) の場合、} \\ \Theta(x, y, n), & y \bmod 2 = 1 \text{ (奇数フィールド) の場合、} \\ \text{消去} & \text{それ以外の場合} \end{cases} \quad [1]$$

【0048】

ここで、 Θ は、オリジナルフレームピクチャを表し、 F は、インターレース式フィールドを表し、 (x, y, n) は、それぞれ画素の水平位置、垂直位置、および時間位置を表す。

【0049】

一般性を失わずに、本開示全体を通じて $n = 0$ が偶数フィールドであると仮定することができる、その結果、上の式 1 は、次のように単純化される。

【数 2】

$$F(x, y, n) = \begin{cases} \Theta(x, y, n), & y \bmod 2 = n \bmod 2 \text{ の場合、} \\ \text{消去} & \text{それ以外の場合} \end{cases} \quad [2]$$

【0050】

デシメーションは、水平次元では行われないので、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。

【0051】

デインターレースの目的は、インターレース式ビデオ（フィールドのシーケンス）を非
インターレース式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。言い換えると、フルフレームピクチャを「回復し」または生成するために、偶数フィールドおよび奇数フィールドを補間する。これは、式 3 によって表すことができる。

10

【数 3】

$$F_o(x, y, n) = \begin{cases} F(x, y, n), & y \bmod 2 = n \bmod 2, \\ F_i(x, y, n), & \text{それ以外の場合} \end{cases} \quad [3]$$

【0052】

ここで、 F_i は、欠けている画素のデインターレーシング結果を表す。

20

【0053】

図 40 は、インターレース式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のに Wmed フィルタリングおよび動き推定を使用するデインターレース 605 の一態様の特定の側面を示すブロック図である。図 40 の上側部分は、現行フィールド、2つの前のフィールド（前の前のフィールドおよび前のフィールド）、および2つの後のフィールド（次のフィールドおよび次の次のフィールド）からの情報を使用して生成できる動き強度マップ（motion intensity map）4002 を示す。動き強度マップ 4002 は、現行フレームを複数の異なる動きレベルに分類し、またはパーティショニングするが、下でさらに詳細に説明する時空間フィルタリングによって生成することができる。いくつかの態様において、動き強度マップ 4002 は、下で式 4 ~ 8 を参照して説明するように、静的な区域、遅い動きの区域、および速い動きの区域を識別するために生成される。時空間フィルタ、たとえば Wmed フィルタ 4004 は、動き強度マップに基づく判断基準を使用して、インターレース式マルチメディアデータをフィルタリングし、時空間的な仮のデインターレースされたフレームを作る。いくつかの態様において、Wmed フィルタリング処理は、 $[-1, 1]$ の水平近傍、 $[-3, 3]$ の垂直近傍、および図 40 に示された5つのフィールド（前の前のフィールド、前のフィールド、現行フィールド、次のフィールド、および次の次のフィールド）によって表される5つの隣接するフィールドの時間近傍を用い、 Z^{-1} は、1フィールドの遅延を表す。現行フィールドに対して相対的に、次のフィールドおよび前のフィールドは、非パリティフィールドであり、次の次のフィールドおよび前の前のフィールドは、パリティフィールドである。時空間フィルタリングに使用される「近傍」は、フィルタリング動作中に実際に使用されるフィールドおよび画素の空間的・時間的位置を指し、たとえば図 6 および 7 に示されているように、「アパーチャ」として図示することができる。

30

40

【0054】

デインターレース 605 は、Wmed フィルタ 4004 によって生成された時空間的な仮のデインターレース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デノイザ（雑音除去フィルタ）4006 をも含むことができる。時空間的な仮のデインターレースされたフレームを雑音除去することによって、特にソースのインターレース式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、後続の動き探索処理がより正確になる。これによって、Wmed ピクチャ内の偶数行と奇数行との間のエイリアス

50

を少なくとも部分的に除去することもできる。デノイザ 4006 は、ウェーブレット収縮 (wavelet shrinkage) およびウェーブレットウィーナーフィルタベースのデノイザを含むさまざまなフィルタとして実施することができる。デノイザは、候補 Wmed フレームが動き補償情報を使用してさらに処理される前にその候補となる Wmed フレームから雑音を除去するのに使用することができ、Wmed フレーム内に存在する雑音を除去し、信号の周波数の内容に関わりなく、存在する信号を保つことができる。ウェーブレットフィルタを含むさまざまなタイプの雑音除去フィルタを使用することができる。ウェーブレットは、空間領域とスケーリング領域との両方で所与の信号を局所化するのに使用される、ある種類の関数である。ウェーブレットの背後にある基本的な発想は、ウェーブレット表現での小さい変化が、対応してオリジナル信号の小さい変化を作るように、異なるスケールまたは異なる分解能で信号を分析することである。

10

【0055】

ウェーブレット収縮またはウェーブレットウィーナーフィルタを、デノイザとして適用することもできる。ウェーブレット収縮は、雑音の多い信号のウェーブレット変換と、それに続き、大きいウェーブレット係数が影響されないままにしながら小さいウェーブレット係数を 0 (またはより小さい値) に縮小することからなる。最後に、逆変換が、推定された信号を獲得するために実行される。

【0056】

雑音除去フィルタリングは、雑音の多い環境で動き補償の精度を高める。ウェーブレット収縮雑音除去は、ウェーブレット変換領域での縮小を用いることができ、通常は、3つのステップすなわち、線形順方向ウェーブレット変換、非線形収縮雑音除去、および線形逆ウェーブレット変換を備える。ウィーナーフィルタは、加法性雑音およびぼけによって劣化した画像を改善するのに使用できる MSE 最適線形フィルタである。そのようなフィルタは、当技術分野で一般に知られており、たとえば、上で参照した「Ideal spatial adaptation by wavelet shrinkage」、および参照によってその全体が特に本明細書に組み込まれている S. P. Ghael、A. M. Sayeed、および R. G. Baraniuk、「Improvement Wavelet denoising via empirical Wiener filtering」、Proceedings of SPIE、vol. 3169、389 ~ 399 頁、San Diego、1997 年 7 月に記載されている。

20

30

【0057】

いくつかの態様において、雑音除去フィルタは、(4, 2) 双直交 (bi-orthogonal) 3 次 B スプラインウェーブレットフィルタの一態様に基づく。1つのそのようなフィルタを、次の順方向変換および逆変換によって定義することができる。

【数 4】

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

【0058】

および

40

【数 5】

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

【0059】

雑音除去フィルタの適用は、雑音の多い環境での動き補償の精度を高めることができる。そのようなフィルタの実施態様が、さらに、参照によってその全体が特に本明細書に組み込まれている、「Ideal spatial adaptation by wavelet shrinkage」、D. L. Donoho および I. M. Johnstone、Biometrika、vol. 8、425 ~ 455 頁、1994 年に記載されて

50

いる。

【0060】

図40の下側部分は、インターレース式マルチメディアデータの動き情報（たとえば、動きベクトル候補、動き推定、動き補償）の決定の一態様を示す。具体的に言うと、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するのに使用され、その後、デインターレースされた現行フレーム4014として図示された結果の「最終的な」プログレッシブフレームを形成するためにWmed仮フレームと組み合わせられる、動き推定および動き補償方式を示す。いくつかの態様において、インターレース式マルチメディアデータの動きベクトル（「MV」）候補（または推定値）は、外部の動き推定器からデインターレースに供給され、双方向動き推定器および補償器（「ME/MC」）4018の開始点を与えるのに使用される。いくつかの態様で、MV候補選択器4022は、処理されているブロックのMV候補について、以前に処理されたブロック、たとえばデインターレースされた前フレーム4020内のブロックのMVなど、隣接ブロックの以前に決定されたMVを使用する。動き補償は、前のデインターレースされたフレーム70および次の（たとえば将来の）Wmedフレーム4008に基づいて、双方向で行うことができる。現行Wmedフレーム4010および動き補償された（「MC」）現行フレーム4016は、コンバイナ4012によってマージされ、または組み合わせられる。結果として生じるデインターレースされた現行フレーム4014は、現在はプログレッシブフレームであるが、デインターレースされた前フレーム4020として使用するためにME/MC 4018に戻して供給され、後続処理のためにデインターレース605の外部にも通信される。

【0061】

インターフィールド補間を備え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を伴うイントラフィールド補間から分離することが可能である。言い換えると、時空間Wmedフィルタリングを、主にイントラフィールド補間のために使用することができ、インターフィールド補間を、動き補償中に実行することができる。これは、Wmed結果のピーク信号対雑音比を下げるが、動き補償が適用された後の視覚的品質は、不正確なインターフィールド予測モード判断からの不良画素がWmedフィルタリング処理から除去されるので、より満足である。

【0062】

適当な逆テレシネまたはデインターレーシング処理の後に、ブロック608で、プログレッシブビデオが、エイリアス抑制および再サンプリング（たとえば、サイズ変更）のために処理される。いくつかの再サンプリング態様では、多フェーズリサンブラが、ピクチャサイズ変更のために実施される。ダウンサンプリングの一例では、オリジナルピクチャとサイズ変更されたピクチャとの間の比を、 p/q とすることができ、ここで、 p および q は、互いに素な整数である。フェーズの総数は、 p である。いくつかの態様での多フェーズフィルタのカットオフ周波数は、約0.5のサイズ変更係数に関して0.6である。カットオフ周波数は、サイズ変更されたシーケンスの高周波応答を高めるために、サイズ変更比と正確には一致しない。これは、不可避免的に多少のエイリアシングを許容する。しかし、人間の目が、ぼけているがエイリアスのないピクチャよりも、先鋭であるが多少エイリアスを有するピクチャを好むことが周知である。

【0063】

図41に、サイズ変更比が3/4である場合の相を示す多フェーズ再サンプリングの例を示す。図41に示されたカットオフ周波数は、やはり3/4である。オリジナル画素は、上の図では垂直軸を用いて図示されている。シンク関数も、フィルタ波形を表すためにこの軸を中心として描かれている。カットオフ周波数が、再サンプリング分配と正確に同一になるように選択されたので、シンク関数の0は、図41で十字を用いて示されているサイズ変更の後の画素の位置とオーバーラップする。サイズ変更の後の画素値を見つけるために、次の方程式に示されているように、オリジナル画素からの寄与を合計することができる。

【数 6】

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

【0064】

ここで、 f_c はカットオフ周波数である。上の 1D 多フェーズフィルタは、水平次元と垂直次元との両方に適用することができる。

【0065】

再サンプリング（サイズ変更）のもう 1 つの態様は、オーバースキャンを考慮に入れることである。NTSC テレビジョン信号では、1 つの画像が、486 本の走査線を有し、デジタルビデオでは、各走査線上に 720 個の画素を有することができる。しかし、サイズと画面フォーマットとの間の不一致に起因して、画像全体のすべてがテレビジョン上で可視であるわけではない。画像のうちで可視でない部分を、オーバースキャンと呼ぶ。

10

【0066】

放送会社が、できる限り多数のテレビジョンで可視の区域に有用な情報を置くのを助けるために、映画テレビジョン技術者協会（SMPTE）は、セーフアクションエリアおよびセーフタイトルエリアと呼ばれるアクションフレームの特定のサイズを定義した。SMPTE 推奨実践 RP 27.3 - 1989（Specifications for Safe Action and Safe Title Areas Test Pattern for Television Systems）を参照されたい。セーフアクションエリアは、SMPTE によって、「すべての重要なアクションが行われなければならない」区域と定義される。セーフタイトルエリアは、「家庭テレビジョン受像機の大多数での可視性を保証するために、すべての有用な情報を閉じこめることができる」区域と定義される。

20

【0067】

たとえば、図 25 を参照すると、セーフアクションエリア 2510 は、画面の中央 90 % を占め、四方に 5 % のボーダーを与える。セーフタイトルエリア 2505 は、画面の中央 80 % を占め、10 % のボーダーを与える。ここで図 26 を参照すると、セーフタイトルエリアは非常に小さいので、画像内により多くのコンテンツを追加するために、一部の局は、白い長方形の窓 2615 の内側であるセーフアクションエリア内にテキストを置く。

30

【0068】

通常、黒いボーダーが、オーバースキャン内で見られる場合がある。たとえば、図 26 では、黒いボーダーが、画像の上側 2620 および下側 2625 に現れている。これらの黒いボーダーは、H. 264 ビデオが動き推定でバウンダリエクステンション（boundary extension）を使用するので、オーバースキャン内で除去することができる。拡大された黒いボーダーが、残差を増やす可能性がある。逆に、境界を 2 % だけカットし、その後、サイズ変更を行うことができる。サイズ変更用のフィルタは、それ相応に生成することができる。トランケーションが、多フェーズ・ダウンサンプリングの前にオーバースキャンを除去するために実行される。

40

【0069】

もう一度図 6 を参照すると、プログレッシブビデオは、次に、ブロック 610 に進み、ここで、デブロック動作およびデリング動作が実行される。2 つのタイプの人為的な結果、すなわち「ブロッキング」および「リングング」が、一般にビデオ圧縮アプリケーションで発生する。ブロッキングの人為的な結果は、圧縮アルゴリズムが各フレームをブロック（たとえば、8 × 8 ブロック）に分割するので発生する。各ブロックは、ある小さい誤差を伴って再構成され、ブロックの縁での誤差は、しばしば、隣接するブロックの縁での誤差に対して明暗を有し、ブロック境界を可視にする。対照的に、リングングの人為的な結果は、画像特徴の縁の周囲のひずみとして現れる。リングングの人為的な結果は、符号化

50

器が高周波 D C T 係数の量子化で破棄する情報が多すぎるので発生する。いくつかの例示的な例で、デブロッキングとデリリングとの両方が、これらの可視の人為的結果を隠すために低域 F I R (有限インパルス応答) フィルタを使用することができる。

【 0 0 7 0 】

デブロッキング処理の一例で、デブロッキングフィルタを、フレームの境界の縁およびデブロッキングフィルタ処理が使用不可にされるすべての縁を除いて、フレームのすべての 4×4 ブロックの縁に適用することができる。このフィルタリング処理は、フレーム構成処理の完了後にマクロブロックを基礎として実行されなければならない、1 フレーム内のすべてのマクロブロックが、マクロブロックアドレスの昇順で処理される。マクロブロックごとに、垂直の縁が、まず、左から右へフィルタリングされ、次に、水平の縁が、上から下へフィルタリングされる。図 3 9 に示されているように、水平方向および垂直方向について、ルマデブロッキングフィルタ処理は、4 つの 16 サンプルの縁に対して実行され、各クロマ成分のデブロッキングフィルタ処理は、2 つの 8 サンプルの縁に対して実行される。前のマクロブロックに対するデブロッキング処理動作によって既に変更されている可能性がある、現行マクロブロックの上および左のサンプル値は、現行マクロブロックに対するデブロッキングフィルタ処理への入力として使用されなければならない、現行マクロブロックのフィルタリング中にさらに変更される可能性がある。垂直の縁のフィルタリング中に変更されたサンプル値を、同一マクロブロックの水平の縁のフィルタリングについて入力として使用することができる。デブロッキング処理を、ルマ成分およびクロマ成分について別々に呼び出すことができる。

【 0 0 7 1 】

デリリング処理の例では、2 D フィルタを、縁の近くの区域を平滑化するために適応的に適用することができる。縁の画素は、ぼけを防ぐためにフィルタリングをほとんどまたは全く受けない。

【 0 0 7 2 】

G O P パーティショナ

デブロッキングおよびデリリングの後に、プログレッシブビデオは、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 パーティショニングには、ショット変化の検出、複雑度マップ (たとえば、時間帯域幅マップ、空間帯域幅マップ) の生成、および適応的 G O P パーティショニングを含めることができる。これらを、それぞれ下で説明する。

【 0 0 7 3 】

A . シーン変化検出

ショット検出は、シーン変化が発生したことを示すデータを、グループオブピクチャ (G O P) 内の 1 フレームが何時示すかを判定することに関する。一般に、G O P 内で、フレームは、2 つまたは 3 つ (あるいはそれ以上) の隣接するフレームで大きい変化を有しない場合があり、あるいは、遅い変化または速い変化がある場合がある。もちろん、これらのシーン変化分類を、必要な、特定の応用例に依存する、より高いレベルの変化にさらに分解することができる。

【 0 0 7 4 】

ショット変化またはシーン変化の検出は、ビデオの効率的な符号化に重要である。通常、G O P が大きく変化してはいない時に、複数の予測フレームがその後に続くその G O P の先頭の I フレームは、ビデオを十分に符号化することができ、その結果、ビデオの後続の復号および表示が視覚的に許容できるものになる。しかし、シーンが、突然にまたはゆっくりとのいずれであれ、変化しつつある時には、追加の I フレームおよびより少ない予測符号化 (P フレームおよび B フレーム) が、後に復号化される視覚的に許容できる結果を作るために必要になる可能性がある。

【 0 0 7 5 】

既存の符号化システムの性能を改善するショット検出および符号化のシステムおよび方法を、下で説明する。そのような態様は、プリプロセッサ 2 2 6 (図 7) の G O P パーティショナ 6 1 2 内で実施することができ、あるいは、プリプロセッサと共にまたはプリブ

ロセッサなしで動作できる符号化器デバイス内に含めることができる。そのような態様は、突然のシーン変化が発生したか、シーンがゆっくり変化してるか、あるいはビデオ符号化を特に複雑にする可能性があるシーン内のカメラフラッシュライトがあるかどうかを判定するために、ビデオデータの隣接フレームの間の統計的比較を含む統計（またはメトリックス）を利用する。この統計は、プリプロセッサから取得可能であり、その後、符号化デバイスに送ることができ、あるいは、符号化デバイス内で生成することができる（たとえば、動き補償を実行するように構成されたプロセッサによって）。結果として生じる統計は、シーン変化の検出の判断を助ける。トランスコーディングを行うシステムには、しばしば、適切なプリプロセッサまたは構成可能なプロセッサが存在する。プリプロセッサが、動き補償によって助けられるデインターレーシングを実行する場合には、動き補償の統計が、利用可能であり、直ちに使用できる。そのようなシステムでは、ショット検出アルゴリズムが、システムの複雑さをわずかに高める場合がある。

10

【 0 0 7 6 】

本明細書で説明するショット検出器の例示的な例は、前のフレーム、現行フレーム、および次のフレームからの統計を利用することだけを必要とし、したがって、非常に低いレイテンシを有する。ショット検出器は、突然のシーン変化、クロスフェーディングおよび他の遅いシーン変化、ならびにカメラフラッシュライトを含む、複数の異なるタイプのショット事象を区別する。異なるタイプのショット事象を符号化器内で異なる戦略を用いて判定することによって、符号化器の効率および視覚的品質が高められる。

20

【 0 0 7 7 】

シーン変化検出は、ビデオ符号化システムが固定された間隔でIフレームを挿入することによってインテリジェントにビットを節約するために、任意のビデオ符号化システムに使用することができる。いくつかの態様において、プリプロセッサによって取得されるコンテンツ情報（たとえば、メタデータに組み込まれるかプリプロセッサ226によって計算されるかのいずれか）を、シーン変化検出に使用することができる。たとえば、コンテンツ情報に応じて、下で説明するしきい値および他の判断基準を、異なるタイプのビデオコンテンツについて動的に調整することができる。

【 0 0 7 8 】

ビデオ符号化は、通常、構造化されたグループオブピクチャ（GOP）に対して動作する。GOPは、通常、イントラ符号化されたフレーム（Iフレーム）から始まり、これに一連のP（予測）フレームまたはB（双方向）フレームが続く。通常、Iフレームは、そのフレームを表示するのに必要なすべてのデータを格納することができ、Bフレームは、先行するフレームおよび続くフレーム内のデータに頼り（たとえば、前のフレームから変化したデータまたは次のフレーム内のデータと異なるデータだけを含む）、Pフレームは、前のフレームから変化したデータを含む。一般的な使用法で、Iフレームは、符号化されたビデオ内でPフレームおよびBフレームと共に散在する。サイズ（たとえば、フレームを符号化するのに使用されるビット数）に関して、Iフレームは、通常、Pフレームよりはるかに大きく、Pフレームは、Bフレームより大きい。効率的な符号化、伝送、および復号化処理のために、GOPの長さは、大きいIフレームからの効率的な損失を減らすのに十分に大きく、符号化器と復号化器との間の不一致またはチャネル障害と戦うのに十分に短くしなければならない。さらに、Pフレーム内のマクロブロック（MB）を、同一の理由からイントラ符号化することができる。

30

40

【 0 0 7 9 】

シーン変化検出は、しばしば不必要なIフレームを固定間隔で挿入するのではなく、ビデオ符号化器が正しいGOP長を判定し、そのGOP長に基づいてIフレームを挿入するために使用することができる。実用的なストリーミングビデオシステムでは、通信チャネルは、通常、ビット誤りまたはパケット消失によってそこなわれる。IフレームまたはIMBが置かれる場所は、復号化されたビデオの品質および視聴経験に大きく影響する可能性がある。1つの符号化方式は、同一位置の以前のピクチャまたはピクチャ部分からの大きい変化を有するピクチャまたはピクチャの部分に、イントラ符号化されたフレームを

50

使用することである。通常、これらの領域は、動き推定を用いて有効に効率的に予測することはできず、符号化は、そのような領域がインターフレーム符号化技法（たとえば、BフレームおよびPフレームを使用する符号化）から除外される場合に、より効率的に行うことができる。チャネル障害の文脈では、これらの領域は、誤り伝搬から損害を受ける可能性が高いが、この誤り伝搬は、イントラフレーム符号化によって減らすか除去する（またはほぼ除去する）ことができる。

【0080】

GOPビデオの諸部分を、複数のカテゴリに分類することができ、ここで、各領域は、特定の実装に依存する可能性のあ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。例として、ビデオを、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および他の遅いシーン変化、ならびにカメラフラッシュライトに分類することができる。

10

【0081】

突然のシーン変化は、通常はカメラ動作によって引き起こされる、前フレームから大きく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ので、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されなければならない。

【0082】

クロスフェーディングおよび他の遅いシーン変化は、通常はカメラショットのコンピュータ処理によって引き起こされる、シーンの遅い切替を含む。2つの異なるシーンの徐々のブレンドは、人間の目にはより満足に見える可能性があるが、ビデオ符号化に対して課題を提示する。動き補償は、これらのフレームのビットレートを効果的に下げることができず、より多くのイントラMBが、これらのフレームについて更新される可能性がある。

20

【0083】

カメラフラッシュライトまたはカメラフラッシュ事象は、フレームのコンテンツがカメラフラッシュを含む時に発生する。そのようなフラッシュは、持続時間において比較的短く（たとえば、1フレーム）、極端に明るく、フラッシュを描いたフレーム内の画素は、隣接フレーム上の対応する区域に対して相対的に異常に高い輝度を示すようになっている。カメラフラッシュライトは、ピクチャの輝度を突然にすばやくシフトする。通常、カメラフラッシュライトの持続時間は、通常は44msと定義される人間の視覚系(HVS)の時間マスキング持続時間より短い。人間の目は、明るさのこれらの短いバーストの品質に敏感ではなく、したがって、これらを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、動き補償を用いて有効に処理することができず、将来のフレームに関する悪い予測候補なので、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分類されたシーンは、「人工的に」高い輝度のゆえに他のフレームの予測に使用してはならず、他のフレームは、同一の理由からこれらのフレームを予測するのに効果的に使用することができない。識別されたならば、これらのフレームは比較的大量の処理を必要とする可能性があるので、これらのフレームを除外することができる。1つのオプションは、カメラフラッシュライトフレームを除去し、その代わりにDC係数を符号化することであり、そのような解決策は、単純であり、計算的に高速であり、多数のビットを節約する。

