

HDR 画像生成のためのノイズの低減を考慮した多重露光画像統合法

松岡 諒[†] 神納 貴生^{††} 奥田 正浩[†]

Multiple Exposure Image Fusion with Noise Reduction for HDR Acquisition

Ryo MATSUOKA[†], Takao JINNO^{††}, and Masahiro OKUDA[†]

あらまし 暗部をもつシーンの光情報を保存する HDR (High Dynamic Range) 画像を高感度撮影した多重露光画像を統合して生成する場合、高感度撮影により発生するノイズが無視できなくなる。また、HDR 画像の生成に用いる LDR (Low Dynamic Range) 画像の階調不足による量子化誤差が擬似エッジを生む。これらは HDR 画像をトーンマッピングにより階調削減した場合に特に顕著となる。本論文では比較的ノイズの少ない高露光画像を基に低露光画像の暗部ノイズを軽減し、また、暗部ノイズと擬似エッジを選択的に軽減できる新たな重み付け手法により HDR 画像を生成する多重露光画像統合法を提案する。本論文では、暗部ノイズ除去、暗部ノイズと擬似エッジの選択的軽減、の二種類についてそれぞれ実験を行い、その結果を従来の手法と比較することで提案法の有効性を示す。

キーワード HDR 画像, ノイズ除去, 多重露光画像統合, 重み関数

1. ま え が き

人間の視覚特性 (Human Visual System (HVS)) は様々な観測条件における光の強度に適応し、高いダイナミックレンジの光を知覚することができる。一方、最近のデジタルカメラで使用されている CCD や CMOS センサーのダイナミックレンジは、現実シーンの知覚範囲をカバーすることができない。また、自然シーンの高いダイナミックレンジの放射照度を捉え、輝度値を保存することは多くのアプリケーションで非常に重要となっている。例えば、高ダイナミックレンジの光情報を取得できれば、強い逆光の下で人を撮影したり、トンネルから外側を撮影するとき、露出アンダー若しくは露出オーバーの問題に対処することが可能となる。そこで、自然シーンの高いダイナミックレンジを保持するために考案されたのが High Dynamic Range Image (HDR 画像) である。HDR 画像は人間の視覚特性と同等のダイナミックレンジを保持することができ、CG の分野において、輝度情報を基にした

画像を用いた高品質なレンダリング手法で一般的に使われている [1] 他、高精細写真現像、車載カメラ、監視カメラ等に应用されている。

ここ十年間で、多重露光画像を基にした HDR 画像を生成するための多くの手法が提案されてきた [2]～[8]。HDR 画像は幾つかの多重露光画像を撮影し合成することによって構築される。多くの手法では、高いダイナミックレンジを得るために、低い露光量から、高い露光量の複数枚の写真を撮影する。ダイナミックレンジは通常、画像の暗所と明所の強度の比で定義され、暗所においてはノイズに支配されない最も低い値と定義するのが一般的である。また、画像センサーによって捕捉された画像は、暗電流ノイズやショットノイズのような複数のノイズを含んでいる。そのため、高いダイナミックレンジを取得するためには暗部のノイズ除去が重要となる。

多重露光の統合処理は本質的にノイズを減らす働きをもっている。しかしながら、十分にノイズを除去することはできず、暗い領域において予期しないノイズが現れてしまう。特に暗い照明条件の下、手もちカメラで写真を撮影するとき、暗い領域での手ぶれを回避するために高い ISO 感度の設定が必須となるが、この場合、ダイナミックレンジの低下を招いてしまう。このノイズは、HDR 画像に影響を及ぼし、暗い領域に

[†] 北九州市立大学, 北九州市
The University of Kitakyushu, Kitakyushu-shi, 808-0135
Japan

^{††} 豊橋技術科学大学, 豊橋市
Toyohashi University of Technology, Toyohashi-shi, 441-
8580 Japan

において顕著となる．HDR 画像は一般的な出力デバイスで出力する際、そのダイナミックレンジに合わせるために、トーンマッピング処理を行うが、多くのトーンマッピング手法は暗部を強調する効果をもち、暗部におけるセンサーノイズが強調される．一方、明部では多重露光画像の低い階調数が原因で、量子化誤差が無視できなくなることがある．

本論文では、暗い領域をもつシーンでもノイズの少ない HDR 画像の生成を実現する新たな多重露光画像生成手法を提案する．提案法では、比較的ノイズの少ない高露光画像を基に低露光画像の暗部ノイズの軽減を実現し、多重露光画像のための wavelet-shrinkage と、統合のための新たな重み付け手法によりノイズ除去を達成する．具体的には、多重露光画像を逆カメラレスポンスカーブにより線形化した後、多重露光画像間でのノイズ除去を行い、重み付けを行う．重み付けされた画像は、シフト不変 wavelet 変換により、低周波数成分と高周波数成分に変換され、高周波数成分の wavelet 係数は多重露光画像統合用に最適化した wavelet-shrinkage によりしきい値判定を行い、wavelet 領域で画像を統合する．また、提案法では統合の際の新たな重み付け手法を提案する．提案する重み関数は多重露光画像統合の生成過程で暗部ノイズと量子化ノイズを選択的に重み付けすることが可能となる．

