

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第6312775号
(P6312775)

(45) 発行日 平成30年4月18日(2018.4.18)

(24) 登録日 平成30年3月30日(2018.3.30)

(51) Int.Cl. F I
HO4N 19/85 (2014.01) HO4N 19/85

請求項の数 13 外国語出願 (全 28 頁)

(21) 出願番号	特願2016-213145 (P2016-213145)	(73) 特許権者	507236292
(22) 出願日	平成28年10月31日(2016.10.31)		ドルビー ラボラトリーズ ライセンシング
(62) 分割の表示	特願2015-563023 (P2015-563023) の分割		グ コーポレイション
原出願日	平成26年6月16日(2014.6.16)		アメリカ合衆国 94103 カリフォル
(65) 公開番号	特開2017-79471 (P2017-79471A)		ニア州 サンフランシスコ マーケット
(43) 公開日	平成29年4月27日(2017.4.27)	(74) 代理人	100101683
審査請求日	平成28年12月16日(2016.12.16)		弁理士 奥田 誠司
(31) 優先権主張番号	61/836,044	(74) 代理人	100155000
(32) 優先日	平成25年6月17日(2013.6.17)		弁理士 喜多 修市
(33) 優先権主張国	米国 (US)	(74) 代理人	100135703
(31) 優先権主張番号	61/951,914		弁理士 岡部 英隆
(32) 優先日	平成26年3月12日(2014.3.12)	(74) 代理人	100188813
(33) 優先権主張国	米国 (US)		弁理士 川喜田 徹

最終頁に続く

(54) 【発明の名称】 エンハンスダイナミックレンジ信号の階層符号化のための適応的再構成

(57) 【特許請求の範囲】

【請求項 1】

ビデオ符号化の際にクロマに関連するアーチファクトを低減する方法であって、輝度成分および2つの色度成分を含むある色空間における入力信号にアクセスする工程と、

前記入力信号の前記色空間における白色点の元の白色点色度座標(D_u、D_v)を決定する工程と、

色度変換関数を用いて、前記入力信号を、2つの変換された色度成分を含む変換後の色空間における別の信号に変換する工程であって、前記色度変換関数は、前記元の白色点色度座標を前記変換後の色度色空間における所定の別の1組の座標にシフトさせる、工程と

10

、前記別の信号を符号化することにより符号化ビットストリームを生成する工程と、を包含し、

前記方法は1つ以上のプロセッサによって実行される方法。

【請求項 2】

前記入力信号の前記色空間は、知覚的量子化された輝度成分(Luma)ならびにu'およびv'色度成分を含む、請求項1に記載の方法。

【請求項 3】

前記色度変換関数は、

【数 3 2】

$$C0_t = (C0 - Dc)a0 + b0 \quad ,$$

$$C1_t = (C1 - Dv)a1 + b1 \quad ,$$

を算出することを含み、ここで、C 0 および C 1 は、入力色度値を表し、C 0_t および C 1_t は、変換された色度値を表し、D c および D v は、前記元の白色点座標値を表し、a 1 および a 2 は定数であり、b 1 および b 2 は、前記白色点座標の前記所定の別の組の座標値を表す、請求項 1 に記載の方法。

10

【請求項 4】

b 0 および b 1 は 0 . 5 に等しく、a 1 および a 2 は 1 . 0 から 1 . 5 の間である、請求項 3 に記載の方法。

【請求項 5】

前記色度変換関数は、

【数 3 3】

$$C0_t = g(Luma) (C0 - Dc)a0 + b0 \quad ,$$

$$C1_t = g(Luma) (C1 - Dv)a1 + b1 \quad ,$$

20

を算出することを含み、ここで、C 0 および C 1 は、入力色度値を表し、C 0_t および C 1_t は、変換された色度値を表し、D c および D v は、前記元の白色点座標値を表し、a 1 および a 2 は定数であり、b 1 および b 2 は、前記白色点座標の前記所定の別の組の座標値を表し、g (L u m a) は前記輝度成分の関数を表す、請求項 1 に記載の方法。

【請求項 6】

前記符号化工程の前に、前記別の信号の色度成分をサブサンプリングする工程をさらに包含する、請求項 1 に記載の方法。

【請求項 7】

前記白色点の、前記所定の別の組の座標は、前記変換後の色度色空間の中心にある、請求項 1 に記載の方法。

30

【請求項 8】

前記色度変換関数は、

【数 3 4】

$$C0_t = g(I) (C0 - a) + a \quad ,$$

$$C1_t = g(I) (C1 - a) + a \quad ,$$

を算出することを含み、ここで、C 0 および C 0 は、入力色度値を表し、C 0_t および C 0_t は、変換された色度値を表し、g (I) は、前記入力信号の前記輝度成分の関数を表し、a は定数であり、ここで、C は知覚的量子化された I P T 色空間 (I P T - P Q) 内の信号の色差成分を含み、I は前記知覚的量子化された I P T 色空間内の信号の輝度成分を含む、請求項 1 に記載の方法。

40

【請求項 9】

デコーダにおいて、変換後の色空間において符号化されたビデオ信号を復号化する方法であって、

輝度成分および 2 つの変換された色度成分を含む変換後の色空間における符号化されたビットストリームを受信する工程であって、前記 2 つの変換された色度成分は、色度変換関数および一組の元の白色点色度座標に基づいてエンコーダによって変換されたものであ

50

る、工程と、

前記符号化されたビットストリームを復号化することによって復号化信号を生成する工程と、

前記元の白色点色度座標および前記色度変換関数の逆関数を用いて、前記復号化信号を2つの元の色度成分を有する色空間における出力信号に変換する工程であって、前記エンコーダにおいて、前記色度変換関数は、前記元の白色点色度座標を、前記変換後の色度色空間における所定の別の組の座標にシフトさせる、工程と、

を包含し、

前記方法は、1つ以上のプロセッサによって実行される方法。

【請求項10】

前記色度変換関数は、

【数35】

$$C0_t = (C0 - Dc)a0 + b0 \quad ,$$

$$C1_t = (C1 - Dv)a1 + b1 \quad ,$$

を算出することを含み、ここで、 $C0$ および $C1$ は、入力色度値を表し、 $C0_t$ および $C1_t$ は、変換された色度値を表し、 Dc および Dv は、前記元の白色点色度座標を表し、 $a1$ および $a2$ は定数であり、 $b1$ および $b2$ は、前記白色点座標の前記所定の別の組の座標値を表す、請求項9に記載の方法。

【請求項11】

前記色度変換関数は、

【数36】

$$C0_t = g(Luma) (C0 - Dc)a0 + b0 \quad ,$$

$$C1_t = g(Luma) (C1 - Dv)a1 + b1 \quad .$$

を算出することを含み、ここで、 $C0$ および $C1$ は、入力色度値を表し、 $C0_t$ および $C1_t$ は、変換された色度値を表し、 Dc および Dv は、前記元の白色点色度座標を表し、 $a1$ および $a2$ は定数であり、 $b1$ および $b2$ は、前記白色点座標の前記所定の別の組の座標値を表し、 $g(Luma)$ は、前記入力信号の前記輝度成分の関数を表す、請求項9に記載の方法。

【請求項12】

ビデオ符号化の際に色に関連するアーチファクトを低減する装置であって、

輝度成分および2つの色度成分を含むある色空間における入力信号を受信する入力段(input)と、

プロセッサおよびメモリであって、

前記入力信号の前記色空間における白色点の元の白色点色度座標(Du 、 Dv)を決定し、

色度変換関数を用いて、前記入力信号を、2つの変換された色度成分を含む変換後の色空間における別の信号に変換し、ここで、前記色度変換関数は、前記元の白色点色度座標を前記変換後の色度色空間における所定の別の組の座標にシフトさせ、

前記別の信号を符号化することにより符号化ビットストリームを生成する、プロセッサおよびメモリと、

前記符号化ビットストリームを出力する出力段(output)と、

を備える装置。

【請求項13】

請求項1に記載の方法を1つ以上のプロセッサで実行するためのコンピュータが実行可

10

20

30

40

50

能な命令を格納した非一時的コンピュータ可読記憶媒体。

【発明の詳細な説明】

【技術分野】

【0001】

本願は、2013年6月17日出願の米国特許仮出願第61/836,044号、2014年3月12日出願の米国特許仮出願第61/951,914号、および、2014年5月23日出願の米国特許仮出願第62/002,631号に基づく優先権を主張するものであり、これらの出願の開示内容の全てを本願に援用する。

【0002】

本願は、2014年3月25日出願の国際出願第PCT/US2014/031716号にも関連し得るものであり、この出願の開示内容の全てを本願に援用する。 10

【0003】

本発明は、広義には、ビデオ画像に関する。より具体的には、本発明のある実施形態は、階層符号化および復号化のための、ハイまたはエンハンスト(enhanced)ダイナミックレンジを有する画像の適応的再構成(adaptive reshaping)に関する。

【背景技術】

【0004】

本明細書において、用語「ダイナミックレンジ」(DR)は、画像においてある範囲の強度(例えば、輝度、ルマ)(例えば、最暗部(黒)から最明部(白)まで)を知覚する人間の心理視覚システム(HVS)の能力に関連し得る。この意味では、DRはシーン-リファード(scene-referred)の強度に関する。DRはまた、ディスプレイデバイスが特定の幅を有する強度範囲を妥当にまたは近似的に描画する能力にも関連し得る。この意味では、DRは、ディスプレイ-リファード(display-referred)の強度に関する。本明細書中の任意の箇所において、ある特定の意味が特に明示的に指定されている場合を除いて、この用語はどちらの意味としても(例えば、区別なく)使用できるものとする。 20

