

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5078837号
(P5078837)

(45) 発行日 平成24年11月21日(2012.11.21)

(24) 登録日 平成24年9月7日(2012.9.7)

(51) Int.Cl. F I
HO4N 7/30 (2006.01) HO4N 7/133 Z

請求項の数 9 (全 16 頁)

(21) 出願番号	特願2008-266831 (P2008-266831)	(73) 特許権者	000001007
(22) 出願日	平成20年10月15日(2008.10.15)		キヤノン株式会社
(65) 公開番号	特開2009-135902 (P2009-135902A)		東京都大田区下丸子3丁目30番2号
(43) 公開日	平成21年6月18日(2009.6.18)	(74) 代理人	100076428
審査請求日	平成23年2月17日(2011.2.17)		弁理士 大塚 康德
(31) 優先権主張番号	特願2007-280940 (P2007-280940)	(74) 代理人	100112508
(32) 優先日	平成19年10月29日(2007.10.29)		弁理士 高柳 司郎
(33) 優先権主張国	日本国(JP)	(74) 代理人	100115071
			弁理士 大塚 康弘
		(74) 代理人	100116894
			弁理士 木村 秀二
		(74) 代理人	100130409
			弁理士 下山 治
		(74) 代理人	100134175
			弁理士 永川 行光

最終頁に続く

(54) 【発明の名称】 符号化装置、符号化装置の制御方法及びコンピュータプログラム

(57) 【特許請求の範囲】

【請求項1】

入力画像を複数のブロックに分割して符号化する符号化装置であって、
 前記複数のブロックのそれぞれについて画像の特徴を判定し、視覚的劣化が目立ちやすいブロックであるか否かを検出する検出手段と、
 前記検出手段による検出結果に基づいて、前記ブロックごとの量子化パラメータを決定する決定手段と、
 前記ブロックの直交変換と、決定された前記量子化パラメータを用いた量子化とを行う変換手段と、
 前記変換手段による変換結果を、可変長符号化する符号化手段と、
 前記変換手段による変換結果を逆変換して局所復号画像を生成する逆変換手段と、
 前記画像の特徴を判定するための判定基準を変化させるための参照値を算出する算出手段と、
 を備え、
 前記算出手段は、前記入力画像と前記局所復号画像とに基づいてPSNRを演算し、演算された前記PSNRの値に基づく画像の劣化度合いを示す前記参照値を出力し、
 前記検出手段は、
 画像の特徴毎に設定された閾値を前記判定基準として用いて判定を行って前記視覚的劣化が目立ちやすいブロックを検出し、
 前記閾値の値は画像の特徴毎に前記参照値がとりえる各値と対応づけて設定され、前

10

20

計算出された参照値に応じた閾値が前記判定基準として用いられることを特徴とする符号化装置。

【請求項 2】

前記視覚的劣化が目立ちやすいブロックが、前記入力画像において平坦部を構成するブロックである場合に、

前記検出手段は、前記ブロック内の画素の分散値を算出し、該分散値の大きさに基づいて、該ブロックが平坦部を構成するブロックであるかを判定して、前記視覚的劣化が目立ちやすいブロックの検出を行うことを特徴とする請求項 1 に記載の符号化装置。

【請求項 3】

前記検出手段は、前記判定基準として第 1 の閾値を用いて、該第 1 の閾値と前記分散値との比較により前記ブロックが平坦部を構成するブロックであるかを判定することを特徴とする請求項 2 に記載の符号化装置。

10

【請求項 4】

前記視覚的劣化が目立ちやすいブロックが、前記入力画像においてエッジを構成するブロックである場合に、

前記検出手段は、前記ブロックをサブブロックに更に分割して、該サブブロック間での画素の分散値の差分値を算出し、該差分値の大きさに基づいて、該ブロックがエッジを構成するブロックであるかを判定して、前記視覚的劣化が目立ちやすいブロックの検出を行うことを特徴とする請求項 1 に記載の符号化装置。

【請求項 5】

20

前記検出手段は、前記判定基準として第 2 の閾値を用いて、該第 2 の閾値と前記差分値との比較により前記ブロックがエッジを構成するブロックであるかを判定し、

前記参照値が示す度合いが予め定めた度合いよりも高くなる場合には、前記第 2 の閾値の値を増加させることを特徴とする請求項 4 に記載の符号化装置。

【請求項 6】

前記視覚的劣化が目立ちやすいブロックが、前記入力画像において肌色部分を構成するブロックである場合に、

前記検出手段は、前記ブロック内の肌色画素の個数を算出し、該個数の大きさに基づいて、該ブロックが肌色部分を構成するブロックであるかを判定して、前記視覚的劣化が目立ちやすいブロックの検出を行うことを特徴とする請求項 1 に記載の符号化装置。

30

【請求項 7】

前記検出手段は、前記判定基準として第 3 の閾値を用いて、該第 3 の閾値と前記個数との比較により前記ブロックが肌色部分を構成するブロックであるかを判定し、

前記参照値が前記画像の劣化度合いを示す場合に、該劣化度合いが予め定めた度合いよりも高くなる場合には前記第 3 の閾値の値を減少させることを特徴とする請求項 6 に記載の符号化装置。

【請求項 8】

入力画像を複数のブロックに分割して符号化する符号化装置の制御方法であって、

検出手段が、前記複数のブロックのそれぞれについて画像の特徴を判定し、視覚的劣化が目立ちやすいブロックであるか否かを検出する検出工程と、

40

決定手段が、前記検出工程における検出結果に基づいて、前記ブロックごとの量子化パラメータを決定する決定工程と、

変換手段が、前記ブロックの直交変換と、決定された前記量子化パラメータを用いた量子化とを行う変換工程と、

符号化手段が、前記変換工程における変換結果を、可変長符号化する符号化工程と、

逆変換手段が、前記変換工程における変換結果を逆変換して局所復号画像を生成する逆変換工程と、

算出手段が、前記画像の特徴を判定するための判定基準を変化させるための参照値を算出する算出工程と、

を備え、

50

前記算出工程では、前記入力画像と前記局所復号画像とに基づいてP S N Rを演算し、演算された前記P S N Rの値に基づく画像の劣化度合いを示す前記参照値を出力し、

前記検出工程では、

画像の特徴毎に設定された閾値を前記判定基準として用いて判定を行って前記視覚的劣化が目立ちやすいブロックを検出し、

前記閾値の値は画像の特徴毎に前記参照値がとりえる各値と対応づけて設定され、前記算出された参照値に応じた閾値が前記判定基準として用いられる

ことを特徴とする符号化装置の制御方法。

【請求項9】

コンピュータを、請求項1乃至7のいずれか1項に記載の符号化装置の各手段として機能させるためのコンピュータプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、符号化装置、符号化装置の制御方法及びコンピュータプログラムに関する。