本論文では、2. に従来の多重露光画像統合法について概要を説明し、3. に暗部領域のノイズを除去する新たな多重露光画像統合法について説明し、4. では、暗部ノイズと量子化ノイズをより一層減少させるような新たな重み付け手法について説明する．5. では、提案統合手法での結果を従来法と比較し、また、提案重み付け手法の効果を示し、提案法のアルゴリズムの妥当性を示す．

2. 多重露光画像統合法

2.1 概 要

図 1 に提案法の処理手順の概略を示す．ここではカラー多重露光画像を入力とする．提案法ではカラー画

像の RGB チャンネルそれぞれにおいて同じ処理を行う．初めに、各入力画像のシーン輝度に対する非線形性を補正し、画素値をシーン輝度に対して線形となるように輝度補正した上で、多重露光画像間でノイズ除去を行う．そして、シーン輝度に対する誤差を最小にするように構成された重み関数を、それぞれの画像に掛け、その重み付けされた画像をシフト不変 wavelet 変換する．wavelet 領域において、高周波 wavelet 係数にしきい値判定処理が施され、その後、多重露光画像は統合される．HDR 画像は、逆シフト不変 wavelet 変換を行うことで復元される．各ステップの詳細については以下の章で説明する．

2.2 従 来 法

放射照度 R と、センサーによって測定する光量 Y の関係は式 (1) によって表現することができる [4]．

$$Y = R \cdot t, \quad (1)$$

ここで t は露出時間である．

大半のカメラセンサーでは、捕捉された信号 Y は画素値 i に非線形変換される．画素値 i は、RGB いずれかのチャンネルの値を表し、一般的には 8bit に量子化される．本論文では多重露光画像の画素値を $[0, 1]$ と正規化して扱う．放射照度を正確に推定するために、この非線形変換の逆関数を近似することによって非線形性を補正する必要がある．非線形変換はカメラレスポンスカーブと呼ばれ、次のように定義される．

$$i = g(Y) \quad (2)$$

式 (1), (2) より放射照度 R は以下のとおり導出される．

$$R = f(i) / t \quad (3)$$

ここで、 $f(x)$ は逆カメラレスポンスカーブであり、 $f(x) = g^{-1}(x)$ で定義できる．

本論文では既存のカメラキャリブレーション手法の中から、Mitsunaga らの方法 [4] を使用し、関数 f を推定している．この手法で関数 f は多重露光画像と、画像間の露光比から推定され、低次数の多項式で近

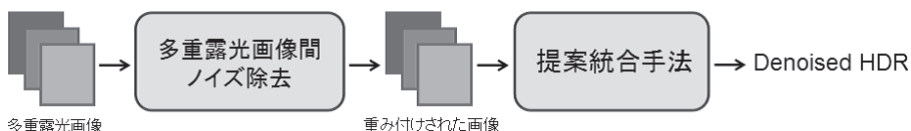


図 1 提案法のフローチャート

Fig. 1 Our method flow.

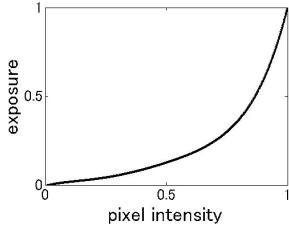


図 2 逆カメラレスポンスカーブ
Fig. 2 Inverse camera response curve.

似される．この手法により推定された CANON EOS 20D の関数 f を図 2 に例示する．また，ここで関数 f は RGB それぞれ推定されるが，これらの関数 f はほぼ同様の応答関数を示すため， G チャンネルの関数のみを例示している．本論文において，多重露光画像はカメラの露出時間を変化させて撮影している．このとき，露出時間以外の設定は固定する．

従来手法において多重露光画像の統合は次式により画素単位で行われる．

$$I_m = \frac{\sum_{n=1}^N \omega_m(n) R_m(n)}{\sum_{n=1}^N \omega_m(n)}, \quad (4)$$

N は画像枚数であり， $R_m(n)$ は n 枚目 m チャンネルの露光画像の放射照度であり， $\omega_m(n)$ はその重みである．ここで，重みは飽和領域を除外するように設計された重み関数により画素単位で導出される．本論文では，全画素で同じ処理を行うため以降全ての式において画素の添え字を省略する．

3. ノイズを考慮した多重露光画像統合法

3.1 多重露光画像間でのノイズ除去

本章では，ノイズ除去を実現するための wavelet 領域における新たな多重露光画像統合手法について説明する．まず，多重露光画像間の揺らぎを除去する画像間でのノイズ除去について説明し，その後，wavelet-shrinkage による多重露光画像の統合について説明する．そして，提案法の重み関数の設計については次章で詳しく述べる．