【0005】

本明細書において、ハイダイナミックレンジ(HDR)という用語は、人間の視覚システム(HVS)において14~15桁ほどにわたるDR幅に関する。例えば、しっかりと順応した、実質的に正常(例えば、統計的な意味、バイオメトリック的な意味、または、眼科的な意味の1つ以上において)な人間は、約15桁にわたる強度範囲を有する。順応した人間は、ほんの少量の光子のうす暗い光源を知覚し得る。また、同じ人間が、砂漠、海または雪上における真昼の太陽の、目が痛くなるほどの明るい強度を知覚し得る(また、怪我をしないような短時間であるとしても太陽を見ることもできる)。ただし、この範囲は、「順応した人間」、例えば、そのHVSをリセットして調節する時間を経た人間に当てはまるものである。 30

【0006】

これに対して、人間が、広範囲の強度範囲を同時に知覚し得るDRは、HDRに対して幾分端折られ得る。本明細書において、エンハンストダイナミックレンジ(EDR)または視覚ダイナミックレンジ(VDR)という用語は、個別にまたは区別なく、HVSによって同時に知覚可能なDRに関連し得る。本明細書において、EDRは、5~6桁にわたるDRに関連し得る。従って、真のシーンリファードのHDRに対して幾分狭いものの、EDRは、広いDR幅を示す。 40

【0007】

実際には、画像は、1つ以上の色成分(例えば、輝度Yおよび色差CbおよびCr)を有し、各色成分はnビット/ピクセルの精度で表される(例えば、n=8)。輝度ダイナミックレンジとビット深度は同じものではないが、しばしば相互に関連性がある。n=8の画像(例えば、カラー24ビットJPEG画像)は標準的なダイナミックレンジの画像とされ、n>8の画像はエンハンストダイナミックレンジの画像とされる。EDRおよび 50

HDR画像は、"Industrial Light and Magic"が開発したOpenEXRファイルフォーマットなどの高精度の(例えば、16ビット)浮動小数点フォーマットを用いて、保存および配信され得る。

【0008】

ビデオ信号は、ビット深度、色空間、色域および解像度などの複数のパラメータによって特徴付けられる。最新のテレビおよびビデオ再生デバイス(例えば、ブルーレイプレーヤー)は、標準解像度(例えば、720×480i)および高精細(HD)(例えば、1920×1080p)を含む様々な解像度をサポートしている。超高精細(UHD)は、少なくとも3,840×2,160の解像度("4K UHD")を有し、7680×4320("8K UHD")まで上げられるオプションを有する次世代解像度フォーマットである。超高精細は、UltraHD、UHDTVまたはスーパーハイビジョンとも呼ばれることがある。本明細書において、UHDは、HD解像度よりも高い任意の解像度を指す。

10

【発明の概要】

【発明が解決しようとする課題】

【0009】

新しいHDRまたはUHD符号化およびディスプレイ技術だけでなく、旧式の8ビットの再生デバイスとの下位互換性をサポートするために、上流側デバイスから下流側デバイスへとUHDおよびHDR(またはEDR)ビデオデータを送信する際に、複数フォーマットが用いられ得る。あるEDRストリームに対して、あるデコーダは、1セットの8ビットレイヤを用いて、コンテンツのHD SDRまたはEDRバージョンを再構築し得る。新型のデコーダは、従来の8ビットよりも高いビット深度で符号化された第2のセットのレイヤを用いて、コンテンツのUHD EDRバージョンを再構築し、それをより高性能なディスプレイ上に描画し得る。本願発明者も理解しているように、EDRビデオの符号化および配信のためにより向上した技術が望まれる。

20

【0010】

本節に記載されている手法は、探求し得る手法ではあるが、必ずしもこれまでに着想または探求されてきた手法ではない。従って、特に反対の記載がない限り、本節に記載された手法のいずれも、本節に記載されているという理由だけで従来技術としての適格性を有すると考えるべきではない。同様に、特に反対の記載がない限り、1以上の手法に関して特定される問題が、本節に基づいて、いずれかの先行技術において認識されたことがあると考えるべきではない。

30

【図面の簡単な説明】

【0011】

【図1A】図1Aは、本発明のある実施形態による、EDR階層符号化のフレームワーク例を示す図である。

【図1B】図1Bは、本発明のある実施形態による、EDR階層復号化のフレームワーク例を示す図である。

【図2】図2は、本発明のある実施形態による、関数パラメータaを決定する、べき関数に基づくEDR信号再構成関数の一例を示す図である。

40

【図3】図3は、本発明のある実施形態による、EDR入力の順方向構成関数の最適な指数を決定する処理例を示す図である。

【図4】図4は、本発明のある実施形態による、EDR符号語の順方向マッピングを決定する処理例を示す図である。

【図5】図5は、入力EDR符号語(v_c)を、本発明のある実施形態に従って算出される、ブロックベースのスケーリング係数($k(v_c)$)に中間マッピングする一例を示す図である。

【図6】図6は、本発明のある実施形態による、入力EDR符号語を最終出力である再構成済みシンボルにマッピングする一例を示す図である。

【図7】図7は、本発明のある実施形態に従って算出される、逆方向マッピングの一例を

50

示す図である。

【図8】図8Aおよび図8Bは、本発明のある実施形態による色度範囲スケーリングの一例を示す図である。

【図9】図9は、本発明のある実施形態による、符号化および復号化パイプラインの一例を示す図である。

【発明を実施するための形態】

【0012】

エンハンスダイナミックレンジ(EDR)を有するビデオ画像の階層符号化のための、適応的再構成技術を本明細書に説明する。以下の説明においては、便宜上、本発明を完全に理解できるように、多数の詳細事項を説明する。ただし、これらの詳細事項が無くても本発明を実施可能であることは明白であろう。他方、本発明の説明を不必要に煩雑にしたり、不明瞭にしたり、難読化したりしないように、周知の構造およびデバイスの細かな詳細までは説明しない。

10

【0013】

同様の部材に同様の参照符号を付した添付図面の各図において、本発明のある実施形態を限定する事なく例示する。

【0014】

概要

本明細書に記載の実施形態例は、効率的な階層符号化を行うための、ハイまたはエンハンスダイナミックレンジを有するビデオ画像の適応的再構成に関する。エンコーダは、レイヤ表現で符号化されるための入力エンハンスダイナミックレンジ(EDR)画像を受信する。入力画像は、ガンマ符号化されていてもよいし、利用可能なビデオエンコーダの1つ以上でサポートされていないビット深度フォーマットを用いて知覚的符号化されていてもよい。入力画像を1つ以上の量子化レイヤに再マッピングし、これにより、利用可能なビデオエンコーダを用いた圧縮に適した出力符号語を生成する。

20

【0015】

ある実施形態において、再マッピングは、単一の関数パラメータを用いたべき関数に基づいている。入力EDR画像中の各ブロックについてブロックベースの複雑性尺度を算出し、その後、量子化画像における量子化に起因する歪みの量を評価することにより、最適な関数パラメータを求める技術を説明する。

30

【0016】

別の実施形態においては、標準偏差のようなブロックベースの複雑性尺度と、ブロックベースの線形量子化モデル(各画像ブロックについて別々の最適な量子化器スケーラを求める)とを用いて、最適マッピングを生成する。この別々の最適スケーラを組み合わせることにより、各入力符号語について包絡線の傾きを求め、入力および出力符号語間の最適な順方向マッピング関数を包絡線の傾きに基づいて求める。逆方向マッピング関数を、ルックアップテーブルとしてデコーダに送信してもよいし、区分的多項式近似を用いて近似してもよい。

【0017】

別の実施形態において、ある逆方向マッピングルックアップテーブルが与えられると、多項式近似技術を用いて逆方向マッピング関数が近似される。

40

【0018】

デコーダにおいて、符号化ビットストリームレイヤを復号化することによって復号化ビデオレイヤを生成し、これら復号化ビデオレイヤを再結合して1つの復号化信号を生成する。そして、エンコーダの再構成(reshaping)またはマッピング関数を定義する受信パラメータが与えられると、復号化信号を逆方向マッピングして、これにより、エンコーダからデコーダに送信される元のEDR信号の推定結果を生成する。

【0019】

別の実施形態においては、所望の白色点の座標が変換後の色度範囲の中心に概ねシフトされるように、入力ビデオ信号の色度色成分を変換してもよい。

50

【 0 0 2 0 】

ビデオ信号再構成および階層分解のためのフレームワーク例
階層符号化および復号化

HDTV、セットトップボックスまたはブルーレイプレーヤーのような既存のディスプレイおよび再生デバイスは、典型的に、最大1080pHD解像度（例えば、1920×1080 60フレーム/秒）までの信号をサポートしている。家庭用途では、このような信号は、現在、典型的には色差成分が輝度成分よりも低い解像度を有するような輝度-色差フォーマット（例えば、YCbCrまたはYUV4:2:0色フォーマット）で、各色成分について8ビット/ピクセルのビット深度を用いて典型的に圧縮される。8ビットの深度およびそれに対応する低ダイナミックレンジのために、これらの信号は典型的には、標準ダイナミックレンジ（SDR）を有する信号と呼ばれる。超高精細（UHD）などの新しいテレビ規格が開発される中、エンハンスド（enhanced）解像度および/またはエンハンスドダイナミックレンジを用いて信号を符号化することが望まれ得る。

