

視覚モデルに基づいた高ダイナミックレンジ動画像の見え再現

平井 経太[†] 中口 俊哉[†] 津村 徳道[†] 三宅 洋一^{†, ††}

Appearance Reproduction of High Dynamic Range Video based on Human Visual Models

Keita HIRAI[†], Toshiya NAKAGUCHI[†], Norimichi TSUMURA[†],
and Yoichi MIYAKE^{†, ††}

あらまし 本論文では動画像の見えをより現実の風景のように再現することを目的とし、そのための画像改善手法を提案する。カメラはダイナミックレンジの制限により黒つぶれ・白飛びを生じることがあり、現実に近い見えを再現する為には、より広い輝度レンジで多くの階調情報を取得する必要がある。更に視覚の影響により、現実の風景とディスプレイ上での動画像の見えが異なるため、視覚特性を考慮した色再現処理も行わなければならない。そこで本論文では広い輝度レンジ、多くの階調情報を持つ高ダイナミックレンジ(High Dynamic Range, HDR)動画像を作成し、それらに視覚特性を考慮した色再現処理を行う手法を提案する。本手法を用いることで、ビデオカメラで撮影された動画像から色再現、階調再現がより現実の見えに近い動画像を取得することができた。
キーワード 動画像処理, 高ダイナミックレンジ画像, トーンマッピング, 視覚特性モデル, 絶対輝度

1. まえがき

近年、携帯電話、パソコン、DVD 等、デジタル動画像に接する機会が増加している。特に、デジタルビデオカメラの普及により手軽に動画像を記録・再生することが可能になった。しかし、撮影した動画像と我々が実際に見た風景の見えが違う事は多い。例えば、長いトンネルを抜け、一瞬目が眩んだ後に目前に広がる雪景色に感動することはあるが、カメラやディスプレイの性能の制限により、そのような風景をそのまま記録し、再生することは出来ない。そこで、実際に見た風景と撮影した動画像のディスプレイ上での見えを一致させることは重要であり、本研究ではそのような画像改善を目的とする。

現実感のある画像再現の研究として、高ダイナミックレンジ(High Dynamic Range, HDR)画像の取得[2][7]やトーンマッピングオペレータ(Tone Mapping Operator, TMO)またはトーンリプロダクションオペ

レータ(Tone Reproduction Operator)と呼ばれる手法を用いた画像処理が挙げられる。これらの研究の発展により、現在では特別な CCD センサ(HDR カメラ)や表示デバイス[14]を用いず、安価で容易に実際の風景の見えに近い画像が得られるようになった。

HDR 画像取得の代表的な方法として、露光を変えた複数の画像から1枚の画像を合成する手法が挙げられる[2][7]。この手法を用いることで、カメラのダイナミックレンジの制限を越えて、自然界の広い輝度レンジに対する多くの階調情報を持った HDR 静止画像または HDR 動画像が得られる。

また、ディスプレイ上に HDR 画像の階調再現を行う方法として数多くの TMO が提案されている[3]。TMO の有効性を評価する論文[9][17]も発表されているが、未だ明確な評価基準は定まっていない。一方、Tumblin らは現実的な画像の見えを再現する為には視覚特性を考慮した TMO がこそ有用であると提唱している[16]。視覚特性モデルを利用した TMO として、iCAM[4][6]や Pattanaik らの手法[12]があげられる。iCAM は静止画像のみを対象としている。一方 Pattanaik らの手法では動画像を対象としているが、視覚の順応を考慮するため、HDR 動画像と共に入力パラメータとしてシーンの絶対的な輝度

[†] 千葉大学大学院自然科学研究科, 千葉市
Graduate School of Science and Technology, Chiba University,
1-33, Yayoi-cho, Inage-ku, Chiba, 263-8522 Japan

^{††} 千葉大学フロンティアメディカル工学研究開発センター,
千葉市
Research Center for Frontier Medical Engineering, Chiba
University, 1-33, Yayoi-cho, Inage-ku, Chiba, 263-8522 Japan

情報を伴う必要がある。そのため Pattanaik らの手法は CG などの輝度情報が既知な HDR 動画像に広く利用されてきた。

しかし、カメラなどの入力系からディスプレイなどの出力系までを考慮し、動画像の見えを現実のように再現する手法について、今まで議論されていなかった。そのため、カメラで撮影された動画像に対し視覚特性を考慮した TMO を適用する手法は現在提案されていない。

そこで本論文では、一般的なカメラで撮影した動画像のディスプレイ上の見えをより現実の風景の見えに近づけるための基本的な概念を述べ、その手段として HDR 動画像の作成手法と視覚の時空間特性に基づいた TMO を提案する。本手法の流れを 2

章で述べ、動画像データと視覚特性を考慮する際重要となる絶対的な輝度情報の取得方法を 3 章で提案する。4 章では輝度情報をもった HDR 動画像の作成について、5 章では時空間的視覚特性を考慮した TMO を提案する。

2. 本手法の流れ

研究の基本的な概念と提案する処理手順について述べる。図 1 は実際の風景の見え(図 1(a)), カメラで撮影した動画像に対する一般的な処理手順とディスプレイ上での見え(図 1(b)), 本手法による処理手順とその動画像の見え(図 1(c))をモデル化し、比較した図である。一般的な処理手順では実際の風景をカメラで撮影し、その動画像データをディスプ

