

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第6698077号

(P6698077)

(45) 発行日 令和2年5月27日 (2020.5.27)

(24) 登録日 令和2年4月30日 (2020.4.30)

(51) Int. Cl. F I
HO 4 N 19/126 (2014.01) HO 4 N 19/126
HO 4 N 19/14 (2014.01) HO 4 N 19/14
HO 4 N 19/176 (2014.01) HO 4 N 19/176
HO 4 N 19/196 (2014.01) HO 4 N 19/196

請求項の数 30 (全 36 頁)

(21) 出願番号	特願2017-513750 (P2017-513750)	(73) 特許権者	507030416
(86) (22) 出願日	平成27年9月3日 (2015.9.3)		ユークリッド・ディスカバリーズ・エルエ
(65) 公表番号	特表2017-532858 (P2017-532858A)		ルシー
(43) 公表日	平成29年11月2日 (2017.11.2)		E U C L I D D I S C O V E R I E S ,
(86) 国際出願番号	PCT/US2015/048353		L L C
(87) 国際公開番号	W02016/040116		アメリカ合衆国, マサチューセッツ州 O
(87) 国際公開日	平成28年3月17日 (2016.3.17)		1 7 4 2 , コンコード, スイート 2 1 2
審査請求日	平成30年8月30日 (2018.8.30)		, モニュメント スクエア 30
(31) 優先権主張番号	62/049, 342	(74) 代理人	100087941
(32) 優先日	平成26年9月11日 (2014.9.11)		弁理士 杉本 修司
(33) 優先権主張国・地域又は機関	米国 (US)	(74) 代理人	100086793
(31) 優先権主張番号	62/078, 181		弁理士 野田 雅士
(32) 優先日	平成26年11月11日 (2014.11.11)	(74) 代理人	100112829
(33) 優先権主張国・地域又は機関	米国 (US)		弁理士 堤 健郎

最終頁に続く

(54) 【発明の名称】 モデルベースの映像符号化用の知覚的最適化

(57) 【特許請求の範囲】

【請求項 1】

複数の映像フレームを符号化する方法であって、

前記映像フレームは、互いに重なり合わないターゲットブロックを有しており、

当該方法は、

重要度マップが量子化を調整することによって各映像フレーム内の符号化すべき各ターゲットブロックの符号化品質に影響を与えるように、前記重要度マップを用いて前記複数の映像フレームを符号化する過程、

を備え、前記重要度マップが：

時間的情報及び空間的情報を用いて当該重要度マップを設定すること；ならびに、

(i) 当該重要度マップが高い数値をとるブロックであるほど、ブロック量子化パラメータ (Q P) がフレーム量子化パラメータ $Q P_{frame}$ に比べて小さくされることで、これらのブロックについてはより高い品質となるように、かつ、(i i) 当該重要度マップが低い数値をとるターゲットブロックであるほど、前記ブロック量子化パラメータが前記フレーム量子化パラメータ $Q P_{frame}$ に比べて大きくされることで、これらのブロックについてはより低い品質となるように、計算によって、前記複数の映像フレームのうちのある映像フレームのどの部分が人間の知覚にとって最も気付き易いのかを当該重要度マップに示させること；

によって構成され、前記時間的情報が、

時間的に周期的な刺激に対する人間の視覚系の応答を測定してどのターゲットブロック

10

20

が観測者である人間にとって時間的に最も気付き易いかを示す時間的コントラスト感度関数 (T C S F)、および、

どのターゲットブロックが前景データに相当するかを示す真の動きベクトルマップ (T M V M)

により提供されて、前記 T C S F は、前景データとして特定されたターゲットブロックについてのみ有効とされ、

前記 T C S F の時間的周波数は、動きベクトルの大きさとフレームレートとを用いて求められる速度の近似を、前記ターゲットブロックとその参照ブロックとの間の色空間領域における平均構造的類似度 (S S I M) を用いて求められる波長の近似で除することによって算出される、方法。

10

【請求項 2】

請求項 1 に記載の方法において、前記空間的情報が、ルールに基づく空間的複雑度マップ (S C M) により提供されて、その最初のステップが、前記フレーム内のどのターゲットブロックが当該フレーム内の平均ブロック分散 var_{frame} よりも大きい分散を有するかを決定することであり、

前記平均ブロック分散 var_{frame} よりも大きい分散を有するブロックに対して、前記フレーム量子化パラメータ QP_{frame} よりも高い量子化パラメータ (Q P) 値を振り当て、このブロック量子化パラメータ (Q P) の振当量 QP_{block} は、そのブロックのブロック分散 var_{block} 、前記平均ブロック分散 var_{frame} 、量子化パラメータ上限 QP_{max} および前記フレーム量子化パラメータ QP_{frame} を用いて、

$$QP_{block} = ((var_{block} - var_{frame}) / var_{block}) \times (QP_{max} - QP_{frame}) + QP_{frame}$$
 となるように線形的に増減される、方法。

20

【請求項 3】

請求項 2 に記載の方法において、前記フレーム内の平均ブロック分散 var_{frame} よりも分散の大きいブロックは、そのブロック量子化パラメータ (Q P) である前記振当量 QP_{block} が、前記 T M V M がターゲットブロックを前景データとして特定し且つ前記 T C S F のこのブロックについてのコントラスト感度対数値が 0 . 5 未満である場合には前記振当量 QP_{block} が 2 増加するように、前記 T C S F 及び前記 T M V M によりさらに洗練化される、方法。

30

【請求項 4】

請求項 2 に記載の方法において、前記 S C M は、さらに、ブロック平均輝度が 170 超の輝度か又は 60 未満の輝度のターゲットブロックのブロック量子化パラメータである前記振当量 QP_{block} が QP_{max} に調節し直される輝度マスキングを含む、方法。

【請求項 5】

請求項 2 に記載の方法において、前記 S C M は、さらに、前記符号化された映像の品質レベルに基づく前記量子化パラメータ上限 QP_{max} の動的な決定を含み、

この動的な決定では、イントラ (I) フレーム内のターゲットブロックの平均構造的類似度 (S S I M) 算出結果を前記 I フレームの平均ブロック分散 var_{frame} と共に用いて、品質が測定され、

40

前記測定された品質が低いほど、前記量子化パラメータ上限 QP_{max} の数値が前記フレーム量子化パラメータ QP_{frame} により近づくように減らされる、方法。

【請求項 6】

請求項 2 に記載の方法において、分散が 60 未満であるブロックに対して、これらの領域における高品質符号化を確実にするために、前記ブロック分散が小さいほど前記振当量 QP_{block} の数値が低くなるように、決められた低い量子化パラメータ (Q P) の値である前記振当量 QP_{block} が振り当てられる、方法。

【請求項 7】

請求項 6 に記載の方法において、分散が 60 未満であるブロックに対する前記量子化パラメータ (Q P) の値である前記振当量 QP_{block} は、最初に、I フレームの当該ブ

50

ロックについて決められ、その後、Pフレーム及びBフレームにおける当該ブロックに相当するブロックについてはIフレームのフレームQP値のPフレームのフレームQP値に対する比率である $i p r a t i o$ パラメータ及びPフレームのQP値の、BフレームのQP値に対する比率である $p b r a t i o$ パラメータを用いて決められる、方法。

【請求項8】

請求項6に記載の方法において、分散が60以上で、平均ブロック分散以下であるブロックは、当該ブロックについて、

前記ブロック量子化パラメータ(QP)の初めの推定値である前記振当量 $Q P_{b l o c k}$ が現在のブロックの左、左上、右および右上の既に符号化済みの近傍ブロックの量子化パラメータ(QP)の値を平均することによって算出されて、且つ、

前記現在のブロックの前記SSIMの推定 $S S I M_{e s t}$ が前記現在のブロックの左、左上、右および右上の既に符号化済みの近傍ブロックのSSIM値から算出されて、且つ、

$S S I M_{e s t}$ が0.9未満の場合、品質向上が必要か否かの判定において品質向上が必要と判定され、前記算出された振当量 $Q P_{b l o c k}$ の数値が2減少されるように、品質向上の処理がなされる、方法。

【請求項9】

請求項8に記載の方法において、前記品質向上は、前記TMVMにより前景データとして特定されて且つ前記TCSFのコントラスト感度対数値が0.8超であるブロックにのみ適用される、方法。

【請求項10】

請求項1に記載の方法において、前記TCSFは、現在のフレームについての当該TCSFが符号化された過去のフレームにおけるTCSFマップの重み付き平均であるように且つより時間的に近い過去の、符号化された過去のフレームがより大きい重み付けを受けるように、複数のフレームにわたって算出される、方法。

【請求項11】

請求項1に記載の方法において、前記TMVMは、前景データの場合にのみ1に設定される、方法。

【請求項12】

請求項11に記載の方法において、前景データは、所与のターゲットブロックについてのエンコーダ動きベクトルと当該ブロックについてのグローバル動きベクトルとの差分を算出し、特定値より大きい差分を有するブロックが前景データであると判断されることによって特定される、方法。

【請求項13】

請求項12に記載の方法において、前景データとして特定されたデータブロックについて、前記グローバル動きベクトルから前記エンコーダ動きベクトルが減算されることによって差分動きベクトルを得て、前記TCSFの時間的周波数を算出するのに用いられる速度 v が、この差分動きベクトルの大きさ $|D M V|$ 、この差分動きベクトルにより指し示される参照フレームと現在のフレームとの間のフレームの数 N およびフレームレートを用いて、 $v = |D M V| \times \text{フレームレート} / N$ と算出される、方法。

【請求項14】

請求項1に記載の方法において、前記TCSFは、エンコーダからの動きベクトル $M V$ を用いて前記TCSFへの入力である時間的周波数を算出するための速度 v を $v = |M V| \times \text{フレームレート} / N$ (式中、フレームレートはその映像が生成された1秒当たりのフレームの数であり、 N は前記動きベクトルにより指し示される参照フレームと現在のフレームとの間のフレームの数である) より算出する、方法。

【請求項15】

請求項2に記載の方法において、前記重要度マップが前記時間的情報及び前記空間的情報で設定されたものである場合、当該重要度マップは統合化された重要度マップ(TCSFおよびSCMの両方からの情報を含む重要度マップ)である、方法。

【請求項 16】

映像データを符号化するシステムであって、

重要度マップを用いて複数の映像フレームを符号化するコーデックであって、当該映像フレームは、互いに重なり合わないターゲットブロックを有している、コーデック、

を備え、前記重要度マップは、量子化を調整することによって各映像フレーム内の符号化すべき各ターゲットブロックの符号化品質に影響を与えるように構成されており、

前記重要度マップが：

時間的情報及び空間的情報を用いて当該重要度マップを設定すること；ならびに、

(i) 当該重要度マップが高い数値をとるブロックであるほど、ブロック量子化パラメータ (Q P) がフレーム量子化パラメータ $Q P_{frame}$ に比べて小さくされることで、これらのブロックについてはより高い品質となるように、かつ、(i i) 当該重要度マップが低い数値をとるターゲットブロックであるほど、前記ブロック量子化パラメータが前記フレーム量子化パラメータ $Q P_{frame}$ に比べて大きくされることで、これらのブロックについてはより低い品質となるように、計算によって、前記複数の映像フレームのうちのある映像フレームの、人間の知覚にとって最も気付き易い部分を当該重要度マップに示させること；

によって構成され、前記時間的情報が、

時間的に周期的な刺激に対する人間の視覚系の応答を測定してどのターゲットブロックが観測者である人間にとって時間的に最も気付き易いかを示す時間的コントラスト感度関数 (T C S F) および

どのターゲットブロックが前景データに相当するかを示す真の動きベクトルマップ (T M V M) ；

により提供されて、前記 T C S F は、前景データとして特定されたターゲットブロックについてのみに有効とされ、

前記 T C S F の時間的周波数は、動きベクトルの大きさとフレームレートとを用いて求められる速度の近似を、前記ターゲットブロックとその参照ブロックとの間の色空間領域における平均構造的類似度 (S S I M) を用いて求められる波長の近似で除することによって算出される、システム。

【請求項 17】

請求項 16 に記載のシステムにおいて、前記空間的情報が、ルールに基づく空間的複雑度マップ (S C M) により提供されて、その最初のステップが、前記フレーム内のどのターゲットブロックが当該フレーム内の平均ブロック分散 var_{frame} よりも大きい分散を有するかを決定することであり、

前記平均ブロック分散 var_{frame} よりも大きい分散を有するブロックに対して、前記フレーム量子化パラメータ $Q P_{frame}$ よりも高い量子化パラメータ (Q P) 値を振り当て、このブロック量子化パラメータ (Q P) の振当量 $Q P_{block}$ は、そのブロックのブロック分散 var_{block} 、前記平均ブロック分散 var_{frame} 、量子化パラメータ上限 $Q P_{max}$ および前記フレーム量子化パラメータ $Q P_{frame}$ を用いて、 $Q P_{block} = ((var_{block} - var_{frame}) / var_{block}) \times (Q P_{max} - Q P_{frame}) + Q P_{frame}$ となるように線形的に増減される、システム。

【請求項 18】

請求項 17 に記載のシステムにおいて、前記フレーム内の平均ブロック分散 var_{frame} よりも分散の大きいブロックは、そのブロック量子化パラメータ (Q P) である前記振当量 $Q P_{block}$ が、前記 T M V M がターゲットブロックを前景データとして特定し且つ前記 T C S F のこのブロックについてのコントラスト感度対数値が 0 . 5 未満である場合には前記振当量 $Q P_{block}$ が 2 増加するように、前記 T C S F 及び前記 T M V M によりさらに洗練化される、システム。

【請求項 19】

請求項 17 に記載のシステムにおいて、前記 S C M は、さらに、ブロック平均輝度が 1

70 超の輝度か又は 60 未満の輝度のターゲットブロックのブロック量子化パラメータである前記振当量 QP_{block} が QP_{max} に調節し直される輝度マスキングを含む、システム。

【請求項 20】

請求項 17 に記載のシステムにおいて、前記 SCM は、さらに、符号化された映像の品質レベルに基づく前記量子化パラメータ上限 QP_{max} の動的な決定を含み、

この動的な決定では、イントラ (I) フレーム内のターゲットブロックの平均構造的類似度 (SSIM) 算出結果を前記 I フレームの平均ブロック分散 var_{frame} と共に用いて、品質が測定され、

測定された品質が低いほど、前記量子化パラメータ上限 QP_{max} の数値が前記フレーム量子化パラメータ QP_{frame} により近づくように減らされる、システム。

10

【請求項 21】

請求項 17 に記載のシステムにおいて、分散が 60 未満であるブロックに対して、これらの領域における高品質符号化を確実にするために、前記ブロック分散が小さいほど前記振当量 QP_{block} の数値が低くなるように、決められた量子化パラメータ (QP) の値である前記振当量 QP_{block} が振り当てられる、システム。

【請求項 22】

請求項 21 に記載のシステムにおいて、分散が 60 未満であるブロックに対する前記低い量子化パラメータ (QP) の値である前記振当量 QP_{block} は、最初に、I フレームについては決められ、その後、P フレーム及び B フレームにおける当該ブロックに相当するブロックについては I フレームのフレーム QP 値の P フレームのフレーム QP 値に対する比率である $ipratio$ パラメータ及び P フレームの QP 値の、B フレームの QP 値に対する比率である $pbratio$ パラメータを用いて決められる、システム。