10

【 0 0 2 1 】

ビデオ画像は、典型的には、人間の視覚システムの特徴を補うためにガンマ符号化される。例えば、ITU-R Rec 2020は、推奨されるUHD TV信号のガンマ符号化を定義している。EDR画像については、知覚量子化（PQ）は、従来のガンマ符号化よりも良い代替法であり得る。人間の視覚システムは、光レベルの増大に対して非常に非線形的に反応する。人間が刺激を見る能力は、その刺激の輝度、その刺激の大きさ、その刺激を構成する空間周波数、および、その刺激を見ている瞬間までに目が順応した輝度レベルに影響される。知覚的量子化器関数は、線形入力グレイレベルを、人間の視覚システムにおけるコントラスト感度閾値によりマッチした出力グレイレベルにマッピングする。PQマッピング関数の一例は、J.S. Millerらによる2012年12月06日出願の「Perceptual luminance nonlinearity-based image data exchange across different display capabilities」という表題のPCT出願シリアル番号第PCT/US2012/068212号（以下、「'212出願」と呼ぶ）に記載されており、この出願の開示内容を全て本願に援用する。上記PCT出願において、ある固定刺激サイズに対して、それぞれの輝度レベル（即ち、刺激レベル）について、最高感度の適応レベルおよび最高感度の空間周波数（HVSモデルによる）に応じて、その輝度レベルにおける最小可視コントラストステップを選択する。物理的な陰極線管（CRT）装置の応答曲線を表しており、人間の視覚システムの応答の仕方に対して非常に大まかな類似性を偶然有し得る従来のガンマ曲線と比較して、「'212出願において決定されているPQ曲線は、比較的シンプルな関数モデルを用いながら人間の視覚システムの本当の視覚応答を模擬している。

20

30

【 0 0 2 2 】

本願にその開示内容が援用される2013年3月26日出願の「Encoding perceptually-quantized video content in multi-layer VDR coding」という表題の米国特許仮出願シリアル番号第61/805,388号（以下「'388出願」と呼ぶ）（2014年3月25日にPCT/US2014/031716としても出願されている）において、発明者らは、2レイヤエンコーダを用いたPQ符号化EDR画像データの効率的な符号化および送信を行うための、画像再構成技術を記載している。本願は、シングルレイヤまたはマルチレイヤエンコーダの両方を用いた、EDRデータの符号化に適用可能な新規なマッピングまたは再構成技術を記載することによって「'388出願を発展させたものである。

40

【 0 0 2 3 】

図1Aは、本発明のある実施形態による、EDR階層符号化のフレームワーク例を示す。入力信号（102）は、ガンマまたはPQ符号化され得るEDRピクセル値を有する、ビデオフレームシーケンスを含む。システムは、少なくとも1つのベースレイヤ（BL）

50

ビデオエンコーダ(120-0)を含み、合計L個の符号化レイヤに対して、1つ以上、最大L-1個のエンハンスメントレイヤ(EL)ビデオエンコーダ(120-1, 120-L-1)を含み得る。例えば、L=2の場合、システムは、1つのデュアルレイヤエンコーダを含む。ビデオエンコーダ(120)は全て同じでもよく、または異なってもよく、MPEG-2、MPEG-4、part 2、H.264(またはAVC)、H.265(またはHEVC)等のような既知の、または未来のビデオ圧縮用符号化フォーマットを実行する。さらに、あるレイヤのビデオエンコーダが、別のレイヤのビデオエンコーダとは異なるビット深度をサポートし得る。例えば、限定はしないが、実施形態は、以下の構成を有していてもよい。

- ・ 少なくとも10ビットのビット深度をサポートし、好ましくは12ビット以上をサポートする、シングルレイヤのHEVCエンコーダ。

10

- ・ デュアルレイヤエンコーダであって、両方のエンコーダが同じフォーマット(例えばH.264)を用いて符号化を行い、両方のエンコーダが同じまたは異なるビット深度(例えば8ビットおよび10ビット)をサポートするようなデュアルレイヤエンコーダ。

- ・ デュアルレイヤエンコーダであって、2つのエンコーダは異なる符号化フォーマットを用いて符号化を行い、それぞれが異なるビット深度(例えば8ビットおよび10ビット以上)をサポートし得るような、デュアルレイヤエンコーダ。

- ・ マルチレイヤエンコーダであって、少なくとも1つのエンコーダが8ビットMPEG-2エンコーダであり、少なくとも他方のエンコーダがHEVCまたはH.264エンコーダである、マルチレイヤエンコーダ。

20

ビデオエンコーダ(120)は、全て、単一のプロセッサまたは1つ以上のプロセッサによって実施され得る。

【0024】

ある実施形態によれば、信号再構成モジュール(110)は、入力EDR信号(vと表記する)(102)を量子化して、最大サポートビット深度などのビデオエンコーダ(120)の特性により良く適合する、信号s(112)を得る。本明細書において、「再構成(reshaping)」、「量子化」および「(順方向)マッピング」の用語は、入力信号を第1のダイナミックレンジから(典型的には第1のダイナミックレンジよりも低い)第2のダイナミックレンジの出力信号へとマッピングする同様な機能を指し、互いに

30

【0025】

B_l がl番目のレイヤビデオエンコーダ(120-l、 $l=0, 1, 2, \dots, L-1$)が用いるビット深度を指すものとする、各レイヤは、合計 $N_T = N_0 + N_1 + \dots + N_{L-1}$ 個の符号語に対して、最大 $N_l = 2^{B_l}$ 個の入力符号語をサポートできる。例えば、 $L=1$ (シングルレイヤ)および $B_0=10$ の場合、 $2^{10}=1024$ 個の量子化符号語が存在する。 $L=2$ (デュアルレイヤ)および $B_0=B_1=8$ の場合、 $2^8+2^8=512$ 個の量子化符号語が存在する。 $L=2$ および $B_0=10$ および $B_1=8$ の場合、合計 $2^{10}+2^8=1280$ 個の量子化符号語が存在する。よって、システムは、それぞれが独自のビット深度で動作する複数のビデオ符号化規格の、任意の組み合わせに対応し得る。

【0026】

40

$s = f(v)$ が信号再構成/量子化関数(110)を表すものとする。このような関数の例は詳しく後述する。再構成関数を特定するパラメータが、メタデータ信号(119)として含まれていてもよい。いくつかの実施形態において、メタデータ(119)はメタデータエンコーダ(125)によって符号化され、符号化メタデータ(127)は、適切な逆量子化および復号化のために、図1Bに示したようなデコーダに送信され得る。別の実施形態において、信号再構成(110)は、1群の信号再構成関数を含み、1つ以上のレイヤに対して、または1レイヤ内の1つ以上の色度成分に対して、別々の構成関数を用いてもよい。例えば、ある実施形態において、ベースレイヤ($l=0$)の信号再構成関数は線形関数である一方で、第1のエンハンスメントレイヤ($l=1$)の信号再構成関数は非線形関数または区分線形関数を含んでいてもよい。

50

【0027】

レイヤ分解

ある実施形態において、量子化信号 s (112) のピクセル値を、区分境界 $\{p_i, i = 0, 1, \dots, L\}$ によって定義される L 個の区分に分割するものとする。ただし、 p_0 は典型的には s の最小値 (例えば、 $p_0 = 0$) を表し、

【数1】

$$p_i = \sum_{j=0}^{i-1} N_j \quad \text{for } i = 1, \dots, L \quad (1)$$

10

である。

【0028】

例えば、 $L = 1$ の場合、 $p_0 = 0$ および $p_1 = N_0$ である。このモジュールは、全符号語をベースレイヤに符号化する。

【0029】

$L = 2$ の場合、 $p_0 = 0$ 、 $p_1 = N_0$ 、 $p_2 = N_0 + N_1$ である。ある実施形態において、 $\{p_0, p_1\}$ 間の符号語を有するピクセルはレイヤ0で符号化され、 $\{p_1, p_2\}$ 間の符号語を有するピクセルはレイヤ1で符号化される。一般的に、 L 個のレイヤがあった場合、各レイヤ l に対して、当該レイヤの s_l ピクセルは、以下のように符号化される。

$$s_l = \text{Clip3}(s, p_l, p_{l+1} - 1) - p_l, \\ l = 0, 1, 2, \dots, L - 1 \quad (2)$$

20

ただし、 $d = \text{Clip3}(s, a, b)$ はクリッピング関数を表し、ここで、 $a \leq s \leq b$ であれば $d = s$ であり、 $s < a$ であれば $d = a$ であり、 $s > b$ であれば $d = b$ である。

【0030】

レイヤ分解 (115) 後、ある実施形態においては、各レイヤ s_l (117 - 1、 $l = 0, 1, \dots, L - 1$) をビデオエンコーダ (120) でそれぞれ独立して符号化し、これにより、圧縮ビットストリーム (122) を生成してもよい。'388出願において説明されているように、いくつかの実施形態において、図1Aに示すシステムを改変して、レイヤ間予測をも可能にしてもよい。そのような実施形態においては、予測器を用いて、 $l - 1$ レイヤのピクセル値に基づいて l 番目のレイヤのピクセル値を推定すればよく、その後、 l 番目のレイヤのピクセル値を直接符号化するのではなく、実値および予測値の間の残差を単に符号化して送信する。

30

【0031】

いくつかの実施形態において、符号化ビットストリーム (122)、符号化メタデータ (127) および他のデータ (例えば、オーディオデータ) を多重化して単一のビットストリームにし、単一の多重化されたビットストリーム (図示せず) としてデコーダに送信してもよい。

【0032】

図1Bは、本発明のある実施形態によるEDR階層復号化のフレームワーク例を示す。図1Bに示すように、オーディオ、ビデオおよび補助的なデータ (図示せず) を組み合わせ得る受信ビットストリームを多重分離した後、受信符号化ビットストリーム (122) をそれぞれビデオデコーダ (130) アレイに送る。デコーダ群 (130) はエンコーダ群 (120) に対応し、1つ以上の復号化ビデオ信号 (132) を生成する。信号逆再構成 (inverse reshaping) およびレイヤ結合器 (140) ユニットを用いて、受信した階層信号を結合および逆再構成して信号 (\wedge 付 v) (142) (元のEDR信号 (102) の推定結果を表す) を生成する。ある実施形態において、出力EDR信号 (142) は以下のように生成され得る。