ノイズ等の生じていない理想的な多重露光画像では，逆カメラレスポンスカーブにより非線形性を補正された各画像の放射照度は飽和領域を除いて一致する．しかし，実際は多重露光画像にはノイズが生じ，また，それらから推定した放射照度にもノイズによる誤差（揺らぎ）が発生する．そこで，提案法ではこれらの誤差（揺らぎ）を軽減するために，画像間での shrinkage

処理を行う．ただし，この処理は式 (3) を用いて非線形性を補正した多重露光画像に対して行う．

まずは式 (5)，(6) より，多重露光画像間の和 L と差 H に分解する．

$$L_m(n) = \frac{M(n)M(n+1)(R_m(n) + R_m(n+1))}{2}, \quad (5)$$

$$H_m(n) = \frac{M(n)M(n+1)(R_m(n) - R_m(n+1))}{2}, \quad (6)$$

ここで， $n = 1, \dots, N-1$ で N は画像枚数， n が小さいほうが低露光画像とする．ただし，高露光画像においては無視できるほどノイズが小さいため， $n = N$ 枚目に関しては処理を行わない． $R_m(n)$ は n 枚目の $m = \{R, G, B\}$ チャンネルの露光画像の放射照度を指している．また， $M(n)$ は n 枚目の露光画像の R, G, B チャンネルのいずれかの画素値が 0（黒潰れ）または 1（白飛び）のときに 0，その他は 1 となるマスクである．このマスクの導入により，飽和領域の影響を無視することができる．式 (6) より得られた $H_m(n)$ に対して次式のしきい値判定処理を行う．

$$H_m^*(n) = \begin{cases} 0, & \text{if } |H_{m=G}(n)| < \text{threshold} \\ H_m(n), & \text{otherwise} \end{cases}, \quad (7)$$

ここで，式 (7) より，露光画像間の放射照度の差がしきい値以下の場合誤差であるとし， $H_m^*(n)$ を 0 とする．つまり，ここでのしきい値とはノイズ除去の程度を表しており，ユーザが設定するパラメータである．この処理により，画像間のノイズ（揺らぎ）を除去している．ここで，このしきい値判定は G チャンネルに対してのみ行い，0 と判定された場合， R, B チャンネルも同様に 0 とする．これは， RGB で独立にしきい値処理を行うと色バランスが崩れ色むらが発生するためであり， G チャンネルのしきい値判定結果を， RGB で統一して用いることで，色バランスを保ちながらノイズを除去している．更に，次式により放射照度を復元する．

$$R_m^*(n) = L_m(n) + H_m^*(n). \quad (8)$$

最終的に多重露光画像を統合する際，画像に重みをつけた重み付き線形和で HDR 画像を復元する．この重みに関する詳細は次章で詳しく述べるが，一般に重みは線形化を行う前にあらかじめ導出する必要がある



図 3 多重露光画像間でのノイズ除去: (左図) 低露光画像, (右図) ノイズ除去結果

Fig. 3 By multiple exposure image noise removal: (left) low exposure image, (right) denoising result.

るため、重みに関しても同様にしきい値判定処理を行い、重みを更新しなければならない。そのままではノイズを考慮した重みとなり、統合の際にノイズの影響が残ってしまうからである。そこで、次式により重み係数を更新する。

$$\omega_m^*(n) = \begin{cases} \frac{1}{2}M(n)M(n+1)(\omega_m(n) + \omega_m(n+1)), & \text{if } |H_{m=G}(n)| < \text{threshold} \\ M(n)\omega_m(n). & \text{otherwise} \end{cases} \quad (9)$$

提案法を用いて予備実験を行った結果を図 3 に示す。図 3 より、多重露光画像を用いたノイズ除去の効果が確認できる。これは、多重露光画像の飽和していない領域では画像本来のエッジの強度は一致するため、画像の揺らぎであるノイズを差分をとりしきい値判定することで軽減することができたと考えられる。

3.2 多重露光画像の統合

従来の統合手法では、式 (4) を用いて多重露光画像の統合を行うが、提案法では、wavelet 領域内で多重露光画像の統合を行う。

まず、多重露光画像の画素値を式 (3) より放射照度 $R_m(n)$ に変換した後、あらかじめ提案法の重み関数により導出しておいた重みを次式により乗算する。

$$R'_m(n) = \omega_m^*(n) \cdot R_m^*(n). \quad (10)$$

ここで $n = 1, \dots, N-1$ 枚目までの露光画像に関しては、式 (8) と式 (9) より導出した放射照度と重みを、 $n = N$ では式 (3) とあらかじめ導出しておいた重みをそれぞれ用いる。式 (10) を用いて全画素で重み付き放射照度値を求めた後、各重み付き画像を Haar 型のシフト不変 wavelet 変換によって変換する。そして、多重露光画像は次節で述べる多重露光画像の統合のた