20

【請求項 23】

請求項 17 に記載のシステムにおいて、分散が 60 以上で、平均ブロック分散以下であるブロックは、当該ブロックについて、

前記ブロック量子化パラメータ (QP) の初めの推定値である前記振当量 QP_{block} が現在のブロックの左、左上、右および右上の既に符号化済みの近傍ブロックの量子化パラメータ (QP) の値を平均することによって算出されて、且つ、

前記現在のブロックの前記 SSIM の推定 $SSIM_{est}$ が前記現在のブロックの左、左上、右および右上の既に符号化済みの近傍ブロックの SSIM 値から算出されて、且つ、

30

$SSIM_{est}$ が 0.9 未満の場合、品質向上が必要か否かの判定において品質向上が必要と判定され、前記算出された振当量 QP_{block} の数値が 2 減少されるように、品質向上の処理がなされる、システム。

【請求項 24】

請求項 23 に記載のシステムにおいて、前記品質向上は、前記 TMVM により前景データとして特定されて且つ前記 TCSF のコントラスト感度対数値が 0.8 超であるブロックにのみ適用される、システム。

【請求項 25】

40

請求項 16 に記載のシステムにおいて、前記 TCSF は、現在のフレームについての当該 TCSF が符号化された過去のフレームにおける TCSF マップの重み付き平均であるように且つより時間的に近い過去の、符号化された過去のフレームがより大きい重み付けを受けるように、複数のフレームにわたって算出される、システム。

【請求項 26】

請求項 16 に記載のシステムにおいて、前記 TMVM は、前景データの場合にのみ 1 に設定される、システム。

【請求項 27】

請求項 26 に記載のシステムにおいて、前景データは、所与のターゲットブロックについてのエンコーダ動きベクトルと当該ブロックについてのグローバル動きベクトルとの差

50

分を算出し、特定値より大きい差分を有するブロックが前景データであると判断されることによって特定される、システム。

【請求項 28】

請求項 16 に記載のシステムにおいて、前景データとして特定されたデータブロックについて、前記データブロックについてのグローバル動きベクトルから前記データブロックについてのエンコーダ動きベクトルが減算されることによって差分動きベクトルを得て、前記 T C S F の時間的周波数を算出するのに用いられる速度 v が、この差分動きベクトルの大きさ $|DMV|$ 、この差分動きベクトルにより指し示される参照フレームと現在のフレームとの間のフレームの数 N およびフレームレートを用いて、 $v = |DMV| \times \text{フレームレート} / N$ と算出される、システム。

10

【請求項 29】

請求項 16 に記載のシステムにおいて、前記 T C S F は、前記エンコーダからの動きベクトル MV を用いて前記 T C S F への入力である時間的周波数を算出するための速度 v を $v = |MV| \times \text{フレームレート} / N$ (式中、フレームレートはその映像が生成された 1 秒当たりのフレームの数であり、 N は前記動きベクトルにより指し示される参照フレームと現在のフレームとの間のフレームの数である) より算出する、システム。

【請求項 30】

請求項 17 に記載のシステムにおいて、前記重要度マップが前記時間的情報と前記空間的情報で設定されたものである場合、当該重要度マップは統合化された重要度マップ (T C S F および S C M の両方からの情報を含む重要度マップ) である、システム。

20

【発明の詳細な説明】

【関連出願】

【0001】

本願は、2015年5月7日出願の米国仮特許出願第62/158,523号および2014年11月11日出願の米国仮特許出願第62/078,181号の利益を主張する。本願は、さらに、2014年11月4日出願の米国特許出願第14/532,947号の一部継続出願 (C I P) である。この米国特許出願第14/532,947号は、2014年3月10日出願の米国仮特許出願第61/950,784号および2014年9月11日出願の米国仮特許出願第62/049,342号の利益を主張する。これら参照した特許出願の全教示内容は、参照をもって本明細書に取り入れたものとする。

【背景技術】

30

【0002】

映像圧縮は、デジタル映像データを、記憶時又は伝送時において少ないビット数を用いた形式で表現するプロセスであると考えられる。映像符号化は、映像データの空間的冗長性、時間的冗長性又は色空間冗長性を利用することにより圧縮を達成することができる。典型的に、映像圧縮プロセスは、映像データをフレームの集まりやペルの集まりなどの各部位に分割して、その映像内における冗長な部分を特定し、当該冗長な部分を元々の映像データで求められるよりも少ないビット数で表現し得る。データのこのような冗長性を利用することにより、より大きな圧縮を達成することができる。映像データを符号化フォーマットに変換するにはエンコーダが用いられ得る。そして、デコーダを用いることにより、符号化された映像を元々の映像データに匹敵する形態に変換することができる。エンコーダ/デコーダを実現するものがコーデックと称される。

40

【0003】

標準のエンコーダ (標準エンコーダ) は、符号化にあたって、所与の映像フレームを、互いに重なり合わない複数の符号化単位すなわちマクロブロック (複数の連続するペルからなる矩形領域) に分割する。典型的に、マクロブロック (本明細書では、より包括的に「入力ブロック」や「データブロック」と称される) は、映像フレームの左から右の走査順序や上から下の走査順序で処理される。圧縮は、入力ブロックが符号化済みのデータを用いて予測・符号化されることで達成され得る。入力ブロックを、同じフレーム内の先に符号化されたブロックのうち、当該入力ブロックと空間的に隣接するサンプルを用いて符号化するプロセスは、イントラ予測と称される。イントラ予測は、データにおける空間的

50

冗長性を利用しようとするものである。入力ブロックを、動き推定プロセスを用いて見つけ出された、先に符号化したフレームからの類似する領域を用いて符号化することは、インター予測と称される。インター予測は、データにおける時間的冗長性を利用しようとするものである。動き推定プロセスは、動きベクトルを生成し得る。動きベクトルは、例えば、参照フレーム内の一致する領域の、符号化中の入力ブロックに対する位置を指定する。大抵の動き推定プロセスは、所与の入力ブロックについての動きベクトルの初めの粗推定（および対応する時間的予測）を提供する動き初期推定（初期の動き推定）と、この初めの推定の近傍において局所探索を実行することによってその入力ブロックについての動きベクトルのより正確な推定（および対応する予測）を決定する動き精推定（精細な動き推定）との、2つの主なステップからなる。

10

【0004】

エンコーダは、符号化すべきデータと予測（予測結果）との差分を測定することにより、残差を生成し得る。この残差は、予測されたブロックと元々の入力ブロックとの差分を提供し得る。これらの予測、動きベクトル（インター予測用）、残差および関連データは、空間変換、量子化、エントロピー符号化、ループフィルタなどの他のプロセスと組み合わせられることにより、映像データの効率的な符号（符号化）を生成することができる。量子化及び変換を受けた残差は、処理されて且つ上記予測に組み戻されることで復号化フレームへと組み立てられて、フレームストアに記憶される。このような映像符号化技術の詳細は、当業者であればよく知っている。

【0005】

20

MPEG-2 (H.262) および H.264 (MPEG-4 Part 10 Advanced Video Coding (AVC)) は、比較的低いビットレートで高品質映像表現を達成する、映像圧縮用の2種類のコーデック規格である（以降では、それぞれ MPEG-2、H.264 と称する）。MPEG-2 及び H.264 の符号化基本単位は、 16×16 マクロブロックである。H.264 は、広く普及している最近の映像圧縮規格であると共に、映像データを圧縮するにあたって MPEG-2 の2倍の効率を有すると一般的に考えられている。

【0006】

基礎的な MPEG 規格は、フレーム内の入力ブロックの符号化方法に基づいて3種類のフレーム（又はピクチャ）を規定する。Iフレーム（イントラ符号化ピクチャ）は、そのフレームに存在するデータのみを用いて符号化されるものなので、イントラ予測ブロックのみで構成される。Pフレーム（予測ピクチャ）は、予め復号化されたIフレーム又はPフレーム（参照フレームとも称される）からのデータを用いた前方向予測により符号化される。Pフレームは、イントラブロックおよび（前方向）予測ブロックのいずれも含み得る。Bフレーム（双予測ピクチャ）は、前のフレームと後のフレームの両方からのデータを用いた双方向予測により符号化される。Bフレームは、イントラブロック、（前方向）予測ブロックおよび双予測ブロックのいずれも含み得る。

30

【0007】

参照フレームの特定の集合のことを、Group of Pictures（ピクチャの集まり）（GOP）と称する。GOPは、各参照フレーム内の復号化されたペルのみを含み、入力ブロックやフレームがどのように符号化されたのか（Iフレームなのか、Bフレームなのか、それともPフレームなのか）についての情報を含まない。MPEG-2などの古い映像圧縮規格は、Pフレームを予測するのに1つの参照フレーム（過去のフレーム）を利用し、Bフレームを予測するのに2つの参照フレーム（1つ前のフレームと1つ後のフレーム）を利用する。対照的に、H.264、HEVC (High Efficiency Video Coding) などのより新しい圧縮規格は、Pフレーム及びBフレームの予測に複数の参照フレームを利用することを可能にする。典型的な参照フレームは現在のフレームと時間的に隣接するフレームであるが、これらの規格は、時間的に隣接しないフレームを参照フレームとすることも可能である。

40

【0008】

50

従来のインター予測は、ブロックベースの動き推定・補償 (BBMEC) に基づくものである。BBMEC プロセスは、ターゲットブロック (符号化中の現在の入力ブロック) と予め復号化された参照フレーム内の同サイズの領域との最良の一致を探索する。そのような一致が見つかり、エンコーダは、参照フレーム内でのこの最良の一致の位置へのポインタの役割を果たす動きベクトルを送信し得る。しかし、BBMEC 探索プロセスは、演算上の理由により、探索対象にできる参照フレームの観点からみて時間的に制限されているだけでなく、探索対象にできる近傍領域の観点からみて空間的にも制限されている。これは、「考えられる最良の」一致が常に見つかるとは限らず、また、高速で変化するデータの場合には特にそうであることを意味する。

【0009】

10

最も単純な形式のBBMEC プロセスは、動き推定の初期設定を、(0, 0) 動きベクトルとする。これは、つまり、ターゲットブロックの初めの推定が、参照フレーム内での同位置のブロックであることを意味する。次に、動き精推定が、この領域の局所近傍においてそのターゲットブロックと最良に一致する (すなわち、そのターゲットブロックに対する誤差が最小になる) 領域を探索することによって行われる。この局所探索は、その局所近傍を網羅的にクエリすることによって行われたり、ダイヤモンドサーチや六角形サーチなどといった幾つかの「高速探索」法のうちの任意のものをを用いて行われたりし得る。

【0010】

MPEG-2 の後発バージョン以降の標準コーデックに設けられた、BBMEC プロセスの改良として、拡張予測区域探索 (enhanced predictive zonal search) (EPZS) 法 (非特許文献 1: Tourapis 達による「単・多フレーム動き推定のための拡張予測区域探索 (Enhanced predictive zonal search for single and multiple frame motion estimation) 」) が挙げられる。この EPZS 法は、ターゲットブロックの初めの推定用として、既に符号化済みの近傍ブロックの動きベクトルならびに前の参照フレームにおける同位置のブロック (および近傍) の動きベクトルに基づく、動きベクトル候補の集合を検討する。EPZS 法は、映像の動きベクトル場が幾らかの空間的及び時間的冗長性を有すると考え、よって、ターゲットブロックについての動き推定の初期設定は、近傍ブロックの動きベクトルまたは符号化済みのフレームにおける近隣ブロックからの動きベクトルとするのが合理的であると仮定する。EPZS 法は、それら初めの推定の集合が集まると、近似レート歪み解析によって当該集合を絞り込む。この後に、動き精推定が行われる。

20

30

【0011】

任意の所与のターゲットブロックについて、エンコーダが、選択候補となる複数のインター予測を生成し得る。これらの予測は、複数の予測プロセス (例えば、BBMEC 方式、EPZS 方式、モデルベース方式等) から生じ得る。また、これらの予測は、ターゲットブロックのサブ区分処理に基づいて異なり得る。サブ区分処理では、相異なる動きベクトルがターゲットブロックの相異なるサブ区分に対応付けられ、かつ、それぞれの動きベクトルが参照フレーム内のサブ区分サイズの領域をそれぞれ指し示す。また、これらの予測は、動きベクトルが指し示す参照フレームに基づいて異なり得る。というのも、前述したように、最近の圧縮規格は、複数の参照フレームの使用を可能にするからである。通常、所与のターゲットブロックについての最良の予測の選択は、レート歪み最適化により達成される。レート歪み最適化では、最良の予測は、レート歪み尺度 $D + R$ (式中、歪み D は、ターゲットブロックと予測との誤差であり、レート R は、予測を符号化するコスト (ビット単位) を定量化し、 λ は、スカラー重み付け係数である。) を最小化する予測とされる。

40

【先行技術文献】

【非特許文献】

【0012】

【非特許文献 1】Tourapis, A., 2002, "Enhanced predictive zonal search for single and multiple frame motion estimation," Proc. SPIE 4671, Visual Communications a

50

nd Image Processing, pp. 1069- 1078

【発明の概要】

【発明が解決しようとする課題】

【0013】

過去、B B M E C 予測の限界を回避する目的で、数多くのモデルベース圧縮スキームが提案されてきた。このようなモデルベース圧縮スキーム（この種のスキームとして、M P E G - 4 P a r t 2 規格が恐らく最も良く知られている）は、映像内のオブジェクトや特徴（一般的に「関心成分」と定義される）の検出及び追跡、さらに、これらの特徴／オブジェクトを映像フレームの残りの部分とは別に符号化する方法を利用する。特徴／オブジェクトの検出／追跡は、標準の動き推定プロセスにおける空間的探索と独立して行われるので、特徴／オブジェクトのトラックは、標準の動き推定により得られるものとは異なる集合の予測を生じさせ得る。

10

【0014】

しかし、特徴／オブジェクトに基づくそのようなモデルベース圧縮スキームでは、映像フレームをオブジェクト領域とオブジェクトでない領域と（あるいは、特徴領域と特徴でない領域と）に分割することによる問題に直面する。第一に、オブジェクトのサイズは多種多様であり得るので、オブジェクトのテクスチャ（カラーコンテンツ）だけでなくオブジェクトの形状も符号化する必要がある。第二に、動きを伴うオブジェクトを複数追跡することは困難であり得て、精度の低いトラッキング（追跡）は不正確なセグメンテーション（分割）を引き起こし、通常、低い圧縮性能につながる。第三の問題は、全ての映像コンテンツがオブジェクトや特徴で構成されとは限らないので、オブジェクト／特徴が存在しない場合には、代替りの符号化スキームが必要となる。

20

【0015】

2014年11月4日出願の同時係属中の米国仮特許出願第61/950,784号（本明細書では「'784出願」と称する）は、上記のセグメンテーション問題を回避するモデルベース圧縮スキームを提示している。「'784出願の連続的ブロックトラッカー（連続的ブロック追跡手段）（C B T）は、オブジェクトや特徴を検出せず、オブジェクトや特徴をオブジェクトでない／特徴でない背景と分割する必要性をなくす。むしろ、C B Tは、フレーム - フレーム間の動き推定を連続的なトラックへと組み込むことにより、映像フレーム内の全ての入力ブロック（「マクロブロック」）をあたかも関心領域であるかの如く追跡する。こうすることにより、C B Tは、インター予測を向上させるというデータの high 次モデリング（モデル化）の恩恵を、セグメンテーション問題を回避しつつ享受するように映像内の動きをモデル化する。