40

【数 2】

$$\hat{v} = f^{-1}\left(\sum_{i=0}^{L-1} \hat{s}_i\right) \quad (3)$$

ただし、 $f^{-1}(\)$ は、信号再構成関数 (1 1 0) の逆数 (または逆数の近い近似値) を表し、 $(\wedge \text{付 } s_i)$ は、元の s_i 信号 (1 1 7) の非常に近い近似値を表す再構築されたレイヤ信号 (1 3 2) を表す。図 1 B に示すように、受信レイヤ間においてレイヤ間予測は無い。しかし、ビデオ符号化の分野において公知であるように、容易にこのシステムは拡張することができ、受信した残差信号およびレイヤ間予測を用いて $(\wedge \text{付 } s_i)$ 信号を生成するデコーダとすることができる。

10

【 0 0 3 3 】

べき関数を用いた E D R 信号再構成

‘ 3 8 8 出願に記載のように、P Q 符号化された信号について、ある実施形態においては、信号再構成関数 (1 1 0) は以下のように表される。

【数 3】

$$s_i = \text{round}\left((c_H - c_L)\left(\frac{v_i - v_L}{v_H - v_L}\right)^\alpha + c_L\right) \quad (4)$$

20

ただし、 v_L および v_H は、入力 E D R 信号 (1 0 2) について検討中のカラーチャンネルにおける最小値および最大値を表し、 c_L および c_H は対応する最小および最大出力値を表す。例えば、ある実施形態においては、等式 (1) に定義されるように、 $c_L = 0$ および $c_H = p_L - 1$ である。 の値は一定であるが、フレーム毎に、シーン毎に、または、他の適切な基準に従って適応および変更してもよい。図 2 は、 $\alpha < 1$ についての等式 (4) のべき再構築関数の一例を示す。ある実施形態においては、入力 (1 1 2) が P Q 符号化されている場合 $\alpha > 1$ であり、またそうではなく、入力 (1 1 2) がガンマ符号化されている場合 $\alpha < 1$ である。

【 0 0 3 4 】

‘ 3 8 8 出願には、あるブロック内のピクセルの標準偏差またはそのブロック内の最小と最大ピクセル値の差分に基づいて、ブロック複雑性尺度 (metrics) (M_{EDR}) を用いて、 の最適値を求める方法を記載した。これらの尺度に加えて、あるブロックの空間的または周波数特性のいずれかに基づいて他の複雑性尺度を適用することも可能である。例えば、 M_{EDR} は、ブロック内でのピクセルの分散、ブロックの D C 値、または、そのピクセル値あるいは D C T 係数の別の関数に対応していてもよい。

30

【 0 0 3 5 】

それぞれ N ブロックに区分された (例えば、各ブロックは 8×8 ピクセル) F 枚のビデオフレームを含むビデオシーンについて考える。画像ブロックはオーバーラップしていてもよいが、最適な実施形態においてはオーバーラップしていない。図 3 は、一般化したブロック複雑性尺度 $M_{EDR}(j, n)$ (例えば、限定はしないが、ブロック内のピクセル値の標準偏差) に基づいて最適な を算出する処理を示す。

40

【 0 0 3 6 】

図 3 に示すように、そのシーン内の全フレームについてブロック毎に $M_{EDR}(j, n)$ を算出する工程 (3 0 5) の後、工程 (3 1 0) において、ある基準 (例えば、 $M_{EDR}(j, n) > T$ 但し、T は予め定められた閾値 (例えば 0) である) を満たす全画像ブロックを含む 1 つの集合 (set) S_j を構築する。

【 0 0 3 7 】

工程 (3 1 5)、(3 2 0) および (3 2 5) は、所定の範囲 (例えば、 M_{IN_j} M_{MAX_j}) 内の様々な j 値について算出されるループ (3 2 7) を構成する。例えば、 j は始めに 1 に等しくしておき、その後、元の E D R ビデオデータがどのように

50

符号化されているかによって増減させてもよい。例えば、PQ符号化されたデータの場合には増加し、ガンマ符号化されたデータの場合には低減してもよい。工程(320)においては、与えられた α_j および等式(4)を用いて、入力データEDRを量子化し、量子化された各ブロックについて新たな尺度 $M_{LD}(j, n, \alpha_j)$ を算出してもよい。いくつかの実施形態において、複雑性尺度 M_{LD} は複雑性尺度 M_{EDR} と同一である。他の実施形態においては、これら2つの複雑性尺度は異なってもよい。入力EDRデータを量子化する程、量子化信号(112)の特性はより変化する。理想的には、量子化(110)がなるべく入力を歪ませない方がよい。工程(325)においては、量子化による歪みの尺度を適用して、選択 α_j 値が最適かどうかを突き止めてもよい。例えば、ある実施形態において、

【数4】

$$|\Phi_j| - \sum_{n \in \Phi_j} (M_{LD}(j, n, \alpha_j) > T_\sigma) = 0$$

の場合に、 α_j を最適として選択してもよい。ただし、 T はまた別の所定の閾値(例えば、 $T = 0$)

【0038】

あるシーンの全ブロックを量子化した後、工程(330)において総合的に最適なアルファ値を選択する。例えば、ある実施形態において、 $\alpha_j > 1$ の場合、総合的に最適な α_j は、全最適 α_j 値に対して最小である α_j として選択する。同様に、 $\alpha_j < 1$ の場合、総合的に最適な α_j は、全最適 α_j 値に対して最大である α_j として選択する。

【0039】

いくつかの実施形態において、ビデオエンコーダ(120)による損失の多い圧縮に起因する量子化の影響に対して調節するために、総合的に最適なアルファをさらに調節してもよい(例えば、 $\alpha_j = +$ 、但し、 $\alpha_j > 1$ の場合は α_j は負の値であり、 $\alpha_j < 1$ の場合は α_j は正の値である)。同パラメータ最適化処理は、複数の関数パラメータで特徴付けられる、他の線形または非線形量子化関数および再構築関数に、容易に拡張することができる。

【0040】

表1は、図3に示す処理(所望の $\alpha_j < 1$ である)に基づくある実施形態による、ガンマ符号化値を再構成するためのアルゴリズムの一例を疑似コードで示したものである。

表1：ガンマ符号化EDR信号について最適な α_j を算出する方法

10

20

30

【表 1】

```

//Initialize for each frame j within a scene of F total frames
 $\alpha^{opt}[j] = 1.0$  for all j.

// for each frame
for(j = 0; j < F; j++){
  // STEP 1.a : calculate complexity measurement for each block n in EDR domain
  Calculate  $M_{EDR}(j, n)$  for all values of n
  // STEP 1.b : Identify those blocks that satisfy a complexity criterion
  Construct  $\Phi_j = \{i \mid M_{EDR}(j, n) > T\}$ 

  // STEP 2: Apply the reshaping function and compute updated complexity metrics for
  different  $\alpha$ 
  for(  $\alpha_j = 1.0$ ;  $\alpha_j > MIN\_a$ ;  $\alpha_j = \alpha_j - 0.1$  ) {
    // STEP 2a: quantize EDR using power function with  $\alpha_j$ 
     $s_{j,i} = round((c_H - c_L) \left( \frac{v_{j,i} - v_L}{v_H - v_L} \right)^{\alpha_j} + c_L)$ 

    // STEP 2b: calculate complexity measurement for each block after layer
    decomposition
    Calculate  $M_{LD}(j, n, \alpha_j)$  for all n

    // STEP 2c: calculate whether given  $\alpha_j$  causes near-zero deviation
    if(  $|\Phi_j| - \sum_{n \in \Phi_j} (M_{LD}(j, n, \alpha_j) > T_\sigma) = 0$  )
       $\alpha^{opt}[j] = \alpha_j$ 
    else
      break;
  }
}

// STEP 3: find the maximal  $\alpha^{opt}[j]$  as the final  $\alpha$ 
 $\alpha = \max\{ \alpha^{opt}[j] \}$ 

//Apply optional safety margin

 $\alpha = \alpha + \Delta\alpha,$ 

```

【 0 0 4 1 】

デコーダにおいて、逆構成処理 (1 4 0) は以下のように表すことができる。

【数 5】

$$\hat{v}_i = (v_H - v_L) \left(\frac{\hat{s}_i - c_L}{c_H - c_L} \right)^{1/\alpha} + v_L \quad (5)$$

40

【 0 0 4 2 】

いくつかの実施形態において、等式 (4) のべき関数は、区分的線形多項式、区分的 2 次または 3 次多項式または区分的 B スプラインとして表すことができる。そのような実施の場合、量子化に関連するアーチファクトを避けるために、全区分にわたって、滑らかさおよび単調性に関する制約を与えるべきである。そして、これまでに、または、' 3 8 8 出願において記載されたものと同様の検索方法を適用してもよい。

【 0 0 4 3 】

ブロック適応的な再構成関数

再び、それぞれ N ブロックに区分された (例えば、各ブロックは 8 × 8 ピクセル) F 枚

50

のビデオフレームを含むビデオシーンについて考える。画像ブロックはオーバーラップしていてもよいが、好適な実施形態においてはオーバーラップしていない。図4は、ある実施形態による、入力EDR符号語(102)を再構成済み出力値(112)にマッピングするためのデータフローの一例を示す。

【0044】

先述のように、ブロックベースの複雑性尺度(M_{EDR})を定義できる。ある実施形態において、工程(405)において、限定はしないが、複雑性尺度がブロック内のピクセルの標準偏差(std)に基づいて算出されるものとして考える。なお、フレーム j ($j = 1, 2, \dots, F$)におけるブロック n の標準偏差がゼロである(例えば、 $M_{EDR}(j, n) = std_{jn} = 0$)かどうかを確認することは、このブロック内の最大値(例えば、 $B(j, n)$)とこのブロック内の最小値(例えば、 $A(j, n)$)との差異が0であるかどうかを確認することに等しい。