めの wavelet-shrinkage 処理により wavelet 領域内で統合される。

3.3 wavelet-shrinkage によるノイズ除去

画像の暗部に発生したノイズは、撮影時のカメラレスポンスカーブと HDR 画像作成後の表示の際に用いられるトーンマッピング操作によって強調される。一般的に複数枚の画像を統合する処理は、ランダム性をもったノイズを低減する効果をもつ。しかし、多重露光画像の統合ではカメラレスポンスカーブによりノイズレベルが各露光画像で変化してしまうため、ノイズが低減されづらい。また、生成した HDR 画像は表示のためにトーンマッピング処理されるが、この処理は暗部をもち上げる効果をもつため、残されたノイズが強調されてしまう。

提案統合手法ではノイズを除去するために、多重露光画像用に最適化した wavelet-shrinkage を適用する。式 (10) を用いて生成した重み付き画像を Haar 型のシフト不変 wavelet 変換で変換する。つまり、サブサンプリングなしの wavelet 分解を行う。wavelet 変換では、ローパスフィルタ (L) とハイパスフィルタ (H) は水平及び垂直方向に施される。そして、四つのサブバンド (LL, HL, LH, HH) が生成され、それから、繰り返し LL 帯域を四つのサブバンドに変換していく。wavelet-shrinkage は、wavelet 係数をしきい値処理することによって、wavelet 疎表現を作るという処理である [9]~[12]。

wavelet-shrinkage で一般に用いられる Hard Shrinkage [11] は、次式のコスト関数を最小化することで定義される。

$$\min_h |h|^0 + \lambda(h-l)^2, \quad \lambda > 0, \quad (11)$$

ここで、 l は入力 wavelet 係数であり、 h は出力 wavelet 係数である。また、 λ はノイズ除去具合を調整するパラメータである。式 (11) を最小化する最適な wavelet 係数は次式より導出される。

$$h^* = \begin{cases} 0, & \text{if } |l| < \sqrt{1/\lambda} \\ l, & \text{otherwise} \end{cases}, \quad (12)$$

式 (12) より、しきい値以下の wavelet 係数はノイズであるとみなし、その wavelet 係数を 0 にすることでノイズ除去を実現している。対して提案法の wavelet-shrinkage は、多重露光画像のノイズ除去に加え最適な統合を同時に実現する手法となる。

提案法において、wavelet-shrinkage の最適化問題は、次のコスト関数を最小化することで解くことができる。

$$\min_h E(h) = |h|^0 + \frac{\lambda}{N} \sum_{n=1}^N (h - l(n))^2, \quad (13)$$

ここで、 $l(n)$ は式 (10) を用いて生成した n 枚目の重み付き画像を wavelet 分解し得られるサブバンド (HL, LH, HH) のある wavelet 係数であり、 h は HDR 化された出力 wavelet 係数である。また λ は shrinkage のしきい値を決定するパラメータである。ただし、カラーチャンネルごとに処理を行う。また、簡単のために、添え字 m を省略する。このコスト関数を最小化する最適な wavelet 係数は次式のとおりに閉じた形で導かれる。

$$h^* = \begin{cases} 0, & \text{if } 1 - \lambda \left(\frac{1}{N} \sum_{n=1}^N l(n) \right)^2 > 0 \\ \frac{1}{N} \sum_{n=1}^N l(n), & \text{otherwise} \end{cases} \quad (14)$$

これは一般に用いられる式 (12) 等の wavelet-shrinkage 処理 [9]～[12] とは異なり、ノイズ除去だけではなく多重露光画像の統合を行う式となっている。つまり、統合に最適でない多重露光画像の wavelet 係数は統合せず 0 に、そうでない場合は統合を行う。ここで、 $N = 1$ とすると式 (14) は式 (12) に一致し、式 (14) は wavelet-shrinkage を多重露光画像統合用に一般化したものと捉えられる。なお、式 (14) の導出については付録に示す。

4. 重み関数の設計

4.1 従来法

従来法の統合式である式 (4) に重み関数が導入されているのは、露光不足若しくは過度の露出における飽和領域を加算から除外するためである。従来法では、中間強度を高く、特定の飽和領域の画素値に低い重みを付けるようになっている [2]～[6]。従来の重み関数 [2], [3] について二つの例を図 4 に示す。図 4 の従来の重み関数は、0.5 付近の中間強度は放射照度において最も高い信頼性があるという仮定により設計されている。

4.2 ノイズ除去のための重み付け

CCD や CMOS センサーで発生したノイズは、一般的に信号に依存した項と独立した項でモデル化でき

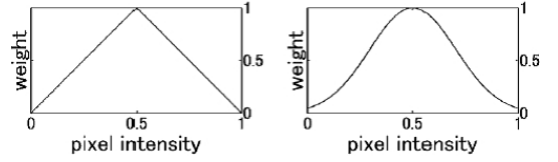


図 4 従来の重み関数: (左図) Hat 型の重み関数, (右図) ガウス型の重み関数

Fig. 4 Conventional two weight functions: (left) Hat-type, (right) Gauss-type.