【背景技術】

【0002】

近年のマルチメディアの発展に伴い様々な動画像圧縮符号化方式が提案されている。その代表的なものに、MPEG-1, 2, 4やH.264といったものがある。これらの圧縮符号化の処理は、動画像に含まれる原画像(画像)をブロックと呼ばれる所定の領域に分割し、この分割したブロックを単位にして動き補償予測やDCT変換処理を施すものである。また、動き補償予測を行う場合、既に符号化済みの画像データを局所復号化して得られた画像を参照画像としていることで、符号化を行う際にも復号化処理が必要となる。

【0003】

また、MPEG方式に準拠して画像の圧縮、符号化を行う場合、その符号量は、画像自体の特性である空間周波数特性やシーン及び量子化スケール値に応じて大きく異なる場合が多い。このような符号化特性を有する符号化装置を実現する上で良好な画質の復号画像を得ることができるようにするための重要な技術が符号量制御である。

【0004】

この符号量制御アルゴリズムの1つとして、TM5(Test Model 5)が一般的に使用されている。このTM5による符号量制御アルゴリズムは、以下に説明する3つのステップから構成される。TM5では、GOP(Group Of Picture)毎にビットレートが一定になるように、以下に挙げる3ステップで符号量が制御される。

【0005】

(STEP1)

今から符号化を行うピクチャの目標符号量を決定する。現在のGOPにおいて利用可能な符号量であるRgopが以下の(1)式により演算される。

$$R_{gop} = (n_i + n_p + n_b) * (bits_rate / picture_rate) \quad \dots (1)$$

ここで、 n_i , n_p , n_b はそれぞれI、P、Bピクチャの現GOPにおける残りのピクチャ数であり、 $bits_rate$ は目標ビットレート、 $picture_rate$ はピクチャレートを表す。

【0006】

更に、I、P、Bピクチャ毎に符号化結果からピクチャの複雑度を以下の(2)式で求めている。

$$\begin{aligned} X_i &= R_i * Q_i \\ X_p &= R_p * Q_p \quad \dots (2) \\ X_b &= R_b * Q_b \end{aligned}$$

ここで、 X_i 、 X_p 、 X_b はコンプレキシティ(Complexity)とも呼ばれる。また、 R_i 、 R_p 及び R_b はそれぞれI、P、Bピクチャを符号化した結果得られる符号量である。さらに、 Q_i 、 Q_p 及び Q_b はそれぞれI、P、Bピクチャ内のすべてのマクロブロックにおける

10

20

30

40

50

Qスケールの平均値である。式(1)及び式(2)から、I、P、Bピクチャそれぞれについての目標符号量 T_i 、 T_p 及び T_b は、以下の(3)式で求めることができる。

【0007】

$$T_i = \max\left\{\left(\frac{R_{gop}}{1 + \left(\frac{N_p \cdot X_p}{X_i \cdot K_p}\right) + \left(\frac{N_b \cdot X_b}{X_i \cdot K_b}\right)}\right), \left(\frac{\text{bit_rate}}{8 \cdot \text{picture_rate}}\right)\right\}$$

$$T_p = \max\left\{\left(\frac{R_{gop}}{N_p + \left(\frac{N_b \cdot K_p \cdot X_b}{K_b \cdot X_p}\right)}\right), \left(\frac{\text{bit_rate}}{8 \cdot \text{picture_rate}}\right)\right\}$$

$$T_b = \max\left\{\left(\frac{R_{gop}}{N_b + \left(\frac{N_p \cdot K_b \cdot X_p}{K_p \cdot X_b}\right)}\right), \left(\frac{\text{bit_rate}}{8 \cdot \text{picture_rate}}\right)\right\}$$

・・・(3)

ただし、 N_p 及び N_b は現GOP内のそれぞれP及びBピクチャの残りの枚数、また定数 $K_p=1.0$ 及び $K_b=1.4$ である。

【0008】

(STEP2)

I、P及びBピクチャ毎に3つの仮想バッファを使用し、式(3)で求めた目標符号量と発生符号量との差分を管理する。仮想バッファのデータ蓄積量をフィードバックし、そのデータ蓄積量に基づいて実際の発生符号量が目標符号量に近づくように、次にエンコードするマクロブロックについて、Qスケールの参照値が設定される。例えば、現在のピクチャタイプがPピクチャの場合には、目標符号量と発生符号量との差分は、次の(4)式に従う演算処理により求めることができる。

【0009】

$$d_{p,j} = d_{p,0} + B_{p,j-1} - \left(\frac{T_p \cdot (j-1)}{MB_cnt}\right) \cdot \dots \quad (4)$$

ここで、添字 j はピクチャ内のマクロブロックの番号であり、 $d_{p,0}$ は仮想バッファの初期フルネスを示し、 $B_{p,j}$ は j 番目のマクロブロックまでの総符号量、 MB_cnt はピクチャ内のマクロブロック数を示す。次に $d_{p,j}$ (以後、「 d_j 」と記載する。) を用いて、 j 番目のマクロブロックにおけるQスケールの参照値を求めると、(5)式のようになる。

【0010】

$$Q_j = (d_j \cdot 31) / r \quad \dots \quad (5)$$

$$\text{ここで、} r = 2 \cdot \text{bits_rate} / \text{picture_rate} \quad \dots \quad (6)$$

である。

【0011】

(STEP3)

視覚特性、即ち、復号画像の画質が良好になるように、エンコード対象のマクロブロックの空間アクティビティに基づいて、量子化スケールを最終的に決定する処理を実行する。

【0012】

$$ACT_j = 1 + \min(vblk1, vblk2, \dots, vblk8) \quad \dots \quad (7)$$

(7)式中において、 $vblk1 \sim vblk4$ はフレーム構造のマクロブロックにおける 8×8 のサブブロックにおける空間アクティビティを示す。また、 $vblk5 \sim vblk8$ はフィールド構造のマクロブロックにおける 8×8 サブブロックの空間アクティビティを示す。ここで、空間アクティビティの演算は次の(8)、(9)式により求めることが可能である。

【0013】

$$vblk = (P_i - P_{bar})^2 \quad \dots \quad (8)$$

$$P_{bar} = (1/64) \cdot \sum P_i \quad \dots \quad (9)$$

ここで、 P_i は i 番目のマクロブロックにおける画素値であり、式(8)、(9)中のは $i=1 \sim 64$ の演算である。次に(7)式で求めた ACT_j を以下の(10)式によって正規化を行う。