10

【0045】

再構成関数(110)が区分的線形直線で構築されているとすると、入力 v_i に対して、ローカル量子化器は以下のように表すことができる。

【数6】

$$v_i \in [A(j, n) \ B(j, n)]$$

【数7】

$$s_i = \text{round}(k(j, n)(c_H - c_L) \left(\frac{v_i - v_L}{v_H - v_L} \right) + c_L) \quad (6)$$

20

ただし、 $k(j, n)$ は、 j 番目のフレームの n 番目のブロックにおける量子化器の傾きを調節するためのスケーリング係数である。

【0046】

工程(410)において、 Φ が、ブロック尺度がある基準を満たす全ブロックの集合を表すものとする。例えば、 Φ が、量子化前に非ゼロの標準偏差を有する全ブロックの集合を表すものとする。すなわち、

30

【数8】

$$\Phi = \{(j, n) \mid B(j, n) - A(j, n) > 0\} \quad (7)$$

【0047】

ある実施形態において、閾値 T_{th} (ただし、限定はしないが、 $T_{th} \in [0, 1]$)が与えられ、ブロックの最小および最大ピクセル値が与えられると、工程(415)において、最適な $k(j, n)$ は以下のように求められ得る。

【数9】

$$k(j, n) = T_{th} \frac{v_H - v_L}{(c_H - c_L)(B(j, n) - A(j, n))} \quad (8)$$

40

【0048】

データ $\{A(j, n), B(j, n), k(j, n)\}$ が与えられるとき、この3つの値は、区分 $[A(j, n), B(j, n)]$ 内で、量子化器は、少なくとも $k(j, n)$ の傾きを有するべきであることを示す。特定のEDR符号(例えば、 v_c)が複数 $[A(j, n), B(j, n)]$ 区分に属し得るとするならば、EDR符号語 v_c について、各符号語について全ブロックを満足する最大の傾きを求める必要がある。

【0049】

50

(v_c) は、符号語 v_c をカバーする全ブロックにわたる全区分の集合を表すものとする。すなわち、

【数 1 0】

$$\theta(v_c) = \{(j, n) \mid A(j, n) \leq v_c \leq B(j, n), (j, n) \in \Phi\} \quad (9)$$

すると、工程 (4 2 0) において、符号語 v_c において必要な傾きは、(v_c) 集合に属するブロック内の全ての最適な傾きの包絡線として求められ得る。すなわち、

【数 1 1】

$$k(v_c) = \max_{(j, n) \in \theta(v_c)} k(j, n) \quad (10)$$

10

【0 0 5 0】

工程 (4 2 5) において、これらの包絡線の傾き全ての合計を以下のように表記するものとする、

【数 1 2】

$$k = \sum_{v_c} k(v_c) \quad (11)$$

各 v_c 符号語について、限定はしないが、工程 (4 3 0) において、累積傾斜関数は以下のように定義されてもよい。

20

【数 1 3】

$$K(v_c) = \sum_{i=0}^{v_c} k(i) \quad (12)$$

全符号語が範囲 [c_L, c_H] 内にマッピングされることを担保するためには、 v_c 符号語から s_i 値へのマッピングは、以下を用いて算出してもよい。

【数 1 4】

$$\bar{K}(v_c) = \text{round}\left(\frac{c_H - c_L}{k} K(v_c)\right) \quad (13)$$

30

等式 (13) が与えられたとき、入力 v_c 値と出力 s_i 値の間のマッピングは、順方向マッピングルックアップテーブル (例えば、 $LUT_F(v_c)$) を用いて算出してもよい。ある実施形態においては、このテーブルは、データと一緒に保存してもよいし、画像メタデータ (119) の一部としてデコーダに送信し、デコーダが逆方向マッピング処理を再構築し得るようにしてもよい。

【数 1 5】

$$LUT_F(v_c) = \bar{K}(v_c)$$

40

【0 0 5 1】

ある実施形態例において、表 2 は、疑似コードを用いて、図 4 に示すマッピング処理をまとめたものである。

【表 2】

```

// STEP 1: find max and min value
for(j = 0; j < F; j++){
  for(n = 0; n < N; n++){
    // STEP 1.a : find the maximal value for each block n in EDR domain
    Calculate B(j, n)
    // STEP 1.b : find the minimal value for each block n in EDR domain
    Calculate A(j, n)
  }
  end
end

```

10

```

// STEP 2 : find block with (B(j, n) - A(j, n)) > 0; i.e. non-zero std block index
Construct Φ = {(j, n) | B(j, n) - A(j, n) > 0}
for each block in Φ
  // STEP 2.a: calculate optimal slope
  
$$k(j, n) = T_{th} \frac{v_H - v_L}{(c_H - c_L)(B(j, n) - A(j, n))}$$

end

```

20

```

// STEP 3: find optimal slope
for each valid EDR code word v_c
  
$$\theta(v_c) = \{(j, n) | A(j, n) \leq v_c \leq B(j, n), (j, n) \in \Phi\}$$

  
$$k(v_c) = \max_{(j, n) \in \theta(v_c)} k(j, n)$$

end

```

30

```

// STEP 4: find optimal EDR LUT by accumulating the slope and taking normalization

$$K(v_c) = \sum_{i=0}^{v_c} k(i) \text{ for all } v_c, \text{ then}$$


$$\bar{K}(v_c) = \text{round}\left(\frac{c_H - c_L}{k} K(v_c)\right).$$


```

【0052】

いくつかの実施形態において、等式(12)では、代替的な関数を用いて累積傾斜関数 $k(v_c)$ を算出してもよい。例えば、 $k(v_c)$ 値は、合算する前に、以下のようにフィルタリングまたは重み付けされてもよい。

【数16】

$$k(v_c) = \sum_{i=-u}^u w_i k(v_c + i) \quad (14)$$

40

ただし、 w_i は、所定の重み、または、 $(2u + 1)$ フィルタタップ（例えば、 $u = 2$ であって、フィルタ係数はローパスフィルタのそれと対応する）を有するフィルタのフィルタ係数を表す。

【0053】

図5は、 $[c_L, c_H]$ 範囲を $[0, 255]$ とした場合の、フレーム群のテストシーケンスについての、 v_c 符号語に対する包絡線傾き $k(v_c)$ 値のプロットの一例を示す。

【0054】

図5に示すデータに対して、図6は、 v_c に対する $K(v_c)$ マッピングの一例を示す。

【0055】

50

逆再構成

デコーダにおいて、等式(13)の付K(v_c)値が与えられると、逆量子化器または再構成関数(140)は以下のようにして求められ得る。

各復号化符号語(\hat{s}_c)に対して、

【数17】

$$\omega(\hat{s}_c) = \{i \mid \bar{K}(v_i) = \hat{s}_c\} \quad (15)$$

とすると、

【数18】

$$\bar{K}^{-1}(\hat{s}_c) = \frac{\sum_{i \in \omega(\hat{s}_c)} v_i}{|\omega(\hat{s}_c)|} \quad (16)$$

となる。

【0056】

換言すれば、量子化領域(\hat{s}_c)における所与の符号語について、対応する推定EDR符号語(\hat{v}_c)は、まず、量子化値(\hat{s}_c)を有する全ピクセルをグループ分けし、対応するEDR符号語を見つけ、そして、集めた全EDR符号語の平均をとることによって構築される。等式(16)から、逆ルックアップテーブルLUT_B(\hat{s}_c)を構築して、データと一緒に保存してもよいし、または例えばメタデータ(119)の一部としてこれをデコーダに送信してもよい。

【数19】

$$\text{LUT}_B(\hat{s}_c) = \bar{K}^{-1}(\hat{s}_c)$$

【0057】

図6に示すデータに対して、図7は、

【数20】

$$\hat{s}_c \text{ に対する } \bar{K}^{-1}(\hat{s}_c)$$

マッピングまたは逆再構成(140)の一例を示す。

【0058】

ある実施形態において、

【数21】

$$\bar{K}^{-1}(\hat{s}_c)$$

という関係によって定義されるマッピングは、メタデータ(119、127)を用いてデコーダに送信してもよい。このような手法は、ビットレートオーバーヘッドの観点から高コスト過ぎるかもしれない。例えば、8ビットデータの場合、ルックアップテーブルは255個のエントリを含んでおり、これがシーンチェンジのたびに送信されてもよい。他の実施形態においては、逆方向マッピングを区分的多項式近似に置き換えてもよい。このような多項式は典型的には1次および2次多項式を含むが、高次多項式またはBスプラインを用いることも可能である。あるレイヤ l ($l = 0, 1, \dots, L-1$)に対する1つのLUTを近似する多項式の数、利用可能な帯域幅および処理の複雑度によって変わり得る。ある実施形態において、エンハンスメントレイヤが単一の区分を使用するのに対して、ベースレイヤは最大8区分を使用する。