30

40

【0084】

フレームの上のカテゴリのいずれかが検出される時には、ショット事象が宣言される。ショット検出は、符号化品質を改善するのに有用であるだけでなく、ビデオコンテンツ探索およびインデクシングを識別するのをも助けることができる。シーン検出処理の1つの例示的態様を、下で説明する。この例では、ショット検出処理は、まず、ショット検出について処理される選択されたフレームの情報またはメトリックスを計算する。このメトリックスには、ビデオの双方向の動き推定および補償処理からの情報、ならびに他の輝度ベースのメトリックスを含めることができる。

50

【 0 0 8 5 】

双方向の動き推定 / 補償を実行するために、ビデオシーケンスを、現行フレームのすべての 8×8 ブロックを 1 つは過去の、1 つは将来の 2 つの最も隣接する近傍フレーム内のブロックと照合する双方向の動き補償器を用いて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ついて動きベクトルおよび差分メトリックを作る。図 29 は、過去のフレーム P および将来の（または次の）フレーム N に一致する現行フレーム C の画素の例を示し、一致する画素への動きベクトル（過去の動きベクトル MV_P および将来の動きベクトル MV_N ）を示す図である。双方向の動きベクトルの生成および関連する符号化の一般的な説明は、一般に、下で、図 32 を参照して説明する。

【 0 0 8 6 】

双方向の動き情報（たとえば、対応する隣接フレーム内の MB（最良一致）を識別する動き情報を決定した後に、次のフレームおよび前のフレームに対する現行フレームのさまざまな比較によって、追加のメトリックスを生成することができる（たとえば、GOP パーティション 6 1 2 内の動き補償器または別の適切な構成要素によって）。動き補償器は、すべてのブロックの差分メトリックを作ることができる。差分メトリックは、二乗差の合計（SSD）または絶対差の合計（SAD）とすることができる。一般性を失わずに、本明細書では例として SAD を使用する。

【 0 0 8 7 】

すべてのフレームについて、「コントラスト比」とも称する SAD 比を、下のよう

【 数 7 】

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

【 0 0 8 8 】

ここで、 SAD_P および SAD_N は、それぞれ順方向の差分メトリックおよび逆方向の差分メトリックの絶対差の合計である。分母が、「ゼロによる除算」のエラーを防ぐために、小さい正の数を含むことに留意されたい。分子も、分母の 1 の影響の釣り合いをとるためにを含む。たとえば、前のフレーム、現行フレーム、および次のフレームが同一である場合に、動き探索は、 $SAD_P = SAD_N = 0$ をもたらさなければならない。この場合に、0 または無限大ではなく上の計算は $= 1$ を生成する。

【 0 0 8 9 】

輝度ヒストグラムを、すべてのフレームについて計算することができる。通常、マルチメディア画像は、8 ビットの輝度深さ（たとえば、「ビン」の個数）を有する。いくつかの態様による輝度ヒストグラムの計算に使用される輝度深さには、ヒストグラムを取得するために 16 に設定されることができる。他の態様では、輝度深さに、処理されるデータのタイプ、使用可能な計算能力、または他の所定の判断基準に依存するものとして設定される適切な数に設定されることができる。いくつかの態様において、輝度深さは、データの内容など、計算されたメトリックまたは受け取られたメトリックに基づいて動的に設定されることができる。

【 0 0 9 0 】

下の方程式は、輝度ヒストグラム差分（ラムダ）を計算することの 1 つの例を示す。

【 数 8 】

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

【 0 0 9 1 】

ここで、 N_{Pi} は、前のフレームの第 i ビン内のブロックの個数であり、 N_{Ci} は、現行

フレームの第 i ビン内のブロックの個数であり、 N は、1 フレーム内のブロックの総数である。前のフレームおよび現行フレームの輝度ヒストグラム差分が、完全に異なる（または互いに素である）場合には、 $\lambda = 2$ である。

【 0 0 9 2 】

この情報を使用して、フレーム差分メトリック (D) が、次のように計算される。

【 数 9 】

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

10

【 0 0 9 3 】

ここで、 A は、アプリケーションによって選択される定数であり、

【 数 1 0 】

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

【 0 0 9 4 】

である。

【 0 0 9 5 】

20

選択された（現行の）フレームは、フレーム差分メトリックが式 9 に示された判断基準を満足する場合に、突然のシーン変化と分類される。

【 数 1 1 】

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

【 0 0 9 6 】

ここで、 A は、アプリケーションによって選択される定数であり、 T_1 は、しきい値である。

30

【 0 0 9 7 】

一つの例示的なシミュレーションは、 $A = 1$ および $T_1 = 5$ を設定することによって、よい検出性能が達成されることを示す。現行フレームが、突然のシーン変化のフレームである場合には、 γ_C を大きくしなければならず、 γ_P を小さくしなければならない。比

【 数 1 2 】

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

【 0 0 9 8 】

40

を単独の γ_C の代わりに使用することができ、その結果、このメトリックが、コンテキストのアクティビティレベルに対して正規化されるようになる。

【 0 0 9 9 】

上の判断基準が、非線形の形で輝度ヒストグラム差分ラムダ (λ) を使用することに留意されたい。図 16 は、 $\lambda^*(2\lambda + 1)$ が凸関数であることを示す。 λ が小さい（たとえば、0 に近い）場合に、 $\lambda^*(2\lambda + 1)$ は、かろうじてプリエンファシスである。 λ が大きくなればなるほど、この関数によってより多くのエンファシスが行われる。このプリエンファシスを用いると、1.4 より大きいすべての λ について、突然のシーン変化は、しきい値 T_1 に 5 がセットされる場合に検出される。

【 0 1 0 0 】

50

現行フレームは、シーン強度メトリック D がある個数の連続するフレームについて式 5 に示された判断基準を満足する場合に、クロスフェーディングまたは遅いシーン変化であると判定される。

【数 1 3】

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

【0 1 0 1】

ここで、 T_1 は、上で使用されたものと同一のしきい値であり、 T_2 は、もう 1 つのしきい値である。

10

【0 1 0 2】

フラッシュライト事象は、通常、輝度ヒストグラムをより明るい側にシフトさせる。この例示的態様のカメラでは、現行フレームがカメラフラッシュライトを備えるかどうかを判定するのに輝度ヒストグラム統計が使用される。ショット検出処理は、式 1 1 および 1 2 に示されているように、現行フレームマイナスの輝度が前のフレームの輝度があるしきい値 T_3 だけ超えて明るく、現行フレームの輝度が次のフレームの輝度をしきい値 T_3 だけ超えて明るいかどうかを判定することができる。

【数 1 4】

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

20

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

【0 1 0 3】

上の判断基準が満足されない場合には、現行フレームは、カメラフラッシュライトを備えるものとしては分類されない。その判断基準が満足される場合には、ショット検出処理は、下の式に示されているように、逆方向の差分メトリック SAD_P および順方向の差分メトリック SAD_N が、あるしきい値 T_4 より大きいかどうかを判定する。

【数 1 5】

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

30

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

【0 1 0 4】

ここで

【数 1 6】

$$\bar{Y}_C$$

40

【0 1 0 5】

は、現行フレームの平均輝度であり、

【数 1 7】

$$\bar{Y}_P$$

【0 1 0 6】

は、前のフレームの平均輝度であり、

【数 18】

$$\bar{Y}_N$$

【0107】

は、次のフレームの平均輝度であり、SADPおよびSADNは、現行フレームに関連する順方向の差分メトリックおよび逆方向の差分メトリックである。

【0108】

ショット検出処理は、まず現行フレームの輝度が前のフレームの輝度および次のフレームの輝度より大きいかどうかを判定することによって、カメラフラッシュ事象を判定する。そうでない場合には、そのフレームは、カメラフラッシュ事象ではないが、そうである場合には、カメラフラッシュ事象である可能性がある。次に、ショット検出処理は、逆方向の差分メトリックがしきい値 T_3 より大きいかどうか、および順方向の差分メトリックがしきい値 T_4 より大きいかどうかを評価することができ、これらの条件の両方が満足される場合には、ショット検出処理は、現行フレームをカメラフラッシュライトを有するものとして分類する。その判断基準が満足されない場合には、そのフレームは、どのタイプのショット事象としても分類されず、あるいは、そのフレームに、そのフレームに対して行うべき符号化（たとえば、フレームを捨てる、Iフレームとして符号化する）を識別するデフォルト分類を与えることができる。

【0109】

T_1 、 T_2 、 T_3 、および T_4 のいくつかの例示的な値は、上に示されている。通常、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介して選択される。いくつかの態様において、しきい値 T_1 、 T_2 、 T_3 、および T_4 のうちの1つまたは複数、事前に決定され、そのような値は、符号化デバイス内のショット分類器に組み込まれる。いくつかの態様において、しきい値 T_1 、 T_2 、 T_3 、および T_4 のうちの1つまたは複数、ショット分類器に供給される情報（たとえば、メタデータ）を使用することに基づいて、また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（たとえば、動的に）設定することができる。

【0110】

ショット検出情報を使用するビデオの符号化は、通常、符号化器内で実行されるが、ショット検出の開示の完全を期してここで説明する。図30を参照すると、符号化処理301は、フレームのシーケンス内の検出されたショットに基づいてビデオを符号化するのにショット検出情報を使用することができる。処理301は、ブロック303に進み、現行フレームが突然のシーン変化として分類されるかどうかを調べるために検査する。そうである場合には、ブロック305で、現行フレームをIフレームとして符号化することができ、GOP境界を判定することができる。そうでない場合には、処理301はブロック307に進み、現行フレームが遅く変化するシーンの一部として分類される場合には、ブロック309で、現行フレームおよび遅く変化するシーン内の他のフレームを予測フレーム（たとえば、PフレームまたはBフレーム）として符号化することができる。処理301は、その後、ブロック311に進み、ここで、現行フレームがカメラフラッシュを備えるフラッシュライトシーンとして分類されるかどうかを検査する。そうである場合には、ブロック313で、そのフレームを空間処理、たとえば除去またはフレームのDC係数の符号化のために識別することができ、そうでない場合には、現行フレームの分類は行われておらず、現行フレームを、他の判断基準に従って符号化するか、Iフレームとして符号化するか、捨てることができる。

【0111】

上で説明した態様では、圧縮されるフレームとそれに隣接する2つのフレームとの間の差の量が、フレーム差分メトリックDによって示される。大きい量の一方の輝度の変化が検出される場合には、これは、フレーム内のクロスフェード効果を意味する。クロスフェードが顕著であればあるほど、Bフレームを使用することによって多くの利益を達成す

ることができる。いくつかの態様では、下の方程式に示される、変更されたフレーム差分メトリックが使用される。

【数 19】

$$D_1 = \begin{cases} \left(1 - \alpha + 2\alpha \frac{|d_P - d_N|}{d_P + d_N}\right) \times D, & Y_P - \Delta \geq Y_C \geq Y_N + \Delta \text{ または } Y_P + \Delta \leq Y_C \leq Y_N - \Delta \text{ の場合、} \\ D, & \text{それ以外の場合} \end{cases}$$

[15]

10

【0112】

ここで、 $d_P = |Y_C - Y_P|$ および $d_N = |Y_C - Y_N|$ は、現行フレームと前のフレームとの間のルマ差分、ならびに現行フレームと次のフレームとの間のルマ差分であり、 α は、実装に依存するので通常の実験で決定されることのできる定数を表し、 α は、0と1との間の値を有する重み付け変数である。

【0113】

B：帯域幅マップ生成

プリプロセッサ226（図6）を、マルチメディアデータの符号化に使用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。いくつかの態様では、符号化器228（図7）内のコンテンツ分類モジュール712が、その代わりに帯域幅マップを生成する。

20

【0114】

人間の視覚的品質Vは、符号化の複雑度Cと割り当てられたビットB（帯域幅とも称する）との両方の関数とすることができる。図15は、この関係を示すグラフである。符号化の複雑度メトリックCが、人間の視覚の観点からの空間周波数および時間周波数を考慮することに留意されたい。人間の目がより敏感である歪みについて、複雑度の値は、それに対応してより高い。通常は、Vが、Cにおいて単調に減少し、Bにおいて単調に増加すると仮定することができる。

【0115】

一定の視覚的品質を達成するために、真下の2つの方程式で表される判断基準を満足する帯域幅（ B_i ）が、符号化されるi番目のオブジェクト（フレームまたはMB）に割り当てられる。

30

【数20】

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

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

【0116】

真上の2つの方程式では、 C_i が、i番目のオブジェクトの符号化の複雑度であり、Bが、使用可能な総帯域幅であり、Vが、オブジェクトについて達成される視覚的品質である。人間の視覚的品質は、方程式として定式化することがむずかしい。したがって、上の方程式の組は、正確には定義されない。しかし、3Dモデルがすべての変数において連続であると仮定する場合に、帯域幅比（ B_i / B ）を、（C, V）対の近傍の中で変更されないものとして扱うことができる。帯域幅比 β_i は、下に示す方程式で定義される。

40

【数21】

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

【0117】

次に、ビット割当を、次の方程式で表されるように定義することができる。

50

【数 2 2】

$$\beta_i = \beta(C_i) \quad [19]$$

$$1 = \sum_i \beta_i \quad (C_i, V) \in \delta(C_0, V_0) \text{ について}$$

【0 1 1 8】

ここで、 δ は、「近傍」を示す。

【0 1 1 9】

符号化の複雑度は、空間と時間との両方の人間の視覚的感度によって影響される。ジロ 10
 の人間の視覚モデルが、空間的な複雑度を定義するのに使用できるモデルの例である。こ
 のモデルは、局所空間周波数および周囲照明を考慮する。結果のメトリックを、 $D_{c s a t}$
 t と呼ぶ。処理の前処理点では、ピクチャがイントラ符号化されるのか、インター符号化
 されるのかが、未知であり、この両方の帯域幅比が、生成される。ビットは、異なるビデ
 オオブジェクトの $I N T R A$ の間の比に従って割り当てられる。イントラ符号化される
 ピクチャについて、帯域幅比は、次の方程式で表される。

【数 2 3】

$$\beta_{I N T R A} = \beta_{0 I N T R A} \log_{10} (1 + \alpha_{I N T R A} Y^2 D_{c s a t}) \quad [20]$$

20

【0 1 2 0】

上の方程式では、 Y は、マクロブロックの平均輝度成分であり、 $I N T R A$ は、輝度
 の二乗の重み付け係数であり、 $D_{c s a t}$ 項がそれに続き、 $0 I N T R A$ は、

【数 2 4】

$$1 = \sum_i \beta_i$$

【0 1 2 1】

を保証するための正規化係数である。たとえば、 $I N T R A = 4$ の値が、よい視覚的品 30
 質を達成する。ビデオの特定のコンテンツの所望のよい視覚的品質レベルに対応する値を
 $I N T R A$ に設定するために、コンテンツ情報（たとえば、コンテンツ分類）を使用す
 ることができる。一例で、ビデオコンテンツが「語り手の顔」ニュースブロードキャスト
 を備える場合には、情報画像またはビデオの表示可能部分がオーディオ部分より重要性が
 少ないと考えることができるので、視覚的品質レベルを、より低く設定することができ、
 より少ないビットを、データを符号化するために割り当てることができる。もう 1 つの例
 で、ビデオコンテンツが、スポーツイベントを備える場合には、表示される画像が見る人
 にとってより重要である可能性があるので、より高い視覚的品質レベルに対応する値を
 $I N T R A$ に設定するためにコンテンツ情報を使用することができ、したがって、より多 40
 くのビットを、データを符号化するために割り当てることができる。

【0 1 2 2】

この関係を理解するために、帯域幅が、符号化の複雑度に関して対数的に割り当てられ
 ることに留意されたい。輝度の二乗の項 Y^2 は、より大きい大きさを有する係数が符号化
 により多くのビットを使用するという事実を反映する。対数が負の値になるのを防ぐため
 に、1 が、括弧内の項に加算される。他の底を有する対数を使用することもできる。

【0 1 2 3】

時間的な複雑度は、絶対差分の合計（SAD）などのフレーム差分メトリックと共に動
 きの量（たとえば、動きベクトル）を考慮に入れて 2 つの連続するフレームの間の差を測
 定するフレーム差分メトリックの測定によって決定される。

【0 1 2 4】

50

インター符号化されるピクチャのビット割当は、空間的な複雑度ならびに時間的な複雑度を考慮することができる。これを、下に表す。

【数 25】

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

【0125】

上の方程式では、 MV_P および MV_N は、現行 MB の順方向の動きベクトルおよび逆方向の動きベクトルである（図 29 を参照されたい）。イントラ符号化される帯域幅の式の Y^2 が、二乗差の合計（SSD）に置換されていることに留意されたい。上の方程式の $MV_P + MV_N$ の役割を理解するために、滑らかで予測可能な動き（小さい $MV_P + MV_N$ ）を受けている区域は、注意を引き、目によって追跡できるが、通常は静止領域より多くのひずみを許容することができないという人間の視覚系の特性に留意されたい。しかし、速いまたは予測不能な動き（大きい $MV_P + MV_N$ ）を受けている区域は、追跡できず、大きい量子化を許容することができる。実験は、 $I_{INTER} = 1$ 、 $= 0.001$ が、よい視覚的品質を達成することを示す。

【0126】

C. 適応 GOP パーティショニング

プリプロセッサ 226 によって実行できる処理のもう 1 つの例示的な例において、図 6 の GOP パーティショナ 612 は、一緒に符号化されるグループオブピクチャの構成を適応的に変更することもでき、これを MPEG2 を使用する例を参照して述べる。いくつかのより古いビデオ圧縮標準規格（たとえば、MPEG2）は、GOP が規則的な構造を有することを要求しない（それを強要することはできるが）。MPEG2 シーケンスは、必ず I フレームすなわち前のピクチャを参照せずに符号化されたフレームから始まる。MPEG2 GOP フォーマットは、通常、I フレームに続く P ピクチャすなわち予測ピクチャの GOP 内の間隔を固定することによって、符号化器で事前に配置される。P フレームは、前の I ピクチャまたは P ピクチャから部分的に予測されたピクチャである。先頭の I フレームと後続の 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、基準として「個別に」または「同時に」のいずれかで前および次の I ピクチャまたは P ピクチャを使用することができる。I フレームを符号化するのに必要なビット数は、平均して、P フレームを符号化するのに必要なビット数を超え、同様に、P フレームを符号化するのに必要なビット数は、平均して、B フレームに必要なビット数を超える。スキップされたフレームが使用される場合に、スキップされたフレームは、その表現にビットを必要としないはずである。

【0127】

P フレームおよび B フレームの使用、およびより最近の圧縮アルゴリズムにおける、ビデオを表すのに必要なデータのレートを下げるためのフレームのスキップの使用の基礎になる概念が、時間的な冗長度の除去である。時間的な冗長度が高いすなわちピクチャからピクチャへの変化がほとんどない時に、P ピクチャ、B ピクチャ、またはスキップされたピクチャの使用は、より早期に復号化された I ピクチャまたは P ピクチャが後に他の P ピクチャまたは B ピクチャを復号化するのに基準として使用されるので、ビデオストリームを効率的に表す。

【0128】

適応的 GOP パーティショニングは、この概念を適応的に使用することに基づく。フレームの間の差が、定量化され、適切なテストが定量化された差に対して実行された後に、I フレーム、P フレーム、B フレーム、またはスキップされたフレームによってピクチャを表すという判断が自動的に行われる。適応的な構造は、固定された GOP 構造では利用可能でない利益を有する。固定された構造は、コンテンツの変化がほとんど発生していない可能性を無視し、適応的な手順は、各 I フレームと P フレームとの間または 2 つの P フレームの間に、はるかにより多くの B フレームを挿入することを可能にし、これによって

、フレームのそのシーケンスを適切に表すのに必要なビット数を減らす。逆に、ビデオコンテンツの変化が大きい時には、Pフレームの効率は、予測フレームと基準フレームとの間の差が大きすぎるので、大幅に低下する。これらの条件の下では、一致するオブジェクトが、動き探索領域から外れる可能性があり、あるいは、一致するオブジェクトの間の類似性が、カメラアングルの変化によって引き起こされる歪みに起因して低下する。その点で、PフレームまたはIフレームおよびそれに隣接するPフレームは、互いにより近くなるように選択されなければならない、より少数のBフレームを挿入しなければならない。固定されたGOPは、その調整を行うことができない。

【0129】

本明細書で開示されるシステムでは、これらの条件が、自動的に感知される。GOPの構造は、柔軟であり、コンテンツのこれらの変化に適応するようにされる。このシステムは、距離と同一の加法特性を有する、フレームの間の距離の尺度と考えることができるフレーム差分メトリックを評価する。概念上、インターフレーム距離 d_{12} および d_{23} を有するフレーム F_1 、 F_2 、および F_3 を与えられれば、 F_1 と F_3 との間の距離は、少なくとも $d_{12} + d_{23}$ と考えられる。フレーム割当は、この距離に類似したメトリックを基礎として行われる。

【0130】

GOPパーティションは、フレームが受け取られる時にそれらのフレームにピクチャタイプを割り当てることによって動作する。ピクチャタイプは、各ブロックの符号化に必要な可能性がある予測の方法を示す。

【0131】

Iピクチャは、他のピクチャを参照せずに符号化される。Iピクチャは、独立なので、データストリーム内の復号化を開始できるアクセス点を提供する。I符号化タイプは、先行フレームへの「距離」がシーン変化しきい値を超える場合に、フレームに割り当てられる。

【0132】

Pピクチャは、動き補償された予測に前のIピクチャまたはPピクチャを使用することができる。Pピクチャは、符号化の基礎として、予測されるブロックから変位されているものとして前のフィールドまたはフレーム内のブロックを使用する。基準ブロックが、考慮中のブロックから減算された後に、残差ブロックが、通常は空間的な冗長度の除去のために離散コサイン変換を使用して符号化される。P符号化タイプは、あるフレームとPフレームに割り当てられた最後のフレームとの間の「距離」が、通常は第1しきい値より小さい第2しきい値を超える場合に、そのフレームに割り当てられる。

【0133】

Bフレームピクチャは、上で説明したように、動き補償に前および次のPピクチャまたはIピクチャを使用することができる。Bピクチャ内のブロックを、順方向、逆方向、または双方向で予測することができ、あるいは、他のフレームを参照せずにイントラ符号化することができる。H.264では、基準ブロックを、同数のフレームからの32個ものブロックの線形組合せとすることができる。フレームを、IタイプまたはPタイプになるように割り当てることができない場合に、そのフレームから直前の先行フレームまでの「距離」が通常は第2しきい値より小さい第3しきい値を超えるならば、そのフレームは、Bタイプになるように割り当てられる。

【0134】

フレームを、Bフレーム符号化されるようになるように割り当てることができない場合には、そのフレームには、「スキップフレーム」状況が割り当てられる。このフレームは、事実上前のフレームのコピーなので、スキップすることができる。

【0135】

表示順で隣接するフレームの間の距離を定量化するメトリックを評価することは、この処理の最初に行われる部分である。このメトリックは、上で言及した距離であり、これを用いて、すべてのフレームが、その正しいタイプについて評価される。したがって、Iフ