る (例えば [13])。ここではノイズモデルを次のように定義する。

$$Y = x + a_1\sigma_1 + a_2\sigma_2 \cdot x, \quad (15)$$

ここで、 x はノイズのない理想的な信号であり、 Y はセンサーの出力である。また、 σ_1 と σ_2 はガウシアンノイズである。 a_1 と a_2 はセンサーによって特徴づけられるパラメータである。ここで、第 2 項は信号に依存しない独立したノイズであり、第 3 項は信号に依存したノイズである。センサーに入力された信号はカメラ内の処理により非線形変換され、量子化される。つまりカメラの出力を y とすると、

$$y = g(Y) + \sigma_q$$

と記述できる。ここで σ_q は量子化誤差である。本手法では、この出力 y に逆カメラレスポンスカーブをかけるため、

$$f(y) = f(g(Y) + \sigma_q),$$

となる。

ここで得られる信号 $f(y)$ と理想的な信号 x との誤差を以下のとおりに近似する。

$$\begin{aligned} f(g(Y) + \sigma_q) - x &\approx f(g(Y)) + f'(g(Y))\sigma_q - x \\ &= Y + f'(g(Y))\sigma_q - x \\ &= a_1\sigma_1 + a_2\sigma_2 \cdot x + f'(g(Y))\sigma_q \end{aligned} \quad (16)$$

ここで逆レスポンスカーブが理想的であると仮定しており、カメラレスポンスカーブ g は、 f により打ち消される。つまり $f(g(Y)) = Y$ となる。これを露出時間 t で除算することによりノイズ項 n を得る。

$$n = (a_1\sigma_1 + a_2\sigma_2 \cdot x + f'(g(Y))\sigma_q) / t. \quad (17)$$

提案法では、重み関数は次式のようにノイズの逆関数

で定義する.

$$\begin{aligned}\omega_0(i) &= m(i) \cdot \frac{1}{((a_1\sigma_1 + a_2\sigma_2 \cdot x + f'(g(Y))\sigma_q)/t)} \\ &\approx m(i) \cdot \frac{1}{((a_1\sigma_1 + a_2\sigma_2 \cdot Y + f'(g(Y))\sigma_q)/t)} \\ &= m(i) \cdot \frac{1}{((a_1\sigma_1 + a_2\sigma_2 \cdot f(i) + f'(i)\sigma_q)/t)}\end{aligned}\quad (18)$$

$m(i)$ は画素値 i が 0 または 1 のときに 0, その他は 1 となるマスク関数である. ここでは未知の信号 x のかわりに $x \approx Y$ で近似する. 本論文は, 一般に市販されているカメラによる多重露光画像の統合を前提としており, 市販のカメラはカメラ内部で様々な画像処理が施されている. そこで, ほとんど処理されていない RAW 画像を用いても, キャリブレーションにより式 (18) のパラメータの推定は困難である. また, 特殊な装置を用いてカメラの内部処理前の正確なセンサー出力を測定することができたとしても, これらのパラメータの真値を正確に求めることは困難である. そこで, 式 (18) を次式で表す.

$$\omega(i) = m(i) \cdot \frac{1}{(b_1 + b_2 f(i) + b_3 f'(i))/t} \quad (19)$$

ここで, 分母の第 1 項と第 2 項はセンサーノイズ, 第 3 項は量子化ノイズを制御する項といえ, パラメータ b_1, b_2, b_3 を変化させることで, センサーノイズと量子化ノイズの除去効果に重みを付けることが可能となる. ただし, 画像に応じてパラメータを変化させることは実用的でないため, センサーノイズ, 量子化ノイズの軽減において多くのシーンで良好な結果を得られたパラメータを固定して用いる. 本論文の暗部ノイズの除去実験では, $b_1 = 0.001, b_2 = 0.99, b_3 = 0.01$ を用いる. b_3 に対して相対的に b_2 に大きい値を用いる

ことで, センサーノイズを軽減する重み関数となっている.

5. 実験及び考察

本章では提案法のアルゴリズムの妥当性を評価するため, 従来の統合手法との比較を行った. また, 提案法のノイズ除去性能の確認のため, 従来のノイズ除去法である, Bilateral Filter [14], 及び BM3D [15] との比較も併せて行った.