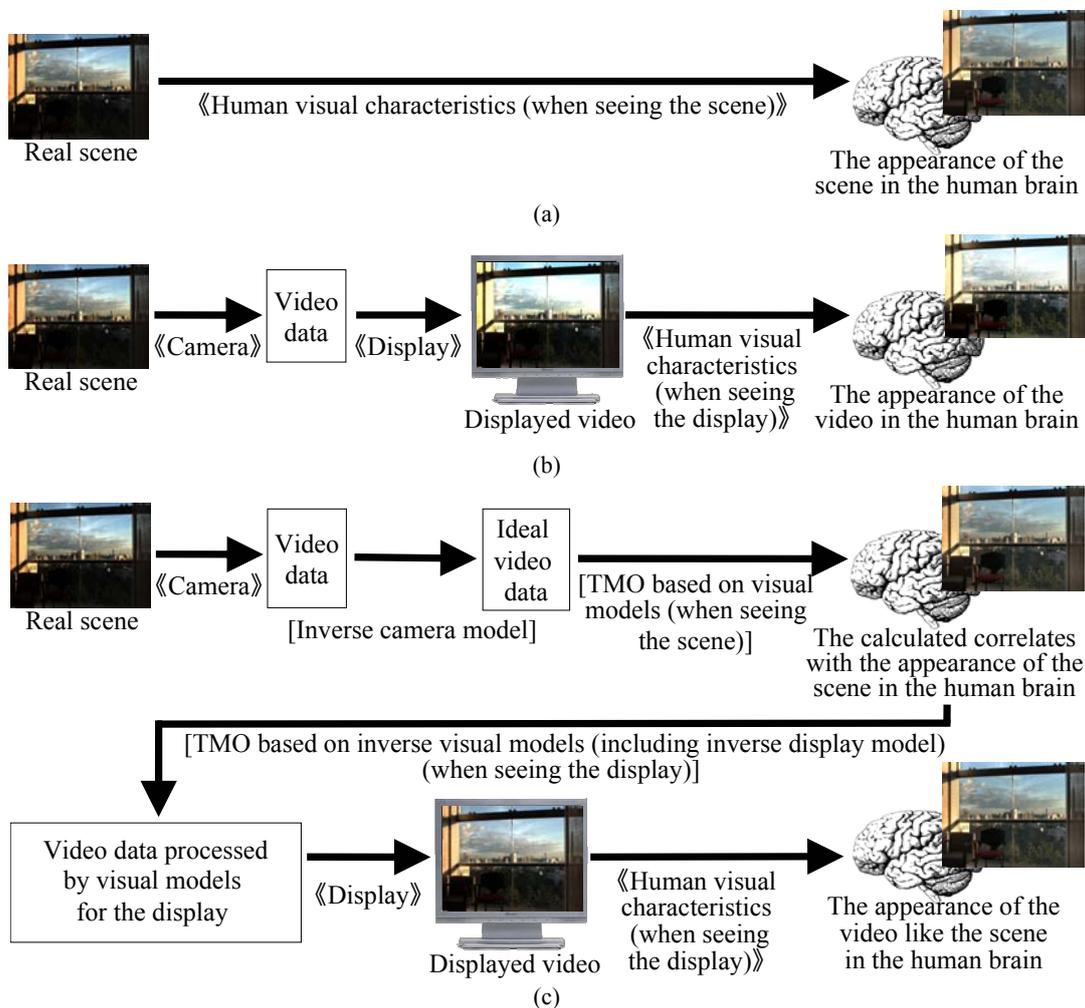


図 1 風景または動画像の見えに至る流れ

(a)実際の風景を見る時の流れ (b)カメラで撮影した動画像を見る時の流れ (c)提案する処理手順

Fig.1 The flows when we see real scenes or videos (a)The flow when we see real scenes

(b)The flow when we see videos on a display, which are taken by a video camera (c)The flow of our proposed method

レイなどのデバイスに表示する。そのため、視覚を通した風景の見え(図 1(a))と動画像の見え(図 1(b))は一致しない。

一方、本手法ではまず撮影時のカメラモデルを用いて、現実の風景と同じ情報を持った動画像(理想動画像データ(Ideal video data))を作成する。本研究では視覚特性を考慮した TMO を適用することを目的とするため、絶対的な輝度情報をもった HDR 動画像(Photometric High Dynamic Range 動画像: PHDR 動画像)を理想動画像とする。更に、図 1(c)のように視覚特性をモデル化した TMO(ディスプレイの特性を考慮した処理(図中 inverse display model)も含む)を PHDR 動画像に適用する。これらの手順により、ディスプレイ上での動画像の見えを現実の風景の見えに一致するよう再現する。

3. 動画像の取得と輝度情報の算出

本研究では動画像データを取得する際、撮影した風景の絶対的な輝度情報が必要となる。従来の HDR 画像取得方法[2][7]では階調性やダイナミックレンジを高めることはできるが、画像内の絶対輝度値は得られない。絶対輝度値が必要な場合、作成した HDR 画像上の点の輝度を実際に測定し、画素値との関係をキャリブレーションする必要がある。一方、本研究では一般的なカメラで使用されている撮影パラメータを利用した方法で絶対的な輝度情報を算出する手法を提案する。

3.1 カメラ情報の取得

人の視覚に入射する光量は瞳孔によって調節されるが、カメラに入射する露光量は絞りとシャッタースピードにより制御される。例えば、図 2 のように明るい太陽光下では絞りの開口度合いは小さくなり、暗いトンネル内では大きくなる。このように、撮影時の絞り値とシャッタースピードの値を知ることが出来れば、動画像内の風景の絶対輝度が算出できる。

本研究では、撮影パラメータが取得可能なカメラを用いて動画像を撮影した。最近の家庭用ビデオカメラでは、映像と共にパラメータ情報が記録されている製品もあり、この方法は十分一般的と考えられる。

3.2 輝度情報の算出

輝度情報の具体的な算出方法について述べる。まずカメラをガンマ補正の影響を除いた状態で自動

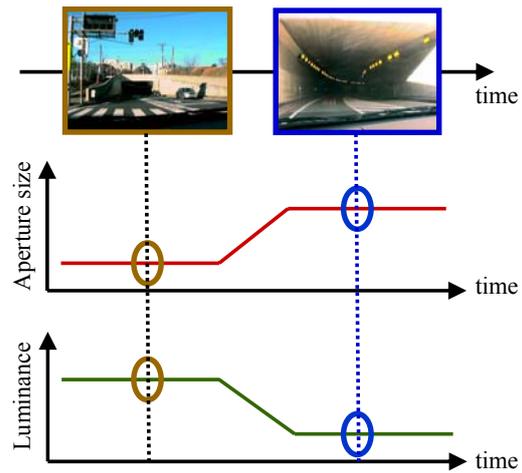


図 2 絞り値と輝度値の関係

Fig.2 The relationship between apertures and scene luminance

露出に設定し、その時の絞り・シャッタースピード・ISO 感度の値を各フレーム毎に取得する。

次に日本工業規格のデジタルカメラ用画像フォーマット規格として用いられている APEX 規格[8]に基づいた計算方法で輝度値を得る。以下はその算出式である。

$$\log_2 \frac{B}{3.42} = 2 \log_2 F - \log_2 T - \log_2 \frac{S}{3.125} \quad (1)$$