【0014】

$$N_ACT_j = (2 \cdot ACT_j + AVG_ACT) / (ACT_j + AVG_ACT) \quad \dots \quad (10)$$

ここで、 AVG_ACT は以前に符号化したピクチャにおける ACT_j の参照値であり、最終的に量子化スケール(Qスケール値) $MQUANT_j$ は以下の(11)式により求められる。

【0015】

10

20

30

40

50

$$MQANT_j = Q_j * N_ACT_j \quad \dots (11)$$

以上のT M 5のアルゴリズムによれば、S T E P 1の処理によりIピクチャに対して多くの符号量を割り当てており、更にピクチャ内においては視覚的に劣化が目立ちやすい平坦部(空間アクティビティが低い)に符号量が多く配分されるようになる。よって、予め定めたビットレート内で、画質の劣化を抑えた符号量制御ならびに量子化制御を行うことが可能となる。

【0016】

また、T M 5と同様に画像の特徴に応じて量子化制御を行う手法は他にも提案されており、視覚的向上を実現できている(特許文献1参照)。

【特許文献1】特開平11-196417号

10

【発明の開示】

【発明が解決しようとする課題】

【0017】

前述したT M 5方式は、マクロブロック単位で特徴を抽出し、その特徴に基づいて量子化パラメータを変化させる適応量子化を行うことで、あらかじめ定められた目標符号量になるような量子化制御を行っている。

【0018】

また特許文献1では、複雑度が大きく量子化パラメータを上げるブロックの数が少ない場合、発生符号量が増大するにもかかわらず複雑度が高いブロックの劣化が目立つので、適応量子化を行わないという制御を行っている。複雑度が小さく量子化パラメータを下げるブロックが多い場合も同様である。

20

【0019】

しかしながら、上記アルゴリズムではフレーム間での適応量子化の有り無しがあるため画質がばたつく問題があった。

【0020】

また、量子化パラメータを下げるブロックが多い画像でも、ブロックの持つ特徴によっては平坦部のように量子化パラメータを下げても符号量が増大しないブロックも存在する。従って、本来適応量子化を行うべき画像に対して適応量子化が行えないという問題も存在する。

【0021】

30

本発明は、上述した問題点に鑑みたものであり、符号化画像の劣化度合いとブロックが持つ特徴を考慮した適応量子化を行うことで、画質劣化を抑えることを可能にした符号化技術を提供することを目的とする。

【課題を解決するための手段】

【0022】

上記課題を解決するための本発明は、入力画像を複数のブロックに分割して符号化する符号化装置であって、

前記複数のブロックのそれぞれについて画像の特徴を判定し、視覚的劣化が目立ちやすいブロックであるか否かを検出する検出手段と、

前記検出手段による検出結果に基づいて、前記ブロックごとの量子化パラメータを決定する決定手段と、

40

前記ブロックの直交変換と、決定された前記量子化パラメータを用いた量子化とを行う変換手段と、

前記変換手段による変換結果を、可変長符号化する符号化手段と、

前記変換手段による変換結果を逆変換して局所復号画像を生成する逆変換手段と、

前記画像の特徴を判定するための判定基準を変化させるための参照値を算出する算出手段と、

を備え、

前記算出手段は、前記入力画像と前記局所復号画像とに基づいてP S N Rを演算し、演算された前記P S N Rの値に基づく画像の劣化度合いを示す前記参照値を出力し、

50

前記検出手段は、

画像の特徴毎に設定された閾値を前記判定基準として用いて判定を行って前記視覚的劣化が目立ちやすいブロックを検出し、

前記閾値の値は画像の特徴毎に前記参照値がとりえる各値と対応づけて設定され、前記算出された参照値に応じた閾値が前記判定基準として用いられることを特徴とする。

【発明の効果】

【0023】

本発明によれば、符号化画像の劣化度合いとブロックが持つ特徴を考慮した適応量子化を行うことで、画質劣化を抑えることを可能にした符号化技術を提供できる。

【発明を実施するための最良の形態】

【0024】

以下、添付する図面を参照して発明の実施形態を説明する。

【0025】

第1の実施形態を図1から図3に基づいて説明する。図1は、本実施形態に対応する符号化装置の全体構成の一例を示すブロック図である。図2、図3は、本実施形態に対応する特徴抽出を説明するための図である。図4は、PSNRについて説明するための図である。図5A乃至図5Cは、PSNRと検出閾値の関係を表した図である。

【0026】

図1は、発明の実施形態に対応する符号化装置を示すが、該装置は、例えばデジタルビデオカメラのような映像音声信号記録装置として実現できる。また、符号化方式としては、例えばMPEG (Moving Pictures of Experts Group) やH.264/AVC (Advanced Video Coding) に対応する。

【0027】

図1において、入力信号100は本符号化装置への入力信号であり、動画像が所定のブロックに分割された状態で入力される。該ブロックは、例えばMPEGでは 16×16 、 8×8 を始めとしたブロックであり、符号化方式に応じてサイズが決まる。なお今後説明の中では当該ブロックのことを「マクロブロック」と呼ぶこととする。

【0028】

符号化装置は、フレーム並べ替え部101、加減算部102、直交変換(DCT)部103、量子化部104、逆量子化部105、逆直交変換(逆DCT)部106、動き予測・動き補償部107を含む。また、加減算部108、フレームメモリであるビデオバッファ109、可変長符号化部110、符号量制御部111、量子化制御部112、特徴抽出部113、本符号化装置からのストリームを一時的に保持するバッファ114を含む。更に、入力画像と局所復号画像からPSNR (Peak Signal to Noise Ratio) を算出するPSNR算出部115、本符号化装置からの出力信号116を含む。PSNRとは、符号化による画像の劣化度合いを示す指標であり、dB (デシベル) と呼ばれる単位で表現される。その算出方法は、入力画像と局所復号画像との差分の二乗和を用いて行われ、算出結果であるPSNR値が大きいときは符号化劣化が少なく、PSNR値が小さいときは符号化劣化が多いことを意味する。なお、本実施形態ではマクロブロック単位にPSNRを算出しており、画像内に含まれる全マクロブロックの合計値を観測PSNRとして用いている。

【0029】

以下、本実施形態に対応する符号化装置の動作を説明する。入力信号100は、フレーム並べ替え部101において符号化ピクチャタイプに応じた並べ替えが行われる。並べ替えが終わった画像はブロックに分割された後、符号化ピクチャがフレーム内符号化(イントラ符号化)方式の時は、直交変換部103においてマクロブロック内の信号に対して直交変換処理を行う。量子化部104では、直交変換結果として得られたDCT係数を量子化する。