30

【0016】

他のモデルベース圧縮アプローチとして、映像データのコンテンツに対する人間の視覚系（H V S）の応答を、映像フレームのどの部分が人間の知覚にとって最も気付き易いのかを示す重要度マップとしてモデル化するものが挙げられる。重要度マップは、映像フレーム内のそれぞれの入力ブロック又はデータブロックについて数値を取る。また、所与のどのブロックについての重要度マップ値（重要度マップの数値）も、映像をとおしてフレーム - フレーム間で変化し得る。一般的に、重要度マップは、より高い数値がより重要なデータブロックを示すように定義される。

40

【0017】

重要度マップの一種として、時間的コントラスト感度関数（T C S F）（de Lange, H., 1954, "Relationship between critical flicker frequency and a set of low frequency characteristics of the eye（臨界融合周波数と眼の低周波数特性のセットとの関係）," J. Opt. Soc. Am., 44:380-389）が挙げられる。T C S Fは、周期的な刺激に対するH V Sの応答を時間的に測定し、データ内の特定の時間的特性が観測者である人間にとって気付き易いものであることを明らかにする。これらの時間的特性がデータ内の動きに関連付けられて、T C S Fは、データ内で最も気付き易い種類の動きが極めて高い時間的周波数および極めて低い時間的周波数のいずれにも該当しない「中程度」の動きであるこ

50

とを予測する。

【0018】

留意すべき重要な点として、TCSFが、正確な時間的コントラスト値を生成するのに、映像内の動きを伴うコンテンツの速度の正確な測定を必要とすることが挙げられる。このような速度は、カメラの動きおよび/またはオブジェクトの動きによる映像コンテンツの正味の（明らかな）動きを表す、オブティカルフローを算出することによって近似を求めることが可能である。しかし、標準の映像エンコーダの大半は、オブティカルフローを正確に算出するよりも、圧縮効率を最適化する動き推定プロセスを採用している。

【0019】

他種の重要度マップとして、空間的コントラスト感度に基づくものが挙げられ、これは、明るさ、エッジ、空間的周波数、色などの空間的特性に対するHVSの応答を測定する。空間的コントラスト感度関数（SCSF）（例えば、Barten, P., 1999, Contrast Sensitivity of the Human Eye and Its Effects on Image Quality（人間の眼のコントラスト感度および画像品質へのその影響）, SPIE Press等を参照のこと）は、単にコントラスト感度関数（CSF）としても知られており、HVSにとって顕著である空間的コントラストを測定する。SCSFは、JPEG2000画像圧縮規格において、画像圧縮アーチファクトを低減する目的での適用が成功を収めている。オブジェクトや特徴についても、空間的コントラスト手法の支援（例えば、空間的周波数勾配により示されるエッジの存在等）によって典型的に検出される。空間的コントラスト感度は画像圧縮（例えば、JPEG2000コーデック等）においては研究・利用されてきたものの、また、オブジェクト・特徴の検出に基づく映像圧縮プロセスが数多く提案されてきたものの、TCSFで表される時間的コントラスト感度がこれまで映像圧縮に適用されることはなかった。

【課題を解決するための手段】

【0020】

開示する幾つかの発明的実施形態は、映像符号化の品質を向上させるように、重要度マップを映像圧縮に適用する。例示的な一実施形態では、標準の映像符号化処理ストリーム内の時間的周波数が、色空間領域における構造的類似度（SSIM）を用いて波長の近似を求めることにより、かつ、エンコーダの動きベクトル（エンコーダ動きベクトル）を用いて速度の近似を求めることにより算出される。次に、時間的周波数が、時間的コントラスト感度関数（TCSF）への入力としての役割を果たす。TCSFは、全てのデータブロックについて算出され得て、これにより、映像フレームのどの領域が観測者である人間にとって最も気付き易いのかを示す時間的重要度マップを生成し得る。

【0021】

例示的なさらなる実施形態では、エンコーダにより生成された動きベクトルの相対品質についての情報が、符号化プロセスにおける種々の時点で算出され得て、真の動きベクトルマップを生成するのに用いられ得る。真の動きベクトルマップは、それぞれのターゲットブロックについて、その動きベクトルがどれほど信頼できるのかを出力する。0または1の数値を取るこの真の動きベクトルマップは、動きベクトルが正確でないターゲットブロック（すなわち、真の動きベクトルマップが0であるターゲットブロック）にはTCSFが適用されないようにTCSFを洗練化するマスクとして用いられ得る。

【0022】

さらなる実施形態では、空間的複雑度マップ（SCM）が、所与のターゲットブロックの、その近傍に対する空間的コントラストを決定するように、ブロック分散（ブロック内分散）、ブロック輝度、エッジ検出などの尺度から算出され得る。他の実施形態では、SCMからの情報が、複合的な統合化された重要度マップを得るためにTCSFと組み合わせられ得る。この統合化された重要度マップにおける空間的および時間的コントラスト情報の組合せは、人間の視覚的応答の両側面を効果的に両立させる。

【0023】

例示的な一実施形態では、統合化された重要度マップ（TCSFおよびSCMの両方からの情報を含む重要度マップ）が、標準のレート歪み尺度であるD + Rのうちの歪み部

分を重み付けるのに用いられる。これにより、それぞれのターゲットブロックの知覚的相対重要度に合ったソリューション（重要度マップがその最大値に近いときには低歪みソリューションで、重要度マップがその最小値に近いときには低レートソリューション）へと重み付けされた、改変されたレート歪み最適化が得られる。代替的な一実施形態では、上記の目的に、T C S F 又は S C M が独立して用いられ得る。

【 0 0 2 4 】

例示的な他の実施形態では、（真の動きベクトルによる洗練化を伴う）T C S F および S C M が、エンコードのブロックレベル量子化を調整するように用いられ得る。前記重要度マップが高い数値をとるターゲットブロックでは、量子化パラメータがフレーム量子化パラメータに比べて小さくされることで、これらのブロックについては高い品質が得られる。前記重要度マップが低い数値をとるターゲットブロックでは、量子化パラメータがフレーム量子化パラメータに比べて大きくされることで、これらのブロックについては低い品質が得られる。代替的な一実施形態では、上記の目的に、T C S F 又は S C M が独立して用いられ得る。

【 0 0 2 5 】

T C S F は、インター予測を組み込み且つ動きベクトル（映像内のコンテンツの速度の近似を求めるのに T C S F により利用される）を生成するエンコードであればどのようなエンコードの場合にも算出可能であるが、映像圧縮への T C S F の適用は、どの動きベクトルが真の動きベクトルであるのかを正確に決定可能な「784出願の連続的ブロックトラッカー（C B T）」などの、モデルベース圧縮フレームワークにおいて最も効果的となる。前述したように、標準の映像エンコードの大半は、真の動きを反映するよりも圧縮効率を最適化する動きベクトルを算出する。対照的に、C B T は、高い圧縮効率に適した動きベクトルと T C S F の効果を最大化するモデル化情報との両方を提供する。

【 0 0 2 6 】

例示的な一部の発明的実施形態は、得られたビットストリームが、ブロックベースの動き推定を用いて且つその後に残差信号の変換、量子化及びエントロピー符号化が続けられる任意の映像圧縮規格に準拠したものとなるように構築される。そのような映像圧縮規格は、M P E G - 2、H . 2 6 4 および H E V C を含むが、これらに限定されるわけではない。本発明は、ブロックベースでない非標準の映像エンコードであっても、インター予測を組み込み且つ動きベクトルを生成するものであれば、そのような映像エンコードにも適用可能である。

【 0 0 2 7 】

例示的な一部の実施形態は、映像データを符号化する方法及びシステム、ならびにこれを実現するための任意のコーデック（エンコードおよびデコード）を含み得る。複数の映像フレームであって、当該映像フレームが互いに重なり合わないターゲットブロックを有する複数の映像フレームが、エンコードにより処理され得る。前記複数の映像フレームは、重要度マップを用いて、当該重要度マップが量子化を改変（調整）することによって各映像フレーム内の符号化すべき各ターゲットブロックの符号化品質に変更を加えるように、前記エンコードにより符号化され得る。

【 0 0 2 8 】

前記重要度マップは、時間的情報と空間的情報の少なくとも一方を用いて構成され得る。時間的情報と空間的情報との両方が用いられた場合、前記重要度マップは統合化された重要度マップと見なされる。前記重要度マップは、前記複数の映像フレームのうちのある映像フレームの、人間の知覚にとって最も気付き易い部分を示す（特定する、または表す）ように設定され得る。具体的に述べると、前記重要度マップが高い数値をとるブロックでは、ブロック量子化パラメータ（Q P）がフレーム量子化パラメータ $Q P_{frame}$ に比べて小さくされることで、これらのブロックについては高い品質が得られる。また、前記重要度マップが低い数値をとるターゲットブロックでは、前記ブロック量子化パラメータが前記フレーム量子化パラメータ $Q P_{frame}$ に比べて大きくされることで、これらのブロックについては低い品質が得られる。

【 0 0 2 9 】

前記空間的情報は、ルールに基づく空間的複雑度マップ (S C M) により提供され得て、その最初のステップは、前記フレーム内のどのターゲットブロックが当該フレーム内の平均ブロック分散 var_{frame} よりも大きい分散を有するかを決定することである。平均ブロック分散 var_{frame} よりも大きい分散を有するブロックに対して、前記フレーム量子化パラメータ QP_{frame} よりも高い QP 値が振り当てられ得て、このブロック QP の振当量 QP_{block} は、そのブロック分散 var_{block} が var_{frame} よりもいかなる程度大きいかに従って、 QP_{frame} と量子化パラメータ上限 QP_{max} との間で線形的に増減される。

【 0 0 3 0 】

10

好ましくは、前記時間的情報は、どのターゲットブロックが観測者である人間にとって時間的に最も気付き易いかを示す時間的コントラスト感度関数 (T C S F)、および、どのターゲットブロックが前景データに相当するかを示す真の動きベクトルマップ (T M V M) により提供され得る。なお、前記 T C S F は、前景データとして特定されたターゲットブロックについてのみ有効とされ得る。

【 0 0 3 1 】

分散の大きい (high-variance) ブロックは、そのブロックの QP の振当量 QP_{block} が、前記 T M V M がターゲットブロックを前景データとして特定し且つ前記 T C S F のこのブロックについてのコントラスト感度対数値が 0 . 5 未満である場合には QP_{block} が 2 増加するように、前記 T C S F 及び前記 T M V M によりさらに洗練化され得る。

20

【 0 0 3 2 】

前記 S C M は、極めて明るい (1 7 0 超の輝度) か又は極めて暗い (6 0 未満の輝度) ターゲットブロックのブロック量子化パラメータ QP_{block} が QP_{max} に調節し直される輝度マスキングを含み得る。前記 S C M は、符号化された映像の品質レベルに基づく QP_{max} の動的な決定を含み得て、この動的な決定では、イントラ (I) フレーム内のターゲットブロックの平均構造的類似度 (S S I M) 算出結果をこれらフレームの平均ブロック分散 var_{frame} と共に用いて、品質が測定され、測定された品質が低いと、前記量子化パラメータ上限 QP_{max} の数値が前記フレーム量子化パラメータ QP_{frame} に幾分近づくように減らされる。

30

【 0 0 3 3 】

分散の極めて小さい (very-low-variance) ブロックに対して、これらの領域における高品質符号化を確実にするために、前記ブロック分散が小さいほど前記振当量 QP_{block} の数値が低くなるように (、かつ、品質が高くなるように) 決められた低い量子化パラメータ (QP) の値である前記振当量 QP_{block} が振り当てられ得る。分散の極めて小さいブロックに対する前記低い量子化パラメータ (QP) の値である前記振当量 QP_{block} は、最初に、I フレームについて決められ、その後、P フレーム及び B フレームについては $i p r a t i o$ パラメータ及び $p b r a t i o$ パラメータを用いて決められ得る。分散は小さいが、分散が極めて小さいとは見なさないブロックは、当該ブロックについて品質向上が必要か否かを判定するために、前記ブロック量子化パラメータ (QP) の初めの推定値である前記振当量 QP_{block} が現在のブロックの左、左上、右および右上の既に符号化済みの近傍ブロックの量子化パラメータ (QP) の値を平均することによって算出されるように調べられる。前記現在のブロックの前記 S S I M の推定 $S S I M_{est}$ が、前記現在のブロックの左、左上、右および右上の既に符号化済みの近傍ブロックの S S I M 値から算出され得る。 $S S I M_{est}$ が 0 . 9 未満の場合、前記振当量 QP_{block} の数値が 2 減少され得る。

40

【 0 0 3 4 】

一部の実施形態において、前記品質向上は、前記 T M V M により前景データとして特定されて且つ前記 T C S F のコントラスト感度対数値が 0 . 8 超であるブロックにのみ適用される。前記 T M V M は、前景データの場合にのみ 1 に設定され得る。

50

【 0 0 3 5 】

一部の実施形態において、前記 T C S F の時間的周波数は、前記ターゲットブロックとその参照ブロックとの間の色空間領域における S S I M を用いて波長の近似を求めて且つ動きベクトルの大きさ（動きベクトル大きさ）とフレームレートとを用いて速度の近似を求めることによって算出される。

【 0 0 3 6 】

前記 T C S F は、現在のフレームについての当該 T C S F が最近のフレームにおける T C S F マップの重み付き平均であるように且つより最近のフレームがより大きい重み付けを受けるように、複数のフレームにわたって算出され得る。

【 0 0 3 7 】

前景データは、所与のターゲットブロックについてのエンコーダ動きベクトルと当該ブロックについてのグローバル動きベクトルとの差分を算出し、十分に大きい差分を有するブロックが前景データであると判断されることによって特定され得る。

【 0 0 3 8 】

前景データとして特定されたデータブロックについて、前記グローバル動きベクトルから前記エンコーダ動きベクトルが減算されることによって差分動きベクトルを得ることがあり得て、この差分動きベクトルの大きさが前記 T C S F の時間的周波数を算出するのに用いられる。

【 0 0 3 9 】

映像データを処理するコンピュータに基づく方法、映像データを処理するコーデック（エンコーダおよびデコーダ）、ならびに映像データを処理するその他のコンピュータシステム及び装置が、本発明の前述した原理を具現化し得る。

【 0 0 4 0 】

前述の内容は、添付の図面に示す本発明の例示的な実施形態についての、以下のより詳細な説明から明らかになる。図面では、異なる図をとおして同一の参照符号が同一の構成／構成要素を指すものとする。図面は必ずしも縮尺どおりではなく、むしろ、本発明の実施形態を示すことに重点が置かれている。

【図面の簡単な説明】

【 0 0 4 1 】

【図 1】標準のエンコーダの構成を示すブロック図である。

【図 2】一般的なエンコーダの場合のインター予測に伴うステップを示すブロック図である。

【図 3】連続的ブロック追跡による動き初期推定に伴うステップを示すブロック図である。

【図 4】連続的ブロック追跡と拡張予測区域探索との組合せによる統合化された動き推定を示すブロック図である。

【図 5】Wooten 達による時間的コントラスト感度関数の最近の測定結果（2010 年）を示すプロットである。

【図 6】本発明の一実施形態における、C I E 1 9 7 6 L a b 色空間における構造的類似度（S S I M）の算出の様子を示すブロック図である。

【図 7】本発明の一実施形態における、映像符号化の知覚的品質を向上させるための知覚的統計量の一般的な適用の様子を示すブロック図である。

【図 8 A】本発明の一実施形態における、連続的ブロック追跡によるインター予測を改変して映像符号化の知覚的品質を向上させるための知覚的統計量の適用の様子を示すブロック図である。