5.1 実験用画像及びパラメータ設定

本実験では CANON EOS 20D を用いて撮影した多重露光画像群を複数シーン用意した. これらの多重露光画像は絞り値を固定し, 露出時間のみを変化させており, 露光画像は低露光, 中露光, 高露光の 3 枚とした. 本論文は手ぶれしない ISO 感度設定での撮影を想定しているため, ISO 設定は ISO1600 とし, 本実験の評価の基準となる Ground truth 画像 (HDR 画像) の生成用に各露光 15 枚ずつ (各シーン, 合計 $3 \times 15 = 45$ 枚) 撮影した.

図 5 に, 本実験で用いたシーン (scene1, scene2) を例示する. 各シーンの多重露光画像の露光時間は, scene1 では低露光画像から 1/8sec, 1/2sec, 2sec, scene2 では, 1/50sec, 1/13sec, 1/3sec である. HDR 画像は非常に高いダイナミックレンジをもつことから, 既存のディスプレイでは表示することが難しく通常トーンマッピングによってダイナミックレンジ圧縮する必要がある. そのため, 本文で例示する HDR 画像は全てトーンマッピングを施した LDR 画像である.

本実験においては, 各露光 15 枚を平均した多重露光画像を従来の統合手法により統合し生成された HDR 画像を Ground truth とする. 高感度撮影した画像を従来の統合手法により統合し生成された HDR 画像 (Noisy image) にはノイズが発生している. そこで,



図 5 ダイナミックレンジ圧縮した HDR 画像 (scene1, scene2)
Fig. 5 Tonemapped HDR sample images (scene1, scene2).

従来のノイズ除去法である, Bilateral Filter [14], 及び BM3D [15] を Noisy image, つまり HDR 画像に対して適用し除去を行った結果との比較を行う. 従来法のパラメータはノイズ除去とエッジ保存, 二つの基準において最も良い画質になるように設定している. ただし, Bilateral Filter [14] に関しては, 弱い平滑化ではノイズ除去効果が少なかったため, エッジ保存よりもノイズ除去に重きを置いてパラメータを設定している. 提案法のパラメータは手動で調整し, HDR 画像をトーンマッピングした LDR 画像のノイズ除去とエッジ保存の効果が最も高くなる値を, PSNR, SSIM [17], LDR 画像の視覚評価を基に設定した. 式 (7) のしきい値は 0.0008 から 0.0020 の値を用いており, しきい

値判定により 0 となる割合は, 平均で約 33% となった. 式 (7) は従来の Hard Shrinkage [11] を用いているため同様のロバスト性があると考えられる.

5.2 定性評価

ここで, 図 6 に従来, 提案統合手法の結果と従来のノイズ除去処理によりノイズ除去を施した HDR 画像の結果を示す. ここでは, MATLAB の tonemap 関数によりトーンマッピング処理を施した LDR 画像を示している. このトーンマッピング関数は, Ward らのコントラストに制限を付けたヒストグラム均等化アルゴリズム [16] を基に設計されている. 図 6 より, 従来法の Bilateral Filter [14] ではどの画像でもノイズの軽減が確認できるが, 画像本来のエッジがボケて