10

20

30

40

50

レームと隣接するPフレームとの間または2つの連続するPフレームの間の間隔を、可変とすることができる。このメトリックの計算は、ブロックベースの動き補償器を用いてビデオフレームを処理することによって開始され、ブロックは、通常は16×16画素からなるビデオ圧縮の基礎単位であるが、8×8、4×4、および8×16などの他のブロックサイズが可能である。2つのデインターレースされたフィールドからなるフレームについて、動き補償は、フィールドを基礎として行うことができ、基準ブロックの探索は、フレームではなくフィールド内で行われる。現行フレームの第1フィールド内のブロックについて、順方向の基準ブロックは、それに続くフレームのフィールド内で見つけれ、同様に、逆方向の基準ブロックは、現行フィールドに直接に先行するフレームのフィールド内で見つけれ。現行ブロックは、補償されたフィールドに組み立てられる。この処理は、そのフレームの第2フィールドに継続される。2つの補償されたフィールドが、順方向の補償されたフレームおよび逆方向の補償されたフレームを形成するために組み合わせられる。

10

【0136】

逆テレシネ606で作成されたフレームについて、基準ブロックの探索は、再構成されたフィルムフレームだけなので、フレームのみを基礎とする。順方向および逆方向の、2つの基準ブロックおよび2つの差が、見つけれ、順方向および逆方向の補償されたフレームにもつながる。要約すると、動き補償器は、すべてのブロックについて動きベクトルおよび差分メトリックスを作るが、ブロックは、デインターレース605の出力が処理される場合には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にはフィルムフレームの一部である。メトリックの差が、検討されているフィールドまたはフレームの中のブロックと、順方向の差分または逆方向の差分のどちらが評価されているかに応じて、前のフィールドもしくはフレーム内、または直接に続くフィールドもしくはフレーム内のいずれであれ、最良一致ブロックとの間で評価されることに留意されたい。輝度値だけが、この計算に入力される。

20

【0137】

したがって、動き補償ステップは、2組の差を生成する。これらは、輝度の現在値のブロックと、時間的に現在のフレームの直後および直前のフレームからとられた基準ブロック内の輝度との間の差である。順方向の差分の各々および逆方向の差分の各々の絶対値が、画素ごとに判定され、それぞれが、フレーム全体にわたって別々に合計される。1つのフレームを備えるデインターレースされたNTSCフィールドが処理される時には、両方のフィールドが、2つの合計に含まれる。この形で、 SAD_p および SAD_N すなわち順方向の差分および逆方向の差分の合計された絶対値が見つけれ。

30

【0138】

すべてのフレームについて、 SAD 比は、次の関係を使用して計算される。

【数26】

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

【0139】

ここで、 SAD_p および SAD_N は、それぞれ、順方向の差分および逆方向の差分の合計された絶対値である。「ゼロによる除算」のエラーを防ぐために、分子に小さい正の数が加算される。同様な項が、分母に加算され、 SAD_p または SAD_N のいずれかが0に近い時に、その感度をさらに下げる。

40

【0140】

代替の態様では、差を、SSDすなわち二乗差の合計およびSADすなわち絶対差の合計またはSATDとすることができ、SATDでは、画素値のブロックが、ブロック要素内の差をとる前にそれらのブロックに2次元離散コサイン変換を適用することによって変換される。合計は、アクティブビデオの区域にわたって評価されるが、他の態様では、より小さい区域を使用することができる。

50

【 0 1 4 1 】

受け取られたままの（動き補償されていない）すべてのフレームの輝度ヒストグラムも、計算される。このヒストグラムは、輝度値のブロックが使用可能であった場合にその輝度値のブロックに 2 次元離散コサイン変換を適用したことの結果である係数の 16×16 配列内の DC 係数すなわち (0, 0) 係数に影響する。これと同等に、 16×16 ブロック内の輝度の 256 個の値の平均値を、ヒストグラム内で使用することができる。その輝度深さが 8 ビットである画像について、ピンの個数は 16 に設定される。次のメトリックは、ヒストグラム差分を評価する。

【 数 2 7 】

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

10

【 0 1 4 2 】

上の式では、 N_{Pi} は、第 i ピン内の前のフレームからのブロックの個数であり、 N_{Ci} は、第 i ピンに属する現行フレームからのブロックの個数であり、 N は、1 フレーム内のブロックの総数である。

【 0 1 4 3 】

これらの中間結果は、次のように現行フレームの差分メトリックを形成するために組み立てられる。

20

【 数 2 8 】

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

【 0 1 4 4 】

ここで、 γ_C は、現行フレームに基づく SAD 比であり、 γ_P は、前フレームに基づく SAD 比である。シーンが滑らかな動きを有し、そのルマヒストグラムがほとんど変化しない場合には、 $D = 1$ である。現行フレームが突然のシーン変化を表示する場合には、 γ_C が大きくなり、 γ_P が小さくならなければならない。 γ_C 単独ではなく比

30

【 数 2 9 】

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

【 0 1 4 5 】

が使用され、その結果、このメトリックが、コンテキストのアクティビティレベルに対して正規化されるようになる。

【 0 1 4 6 】

図 4 2 に、フレームに圧縮タイプを割り当てる処理を示す。D すなわち式 1 9 で定義された現行フレームの差が、フレーム割当に関して行われる判断の基礎である。判断ブロック 4 2 0 2 に示されているように、考慮中のフレームがシーケンスの最初のフレームである場合に、YES とマークされたブロック 4 2 0 6 への判断経路に従い、これによって、そのフレームが I フレームであると宣言する。ブロック 4 2 0 8 で、累算されたフレーム差分に 0 を設定し、この処理は（ブロック 4 2 1 0 で）開始ブロックにリターンする。考慮中のフレームがシーケンスの最初のフレームではない場合に、その判断が行われたブロック 4 2 0 2 からの NO とマークされた判断経路に従い、テストブロック 4 2 0 4 で、現行フレームの差をシーン変化しきい値に対してテストする。現行フレームの差がそのしきい値より大きい場合には、ブロック 4 2 0 6 への YES とマークされた判断経路に従い、これは、やはり I フレームの割当につながる。

40

【 0 1 4 7 】

50

現行フレームの差がシーン変化しきい値より小さい場合には、ブロック 4 2 1 2 への N O 経路に従い、ここで、現行フレームの差を、累算されたフレーム差分に加算する。この流れ図を介して判断ブロック 4 2 1 4 で継続して、累算されたフレーム差分を、一般にシーン変化しきい値より小さいしきい値 t と比較する。累算されたフレーム差分が t より大きい場合には、制御はブロック 4 2 1 6 に転送され、フレームは P フレームになるように割り当てられ、その後、ステップ 4 2 1 8 で、累算されたフレーム差分を 0 にリセットする。累算されたフレーム差分が t より小さい場合には、制御はブロック 4 2 1 4 からブロック 4 2 2 0 に転送される。ここでは、現行フレームの差を、 t より小さい と比較する。現行フレームの差が より小さい場合には、ブロック 4 2 2 2 で、そのフレームは、スキップされるように割り当てられ、この処理はリターンし、現行フレームの差が より大きい場合には、ブロック 4 2 2 6 で、そのフレームは、B フレームになるように割り当てられる。

10

【 0 1 4 8 】

符号化器

戻って図 2 を参照すると、トランスコーダ 2 0 0 は、処理されたメタデータおよび生ビデオをプリプロセッサ 2 2 6 から受け取る符号化器 2 2 8 を含む。メタデータには、元々はソースビデオ 1 0 4 内で受け取られたすべての情報およびプリプロセッサ 2 2 6 によって計算されたすべての情報を含めることができる。符号化器 2 2 8 は、第 1 パス符号化器 2 3 0、第 2 パス符号化器 2 3 2、および再符号化器 2 3 4 を含む。符号化器 2 2 8 は、第 2 パス符号化器 2 3 2 から第 1 パス符号化器 2 3 0、再符号化器 2 3 4、ならびにプリプロセッサ 2 2 6 に情報（たとえば、メタデータ、誤り弾力性情報、コンテンツ情報、符号化ビットレート情報、ベースレイヤおよびエンハンスメントレイヤバランス情報、および量子化情報）を供給できるトランスコーダ制御 2 3 1 から入力を受け取る。符号化器 2 2 8 は、プリプロセッサ 2 2 6 から受け取ったコンテンツ情報および / または符号化器 2 2 8 自体によって、たとえばコンテンツ分類モジュール 7 1 2（図 7）によって生成されたコンテンツ情報を使用して、受け取ったビデオを符号化する。

20

【 0 1 4 9 】

図 2 に示された符号化器 2 2 8 のために使用できる例示的な 2 パス符号化器に含めることができる機能モジュールのブロック図を図 7 に示す。機能モジュールのさまざまな態様が図 7 に示されているが、図 7 および本明細書の説明は、必ずしも、符号化器に組み込むことのできるすべての機能性には対処しない。したがって、機能モジュールのある種の態様が、以下のベースレイヤ符号化およびエンハンスメントレイヤ符号化の議論に続いて、以下で説明される。

30

【 0 1 5 0 】

ベースレイヤ符号化およびエンハンスメントレイヤ符号化

符号化器 2 2 8 は、プリプロセッサ 2 2 6 からの生ビデオおよびメタデータを本明細書でベースレイヤとも称する符号化されたデータの第 1 グループおよび本明細書でエンハンスメントレイヤとも称する符号化されたデータの 1 つまたは複数の追加的なグループに符号化することができる S N R スケラブル符号化器とすることができる。符号化アルゴリズムは、復号化された時に、両方のレイヤが復号化に使用可能な時に復号化器で組み合わせることができるベースレイヤ係数およびエンハンスメントレイヤ係数を生成する。両方のレイヤが使用可能ではない時には、ベースレイヤの符号化が、ベースレイヤを単一のレイヤとして復号化することを可能にする。

40

【 0 1 5 1 】

そのようなマルチレイヤ符号化処理の一態様を、図 3 1 を参照して説明する。ブロック 3 2 1 で、完全にイントラ符号化されたマクロブロック（イントラ符号化された M B）を用いて、I フレームを符号化する。H . 2 6 4 では、I フレーム内のイントラ符号化された M B は、かなりの量の符号化利得をもたらす完全に活用された空間的な予測を用いて符号化される。2 つのサブモードすなわち、I n t r a 4 x 4 および I n t r a 1 6 x 1 6 がある。ベースレイヤが、空間的な予測によって与えられる符号化利得を利用する場合に

50

は、ベースレイヤは、エンハンスメントレイヤの符号化および復号化の前に符号化され、復号化される必要がある。Iフレームの2パス符号化および復号化が使用される。ベースレイヤでは、ベースレイヤ量子化パラメータ Q_P が、変換係数に粗い量子化ステップサイズを与える。オリジナルフレームと再構成されたベースレイヤフレームとの間の画素ごとの差が、エンハンスメントレイヤで符号化される。エンハンスメントレイヤは、より微細な量子化ステップサイズを与える量子化パラメータ Q_P を使用する。図2の符号化器228などの符号化ための手段が、ブロック321で符号化を実行することができる。

【0152】

ブロック323で、符号化器は、処理されているGOP内のPフレームおよび/またはBフレームのベースレイヤデータおよびエンハンスメントレイヤデータを符号化する。符号化器228などの符号化ための手段が、ブロック323で符号化を実行することができる。ブロック325で、符号化処理は、符号化すべきPフレームまたはBフレームがまだあるかどうかを検査する。SNRスケーラブル符号化器228などの符号化ための手段が、行為325を実行することができる。PフレームまたはBフレームがまだ残っている場合には、GOP内のすべてのフレームが符号化され終わるまで、ステップ323を繰り返す。PフレームおよびBフレームは、インター符号化されたマクロブロック（インター符号化されたMB）からなるが、以下で述べるように、PフレームおよびBフレーム内にイントラ符号化されたMBがあるものとすることができる。

【0153】

復号化器が、ベースレイヤデータとエンハンスメントレイヤデータを区別するために、符号化器228は、ブロック327で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タイプには、たとえば、レイヤの個数を識別するデータ、ベースレイヤとしてレイヤを識別するデータ、エンハンスメントレイヤとしてレイヤを識別するデータ、レイヤの間の相互関係を識別するデータ（レイヤ2がベースレイヤ1のエンハンスメントレイヤである、あるいはレイヤ3がエンハンスメントレイヤ2のエンハンスメントレイヤであるなど）、または複数のエンハンスメントレイヤの系列内の最後のエンハンスメントレイヤとしてレイヤを識別するデータが含まれる。オーバーヘッド情報は、関係するベースレイヤデータおよび/またはエンハンスメントレイヤデータに接続されたヘッダに含めるか、別々のデータメッセージに含めることができる。図2の符号化器228などの符号化ための手段が、ブロック327でこの処理を実行することができる。

【0154】

単一レイヤの復号化を有するために、2つのレイヤの係数が、逆量子化の前に組み合わせられる。したがって、この2つのレイヤの係数は、相互に作用して生成されなければならない、さもなければ、これが、かなりの量のオーバーヘッドを導入する可能性がある。オーバーヘッド増加の1つの理由は、ベースレイヤ符号化およびエンハンスメントレイヤ符号化が、異なる時間的な基準を使用し得ることである。ベースレイヤとエンハンスメントレイヤとの両方が使用可能である時に、逆量子化の前に復号化器で組み合わせることができる、ベースレイヤ係数およびエンハンスメントレイヤ係数を生成するために、アルゴリズムが必要である。それと同時に、このアルゴリズムは、エンハンスメントレイヤが使用可能でない時、またはたとえば電力節約などの理由から復号化器がエンハンスメントレイヤを復号化しないと判断する時に、許容できるベースレイヤビデオをもたらさなければならない。そのような処理の例示的な例の詳細を、すぐ下の標準予測符号化の短い議論の文脈で、以下にさらに述べる。

【0155】

Pフレーム（または、任意のインター符号化されたセクション）は、現行ピクチャ内の領域と基準ピクチャ内の最良一致予測領域との間の時間的な冗長度を活用することができる。基準フレーム内の最良一致予測領域の位置を、動きベクトル内で符号化することができる。現行領域と最良一致基準予測領域との間の差は、残差（または予測誤差）として知られる。

【0156】

図 3 2 は、たとえば M P E G - 4 での、P フレーム構成処理の例の図示である。処理 3 3 1 は、図 3 1 のブロック 3 2 3 で行うことができる例の処理のより詳細な図示である。処理 3 3 1 は、5 × 5 個のマクロブロックから構成された現行ピクチャ 3 3 3 を含み、ここで、この例のマクロブロックの個数は、任意である。マクロブロックは、16 × 16 個の画素から構成される。画素は、1 つの 8 ビット輝度値 (Y) および 2 つの 8 ビットクロミナンス値 (C r および C b) によって定義することができる。M P E G では、Y 成分、C r 成分、および C b 成分を、4 : 2 : 0 フォーマットで格納することができ、ここで、C r 成分および C b 成分は、X 方向および Y 方向で 2 倍ダウンサンプリングされる。したがって、各マクロブロックは、256 個の Y 成分、64 個の C r 成分、および 64 個の C b 成分からなる。現行ピクチャ 3 3 3 のマクロブロック 3 3 5 は、現行ピクチャ 3 3 3 と異なる時点の基準ピクチャ 3 3 7 から予測される。Y 値、C r 値、および C b 値に関して、符号化される現行マクロブロック 3 3 5 に最も近い最良一致マクロブロック 3 3 9 を突き止めるために、基準ピクチャ 3 3 7 内で探索を行う。基準ピクチャ 3 3 7 内の最良一致マクロブロック 3 3 9 の位置は、動きベクトル 3 4 1 内で符号化される。基準ピクチャ 3 3 7 は、復号化器が現行ピクチャ 3 3 3 の構成の前に再構成する I フレームまたは P フレームとすることができる。最良一致マクロブロック 3 3 9 が、現行マクロブロック 3 3 5 から減算され (Y 成分、C r 成分、および C b 成分のそれぞれの差が、計算される)、残差 3 4 3 がもたらされる。残差 3 4 3 は、2D 離散コサイン変換 (D C T) 3 4 5 を用いて符号化され、その後、量子化 3 4 7 される。量子化 3 4 7 は、たとえば低周波係数により多くのビットを割り当てながら高周波係数により少ないビットを割り当てることによって、空間圧縮を実現するために実行することができる。残差 3 4 3 の量子化された係数は、動きベクトル 3 4 1 および基準ピクチャ 3 3 3 を識別する情報と一緒に、現行マクロブロック 3 3 5 を表す符号化された情報である。この符号化された情報を、将来の使用のためにメモリに格納し、たとえば誤り訂正または画像の質の向上のために操作し、あるいはネットワーク 3 4 9 を介して伝送することができる。

【 0 1 5 7 】

残差 3 4 3 の符号化され量子化された係数は、符号化された動きベクトル 3 4 1 と共に、後続の動き推定および動き補償の基準フレームの一部として使用するために符号化器内で現行マクロブロック 3 3 5 を再構成するのに使用することができる。符号化器は、この P フレーム再構成のために復号化器の手順をエミュレートすることができる。復号化器のエミュレーションは、符号化器と復号化器との両方が同一の基準ピクチャを処理することをもたす。将来のインター符号化のために符号化器内で行われるものであれ、復号化器内で行われるものであれ、再構成処理をここで提示する。P フレームの再構成は、基準フレーム (または、参照されているピクチャもしくはフレームの一部) が再構成された後に開始することができる。符号化され量子化された係数が、逆量子化 3 5 1 され、その後、2D 逆 D C T すなわち I D C T 3 5 3 が実行され、復号化されたまたは再構成された残差 3 5 5 をもたす。符号化された動きベクトル 3 4 1 が、復号化され、既に再構成されている基準ピクチャ 3 3 7 内の既に再構成されている最良一致マクロブロック 3 5 7 を突き止めるのに使用される。次に、再構成された残差 3 5 5 が、再構成されたマクロブロック 3 5 9 を形成するために、再構成された最良一致マクロブロック 3 5 7 に加算される。再構成されたマクロブロック 3 5 9 を、メモリに格納するか、独立にまたは他の再構成されたマクロブロックと共にピクチャ内に表示するか、あるいは画像の質を高めるためにさらに処理することができる。

【 0 1 5 8 】

B フレーム (または双方向予測を用いて符号化された任意のセクション) は、現行ピクチャ内の領域と前のピクチャ内の最良一致予測領域と後続ピクチャ内の最良一致予測領域との間の時間的な冗長度を活用することができる。後続の最良一致予測領域および前の最良一致予測領域は、組み合わされた双方向予測領域を形成するために組み合わされる。現行ピクチャ領域と最良一致の組み合わされた双方向予測領域との間の差が、残差 (または予測誤差) である。後続の基準ピクチャ内の最良一致予測領域および前の基準ピクチャ内の

最良一致予測領域の位置を、２つの動きベクトル内で符号化することができる。

【 0 1 5 9 】

図 3 3 に、符号化器 2 2 8 によって実行できるベースレイヤ係数およびエンハンスメントレイヤ係数を符号化する符号化器の処理の例を示す。ベースレイヤおよびエンハンスメントレイヤは、S N R スケーラブルビットストリームを提供するために符号化される。図 3 3 に、図 3 1 のステップ 3 2 3 で行われるインター M B 残差係数の符号化の例を示す。しかし、類似する方法を、イントラ M B 係数の符号化にも使用することができる。図 2 の符号化器の構成要素 2 2 8 などの符号化ための手段が、図 3 3 に示された処理および図 3 2 のステップ 3 2 3 を実行することができる。オリジナルの（符号化される）ビデオデータ 4 0 6（ビデオデータは、この例ではルマ情報およびクロマ情報を備える）が、ベースレイヤ最良一致マクロブロックループ 3 0 2 およびエンハンスメントレイヤ最良一致マクロブロックループ 3 6 5 に入力される。両方のループ 3 6 3 および 3 6 5 の目的は、それぞれ加算器 3 6 7 および 3 6 9 で計算される残差を最小にすることである。ループ 3 6 3 および 3 6 5 を、図示のように並列にまたは順次実行することができる。ループ 3 6 3 および 3 6 5 は、最良一致マクロブロックとオリジナルデータ 3 6 1 との間の残差を最小にする最良一致マクロブロックを識別するために、それぞれ、基準フレームを含むバッファ 3 7 1 および 3 7 3 を探索するロジックを含む（バッファ 3 7 1 および 3 7 3 は、同一のバッファとすることができる）。ループ 3 6 3 およびループ 3 6 5 の残差は、ベースレイヤループ 3 6 3 が一般にエンハンスメントレイヤループ 3 6 5 より粗い量子化ステップサイズ（より大きい Q P 値）を利用するので、異なる。変換ブロック 3 7 5 および 3 7 7 は、各ループの残差を変換する。

【 0 1 6 0 】

次に、変換された係数が、選択器 3 7 9 内でベースレイヤ係数およびエンハンスメントレイヤ係数へと解析される。選択器 3 7 9 の解析は、下で述べるように複数の形をとることができる。解析技法の 1 つの共通の特徴は、エンハンスメントレイヤ係数 C'_{enh} が、ベースレイヤ係数 C'_{base} に対する差分改善になるように計算されることである。ベースレイヤに対する改善になるようにエンハンスメントレイヤを計算することによって、復号化器が、ベースレイヤ係数を単独で復号化し、画像の合理的な表現を有すること、またはベースレイヤ係数およびエンハンスメントレイヤ係数を組み合わせ、画像の改善された表現を有することが可能になる。次に、選択器 3 7 9 によって選択された係数が、量子化器 3 8 1 および 3 8 3 によって量子化される。量子化された係数

【数 3 0】

$$\tilde{C}'_{base}$$

【 0 1 6 1 】

および

【数 3 1】

$$\tilde{C}'_{enh}$$

【 0 1 6 2 】

（それぞれ量子化器 3 8 1 および 3 8 3 を用いて計算される）を、メモリに格納するか、ネットワークを介して復号化器に伝送することができる。

【 0 1 6 3 】

復号化器でのマクロブロックの再構成と一致するために、逆量子化器 3 8 5 が、ベースレイヤ残差係数を逆量子化する。逆量子化された残差係数が、逆変換 3 8 7 され、バッファ 3 7 1 内で見つかる最良一致マクロブロックに加算 3 8 9 され、復号化器で再構成されるものと一致する再構成されたマクロブロックがもたらされる。量子化器 3 8 3、逆量子化器 3 9 1、逆変換器 3 9 3、加算器 3 9 7、およびバッファ 3 7 3 は、ベースレイヤル

ープ 3 6 3 内で行われたものに類似するエンハンスメントループ 3 6 5 内の計算を実行する。さらに、加算器 3 9 3 は、エンハンスメントレイヤの再構成に使用される逆量子化されたエンハンスメントレイヤ係数およびベースレイヤ係数を組み合わせるのに使用される。エンハンスメントレイヤの量子化器および逆量子化器は、一般に、ベースレイヤより微細な量子化ステップサイズ（より小さい Q P ）を利用する。

【 0 1 6 4 】

図 3 4、3 5、および 3 6 に、図 3 3 の選択器 3 7 9 内で使用できるベースレイヤ係数の選択器の処理およびエンハンスメントレイヤ係数の選択器の処理の例を示す。図 2 の符号化器 2 2 8 などの選択のための手段が、図 3 4、3 5、および 3 5 に示された処理を実行することができる。図 3 4 を例として使用すると、変換された係数は、次の方程式に示されているように、ベースレイヤ係数およびエンハンスメントレイヤ係数へと解析される。

【数 3 2】

$$C'_{base} = \begin{cases} 0, & \text{if } C_{base} \text{ および } C_{enh} \text{ の符号が反対である場合} \\ \min(C_{base}, C_{enh}), & \text{それ以外場合} \end{cases} \quad [25]$$