【図 8 B】重要度マップを用いてブロック量子化を改変して符号化するプロセスの一例を示すブロック図である。

【図 9 A】各実施形態が配備されるコンピュータネットワーク環境の概略図である。

【図 9 B】図 9 A のネットワークにおけるコンピュータノードのブロック図である。

【発明を実施するための形態】

【 0 0 4 2 】

本明細書で引用する全ての特許公報、全ての特許出願公報及び全ての刊行物の全教示内容は、参照をもって本明細書に取り入れたものとする。以下では、本発明の例示的な実施形態について説明する。

【 0 0 4 3 】

本発明は、標準の様々な符号化に適用可能である。以下では、特記しない限り、「従来」や「標準」という語（しばしば、「圧縮」、「コーデック」、「符号化（encoding）」や「エンコーダ」と共に用いられる）は、MPEG-2、MPEG-4、H.264またはHEVCのことを指し得る。「入力ブロック」は、一般性を失うことなく、エンコーダの符号化基本単位のことを指すものとし、しばしば、「データブロック」や「マクロブロック」と同じ意味で称され得る。符号化中の現在の入力ブロックは、「ターゲットブロック」と称される。

10

【 0 0 4 4 】

< 連続的ブロック追跡による映像符号化及びインター予測 >

符号化プロセスは、映像データを、圧縮フォーマットつまり符号化フォーマットに変換するものであり得る。同様に、解凍つまり復号化プロセスは、圧縮された映像を、圧縮される前のつまり未処理のフォーマットに変換するものであり得る。映像圧縮・解凍プロセスは、一般的にコーデックと称されるエンコーダおよびデコーダのペアとして実現され得る。

【 0 0 4 5 】

20

図1は、標準の変換ベース且つ動き補償のエンコーダのブロック図である。図1のエンコーダは、ソフトウェア環境、ハードウェア環境又はこれらの組合せで実現され得る。このエンコーダは、任意の組合せの構成要素を備え得る。これらの構成要素は、インター予測手段20に出力する動き推定手段15、イントラ予測手段30、変換・量子化手段60、逆変換・量子化手段70、ループ内フィルタ80、フレームストア85およびエントロピー符号化手段90を含むが、これらに限定されるわけではない。上記の予測手段（インター予測とイントラ予測との両方）の目的は、所与の入力映像ブロック10（略して「入力ブロック」、あるいは、「マクロブロック」又は「データブロック」）についての最良の予測信号40を生成することである。入力ブロック10から予測信号40が減算されることによって予測残差50を生成し、この予測残差50が変換・量子化60を受ける。その後、この残差の量子化係数65がエントロピー符号化手段90へと渡されて、エントロピー符号化手段90が圧縮ビットストリームへと符号化する。量子化係数65は逆変換・量子化手段70にも渡されて、その結果得られる信号（前記予測残差の近似）が予測信号40に組み戻されて、これによって入力ブロック10についての再構成信号75を生成する。再構成信号75はデブロッキングフィルタなどのループ内フィルタ80に通され得て、この（場合によってフィルタリング済みである）再構成信号がフレームストア85の一部となる。フレームストア85は、今後の入力ブロックの予測を支援する。図1に示すエンコーダの各構成要素の機能は、当業者であればよく知っている。

30

【 0 0 4 6 】

図2に、標準のインター予測（図1の符号30）における各種ステップを示す。インター予測の目的は、新たなデータを、前のフレームからの予め復号化されたデータを用いて、当該データにおける時間的冗長性を利用して符号化することである。インター予測では、現在符号化中のフレーム（ターゲットフレームとも称される）からの入力ブロック10が、図1のフレームストア85に記憶された、予め復号化された参照フレーム内の同サイズの領域から「予測」される。符号化中のフレーム内の入力ブロックの位置と参照フレーム内のこれと一致する領域の位置との（x、y）のずれを示す二成分ベクトルは、動きベクトルと称される。このように、動き推定のプロセスは、符号化すべき入力ブロックを参照フレーム内のこれと一致する領域と最良に結び付ける動きベクトルを決定することを伴う。

40

【 0 0 4 7 】

50

大半のインター予測プロセスは、所与の入力ブロックについての「良好」な動きベクトル 1 1 5 の 1 つ以上の粗推定を生成する動き初期推定（図 2 の符号 1 1 0）から始まる。この後に、任意で、複数の動きベクトル候補が近似レート歪み尺度を用いて単一の候補に低減され得る動きベクトル候補フィルタリングステップ 1 2 0 が続けられる。レート歪み解析では、最良の動きベクトル候補（予測）は、レート歪み尺度 $D + R$ （式中、歪み D は、入力ブロックとこれと一致する領域との誤差であり、レート R は、予測を符号化するコスト（ビット単位）を定量化し、 R は、スカラー重み付け係数である。）を最小化するものが選ばれる。実際のレートコストは、テクスチャビットと動きベクトルビットとの 2 種類の成分を含む。テクスチャビットは、残差信号（入力ブロックから予測を減算したものの）の量子化変換係数を符号化するのに必要なビットの数であり、動きベクトルビットは、動きベクトルを符号化するのに必要なビットの数である。通常、動きベクトルは、既に符号化済みの動きベクトルに対して差分符号化される。エンコーダにおける初期の段階ではテクスチャビットが利用可能でないので、レート歪み尺度のうちのレート部分は、動きベクトルビットにより近似で求められる。一方で、動きベクトルビットは、差分動きベクトルの大きさに依存する動きベクトルペナルティ係数として近似される。したがって、動きベクトル候補フィルタリングステップ 1 2 0 において、この近似レート歪み尺度が、単一の「最良」の初めの動きベクトル又はより少数の集合の「最良」の初めの動きベクトル 1 2 5 を選び出すのに用いられる。次に、そのような初めの動きベクトル 1 2 5 は、動き精推定 1 3 0 により洗練化（refine（さらに改良））される。動き精推定 1 3 0 は、それぞれの初めの推定の近傍において局所探索を行うことにより、その入力ブロックについての動きベクトル（および対応する予測）のより正確な推定を決定する。通常、この局所探索の後に、整数値動きベクトルが内挿により $1/2$ 又は $1/4$ ピクセル精度まで洗練化されるサブピクセル洗練化が続けられる。動き精推定ブロック 1 3 0 は、洗練化済みの動きベクトル 1 3 5 の集合を生成する。

【0048】

次に、動き精ベクトル 1 3 5 が与えられると、モード生成手段 1 4 0 がそのエンコーダが採用し得る符号化モードに基づいて予測候補 1 4 5 の集合を生成する。このようなモードは、コーデックによって異なる。符号化モードが異なるというのは、インターレース対プログレッシブ（フィールド対フレーム）動き推定、参照フレームの方向（前方向予測、後方向予測、双予測）、参照フレームのインデックス（複数の参照フレームを可能にする H. 264、HEVC などのコーデックの場合）、インター予測対イントラ予測（良好なインター予測が存在しない場合にイントラ予測へと戻ることを可能にする一部のシナリオ）、異なる量子化パラメータ、および入力ブロックの異なるサブ区分である（ただし、これらに限定されるわけではない）。予測候補 1 4 5 の集合の全体が、単一の最良の候補を決定するように「最終」のレート歪み解析 1 5 0 を受ける。「最終」のレート歪み解析では、正確なレート歪み尺度 $D + R$ が用いられて、歪み部分用の予測誤差 D （通常、二乗誤差和（SSE）として算出）およびレート部分用の実際の符号化ビット R （図 1 のエントローピー符号化 90 からのもの）を算出する。最終の予測 1 6 0（つまり図 1 の符号 40）は、全ての候補のなかで最小のレート歪みスコア $D + R$ を有する予測であり、この最終の予測が、その動きベクトル及び他の符号化パラメータと共にエンコーダにおける後続のステップへと渡される。

【0049】

図 3 に、インター予測時に、連続的ブロック追跡（CBT）による動き初期推定がどのようにして行われ得るのかを示す。CBT は、ターゲットフレームと時間的予測が導き出される参照フレームとの間に、複数のフレーム分のギャップが存在する場合に有用である。MPEG-2 の場合、IBBPBBP（イントラ予測 I フレーム、双予測 B フレームおよび前方向予測 P フレームからなる）の典型的な GOP 構造は、現在のフレームから最大 3 フレーム分離れた参照フレームを可能にする（その理由は、MPEG-2 では B フレームが参照フレームとして機能できないからである）。符号化すべき各フレームにつき複数の参照フレームを可能にする H. 264 や HEVC では、上記と同じ GOP 構造でも、現

10

20

30

40

50

在のフレームから6フレーム分以上離れた参照フレームを可能にする。より長いGOP構造（例えば、Pフレーム同士の上に7つのBフレーム等）であれば、参照フレームは、ターゲットフレームからなおいっそう離れたものとなり得る。現在のフレームと参照フレームとの間に複数のフレーム分のギャップが存在するとき、連続的追跡は標準の時間的予測手法では捉えられないデータ内の動きをエンコーダが捉えることを可能にするので、CBTはより優れた時間的予測を生成することを可能にする。

【0050】

CBTでの最初のステップは、フレーム-フレーム間追跡（図3の符号210）を実行することである。所与のフレーム内の入力ブロック10ごとに、フレームバッファ205内の前のフレームへの後方向と当該フレームバッファ内の次のフレームへの前方向との両方の動きベクトルが算出される。一実施形態において、フレーム-フレーム間追跡は、再構成された参照フレームではなく元々のソース映像からのフレームに作用する。これは、ソース映像のフレームが量子化や他の符号化アーチファクトによって劣化していないことから、ソース映像のフレームに基づく追跡は映像における真の動き場をより正確に表すので有利だからである。フレーム-フレーム間追跡は、従来のブロックベースの動き推定（BBME）又は階層的な動き推定（HME）を用いて行われ得る。

【0051】

フレーム-フレーム間追跡の結果は、フレーム内のそれぞれの入力ブロックごとに対する、フレームバッファ205内の一番最近のフレーム内での最良に一致する領域と、フレームバッファ205内の一番最近のフレームのブロックごとに対する、現在のフレーム内での最良に一致する領域とを表す、フレーム-フレーム間動きベクトル215の集合である。次に、連続的追跡220が、利用可能なフレーム-フレーム間追跡情報を集約することにより、複数の参照フレームにわたってそれぞれの入力ブロックについての連続的なトラックを生成する。連続的追跡を行う方法の詳細については、全内容を参照をもって本明細書に取り入れた「784出願に記載されている。連続的追跡220の出力は、符号化中の現在のフレーム内の全ての入力ブロックを過去の参照フレーム内のこれらと一致する領域へと追跡する連続的ブロック追跡（CBT）動きベクトル225である。CBTの場合、これらのCBT動きベクトルが初めの動きベクトル（図2の符号125）となり且つ前述したように動き精推定（図2の符号130）により洗練化され得る。

【0052】

図4に、本発明の一実施形態において、CBTがどのようにしてEPZS法と組み合わせられることで統合化された動き推定プロセスを作り出し得るのかを示す。図4では、CBTが動き初期推定110のためにフレーム-フレーム間追跡210及び連続的追跡220により動きベクトルを生成した後、動き精推定130のための局所探索及びサブピクセル洗練化250が続けられる。EPZSが候補生成手段230により初めの動きベクトルを生成した後、先に詳述したような近似レート歪み解析によってフィルタリングを行う候補フィルタリング手段240が続けられる。この後に、さらに、局所探索及びサブピクセル洗練化260による動き精推定130が続けられる。このようにして得られたCBT動きベクトル255およびEPZS動きベクトル265の両方が、総合的な「最良」のインター予測を決定するために残りのインター予測ステップ（図2のモード生成140及び最終のレート歪み解析150）へと渡される。

【0053】

代替的な一実施形態では、図4のCBT動きベクトル候補255及びEPZS動きベクトル候補265に、さらなる候補が追加され得る。このような候補は、ランダム動きベクトル、（0，0）動きベクトル、およびいわゆる「中央値予測子（median predictor）」を含む（ただし、これらに限定されるわけではない）。ランダム動きベクトルには、その局所近傍において最良の候補を見つけ出すように動き精推定130が適用されてもよい。（0，0）動きベクトルは、EPZSの初めの候補のうちの一つであるが、EPZS候補フィルタリング（図4の符号240）後の時点で常に選出されているとは限らず、仮に候補フィルタリング後の時点で選出されていたとしても、動き精推定130によって（

(0, 0)以外の動きベクトルが出力される可能性がある。(動き精推定を受けない)(0, 0)動きベクトルを最終のレート歪み解析用の候補として明示的に含めることは、少なくとも1つの、大きさの小さい且つ「動きの小さい」候補が検討されることを確実にする。同様に、「中央値予測子」もEPZSの初めの候補のうちの一つであるが、EPZS候補フィルタリング(図4の符号240)後の時点で常に選出されているとは限らない。中央値予測子は、現在符号化中のデータブロックの左、上および右上のデータブロックで予め算出された動きベクトルの、中央値として定義される。(動き精推定を受けない)中間値予測子を最終のレート歪み解析用の候補として明示的に含めることは、映像フレームのうちの空間的に均質な(「フラット」な)領域を符号化するうえで特に有益となり得る。つまり、代替的なこの実施形態では、5種類以上の動きベクトル候補(CBT由来の動きベクトル、EPZS由来の動きベクトル、ランダム動きベクトル由来の動きベクトル、(0, 0)動きベクトル、および中央値予測子を含む(ただし、これらに限定されるわけではない))が、残りのインター予測ステップ(図2のモード生成140及び最終のレート歪み解析150)へと渡され得る。

【0054】

<映像符号化用の重要度マップの算出>

知覚的統計量が、映像フレームのどの領域が人間の視覚系(HVS)にとって重要なかを示す重要度マップを算出するのに用いられ得る。

【0055】

知覚的統計量の一例として、時間的に周期的な刺激に対する人間の視覚系(HVS)の応答をモデル化する、いわゆる時間的コントラスト感度関数(TCSF)が挙げられる。背景技術の欄で述べたように、TCSFの概念は(「時間変調伝達関数」として紹介された)1950年代から存在しているが、これまで映像圧縮に適用されることはなかった。図5に、TCSFの最近の測定結果(Wooten, B. 達による2010, "A practical method of measuring the temporal contrast sensitivity function (時間的コントラスト感度関数を測定する実用的な方法)," Biomedical Optical Express, 1(1):47-58)を、周波数の対数の関数としての時間的コントラスト感度の対数(横軸に周波数の対数、縦軸に時間的コントラスト感度の対数)のかたちで示す。測定データ点(図5の丸印)は、3次多項式(図5の実線)を用いてフィッティングされている。なお、後述の全てのTCSF算出に、このフィッティングを用いている。TCSFは、人間の視覚系(HVS)が中程度の周波数域に対して最大の応答を示す一方で、低周波数域ではHVSの応答がわずかに低下し且つ高周波数域では急激に低下するものと予想する。