10

20

30

40

50

【 0 0 5 9 】

表 3 は、ある実施形態において、2 次多項式を用いてデコーダ L U T テーブルを近似するアルゴリズムの一例を示す。

表 3 : 2 次多項式を用いた L U T 近似

【 表 3 - 1 】

```

// Input :
//   LUTB : Inverse reshaping lookup table
//   ε      : maximal error tolerance between LUT values and approximated values
//   Nl    : number of code words in layer l
//   pl    : partition code word separating layer l-1 and l
// Output:
//   mx(t) : pivot points for input
//   my(t) : corresponding output value
//   a(i,j) : Polynomial coefficients

T = 0; // counter to count number of pivot points
Len_LUT = Nl; // number of entries in the lth layer
LUT[ x ] = LUTB[ x + pl ]; // segment of backward resaper used in lth layer
mx(0) = 0; // first pivot point ; input value
my(0) = LUTB[pl]; // first pivot point ; output value
converge = 0; // flag to check convergence

while( ! converge ){
  found_one_seg = 0;
  cnt = 1;
  // find one segment which won't exceed the maximal tolerance error
  while( ~found_one_seg ){
    // STEP 1 : move pivot candidate to next feasible point
    m = mx(T) + cnt;

    // STEP 2: approximate value from LUT[mx(T)] to LUT[m]
    // by a 2nd order polynomial
    f( T , i ) = a( T , 2 ) * i2 + a( T , 1 ) * i + a( T , 0 )

    where a( T , 2 ), a( T , 1 ), and a( T , 0 ) are the fitted polynomial coefficients

    // Compute the maximum approximation error
    for( i = mx(T) ; i <= m ; i++){
      εi = f( T , i ) - LUT( i )
    }
    εmax = max{ εi , i = mx(T) , ..., m }

    // STEP 3.(a) if max error exceeds tolerance, take the previous one
    if( ( εmax > ε ) && ( T < Len_LUT ) ){
      found_one_seg = 1;
      T++;

```

【表 3 - 2】

```

        mx(T) = m - 1;
        my(T) = LUT(mx(T));
    }
    // STEP 3.(b) reach the last entry
    else if( ( εmax <= ε ) &&( m == Len_LUT ) ){
        found_one_seg = 1;
        T++;
        mx(T) = m;
        my(T) = LUT(mx(T));
        converge = 1
    }

    // STEP 4: increase the counter
    cnt ++;
}
}

```

【 0 0 6 0 】

表 3 に示すように、ある実施形態において、近似処理の入力は、元のルックアップテーブル（例えば、等式（16）を用いて算出）と、LUT中の値および多項式近似により生成される値の間の許容可能な誤差範囲と、利用可能な符号語数と、その最初の符号語値（等式（1）参照）とを含む。出力は、各多項式のエンドポイント（ピボットポイントとも呼ぶ）と、多項式係数とを含み得る。

【 0 0 6 1 】

最初のピボットポイントからスタートして、アルゴリズムは、限定はしないが2次多項式を用いて、利用可能な符号語の最大範囲のフィッティングを試みる。平均二乗誤差多項式フィット等のような、任意の既知の多項式フィットアルゴリズムを用いてよい。

【 0 0 6 2 】

算出最大誤差が入力許容範囲を超える場合、最適な多項式のパラメータを保存して、新たな多項式の検索を開始し、これをLUTテーブル全体がマッピングされるまで行う。

【 0 0 6 3 】

いくつかの実施形態において、1つのLUTを近似するのに使用できる多項式数が固定値（例えば8）に制限される場合がある。その場合、より大きな誤差許容範囲をアルゴリズムに組み込んでよい。

【 0 0 6 4 】

表3の方法はまた、より高次の多項式、Bスプラインまたは近似関数の組み合わせなどの他の近似関数に対応するようにも容易に改変可能である。

【 0 0 6 5 】

知覚的に均質な色空間における符号化

ビデオ信号は典型的にはよく知られたRGB色空間に描画されるが、MPEG-2、H.264（AVC）、H.265（HEVC）等のようなほとんどのビデオ圧縮規格は、これと対立するYCbCrまたはYUVのような色空間で使用するように最適化されている。これらの色空間は、8～10ビットの標準ダイナミックレンジ（SDR）ビデオを符号化および送信するには十分であるが、ビット/ピクセル要件の観点からは、EDRビデオを符号化および送信する際に最も効率が良いとはいえない。例えば、Lu'v'およびLog(L)u'v'色空間がこれまでに提唱されている。

【 0 0 6 6 】

発明者らの理解によれば、知覚的に均質な色空間における信号の符号化は、ビデオコーデックによる処理の前のu'およびv'色度データを追加的に処理することによって利点を有し得る。例えば、ある実施形態において、エンコーダにおいて、このような処理は、

信号再構築処理 (1 1 0) の一部として入力信号 (1 0 2) 上で行われ得る。

【 0 0 6 7 】

白色点変換

ある実施形態において、線形 X Y Z からルマ u ' v ' 色空間への変換は、以下の工程を含み得る。

- a) 白色点の座標 (例えば、D 6 5) を定義し、
- b) $L u m a = f (Y)$ について解を求め、
- c) X、Y および Z から u ' および v ' について解を求める。

【 0 0 6 8 】

本明細書において、関数 $f (Y)$ は、L (または L ')、 $L o g (L)$ 等のような明度に関連する関数を表している。好適な実施形態において、 $f (Y)$ は、' 2 1 2 出願に記載したような知覚的量子化 (P Q) マッピング関数を表していてもよい。

10

【 0 0 6 9 】

ある実施形態においては、白色点は、u ' および v ' 座標を用いて D 6 5 (6 5 0 0 K) と定義されてもよい。

$$D u = d 6 5 u = 0 . 1 9 7 8 3 0 0 0 6 6 4 2 8 ;$$

$$D v = d 6 5 v = 0 . 4 6 8 3 1 9 9 9 4 9 3 8 8 ;$$

【 0 0 7 0 】

ある実施形態において、u ' および v ' は以下のように求めることができる。

【 数 2 2 】

20

$$u' = \frac{4X}{X+15Y+3Z} \quad (17a)$$

$$v' = \frac{9Y}{X+15Y+3Z} \quad (17b)$$

$$\text{if } (X+15Y+3Z) \neq 0, \quad \text{and}$$

$$u' = D u \quad (17c)$$

$$v' = D v \quad (17d)$$

30

$$\text{if } (X+15Y+3Z) = 0$$

【 0 0 7 1 】

逆演算処理は以下を含む。

- a) 白色点の座標 (例えば、D 6 5) を定義し、
- b) $Y = f^{-1} (L u m a)$ について解を求め、
- c) u ' および v ' から X および Z について解を求める。

【 0 0 7 2 】

例えば、' 2 1 2 出願による知覚的量子化関数を用いる実施形態においては、対応する逆 P Q マッピングを適用して Y ピクセル値を生成することが可能である。

40

【 0 0 7 3 】

ある実施形態において、X および Z は以下のように求めることができる。

【数 2 3】

$$X = Y \frac{9u'}{4v'} , \quad (18a)$$

$$Z = Y \frac{12-3u'-20v'}{4v'} , \quad (18b)$$

if $v' \neq 0$, and

$$X = Z = Y, \text{ if } v' = 0. \quad (18c)$$

10

【0 0 7 4】

図 8 A は、 $u' v'$ 色度空間における白色点 (8 0 5) (例えば、D 6 5) の従来のマッピングを示している。図 8 A に示すように、 u' および v' 色度値の範囲は、概ね、 u' については (0、0.623) であり、 v' については (0、0.587) である。図 8 A に示すように、D 6 5 白色点は、 $u' v'$ 信号表記では中心ではない。これにより、色差成分をサブサンプリングおよびアップサンプリングして、元の 4 : 4 : 4 色フォーマットとビデオ符号化において典型的に用いられる 4 : 2 : 0 または 4 : 2 : 2 色フォーマットとの間の変換を行った後に、色ズレが生じ得る。このような色ズレを軽減するために、色度値に変換関数を適用することが推奨される。ある実施形態において、この変換関数によって、白色点を、概ね、変換後の $u' v'$ の中心にずらす、白色点は、クロマ誤差を引きつけたい任意の他の色値に置き換えられてもよい。例えば、変換後の値 u'_t および v'_t が (0、1) の範囲内にある場合、以下のマッピングを適用してもよい。

20

【数 2 4】

$$u'_t = (u' - Du)a_1 + b_1 , \quad (19a)$$

$$v'_t = (v' - Dv)a_2 + b_2 , \quad (19b)$$

ただし、 Du および Dv は選択された白色点の元の u' および v' 座標を表し、 (b_1, b_2) は変換後の色空間における白色点の所望位置の座標を決定し、 a_i ($i = 1, 2$) は、所望の変換点および u' および v' の最小および最大値に基づいて算出される定数である。ある実施形態において、変換パラメータ (例えば、 a_1 および a_2) は、全体のビデオシーケンスについて固定であってもよい。別の実施形態においては、入力コンテンツの色度範囲が変化することを利用するために、シーン毎またはフレーム毎に変換パラメータを算出してもよい。

30

【0 0 7 5】

図 8 B は、ある実施形態例による変換後の色度空間における白色点 (8 0 5) のマッピングを示す。図 8 B において、元の u' および v' 色度値は、選択点 (例えば、D 6 5) が概ね変換後の色度空間の中心 (0.5、0.5) になるように置き換えられる。例えば、変換後の値 u'_t および v'_t が (0、1) 内であれば、 $b_1 = b_2 = 0.5$ について、ある実施形態において、以下のマッピングを適用してもよい。

40

【数 2 5】

$$u'_t = (u' - Du)1.175 + 0.5 , \quad (20a)$$

$$v'_t = (v' - Dv)1.105 + 0.5 , \quad (20b)$$

ただし、 Du および Dv は、選択された白色点の u' および v' 座標を表す。この変換の結果、色誤差は、色相ずれではなく、脱飽和に見えるであろう。 u' および v' 色度値に非線形関数を適用して、これと同じ変換を実現することも可能であることは当業者であれば理解できるであろう。このような非線形関数は、符号化および量子化による色誤差の可視性をさらに低減するために、ほぼ中間色である色に対してより高い精度を割り当てて

50

もよい。

【0076】

クロマエントロピーの低減

u' および v' ピクセル成分に光度の関数を掛ければ、クロマの詳細の可視性をより改善できるかもしれない。例えば、ある実施形態において、変換された色度値は以下のように求めてもよい。