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

【 0 1 6 5 】

ここで、「min」関数は、2 つの引数の数学的な最小値または最小の大きさのいずれかとすることができる。式 2 5 は、ブロック 4 0 1 として図示され、式 2 6 は、図 3 4 で加算器 5 1 0 として図示されている。式 2 6 では、 Q_b が、ベースレイヤ量子化器 3 8 1 を表し、 Q_b^{-1} が、ベースレイヤの逆量子化器 3 8 5 を表す。式 2 6 は、エンハンスメントレイヤ係数を、式 2 5 を用いて計算されたベースレイヤ係数の差分改善に変換する。

【 0 1 6 6 】

図 3 5 は、ベースレイヤおよびエンハンスメントレイヤの係数の選択器 3 7 9 のもう 1 つの例を示す。この例では、ブロック 4 0 5 に含まれる方程式（.）が、次式を表す。

【数 3 3】

$$C'_{base} = \begin{cases} C_{base}, & |Q_b^{-1}Q_b(C_{base}) - C_{enh}| < |C_{enh}| \text{ の場合} \\ 0, & \text{それ以外の場合} \end{cases} \quad [27]$$

【 0 1 6 7 】

加算器 4 0 7 は、次の 2 つの方程式に示されているようにエンハンスメントレイヤ係数を計算する。

【数 3 4】

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

【 0 1 6 8 】

ここで、 C'_{base} は、式 2 7 によって与えられる。

【 0 1 6 9 】

図 3 6 は、ベースレイヤおよびエンハンスメントレイヤの選択器 3 7 9 のもう 1 つの例の図示である。この例では、ベースレイヤ係数は変更されず、エンハンスメントレイヤは、量子化 / 逆量子化されたベースレイヤ係数とオリジナルのエンハンスメントレイヤ係数との間の差と等しい。

【 0 1 7 0 】

ベースレイヤおよびエンハンスメントレイヤの残差係数に加えて、復号化器は、MB がどのように符号化されているかを識別する情報を必要とする。図 2 の符号化器の構成要素

10

20

30

40

50

228などの符号化ための手段は、たとえば、マクロブロック（またはサブマクロブロック）がイントラ符号化またはインター符号化されたとして識別され（たとえば順方向、逆方向、または双方向を含む、どのタイプのインター符号化かをも識別する）、インター符号化された部分がどのフレームに対して参照されるかを識別するMBマップなど、イントラ符号化された部分およびインター符号化された部分の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。例の態様では、MBマップおよびベースレイヤ係数が、ベースレイヤ内で符号化され、エンハンスメントレイヤ係数が、エンハンスメントレイヤ内で符号化される。

【0171】

PフレームおよびBフレームは、イントラ符号化されたMBならびにインターMBを含むことができる。ハイブリッドビデオ符号化器が、PフレームまたはBフレーム内のあるマクロブロックをイントラ符号化されたMBとして符号化すると判断するために速度歪み（RD）最適化を使用することが、一般的である。イントラ符号化されたMBがエンハンスメントレイヤのインターMBに依存しない単一レイヤの復号化を有するために、すべての隣接するインターMBは、ベースレイヤのイントラ符号化されたMBの空間的な予測には使用されない。エンハンスメントレイヤの復号化について計算の複雑度を変更されないままに保つために、ベースレイヤのPフレームまたはBフレーム内のイントラ符号化されたMBについて、エンハンスメントレイヤでの改善をスキップすることができる。

【0172】

PフレームまたはBフレーム内のイントラ符号化されたMBは、インターMBよりはるかに多数のビットを必要とする場合がある。この理由のために、PフレームまたはBフレーム内のイントラ符号化されたMBは、より高いQPでベースレイヤ品質で符号化することしかできない。これは、ビデオ品質のある劣化をもたらすが、この劣化は、上で述べたようにベースレイヤおよびエンハンスメントレイヤ内のインターMB係数を用いて後のフレームで改善される場合に、人目をひかないに違いない。2つの理由が、この劣化を人目をひかないものにする。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を改善することである。第1フレームから第2フレームへ位置を変更するオブジェクトがあると、第1フレーム内のいくつかの画素は、第2フレーム内で不可視であり（隠される情報）、第2フレーム内のいくつかの画素は、初めて可視になる（覆いを取られる情報）。人間の目は、覆いを取られる画像情報および隠される画像情報に敏感ではない。したがって、覆いを取られる情報について、それがより低い品質で符号化される場合であっても、目は、その差を見分けることができない。同一の情報が、続くPフレームに留まる場合に、エンハンスメントレイヤはより低いQPを有するので、続くPフレームが、エンハンスメントレイヤでその情報を改善できる可能性が高い。

【0173】

PフレームまたはBフレームにイントラ符号化されたMBを導入するもう1つの一般的な技法は、イントラリフレッシュとして知られる。この場合に、標準R-D最適化が、いくつかのMBをインター符号化されたMBにしなければならないと命ずる場合であっても、それらのMBは、イントラ符号化されたMBとして符号化される。これらのイントラ符号化されたMBは、ベースレイヤに含まれるが、QP₀またはQP₁のいずれかを用いて符号化することができる。QP₀がベースレイヤに使用される場合には、エンハンスメントレイヤでの改善は不要である。QP₀がベースレイヤに使用される場合には、改善が必要である場合があり、さもなければ、エンハンスメントレイヤで、品質の低下が顕著になる。インター符号化は、符号化効率の意味でイントラ符号化より効率的なので、エンハンスメントレイヤでのこれらの改善は、インター符号化される。この形では、ベースレイヤ係数は、エンハンスメントレイヤに使用されない。したがって、品質は、新しい動作を導入せずにエンハンスメントレイヤで改善される。

【0174】

Bフレームは、一般に、それが提供する高い圧縮品質のゆえに、エンハンスメントレイ

10

20

30

40

50

ヤで使用される。しかし、Bフレームが、Pフレームのイントラ符号化されたMBを参照しなければならない場合がある。Bフレームの画素が、エンハンスメントレイヤ品質で符号化されなければならない場合に、これは、上で述べたように、Pフレームのイントラ符号化されたMBのより低い品質に起因して、多すぎるビットを必要とする場合がある。上で述べたHVSの特質を利用することによって、BフレームMBを、Pフレームのより低い品質のイントラ符号化されたMBを参照する時に、より低い品質で符号化することができる。

【0175】

PフレームまたはBフレーム内のイントラ符号化されたMBの1つの極端な例が、PフレームまたはBフレーム内のすべてのMBが、符号化されるビデオ内のシーン変化の存在に起因してイントラモードで符号化される時である。この場合に、フレーム全体を、ベースレイヤ品質で、エンハンスメントレイヤでの改善なしで符号化することができる。シーン変化がBフレームで発生する場合に、Bフレームがエンハンスメントレイヤでのみ符号化されると仮定すると、そのBフレームを、ベースレイヤ品質で符号化するか、単純に捨てることができる。シーン変化がPフレームで発生する場合に、変更が不要である可能性があるが、そのPフレームを、捨てるか、ベースレイヤ品質で符号化することができる。スケーラブルレイヤ符号化は、本願の譲受人が所有し、参照によってその全体が本明細書に組み込まれている、同時係属の米国特許出願[弁理士整理番号第050078号]、名称「SCALABLE VIDEO CODING WITH TWO LAYER ENCODING AND SINGLE LAYER DECODING」にさらに記載されている。

【0176】

符号化器の第1パス部分

図7に、図2の符号化器228の例示的な例を示す。図示のブロックは、符号化器228内に含めることができるさまざまな符号化器の処理を示す。この例では、符号化器228は、境界線704の上の第1パス部分706および線704の下の第2パス部分706(図2の第2パス符号化器232および再符号化器234の機能を含む)を含む。

【0177】

符号化器228は、プリプロセッサ226からメタデータおよび生ビデオを受け取る。メタデータには、ビデオのコンテンツ情報に関するメタデータを含む、プリプロセッサ226によって受け取られたか計算されたすべてのメタデータを含めることができる。符号化器228の第1パス部分702は、その機能の観点から下記で説明される第1パス符号化702に含めることができる例示的処理を図解する。当業者が知るとおり、そのような機能性は、さまざまな形(たとえば、ハードウェア、ソフトウェア、ファームウェア、またはこれらの組合せ)で実施することができる。

【0178】

図7に、適応的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 inputs を供給する。第1パス部分702には、メタデータおよびビデオを受け取り、そのビデオに関するコンテンツ情報を判定するように構成されたコンテンツ分類モジュール712をも含めることができる。コンテンツ情報を、メタデータおよびビデオをも受け取るレート制御ビット割当モジュール714に供給することができる。制御ビット割当モジュール714は、レートビット制御情報を決定し、これをモード判断モジュール715に供給する。コンテンツ情報およびビデオを、イントラ符号化された歪み情報をモード判断モジュール715ならびにベースおよびエンハンスメントレイヤ用のスケーラビリティ・レート歪みモジュール718に供給するイントラモデル(歪み)モジュール716に供給することができる。ビデオおよびメタデータは、インター符号化された歪み情報をベースおよびエンハンスメントレイヤ用のスケーラビリティ・レート歪みモジュール718に供給する動き推定(歪み)モジュール720に供給される。ベースおよびエンハンスメントレイヤ用のスケーラビリティ・レート歪みモジュール718は、動き

推定モジュール 720 およびイントラモデル歪みモジュール 716 からの歪み推定値を使用してスケラビリティ・レート歪み情報を決定し、このスケラビリティ・レート歪み情報は、モード判断モジュール 715 に供給される。モード判断モジュール 715 は、スライス / MB 順序付けモジュール 722 から入力を受け取る。スライス / MB 順序付けモジュール 722 は、誤り弾力性モジュール 740 (第 2 パス部分 706 内に図示) から入力を受け取り、ビデオの独立に符号化可能な部分 (スライス) を誤り弾力性に関するアクセス単位境界に整列することに関する情報をモード判断モジュール 715 に供給する。モード判断モジュール 715 は、その入力に基づいて符号化モード情報を判定し、「最良」符号化モードを第 2 パス部分 706 に供給する。そのような第 1 パス部分 702 符号化のいくつかの例のさらなる例示的な説明を、下で説明する。

10

【0179】

上で述べたように、コンテンツ分類モジュール 712 は、プリプロセッサ 226 によって供給されるメタデータおよび生ビデオを受け取る。いくつかの例で、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、そのコンテンツ情報をコンテンツ分類モジュール 712 に供給し (たとえば、メタデータ内で)、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コンテンツ分類を判定するのに、そのコンテンツ情報を使用することができる。いくつかの他の態様において、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さまざまなコンテンツ情報を決定するように構成され、コンテンツ分類を決定するようにも構成され得る。

20

【0180】

コンテンツ分類モジュール 712 を、異なるタイプのコンテンツを有するビデオの異なるコンテンツ分類を判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の符号化の諸態様、たとえば、量子化パラメータを決定するためのビットレート (たとえば、ビット割当) の決定、動き推定、スケラビリティ、誤り弾力性、チャンネルにまたがる最適マルチメディアデータ品質の維持、および高速なチャンネルの切替方式 (たとえば、高速なチャンネルの切替を可能にするための周期的な I フレームの強制に使用される異なるパラメータをもたらすことができる。一例によれば、符号化器 228 は、コンテンツ分類に基づいてレート歪み (R - D) 最適化およびビットレート割当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く所望のビットレートに対応する所与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。また、マルチメディアデータのコンテンツを分類する (たとえば、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、受け取るデバイスのディスプレイ上での通信されたマルチメディアデータの結果の知覚品質が、ビデオコンテンツに依存するようにされる。

30

【0181】

コンテンツを分類するためにコンテンツ分類モジュール 712 が経験する手順の例として、図 9 に、コンテンツ分類モジュール 712 がそれによって動作できる例示的処理を示す処理 900 を示す。図示されているように、処理 900 は、入力ブロック 902 で開始され、ここで、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およびメタデータを受け取る。次に、処理 900 は、ブロック 904 に進み、ここで、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情報および時間的情報を判定する。いくつかの態様で、空間的情報および時間的情報は、空間的マスキングおよび時間的マスキング (たとえば、フィルタリング) によって判定される。空間的情報および時間的情報は、シーン変化データおよび動きベクトル (MV) 平滑化を含むメタデータに基づいて判定することができる。次に、処理 900 は、ブロック 912 に進み、ここで、空間的な複雑度の推定、時間的な複雑度の推定、および感度推定を実行する。次に、処理 900 は、ブロック 916 に進み、ここで、マルチメディアデータのコンテンツを、ブロック 904 および 912 で判定された空間的データ、時間的データ、および感度データの結果に基づいて分類する。また、ブロック 916 で、特定のレート歪み (R - D) 曲線を選択することができる。かつ / または R - D 曲線データを更新することができる。次に、処理 90

40

50

0 は、出力ブロック 9 1 8 に進み、ここで、出力には、複雑度 - 歪みマップまたは空間的アクティビティおよび時間的アクティビティを示す値（たとえば、コンテンツ分類）ならびに / あるいは選択された R - D 曲線を含めることができる。戻って図 7 を参照すると、コンテンツ分類モジュール 7 1 2 は、出力を、上で述べたレート制御ビット割当モジュール 7 1 4、イントラモデル（歪み）モジュール 7 1 6、および I フレームインスタンス化モジュール 7 0 8 に供給する。

【 0 1 8 2 】

コンテンツ情報

コンテンツ分類モジュール 7 1 2 を、以下でさらに説明する、空間的な複雑度、時間的な複雑度、コントラスト比値、標準偏差、およびフレーム差分メトリックを含むさまざまなコンテンツ関連メトリックを含むさまざま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。

10

【 0 1 8 3 】

コンテンツ分類モジュール 7 1 2 を、マルチメディアデータの空間的な複雑度および時間的な複雑度を判定し、また、テクスチャ値を空間的な複雑度に関連付け、動き値を時間的な複雑度に関連付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 7 1 2 は、プリプロセッサ 2 2 6 から、符号化されるマルチメディアデータのコンテンツに関する前処理されたコンテンツ情報を受け取り、あるいはその代わりに、コンテンツ情報を計算するようにプリプロセッサ 2 2 6 を構成することができる。上で説明したように、コンテンツ情報には、たとえば、1 つまたは複数の $D_{c s a t}$ 値、コントラスト比値、動きベクトル（ MV ）、および絶対差の合計（ SAD ）を含めることができる。

20

【 0 1 8 4 】

一般に、マルチメディアデータは、画像またはフレームの 1 つまたは複数のシーケンスを含む。各フレームを、処理のために画素のブロックに分解することができる。空間的な複雑度は、1 フレーム内の空間的な詳細のレベルの尺度を全般的に記述する広義の用語である。輝度およびクロミナンスの主に平坦なまたは変化しないまたは少し変化する区域を有するシーンは、低い空間的な複雑度を有する。空間的な複雑度は、ビデオデータのテクスチャに関連する。空間的な複雑度は、この態様では、局所空間周波数および周囲照明の関数としてブロックごとに計算される、 $D_{c s a t}$ と呼ばれる人間の視覚感度メトリックに基づく。当業者は、人間の視覚系を利用するために視覚的画像の空間周波数パターン、照明特性、およびコントラスト特性を使用する技法を知っている。複数の感度メトリックが、人間の視覚系の知覚限度を利用するために既知であり、本明細書で説明される方法と共に使用することができる。

30

【 0 1 8 5 】

時間的な複雑度は、フレームのシーケンス内のフレームの間で参照されるマルチメディアデータ内の動きのレベルの尺度を全般的に記述するのに使用される広義の用語である。動きがほとんどまたは全くないシーン（たとえば、ビデオデータのフレームのシーケンス）は、低い時間的な複雑度を有する。時間的な複雑度は、マクロブロックごとに計算することができ、 $D_{c s a t}$ 値、動きベクトル、およびあるフレームと別のフレーム（たとえば、基準フレーム）との間の絶対画素差の合計に基づくものとすることができる。

40

【 0 1 8 6 】

フレーム差分メトリックは、予測器と現行マクロブロックとの間の絶対差の合計（ SAD ）として表される残差エネルギーと共に、動きの量を考慮に入れた 2 つの連続するフレームの間の差の尺度（たとえば、動きベクトルすなわち MV ）を与える。フレーム差は、双方向または単一方向の予測効率の尺度をも提供する。

【 0 1 8 7 】

動き補償されたデインターレーシングを潜在的に実行するプリプロセッサから受け取られた動き情報に基づくフレーム差分メトリックの一例は、次の通りである。デインターレーサは、双方向動き推定を実行し、したがって、双方向動きベクトルおよび SAD 情報が、使用可能である。各マクロブロックの SAD_MV によって表されるフレーム差は、次

50

のように導出することができる。

【数 3 5】

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

【0188】

ここで、 $\text{MV} = \text{Square_root}(\text{MV}_x^2 + \text{MV}_y^2)$ 、 $\text{SAD} = \min(\text{SAD}_N, \text{SAD}_P)$ であり、 SAD_N は、逆方向の基準フレームから計算されたSADであり、 SAD_P は、順方向の基準フレームから計算されたSADである。

【0189】

フレーム差を推定するもう1つの手法を、上で式6～8を参照して説明した。SAD比（またはコントラスト分配）を、前に上で式6で説明した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、式7を使用して計算される。フレーム差分メトリックDは、式8に示されているように計算することができる。

【0190】

1つの例示的な例で、コントラスト比およびフレーム差分メトリックは、次の形で、所与のビデオシーケンス内の特徴を信頼できる形で予測できるビデオコンテンツ分類を入手するのに利用される。ここでは符号化器228内で発生するものとして説明するが、コンテンツ分類（または他のコンテンツ情報）を決定し、そのコンテンツ分類をメタデータを介して符号化器228に渡すように、プリプロセッサ226を構成することもできる。下の例で説明する処理は、コンテンツを、R-D曲線ベースの分析から得られる分類と同様の8つの可能なクラスに分類する。この分類処理は、スーパーフレームごとに、シーンの複雑度およびそのスーパーフレーム内でのシーン変化発生回数に依存して、0と1との間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値およびフレーム差分値からコンテンツ分類メトリックを得るために、スーパーフレームごとに次のステップ(1)～(5)を実行することができる。

【0191】

1. マクロブロックコントラスト値から平均フレームコントラストおよびフレームコントラスト偏差を計算する。

【0192】

2. それぞれ40および5である、シミュレーションから得られた値を使用して、フレームコントラスト値およびフレーム差分値を正規化する。

【0193】

3. たとえば一般化された方程式

【数 3 6】

$$\begin{aligned} \text{CCMetric} &= \text{CCW1} * \text{I_Frame_Contrast_Mean} + \text{CCW2} \\ &\quad * \text{Frame_Difference_Mean} - \\ &\quad \text{CCW3} * \text{I_Contrast_Deviation}^2 * \exp(\text{CCW4} * \text{Frame_Difference_Deviation}^2) \end{aligned}$$

[30]

【0194】

を使用して、コンテンツ分類メトリックを計算する。ここで、CCW1、CCW2、CCW3、およびCCW4は、重み付け係数である。この例では、値は、CCW1にいて0.2、CCW2について0.9、CCW3について0.1、CCW4について-0.00009になるように選択される。

【0195】

4. スーパーフレーム内のシーン変化の回数を判定する。一般に、スーパーフレームとは、特定の時間期間内に表示できるピクチャまたはフレームのグループを指す。通常、こ

10

20

30

40

50

の時間期間は1秒である。いくつかの態様において、スーパーフレームは、30個のフレームを備える(30/fpsビデオの場合)。他の態様において、スーパーフレームは、24個のフレームを備える(24/fpsビデオ)。シーン変化の回数に応じて、次のケースのうちの1つが実行される。

【0196】

(a) シーン変化なし スーパーフレーム内にシーン変化がない時には、メトリックは、次の方程式に示されているように、完全にフレーム差分値だけに依存する。

【数37】

$$\text{CCMetric} = (\text{CCW2} + (\text{CCW1}/2)) * \text{Frame_Difference_Mean} - (\text{CCW3} - (\text{CCW1}/2)) * 1 * \exp(-\text{CCW4} * \text{Frame_Difference_Deviation}^2)$$

10

[31]

【0197】

(b) 単一のシーン変化 スーパーフレーム内で観察された単一のシーン変化がある時には、下に示されているように、デフォルト方程式がメトリックの計算に使用される。

【数38】

$$\text{CCMetric} = \text{CCW1} * \text{I_Frame_Contrast_Mean} + \text{CCW2} * \text{Frame_Difference_Mean} - \text{CCW3} * \text{I_Contrast_Deviation}^2 * \exp(\text{CCW4} * \text{Frame_Difference_Deviation}^2)$$

20

[32]

【0198】

(c) 2つのシーン変化 所与のスーパーフレーム内に最大で2つのシーン変化があることが観察される時には、次の方程式に示されているように、最後のスーパーフレームに最初のスーパーフレームより大きい重みを与える。というのは、いずれにせよ最初のスーパーフレームが後者よりすばやくリフレッシュされるからである。

【数39】

$$\text{CCMetric} = 0.1 * \text{I_Frame_Contrast_Mean1} + \text{CCW1} * \text{I_Frame_Contrast_Mean2} + (\text{CCW2} - 0.1) * \text{Frame_Difference_Mean} - \text{CCW3} * \text{I_Contrast_Deviation1}^2 * \text{I_Contrast_Deviation2}^2 * \exp(\text{CCW4} * \text{Frame_Difference_Deviation}^2)$$

30

[33]

【0199】

(d) 3つ以上のシーン変化 所与のスーパーフレームが3つ以上のIフレーム(たとえば、N個)を有することが観察される場合には、次の方程式に示されているように、最後のIフレームが、より大きい重みを与えられ、他のすべてのIフレームは、0.05の重みを与えられる。

40

【数40】

$$\text{CCMetric} = 0.05 * \text{I_Frame_Contrast_Mean}_{(1...N-1)} + \text{CCW1} * \text{I_Frame_Contrast_Mean}_{(N)} + (\text{CCW2} - (0.05 * (N-1))) * \text{Frame_Difference_Mean} - \text{CCW3} * \text{I_Contrast_Deviation}_{(N)}^2 * \text{I_Contrast_Deviation}_{(1...N-1)}^2 * \exp(\text{CCW4} * \text{Frame_Difference_Deviation}^2)$$

[34]

50

【0200】

5. フレーム差分の平均値が0.05未満の時の小さい動きのシーンの場合に、メトリックについて補正を使用することができる。(CCOFFSET)0.33のオフセットを、CCMetricに加算する。

【0201】

コンテンツ分類モジュール712は、マクロブロック(またはビデオデータの指定された量)の空間的な複雑度を示す値を決定するのに、D_{class}値、動きベクトル、および/または絶対差の合計を使用する。時間的な複雑度は、フレーム差分メトリック(動きベクトルを用いて動きの量を考慮に入れた2つの連続するフレームの間の差およびそれらのフレームの間の絶対差の合計)の測定によって決定される。

10

【0202】

いくつかの態様において、コンテンツ分類モジュール712を、帯域幅マップを生成するように構成することができる。たとえば、プリプロセッサ226が帯域幅マップを生成しない場合に、帯域幅マップ生成を、コンテンツ分類モジュール712によって実行することができる。