【0056】

映像圧縮へのTCSFの適用には、TCSFへの入力である時間的周波数(図5の横軸)を算出する方法が必要となる。周波数を算出するための本発明の一実施形態にかかる方法の一つを、次で説明する。周波数 f は、 $f = v / \lambda$ (式中、 v は速度で、 λ は波長である)により与えられる。一実施形態において、任意のデータブロックのコンテンツの速度 v (ピクセル/秒単位)は、エンコーダにより生成された動きベクトル(例えば、図2の符号135、図3の符号215, 225、図4の符号255, 265等)の大きさから $v = |MV| \times \text{フレームレート} / N$ (式中、 $|MV|$ は、そのデータブロックの動きベクトルの大きさであり、フレームレートはその映像が生成された1秒当たりのフレームの数であり、 N は動きベクトルにより指し示される参照フレームと現在のフレームとの間のフレームの数である)として算出され得る。

【0057】

波長 λ の適切な近似は、CIE 1976 Lab色空間(www://en.wikipedia.org/wiki/Lab_color_space)において算出される構造的類似度(SSIM)(Wang, Z. 達による2004, "Image quality assessment: From error visibility to structural similarity (画像品質評価: 誤差可視度から構造的類似度まで)," IEEE Trans, on Image Processing, 13(4):600-612)の算出結果から導き出され得る。図6に、Lab色空間におけるSSIMの算出の様子を示す。SSIMは、ターゲットブロック300(符号化すべき現在

のデータブロック)とその動きベクトルが指し示す参照ブロック310との間で算出される。通常、エンコーダにより処理される映像データはYUV420などの標準の空間で表現されるので、次のステップは、それらターゲットブロック(符号320)および参照ブロック(符号330)の両方を一般的に文献に記載されている任意の手法を用いてCIE 1976 Lab空間に変換することである。次に、Lab空間におけるこれらのターゲットブロックと参照ブロックとの間の誤差E(符号340)が、

【0058】

【数1】

$$\Delta E = \sqrt{(L_T - L_R)^2 + (a_T - a_R)^2 + (b_T - b_R)^2}$$

10

【0059】

(式中、添字Tは「ターゲットブロック」を意味し、添字Rは「参照ブロック」を意味する)として算出される。最後に、誤差Eと同次元のゼロ行列との間のSSIM360が、データの色空間変化の尺度を示すものとして算出される。初めに定まるSSIMは、-1~1の数値を取り、数値の1は完全な類似(空間的な差異がないこと)を示す。SSIMを波長に変換する目的に、0~1の数値を取る空間的非類似度DSSIM=(1-SSIM)/2を使用することが可能であり得て、0は短い波長(最大の空間的類似度)に相当し、1は長い波長(最小の空間的類似度)に相当する。SSIMをピクセル単位に変換するには、SSIMの数値を、算出対象のブロックのピクセルの数で乗算することが可能であり得る。一実施形態では、SSIMのブロックサイズが8×8なので、DSSIM値に64が乗算される。この場合、周波数の最終的な算出結果は、

20

$$f = |MV| \times \text{フレームレート} / (N \times 64 \times (1 - \text{SSIM}) / 2)$$

により与えられる。

【0060】

所与のターゲットブロックについての周波数が算出されれば、このブロックについてのTCSF値が、図5の曲線フィット(実線)から決定可能であり得る。TCSFは、log₁₀スケールで0~1.08または絶対スケールで1~11.97の数値を取る。フレーム内の相異なるブロックが相異なるTCSF値を取ることににより、フレーム内の全てのブロックにわたるTCSF値の集約集合(aggregate set)が重要度マップを形成し、高い数値は時間的コントラストの観点からみて知覚的に重要なブロックを示し且つ低い数値は知覚的に重要でないブロックを示す。

30

【0061】

さらなる実施形態では、最近のフレームからのTCSFの数値が、TCSFベースの重要度マップがフレーム間で変動し過ぎることのないようにそれぞれのデータブロックにつき平均化され得る。例えば、平均TCSF $TCSF_{avg} = 0.7 \times TCSF_{cur} + 0.3 \times TCSF_{prev}$ (式中、TCSF_{cur}は現在のフレームからのTCSF値であり、TCSF_{prev}は一番最近符号化された過去のフレームからのTCSF値である)が挙げられ得る。TCSFの計算は、このように平均化されることでよりロバスト(頑健)になる。

40

【0062】

さらなる実施形態では、エンコーダにより生成された動きベクトルの相対品質についての情報が、符号化プロセスにおける種々の時点で算出され得て、真の動きベクトルマップ(TMVM)を生成するのに用いられ得る。真の動きベクトルマップ(TMVM)は、それぞれのデータブロックについて、その動きベクトルがどれほど信頼できるのかを出力する。0または1の数値を取るこの真の動きベクトルマップは、動きベクトルが正確でないデータブロック(すなわち、TMVM値が0であるデータブロック)にはTCSFが適用されないようにTCSFを洗練化するマスクとして用いられ得る。

50

【 0 0 6 3 】

一実施形態において、動きベクトルの正確さは、所与の映像フレームについてのグローバル動きモデルを推定し且つこの動きモデルを当該フレーム内のそれぞれのデータブロックに適用することによってそれぞれのデータブロックについてのグローバル動きベクトルを決定してから且つこのグローバル動きベクトルをそのデータブロックについてのエンコードの動きベクトル（エンコード動きベクトル）と比較することにより、決定され得る。グローバル動きは、そのフレームからの符号化動きベクトルの集約集合であって、6つのパラメータ又は8つのパラメータのアフィン動きモデルにフィッティングされた集約集合から推定され得る。所与のデータブロックについてグローバル動きベクトルとエンコード動きベクトルとが同一である（又は類似する）場合、エンコード動きベクトルが正確であると見なされる（そして、そのデータブロックについて $T M V N = 1$ となる）。それら2つのベクトルが同一でない場合、それらの（二乗誤差和（ $S S E$ ）又は差分絶対値和（ $S A D$ ）で測定された）予測誤差同士を比較してもよい。一方の誤差が小さくて他方の誤差が大きい場合には、誤差が小さいほうの動きベクトルが、符号化に使われて且つ正確と見なされる（ $T M V M = 1$ ）。

10

【 0 0 6 4 】

代替的な一実施形態では、所与のデータブロックについてグローバル動きベクトルとエンコード動きベクトルとの差分の大きさが、そのデータブロックが前景データであること（これは、そのデータブロック内のコンテンツが、フレームの残りの部分（背景）と異なる動きを伴っていることを意味する）を特定するのに用いられる。この実施形態では、 $T M V M$ が1に設定されて、 $T C S F$ が前景データの場合にのみ適用される。さらなる実施形態では、前景データとして特定されたデータブロックについて、グローバル動きベクトルからエンコード動きベクトルが減算されることによって差分動きベクトルを得て、（エンコード動きベクトルではなくて）この差分動きベクトルの大きさが $T C S F$ の周波数を算出するのに用いられる（前述の式において、 $| M V |$ を $| D M V |$ （ $D M V$ は差分動きベクトルである）に置き換える）。

20

【 0 0 6 5 】

他の実施形態では、動きベクトル対称度が、 $T M V M$ を洗練化するのに用いられ得る。動きベクトル対称度（Bartels, C.及びde Haan, G.による2009, "Temporal symmetry constraints in block matching（ブロックマッチングにおける時間的対称度制約）," Proc. IEEE 13th Int'l. Symposium on Consumer Electronics, pp. 749-752）は、動き推定の時間的方向を切り替えたときに互いに対をなす、動きベクトルのペアの相対対称度として定義され、算出された動きベクトルの品質の尺度となる（対称度が高ければ高いほど、動きベクトルの品質が優れている）。「対称度誤差ベクトル」は、前方向動き推定により得られた動きベクトルと後方向動き推定により得られた動きベクトルとの差分として定義される。動きベクトル対称度が低いこと（対称度誤差ベクトルが大きいこと）は、しばしば、オクルージョン（あるオブジェクトが別のオブジェクトの前側を動くことにより、その背景オブジェクトを隠したり露わにしたりすること）、オブジェクトの動きが映像フレーム上に又は映像フレーム外になること、照明変化など（いずれも、正確な動きベクトルを導き出すことを困難にする）の複雑な現象が存在することの指標となる。

30

40

【 0 0 6 6 】

一実施形態では、対称度誤差ベクトルの大きさが符号化中のデータブロックの範囲の半分よりも大きい場合（例えば 16×16 マクロブロックのときには、大きさが（8, 8）ベクトルよりも大きい場合）に、低対称度（対称度が低い）と判断される。他の実施形態では、対称誤差ベクトルの大きさが追跡プロセス時に導き出された動きベクトル統計量に基づく閾値（例えば、現在のフレーム又は最近のフレーム同士の所与の組合せにおける、動きベクトル大きさ（動きベクトルの大きさ）の平均値に動きベクトル大きさの標準偏差の倍数を足したものと等）よりも大きい場合に、低対称度（対称度が低い）と判断される。一実施形態では、動きベクトルが上記の定義で低対称度を有するデータブロックに対して $T M V M$ 値 = 0 が自動的に振り当てられて、他のデータブロックはグローバル動きベクトル

50

ルとエンコーダ動きベクトルとの比較に由来するそれまでのTMVM値を維持する。

【0067】

フラットなブロックは、高い空間的コントラスト感度を有するものの、動きベクトルを算出する際によく知られている開口問題 (aperture problem)

【0068】

【数2】

(http://en.wikipedia.org/wiki/Motion_perception#The_aperture_problem を参照のこと)

【0069】

が原因となり、信頼できない動きベクトルを生じる傾向にある。フラットなブロックは、例えば、エッジ検出プロセス (データブロック内においてエッジが検出されなかった場合にフラットなブロックであると判断される) を用いて、あるいは、データブロックの分散を閾値と比較すること (この閾値よりも小さい分散がフラットなブロックを示す) によって検出され得る。一実施形態では、ブロックのフラットさが、前述のように算出されたTMVM値を変更するように用いられ得る。例えば、フラットなブロックであると検出されたブロックには、TMVM値 = 0 が振り当て直され得る。

【0070】

一実施形態では、TMVMが、信頼できる動きベクトルを有するか否かに影響されるTCSFを洗練化するマスクとして用いられ得る。TMVMの数値は0又は1なので、あるブロックについてのTMVM値を、そのブロックについてのTCSF値へとブロック毎に乗算することには、TCSFをマスクする効果がある。TMVM値が0であるブロックの場合、TCSFの算出に必要な動きベクトルが信頼できないことになるので、そのTCSFが「無効」にされる。TMVM値が1であるブロックの場合、TCSF算出結果が信頼できるとみなされて且つこれまでに述べた任意の手法が確信して利用される。

【0071】

他の実施形態では、時間的コントラストマップ (前述のTCSF) の代わりに、あるいは、これに加えて、空間的コントラストマップが生成され得る。本発明では、空間的コントラスト (これの反対は「空間的複雑度」と称される) を測定するのに、単純な尺度が用いられる。一実施形態では、データの輝度成分と色差成分との両方について測定されるブロック分散が、所与の入力ブロックの空間的複雑度を測定するのに用いられる。分散の大きい入力ブロックは、空間的に複雑であり且つHVSにとって気付き難いと考えられるので、その空間的コントラストは小さいことになる。

【0072】

他の実施形態では、データの輝度成分について測定されるブロック輝度が、空間的複雑度の分散測定結果を洗練化するのに用いられる。分散は小さい (空間的複雑度が低い、空間的コントラストが大きい) が極めて明るい又は極めて暗い入力ブロックは、空間的コントラストが小さいと自動的に見なされて且つその前に大きいと測定された空間的コントラストを上書きする。その理由は、極めて暗い領域や極めて明るい領域が、HVSにとって気付き難いからである。所与のブロックを極めて明るい又は極めて暗いと分類するための輝度閾値は、その時々用途に特化したものとされるが、8ビットの映像の場合の典型的な数値は、極めて明るい「170超」で、極めて暗いが「60未満」である。

【0073】

上記のようにブロック輝度により改変されたブロック分散が、空間的コントラストの観点からHVSにとっての気付き易さの高い領域及び気付き易さの低い領域を示す空間的コントラストマップ (SCM) を形成するために、映像フレームの全ての入力ブロックについて算出され得る。

【0074】

一実施形態では、SCMが、(TMVMにより洗練化された)TCSFと組み合わせられることによって、統合化された重要度マップを形成し得る。この統合化されたマップは、

10

20

30

40

50

例えば、SCMおよびTCSFの両方を適宜正規化したうえで、所与のブロックについてのSCM値をそのブロックについてのTCSF値へとブロック毎に乗算することによって形成され得る。他の実施形態では、TCSFの代用としてSCMが使用され得る。他の実施形態では、SCMが、TCSFを洗練化するのに用いられ得る。例えば、複雑度の高いブロックではSCM値がそのブロックについてのTCSF値を上書きし得て、複雑度の低いブロックではそのブロックについてのTCSF値が直接使用され得る。

【0075】

<映像符号化への重要度マップの適用>

前述した重要度マップは、一般的なエンコーダ(図2)及びCBTエンコーダ(図3)のいずれの映像符号化プロセスにも適用され得て、符号化ビットストリームの品質を向上させる。

10

【0076】

図7に、映像符号化への重要度マップの一般的な適用の様子を示す。入力映像フレーム5およびフレームストア85が、知覚的統計量390を生成するように使用される。そして、知覚的統計量390が、前述したような(TMMにより洗練化された)TCSFおよび/またはSCMの重要度マップ400を形成するように適用される。知覚的統計量390は、動きベクトル大きさ、ブロック分散、ブロック輝度、エッジ検出、およびグローバル動きモデルパラメータを含み得る(ただし、これらに限定されるわけではない)。入力映像フレーム5およびフレームストア85は、さらに、符号450での映像フレームの符号化へと通常どおり入力される。その符号化は、通常の符号化ステップ(図1の動き推定15、インター予測20、イントラ予測30、変換・量子化60およびエントロピー符号化90)を含む。ただし図7では、符号化450が、後述の方法で重要度マップ400により機能拡張される。

20

【0077】

図8Aに、CBTを用いた映像符号化を向上するための、重要度マップの具体的な適用の様子を示す。図8Aには、CBTからのフレーム-フレーム間追跡210ステップ及び連続的追跡220ステップによる動き初期推定(図2の符号110)が示されている。そして、動き精推定130が、既述したものと同一局所探索及びサブピクセル洗練化の動き精推定ステップ(図4の符号250)で、グローバルCBT動きベクトル225に適用される。ここでも、その後に、エンコーダが採用し得る符号化モードに基づいて予測候補145の集合を生成するモード生成手段140が続く。図4と同じく、EPZSおよびモデルベースでない他の候補(例えば、(0,0)動きベクトル、中央値予測子等)も、統合化された動き推定フレームワークの一部として並行して生成され得る(図8Aでは、図示を簡略化するためにこれら他の候補の図示を省略している)。図8Aにおいても、CBT候補のあらゆる符号化モードおよび場合によってはモデルベースでない他の候補のあらゆる符号化モードも含め、予測候補145の集合の全体が、単一の最良の候補を決定するように「最終」のレート歪み解析155を受ける。「最終」のレート歪み解析では、正確なレート歪み尺度 $D + R$ が用いられて、歪み部分用の予測誤差 D およびレート部分用の実際の符号化ビット R (図1のエントロピー符号化90からのもの)を算出する。最終の予測160(または図1の符号40)が、その動きベクトル及び他の符号化パラメータと共にエンコーダにおける後続のステップへと渡される。