【0030】

また、符号化ピクチャがフレーム間符号化ピクチャ(インター符号化)方式の時は、す

10

20

30

40

50

で符号化されたピクチャに対して、逆量子化部 105 で逆量子化、逆直交変換部 106 で逆直交変換処理を行う。このような逆変換を施すことでローカルデコード（局所復号画像）画像を生成する。更に、符号化しようとするピクチャとの動き予測ならびに動き補償を動き予測、動き補償部 107 で行い、ローカルデコード画像との差分値を加減算部 102 で算出する。差分値には、直交変換部 103 において直交変換処理が施され、量子化部 104 で DCT 係数が量子化される。

【0031】

フレーム内符号化、フレーム間符号化に関わらず量子化部 104 で量子化された量子化信号は、可変長符号化部 110 で符号化され、符号化された信号が出力信号 116 として出力される。

10

【0032】

次に、符号量制御部 111、量子化制御部 112 について説明する。符号量制御部 111 は、GOP 内の各ピクチャに対する割り当てビット量を、割り当て対象ピクチャを含め GOP 内でまだ符号化されていないピクチャに対するビット量を基に配分する。この配分を GOP 内の符号化ピクチャ順に繰り返し、ピクチャごとにピクチャ目標符号量を設定する。符号量制御部 111 はまた、特徴抽出部 113 に対して、特徴毎の検出閾値を提供する。

【0033】

次に、量子化制御部 112 は、量子化パラメータを決定する。その際、各ピクチャに対する目標符号量と実際の発生符号量とを一致させるため、仮想バッファの容量を基に量子化スケールの参照値を、可変長符号化部 110 から出力されるマクロブロック単位の発生符号量としてフィードバック制御で求める。量子化部 104 で使用する量子化パラメータは、量子化スケールの参照値に対して特徴抽出部 113 で算出するアクティビティを基に式 (11) を用いて決定される。なお以上の動作は背景技術で述べたステップ 1 ~ 3 に相当する。

20

【0034】

これ以降、特徴抽出部 113 における処理について、図 2 及び図 3 を参照して説明する。特徴抽出部 113 は、符号量制御部 111 から提供された検出閾値（画像の特徴を判定するための判定基準に相当）に基づいて画像の特徴を判定し、各マクロブロックが視覚的劣化が目立ちやすいブロックであるか否かを検出する。本実施形態では、視覚的劣化が目立ちやすいブロックとして、平坦部、エッジ部、肌色部を例として説明する。なお、以下では、平坦部、エッジ部、肌色部の全てについて説明するが、いずれか 1 つ、或いは、複数の組み合わせにおいて発明を実施してもよい。

30

【0035】

図 2 に示すように、特徴抽出部 113 は、平坦検出部 201、エッジ検出部 203 及び肌色検出部 205 を備える。各検出部には平坦検出用閾値 202 ($th1F$ 、 $th2F$)、エッジ検出用閾値 204 ($th1E$ 、 $th2E$) 及び肌色検出用閾値 206 ($th1S$ 、 $th2S$) がそれぞれ符号量制御部 111 から入力される。各閾値は、マクロブロックが平坦部、エッジ部或いは肌色部に属するかを判断するための判定基準となる。また、最小値算出部 207 は、各検出部で算出されたレベルに応じて正規化前アクティビティを算出する。正規化部 208 は、正規化前アクティビティに対して所定の制御感度（リアクションパラメータ）を用いて正規化アクティビティを算出する。各閾値は変動可能である。

40

【0036】

また、図 3 は各検出部の強さをレベル化するための関数を説明するための図であり、図 3 (a) は平坦検出部用の関数 301、図 3 (b) はエッジ検出部用の関数 302、図 3 (c) は肌色検出部用の関数 303 を示している。

【0037】

ブロック単位に特徴抽出部 113 に入力される画像信号は各検出部に入力され閾値に応じてレベル化される。各検出部の処理は以下の通りである。

【0038】

50

まず、平坦検出部 201 は、ブロック信号に対して分散値を算出し、その分散値を閾値と比較し、 $th1F$ 以下であればレベルを 1 とし、 $th2F$ 以上であればレベルを無限大とする。ここで、分散値は、ブロック内における画素値のバラツキ度合いを示す値である。また、レベルが低い値を取るほど、ブロック内の画素値の分散が少なく、該ブロックが平坦画像の傾向が高いことを示す。分散値が閾値 $th1F$ から $th2F$ の間の値であれば ($th1F, 1$) と ($th2F, pre_avg$) を結ぶ関数からレベルを算出する。なお「 pre_avg 」とは 1 フレーム前に符号化したピクチャにおけるアクティビティの平均値である。また、第 1 の閾値である $th1F$ 及び $th2F$ はフレーム単位に CPU である符号量制御部 111 からセットされた平坦検出用閾値 202 (第 1 の閾値) である。

【0039】

10

エッジ検出部 203 はブロック信号をさらに 8×8 のサブブロックに分割して、サブブロック毎に分散値を算出する。次に、分散値の最大値と最小値との差分を算出し、その値を閾値と比較し $th2E$ 以上であればレベルを 1 とし、 $th1E$ 以下であればレベルを無限大とする。ここで、レベルが低い値を取るほど、サブブロック間での分散値の差分が大きく、該ブロックがエッジ画像の傾向が高いことを示す。差分が閾値 $th1E$ から $th2E$ の値をとれば ($th1E, pre_avg$) と ($th2, 1$) を結ぶ関数からレベルを算出する。なお「 pre_avg 」とは、1 フレーム前に符号化したピクチャにおけるアクティビティの平均値である。また、閾値 $th1E$ 及び $th2E$ は、フレーム単位に CPU である符号量制御部 111 からセットされたエッジ検出用閾値 204 (第 2 の閾値) である。

【0040】

20

肌色検出部 205 はブロック信号に対して輝度成分と色差成分から肌色画素の個数をカウントする。その個数を閾値と比較し、 $th2S$ 以上であればレベルを 1 とし、 $th1S$ 以下であれば無限大とする。ここで、レベルが低い値を取るほど、肌色の画素数が多く、該ブロックが肌色画像の傾向が高いことを示す。個数が閾値 $th1S$ から $th2S$ の間の値であれば ($th1S, pre_avg$) と ($th2S, 1$) を結ぶ関数からレベルを算出する。なお「 pre_avg 」とは 1 フレーム前に符号化したピクチャにおけるアクティビティの平均値である。また、閾値 $th1S$ 及び $th2S$ は、フレーム単位に CPU である符号量制御部 111 からセットされた肌色検出用閾値 206 (第 3 の閾値) である。