【0203】

テクスチャ値および動き値の判定

マルチメディアデータ内のマクロブロックごとに、コンテンツ分類モジュール712は、テクスチャ値を空間的な複雑度に関連付け、動き値を時間的な複雑度に関連付ける。テクスチャ値は、マルチメディアデータのルミネセンス値に関連し、低いテクスチャ値は、データの隣接する画素のルミネセンス値の小さい変化を示し、高いテクスチャ値は、データの隣接する画素のルミネセンス値の大きい変化を示す。テクスチャ値および動き値が一旦計算されたならば、コンテンツ分類モジュール712は、動き情報とテクスチャ情報と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、分類されるビデオデータのテクスチャを、相対テクスチャ値、たとえば、一般にマクロブロックの輝度値の複雑度を示す「低」テクスチャ、「中」テクスチャ、または「高」テクスチャに関連付ける。また、コンテンツ分類モジュール712は、分類されるビデオデータについて計算された動き値を、相対動き値、たとえば、一般にマクロブロックの動きの量を示す「低」動き、「中」動き、または「高」動きに関連付ける。代替の態様では、動きおよびテクスチャに関するより少数またはより多数のカテゴリを使用することができる。次に、コンテンツ分類メトリックが、関連するテクスチャ値および動き値を考慮することによって決定される。

20

30

【0204】

図8に、テクスチャ値および動き値をコンテンツ分類にどのように関連付けるかを示す分類チャートの例を示す。当業者は、たとえばルックアップテーブルにおいて、またはデータベースにおいてなど、そのような分類チャートを実施する多数の形に馴染みがある。分類チャートは、ビデオデータコンテンツの所定の評価に基づいて生成される。ビデオデータ分類を決定するために、「低」、「中」、または「高」のテクスチャ値(「x軸」上)が、「低」、「中」、または「高」の動き値(「y軸」上)と相互参照される。交差するブロックに示されたコンテンツ分類が、ビデオデータに割り当てられる。たとえば、「高」のテクスチャ値および「中」の動き値は、7という分類をもたらす。図8は、この例では8つの異なるコンテンツ分類に関連する相対的なテクスチャ値および動き値のさまざまな組合せを示す。いくつかの他の態様では、より多数またはより少数の分類を使用することができる。コンテンツ分類の例示的態様のさらなる説明が、本願の譲受人に譲渡され、参照によって本明細書に組み込まれている、同時係属の2006年3月11日出願の米国特許出願第11/373,577号、名称「Content Classification for Multimedia Processing」に開示されている。

40

【0205】

レート制御ビット割当

50

本明細書で説明するように、マルチメディアデータのコンテンツ分類は、符号化アルゴリズム内で、ビデオの一定の知覚品質を維持しながらビット管理を効果的に改善するのに使用することができる。たとえば、分類メトリックは、アルゴリズム内で、シーン変化検出、符号化ビットレート割当制御、およびフレームレートのアップコンバージョン（FRUC）に使用することができる。コンプレッサ/デコンプレッサ（コーデック）システムおよびデジタル信号処理アルゴリズムは、ビデオデータ通信に一般に使用され、帯域幅を節約するように構成することができるが、品質と帯域幅節約との間にトレードオフがある。最良のコーデックは、ビデオ品質の最小の劣化をもたらしながら最大の帯域幅節約を実現する。

【0206】

1つの例示的な例で、レート制御ビット割当モジュール714は、ビットレート（たとえば、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）を決定するのに、ならびに符号化器228の他の処理および構成要素による使用のためにビットレートを格納するのに、コンテンツ分類を使用する。ビデオデータの分類から決定されるビットレートは、一貫した品質レベルでマルチメディアデータを供給しながら帯域幅を節約するのに助けることができる。一態様において、異なるビットレートを、8つの異なるコンテンツ分類のそれぞれに割り当てることができ、その後、そのビットレートが、マルチメディアデータの符号化に使用される。結果として生じる効果は、マルチメディアデータの異なるコンテンツ分類が、符号化に異なるビット数を割り当てられるが、知覚される品質が、ディスプレイ上で見られる時に同様であるが一貫することである。

【0207】

一般に、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、より高いレベルの動きおよび/またはテクスチャを示し、符号化される時により多くのビットを割り当てられる。より低い分類（より少ないテクスチャおよび動きを示す）を有するマルチメディアデータは、より少ないビットを割り当てられる。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、ビットレートを、マルチメディアデータを見ることに関する選択されたターゲットの知覚される品質レベルに基づいて決定することができる。マルチメディアデータ品質の決定は、人間がその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。いくつかの代替の態様において、マルチメディアデータの品質の推定は、たとえば信号対雑音比アルゴリズムを使用する、自動試験システムによって行うことができる。一態様において、標準品質レベルの組（たとえば、5つ）および各特定の品質レベルを達成するのに必要な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ついて事前に決定される。品質レベルの組を決定するために、特定のコンテンツ分類のマルチメディアデータを、あるビットレートを使用して符号化された時のマルチメディアデータの視覚的に知覚される品質の数値表示を提供する平均意見得点（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から5までの範囲内の単一の数として表すことができ、ここで、1は、最低の知覚される品質であり、5は、最高の知覚される品質である。他の態様では、MOSが、5つより多くのまたは5つより少ない品質レベルを有することができ、各品質レベルの異なる記述を、使用することができる。

【0208】

マルチメディアデータ品質の決定は、人間がその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。いくつかの代替の態様において、マルチメディアデータ品質の推定は、たとえば信号対雑音比アルゴリズムを使用する、自動試験システムによって行うことができる。一態様で、標準品質レベルの組（たとえば、5つ）および各特定の品質レベルを達成するのに必要な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ついて事前に決定される。

【0209】

あるコンテンツ分類のマルチメディアデータの、視覚的に知覚される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（たとえば所望の）品質レベルを選択するこ

10

20

30

40

50

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のに使用される目標品質レベルは、事前に選択するか、ユーザによって選択されるか、ユーザからもしくは別の処理からの入力が必要とする自動処理もしくは半自動処理を介して選択されるか、あるいは所定の判断基準に基づいて符号化デバイスもしくは符号化システムによって動的に選択されることができる。目標品質レベルは、たとえば、符号化アプリケーションのタイプまたはマルチメディアデータを受け取るクライアントデバイスのタイプに基づいて選択することができる。

【0210】

図7に示された例では、レート制御ビット割当モジュール714は、コンテンツ分類モジュール712からのデータとプリプロセッサ226から直接のメタデータとの両方を受け取る。レート制御ビット割当モジュール714は、符号化器228の第1パス部分に存在し、レート制御微調整モジュール738は、第2パス部分706に存在する。この2パスのレート制御の態様は、第1パス（レート制御ビット割当モジュール714）が、1つのスーパーフレームのルックアヘッドを用いてコンテキストに適應したビット割当を実行し（たとえば、256 kbpsの長期平均ビットレートを目標値にして）、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）が、2レイヤのスケラビリティのために第1パスの結果を改善し、レート適應を実行するように構成される。レート制御は、（1）GOPレベル（1つのGOPの内部で不均一になるようにIフレーム、Pフレーム、Bフレーム、およびFフレームのビット分配を制御する）、（2）スーパーフレームレベル（最大スーパーフレームサイズに対するハード限界を制御する）、（3）フレームレベル（コンテンツ情報（たとえば、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な複雑度および時間的な複雑度に従ってビット要件を制御する）、および（4）マクロブロックレベル（コンテンツ情報（たとえば、コンテンツ分類）に基づく空間的な複雑度のマップおよび時間的な複雑度のマップに基づいてマクロブロックのビット割当を制御する）という4つのレベルで動作する。

【0211】

レート制御モジュール714の動作の例示的な流れ図を、図10に示す。図10に示されているように、処理1000は、入力1002ブロックで開始される。レート制御モジュール714は、必ずしもすべてが図7によって示されてはいないさまざまな入力を受け取る。たとえば、入力情報には、プリプロセッサ226からのメタデータ、目標ビットレート、符号化器のバッファサイズ（あるいは、同等物として、レート制御に関する最大遅延時間）、初期のレート制御の遅延、およびフレームレート情報を含めることができる。さらなる入力情報に、たとえば、最大スーパーフレームサイズ、GOPの長さおよびP/Bフレーム分布（シーン変化情報を含む）、所望のベースレイヤおよびエンハンスメントレイヤの配置、将来の30フレームに関するGOP内のピクチャの複雑度 - 歪みメトリックを含む、グループオブピクチャ（GOP）レベルでの入力を含めることができる。他の入力情報は、現行ピクチャの複雑度 - 歪みマップ（コンテンツ分類モジュール712から受け取られる）、量子化パラメータ（QP）、および過去の30フレーム（スライディングウィンドウにおさまる）のビット分割を含むピクチャレベルでの入力を含む。最後に、マクロブロック（MB）レベルでの入力は、たとえば、基準ピクチャ内の並べて置かれたマクロブロック（MB）の平均絶対差（MAD）および量子化後（スキップされたか否かに関わりなく）のマクロブロックの符号化されたブロックパターン（CBP）を含む。

【0212】

ブロック1002での入力の後に、処理1000は、ビットストリームの符号化に関する初期化のためにブロック1004に進む。それと同時に、バッファ初期化1006が実行される。次に、ブロック1008に示されているように、GOPを初期化し、GOPビット割当1010が、初期化の一部として受け取られる。GOP初期化の後に、流れは、ブロック1012に進み、ここで、スライスを初期化する。この初期化は、ブロック1014によって示されるように、ヘッダビットの更新を含む。ブロック1004、1008、および1012の初期化が実行された後に、基本単位またはマクロブロック（MB）の

レート制御 (RC) が、ブロック 1016 によって示されているように実行される。ブロック 1016 でのマクロブロックのレート制御の決定の一部として、入力が、符号化器 228 内のインターフェースを介して受け取られる。この入力には、マクロブロック (MB) ビット割当 1018、二次モデルパラメータの更新 1020、および中央値からの中央値絶対偏差 (「MAD」、ばらつきの堅牢な推定値) パラメータの更新 1022 を含めることができる。次に、処理 1000 は、1 つのピクチャを符号化した後の動作 1024 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け取ることを含む。次に、処理 1000 は、出力ブロック 1028 に進み、ここで、レート制御モジュール 714 は、図 7 に示されたモード判断モジュール 715 によって使用される各マクロブロック MB の量子化パラメータ QP を出力する。

10

【0213】

動き推定

動き推定モジュール 720 は、プリプロセッサ 226 からメタデータおよび生ビデオの入力を受け取り、ブロックサイズ、動きベクトル歪みメトリックス、および基準フレーム識別子を含めることができる出力をモード判断モジュール 715 に供給する。図 11 に、動き推定モジュール 720 の例示的動作を示す。図に示されているように、処理 1100 は、入力 1102 から開始される。フレームレベルでは、モジュール 720 は、基準フレーム ID および動きベクトルの入力を受け取る。マクロブロックレベルでは、入力 1102 は、入力画素および基準フレーム画素を含む。処理 1100 は、ステップ 1104 に継続し、ここで、カラー動き推定 (ME) および動きベクトル予測を実行する。この処理を実行するために、MPEG-2 動きベクトルおよびルマ動きベクトル MV 1106、動きベクトル平滑化 1108、ならびに非因果的動きベクトル (non-causal motion vector) 1110 を含むさまざまな入力が、受け取られる。次に、処理 1100 は、ブロック 1112 に進み、ここで、六角形探索法または菱形探索法などの動きベクトル探索アルゴリズムまたは動きベクトル探索方法論が実行される。ブロック 1112 での処理への入力には、ブロック 1114 によって示されているように、絶対差の合計 (SAD)、二乗差の合計 (SSD)、および / または他のメトリックを含めることができる。動きベクトル探索が実行されたならば、処理 1100 は、終了ブロック 1116 に進み、ここで終了処理が実行される。その後、処理 1000 は、ブロックサイズ、動きベクトル (MV)、歪みメトリックス、および基準フレーム識別子の出力をもたらす出力ブロック 1118 で終了する。

20

30

【0214】

ベースレイヤおよびエンハンスメントレイヤのスケラビリティ R-D

図 13 に、スケラビリティ R-D モジュール 718 によって実行できるスケラビリティ処理 1300 の例示的な流れ図を示す。処理 1300 は、開始ブロック 1302 で開始され、ブロック 1304 に進み、ここで、スケラビリティ R-D モジュール 718 は、動き推定モジュール 720 から入力を受け取り、動き推定を実行する。動き推定は、ブロック 1306 によって示されているように、ベースレイヤ基準フレーム、エンハンスメントレイヤ基準フレーム、および符号化されるオリジナルフレームの入力に頼る。そのような情報は、GOP パーティション 612 によって計算し、たとえばメタデータを介してスケラビリティ R-D モジュール 718 に通信することができる。処理 1300 は、ブロック 1308 に進んで、ベースレイヤデータおよびエンハンスメントレイヤ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 に示されているようにベースレイヤ符号化を実行し、これにブロック 1312 のエンハンスメントレイヤ符号化が続く。エンハンスメントレイヤの符号化は、ブロック 1314 によって示されるように入力としてレイヤ間予測のためにベースレイヤ符号化結果を使用することができ、したがって、時間的には、ベースレイヤ符号化の後に実行される。これは、同時係属の米国特許出願 [弁理士整理番号第 050078 号]、名称「SCALABLE VIDEO CODING WITH TWO LAYER ENCODING AND SINGLE LAYE

40

50

R D E C O D I N G」にさらに記載されている。符号化が完了した後に、処理 1 3 0 0 は、ブロック 1 3 1 6 で終了する。

【 0 2 1 5 】

スライス / マクロブロック順序付け

第 1 パス部分 7 0 2 は、第 2 パス部分内の誤り弾力性モジュール 7 4 0 から入力を受け取り、モード判断モジュール 7 1 5 にスライス整列情報を供給するスライス / マクロブロック順序付けモジュール 7 2 2 をも含む。スライスとは、独立に復号化可能な（エントロピ復号化）符号化されたビデオデータの大きい塊である。アクセス単位（A U）は、必ず正確に 1 つの主符号化ピクチャを含む N A L 単位の組をそれぞれが含む符号化されたビデオフレームである。主符号化ピクチャに加えて、アクセス単位には、1 つまたは複数の冗長な符号化ピクチャまたはスライスを含まない他の N A L 単位または符号化ピクチャのスライスデータパーティションをも含めることができる。アクセス単位の復号化は、必ず、復号化されたピクチャをもたらす。

10

【 0 2 1 6 】

フレームを、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されたブロック（T D Mカプセルと称する）とすることができる。スーパーフレームは、時間の 1 単位（たとえば、1 秒）に対応し、4 つのフレームを含む。時間領域でのフレーム境界へのスライス境界および A U 境界の整列は、破壊されたデータの最も効率的な分離および局所化をもたらす。深いフェード中に、T D Mカプセル内の連続データのほとんどが、誤りによって影響される。時間ダイバーシティに起因して、残りの T D Mカプセルは、変更されないままであることの高い確率を有する。破壊されていないデータは、影響された T D Mカプセルから失われたデータを回復し、隠すのに利用することができる。同様の論理が、周波数領域多重化（F D M）にあてはまり、ここでは、データシンボルが変調する周波数サブキャリア内の分離を介して周波数ダイバーシティが達成される。さらに、同様の論理が、しばしば無線ネットワーク内で適用される空間的（送信器アンテナおよび受信器アンテナの分離を介する）および他の形のダイバーシティにあてはまる。

20

【 0 2 1 7 】

スライスおよび A U をフレームに整列するために、外側符号（F E C）符号ブロック作成および M A C 層カプセル化も整列しなければならない。図 2 0 に、スライスおよび A U 内の符号化されたビデオデータまたはビデオビットストリームの編成を示す。符号化されたビデオは、1 つまたは複数のビットストリーム内で、たとえば、階層化されたビデオ符号化が適用される場合にベースレイヤビットストリームおよびエンハンスメントレイヤビットストリーム内で構成することができる。

30

【 0 2 1 8 】

ビデオビットストリームは、図 2 0 でフレーム 1 ' 2 0 0 5、フレーム 3 ' 2 0 1 0、およびフレーム M ' 2 0 1 5 によって示された A U を備える。A U は、スライス 1 2 0 2 0、スライス 2 2 0 2 5、およびスライス N 2 0 3 0 によって示される、データのスライスから成り立つ。スライスの初めのそれぞれは、開始符号によって識別され、ネットワーク適応に供給される。一般に、I フレーム符号化された A U すなわちイントラ符号化された A U は、大きく、これに P フレームまたは順方向に予測されたフレームが続き、これに B フレームが続く。A U を複数のスライスに符号化することは、スライスにまたがる空間的な予測が制限され、スライスヘッダもオーバーヘッドに寄与するので、符号化ビットレートに関してかなりのオーバーヘッドコストをこうむる。スライス境界は、再同期点なので、連続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。というのは、P L P が破壊される時に、誤りが、P L P 内のスライスに閉じこめられるが、P L P が複数のスライスまたは複数のスライスの一部を含む場合には、誤りが、その P L P 内のすべてのスライスまたはスライスの部分に影響するからである。

40

【 0 2 1 9 】

I フレームは、通常は大きく、たとえば数十キロビット程度なので、複数のスライスに起因するオーバーヘッドは、I フレームの総サイズまたは総ビットレートの大きい比率で

50

はない。また、イントラ符号化されるAU内により多くのスライスを有することは、よりよくより頻繁な再同期およびより効率的な空間誤りの隠蔽を可能にする。また、Iフレームは、PフレームおよびBフレームがIフレームから予測されるので、ビデオビットストリーム内で最も重要な情報を運ぶ。Iフレームは、チャネル獲得のランダムアクセス点としても働く。

【0220】

ここで図21を参照すると、Iフレームをフレーム境界に、およびI AUを有するスライスをもフレーム境界に注意深く整列することは、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る場合に、フレーム2 2110に属するスライスは、フレーム2 2110がフレーム1 2105からの大きな時間分離を有するので手付かずのままになる高い確率を有する誤り回復を再同期および誤りの隠蔽を介して実行することができる）。

10

【0221】

Pフレームは、通常、2～3キロビット程度のサイズを有するので、Pフレームのスライスおよび整数個数のPフレームをフレーム境界に整列することは、効率の有害な損失なしで誤り弾力性を可能にする（Iフレームに関する理由と同様の理由から）。時間的な誤りの隠蔽を、そのような態様で実施することができる。代替案では、連続するPフレームを、異なるフレームに到着するように分散させることが、Pフレームの間で追加の時間ダイバーシティをもたらし、これは、時間的な隠蔽が、以前に再構成されたIフレームまたはPフレームからの動きベクトルおよびデータに基づくからであるものとして行うことができる。Bフレームは、極端に小さい（数百ビット）から適度に大きい（2～3千ビット）とすることができる。したがって、整数個数のBフレームをフレーム境界に整列することは、効率の有害な損失なしで誤り弾力性を達成するために望ましい。

20

【0222】

モード判断モジュール

図12に、モード判断モジュール715の動作のいくつかの例を示す。図示されているように、処理1200は、入力ブロック1202で開始される。1つの例示的な例で、モード判断モジュール715に入力されるさまざまな情報には、スライスタイプ、イントラ4×4コスト、イントラ16×16コスト、イントラUV8×8コスト、イントラY16×16モード、イントラUVモード、動きベクトルデータ（MVD）、量子化パラメータ（QP）、SpPredMB4×4Y、SpPredMB16×16Y、SpPredMB8×8U、SpPredMB8×8V、レート歪みフラグ、生YMB画素、生UMB画素、および生VMB画素が含まれる。次に、処理1200は、ブロック1204符号化の初期化に進み、これは、ブロック1206によって示されるように、符号化器の初期化を指示する入力信号またはインターフェースによって開始することができる。初期化には、許容されるモード（スキップ、直接を含む）を設定すること、モードの重みを設定すること（要求される場合に、デフォルトでは全てのモードについて等しい重みになる）、およびバッファをセットすることを含む。初期化の後に、処理1200は、ブロック1208に進み、ここで、許容されるモードごとのマクロブロック（MB）モードコストの計算、重み付け係数を用いる各MBモードコストの重み付け、および最小MBモードコストモードの選択を含む、モード判断の主処理が実行される。これらの動作に用いられる入力、ブロック1210および1212によって示されるように、動き推定（たとえば、MVDおよび予測）および空間的な予測（たとえば、すべてのイントラコストおよび予測）を含む。モード判断モジュール715にインターフェースされるのが、とりわけ圧縮率を改善する、ブロック1214のエントロピ符号化である。処理1200は、次に、ブロック1216に進み、ここで、符号化器の第2パス部分706に情報を渡すためにバッファを更新する。最後に、処理1200は、ブロック1218に進み、ここで、「最良」符号化モードを符号化器の第2パス部分706に通信することができる。

30

40

【0223】

符号化器の第2パス部分

50

もう一度図 7 を参照すると、符号化器 2 2 8 の第 2 パス部分 7 0 6 は、符号化の第 2 パスを実行する第 2 パス符号化器モジュール 2 3 2 を含む。第 2 パス符号化器 2 3 2 は、モード判断モジュール 7 1 5 から出力を受け取る。第 2 パス符号化器 2 3 2 は、MC / 変換量子化モジュール 7 2 6 およびジグザグ (ZZ) / エントロピ符号化器 7 2 8 を含む。第 2 パス符号化器 2 3 2 の結果は、スケーラビリティモジュール 7 3 0 およびビットストリームパッキングモジュール 7 3 1 に出力され、ビットストリームパッキングモジュール 7 3 1 は、同期化層 2 4 0 (図 2 に図示) を介するトランスコーダ 2 0 0 による伝送のために、符号化されたベースレイヤおよびエンハンスメントレイヤを出力する。図 2 に示されているように、第 2 パス符号化器 2 3 2 および再符号化器 2 3 4 からのベースレイヤおよびエンハンスメントレイヤが、同期化層 2 4 0 によって、ベースレイヤおよびエンハンスメントレイヤを含むパケット化された PES 2 4 2、データ PES 2 4 4 (たとえば、CC および他のテキストデータ)、およびオーディオ PES 2 4 6 に組み立てられることに留意されたい。オーディオ符号化器 2 3 6 が、復号化されたオーディオ情報 2 1 8 を受け取り、その情報を符号化し、符号化された情報 2 3 8 を同期化層 2 4 0 に出力することに留意されたい。

【 0 2 2 4 】

再符号化器

もう一度図 7 を参照すると、符号化器の第 2 パス部分 7 0 6 は、図 2 の再符号化器 2 3 4 に対応する再符号化器 2 3 4 をも含む。再符号化器 2 3 4 は、第 1 パス部分 7 0 2 の出力をも受け取り、MC / 変換量子化 7 2 6 部分および ZZ / エントロピ符号化 7 2 8 部分を含む。さらに、スケーラビリティモジュール 7 3 0 が、再符号化器 2 3 4 に出力する。再符号化器 2 3 4 は、シンクロナイザ (たとえば、図 2 に示された同期化層 2 4 0) への伝送のために、再符号化からビットストリームパッキングモジュール 7 3 1 へ、結果のベースレイヤおよびエンハンスメントレイヤを出力する。図 7 の符号化器 2 2 8 の例は、第 2 パスの符号化を同調するのを助ける (たとえば、圧縮効率を高める) ために第 2 パスの符号化器 2 3 2 内の MC / 変換量子化モジュール 2 3 4 と再符号化器 2 3 4 内の ZZ / エントロピモジュール 7 3 6 との両方にビットストリームパッキングフィールドバックを供給するレート制御微調整モジュール 7 3 8 をも含む。