30

40

【0078】

図8Aにおいて、知覚的統計量390が、フレーム-フレーム間動き追跡210から導き出された動きベクトルから算出され得て、その後、前述したような重要度マップ400を形成するように適用され得る。そして、これらの重要度マップ400が、最終のレート歪み解析155へと入力される。ここでも、知覚的統計量390は、動きベクトル大きさ、ブロック分散、ブロック輝度、エッジ検出、およびグローバル動きモデルパラメータを含み得る(ただし、これらに限定されるわけではない)。

【0079】

一実施形態では、重要度マップが、レート歪み最適化条件を当該重要度マップに応じて

50

改変するように用いられる。標準のエンコーダ（図2を参照のこと）では、所与の入力ブロック10についての予測候補145の集合の全体が、単一の最良の候補を決定するように「最終」のレート歪み解析150を受ける。「最終」のレート歪み解析では、正確なレート歪み尺度 $D + R$ が用いられて、歪み部分用の予測誤差 D およびレート部分用の実際の符号化ビット R （図1のエントロピー符号化90からのもの）を算出する。レート歪み尺度 $D + R$ のスコアが最小である候補が、所与の入力ブロック10についての最終の予測160となる。本発明の一実施形態において、図7又は図8の知覚的に最適化されたエンコーダの場合、符号400で重要度マップ IM が算出されて、かつ、最終のレート歪み解析155が改変されたレート歪み尺度 $D \times IM + R$ を使用する。この改変されたレート歪み尺度では、所与の入力ブロックについての IM 値が歪み項に乗算されて、 IM 値が高ければ高いほど、歪みの小さいソリューションにより大きな重要度が振り当てられる。

というのも、高い IM 値は、対応する入力ブロックが知覚的に重要であることを示すからである。重要度マップは、（場合によっては $TMVM$ 値により洗練化されている） $TCSF$ 、 SCM またはこれらを複合したものを含み得る。

【0080】

さらに実施形態では、上記に加えて、レート歪み尺度における歪み D が、 SSE （二乗誤差和：歪みを算出する「標準」の手法）と YUV 空間において算出された $SSIM$ との重み付け和として算出され得る。重み付けは、映像のうちの最初の幾つか（又は一番最近の幾つか）のフレームにおける平均 $SSIM$ 値 $SSIM_{avg}$ が当該映像のうちの当該最初の幾つか（又は当該一番最近の幾つか）のフレームにおける平均 SSE 値 SSE_{avg} と等しくなる（ $\times SSIM_{avg} = SSE_{avg}$ ）ように適応的に算出され得る。つまり、それぞれの入力ブロックについて、改変されたレート歪み尺度は、 $(SSE + \times SSIM) \times IM + 2R$ （式中、 R 項の前にある乗算係数2は、歪み項が2つあることを意味している）となる。歪み測定に $SSIM$ を含めることは、 $SSIM$ がデータの構造的情報に相当することから、レート歪み最適化において HVS 知覚が占める割合をなおいっそう増やすことになる。

【0081】

他の実施形態では、重要度マップ（例えば、 $TMVM$ による洗練化を伴う $TCSF$ 、 SCM 等）が、レート歪み最適化を改変するのに加えて（あるいは、この代わりに）、エンコーダのブロックレベル量子化を改変するように用いられ得る。量子化は、所与の入力ブロックが符号化される相対品質を制御するものである。つまり、高度に量子化されたデータは低品質の符号化済み出力となり、低度量子化されたデータは高品質の符号化済み出力となる。量子化の量は、量子化パラメータ QP によって制御される。標準のエンコーダは、異なるフレームタイプに異なる QP 値 QP_{frame} を振り当てて、 I フレームは最小の QP （最も高品質）で符号化されて且つ B フレームは最大の QP （最も低品質）で符号化されて且つ P フレームは中間の QP （中間の品質）で符号化される。

【0082】

つまり、上記の手法は、複数の映像フレーム（当該映像フレームは、互いに重なり合わないターゲットブロックを有している）を、重要度マップを用いて、各映像フレーム内の各ターゲットブロックの量子化を改変すること（そして、これによってその符号化品質に影響を与えること）によって符号化する方法を提示している。そのような重要度マップは、時間的情報（ $TMVM$ による洗練化を伴う $TCSF$ ）、空間的情報、またはこれら2種類の組合せ（すなわち、統合化された重要度マップ）を用いて設定され得る。重要度マップは各映像フレームのどの部分が人間の知覚にとって最も気付き易いのかを示すものなので、前記重要度マップの数値は、それぞれのターゲットブロックについての QP を、（ i ）当該重要度マップが高い数値をとるブロックでは、ブロック QP が QP_{frame} に比べて小さくされることで、これらのブロックについては高い品質となるように、かつ、（ ii ）当該重要度マップが低い数値をとるブロックでは、前記ブロック QP が前記フレーム量子化パラメータ QP_{frame} に比べて大きくされることで、これらのブロックについては低い品質となるように、変更するのが望ましい。

【0083】

図8Bに、重要度マップ400を用いて符号化時の量子化を改変するプロセスの一例を示す。符号400では、知覚的統計量390から導き出された時間的情報および/または空間的情報を用いて重要度マップが構成/形成され得る。時間的情報は、例えば、どのターゲットブロックが観測者である人間にとって時間的に最も気付き易いのかを示す時間的コントラスト感度関数(TCSF)、および、どのターゲットブロックが前景データに相当するのかを示す真の動きベクトルマップ(TMVM)により提供され得て、前記TCSFは、前景データとして特定されたターゲットブロックについてのみ有効とされる。空間的情報は、例えば、ルールに基づく空間的複雑度マップ(SCM)により提供され得る。

【0084】

そして、重要度マップ400は、符号化450内の量子化ステップ430を上記のように改変するのに用いられる。当該重要度マップが高い数値を取るブロックでは、ブロック量子化パラメータ(QP)が、フレーム量子化パラメータ QP_{frame} に対して減らされて、これらのブロックについては高い符号化品質が得られる。当該重要度マップが低い数値を取るブロックでは、前記ブロック量子化パラメータが、前記フレーム量子化パラメータ QP_{frame} に対して増やされて、これらのブロックについては低い符号化品質が得られる。重要度マップからの情報を用いることにより、量子化は、各映像フレーム内の符号化すべき各ターゲットブロックの符号化品質を向上させるように改変され得る。

【0085】

一実施形態では、所与のフレームについてのTCSFマップが、フレームQPをブロック毎に調節するように用いられ得る。ブロックQP、 QP_{block} を算出する方法の一つは、(Li, Z.達による2011, "Visual attention guided bit allocation in video compression (映像圧縮における、視覚注意によって導かれたビット振当), J. of Image and Vision Computing, 29(1): 1-14)の手法に従って、その調節量をフレームにおけるTCSFマップ全体と関連付けることである。その結果得られる式は、 $QP_{block} = (TCSF_{frame} / (TCSF_{block} \times M)) \times QP_{frame}$ (式中、 $TCSF_{frame}$ はフレーム内の全てのブロックについてのTCSF値の合計であり、 QP_{frame} はフレームQPであり、Mはフレーム内のブロックの数である)により与えられる。さらなる実施形態では、乗算係数($TCSF_{frame} / (TCSF_{block} \times M)$)が、 QP_{block} の最終的な数値が QP_{frame} に対して大きくなり過ぎたり小さくなり過ぎたりすることのないように増減され得る。

【0086】

代替的な一実施形態では、TCSFマップによるQPのブロック毎の調節が、そのフレームについてのTCSFマップ全体を参照することなく達成され得る。この実施形態では、 QP_{block} の計算がより単純になる： $QP_{block} = QP_{frame} / TCSF_{block}$ 。一実施形態では、 QP_{block} の得られる数値が、そのフレームについての所定の上限QP値を上回ったり下限QP値を下回ったりしないように範囲限定される(クリップされる)： $QP_{min} \leq QP_{block} \leq QP_{max}$ 。

【0087】

他の実施形態では、SCMの出力が、ルールに基づくアプローチを用いてブロック毎に量子化パラメータを改変するように用いられ得る。この実施形態は、まず、分散の大きいブロックに高いQP値(低品質)を振り当てることから始まる。というのも、高度に複雑な領域は、HVSにとって気付き難いからである。分散の小さいブロックには、低いQP値(高品質)が振り当てられる。というのも、低度に複雑な領域は、HVSにとって気付き易いからである。一実施形態では、所与のブロックについてのQP振当量が、フレームの上限QP値である QP_{max} 及び下限QP値である QP_{min} によって規制されており、かつ、そのフレーム内の他のブロック分散に対する自身のブロック分散に基づき線形的に増減される。代替的な一実施形態では、フレーム全体の平均分散よりも大きい分散を有するブロックのみに、フレームQPである QP_{frame} と QP_{max} との間のQP値が振り当てられて、かつ、その振当量は $QP_{block} = ((var_{block} - var_f$

10

20

30

40

50

$$r a m e / v a r b l o c k)) \times (Q P_{m a x} - Q P_{f r a m e}) + Q P_{f r a m e}$$
 となるように線形的に増減される。この代替的な実施形態において、分散の大きいブロックに対するQP振当量は、TC SFによりさらに洗練化されてもよい。例えば、TMVMでそのブロックが前景データと見なされて且つTC SFのコントラスト感度対数値（図5の縦軸）が0.5未満である（そのブロックが時間的に重要でないことを意味する）場合には、 $Q P_{b l o c k}$ が2だけ増やされる。代替的な一実施形態では、エッジ検出プロセスが適用され得て、エッジを含むブロックのQPが、それまでに振り当てられていた空間的複雑度からのQPを上書きするように $Q P_{m i n}$ に調節され得る。その理由は、エッジが、HVSにとって極めて気付き易いものだからである。さらなる実施形態では、極めて明るい又は極めて暗いブロックのQPが、それまでに振り当てられていた分散及び（場合によっては）エッジ検出からのQPを上書きすることによって $Q P_{m a x}$ に調節し直され得る。その理由は、極めて暗い領域や極めて明るい領域が、HVSにとって気付き難いからである。このプロセスは、輝度マスキングとして知られる。

【0088】

さらなる実施形態では、上記に加えて、分散の大きいブロックについての $Q P_{m a x}$ の数値が、符号化された映像の品質レベルに基づいて動的に決定され得る。その思想は、低品質の符号化では分散の大きいブロックにおける品質低下を許容できないので $Q P_{m a x}$ は $Q P_{f r a m e}$ により近づけるのが望ましい一方、高品質の符号化ではビットを節約するために分散の大きいブロックについての $Q P_{m a x}$ を増やすことを許容できるというものである。符号化の品質は、各I（イントラ）フレーム毎に、平均フレーム分散の $\pm 5\%$ 以内の分散を有するブロックの平均SSIMを算出することによって更新され得て、かつ、SSIM値が高ければ高いほど $Q P_{m a x}$ のより高い数値に対応するようにされる。代替的な一実施形態では、品質指標が平均SSIMと平均フレーム分散との積として算出されるように、平均SSIMがそのフレームの平均分散によって調節される。

【0089】

さらなる実施形態では、上記に加えて、分散の極めて小さいブロック（HVSにとって特に可視的であるフラットな領域に相当）に対して、これらの領域における高品質符号化を確実にするために、決まった低いQP値が振り当てられ得る。例えば、I（イントラ）フレームの場合、0～10の分散を有するブロックに $Q P = 28$ が振り当てられ得て、10～30の分散を有するブロックに $Q P = 30$ が振り当てられ得て、30～60の分散を有するブロックに $Q P = 32$ が振り当てられ得る。それから、P及びBフレーム内のブロックに対するQP振当量が、上記のQPからそれぞれ*i p r a t i o*（*i p*率）パラメータ及び*p b r a t i o*（*p b*率）パラメータを用いて導き出され得る。

【0090】

さらなる実施形態では、上記に加えて、分散の小さいブロック（例えば、60～平均フレーム分散の分散を有するブロック等）に対してフレーム $Q P_{f r a m e}$ が振り当てられて、それから、当該分散の小さいブロックが、さらなる品質向上が必要なのか否かを判定するように調べられる。一実施形態では、ブロックノイズ（blockiness）アーチファクトを、符号化中の現在の（ターゲット）ブロックからの再構成されたピクセル及び元々のピクセルの空間的複雑度及び輝度を符号化済みの周囲のブロック（例えば、左、左上、上、右上（これらが存在する場合）のブロック等）の空間的複雑度及び輝度と比較することによって検出し得る。仮に、ターゲットブロックの再構成されたピクセルの空間的複雑度尺度及び輝度尺度と近傍ブロックの対応する尺度との間には大きな違いがあるものの、そのターゲットブロックの元々のピクセルとその近傍ブロックの元々のピクセルの間には空間的複雑度及び輝度にそのような違いがない場合には、そのターゲットブロックが「ブロックノイズ（blocky）」であると見なされる。この場合、そのブロックのQP値が、当該ブロックの符号化品質を向上させるように減らされる（例えば、2だけ減らされる）。他の実施形態では、ターゲットブロックの推定品質が、符号化済みの周囲のブロック（例えば、左、左上、右、右上（これらが存在する場合）のブロック等）のSSIM値及びQP値を平均化することによって算出される。その平均QP値 $Q P_{a v g}$ が、そのターゲ

10

20

30

40

50

ットブロックについての推定 QP_{block} とされる。平均 $SSIM$ 値 $SSIM_{est}$ が 0.9 未満であると、 $QP_{block} = QP_{avg}$ が 2 だけ減らされてその品質を向上させる。さらなる実施形態において、 $TMVM$ により前景データとして特定されたターゲットブロックは、その $TCSF$ のコントラスト感度対数値（図 5 の縦軸）が 0.8 超である（そのブロックが時間的に重要であることを意味する）場合にのみ、 QP_{block} が 2 だけ減らされる。

【0091】

これまでに説明した方法は、時間的重要度マップ（ $TMVM$ による洗練化を伴うか又は伴わない $TCSF$ ）、空間的重要度マップまたは両方を使用し得る。時間的重要度マップと空間的重要度マップとの両方が用いられた場合、その結果生じるものは、統合化された重要度マップと称される。

【0092】

前述したような知覚的統計量から生成された重要度マップは、動き補償を用いて動きベクトルを生成する映像圧縮フレームであればどのような映像圧縮フレームにも適用可能であり得て、これにより、同じ符号化サイズで視覚的により優れた符号化を作り出すようにレート歪み解析と量子化との両方が向上される。映像圧縮への重要度マップの適用は、既に詳述した連続的ブロックトラッカー（ CBT ）に適用するうえで特殊な適用を必要としない。しかも、 CBT は、どの動きベクトルが真の動きベクトルであるのかを正確に決定するという追加の能力を提供するので、重要度マップは CBT ベースの符号化フレームワークにおいてより効果的である。その具体的な理由として、 CBT のフレーム - フレーム間動きベクトル（図 8 A のフレーム - フレーム間追跡 210 からのもの）が、映像の元々のフレームから生成されたものであって再構成されたフレームから生成されたものではない点が挙げられる。一般的なエンコードの場合の図 2 及び図 7 のフレームストア 85 は符号化プロセスから生成された再構成後のフレームを含むのに対し、図 3、図 4 及び図 8 A のフレームストア 205 は元々の映像フレームを含んでいる。そのため、 CBT のフレーム - フレーム間追跡（図 3、図 4 及び図 8 の符号 210）は映像の真の動きをより良好に追跡することが可能であり、かつ、そのフレーム - フレーム間動きベクトルはより正確な真の動きベクトルマップを生成する。対照的に、一般的なエンコードの動きベクトルは、レート歪み（圧縮）性能を最適化するように選択されており、映像の真の動きを反映しない可能性がある。