【0041】

最後に最小値算出部 207 で、各検出部 201、203 及び 205 で算出されたレベルの中で、最小のものを正規化前アクティビティとし、正規化部 208 で正規化してアクティビティとして算出する。また、各検出部からのレベルがすべて無限大の場合はブロックの分散値を正規化前アクティビティとする。

30

【0042】

本実施形態では、判定基準としての検出用閾値を PSNR 算出部 115 で算出された PSNR に基づいて、以下のように変化させることができる。この処理について図 4 及び図 5A 乃至図 5C を参照して説明する。

【0043】

図 4 は、横軸に時間、縦軸に PSNR を示した図である。点線が予め定められた目標となる PSNR (目標 PSNR) であり、実線が PSNR 算出部 115 で算出した PSNR (観測 PSNR) である。観測 PSNR が目標 PSNR より高ければ画質が良いと判断でき、低ければ画質が悪くないと判断できる。

40

【0044】

PSNR はまた、符号量的な側面から次のような解釈も可能である。観測 PSNR が目標 PSNR より高い場合は、符号化する画像に対して十分な符号量が割り当てられていると同時に符号量的に安定しているので、特徴抽出でより多くのエッジや肌色ブロックを検出してさらに符号量を与えることが可能である。

【0045】

一方、観測 PSNR が目標 PSNR よりも低い場合は、符号化する画像に対して十分な符号量を割り当てられていないと同時に、符号量的に不安定な状態であると言える。その

50

ため、特徴抽出でエッジや肌色のブロックを多く検出し過ぎると、エッジや肌色ブロックに多くの符号量が割り当てられ、高周波を含むブロックへ符号量が割り当てられなくなってしまう。高周波を含むブロックは多少荒く量子化を行っても視覚的劣化は目立ちにくいとされているが、割り当てる符号量が少なすぎると高周波を含むブロックにおいて画質が大幅に低下してしまう。また、バッファ破綻が発生する可能性もある。

【 0 0 4 6 】

そこで、本実施形態では、目標 P S N R と観測 P S N R との差分に応じて各検出部に与えられる閾値を制御することを特徴とする。図 5 は、横軸に目標 P S N R から観測 P S N R を引いた差分 (P S N R の差分)、縦軸に各検出部の閾値を取ったグラフである。

【 0 0 4 7 】

図 5 (a) は、P S N R の差分に対する平坦検出用閾値 2 0 2 の変化の様子を表すグラフ 5 0 1 を示す。ここで横軸は P S N R の差分、縦軸は図 3 (a) の関数 3 0 1 における $t h 1 F$ 、 $t h 2 F$ に対応する平坦検出用閾値である。即ち、閾値 $t h 1 F$ や $t h 2 F$ は、P S N R の差分に応じてグラフ 5 0 1 に対応するように変化する。

【 0 0 4 8 】

図 5 (b) はエッジ検出用閾値 2 0 4 の変化の様子を表すグラフ 5 0 2 を示す。ここで横軸は P S N R の差分、縦軸は図 3 の関数 3 0 2 における $t h 1 E$ 、 $t h 2 E$ に対応するエッジ検出用閾値である。即ち、閾値 $t h 1 E$ や $t h 2 E$ は、P S N R の差分に応じてグラフ 5 0 2 に対応するように変化する。

【 0 0 4 9 】

図 5 (c) は肌色検出用閾値 2 0 6 の変化の様子を表すグラフ 5 0 3 を示す。ここで横軸は P S N R の差分、縦軸は図 3 の関数 3 0 3 における $t h 1 S$ 、 $t h 2 S$ に対応する肌色検出用閾値である。即ち、閾値 $t h 1 S$ や $t h 2 S$ は、P S N R の差分に応じてグラフ 5 0 3 に対応するように変化する。

【 0 0 5 0 】

図 4 の領域 4 0 1 のように、P S N R 算出部 1 1 5 で算出された観測 P S N R が目標 P S N R より高い場合、画質は安定しているがより高画質化を行える。そこで、エッジ部及び肌色部では、より多く検出するように閾値を変動させる。なお、閾値の変動は図 5 に示しているように目標 P S N R と観測 P S N R の離れ具合 (劣化度合い) を考慮して連続的に変化させることが望ましい。また、平坦部に関しては量子化パラメータを変動させたとしても符号量が大きく増加することはないので閾値の変動は行わないようにする。

【 0 0 5 1 】

一方、図 4 の領域 4 0 2 のように、観測 P S N R が目標 P S N R より低く、劣化度合いが高い場合、全体的な画質が良くないためエッジのようなブロックに多くの符号量を与えても画質の向上は見込めない。そこで、エッジ部では検出閾値 ($T h 1 E$ と $T h 2 E$) を上げて強いエッジ、すなわち視覚的に特に目立つエッジのみを検出するようにする。また肌色部に対しては、人間の顔など重要領域が含まれることが多い。そこで、観測 P S N R が目標 P S N R より特に低く、劣化度合いが所定の度合いを超えて高い場合でも、閾値 ($T h 1 S$ と $T h 2 S$) を減少させて肌色部分を多く検出するようにする。この場合、顔以外の部分は劣化が起こっても、肌色部分の画質が向上することで全体的な視覚的印象は良くなる。

【 0 0 5 2 】

なお、閾値の変動は図 5 に示しているように目標 P S N R と観測 P S N R の離れ具合を考慮して連続的に変化させることが望ましい。また、各検出部の方法はここで述べた方法に限ったものではなく、検出を行うための要素となる分散値は周波数変換などで代用可能である。

【 0 0 5 3 】

以上のように、本実施形態に対応する符号化装置によれば、符号化画像の劣化度合いに応じて抽出するブロック数を変化させることで、発生符号量が安定した状態で視覚的劣化の目立ちやすいブロックに対する画質向上を行うことが可能となる。また、符号化画像の

10

20

30

40

50

劣化度合いが大きい場合でも、特に劣化を抑えたいブロックにより多くの符号量を与えるので画質が良いという印象を与えることができる。

【 0 0 5 4 】

[第 2 の実施形態]

第 1 の実施形態は、特徴抽出部 1 1 3 において、各マクロブロックが視覚的劣化の目立ちやすいブロックであるか否かを判断する基準を画像の劣化度合いである PSNR を用いていた。これに対して第 2 の実施形態では、再生時にフリッカが発生しそうな画像であるかに応じて各検出部に与えられる閾値を制御することを特徴とする。

【 0 0 5 5 】

図 6 は、第 2 の実施形態に対応する符号化装置の全体構成の一例を示すブロック図であり、基本的構成は図 1 と同じであるが、図 1 に対して、フリッカを検出するための機構が付加されている。なお、特徴抽出部 6 1 3 の構成は図 1 の特徴抽出部 1 1 3 と同じであるので図 2 を代用する。