【 0 2 2 5 】

誤り弾力性モジュール

図 7 に示された符号化器 2 2 8 の例は、第 2 パス部分 7 0 6 内に誤り弾力性モジュール 7 4 0 をも含む。誤り弾力性モジュール 7 4 0 は、ビットストリームパッキングモジュール 7 3 1 およびスライス / MB 順序付けモジュール 7 2 2 と通信する。誤り弾力性モジュール 7 4 0 は、プリプロセッサ 2 2 8 からメタデータを受け取り、誤り弾力性方式、たとえばフレーム境界へのスライスおよびアクセス単位の整列、予測階層、および適応イントラリフレッシュを選択する。誤り弾力性方式の選択は、メタデータ内で受け取られる情報に基づくものとするか、ビットストリームパッキングモジュール 7 3 1 およびスライス / MB 順序付けモジュール 7 2 2 から誤り弾力性モジュールに通信される情報からとすることができる。誤り弾力性モジュール 7 4 0 は、選択された誤り弾力性処理を実施するために、第 1 パス部分 7 0 2 内のスライス / マクロブロック (MB) 順序付けモジュールに情報を供給する。エラーが起きやすい環境を介するビデオ伝送は、より明瞭でより少ない誤りを充てんされたデータを見るユーザに提示することをもたらしことができる誤り弾力性戦略および誤り弾力性アルゴリズムを使用することができる。以下の誤り弾力性の説明は、任意の個々の既存のまたは将来のアプリケーション層、トランスポート層、および物理層または他の技術あるいはその組合せに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、低いレイテンシおよび高いスループットなどの通信システムの望ましい特性と共に、OSI 層の中で誤り感受性特性および誤り保護機能の理解を一体化する。誤り弾力性処理は、マルチメディアデータのコンテンツ情報、たとえばマルチメディアデータのコンテンツ分類に基づくものとする。主要な利益の 1 つが、フェージングおよびマルチパスチャネル誤りからの回復性である。以下で説明する誤り弾力性手法は、特に

10

20

30

40

50

、符号化器 2 2 8 内（たとえば、特に誤り弾力性モジュール 7 4 0 およびスライス / M B 順序付けモジュール 7 2 2 内）に組み込むことができ、エラーが起きやすい環境でのデータ通信に包括的に拡張できる処理に特に関する。

【 0 2 2 6 】

誤り弾力性

予測ベースのハイブリッド圧縮システムに関して、イントラ符号化されるフレームは、時間的な予測を全く用いずに独立に符号化される。インター符号化されるフレームは、過去のフレーム（P フレーム）および将来のフレーム（B フレーム）から時間的に予測することができる。最良の予測器は、基準フレーム（1 つまたは複数）での探索処理を介して識別することができ、S A D などの歪み測定値が、最良一致を識別するのに使用される。10
現行フレームの予測符号化される領域は、変化するサイズおよび形状のブロック（1 6 × 1 6、3 2 × 3 2、8 × 4 など）またはたとえばセグメント化を介してオブジェクトとして識別される画素のグループとすることができる。

【 0 2 2 7 】

時間的な予測は、通常、多数のフレーム（たとえば、1 0 個から数十個のフレーム）にまたがって延び、あるフレームが I フレームとして符号化される時に打ち切れ、G O P は、通常、I フレーム頻度によって定義される。最大の符号化効率のために、G O P は、1 シーンであり、たとえば、G O P 境界は、シーン境界に整列され、シーン変化フレームは、I フレームとして符号化される。小さい動きでは、シーケンスは、比較的静的な背景を含み、動きは、一般に、前景オブジェクトに制限される。そのような小さい動きのシーケンスのコンテンツの例は、ニュース番組および天気予報番組を含み、この場合に、ほとんどの見られるコンテンツの 3 0 % 超が、この性質を有する。小さい動きのシーケンスでは、領域のほとんどが、インター符号化され、予測フレームは、中間予測フレームを介して I フレームを戻って参照する。20

【 0 2 2 8 】

図 2 2 を参照すると、I フレーム内のイントラ符号化されたブロック 2 2 0 5 は、符号化されたフレーム（または A U）P 1 内のインター符号化されたブロック 2 2 1 0 の予測器である。この例では、これらのブロックの領域は、背景の静止部分である。連続的な時間的な予測を介して、誤りに対するイントラ符号化されたブロック 2 2 0 5 の感度は、高まる。というのは、これが、その「重要性」がより高いことをも暗示する「よい」予測器であるからである。さらに、イントラ符号化されたブロック 2 2 0 5 は、予測チェーンと呼ばれる時間的な予測のこの連鎖のおかげで、ディスプレイ内により長く（この図の例では、このシーンの持続時間の間）持続する。30

【 0 2 2 9 】

予測階層は、最上部（イントラ符号化されたブロック 2 2 0 5）の親および最下部の子を有する、この重要性レベルまたは持続性の尺度に基づいて作成されるブロックのツリーとして定義される。P 1 内のインター符号化されたブロック 2 2 1 5 が、この階層の第 2 レベルであることなどに留意されたい。葉は、予測チェーンを終わらせるブロックである。

【 0 2 3 0 】

予測階層を、コンテンツタイプと関わりなくビデオシーケンスについて作成することができ（ニュースだけではなく音楽およびスポーツなども）、予測ベースのビデオ（およびデータ）圧縮全般に適用可能である（これは、本願に記載のすべての態様にあてはまる）。予測階層が確立されたならば、以下で説明する適応イントラリフレッシュなどの誤り弾力性アルゴリズムを、より効果的に適用することができる。重要性尺度は、誤りに対する符号化されたビットストリームの弾力性を高めるための隠蔽動作および適応イントラリフレッシュの適用を介するなど、誤りからの所与のブロックの回復性に基づくものとしてすることができる。重要性尺度の推定値は、持続性メトリックとも称する、あるブロックが予測器として使用される回数に基づくものとしてすることができる。この持続性メトリックは、予測誤り伝搬を阻止することによって符号化効率を改善するのににも使用される。持続性メト 40
50

リックは、より高い重要性を有するブロックのビット割当をも増やす。

【 0 2 3 1 】

適応イントラリフレッシュ

適応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弾力性技法である。イントラリフレッシュ処理では、いくつかのMBは、標準R-D最適化がそれらをインター符号化されたMBにしなければならないと命ずる場合であっても、イントラ符号化される。AIRは、PフレームまたはBフレームにイントラ符号化されたMBを導入するために、動きに重み付けされたイントラリフレッシュを使用する。ベースレイヤに含まれる、これらのイントラ符号化されたMBは、QP_bまたはQP_eのいずれかをを用いて符号化することができる。QP_eがベースレイヤに使用される場合には、改善をエンハンスメントレイヤで行う必要はない。QP_bがベースレイヤに使用される場合には、改善が適当である場合があり、さもなければ、エンハンスメントレイヤで、品質の低下が顕著になる。インター符号化は、符号化効率の意味でイントラ符号化より効率的なので、エンハンスメントレイヤでのこれらの改善は、インター符号化される。この形では、ベースレイヤ係数は、エンハンスメントレイヤに使用されず、品質は、新しい動作を導入せずにエンハンスメントレイヤで改善される。

10

【 0 2 3 2 】

いくつかの態様においては、適応イントラリフレッシュは、動きに重み付けされたやり方ではなく、またはそれに加えて、マルチメディアデータのコンテンツ情報（たとえば、コンテンツ分類）に基づくものとして行うことができる。たとえば、コンテンツ分類が比較的高い（たとえば、高い空間的な複雑度および時間的な複雑度を有するシーン）場合に、適応イントラリフレッシュが、相対的により多数のイントラ符号化されたMBをPフレームまたはBフレームに導入することができる。その代わりに、コンテンツ分類が比較的低い（低い空間的な複雑度および/または時間的な複雑度を有するより少なく動的なシーンを示す）場合に、適応イントラリフレッシュが、より少数のイントラ符号化されたMBをPフレームおよびBフレームに導入することができる。誤り弾力性を改善するそのようなメトリックおよび方法は、無線マルチメディア通信の文脈だけではなく、データ圧縮およびマルチメディア処理全般に向けて（たとえば、グラフィックスレンダリングに）適用することができる。

20

【 0 2 3 3 】

チャンネル切替フレーム

本明細書で定義されるチャンネル切替フレーム（CSF）は、高速なチャンネルの獲得のために、したがって、ブロードキャスト多重内のストリームの間の高速なチャンネルの変更のために、ブロードキャストストリーム内の適当な位置に挿入されるランダムアクセスフレームを記述する広義の用語である。チャンネル切替フレームは、主フレームが誤りを伴って伝送される場合に使用できる冗長データを提供するので、誤り頑健性をも高める。IフレームまたはH.264のプログレッシブ復号化器リフレッシュフレームなどのプログレッシブIフレームは、通常、ランダムアクセス点として働く。しかし、頻繁なIフレーム（または、シーン持続時間より短い、短いGOP）は、圧縮効率の大幅な低下をもたらす。イントラ符号化ブロックが、誤り弾力性のために必要になる可能性があるので、誤りに対する頑健性を高めながら符号化効率を改善するために、予測階層を介してランダムアクセスおよび誤り弾力性を効果的に組み合わせることができる。

30

40

【 0 2 3 4 】

ランダムアクセス切替および誤り頑健性の改善は、一致して達成することができ、コンテンツ分類などのコンテンツ情報に基づくものとして行うことができる。小さい動きのシーケンスについて、予測チェーンは、長く、スーパーフレームまたはシーンを再構成するのに必要な情報のかなりの部分が、シーンの初めに現れるIフレームに含まれる。チャンネル誤りは、パースト的である傾向があり、フェードが現れ、FECおよび通信路符号化が障害を発生する時には、隠蔽が失敗する大きい残差がある。これは、小さい動き（したがって、低いビットレート）のシーケンスについて特に深刻である。というのは、符号化された

50

データの量が、ビデオビットストリーム内でよい時間ダイバーシティをもたらすのに十分に多くはないからであり、すべてのビットを再構成に有用にする非常に圧縮可能なシーケンスがあるからである。大きい動きのシーケンスは、コンテンツの性質すなわち、すべてのフレーム内のより多くの新しい情報が、独立に復号化可能であり本質的に誤りに対してより柔軟な符号化されたイントラブロックの個数を増やすことに起因して、誤りに対してより頑健である。予測階層に基づく適応イントラリフレッシュは、大きい動きのシーケンスについて高い性能を達成し、性能改善は、小さい動きのシーケンスについて大きくはない。したがって、Iフレームのほとんどを含むチャンネル切替フレームは、小さい動きのシーケンスに関するダイバーシティのよい源である。誤りがスーパーフレームを攻撃する時に、連続するフレームでの復号化は、予測に起因して失われる情報を回復するCSFから開始され、誤り弾力性が達成される。

10

【0235】

相対的に高いコンテンツ分類（たとえば、6～8）を有するシーケンスなどの大きい動きのシーケンスの場合に、CSFは、SF内で持続するブロックすなわち、よい予測器であるブロックからなる可能性がある。CSFの他のすべての領域は、イントラブロックで終わることを暗示する短い予測チェーンを有するブロックなので、符号化する必要がない。したがって、CSFは、それでも、誤りが発生する時に予測に起因して失われる情報から回復するように働く。小さい動きのシーケンスに関するCSFは、Iフレームのサイズと同程度であり、より大きい量子化を介してより低いビットレートで符号化することができ、大きい動きのシーケンスに関するCSFは、対応するIフレームよりはるかに小さい。

20

【0236】

予測階層に基づく誤り弾力性は、良好にスケーラビリティを処理することができ、非常に効率的な階層化された符号化を達成することができる。物理層技術での階層的な変調をサポートするスケーラビリティは、特定の帯域幅比を用いるビデオビットストリームのデータパーティショニングを必要とする場合がある。これらは、必ずしも最適スケーラビリティのための（たとえば、最小のオーバーヘッドを有する）理想的な比であるとは限らない。いくつかの態様で、1:1帯域幅比を有する2層のスケーラビリティが使用される。等しいサイズの2つの層へのビットストリームのパーティショニングは、小さい動きのシーケンスについて、さほど効率的でない場合がある。小さい動きのシーケンスについて、すべてのヘッダおよびメタデータ情報を含むベースレイヤは、エンハンスメントレイヤより大きい。しかし、小さい動きのシーケンスのCSFは、より大きいので、エンハンスメントレイヤ内の残りの帯域幅にうまくおさまる。

30

【0237】

大きい動きのシーケンスは、1:1へのデータパーティショニングを最小のオーバーヘッドで達成できる十分な残差情報を有する。さらに、そのようなシーケンスのチャンネル切替フレームは、大きい動きのシーケンスについてはるかに小さい。したがって、予測階層に基づく誤り弾力性は、大きい動きのシーケンスのスケーラビリティも良好に処理することができる。適度な動きのクリップについて上で述べた概念を拡張することは、これらのアルゴリズムの説明に基づいて可能であり、提案される概念は、ビデオ符号化全般に適用される。

40

【0238】

多重化装置

いくつかの符号化器の態様で、多重化装置は、符号化器によって作られた複数のマルチメディアストリームを符号化するのに使用でき、符号化されたビットをブロードキャストのために準備するのに使用することができる。たとえば、図2に示された符号化器228の例示的態様では、同期化層240が、多重化装置を備える。多重化装置は、ビットレート割当制御をもたらすために実施することができる。推定された複雑度を、多重化装置に供給することができ、この多重化装置は、多重化されたビデオチャンネルについて予想される符号化の複雑度に従ってこれらのビデオチャンネルの集合の使用可能な帯域幅を割り当て

50

ることができ、次に、多重化されたビデオストリームの集合の帯域幅が比較的一定である場合であっても、特定のチャンネルの品質が比較的一定のままになることを可能にする。これは、チャンネルの集合内の1チャンネルが、相対に一定のビットレートおよび可変の視覚的品質ではなく、可変のビットレートおよび相対的に一定の視覚的品質を有することをもたらす。

【0239】

図18は、複数のマルチメディアストリームまたはチャンネル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、それぞれの符号化器1804によって符号化され、符号化器1804は、多重化装置(MUX)1806と通信し、さらに多重化装置1806は、伝送媒体1808と通信する。たとえば、マルチメディアストリーム1802は、ニュースチャンネル、スポーツチャンネル、映画チャンネルなどのさまざまなコンテンツチャンネルに対応するものとして行うことができる。符号化器1804は、マルチメディアストリーム1802をシステムについて指定された符号化フォーマットに符号化する。ビデオストリームの符号化の文脈で説明するが、開示される技法の原理および利益は、たとえばオーディオストリームを含むマルチメディアストリームに全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、多重化装置1806に供給され、多重化装置1806は、さまざまな符号化されたマルチメディアストリームを組み合わせ、組み合わされたストリームを伝送のために伝送媒体1808に送る。

【0240】

伝送媒体1808は、DirecTV(登録商標)などのデジタル衛星通信、デジタルケーブル、有線および無線のインターネット通信、光ネットワーク、セル電話網、ならびに類似物などであるがこれらに限定はされないさまざまな媒体に対応するものとして行うことができる。伝送媒体1808には、たとえば、無線周波数(RF)への変調を含めることができる。通常、スペクトル制約および同様のものに起因して、伝送媒体は、限られた帯域幅を有し、多重化装置1806から伝送媒体へのデータは、相対的に固定ビットレート(CBR)で維持される。

【0241】

通常システムでは、多重化装置1806の出力での固定ビットレート(CBR)の使用は、多重化装置1806に入力される符号化されたマルチメディアまたはビデオストリームもCBRであることを必要とする場合がある。背景で説明したように、ビデオコンテンツを符号化する時のCBRの使用は、通常は望ましくない可変の視覚的品質をもたらす可能性がある。

【0242】

図解されたシステムでは、複数の符号化器1804が、入力データの予想された符号化の複雑度を通信する。1つまたは複数の符号化器1804が、それに応答して多重化装置1806から適合されたビットレート制御を受けることができる。これは、比較的複雑なビデオを符号化すると期待される符号化器1804が、擬似可変ビットレートの形でビデオのこれらのフレームについてより高いビットレートまたはより多くの帯域幅(より多くのビット毎フレーム)を受け取れることを可能にする。これは、マルチメディアストリーム1802を一定の視覚的品質で符号化することを可能にする。比較的複雑なビデオを符号化する特定の符号化器1804によって使用される余分の帯域幅は、符号化器が固定ビットレートで動作するように実施された場合に他のビデオストリーム1804を符号化するのに他の形で使用されたはずのビットに由来する。これは、多重化装置1806の出力を固定ビットレート(CBR)で維持する。

【0243】

個々のマルチメディアストリーム1802が、比較的「バースト的」である、すなわち、使用される帯域幅において変化する可能性があるが、複数のビデオストリームの累積和は、より少なくバースト的である可能性がある。たとえば多重化装置1806によって、比較的複雑なビデオを符号化するチャンネルに再割り当てすることができる、複雑度のより少ないビデオを符号化するチャンネルからのビットレート、これは、組み合わされたビデオ

ストリームの全体としての視覚的品質を高めることができる。

【0244】

符号化器1804は、一緒に符号化され多重化されるビデオフレームの組の複雑度の表示を多重化装置1806に与える。多重化装置1806の出力は、伝送媒体1808について指定されたビットレートを超えない出力をもたらさなければならない。複雑度の表示は、選択されたレベルの品質を実現するために、上で述べたコンテンツ分類に基づくものとして行うことができる。多重化装置1006は、複雑度の表示を分析し、割り当てられたビット数または帯域幅をさまざまな符号化器1004に与え、符号化器1804は、組のビデオフレームを符号化するのにこの情報を使用する。これは、ビデオフレームの組を個別に可変ビットレートにすることを可能にし、なおかつグループとして固定ビットレートを達成する。

10

【0245】

コンテンツ分類は、任意の包括的な圧縮器に関して、マルチメディア全般の品質ベースの圧縮を使用可能にするのにも使用することができる。コンテンツ分類ならびに本明細書で説明する方法および装置を、任意のマルチメディアデータの品質ベースのおよび/またはコンテンツベースのマルチメディア処理に使用することができる。一例が、任意の包括的な圧縮器に関するマルチメディア全般の圧縮での使用である。もう1つの例が、任意の逆圧縮器または復号化器あるいは補間動作、再サンプリング動作、画質向上動作、復元動作、およびプレゼンテーション動作などのポストプロセッサでの圧縮解除または復号においてである。

20

【0246】

ここで図19を参照すると、通常のビデオ通信システムは、通信ネットワークによって接続されたビデオ符号化器およびビデオ復号化器からなるビデオ圧縮システムを含む。無線ネットワークは、エラーが起きやすいネットワークの1つの種類であり、ここで、通信チャネルは、経路損失に加えて、移動体シナリオで、対数正規フェージングまたはシャドウイングおよびマルチパスフェージングを示す。チャネル誤りと戦い、アプリケーション層データの信頼できる通信を実現するために、RF変調器は、インターリーバと畳込み符号化またはターボ符号化などの通信路符号化とを含む順方向の誤り訂正を含む。

【0247】

ビデオ圧縮は、ソースビデオの冗長度を減らし、符号化されたビデオデータの各ビットで搬送される情報の量を増やす。これは、符号化されたビデオの小さい部分であっても、それが失われる時の品質への影響を増やす。ビデオ圧縮システムに固有の空間的な予測および時間的な予測は、損失を悪化させ、誤りを伝搬させ、再構成されたビデオ内の可視の人為的結果をもたらす。ビデオ符号化器の誤り弾力性アルゴリズムおよびビデオ復号化器の誤り回復アルゴリズムは、ビデオ圧縮システムの誤り頑健性を高める。

30

【0248】

通常、ビデオ圧縮システムは、基盤となるネットワークを認知しない。しかし、エラーが起きやすいネットワークでは、リンク/物理層内のFECおよび通信路の符号化と共に誤り保護アルゴリズムをアプリケーション層内に一体化するか緊密に協力させることが、非常に望ましく、システムの誤り性能を高める上で最高の効率をもたらす。

40

【0249】

図14に、フレームを符号化するために符号化器228内で発生し得るレート歪みデータの流れの一例を示す。処理1400は、開始1402で開始され、判断ブロック1404に進み、ここで、プリプロセッサ226からシーン変化検出器の入力1410を受け取り(たとえば、メタデータを介して)、誤り弾力性の入力1406を獲得する。これらの情報が、選択されたフレームがIフレームであることを示す場合に、この処理は、そのフレームをイントラ符号化する。これらの情報が、選択されたフレームがPフレームまたはBフレームであることを示す場合に、この処理は、そのフレームを符号化するのにイントラ符号化および動き推定(インター)符号化を使用する。

【0250】

50

ブロック 1 4 0 4 の条件に関して肯定の条件が発生した後に、処理 1 4 0 0 は、準備ブロック 1 4 1 4 に進み、ここで、比 R に値 $R = R_{qual}$ すなわち R - D 曲線に基づく所望の目標品質を設定する。この設定は、R - D 曲線を備えるデータブロック 1 4 1 6 から受け取られる。次に、処理 1 4 0 0 は、ブロック 1 4 1 8 に進み、ここで、レート制御ビット割当 { Q p i } が、ブロック 1 4 2 0 のコンテンツ分類処理からの画像 / ビデオアクティビティ情報（たとえば、コンテンツ分類）に基づいて実行される。

【 0 2 5 1 】

レート制御ビット割当ブロック 1 4 1 8 は、ブロック 1 4 2 2 の動き推定に使用される。動き推定 1 4 2 2 は、プリプロセッサ 1 4 1 2 からのメタデータ、ブロック 1 4 2 4 からの動きベクトル平滑化 (M P E G - 2 + ヒストリ)、およびブロック 1 4 2 6 からの複数の基準フレーム (因果的 + 非因果的マクロブロック M B) の入力をも受け取ることができる。次に、処理 1 4 0 0 は、ブロック 1 4 2 8 に進み、ここで、イントラ符号化されたモードのレート計算を、レート制御ビット割当 (Q p i) について決定する。次に、処理 1 4 0 0 は、ブロック 1 4 3 0 に進み、ここで、モードおよび量子化パラメータを決定する。ブロック 1 4 3 0 のモード判断は、ブロック 1 4 2 2 の動き推定入力、誤り弾力性 1 4 0 6 入力、およびブロック 1 4 3 2 で決定されるスケーラビリティ R - D に基づいて行われる。モードを判断したならば、流れはブロック 1 4 3 2 に進む。ブロック 1 4 3 0 から 1 4 3 2 への流れが、データが符号化器の第 1 パス部分から第 2 パス部分に渡される時に発生することに留意されたい。

【 0 2 5 2 】

ブロック 1 4 3 2 では、符号化器 2 2 8 の第 2 パスによって変換および量子化を実行する。変換 / 量子化処理は、ブロック 1 4 4 4 に示されているように、調整されまたは微調整される。この変換 / 量子化処理は、レート制御微調整モジュール (図 7) によって影響されるものとして行うことができる。次に、処理 1 4 0 0 は、符号化されたベースレイヤを作るためのジグザグソートおよびエントロピ符号化のブロック 1 4 3 4 に進む。ジグザグソートは、符号化に効率的なフォーマットで量子化されたデータを準備する。エントロピ符号化は、可能な記号の組を表すのに一連のビット符号を使用する圧縮技法である。変換 / 量子化ブロック 1 4 3 2 のエンハンストレイヤの結果は、加算器 1 4 3 6 にも送られ、加算器 1 4 3 6 は、ベースレイヤを減算し、その結果を、前に図 3 1 ~ 3 6 を参照して説明したエンハンストレイヤ用の Z Z / エントロピコード 1 4 3 8 に送る。さらに注目すべきことに、エンハンストレイヤは、真のレートに関するコンテンツ分類 1 4 2 0 の更新ならびにレート制御によって使用されるビットレートの長期履歴および短期履歴を判定する動作のためにフィードバックされる (線 1 4 4 0 真のレートの更新を参照されたい) 。