【0093】

なお、生成された重要度マップは、イントラ予測フレームにも、これまでに述べた手法に従ってイントラ予測モード間のレート歪み最適化を改変するか又はブロックレベル量子化を改変することによって適用可能であり得る。ただし、オールイントラエンコード（全イントラエンコード）の場合には、 $TCSF$ を算出するうえで、映像フレーム内のそれぞれのデータブロックについての動きベクトルを生成するための別個の符号化手段（例えば、図 8 A のフレーム - フレーム間追跡 210 等）が必要となる。

【0094】

< デジタル処理環境 >

本発明の例示的な実装は、ソフトウェア環境でもファームウェア環境でもハードウェア環境でも実現可能であり得る。図 9 A に、そのような環境の一つを示す。少なくとも 1 つのクライアントコンピュータ / デバイス 950（例えば、携帯電話、コンピューティングデバイス等）およびクラウド 960（またはサーバコンピュータもしくはサーバコンピュータのクラスタ）は、アプリケーションプログラムを実行する処理機能、記憶機能、符号化機能、復号化機能および入出力装置などを提供する。

【0095】

また、少なくとも 1 つのクライアントコンピュータ / デバイス 950 は、通信ネットワーク 970 を介して、他のクライアントデバイス / プロセス 950 および少なくとも 1 つのサーバコンピュータ 960 を含む他のコンピューティングデバイスと接続可能であり得る。通信ネットワーク 970 は、リモートアクセスネットワークの一部、グローバルネッ

10

20

30

40

50

トワーク（例えば、インターネット等）の一部、世界規模のコンピュータの集まりの一部、ローカルエリアネットワークの一部、ワイドエリアネットワークの一部、あるいは、現在それぞれのプロトコル（TCP/IP、Bluetooth（登録商標）など）を用いて相互通信するゲートウェイの一部であり得る。それ以外の電子デバイス/コンピュータネットワークアーキテクチャも使用可能である。

【0096】

本発明の実施形態は、映像又はデータ信号情報を符号化、追跡、モデル化、フィルタリング、調整、復号化又は表示する手段を含み得る。図9Bは、そのような映像又はデータ信号情報の符号化を促進するのに用いられ得る、図9Aの処理環境における所与のコンピュータ/コンピューティングノード（例えば、クライアントプロセッサ/デバイス/携帯電話デバイス/タブレット950、サーバコンピュータ960等）の内部構造の図である。各コンピュータ950、960は、コンピュータ又は処理システムの構成要素間のデータ転送に用いられる実在する又は仮想的なハードウェアラインのセットであるシステムバス979を備える。バス979は、コンピュータシステムの相異なる構成要素（例えば、プロセッサ、エンコーダチップ、デコーダチップ、ディスクストレージ、メモリ、入力/出力ポート等）同士を接続する共有の配管のようなものであり、それら構成要素間のデータのやり取りを可能にする。システムバス979には、様々な入出力装置（例えば、キーボード、マウス、ディスプレイ、プリンタ、スピーカ等）をコンピュータ950、960に接続するための入出力装置インターフェース982が取り付けられている。ネットワークインターフェース986は、コンピュータがネットワーク（例えば、図9Aの符号970で示されるネットワーク等）に取り付けられた他の様々なデバイスと接続することを可能にする。メモリ990は、本発明のソフトウェア実装を実現するのに用いられるコンピュータソフトウェア命令992及びデータ994を記憶する揮発性メモリである。

【0097】

ディスクストレージ995は、本発明の一実施形態を実現するのに用いられるコンピュータソフトウェア命令998（等価的には「OSプログラム」）及びデータ994を記憶する不揮発性ストレージである。また、ディスクストレージ995は、映像を圧縮フォーマットで長期的に記憶するのに使用され得る。システムバス979には、さらに、コンピュータ命令を実行する中央演算処理装置984も取り付けられている。なお、本明細書をとおして「コンピュータソフトウェア命令」と「OSプログラム」は互いに等価物とする。

【0098】

一例として、エンコーダは、時間的情報や空間的情報から形成された重要度マップを用いて映像データを符号化するためのコンピュータ読取り可能な命令992により構成され得る。これらの重要度マップは、映像データの符号化/復号化を最適化するための、エンコーダ（又はエンコーダの構成要素）へのフィードバックループを提供するように構成され得る。

【0099】

一実施形態において、プロセッサルーチン992及びデータ994は、エンコーダ（概して符号992で示す）を備えるコンピュータプログラムプロダクトである。このようなコンピュータプログラムプロダクトは、そのエンコーダ用のソフトウェア命令の少なくとも一部を提供する、ストレージ装置994に記憶可能なコンピュータ読取り可能な媒体を含む。

【0100】

コンピュータプログラムプロダクト992は、当該技術分野において周知である任意の適切なソフトウェアインストール方法によってインストール可能なものであり得る。また、他の実施形態において、前記エンコーダの前記ソフトウェア命令の少なくとも一部は、ケーブルおよび/または通信および/または無線接続を介してダウンロード可能なものであり得る。他の実施形態において、エンコーダシステムソフトウェアは、非過渡的なコンピュータ読取り可能な媒体に組み込まれたコンピュータプログラム伝播信号プロダクト9

07 (図9A)であり、当該コンピュータプログラム伝播信号プロダクト907は、実行されると、伝播媒体上の伝播信号(例えば、電波、赤外線波、レーザ波、音波、インターネットなどのグローバルネットワークや他の少なくとも1つのネットワークによって伝播される電気波など)として実現され得る。このような搬送媒体又は搬送信号が、本発明にかかるルーチン/プログラム992用のソフトウェア命令の少なくとも一部を提供する。

【0101】

代替的な実施形態において、前記伝播信号は、伝播媒体によって搬送されるアナログ搬送波またはデジタル信号である。例えば、前記伝播信号は、グローバルネットワーク(例えば、インターネット等)、電気通信網または他のネットワークによって搬送されるデジタル信号であり得る。一実施形態において、前記伝播信号は、所与の期間のあいだ伝播媒体によって送信されるものであり、例えば、数ミリ秒、数秒、数分またはそれ以上の期間のあいだネットワークによってパケットで送信される、ソフトウェアアプリケーション用の命令等であり得る。他の実施形態において、コンピュータプログラムプロダクト992の前記コンピュータ読取り可能な媒体は、コンピュータシステム950が受け取って読み取りし得る伝播媒体である。例えば、コンピュータシステム950は、前述したコンピュータプログラム伝播信号プロダクトの場合のように、伝播媒体を受け取ってその伝播媒体内に組み込まれた伝播信号を特定する。

【0102】

本発明を例示的な実施形態を参照しながら具体的に図示・説明したが、当業者であれば、添付の特許請求の範囲に包含された本発明の範囲を逸脱しない範囲で形態や細部に様々な変更を施せることを理解するであろう。

なお、本発明は、実施の態様として以下の内容を含む。

〔態様1〕

複数の映像フレームを符号化する方法であって、

前記映像フレームは、互いに重なり合わないターゲットブロックを有しており、

当該方法は、

重要度マップが量子化を調整することによって各映像フレーム内の符号化すべき各ターゲットブロックの符号化品質に影響を与えるように、前記重要度マップを用いて前記複数の映像フレームを符号化する過程、

を備え、前記重要度マップが：

時間的情報及び空間的情報を用いて当該重要度マップを設定すること；ならびに、

(i) 当該重要度マップが高い数値をとるブロックでは、ブロック量子化パラメータ(QP)がフレーム量子化パラメータ QP_{frame} に比べて小さくされることで、これらのブロックについては高い品質となるように、かつ、(ii) 当該重要度マップが低い数値をとるターゲットブロックでは、前記ブロック量子化パラメータが前記フレーム量子化パラメータ QP_{frame} に比べて大きくされることで、これらのブロックについては低い品質となるように、計算によって、前記複数の映像フレームのうちのある映像フレームのどの部分が人間の知覚にとって最も気付き易いのかを当該重要度マップに示させること；

によって構成されている、方法。

〔態様2〕

態様1に記載の方法において、前記空間的情報が、ルールに基づく空間的複雑度マップ(SCM)により提供されて、その最初のステップが、前記フレーム内のどのターゲットブロックが当該フレーム内の平均ブロック分散 var_{frame} よりも大きい分散を有するかを決定することであり、

前記平均ブロック分散 var_{frame} よりも大きい分散を有するブロックに対して、前記フレーム量子化パラメータ QP_{frame} よりも高い量子化パラメータ(QP)値を振り当て、このブロック量子化パラメータ(QP)の振当量 QP_{block} は、そのブロック分散 var_{block} が前記平均ブロック分散 var_{frame} よりもいかなる程度大きいかに従って、前記フレーム量子化パラメータ QP_{frame} と量子化パラメータ上

限 $Q P_{max}$ との間で線形的に増減される、方法。

〔態様 3〕

態様 1 に記載の方法において、前記時間的情報が、
どのターゲットブロックが観測者である人間にとって時間的に最も気付き易いかを示す
時間的コントラスト感度関数 (TCSE)、および、

どのターゲットブロックが前景データに相当するかを示す真の動きベクトルマップ (TMVM)

により提供されて、前記 TCSE は、前景データとして特定されたターゲットブロック
についてののみ有効とされる、方法。

〔態様 4〕

態様 2 に記載の方法において、分散の大きいブロックは、そのブロック量子化パラメータ (QP) である前記振当量 $Q P_{block}$ が、前記 TMVM がターゲットブロックを前景データとして特定し且つ前記 TCSE のこのブロックについてのコントラスト感度対数値が 0.5 未満である場合には前記振当量 $Q P_{block}$ が 2 増加するように、前記 TCSE 及び前記 TMVM によりさらに洗練化される、方法。

〔態様 5〕

態様 2 に記載の方法において、前記 SCM は、さらに、極めて明るい (170 超の輝度) か又は極めて暗い (60 未満の輝度) ターゲットブロックのブロック量子化パラメータである前記振当量 $Q P_{block}$ が $Q P_{max}$ に調節し直される輝度マスキングを含む、
方法。

〔態様 6〕

態様 2 に記載の方法において、前記 SCM は、さらに、前記符号化された映像の品質レベルに基づく前記量子化パラメータ上限 $Q P_{max}$ の動的な決定を含み、

この動的な決定では、イントラ (I) フレーム内のターゲットブロックの平均構造的類似度 (SSIM) 算出結果をこれらフレームの平均ブロック分散 var_{frame} と共に用いて、品質が測定され、

前記測定された品質が低いと、前記量子化パラメータ上限 $Q P_{max}$ の数値が前記フレーム量子化パラメータ $Q P_{frame}$ に近づくように減らされる、方法。

〔態様 7〕

態様 2 に記載の方法において、分散の極めて小さいブロックに対して、これらの領域における高品質符号化を確実にするために、前記ブロック分散が小さいほど前記振当量 $Q P_{block}$ の数値が低くなるように (、かつ、品質が高くなるように)、決められた低い量子化パラメータ (QP) の値である前記振当量 $Q P_{block}$ が振り当てられる、方法。

〔態様 8〕

態様 7 に記載の方法において、分散の極めて小さいブロックに対する前記低い量子化パラメータ (QP) の値である前記振当量 $Q P_{block}$ は、最初に、I フレームについて決められ、その後、P フレーム及び B フレームについては $ipratio$ パラメータ及び $pbratio$ パラメータを用いて決められる、方法。

〔態様 9〕

態様 7 に記載の方法において、分散は小さいが、分散が極めて小さいとは見なさないブロックは、当該ブロックについて品質向上が必要か否かを判定するために、

前記ブロック量子化パラメータ (QP) の初めの推定値である前記振当量 $Q P_{block}$ が現在のブロックの左、左上、右および右上の既に符号化済みの近傍ブロックの量子化パラメータ (QP) の値を平均することによって算出されて、且つ、

前記現在のブロックの前記 SSIM の推定 $SSIM_{est}$ が前記現在のブロックの左、左上、右および右上の既に符号化済みの近傍ブロックの SSIM 値から算出されて、且つ

$SSIM_{est}$ が 0.9 未満の場合、前記振当量 $Q P_{block}$ の数値が 2 減少されるように、

10

20

30

40

50

調べられる、方法。

〔態様１０〕

態様９に記載の方法において、前記品質向上は、前記ＴＭＶＭにより前景データとして特定されて且つ前記ＴＣＳＦのコントラスト感度対数値が０．８超であるブロックにのみ適用される、方法。

〔態様１１〕

態様３に記載の方法において、前記ＴＣＳＦの時間的周波数は、前記ターゲットブロックとその参照ブロックとの間の色空間領域におけるＳＳＩＭを用いて波長の近似を求めて且つ動きベクトルの大きさとフレームレートとを用いて速度の近似を求めることによって算出される、方法。

10

〔態様１２〕

態様３に記載の方法において、前記ＴＣＳＦは、現在のフレームについての当該ＴＣＳＦが最近のフレームにおけるＴＣＳＦマップの重み付き平均であるように且つより最近のフレームがより大きい重み付けを受けるように、複数のフレームにわたって算出される、方法。

〔態様１３〕

態様３に記載の方法において、前記ＴＭＶＭは、前景データの場合にのみ１に設定される、方法。

〔態様１４〕

態様１３に記載の方法において、前景データは、所与のターゲットブロックについてのエンコーダ動きベクトルと当該ブロックについてのグローバル動きベクトルとの差分を算出し、十分に大きい差分を有するブロックが前景データであると判断されることによって特定される、方法。

20

〔態様１５〕

態様１４に記載の方法において、前景データとして特定されたデータブロックについて、前記グローバル動きベクトルから前記エンコーダ動きベクトルが減算されることによって差分動きベクトルを得て、この差分動きベクトルの大きさが前記ＴＣＳＦの時間的周波数を算出するのに用いられる、方法。

〔態様１６〕

態様３に記載の方法において、前記ＴＣＳＦは、エンコーダからの動きベクトルから算出される、方法。

30

〔態様１７〕

態様１に記載の方法において、前記重要度マップが前記時間的情報及び前記空間的情報で設定されたものである場合、当該重要度マップは統合化された重要度マップである、方法。

〔態様１８〕

映像データを符号化するシステムであって、

重要度マップを用いて複数の映像フレームを符号化するコーデックであって、当該映像フレームは、互いに重なり合わないターゲットブロックを有している、コーデック、

を備え、前記重要度マップは、量子化を調整することによって各映像フレーム内の符号化すべき各ターゲットブロックの符号化品質に影響を与えるように構成されており、

40

前記重要度マップが：

時間的情報及び空間的情報を用いて当該重要度マップを設定することであって、これら時間的情報と空間的情報とにより設定された重要度マップは、統合化された重要度マップであること；ならびに、

（ｉ）当該重要度マップが高い数値をとるブロックでは、ブロック量子化パラメータ（ＱＰ）がフレーム量子化パラメータ $Q P_{frame}$ に比べて小さくされることで、これらのブロックについては高い品質となるように、かつ、（ｉｉ）当該重要度マップが低い数値をとるターゲットブロックでは、前記ブロック量子化パラメータが前記フレーム量子化パラメータ $Q P_{frame}$ に比べて大きくされることで、これらのブロックについては