【 0 0 5 6 】

図 6 において、フレーム並べ替え部 6 0 1、加減算部 6 0 2、直交変換 (D C T) 部 6 0 3、量子化部 6 0 4、逆量子化部 6 0 5、逆直交変換 (逆 D C T) 部 6 0 6、動き予測・動き補償部 6 0 7 を含む。また、加減算部 6 0 8、フレームメモリであるビデオバッファ 6 0 9、可変長符号化部 6 1 0、符号量制御部 6 1 1、量子化制御部 6 1 2、特徴抽出部 6 1 3、本符号化装置からのストリームを一時的に保持するバッファ 6 1 4 を含む。更に、入力画像と局所復号画像から P S N R (Peak Signal to Noise Ratio) を算出する P S N R 算出部 6 1 5、本符号化装置からの出力信号 1 1 6 を含む。以上は図 1 と同じ構成である。それに加えて、フレーム特徴検出部 6 2 0、フレーム動き検出部 6 2 1、フリッカ検出部 6 2 2 を設けた。なお、特徴抽出部 6 1 3 はマクロブロック単位で特徴を抽出するのに対して、フレーム特徴検出部 6 2 0 はフレーム単位で特徴を検出するものであり表現を分けておく。これらの作用を説明する。

【 0 0 5 7 】

フレーム特徴検出部 6 2 0 は、フレームアクティビティとして、入力端子 6 0 0 からの画像データから今から符号化する画像の複雑度を算出する。第 2 の実施形態では、複雑度として、画像データの交流成分量、好ましくは高周波成分量を採用する。具体的には、1 画面の画像データを所定サイズのブロックに分割し、各ブロックに対して分散を算出する。そして、各ブロックで算出した分散を画像の全ブロック数分加算した結果を、高周波成分量とする。なお、分散でなく、D C T (離散コサイン変換) やアダマール変換といった周波数変換を行い、その周波数成分で代用しても良い。

【 0 0 5 8 】

フレーム動き検出部 6 2 1 は、入力端子 6 0 0 からの画像データを隣接するフレーム間で相関をとり、今から符号化する画像全体がどれだけ動いたかを算出する。具体的には、1 画面の画像を所定サイズのブロックに分割し、そのブロック毎に、隣接する画面間で一方の画像の座標をずらしながら相関が最も高くなる座標ずれ量を算出する。そして、各ブロックで算出した動きベクトル量の画面内の総和をフレーム間動き量とする。なお、このフレーム間動き量は、大局的な動き (グローバルベクトル) を示すものであり、ここで示す方法以外の方法でも算出できる。

【 0 0 5 9 】

また、第 1 の実施形態と同様に、入力画像と局所復号画像とから符号化歪み量を算出するための P S N R 算出部 6 1 5 が存在する。P S N R 算出部 6 1 5 は、先ず、入力端子 6 0 0 からの画像データと局所復号画像データ (加減算部 6 0 8 の出力画像データ) とから、マクロブロック単位の P S N R を算出する。そして、P S N R 算出部 6 1 5 は、マクロブロック毎の P S N R の画面内の総和を、最終的な P S N R として出力する。ここで算出する P S N R は、符号化済みの画像に対するもの、即ち、今から符号化する画像に対して少なくとも 1 つ以上前に入力された画像に対するものである。

【 0 0 6 0 】

10

20

30

40

50

フリッカ検出部 6 2 2 には、フレーム特徴検出部 6 2 0 からのフレームアクティビティ、フレーム動き検出部 6 2 1 からのフレーム間動き量、及び P S N R 算出部 6 1 5 からの P S N R が入力する。フリッカ検出部 6 2 2 は、これらの 3 つのパラメータ値に従い、今から符号化する画像にフリッカが発生しそうかどうかを検出する。

【 0 0 6 1 】

ここで、フリッカの発生メカニズムについて図 7 を用いて以下に述べる。図 7 は、本実施形態に対応する、フリッカの発生メカニズムを説明するための図であり、フレーム内符号化とフレーム間符号化を併用する符号化方式における符号化後の輝度信号を示す。図 7 のグラフ 7 0 1 は、I ピクチャの再生信号レベルの時間変化を示す。横軸は時間（又はフレーム）を示し、縦軸は、再生映像信号レベルを示す。輝度平均で示す点線は映像信号の平均値を示している。これはグラフ 7 0 2、7 0 3 も同様である。ピーク輝度で示す点線は映像信号の輝度のピーク値を示している。これはグラフ 7 0 2、7 0 3 も同様である。フラットな映像信号に重畳するノイズ成分のピーク輝度は、フレーム内符号化によりある程度、再構成可能である。これは、上述した符号量制御アルゴリズムで、I ピクチャに対する符号量割り当てが、他のピクチャタイプに比べて多いことに起因する。

10

【 0 0 6 2 】

図 7 のグラフ 7 0 2 は、P ピクチャ及び B ピクチャのようなフレーム間符号化されたピクチャの再生信号レベルの時間変化を示す。横軸は時間（又はフレーム）を示し、縦軸は、再生映像信号レベルを示す。複雑度の高い画像はフレーム間の相関が低くなるので、通常の符号化では P ピクチャ及び B ピクチャのフレーム間差分情報量が増加する。その結果、符号化による映像信号の劣化が生じ、輝度のピークを再構成できなくなる。グラフ 7 0 2 では、グラフ 7 0 1 に示す I ピクチャに比べ、ピーク輝度に差が生じている。これにより、グラフ 7 0 3 に示すように、動画再生時に輝度フリッカが発生する。横軸は時間（又はフレーム）を示し、縦軸は、再生映像信号レベルを示す。

20

【 0 0 6 3 】

ただし、ピーク輝度差が大きいもの全てが、フリッカとして感じられるわけではない。人間の視覚特性は、動きの大きい画像の劣化よりも動きの小さい画像の劣化を検知しやすい。静止画で平坦部がざわざわしているのが気になるのはそのためである。そのため、動きの大きい画像ではピーク輝度差は検知できず、フリッカとして感じる事が少ない。動きのある画像よりも動きの小さい画像でピーク輝度差が生じると、フリッカとして目立ってしまう。

30

【 0 0 6 4 】

以上がフリッカ発生メカニズムである。フリッカが発生する条件として、符号化後の輝度ピーク値がピクチャタイプ毎に異なること、動きの少ない画像であることは上述した。符号化後の輝度ピーク値が生じる画像の条件は、1) 画像の複雑さが高いこと、及び、2) 符号化画像が劣化していることの二つの条件を満たしていることである。一つ目の条件は、フレーム特徴検出部 6 2 0 で算出した高周波成分量が高いことを意味する。二つ目の条件は、P S N R 算出部 6 1 5 で算出した P S N R が低いことを意味する。