ここで、 B は絶対輝度(cd/m^2)、 F は絞り値(F 値)、 T はシャッタースピード(sec)、 S は ISO 感度である。

カメラの自動露出設定値は内部の反射光式露出計により決定されるが、これは測光点が 18% のグレイになるよう設定されている。よって、撮影したフレーム t の各画素 (x, y) を式(2)により、正規化する。

$$L(x, y, t) = \frac{100 B(t) Y(x, y, t)}{18 Y_{MAX}(t)} \quad (2)$$

ここで、 $B(t)$ はフレーム t において算出した絶対輝度(cd/m^2)、 $Y(x, y, t)$ は CIE XYZ 表色系での Y 値、 $Y_{MAX}(t)$ は同表色系での最大 Y 値、 $L(x, y, t)$ は絶対輝度(cd/m^2)を示している。この方法により各フレーム・各画素における絶対的な輝度値が算出できる。図 3 は入力画像と本手法により算出した絶対輝度分布画像である。一般的な目安として青空は約 $5000 \text{cd}/\text{m}^2$ 、日陰は約 $100 \text{cd}/\text{m}^2$ とされているが[1]、それに近い値が得られていることが分かる。

4. Photometric High Dynamic Range 動画像の作成

複数の静止画像から HDR 画像を作成する場合、カメラの位置を固定し、動かない対象物体を撮影す

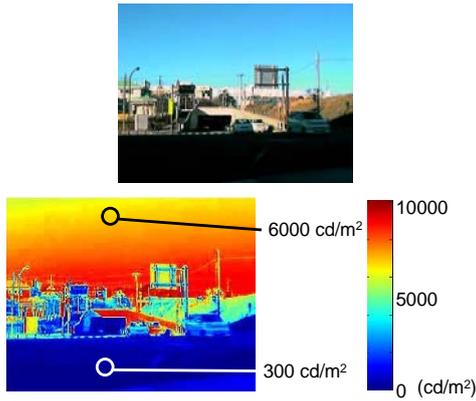


図3 (上段)入力画像
(下段)算出した絶対輝度分布画像

Fig.3 (Top)Input image
(Bottom)Calculated luminance distribution image

る必要がある[2]. しかし、動画像ではそのような条件下での撮影は難しいため、露出設定を変えたフレームを撮影し、それらの間で対応関係を得る必要がある。Kang らは前後フレーム間において、その露光を交互に変えた動画像を撮影し、フレーム間の対応関係を得ることで、HDR 動画像を作成している[7]. また、Sand らは同じ対象物体に対し、露出設定の変えた状態で、2つの動画像を別々に撮影し、その動画像間の対応を求めることにより、HDR 動画像を実現している[13].

本章ではまず、露光とカメラ位置が異なる条件下で撮影された2枚の静止画像から対応関係を求め、HDR 画像を作成する手法を提案する。次にフレーム間の対応関係から PHDR 動画像を作成する。従来手法では意図的に露出設定を変え、動画像を撮影する特殊な方法を用いて HDR 動画像を作成していた。一方、本論文では一般的な方法で動画像を撮影し、その動画像から PHDR 動画像を作成する手法を提案する。

4.1 HDR 画像の作成

HDR 画像作成の流れを図4に示す。画像間の対応を得る一般的な手法では各画像において同一対象物体の画素値は等しいと仮定している。一方、本手法ではカメラを自動露出に設定するため、物体の画素値は変化する。そこで3章で算出した各画素での絶対輝度値を用いて正規化することで、同一対象物体の画素値を等しくする。図5は入力画像、図6は正規化した画像を示している。入力画像の撮影は位置と露出設定が異なる条件になるよう、カメラを

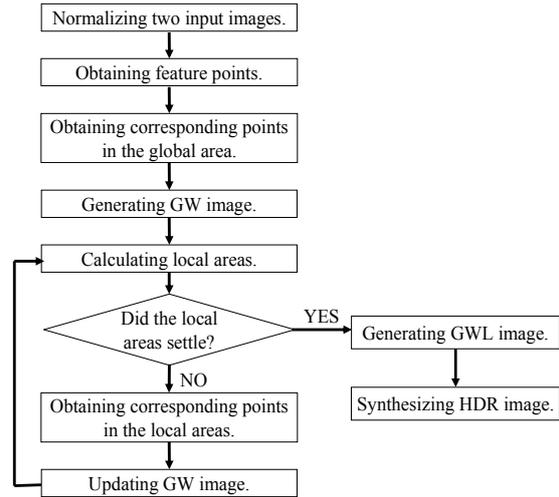


図4 HDR 画像作成の流れ

Fig.4 The flow of generating HDR image

パンした状態で行われた。また図6のように輝度レンジが広い画像(入力画像1, Im1)の最大値を狭い画像(入力画像2, Im2)の最大値と一致させることで、正規化した画像を得る。

次に画像間の対応関係を得る必要があるが、動画像を対象としたとき、カメラの動きなどによる広域での対応関係と画像内の物体の動きや変形などで生じる局所領域での対応関係が必要となる。広域での対応関係のベクトル(オプティカルフロー)はSandら[13]の方法を用いて検出した。これは、特徴点を得る手法として広く知られている Harris オペレータ[5]や動画像のフレーム間の対応関係を得る KLT 法[10][15]をもとにし、よりロバストなオプティカルフローを検出する手法である。ここで、入力画像を広域のオプティカルフローにより変形した画像を Global Warped 画像(GW 画像)と呼ぶ。