50

低い品質となるように、計算によって、前記複数の映像フレームのうちのある映像フレームの、人間の知覚にとって最も気付き易い部分を当該重要度マップに示させること；
によって構成されている、システム。

〔態様 19〕

態様 18 に記載のエンコーダにおいて、前記空間的情報が、ルールに基づく空間的複雑度マップ (SCM) により提供されて、その最初のステップが、前記フレーム内のどのターゲットブロックが当該フレーム内の平均ブロック分散 var_{frame} よりも大きい分散を有するかを決定することであり、

前記平均ブロック分散 var_{frame} よりも大きい分散を有するブロックに対して、前記フレーム量子化パラメータ QP_{frame} よりも高い量子化パラメータ (QP) 値を振り当て、このブロック量子化パラメータ (QP) の振当量 QP_{block} は、そのブロック分散 var_{block} が前記平均ブロック分散 var_{frame} よりもいかなる程度大きいかに従って、前記フレーム量子化パラメータ QP_{frame} と量子化パラメータ上限 QP_{max} との間で線形的に増減される、エンコーダ。

〔態様 20〕

態様 18 に記載のエンコーダにおいて、前記時間的情報が、
どのターゲットブロックが観測者である人間にとって時間的に最も気付き易いかを示す時間的コントラスト感度関数 (TCSE)、および、

どのターゲットブロックが前景データに相当するかを示す真の動きベクトルマップ (TMVM)

により提供されて、前記 TCSE は、前景データとして特定されたターゲットブロックについてのみに有効とされる、エンコーダ。

〔態様 21〕

態様 19 に記載のエンコーダにおいて、分散の大きいブロックは、そのブロック量子化パラメータ (QP) である前記振当量 QP_{block} が、前記 TMVM がターゲットブロックを前景データとして特定し且つ前記 TCSE のこのブロックについてのコントラスト感度対数値が 0.5 未満である場合には前記振当量 QP_{block} が 2 増加するように、前記 TCSE 及び前記 TMVM によりさらに洗練化される、エンコーダ。

〔態様 22〕

態様 19 に記載のエンコーダにおいて、前記 SCM は、さらに、極めて明るい (170 超の輝度) か又は極めて暗い (60 未満の輝度) ターゲットブロックのブロック量子化パラメータである前記振当量 QP_{block} が QP_{max} に調節し直される輝度マスキングを含む、エンコーダ。

〔態様 23〕

態様 19 に記載のエンコーダにおいて、前記 SCM は、さらに、符号化された映像の品質レベルに前記量子化パラメータ上限に基づく QP_{max} の動的な決定を含み、

この動的な決定では、イントラ (I) フレーム内のターゲットブロックの平均構造的類似度 (SSIM) 算出結果をこれらフレームの平均ブロック分散 var_{frame} と共に用いて、品質が測定され、

測定された品質が低いと、前記量子化パラメータ上限 QP_{max} の数値が前記フレーム量子化パラメータ QP_{frame} 近づくように減らされる、エンコーダ。

〔態様 24〕

態様 19 に記載のエンコーダにおいて、分散の極めて小さいブロックに対して、これらの領域における高品質符号化を確実にするために、前記ブロック分散が小さいほど前記振当量 QP_{block} の数値が低くなるように (、かつ、品質が高くなるように)、決められた低い量子化パラメータ (QP) の値である前記振当量 QP_{block} が振り当てられる、エンコーダ。

〔態様 25〕

態様 24 に記載のエンコーダにおいて、分散の極めて小さいブロックに対する前記低い量子化パラメータ (QP) の値である前記振当量 QP_{block} は、最初に、I フレーム

10

20

30

40

50

については決められ、その後、Pフレーム及びBフレームについてはi p r a t i oパラメータ及びp b r a t i oパラメータを用いて決められる、エンコード。

〔態様26〕

態様19に記載のシステムにおいて、分散は小さいが、分散が極めて小さいとは見えないブロックは、当該ブロックについて品質向上が必要か否かを判定するために、

前記ブロック量子化パラメータ(QP)の初めの推定値である前記振当量 QP_{block} が現在のブロックの左、左上、右および右上の既に符号化済みの近傍ブロックの量子化パラメータ(QP)の値を平均することによって算出されて、且つ、

前記現在のブロックの前記SSIMの推定SSIM $_{est}$ が前記現在のブロックの左、左上、右および右上の既に符号化済みの近傍ブロックのSSIM値から算出されて、且つ

、SSIM $_{est}$ が0.9未満の場合、前記振当量 QP_{block} の数値が2減少されるように、

調べられる、システム。

〔態様27〕

態様26に記載のシステムにおいて、前記品質向上は、前記TMVMにより前景データとして特定されて且つ前記TCSFのコントラスト感度対数値が0.8超であるブロックにのみ適用される、システム。

〔態様28〕

態様20に記載のシステムにおいて、前記TCSFの時間的周波数は、前記ターゲットブロックとその参照ブロックとの間の色空間領域におけるSSIMを用いて波長の近似を求めて且つ動きベクトルの大きさとフレームレートとを用いて速度の近似を求めることによって算出される、システム。

〔態様29〕

態様20に記載のシステムにおいて、前記TCSFは、現在のフレームについての当該TCSFが最近のフレームにおけるTCSFマップの重み付き平均であるように且つより最近のフレームがより大きい重み付けを受けるように、複数のフレームにわたって算出される、システム。

〔態様30〕

態様20に記載のシステムにおいて、前記TMVMは、前景データの場合にのみ1に設定される、システム。

〔態様31〕

態様30に記載のシステムにおいて、前景データは、所与のターゲットブロックについてのエンコード動きベクトルと当該ブロックについてのグローバル動きベクトルとの差分を算出し、十分に大きい差分を有するブロックが前景データであると判断されることによって特定される、システム。

〔態様32〕

態様20に記載のシステムにおいて、前景データとして特定されたデータブロックについて、前記グローバル動きベクトルから前記エンコード動きベクトルが減算されることによって差分動きベクトルを得て、この差分動きベクトルの大きさが前記TCSFの時間的周波数を算出するのに用いられる、システム。

〔態様33〕

態様20に記載のシステムにおいて、前記TCSFは、前記エンコードからの動きベクトルから算出される、システム。

〔態様34〕

態様18に記載のシステムにおいて、前記重要度マップが前記時間的情報と前記空間的情報で設定されたものである場合、当該重要度マップは統合化された重要度マップである、システム。

10

20

30

40

【図 5】

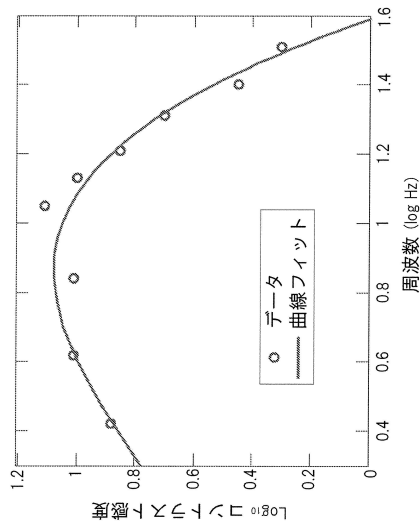


Figure 5: 時間的コントラスト感度関数、傍中心高の対象 [Wooten, 2010]

【図 6】

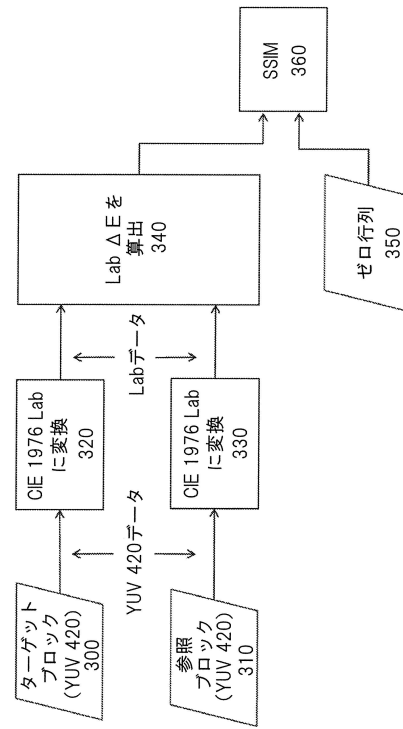


Figure 6: CIE 1976 Lab色空間におけるSSIMの算出

【図 7】

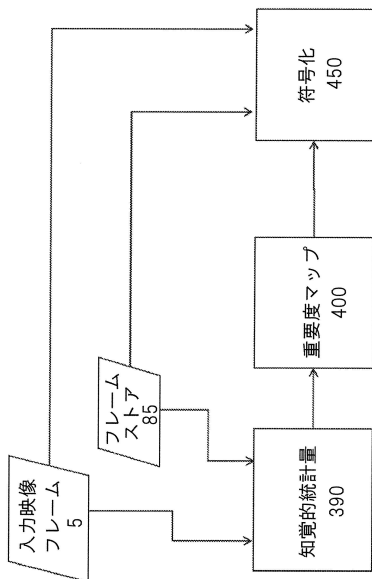


Figure 7: 映像符号化への重要度マップの一般的な適用

【図 8 A】

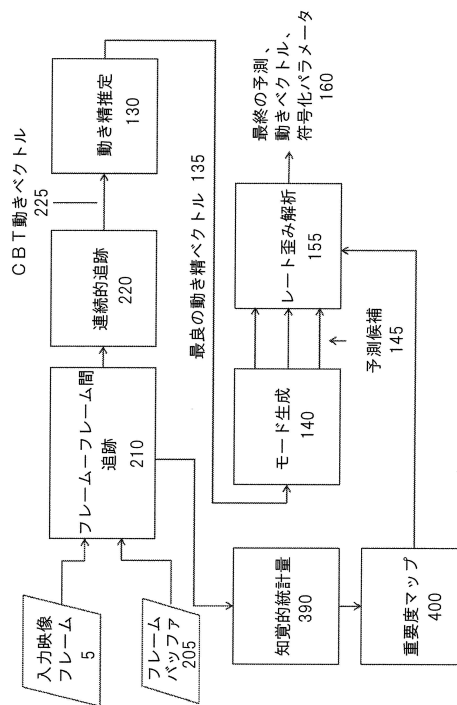


Figure 8A: 連続的ブロック追跡によるインター予測への知覚的統計量の適用

【図 8 B】

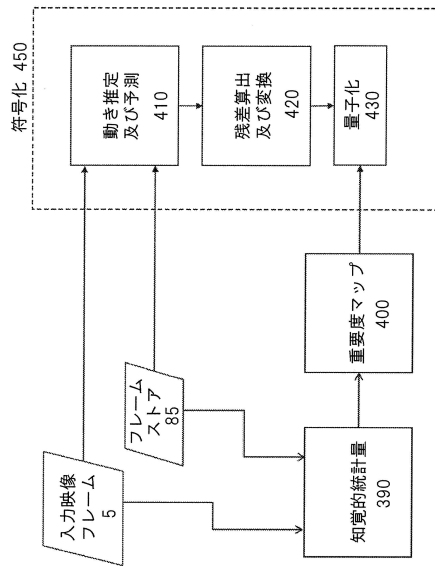


Figure 8B: 映像符号化における量子化を改変するための重要度マップの適用

【図 9 A】

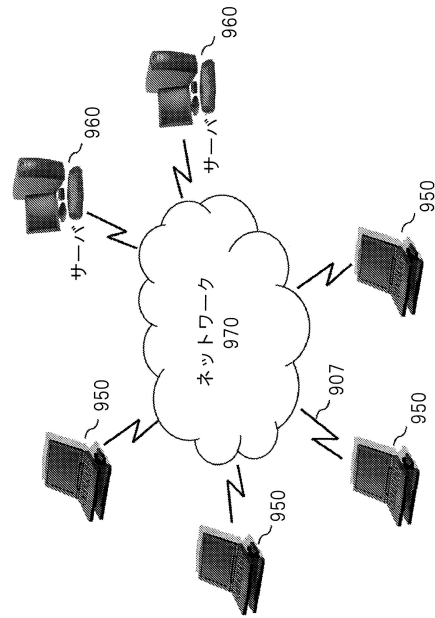


Figure 9A

【図 9 B】

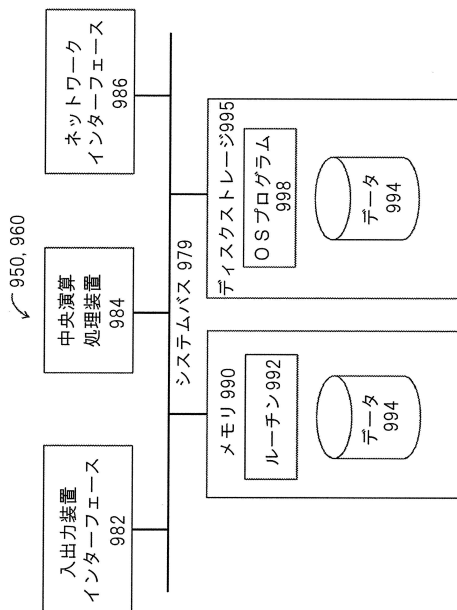


Figure 9B

フロントページの続き

(31)優先権主張番号 14/532,947

(32)優先日 平成26年11月4日(2014.11.4)

(33)優先権主張国・地域又は機関
米国(US)

(31)優先権主張番号 62/158,523

(32)優先日 平成27年5月7日(2015.5.7)

(33)優先権主張国・地域又は機関
米国(US)

(74)代理人 100142608

弁理士 小林 由佳

(74)代理人 100154771

弁理士 中田 健一

(74)代理人 100155963

弁理士 金子 大輔

(74)代理人 100150566

弁理士 谷口 洋樹

(72)発明者 リー・ニゲル

アメリカ合衆国, マサチューセッツ州 02467, チェスナット ヒル, ハーモンド ストリート 727

(72)発明者 パーク・サンソク

アメリカ合衆国, テキサス州 75022, フラワー マウンド, ドレセージ レーン 3620

(72)発明者 トウン・ミョー

アメリカ合衆国, テキサス州 75070, マッキニー, シデーリア ドライブ 10812

(72)発明者 コッケ・デーン・ピー

アメリカ合衆国, ノースカロライナ州 27705, ダラム, マリーゴールド プレイス 18

(72)発明者 リー・ジェユン

アメリカ合衆国, テキサス州 78717, オースティン, オールド クワリー ロード 11125

(72)発明者 ウィード・クリストファー

アメリカ合衆国, マサチューセッツ州 01776, サドベリー, プリムストーン レーン 33

審査官 岩井 健二

(56)参考文献 特表2011-510591(JP, A)

特表2008-507915(JP, A)

特開2006-067302(JP, A)

特開2006-054801(JP, A)

特開2004-248285(JP, A)

特開2002-051345(JP, A)

特開平08-032975(JP, A)

特開平05-103315(JP, A)

国際公開第2014/029315(WO, A1)

Yayu Zheng et al., H.264 ROI Coding Based on Visual Percetion, 2008 5th International Conference on Visual Infformation Engineering (VIE 2008), IET, 2008年 7月, pp.829-834

(58)調査した分野(Int.Cl., DB名)

H 0 4 N 1 9 / 0 0 - 1 9 / 9 8