【 0 2 5 3 】

図 1 7 は、マルチメディア符号化システムの高水準ブロック図である。このマルチメディア符号化システムは、マルチメディアデータを受け取るモジュール 1 7 0 5 によって示される、マルチメディアデータを受け取るための手段を含む。そのようなための手段は、たとえば、トランスコーダ、符号化器、プリプロセッサ、マルチメディアデータを受け取るように構成されたプロセッサ、受話器を備えることができる。より具体的には、受け取るための手段は、さまざまな例でトランスコーダ 2 0 0 を含む、マルチメディアデータを受け取る、本明細書で説明する構成要素およびモジュールを備えることができる。この符号化システムは、マルチメディアデータを符号化するモジュール 1 7 1 0 によって示される、マルチメディアデータを符号化するための手段をも備える。そのような符号化のための手段 1 7 1 0 は、トランスコーダ 2 0 0、符号化器 2 2 8、またはプリプロセッサ 2 2 6 を備えることができる。

【 0 2 5 4 】

図 2 3、2 4、2 7、および 2 8 は、本明細書で説明する諸態様を実施する、マルチメディアデータを符号化する方法を例示する処理の流れ図である。図 2 3 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処理 2 3 0 0 を示す処理の流れ図である。ブロック 2 3 0 5 で、処理 2 3 0 0 は、符号化されたマルチメディアデータを受け取り

、ブロック 2 3 1 0 で、処理 2 3 0 0 は、マルチメディアデータを復号化する。ブロック 2 3 1 5 で、処理 2 3 0 0 は、復号化されたマルチメディアデータに関連するコンテンツ情報を決定する。ブロック 2 3 2 0 で、処理 2 3 0 0 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。

【 0 2 5 5 】

図 2 4 は、コンテンツ情報レベルに基づいてデータ境界を整列するためにマルチメディアデータを符号化する処理 2 4 0 0 を示す処理の流れ図である。ブロック 2 4 0 5 で、処理 2 4 0 0 は、マルチメディアデータに関連するコンテンツ情報を取得し、これは、たとえば、図 7 に示されたプリプロセッサ 2 2 6 または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 で、処理 2 4 0 0 は、コンテンツ情報に基づいてデータ境界を整列するためにマルチメディアデータを符号化する。たとえば、スライス境界およびアクセス単位境界は、符号化される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整列される。その後、符号化されたデータは、後続の処理および/または移動体デバイスへの伝送に使用可能であり、処理 2 4 0 0 は終了する。

【 0 2 5 6 】

図 2 7 は、コンテンツ情報に基づく適応イントラリフレッシュ方式を使用してデータを符号化する処理 2 7 0 0 を示す処理の流れ図である。処理 2 7 0 0 が開始される時に、マルチメディアデータは、取得済みである。ブロック 2 7 0 5 で、処理 2 7 0 0 は、マルチメディアデータのコンテンツ情報を取得する。コンテンツ情報の取得は、たとえば、上で説明したようにプリプロセッサ 2 2 6 またはコンテンツ分類モジュール 7 1 2 によって実行することができる。処理 2 7 0 0 は、ブロック 2 7 1 0 に進み、ここで、適応イントラリフレッシュの誤り弾力性方式を使用してマルチメディアデータを符号化し、適応イントラリフレッシュの誤り弾力性方式は、コンテンツ情報に基づく。ブロック 2 7 1 0 の機能は、符号化器 2 2 8 によって実行することができる。符号化されたデータは、後続の処理および伝送に使用可能にされ、処理 2 7 0 0 は終了する。

【 0 2 5 7 】

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な I フレームを使用してマルチメディアデータを符号化する処理を示す処理の流れ図である。処理 2 8 0 0 が開始される時に、マルチメディアデータは、処理に使用可能である。ブロック 2 8 0 5 で、処理 2 8 0 0 は、マルチメディアデータのコンテンツ情報を取得する。上で説明したように、これは、たとえば、プリプロセッサ 2 2 6 および/または符号化器 2 2 8 によって行うことができる。ブロック 2 8 1 0 で、処理 2 8 0 0 は、コンテンツ情報に基づいて 1 つまたは複数の追加の I フレームを符号化されたデータに挿入するためにマルチメディアデータを符号化する。これは、使用される誤り弾力性方式に応じてベースレイヤまたはエンハンスメントレイヤに I フレームを挿入することによって、誤り弾力性方式に関連して上で説明したように符号化器 2 2 8 によって行うことができる。ブロック 2 8 1 0 の後に、符号化されたデータは、後続の処理および/または移動体デバイスへの伝送に使用可能である。

【 0 2 5 8 】

本明細書で説明した方法を、当業者に既知のさまざまな通信ハードウェア、プロセッサ、およびシステムで実装できることに留意されたい。たとえば、クライアントが本明細書で説明されたように動作するための一般的な要件は、クライアントが、コンテンツおよび情報を表示するディスプレイ、クライアントの動作を制御するプロセッサ、ならびにクライアントの動作に関連するデータおよびプログラムを格納するメモリを有することである。一態様において、クライアントは、セル電話機である。もう 1 つの態様において、クライアントは、通信機能を有するハンドヘルドコンピュータである。もう 1 つの態様において、クライアントは、通信機能を有するパーソナルコンピュータである。さらに、GPS 受話器などのハードウェアを、さまざまな態様を実施するためにクライアントに組み込むことができる。本明細書で開示された諸態様に関連して説明されたさまざまな例示的なロジック、論理ブロック、モジュール、および回路は、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

10

20

30

40

50

ートアレイ（FPGA）または他のプログラマブルロジックデバイス、ディスクリートゲートまたは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、あるいは本明細書に記載の機能を実行するように設計されたこれらの任意の組合せを用いて実施し、また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とすることができるが、代替案では、プロセッサを、任意の従来のプロセッサ、コントローラ、マイクロコントローラ、または状態機械とすることができる。プロセッサは、コンピューティングデバイスの組合せとして、たとえば、DSPおよびマイクロプロセッサの組合せ、複数のマイクロプロセッサ、DSPコアと併用される1つまたは複数のマイクロプロセッサ、または任意の他のそのような構成として実施することもできる。

【0259】

本明細書で開示された諸態様に関連して説明されたさまざまな例示的なロジック、論理ブロック、モジュール、および回路は、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）または他のプログラマブルロジックデバイス、離散的なゲートまたはトランジスタロジック、離散的なハードウェア構成要素、あるいは本明細書に記載の機能を実行するように設計されたこれらの任意の組合せを用いて実施し、また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とすることができるが、代替案では、プロセッサを、任意の従来のプロセッサ、コントローラ、マイクロコントローラ、または状態機械とすることができる。プロセッサは、コンピューティングデバイスの組合せとして、たとえば、DSPおよびマイクロプロセッサの組合せ、複数のマイクロプロセッサ、DSPコアと併用される1つまたは複数のマイクロプロセッサ、または任意の他のそのような構成として実施することもできる。

【0260】

開示された方法および装置は、あるフォーマットで符号化されたビデオデータの別のフォーマットで符号化されたビデオデータへのトランスコーディングを提供し、ここで、符号化は、ビデオデータのコンテンツに基づき、符号化は、誤りに対して弾力的である。本明細書で開示された例に関連して説明された方法またはアルゴリズムは、ハードウェアで直接に、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で、ファームウェアで、またはこれらのうちの複数の組合せで実施することができる。ソフトウェアモジュールは、RAMメモリ、flash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、リムーバブルディスク、CD-ROM、または当技術分野で既知の任意の他の形の記憶媒体に存在することができる。例示的な記憶媒体は、プロセッサが、記憶媒体から情報を読み取り、記憶媒体に情報を書き込むことができるように、プロセッサに結合される。代替案では、記憶媒体を、プロセッサに一体とすることができる。プロセッサおよび記憶媒体が、ASIC内に存在することができる。ASICは、ユーザ端末内に存在することができる。代替案では、プロセッサおよび記憶媒体が、ユーザ端末内の別個の構成要素として存在することができる。

【0261】

上で説明された例は、単に例示的であり、当業者は、今や、本明細書で開示された発明的概念から逸脱せずに、上で説明された例の多数の利用を行い、上で説明された例からの新展開を作ることができる。これらの例に対するさまざまな変更は、当業者には直ちに明白になる可能性があり、本明細書で定義される包括的な原理は、本明細書で説明された新規の諸態様の趣旨または範囲から逸脱せずに、たとえばインスタントメッセージングサービスまたは任意の一般的な無線データ通信アプリケーションで他の例に適用することができる。したがって、本開示の範囲は、本明細書に示された例に限定されることを意図されているのではなく、本明細書で開示された原理および新規な特徴と一貫する最も広い範囲に従わなければならない。単語「例示的」は、本明細書では、もっぱら、「例、実例、または例示としての役割を果たす」を意味するのに使用される。本明細書で「例示的」として説明されたすべての例は、必ずしも、他の例より好ましいまたは有利と解釈してはならない。したがって、本明細書で説明された新規の態様は、添付の特許請求の範囲のみによ

10

20

30

40

50

って定義される。

【図面の簡単な説明】

【0262】

【図1A】異なるビデオフォーマットの間のトランスコーディング用のトランスコードを含むメディアブロードキャストシステムを示すブロック図。

【図1B】マルチメディアデータを符号化し、符号化された第1データグループおよび符号化された第2データグループを供給するように構成された符号化器を示すブロック図。

【図1C】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を示すブロック図。

【図2】図1のシステムのトランスコードの例を示すブロック図。

10

【図3】図2のトランスコード内で使用される解析器の動作を示す流れ図。

【図4】図2のトランスコード内で使用される復号化器の動作を示す流れ図。

【図5】図2のトランスコードによって実行される動作のシーケンスを示すシステムタイミング図。

【図6】図2のトランスコード内で使用できるプリプロセッサの動作のシーケンスおよび機能を示す流れ図。

【図7】図2のトランスコード内で使用できる例示的な2パス符号化器を示すブロック図。

【図8】テクスチャ値および動き値をコンテンツ分類にどのように関連付けるかの一態様を示す分類チャートの例を示す図。

20

【図9】図7の符号化器で使用されるものなどのコンテンツ分類の例示的動作を示す流れ図。

【図10】図7の符号化器と共に使用するためなどのレート制御の動作を示す流れ図。

【図11】図7の符号化器と共に使用するためなどの例示的な動き推定器の動作を示す流れ図。

【図12】図7の符号化器と共に使用するためなどの例示的な動き判断の符号化器の機能の動作を示す流れ図。

【図13】図7の符号化器と共に使用されるスケーラビリティをもたらし例示的動作を示す流れ図。

【図14】たとえば図7の符号化器で発生するレート歪みデータの流れをもたらし例示的動作を示す流れ図。

30

【図15】符号化の複雑度、割り当てられたビット、および人間の視覚的品質の間の関係を示すグラフ。

【図16】非線形シーン検出式を示すグラフ。

【図17】マルチメディアデータを受け取るための手段および受け取られたマルチメディアデータを符号化するための手段を示す、システムを示すブロック図。

【図18】動き推定/補償を使用するデインターレーシング処理を示す図。

【図19】マルチメディア通信システムを示すブロック図。

【図20】エンハンスメントレイヤおよびベースレイヤ内のビデオビットストリームの編成を示す図。

40

【図21】ビデオフレーム境界へのスライスの整列を示す図。

【図22】予測階層を示すブロック図。

【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処理流れ図。

【図24】コンテンツ情報レベルに基づいてデータ境界を整列す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処理の流れ図。

【図25】データのフレームのセーフアクションエリアおよびセーフタイトルエリアを示すグラフィック。

【図26】データのフレームのセーフアクションエリアを示すグラフィック。

【図27】マルチメディアコンテンツ情報に基づく適応イントラリフレッシュを使用して

50

マルチメディアデータを符号化する処理を示す処理の流れ図。

【図28】マルチメディアコンテンツ情報に基づいて冗長Iフレームを使用してマルチメディアデータを符号化する処理を示す処理の流れ図。

【図29】現行フレームと前のフレームとの間の動き補償ベクトル MV_P ならびに現行フレームと次のフレームとの間の動き補償ベクトル MV_N を示す図。

【図30】ショット検出を示す処理の流れ図。

【図31】ベースレイヤおよびエンハンスメントレイヤの符号化を示す処理の流れ図。

【図32】マクロブロックの符号化を示す概略図。

【図33】ベースレイヤおよびエンハンスメントレイヤを符号化するモジュールを示す概略図。

【図34】ベースレイヤおよびエンハンスメントレイヤの係数の選択器の処理の例を示す図。

【図35】ベースレイヤおよびエンハンスメントレイヤの係数の選択器の処理のもう1つの例を示す図。

【図36】ベースレイヤおよびエンハンスメントレイヤの係数の選択器の処理のもう1つの例を示す図。

【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処理の流れ図。

【図38】逆テレシネ処理で可能なシステム判断を示す図。

【図39】デブロッキング処理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す図。

【図40】時空間デインターレーシング処理を示す図。

【図41】1D多フェーズ再サンプリングの例を示す図。

【図42】ビデオストリーミングでの適応GOP構造の例を示す流れ図。

【図1A】

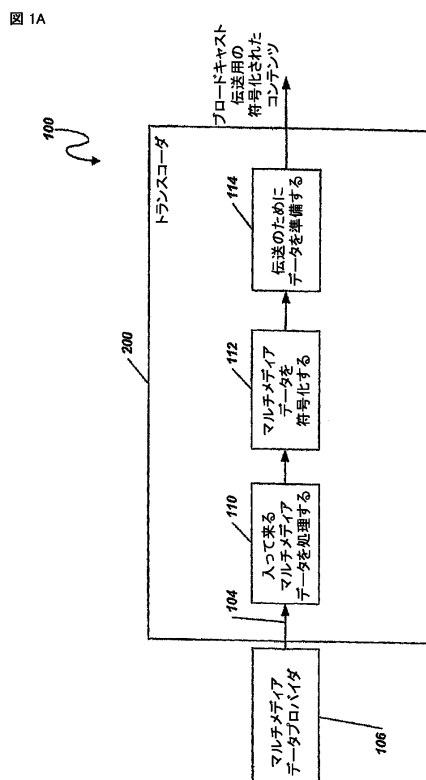
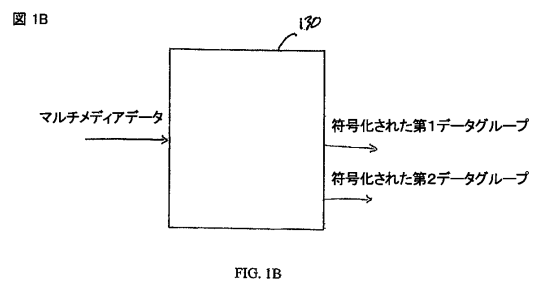
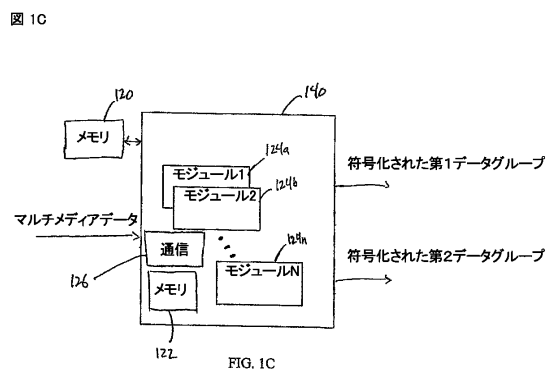


FIG. 1A

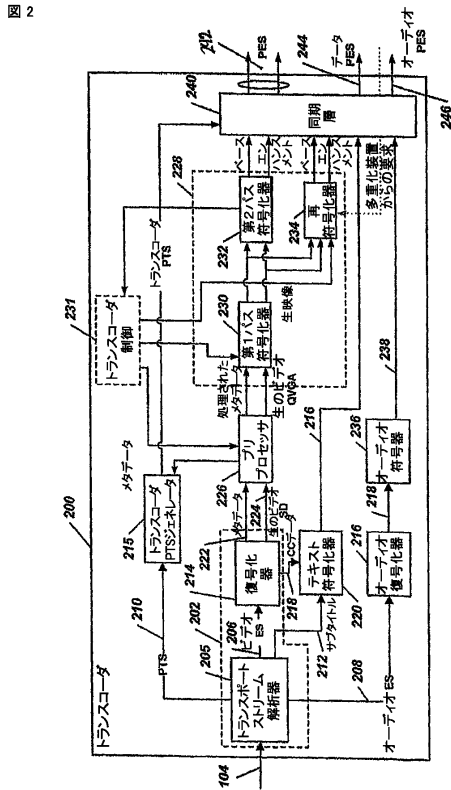
【図1B】



【図1C】



【 図 2 】



【 図 3 】

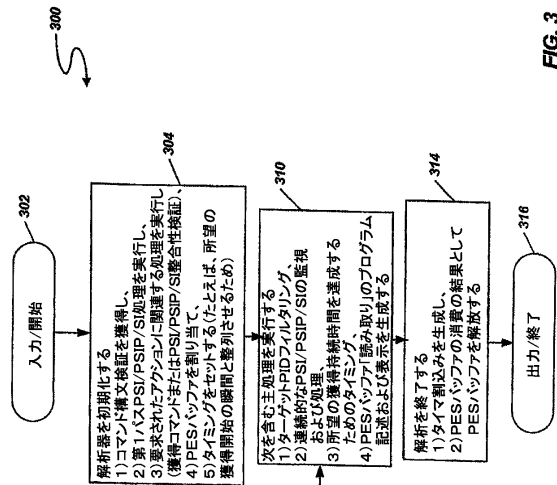
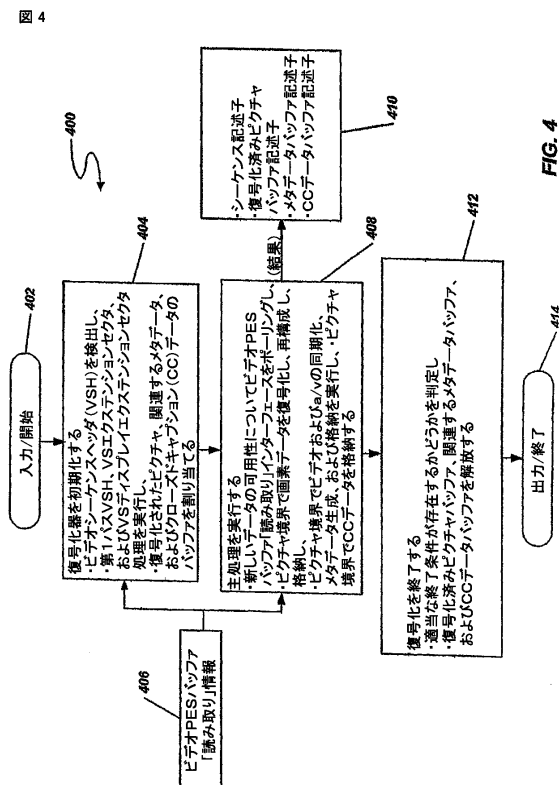


FIG. 3

【 図 4 】



【 図 5 】

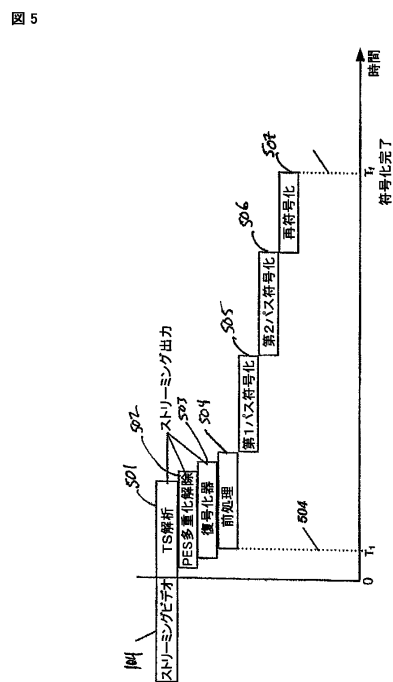


FIG. 5

【図 6】

図 6

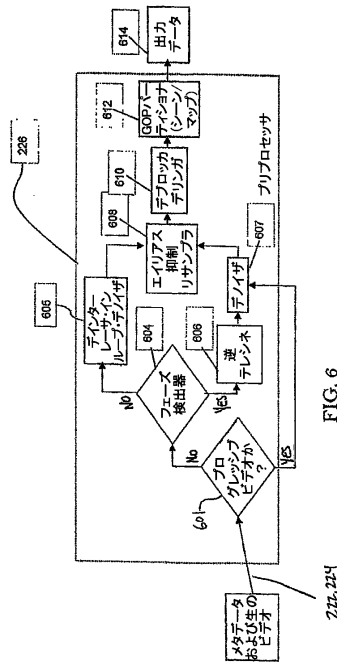


FIG. 6

【図 7】

図 7

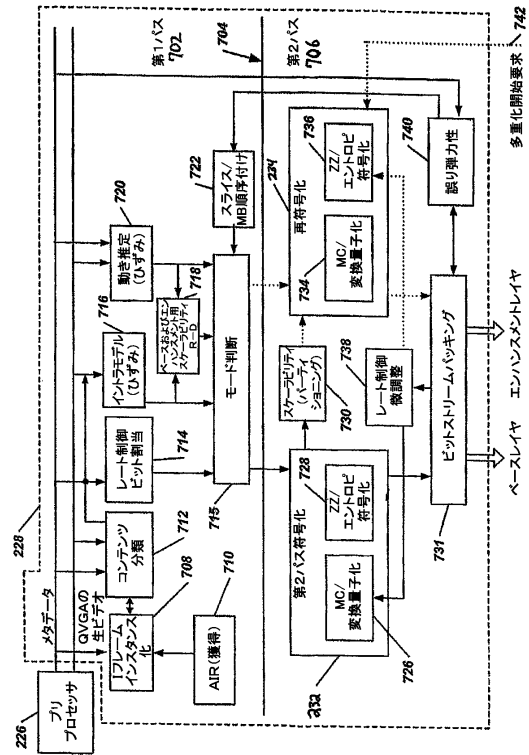


FIG. 7

【図 8】

図 8

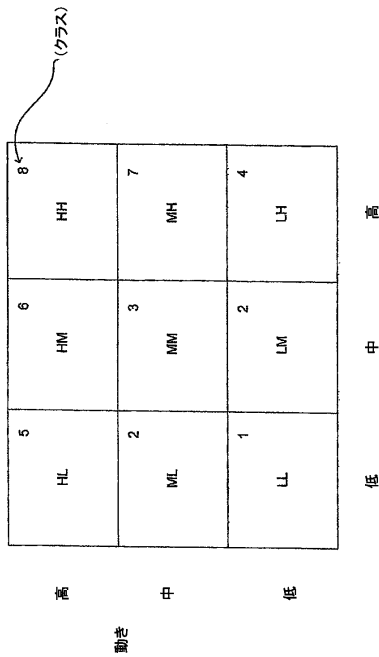
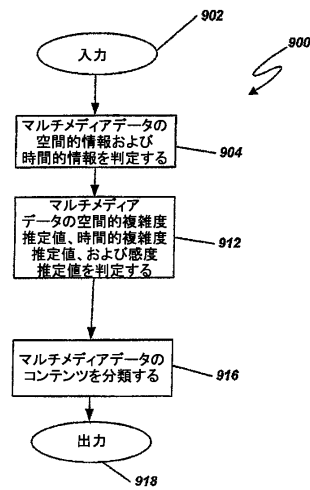