次に、フレーム間で画像の一部に存在する物体が移動した場合を考慮し、GW 画像だけでは対応が一致しないような局所領域内でのオプティカルフローを得る。本研究では式(3)を満たす画素(x,y)を中心とした、3×3 領域を局所領域と定めた。

$$|GWI_{2nor}(x,y) - I_{1nor}(x,y)| > \tau_L \quad (3)$$

ここで、 GWI_{2nor} は正規化された Im2 に対する GW 画像、 I_{1nor} は正規化された Im1、 τ_L は局所領域を満たす閾値である。本研究では経験的に τ_L を最大画素値の 5%とした。この局所領域内で広域と同様にオプティカルフローを求め、画像を変形する。この処理を局所領域の面積が収束するまで繰り返し計算をする。これにより、対応関係を用いて補正した画



図 5 (左)入力画像 1(Im1) (右)入力画像 2(Im2)
Fig.5 (left)Input image 1(Im1) (right)Input image 2(Im2)



図 6 (左)正規化した Im1 (右)正規化した Im2
Fig.6 (left)Normalized Im1 (right)Normalized Im2

像(Global + Local Warped 画像 2, GLW 画像 2)(図 7)を得る. また図 8 はそれぞれ正規化後の Im1 と Im2, Im1 と GLW 画像 2 の差分画像であり, 画素値の差が小さい領域ほど黒く表示されている. このように GLW 画像 2 は Im2 に比べ, Im1 との差分が小さいことから, 位置のずれが補正されていることが確認できる.

最後に, 輝度レンジが狭い画像(GLW 画像 2)と広いレンジの画像(Im1)を用いて, 広い輝度レンジ・多くの階調情報を持つ HDR 画像を合成する. 本研究では, カメラのガンマ補正の影響を除いて画像を撮影している. そのため, 式(4)のように 3.2 節で取得された絶対輝度値を用いて画像を正規化することで, HDR 画像が合成できる.

$$HI(x, y) = \begin{cases} \frac{L_{2MAX}}{L_{1MAX}} GLWI_2(x, y) & \text{if } GLWI_2(x, y) < \tau_H \\ I_1(x, y) & \text{otherwise} \end{cases} \quad (4)$$

ここで, HI は HDR 画像, L_{1MAX} , L_{2MAX} はそれぞれ Im1, Im2 の最大絶対輝度値, I_1 , $GLWI_2$ はそれぞれ Im1, GLW 画像 2, τ_H は GLW 画像 2 を用いる領域の閾値を示し, 最大画素値の 98%とした. これは GLW 画像 2 において, 白とびなどの影響を考慮し, 画素値の上位 2%のレンジを取り除いた領域を使用することを示している. また, カメラ位置の違いから一方の画像のみに撮影されている対象物体が存在する場合がある. そのような領域では, 画像間でオプティカルフローによる対応が取れないため誤差が発生し, 最終的に局所的な誤差領域(図 8(右)の黒以外の領域)として残る. 本研究ではそのような



図 7 入力画像 2 の GLW 画像(GLW 画像 2)
Fig.7 Global + local warped input image 2 (GLW image2)

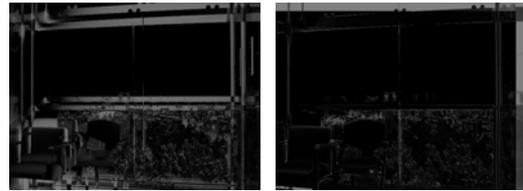


図 8 (左)正規化後の Im1 と Im2 の差分画像 (右) 正規化後の Im1 と GLW 画像 2 の差分画像
Fig.8 (left)The difference image between normalized Im1 and Im2 (right)The difference image between normalized Im1 and GLW image2

領域において I_1 を用いることで, HDR 画像の値を補完した. 図 9 は各画像において HDR 画像の合成に用いる領域を示している. また, 本手法により合成した HDR 画像を図 10(線形圧縮), HDR 画像に TMO として iCAM[4][6]を適用した画像を図 11 に示す. このように, Im 1 では黒つぶれしている領域, Im2 では白飛びしている領域が, 画像間の対応関係を得ることで補正でき, 図 10 のように広い輝度レンジ・多くの階調情報を持ち, また図 11 のように現実的な色・階調再現ができる.

4.2 PHDR 動画像の作成

4.1 節の手法により露出設定が変化したフレーム間の対応を求め, 輝度情報をもつ PHDR 動画像を作成する. 図 12(a)は露出設定の変化が生じたフレームを時系列で並べたもの, 図 12(b)~(d)はそのときの輝度変化と近傍フレームの影響を考慮した PHDR 動画像の作成手順と露出設定が変化した, または変化しないフレームの時間幅を示している.

まず, 露出設定が変化した時(図 12, τ_2 , τ_4)の HDR フレーム作成について述べる. カメラの自動露出機能は最適な値に設定されるまで時間を要するため, 隣接するフレーム間のみで作成された HDR フレームでは広い輝度レンジに対応できない. 広い輝度レンジに対応する方法としては, 露出設定が変化し始



図9 (左)Im1 を用いる領域
(右)GLW 画像 2 を用いる領域
Fig.9 (left)The applied area of Im1
(right)The applied area of GLW image2



図 10 合成された HDR 画像(線形圧縮)
Fig.10 Synthesized HDR image(linearly compression)



図 11 HDR 画像(iCAM 適用後)
Fig.11 HDR image(tone-mapped by iCAM)

めたフレーム(図 12, 時刻 t_a , t_c)と変化が終わるフレーム(時刻 t_b , t_d)から参照用の HDR フレームを時刻 t_a , t_b , t_c , t_d で作成しておき, それをもとに隣接するフレーム間で HDR フレームを作成する方法が考えられる. しかし, この方法ではカメラの動きにより用いる 2 枚のフレームの画像構成が大きく異なり, オプティカルフローの対応が取れない場合がある. そのため, より広い輝度レンジ・多くの階調情報をもつ HDR フレームを作成するためには近傍フレーム全てを考慮する必要がある. 近傍フレーム全てを考慮する方法としては, τ_2 , τ_4 において, 全フレームを同時に処理する方法が挙げられるが, この方法では τ_2 , τ_4 が数十フレームに及ぶ場合, 計算機上での記憶コストが膨大となる. そこで, 本手法では隣接する 2 つのフレーム間の対応から HDR フレームを作成し, これを順方向, 逆方向に処理すること