【数26】

$$u'_t = g(Luma)(u' - Du)a_1 + b_1, \quad (21a)$$

$$v'_t = g(Luma)(v' - Dv)a_2 + b_2, \quad (21b)$$

10

ただし、 $g(Luma)$ はルマチャネルの関数を表している。ある実施形態において、 $g(Luma) = Luma$ である。

【0077】

デコーダにおいて、入力信号は $Luma u'_t v'_t$ と表記され得る。多くの用途において、この信号は、さらなる処理の前に、変換して XYZ 、 RGB または他の色空間に戻さなければならない。ある実施形態例において、 $Luma u'_t v'_t$ から XYZ への変換処理は、以下の工程を含み得る。

a) $Luma$ 符号化を元に戻し、

$$Y = f^{-1}(Luma)$$

20

b) u'_t および v'_t の値をスケーリングする範囲を元に戻して、 u' および v' に戻し、

c) 等式(18)を用いて X および Z を復元する。

【0078】

いくつかの実施形態において、入力信号の $Luma$ 、 u'_t および v'_t 成分は、何らかの色変換を行う前に、(0, 1) 範囲に正規化してもよい。いくつかの実施形態において、ルックアップテーブル、乗算、加算を組み合わせることによって等式(17)~(21)を実行してもよい。例えば、ある実施形態において、

$$Y = f^{-1}(Luma)$$

$$B = 3u',$$

$$C = 20v', \text{ および}$$

$$D = 1 / (4v')$$

30

が、 $Luma$ 、 u' および v' を入力とする3つのルックアップテーブルの出力を表しているとする、等式(18)から、 X および Z の値は、以下のように4つの乗算と2つの加算を用いて算出され得る。

$$Z = (Y * D) * (12 - B - C)$$

および

$$X = (Y * D) * (3 * B)$$

【0079】

例えば、ある実施形態において、10ビット符号化信号について、各LUTは、それぞれターゲットアプリケーションにとって十分な精度(例えば、32ビット)を持つ1024個のエントリを含んでいてもよい。

40

【0080】

図9は、本発明のある実施形態による、符号化および復号化パイプラインの一例を示す。入力信号(902)は、 $RGB4:4:4$ または他の適切な色フォーマットであり得る。工程(910)において、例えば、輝度値に対し知覚的量子化(PQ)マッピング、および色度値に対し等式(17)を用いて、信号(902)を知覚的ルマ $u' v'$ フォーマットに置き換えられる。工程(915)において、白色点が概ね変換された色度空間の中心に位置するように、例えば、等式(19)~(21)に示す変換を適用して、元の $u' v'$ 色度値を変換された色度値 u'_t および v'_t に置き換える。この色変換および変換

50

後 $Luma$ $u'v'$ $4:4:4$ 信号 (例えば、917) は、ビデオエンコーダ (920) による符号化の前に、 $4:2:0$ または $4:2:2$ フォーマットに色サブサンプリングされてもよい (図示せず)。ビデオエンコーダ (920) は、上述のように、信号再構成 (110) およびレイヤ分解 (115) 処理を含んでいてもよい。受信器上では、ビデオデコーダ (930) が復号化信号 (932) を生成する。ビデオデコーダ (930) は、信号逆再構成およびレイヤ結合器 (140) を備えていてもよい。任意の色アップサンプリング (例えば、 $4:2:0$ から $4:4:4$) の後に、(915) の変換演算処理を逆転することによって、逆色度変換工程 (935) が $Luma$ $u'v'$ 信号 (932) を変換して $Luma$ $u'v'$ 信号 (937) に戻してもよい。最後に、 $Luma$ $u'v'$ 信号 (937) を変換して、RGB または他の適切な色空間の出力信号 (942) とし、これを表示またはさらに処理してもよい。

10

【0081】

知覚的量子化された IPT 色空間における符号化

白色点変換は、IPT 色空間または IPT-PQ 色空間 (エンハンストまたはハイダイナミックレンジを有するビデオ信号の符号化に理想的に適していると思われる知覚的量子化された色空間) のような他の色空間にも適用してもよい。IPT-PQ 色空間を最初に記載したのは、本願にその全文が援用される、R. Atkins らによる 2014 年 2 月 13 日出願の「Display management for high dynamic range video」という表題の PCT 出願 PCT/US2014/016304 である。

20

【0082】

本願にその全文が援用される、F. Ebner および M. D. Fairchild による「Development and testing of a color space (ipt) with improved hue uniformity」、Proc. 6th Color Imaging Conference: Color Science, Systems, and Applications, IS&T, Scottsdale, Arizona, 1998 年 11 月、pp. 8-13 (「Ebner 論文」と呼ぶ) に記載の IPT 色空間は、人間の視覚システムにおける錐体間の色差のモデルである。その意味では YCbCr または CIE-Lab 色空間のようであるが、いくつかの科学研究において、これらの空間よりも人間の視覚処理をより良く模擬することが分かっている。CIE-Lab と同様に、IPT は、何らかの参照輝度に対する正規化された空間である。ある実施形態において、正規化は、ターゲットディスプレイの最大輝度に基づいていてもよい。

30

【0083】

本明細書において、用語「PQ」は知覚的量子化を指す。人間の視覚システムは、光レベルの増大に対して非常に非線形的に反応する。人間が刺激を見る能力は、その刺激の輝度、その刺激の大きさ、その刺激を構成する空間周波数、および、その刺激を見ている瞬間までに目が適応した輝度レベルに影響される。好適な実施形態において、知覚的量子化器関数は、線形入力グレイレベルを、人間の視覚システムにおけるコントラスト感度閾値によりマッチした出力グレイレベルにマッピングする。PQ マッピング関数の一例は、

40

【0084】

表 1 は、表示時点においてデジタルビデオ符号値を絶対線形輝度レベルに変換するため

50

の知覚的曲線 E O T F の計算を示している。絶対線形輝度をデジタル符号値に変換するための逆 O E T F の計算も含まれている。

表 1

例示的な等式定義：

D = 知覚的曲線デジタル符号値、S D I - l e g a l 符号無し整数、10 または 12 ビット

b = デジタル信号表記における成分毎のビット数、10 または 12

V = 正規化された知覚的曲線信号値、0 ≤ V ≤ 1

Y = 正規化された輝度値、0 ≤ Y ≤ 1

L = 絶対輝度値、0 ≤ L ≤ 10,000 c d / m²

10

例示的な E O T F 復号化等式：

【数 2 7】

$$V = \frac{D - 4 \cdot 2^{b-10}}{1015 \cdot 2^{b-10}}$$

$$Y = \left(\frac{\max[(V^{1/m} - c_1), 0]}{c_2 - c_3 V^{1/m}} \right)^{1/n}$$

(t1)

$$L = 10,000 \cdot Y$$

20

例示的な O E T F 符号化等式：

【数 2 8】

$$Y = \frac{L}{10,000}$$

$$V = \left(\frac{c_1 + c_2 Y^n}{1 + c_3 Y^n} \right)^m$$

(t2)

$$D = \text{INT}(1015 \cdot V \cdot 2^{b-10}) + 4 \cdot 2^{b-10}$$

(t3)

30

例示的な定数：

【数 2 9】

$$n = \frac{2610}{4096} \times \frac{1}{4} \approx 0.15930176$$

$$m = \frac{2523}{4096} \times 128 = 78.84375$$

$$c_1 = c_3 - c_2 + 1 = \frac{3424}{4096} = 0.8359375$$

40

$$c_2 = \frac{2413}{4096} \times 32 = 18.8515625$$

$$c_3 = \frac{2392}{4096} \times 32 = 18.6875$$

備考：

1. 演算子 I N T は、0 ~ 0.4999... の範囲の小数部に対しては値 0 を返し、

50

0.5 ~ 0.9999... の範囲の小数部に対しては + 1 を返す（即ち、0.5 より大きい小数部は切り上げ）。

2. 定数は全て四捨五入の問題を回避するために、12ビット有理数のちょうどの倍数として定義する。

3. R、GまたはB信号成分は、上記のY信号成分と同じ方法で算出する。

【0085】

IPT - PQ色空間への信号の変換は、以下の工程を含み得る。

a) 入力色空間（例えば、RGBまたはYCbCr）からXYZに信号を変換する。

b) 以下のように、信号をXYZからIPT - PQに変換する。

a. 3 x 3 XYZをLMS行列に適用して、信号をXYZからLMSに変換する。

b. LMS信号の各色成分を知覚的量子化されたLMS信号（L' M' S'またはLMS - PQ）に変換する（例えば、等式（t2）を適用して）。

c. 3 x 3 LMSをIPT行列に適用して、LMS - PQ信号をIPT - PQ色空間に変換する。

【0086】

3 x 3 XYZからLMS、および、L' M' S'（またはLMS - PQ）からIPTの変換行列の例は、Ebner論文に見出すことができる。IPT - PQ信号の色成分（例えば、P'およびT'）が範囲（-0.5、0.5）内にあるとすると、色成分の範囲が実質的に範囲（0、1）内となるように、バイアスa（例えば、a = 0.5）を加えてもよい。例えば、

【数30】

$$P' = P' + a \quad (22a)$$

$$T' = T' + a \quad (22b)$$

【0087】

逆色演算処理は、以下の工程を含み得る。

a) 色成分に加えられた全てのバイアス値を引いておき、

b) 3 x 3 I' P' T'をLMS変換行列に適用して、IPT - PQからLMS - PQに変換し、

c) 逆PQ関数を適用して、LMS - PQからLMSに変換し（例えば、等式（t1）を使用）、

d) 3 x 3 LMSをXYZ変換に適用して、LMSからXYZに変換し、

e) XYZから、選択した装置依存の色空間（例えば、RGBまたはYCbCr）に変換する。

実際には、符号化および/または復号化の際の色変換工程は、予め算出した1 - Dルックアップテーブル（LUTs）を用いて行われてもよい。

【0088】

クロマエントロピーの低減

上述のように、P'およびT'ピクセル成分に光度の関数（例えば、I'）を掛ければ、クロマ的詳細の可視性はさらに改善されるかもしれない。例えば、ある実施形態において、変換された色度値は以下のように求めることができる。