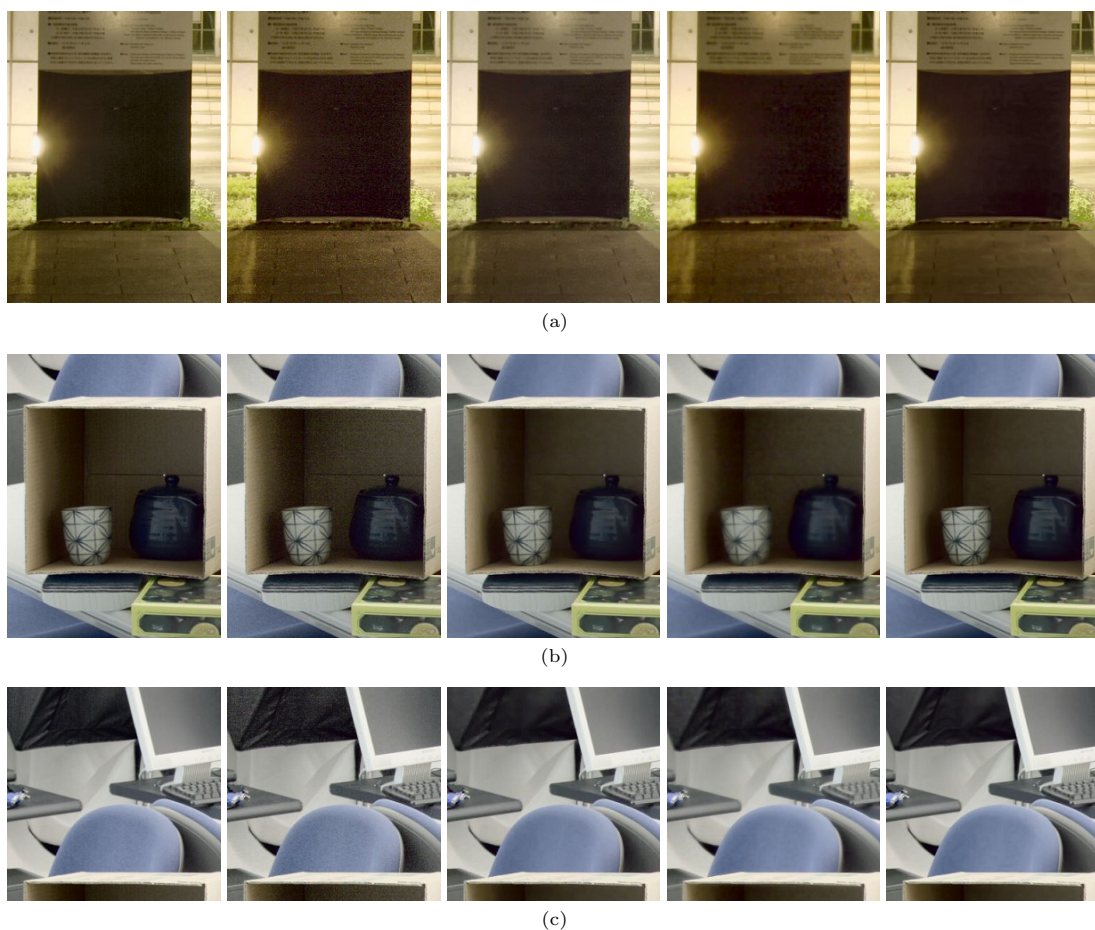


図 6 ダイナミックレンジ圧縮した HDR 画像: (左から) Ground truth, ノイズのある画像, 提案法, Bilateral Filter, BM3D

Fig. 6 Tonemapped HDR images: (from left to right) Ground truth, Noisy image, our method, Bilateral Filter, BM3D.

しまっている．もう一方の BM3D [15] では，明部はシャープになっているが，暗部では少し平滑化が効きすぎている．従来法に対して，提案法では，明部が少しぼやけてはいるものの，暗部の詳細は残されている結果となった．

図 6(a) の看板のフレア付近では，BM3D [15] はシャープだが，提案法では若干擬似エッジが残っている．しかし，看板付け根の草の詳細は，提案法の方が良く残されている．また，BM3D [15] の平滑化が効きすぎている暗部付近，特に看板の右側や地面で，BM3D [15] 特有の幾何学模様のような擬似エッジが確認できる．図 6(b) では，提案法と BM3D [15] の結果では差がほとんどない．しかし，図 6(c) の黒い照明カバーで，提案法では若干エッジの詳細が崩れている印象を受ける．これは，強いエッジ上，若しくは，付近に残っているノイズに関しては，shrinkage では十分に除去できないからだと考えられる．しかし，それ以外に関しては BM3D [15] と同等のノイズ軽減が実現できている．提案統合手法ではノイズ除去の効果が期待でき，特に暗部で良好な効果が得られる．

5.3 定量評価

提案法のノイズ除去性能の定量評価方法としてトーンマッピング後の画像に対する PSNR と，SSIM [17] による比較の二つの方法を用いた．PSNR は次式により算出している．

$$PSNR = 10 \cdot \log_{10} \left(\frac{MAX^2}{\sum (I_{ldr} - I'_{ldr})^2} \right) \quad (20)$$

ここで， I_{ldr} は Ground truth のトーンマッピング画像， I'_{ldr} は高感度撮影した多重露光画像を従来の統合手法，及び提案法により統合し生成された HDR 画像のトーンマッピング後の画像である．ここで，Bilateral Filter [14] 及び BM3D [15] は，Noisy image を処理したあとにトーンマッピングを行い評価している．本節では Ground truth 及び Noisy image の生成に用いた従来統合手法，そして提案統合手法ともに，4.2. のセンサーノイズを軽減するパラメータ設定の提案法の重み関数を用いそれぞれ HDR 画像を生成している．定量評価によりノイズ除去性能の比較を行う際，Noisy image，提案法ともに統合で用いる重みを Ground truth の生成に用いる重みと統一する必要がある．これは，統合で用いる重みが異なる場合，重みの値によって統合後の画素値の強度にノイズによる誤差以外の差が生じ，公平なノイズ除去性能のみの比較を

表 1 各トーンマッピング後の画像の PSNR 比較

Table 1 Value of PSNR, (R.L.T.: Reinhard Local Tonemap, R.G.T.: Reinhard Global Tonemap).