【 0 0 6 5 】

なお、どちらか一方の条件を満たしていなければ、輝度ピーク差が生じるとは言えない。例えば、高周波を多く含んだ画像は画像の複雑さが高く、一つ目の条件は満たす。しかし、ビットレートが高い場合には、符号化画像は劣化していないので輝度ピーク差は生じない。一方、符号化画像が劣化していると、二つ目の条件は満たす。しかし、ビットレートが低い場合、画像の複雑さが低くても符号化画像が劣化する。この場合、輝度ピーク差は生じない。動きの少ない画像の条件は、画像全体の動き量が小さいことであり、フレーム動き検出部 6 2 1 で算出したフレーム間動き量が少ないことを意味する。

40

【 0 0 6 6 】

フリッカ検出部 6 2 2 は、高周波成分量、フレーム間動き量、PSNR の値からフリッカ発生度合いを算出することができる。高周波成分量がその基準値より高く、P S N R がその基準値より低く、フレーム動き量がその基準値より少ない場合、今から符号化しようとす

50

る画像でフリッカが発生する可能性が高いので、フリッカ発生度合いは高くなる。一方、その条件に合致しない場合、フリッカ発生度合いは低くなる。フリッカ発生度合いについて計算方法は問わないが、離散的な値であり、値が大きくなればフリッカ発生度合いが大きくなるものとする。

【 0 0 6 7 】

図 8 は、本実施形態に対応する、フリッカの発生度合いと検出閾値の関係を説明するための図である。図 8 は、横軸にフリッカ発生度合い、縦軸に各検出部の閾値を取ったグラフを示す。図 8 (a) は、フリッカ発生度合いに対する平坦検出用閾値 2 0 2 の変化の様子を表すグラフ 8 0 1 を示す。ここで横軸はフリッカ発生度合い、縦軸は図 3 の関数 3 0 1 における $t h 1 F$ 、 $t h 2 F$ に対応する平坦検出用閾値である。即ち、閾値 $t h 1 F$ や $t h 2 F$ は、フリッカの発生度合いに応じてグラフ 5 0 1 に対応するように変化する。

10

【 0 0 6 8 】

図 8 (b) はフリッカ発生度合いに対するエッジ検出用閾値 2 0 4 の変化の様子を表すグラフ 8 0 2 を示す。ここで横軸はフリッカ発生度合い、縦軸は図 3 の関数 3 0 2 における $t h 1 E$ 、 $t h 2 E$ に対応するエッジ検出用閾値である。即ち、閾値 $t h 1 E$ や $t h 2 E$ は、フリッカの発生度合いに応じてグラフ 5 0 2 に対応するように変化する。

【 0 0 6 9 】

図 8 (c) はフリッカ発生度合いに対する肌色検出用閾値 2 0 6 の変化の様子を表すグラフ 8 0 3 を示す。ここで横軸はフリッカ発生度合い、縦軸は図 3 の関数 3 0 3 における $t h 1 S$ 、 $t h 2 S$ に対応する肌色検出用閾値である。即ち、閾値 $t h 1 S$ や $t h 2 S$ は、フリッカの発生度合いに応じてグラフ 5 0 3 に対応するように変化する。

20

【 0 0 7 0 】

フリッカ検出部 6 2 2 の検出結果から、フリッカ発生度合いが高い場合、劣化が目立つ領域に対して符号量を与え過ぎてしまうと、その反動で高周波成分を含む領域に対する符号量が少なくなってしまう、フリッカが強く発生してしまう。そこで、エッジ部では検出閾値 ($T h 1 E$ と $T h 2 E$) を上げて強いエッジ、すなわち視覚的に特に目立つエッジのみを検出するようにする。また肌色部では、検出閾値 ($T h 1 S$ と $T h 2 S$) を上げて強い肌色、すなわち視覚的に特に目立つ肌色のみを検出するようにする。

【 0 0 7 1 】

一方、フリッカ検出部 6 2 2 の検出結果から、フリッカ発生度合いが低い場合、画像に含まれる高周波成分が少なく、 $P S N R$ も十分高いため、劣化が目立つ領域に対して更に符号量を与えることが可能である。そこで、エッジ部では検出閾値 ($T h 1 E$ と $T h 2 E$) を下げてエッジを多く検出するようにする。また肌色部では、検出閾値 ($T h 1 S$ と $T h 2 S$) を下げて肌色を多く検出するようにする。なお平坦部に関しては、量子化パラメータを変動させたとしても符号量が大きく増加することはないのでフリッカの発生度合いによっては平坦閾値 ($T h 1 F$ と $T h 2 F$) を変化させない。

30

【 0 0 7 2 】

以上のように、本実施形態に対応する符号化装置によれば、符号化画像のフリッカの発生度合いに応じて抽出するブロック数を変化させることで、発生符号量が安定した状態で視覚的劣化の目立ちやすいブロックに対する画質向上を行うことが可能となる。また、符号化画像のフリッカの発生度合いが大きい場合でも、特に劣化を抑えたいブロックにより多くの符号量を与えるので画質が良いという印象を与えることができる。

40

【 0 0 7 3 】

[その他の実施形態]

本発明の目的は、前述した機能を実現するコンピュータプログラムのコードを記録した記憶媒体を、システムに供給し、そのシステムがコンピュータプログラムのコードを読み出し実行することによっても達成される。この場合、記憶媒体から読み出されたコンピュータプログラムのコード自体が前述した実施形態の機能を実現し、そのコンピュータプログラムのコードを記憶した記憶媒体は本発明を構成する。また、そのプログラムのコードの指示に基づき、コンピュータ上で稼働しているオペレーティングシステム ($O S$) など

50

が実際の処理の一部または全部を行い、その処理によって前述した機能が実現される場合も含まれる。

【 0 0 7 4 】

さらに、以下の形態で実現しても構わない。すなわち、記憶媒体から読み出されたコンピュータプログラムコードを、コンピュータに挿入された機能拡張カードやコンピュータに接続された機能拡張ユニットに備わるメモリに書込む。そして、そのコンピュータプログラムのコードの指示に基づき、その機能拡張カードや機能拡張ユニットに備わるCPUなどが実際の処理の一部または全部を行って、前述した機能が実現される場合も含まれる。