【数31】

$$P'_t = g(I')(P' - a) + a, \quad (23a)$$

$$T'_t = g(I')(T' - a) + a, \quad (23b)$$

ただし、g（I'）は、ルマチャネル（I'）の線形または非線形関数を表している。

ある実施形態において、 $g(I') = I'$ である。

【0089】

コンピュータシステム実装例

本発明の実施形態は、コンピュータシステム、電子回路およびコンポーネントで構成されたシステム、マイクロコントローラ、フィールドプログラマブルゲートアレイ (FPGA) または他のコンフィギュラブルまたはプログラマブルロジックデバイス (PLD)、離散時間またはデジタル信号プロセッサ (DSP)、特定用途向け IC (ASIC) などの集積回路 (IC) デバイス、および/または、このようなシステム、デバイスまたはコンポーネントを1つ以上含む装置、を用いて実施し得る。このコンピュータおよび/または IC は、本明細書に記載のようなエンハストダイナミックレンジ (EDR) を有するビデオ画像の階層符号化のための適応的再構成技術に関する命令を行い、制御し、または実行し得る。このコンピュータおよび/または IC は、本明細書に記載の適応的再構成プロセスに関する様々なパラメータまたは値のいずれを演算してもよい。画像およびビデオ実施形態は、ハードウェア、ソフトウェア、ファームウェア、および、その様々な組み合わせで実施され得る。

10

【0090】

本発明の特定の態様は、本発明の方法をプロセッサに行わせるためのソフトウェア命令を実行するコンピュータプロセッサを含む。例えば、ディスプレイ、エンコーダ、セットトップボックス、トランスコーダなどの中の1つ以上のプロセッサは、そのプロセッサがアクセス可能なプログラムメモリ内にあるソフトウェア命令を実行することによって、上記のようなエンハストダイナミックレンジ (EDR) を有するビデオ画像の階層符号化のための適応的再構成技術に関する方法を実装し得る。本発明は、プログラム製品形態で提供されてもよい。このプログラム製品は、データプロセッサによって実行された時に本発明の方法をデータプロセッサに実行させるための命令を含む1セットの、コンピュータが可読な信号を格納する任意の媒体を含み得る。本発明によるプログラム製品は、様々な形態をとり得る。例えば、このプログラム製品は、フロッピーディスク、ハードディスクドライブを含む磁気データ記憶媒体、CD ROM、DVDを含む光学データ記憶媒体、ROM、フラッシュRAMなどを含む電子データ記憶媒体、などの物理的媒体を含み得る。このプログラム製品上のコンピュータ可読信号は、任意に、圧縮または暗号化されていてもよい。

20

30

【0091】

上記においてあるコンポーネント (例えば、ソフトウェアモジュール、プロセッサ、アセンブリ、デバイス、回路など) に言及している場合、そのコンポーネントへの言及 (「手段」への言及を含む) は、そうでないと明記されている場合を除いて、当該コンポーネントの機能を果たす (例えば、機能的に均等である) あらゆるコンポーネント (上記した本発明の実施形態例に出てくる機能を果たす開示構造に対して構造的に均等ではないコンポーネントも含む) を、当該コンポーネントの均等物として、含むものと解釈されるべきである。

【0092】

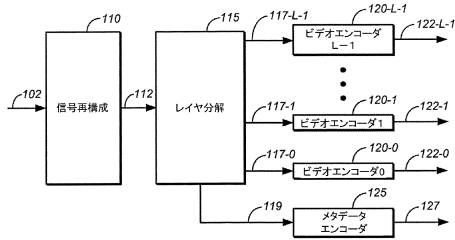
均等物、拡張物、代替物、その他

40

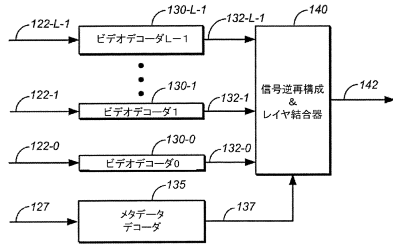
エンハストダイナミックレンジ (EDR) を有するビデオ画像の階層符号化のための適応的再構成技術に関する実施形態例を上述した。この明細書中において、実装毎に異なり得る多数の詳細事項に言及しながら本発明の実施形態を説明した。従って、本発明が何たるか、また、本出願人が本発明であると意図するものを示す唯一且つ排他的な指標は、本願が特許になった際の請求の範囲 (今後出されるあらゆる訂正を含む、特許となった特定請求項) である。当該請求項に含まれる用語に対して本明細書中に明示したあらゆる定義が、請求項内で使用される当該用語の意味を決定するものとする。よって、請求項において明示されていない限定事項、要素、性質、特徴、利点または属性は、その請求項の範囲をいかなる意味においても限定すべきではない。従って、本明細書および図面は、限定的ではなく、例示的であるとみなされるものである。

50

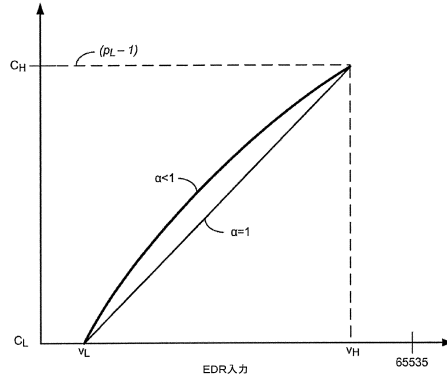
【図1A】



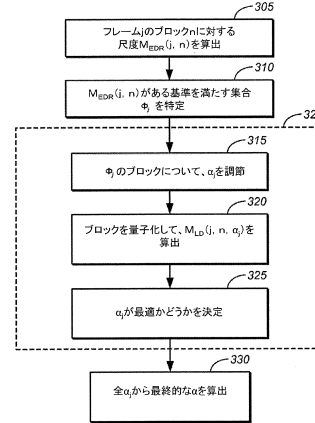
【図1B】



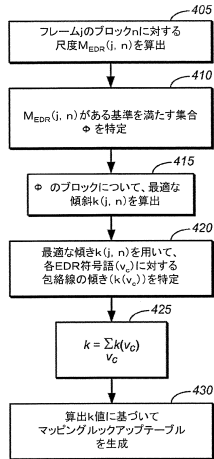
【図2】



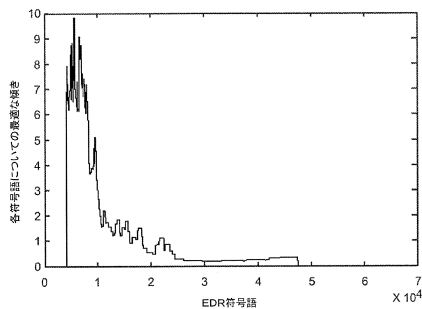
【図3】



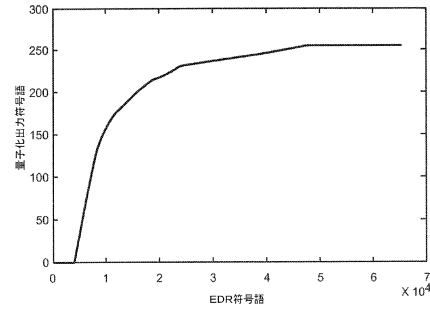
【図4】



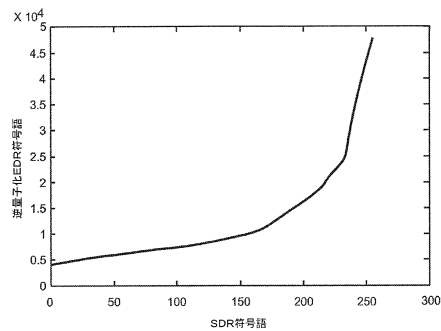
【図5】



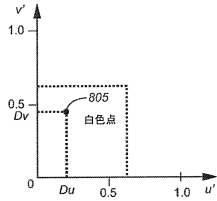
【図6】



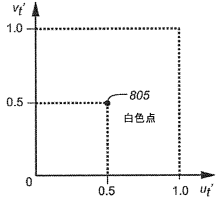
【図7】



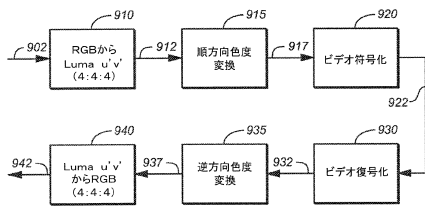
【図 8 A】



【図 8 B】



【図 9】



フロントページの続き

(31)優先権主張番号 62/002,631

(32)優先日 平成26年5月23日(2014.5.23)

(33)優先権主張国 米国(US)

(74)代理人 100202197

弁理士 村瀬 成康

(72)発明者 スウ, グワン ミーン

アメリカ合衆国 94103 カリフォルニア州 サンフランシスコ マーケット ストリート
1275, ドルビー ラボラトリーズ インコーポレイテッド内

(72)発明者 アトキンス, ロビン

アメリカ合衆国 94103 カリフォルニア州 サンフランシスコ マーケット ストリート
1275, ドルビー ラボラトリーズ インコーポレイテッド内

(72)発明者 ミラー, ジョン スコット

アメリカ合衆国 94103 カリフォルニア州 サンフランシスコ マーケット ストリート
1275, ドルビー ラボラトリーズ インコーポレイテッド内

審査官 岩井 健二

(56)参考文献 特表2011-527862(JP, A)

特開2006-071332(JP, A)

特開2005-037996(JP, A)

米国特許出願公開第2012/0013635(US, A1)

(58)調査した分野(Int.Cl., DB名)

H04N 19/00 - 19/98