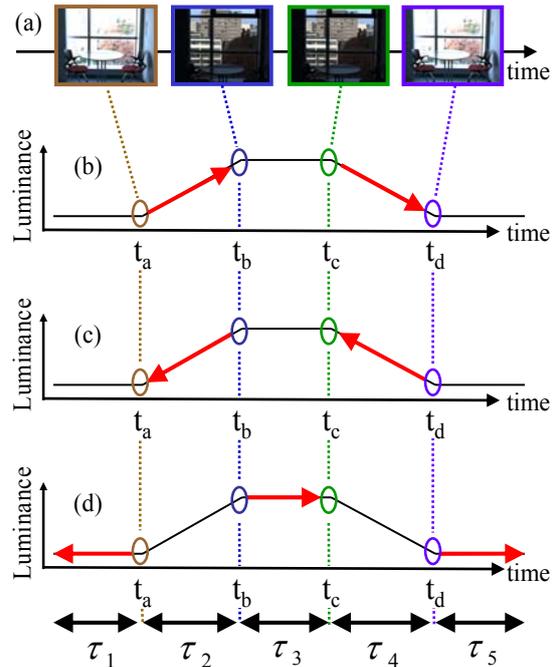


図 12 PHDR 動画の作成手順 (a)露出設定が変化したフレーム (b)順時間軸方向の処理(1 パス目) (c)逆時間軸方向の処理(2 パス目) (d)露出設定が変化しないフレーム間での処理(3 パス目)

Fig.12 The flow of generating PHDR video (a)The frames changing the exposure settings (b)The first processing for generating PHDR video (c)The second processing (d)The third processing

により(図 12(b), (c)), 近傍フレーム全てを考慮した PHDR 動画を作成する.

まず輝度レンジが変化するフレーム間(τ_2 , τ_4)において, 式(5)により順方向の HDR フレーム作成処理(図 12(b))を行う.

$$B'(t) = H(B'(t-1), I(t))$$

$$t = \{t^s + 1, \dots, t^e\}, \quad B'(t^s) = I(t^s) \quad (5)$$

ここで, $H(\text{Im1}, \text{Im2})$ は 4.1 節で述べた入力画像 Im1 と Im2 より HDR 画像を作成する関数, $I(t)$ は時刻 t での入力フレームであり, ここで得られた HDR 画像を $B'(t)$ とおく. t^s は露出設定が変化した時刻(t_a , t_c), t^e は変化が終わる時刻 (t_b , t_d)であり, 常に $t^s < t^e$ を満たす.

得られた HDR 画像 $B'(t^e)$ は露出変化が生じた近傍フレーム全てを考慮しているが, 時刻 t^e 以外のフレームでは全ての近傍フレームを考慮した HDR 画像は作成されていない. そこで, $B'(t^e)$ を参照 HDR フレームとし, 式(6)により逆方向の処理(図 12(c))

を行う。

$$B(t) = H(B'(t), B(t+1))$$

$$t = \{ t^e - 1, \dots, t^e \}, \quad B(t^e) = B'(t^e) \quad (6)$$

ここで、 $B(t)$ は時刻 t でのHDR画像であり、これが露出設定が変化したフレーム間において全ての近傍フレームを考慮したHDRフレームとなる。

τ_2, τ_4 期間のHDRフレーム作成が完了した後に、露出設定が変化しないフレーム間(τ_1, τ_3, τ_5)でHDRフレームを作成する(図12(d))。ここでは、 $B(t_a), B(t_b), B(t_c), B(t_d)$ が得られているので、これらを参照HDRフレームとし、 τ_1, τ_3, τ_5 フレーム群はそれぞれ $B(t_a), B(t_b), B(t_d)$ からHDRフレームを作成する。

ただし、カメラが固定されている時など、露出設定が動画像中で一度も変化しない場合、HDRフレームが作成できない。このような動画像に対しては事前処理として露光を変化させたフレームを撮影し、参照HDRフレームを作成しておく。

本研究では以上の手順により動画像中の全HDRフレームを作成し、これをPHDR動画像とする。

5. 時空間的視覚特性を考慮したTMO

一般的にTMOは空間的に均一または非均一、時間軸に依存または非依存かで大別できる[3]。本研究ではPHDR動画像をディスプレイ上に再現する為に時空間的視覚特性を考慮したTMOが必要となる。そこで本章では、視覚特性を考慮したTMOの概念と従来手法を説明し、それらの手法では実現できなかった時空間的視覚特性を考慮したTMOについて提案する。

5.1 視覚特性を考慮したTMO

Tumblinらは、ディスプレイ上により現実的な映像を表示するためには視覚特性を考慮したTMOが重要であると述べ、その概念(図13)を提唱している[16]。この概念では、まず視覚特性モデル(CIECAM02[11]など)を用いて観察している対象を処理することで、クロマ・明度・明るさ・カラフルネスなどの見えに関する相関量 Q を算出する。これは、照明や表示デバイスが異なる2つの環境下において算出された Q の値が等しいとき、視覚的な見えは一致することを示している。また、視覚モデルの計算は可逆性をもつ。つまり、現実の風景の見えとディスプレイ上での画像の見えを一致させるには、まず風景から視覚特性モデルを用いて Q を算出し、