FIG. 8

【図 9】

図 9



【図 10】

図 10

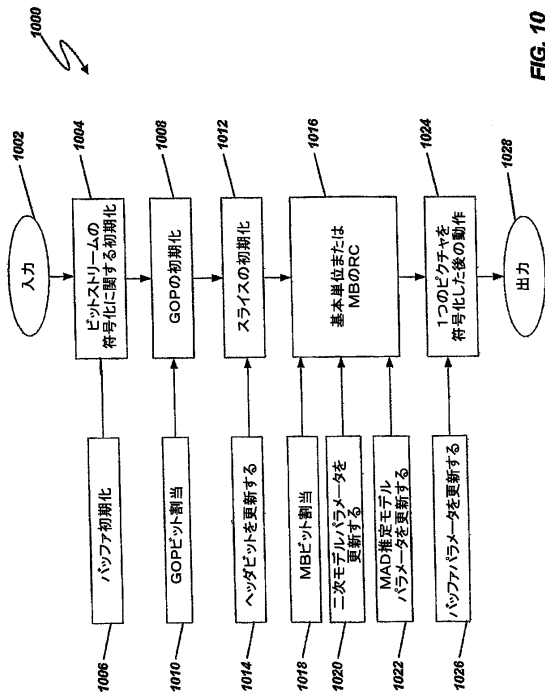


FIG. 10

【図 11】

図 11

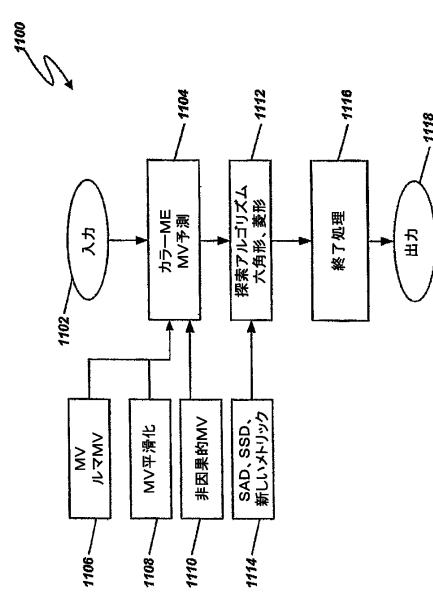


FIG. 11

【図 12】

図 12

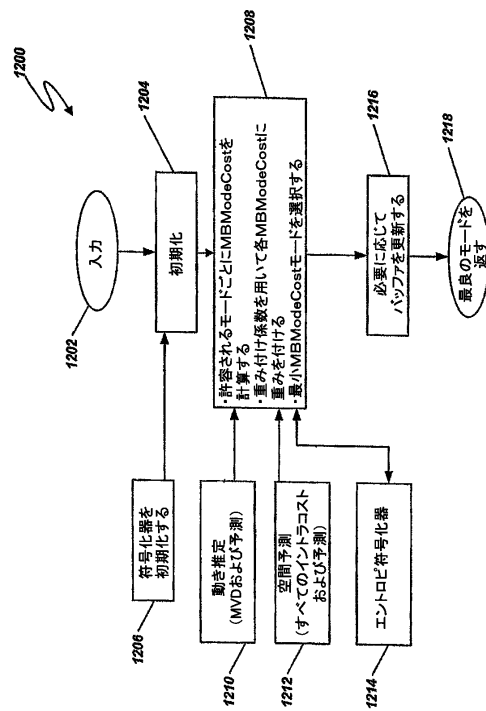


FIG. 12

【図 13】

図 13

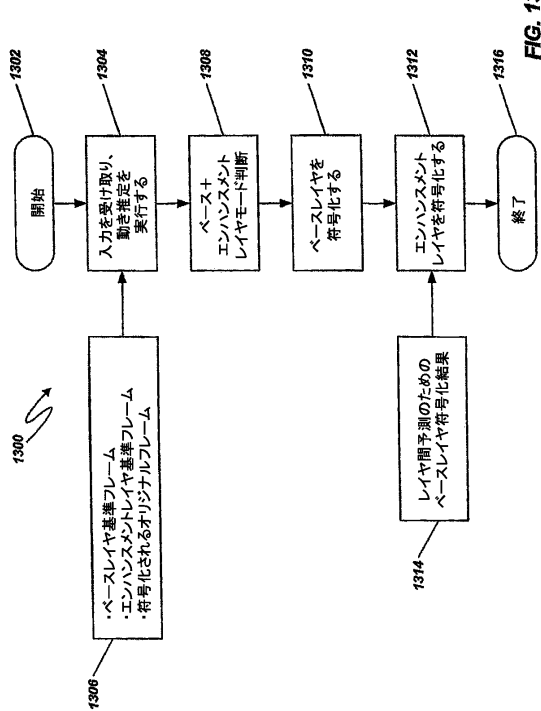
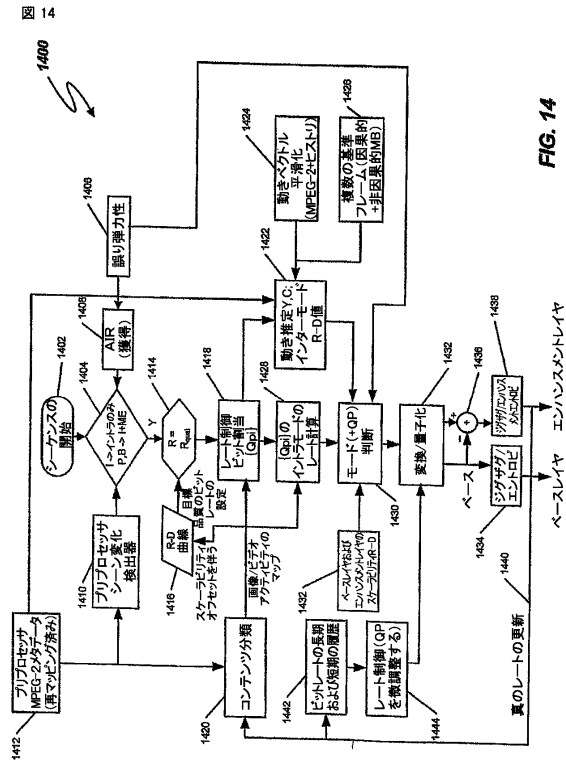


FIG. 13

【図 14】



【図 19】

図 19

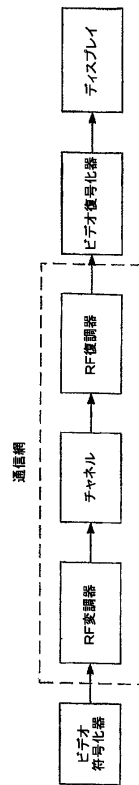


FIG. 19

【図 20】

図 20

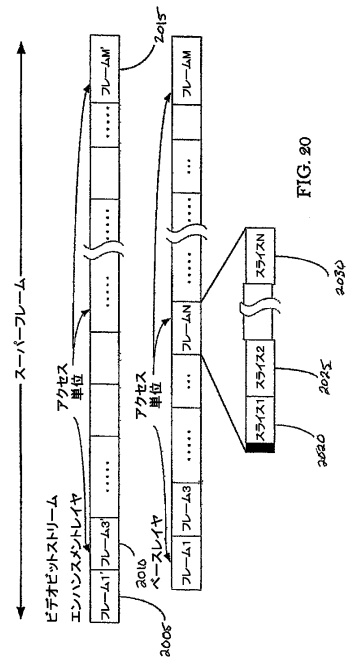


FIG. 20

【図 21】

図 21

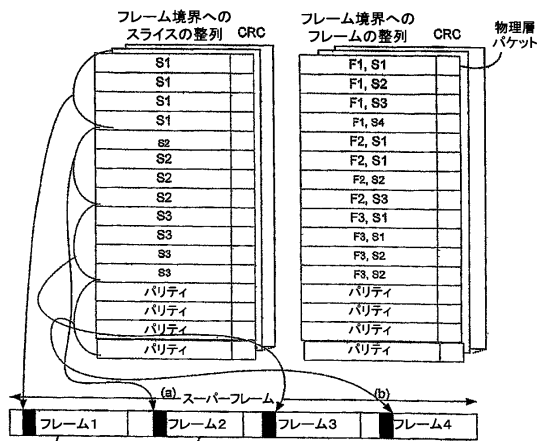


FIG. 21

【図 22】

図 22

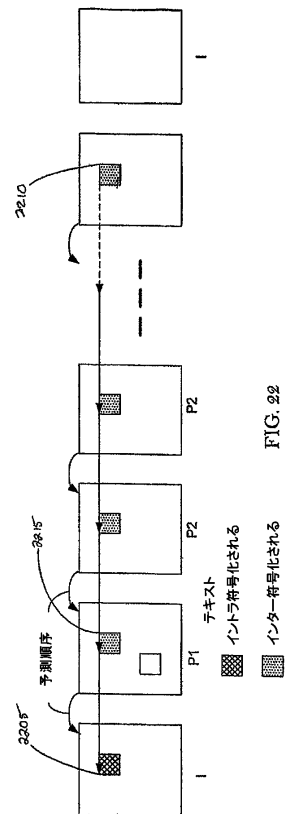
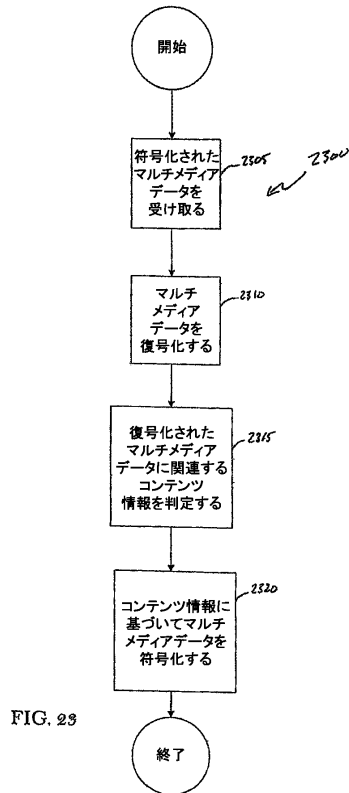


FIG. 22

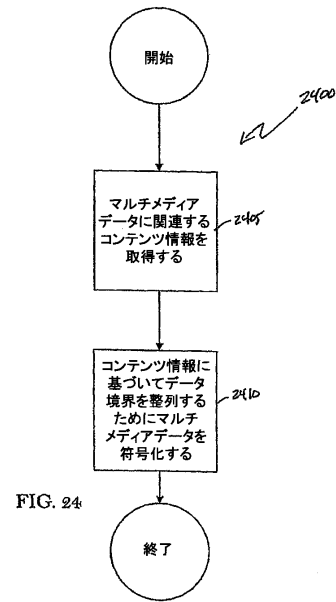
【図 23】

図 23



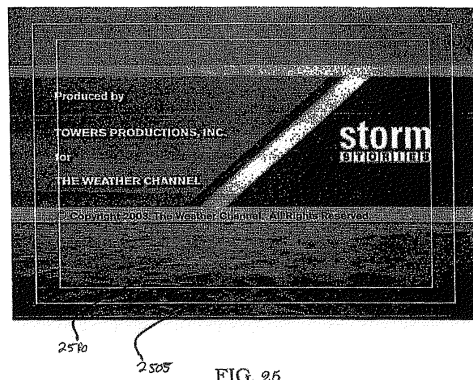
【図 24】

図 24



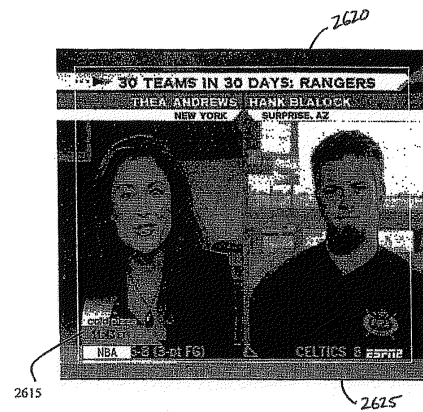
【図 25】

図 25



【図 26】

図 26



【図 27】

図 27

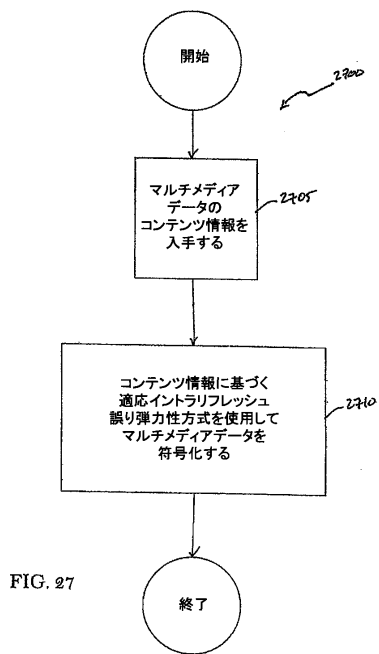


FIG. 27

【図 28】

図 28

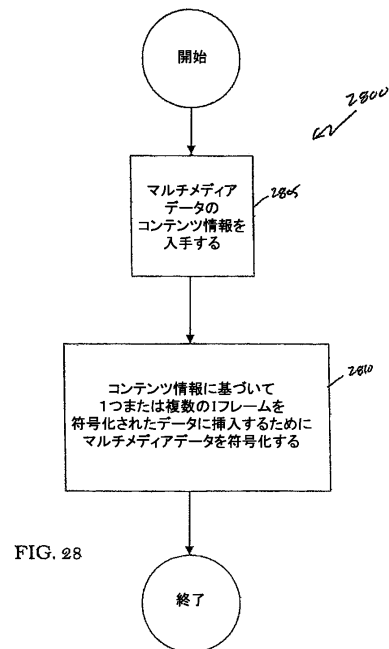


FIG. 28

【図 29】

図 29

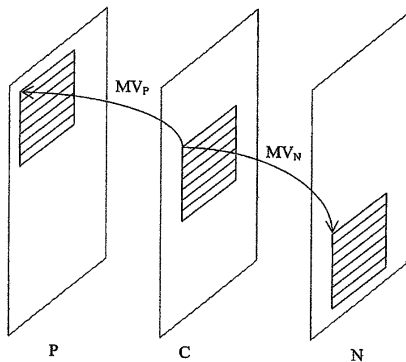


FIG. 29

【図 30】

図 30

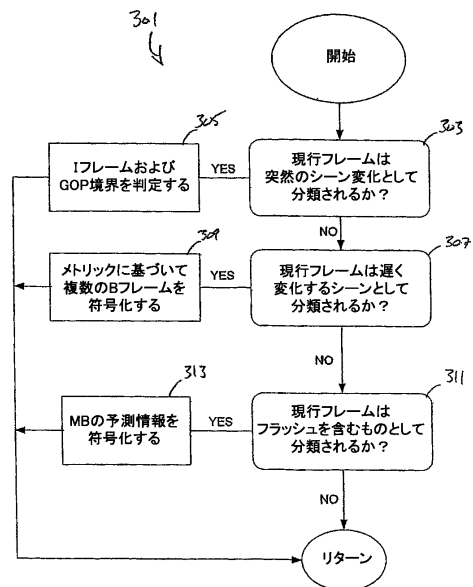


FIG. 30

【図 3 1】

図 31

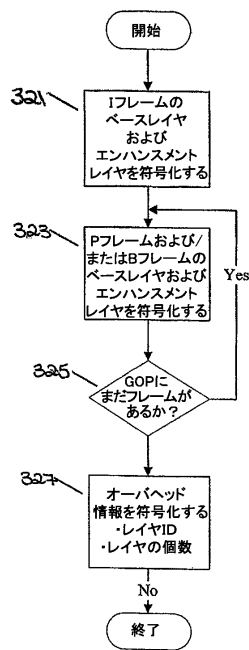


Figure 31

【図 3 2】

図 32

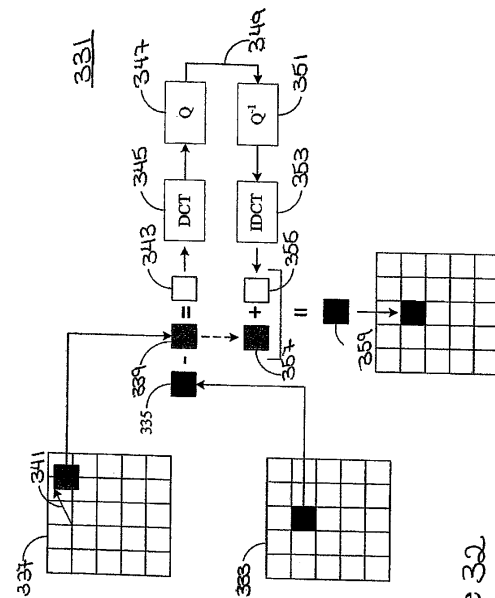


Figure 32

【図 3 3】

図 33

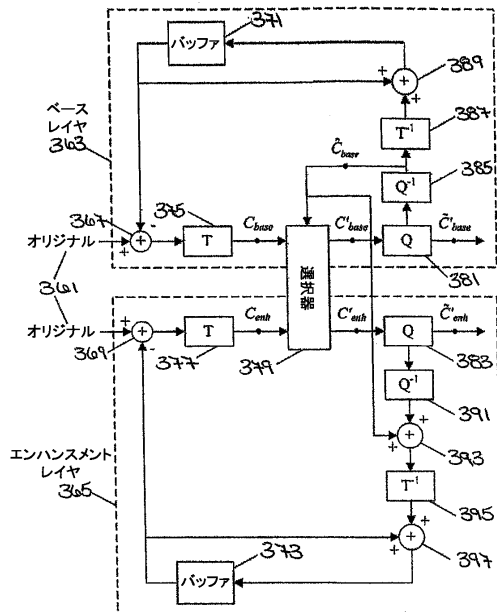


Figure 33

【図 3 4】

図 34

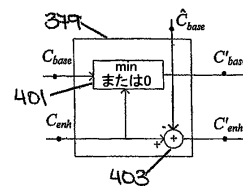


Figure 34

【図 3 5】

図 35

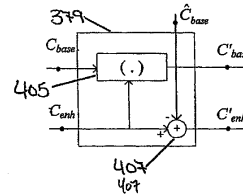


Figure 35

【図 3 6】

図 36

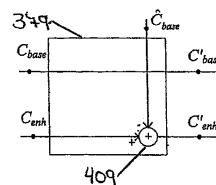
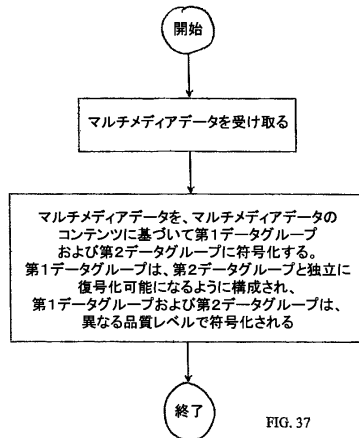


Figure 36

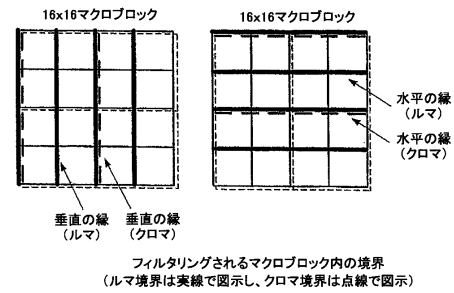
【図 37】

図 37



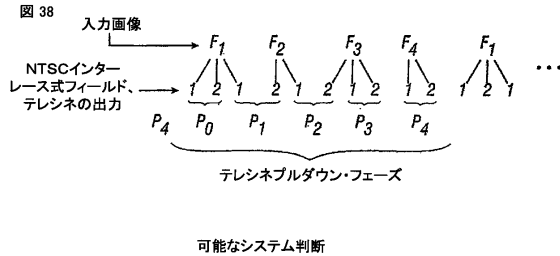
【図 39】

図 39



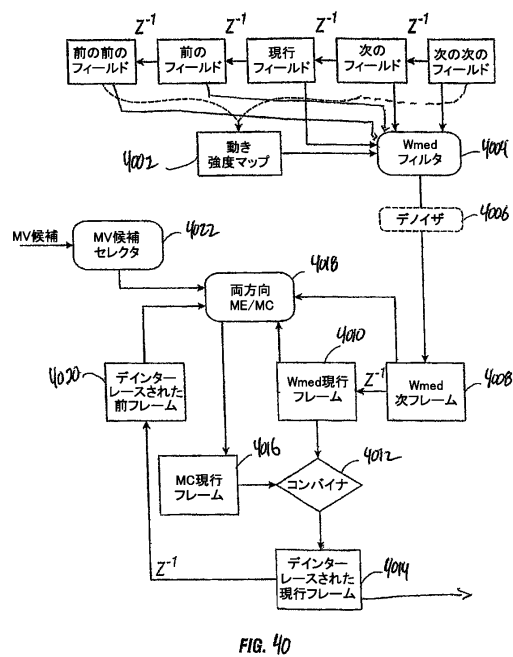
【図 38】

図 38



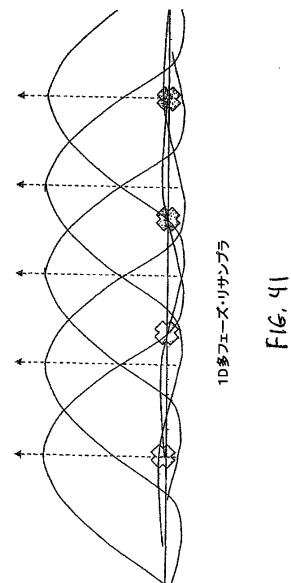
【図 40】

図 40



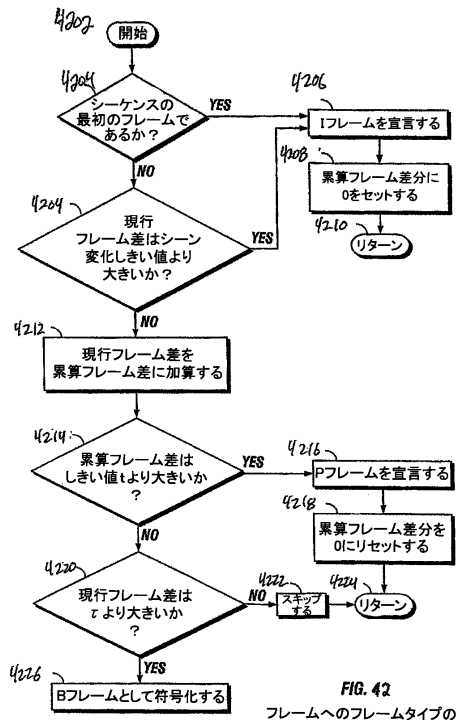
【図 41】

図 41



【図 42】

図 42



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 1 4 5
- (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、カミニト・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 2 2 0 4
- (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 4 1 7

審査官 國分 直樹

- (56)参考文献 特開 2 0 0 1 - 0 4 5 4 9 4 (J P , A)
特開 2 0 0 5 - 1 2 3 7 3 2 (J P , A)
van der Schaar, M. , A hybrid temporal-SNR fine-granular scalability for Internet video
 , IEEE Transactions on Circuits and Systems for Video Technology , IEEE , 2 0 0 1 年 3
月 , Vol.11, No.3 , pp.318-331

- (58)調査した分野(Int.Cl. , D B 名)
H04N7/26-7/50