【 0 0 7 5 】

本発明を上記記憶媒体に適用する場合、その記憶媒体には、先に説明したフローチャートに対応するコンピュータプログラムのコードが格納されることになる。

【 図面の簡単な説明 】

【 0 0 7 6 】

【 図 1 】 発明の第 1 の実施形態に対応する、符号化装置の構成の一例を示すブロック図である。

【 図 2 】 発明の第 1 の実施形態に対応する、特徴抽出部の処理を説明するための図である。

【 図 3 】 発明の第 1 の実施形態に対応する、各検出部の強さをレベル化するための関数を示す図である。

【 図 4 】 発明の第 1 の実施形態に対応する、PSNRについて説明するための図である。

【 図 5 】 発明の第 1 の実施形態に対応する、PSNRと検出閾値の関係を説明するための図である。

【 図 6 】 発明の第 2 の実施形態に対応する、符号化装置の構成の一例を示すブロック図である。

【 図 7 】 発明の第 2 の実施形態に対応する、フリッカの発生のメカニズムを説明するための図である。

【 図 8 】 発明の第 2 の実施形態に対応する、フリッカの発生度合いと検出閾値の関係を説明するための図である。

【 符号の説明 】

【 0 0 7 7 】

- 1 0 0 . . . 入力部 (入力信号)
- 1 0 1 . . . フレーム並べ替え部
- 1 0 2 . . . 加減算部
- 1 0 3 . . . 直交変換部
- 1 0 4 . . . 量子化部
- 1 0 5 . . . 逆量子化部
- 1 0 6 . . . 逆直交変換部
- 1 0 7 . . . 動き予測・動き補償部
- 1 0 8 . . . 加減算部
- 1 0 9 . . . ビデオバッファ (フレームメモリ)
- 1 1 0 . . . 可変長符号化部
- 1 1 1 . . . 符号量制御部
- 1 1 2 . . . 量子化制御部
- 1 1 3 . . . 特徴抽出部
- 1 1 4 . . . バッファ
- 1 1 5 . . . PSNR 算出部
- 1 1 6 . . . 出力部 (出力信号、ストリーム)

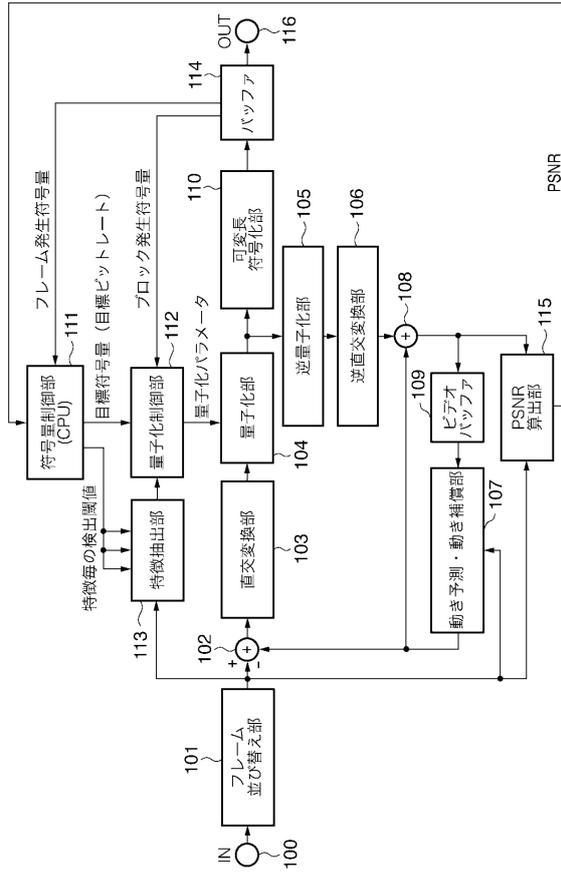
10

20

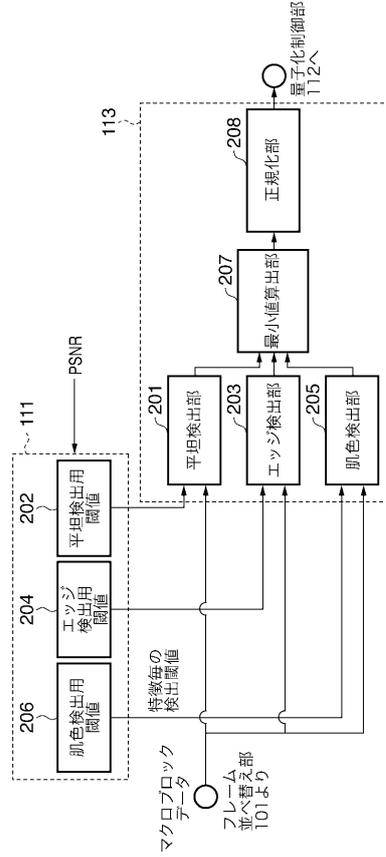
30

40

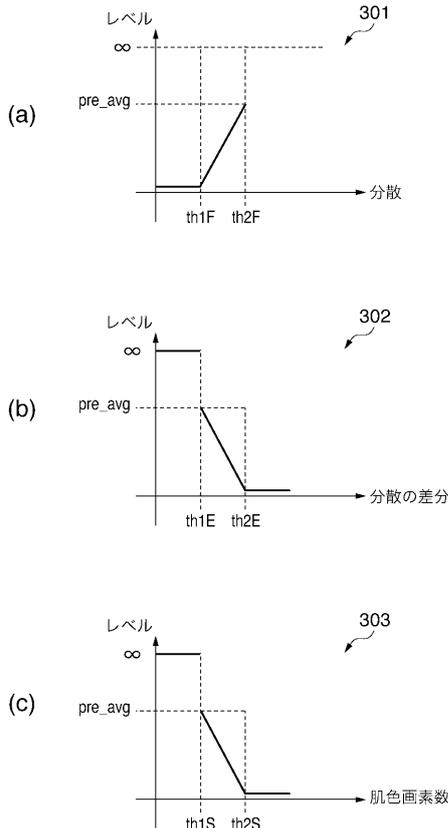
【図1】



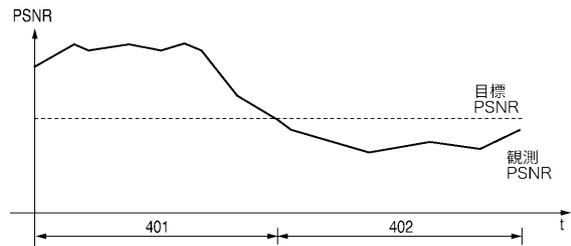
【図2】



【図3】



【図4】



フロントページの続き

(72)発明者 中山 文貴
東京都大田区下丸子3丁目30番2号 キヤノン株式会社内

審査官 坂東 大五郎

(56)参考文献 特開2002-185966(JP,A)
特開2007-067469(JP,A)
特開2002-238060(JP,A)
特開平02-105792(JP,A)

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