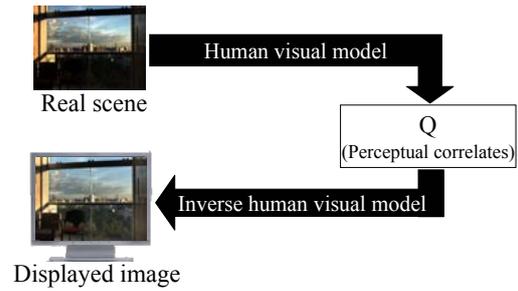


図13 視覚特性モデルに基づいたTMOの概念

Fig.13 The concept of TMO based on human visual models

Q から逆方向処理によりRGB値をもとめ、その値をディスプレイ上に表示すればよい。

iCAM[4][6]やPattanaikらの手法[12]はこの概念に基づいたTMOである。iCAMは静止画像に対し、空間的な処理を行う。また、シーンの輝度情報に影響されるパラメータはフレーム毎に手動で設定する必要があるため、動画像の全フレームに適用することは非現実的である。Pattanaikらの手法は各フレームに対し、シーンの輝度情報に影響されるパラメータは自動的に決定できるが、空間的な処理は考慮されていない。また、この手法では絶対的な輝度情報を既知とし処理を行っているため、何らかの方法で輝度情報を取得する必要がある。

このように従来手法ではカメラで撮影した動画像に適用することは困難であったが、本手法ではこれらの問題を改善し、時空間的視覚特性を考慮したTMOを動画像に適用する手法を提案する。

5.2 時空間的視覚特性を考慮したTMO

本研究では、iCAMを基本とし、輝度情報に依存するパラメータを時間軸で最適に設定することで、PHDR動画像を処理する。

シーンの輝度値が変化するとき、人の視覚は順応するまでに時間がかかる。これは式(7)のように表せる[12]。

$$F_{td} = 1 - e^{-t/r_t}$$

$$L_{td}(x, y, t) = L(x, y, t-1) + F_{td}[L(x, y, t) - L(x, y, t-1)] \quad (7)$$

ここで、 L は式(2)で算出した輝度値、 r_t は視覚が応答するまでの時間(80ms[12])、 L_{td} はシーンの輝度変化に順応するまでの視覚の応答の遅れを考慮した輝度である。

iCAMにおいて、環境光への順応度合いと、シーンの輝度に対する視覚の非線形応答は任意パラメ

ータとなっている。本研究では各フレームにおいて最適な値を定める。

人の視覚は異なる照明下において、その色に順応すること(色恒常性)が知られているが、その順応の度合い D (順応なしの0から完全順応の1.0までの値をとる)はシーンの輝度によって変化する。本研究では順応度合い D を式(8)により算出する[11]。

$$D(x, y, t) = 1 - \left(\frac{1}{3.6} \right) e^{\left(\frac{-L_{id}(x, y, t) + 42}{92} \right)} \quad (8)$$

人の視覚は明るくなるほど感度が鈍るというWeber-Fechnerの法則に従う。シーンの輝度値に対する視覚の順応は非線形応答であり、式(9)により算出できる[11]。本研究ではこの値をiCAMに適用する。

$$F_L(x, y, t) = 0.2 \left\{ \frac{1}{(k+1)} \right\}^4 k + 0.1 \left[1 - \left\{ \frac{1}{(k+1)} \right\}^4 \right]^2 k^{\frac{1}{3}} \quad (9)$$

$$k = 5L_{id}(x, y, t)$$

$$I_a(x, y, t) = \frac{400(F_L(x, y, t)I(x, y, t)/100)^{0.42}}{27.13 + (F_L(x, y, t)I(x, y, t)/100)^{0.42}} + 0.1$$

ここで、 I は輝度への順応前の画像、 I_a は順応後の画像である。

以上の手順により視覚特性を考慮したTMOをPHDR動画像に適用し、ディスプレイへの出力値を得る。

6. 結果

図14はそれぞれ入力動画像、PHDR動画像(線形圧縮)とその輝度分布、処理動画像(5.2節のTMO適用後)である。入力動画像(図14(a))は撮影対象を室内から室外に移したときであり、このような時、カメラは適正露出を自動で算出するが、ダイナミックレンジの制限により、白飛びする領域や影のような黒つぶれ領域が生じる。しかし、本手法を用いることで広い輝度レンジ、多くの階調情報、絶対輝度値をもったPHDR動画像(図14(b), (c))が取得できると共に、処理動画像(図14(d))では、視覚特性を考慮したTMOを適用しているため、入力動画像では白飛びや黒つぶれしていた領域も実際の見えに近い色・階調再現ができています。

一方、カメラ位置が固定された時は、カメラの露出設定は変化せず、PHDR動画像を作成することは出来ない。そのため、事前処理として露光を変化させたフレームを撮影し(図15(a), (b)), 参照HDRフレーム(図15(c))を作成しておく。これにより、入力

動画像(図15(d))では白飛びし、ビルが見えていなかったが、参照HDRフレームを用いた処理動画像(図15(e))では、ビルの色も再現できており、より現実的な見えに近づいたといえる。

ただし、カメラの動きが速い場合、撮影された対象物体の大きさや形がフレーム間で大きく異なる。また一方のフレームでは撮影されている対象が、もう一方のフレームでは手前の物体の影響で撮影されていないこともある。そのような場合、フレーム間での対応がとれないため誤差が大きくなる。このような誤差によるノイズの発生を防ぐため、カメラの速い動きに対応する改善が必要である。

本手法を用いて、PentiumIV 2GHz, 主メモリ512MBの計算機を使用し、640×480のサイズの画像をMATLABにより処理した際の計算時間は、フレーム間の対応関係を取り、HDR画像を生成する処理に平均約15秒、TMO適用に約4秒である。計算量はやや大きいですが、従来法[7][13]と比べれば、妥当なコストだといえる。

7. むすび

本論文では現実の風景の見えと一般的なカメラで撮影した動画像のディスプレイ上での見えを一致させるための基本的な概念を述べ、Photometric High Dynamic Range 動画像の作成手法と時空間的視覚特性を考慮したTMOを提案した。この結果、本手法による処理動画像では、入力動画像に比べより現実の見えに近い色・階調再現ができた。sRGBやICCプロファイルなどを用い、デバイス間の色の整合性を取ることで、本手法を現在市販されているデバイスに実装することは可能と考えられるが、計算量を軽減するためにも、今後はより高速なHDR画像の作成手法、更にはリアルタイム処理の実装を目指す必要がある。

その他の課題としては、上記で挙げたようにフレーム間で対応のとれない領域を補間し、より速いカメラの動きに対応する手法を考える必要がある。また、処理動画像では全体的に色あせた印象や樹木のような細部の鮮やかさが失われた印象もある。これは本手法のTMOが全体的な画像の印象や細部の色再現よりも、現実的な見えを再現することを主な目的としていることが原因と考えられる。そのため、今後の課題としてそれらも考慮したTMOの改善が必要である。

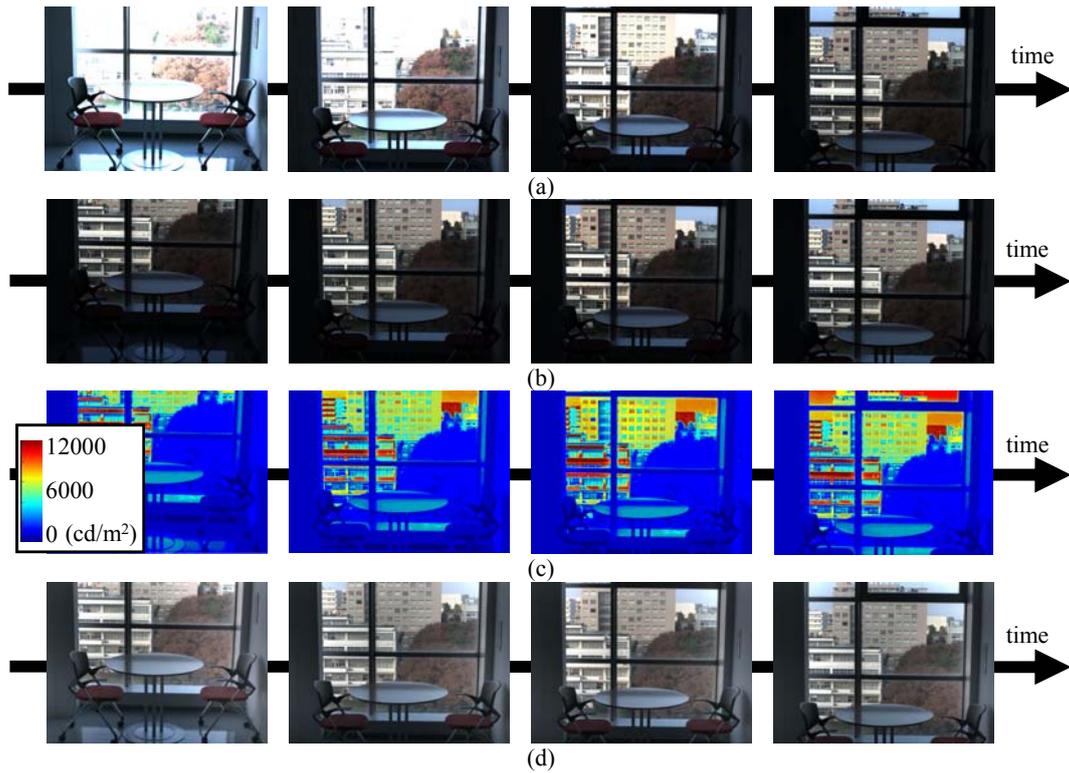


図 14 (a)入力動画像(各フレーム間の時間間隔は 0.5 秒) (b)作成した PHDR 動画像(線形圧縮)
(c)PHDR 動画像の輝度分布 (d)本手法による処理動画像

Fig.14 (a)An input video sequence(The interval time between each frames is 0.5 sec) (b)The generated PHDR video sequence (linearly compression) (c)The luminance distribution of the PHDR video (d)The processed video sequence of our proposed method

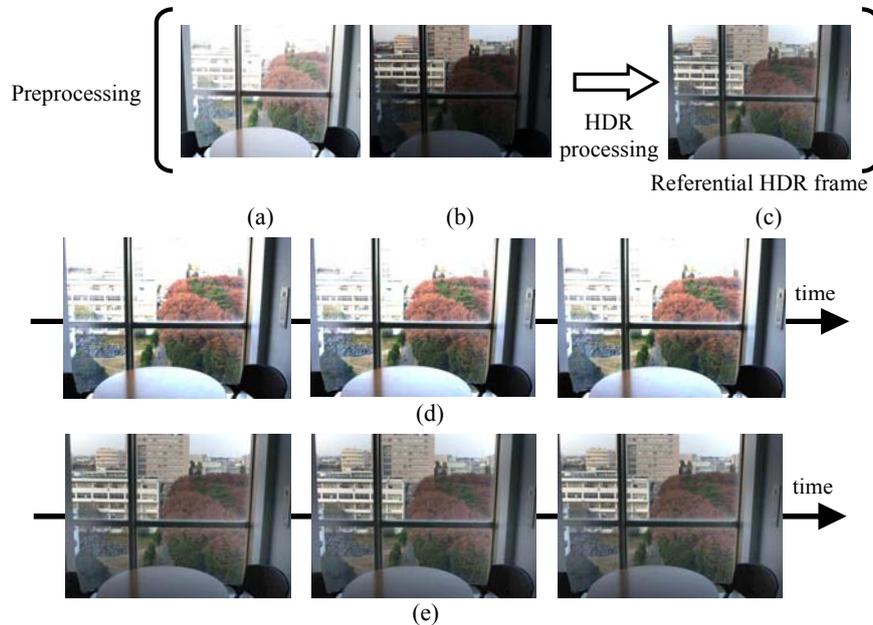


図 15 (a), (b)事前処理として露光を変化させ撮影したフレーム (c)事前処理で作成した参照 HDR フレーム
(d)入力動画像 (e)参照 HDR フレームを用いた処理動画像

Fig.15 (a), (b)The frames taken by changing the exposure settings as preprocessing (c)The referential HDR frame obtained by preprocessing (d)An input video sequence (e)The processed video sequence by using the referential HDR frame

文献

- [1] 大田登, 色再現工学の基礎, コロナ社, 東京, 1997.
- [2] P.E. Debevec, and J. Malik, "Recovering high dynamic range radiance maps from photographs," ACM computer Graphics, vol.31, pp.369-378, 1997 (Proc. SIGGRAPH 1997).
- [3] K. Devlin, "A review of tone reproduction techniques," Technical Report CSTR-02-005, University of Bristol, November. 2002.
- [4] M.D. Fairchild, and G. M. Johnson, "Image appearance model," Proc. IS&T/SPIE 15th Annual Symposium on Electronic Imaging, Vol. 5007, pp.149-160, Santa Clara, Jan. 2003.
- [5] C. Harris, and M. Stephens, "A combined corner and edge detector," Proc. The Fourth Alvey Vision Conf., pp.147-151, Manchester. 1988.
- [6] G. M. Johnson. and M.D. Fairchild, "Rendering HDR image," Proc. IS&T/SID 10th Color Image Conf., pp.23-27, Scottsdale, Nov. 2003.
- [7] S.B. Kang, M. Uyttendaele, S. Winder, and R. Szeliski, "High dynamic range video," ACM Transactions on Graphics, Vol22, Issue3, pp.319-325, 2003 (Proc. SIGGRAPH 2003).
- [8] D. A. Kerr, "Apex - Additive system of Photographic Exposure," <<http://doug.kerr.home.att.net/pumpkin/APEX.pdf>>, (2006/4/24 アクセス).
- [9] P. Ledda, A. Chalmers, T. Troscianko, and H. Seetzen, "Evaluation of tone mapping operators using a high dynamic range display," ACM Transactions on Graphics, Vol.24, Issue3, pp.640-648, 2005 (Proc. SIGGRAPH 2005).
- [10] B.D. Lucas, and T. Kanade, "An iterative image registration technique with an application to stereo vision," Proc. The 7th International Joint Conf. on Artificial Intelligence, pp.674-679, 1981.
- [11] N. Moroney, M.D. Fairchild, R.W.G. Hunt, C.J. Li, M.R. Luo, and T. NewMan, "The CIECAM02 color appearance model," Proc. IS&T/SID 10th Color Image Conference, Scottsdale, pp.23-27, Nov. 2002.
- [12] S.N. Pattanaik, J. Tumblin, H. Yee, and D.P. Greenberg, "Time-dependent visual adaptation for realistic image display," ACM Computer Graphics, pp.47-54, 2000 (Proc. SIGGRAPH 2000).
- [13] P. Sand, and S. Teller, "Video matching," ACM Transactions on Graphics, Vol2, Issue3, pp.592-599, 2004 (Proc. SIGGRAPH 2004).
- [14] H. Seetzen, W. Heidrich, W. Stuerzlinger, G. Ward, L. Whitehead, M. Trentacoste, A. Ghosh, and A. Vorozcovs, "High dynamic range display systems," ACM Transactions on Graphics, Vol2, Issue3, pp.760-768, 2004 (Proc. SIGGRAPH 2004).
- [15] J. Shi, and C. Tomasi, "Good features to track," proc. IEEE 13th Conf. on Computer Vision and Pattern

Recognition, pp.593-600, 1994.

- [16] J. Tumblin, and H. Rushmeier, "Tone reproduction for realistic images," IEEE Computer Graphics and Applications, vol.13, no.6, pp.42-48, Nov. 1993.
- [17] A. Yoshida, V. Blanz, K. Myszkowski, and H. Seidel, "Perceptual evaluation of tone mapping operators with real-world scenes," Proc. IS&T/SPIE 17th Annual Symposium on Electronic Imaging, Vol.5666, pp.192-203, San Jose, Jan. 2005.



平井経太

2005年, 千葉大学工学部 譲歩王画像工学科卒業. 同年同大学院自然科学研究科修士課程に入学. 現在に至る. 主に人間の視覚特性を考慮した入力・出力動画の画質評価・改善に関する



中口俊哉

2003年, 上智大学大学院博士後期課程修了. 博士(工学). 現在, 千葉大学工学部情報画像工学科 助手. 2001-2003年, 日本学術振興会特別研究員. 画像解析, 画質評価, 医用画像処理, 組合せ最適化に関する研究に従事. IEEE, 電子情報通信学会, IS&T, 日本写真学会, 各会員.



津村徳道

1995年, 大阪大学大学院博士後期課程修了. 博士(工学). 現在, 千葉大学工学部情報画像工学科 助教授. 医用画像工学, 計測画像工学, 色再現工学, 質感再現工学の研究に従事. IS&T, 日本光学会, 日本写真学会, 日本医用画像学会, 電子情報通信学会, 各会員



三宅洋一

1968年, 千葉大学大学院修士課程修了, 1978年工博(東工大), 京都工繊大助手, 助教授, スイス連邦工科大研究員, 千葉大助教授を経て1989年教授, 2003年より千葉大フロンティアメディカル工学研究開発センター長. ロチェスター大光学研究所客員教授, 東工大教授(併任), 日本写真学会会長, 日本鑑識科学技術学会(現日本法科学技術学会)理事長, IS&T 副会長などを歴任. IS&T フェロー, 名誉会員, Charles E Ives Award など多数受賞. 著書: デジタルカラー画像の解析評価(東大出版会), 編著: 分光画像処理入門(東大出版会) など多数.

Appearance Reproduction of High Dynamic Range Video based on Human Visual Models

Keita HIRAI[†], Toshiya NAKAGUCHI[†], Norimichi TSUMURA[†], and Yoichi MIYAKE^{†, ††}

[†] Graduate School of Science and Technology, Chiba University, 1-33, Yayoi-cho, Inage-ku, Chiba, 263-8522 Japan

^{††} Research Center for Frontier Medical Engineering, Chiba University, 1-33, Yayoi-cho, Inage-ku, Chiba, 263-8522 Japan

Abstract In order to match precisely the appearance of the recorded video sequence with that of real scenes, the limited dynamic range of a camera, that of a display and human visual characteristics have to be considered. In this paper, we propose a method which can record the high dynamic range video sequence with absolute luminance value of the real scenes, and also describe a new tone mapping method based on spatio-temporal human visual characteristics. Experimental results show that proposed methods reproduce the appearance of real scenes from the recorded video sequence.

Key Words Video processing, High dynamic range imaging, Tone mapping, Human visual models, Absolute luminance