Tone-mapping		R.L.T.	R.G.T.	CLAHE	MATLAB
Noisy image	(scene1)	29.87	27.90	34.09	27.97
	(scene2)	38.83	38.45	31.95	35.68
our method	(scene1)	37.91	35.36	36.02	30.22
	(scene2)	43.49	41.51	33.80	37.49
Bilateral Filter	(scene1)	27.09	24.22	34.23	25.38
	(scene2)	36.90	36.55	35.09	36.03
BM3D	(scene1)	32.79	33.69	36.78	30.97
	(scene2)	38.74	40.29	34.10	37.40

表 2 各トーンマッピング後の画像の SSIM 比較

Table 2 Value of SSIM, (R.L.T.: Reinhard Local Tonemap, R.G.T.: Reinhard Global Tonemap).

Tone-mapping		R.L.T.	R.G.T.	CLAHE	MATLAB
Noisy image	(scene1)	0.993	0.993	0.981	0.989
	(scene2)	0.997	0.997	0.958	0.994
our method	(scene1)	0.997	0.996	0.984	0.993
	(scene2)	0.998	0.998	0.962	0.995
Bilateral Filter	(scene1)	0.986	0.985	0.975	0.983
	(scene2)	0.997	0.997	0.966	0.993
BM3D	(scene1)	0.996	0.995	0.983	0.993
	(scene2)	0.998	0.998	0.967	0.995

定量的に行うことができないためである．本実験で使用するトーンマッピングは，Reinhard らのローカル・グローバルトーンマッピング [18]，CLAHE を基にしたコントラスト強調トーンマッピング [19]，MATLAB の tonemap 関数である．結果を表 1，表 2 に示す．一般に HDR 画像はトーンマッピングを介して表示されることがほとんどであり，また，トーンマッピングが強調する暗部のノイズを低減することが本手法の目的であるため，LDR 画像のみの比較を行っている．なお，HDR 画像の SNR も併せて求めたが，SNR は高い数値が得られ本手法と従来法でほとんど違いが見られなかったためここでは割愛する．今回の実験で用いた画像では，センサーノイズに対し HDR 画像のダイナミックレンジが非常に高く，画像の相対エネルギーが大きくなったためだと考えられる．また，近年 HDR 画像の定量評価手法として注目を集めている HDR-VDP-2 [20] による評価も行ったが同じくほとんど違いが見られなかった．これは，HDR-VDP-2 [20] は暗部の微細なノイズの評価には適していないためだと考えられる．

表 1，表 2 より，提案法は Reinhard [18] らのトーンマッピングで従来法よりも良好な結果となっている．そこで，図 7 に Reinhard [18] らのトーンマッピング

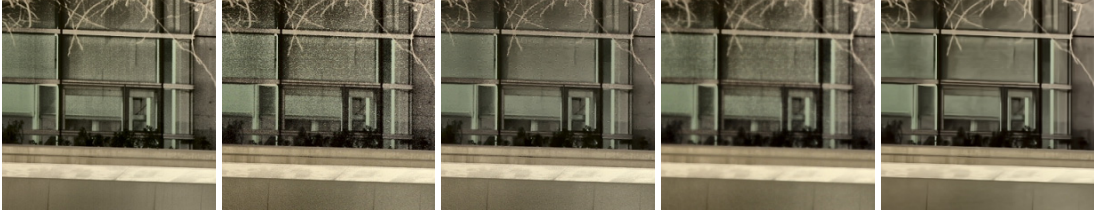


図 7 Reinhard [18] らのローカルトーンマッピングによりトーンマッピングされた HDR 画像: (左から) Ground truth, ノイズのある画像, 提案法, Bilateral Filter, BM3D

Fig. 7 Tonemapped HDR images by Reinhard Local Tonemap: (from left to right) Ground truth, Noisy image, our method, Bilateral Filter, BM3D.

結果を示す. 図 7 より, BM3D [15] では平滑化が効きすぎているため画像右側の壁のテクスチャの詳細がなくなっているのに対し, 提案法ではノイズを除去しつつ, テクスチャやエッジを保存できていることが分かる. また, Bilateral Filter [14] よりも鮮明な結果となっている. このトーンマッピングは, 高輝度になればなるほど受光感度が鈍感になるという人間の視覚特性を考慮し, 低輝度域の情報を残し, 人間が知覚しにくい高輝度側を強く圧縮するトーンマッピングとなっている. そのため, 暗部ノイズを強調する. 提案法は従来法に比べ画像の詳細を保持しつつ暗部ノイズを除去することができたため最も良好な結果となったと考えられる. CLAHE [19] と MATLAB のトーンマッピングは, 局所コントラストを強調し画像の微細な詳細情報を強調するため, 全輝度域で等しく平滑化された結果となった BM3D [15], Bilateral Filter [14] と, 提案法を比較したとき, 提案法は若干数値が劣る場合がある. これは, 提案法は暗部のノイズの軽減に重きを置いており, 高輝度域のノイズ除去をあまり考慮していないため高輝度側のノイズが強調され数値に影響したのではないかと考えられる. しかし, 高輝度側に生じるノイズに関しては, 人間の視覚特性を考慮した一般的なトーンマッピングの場合あまり問題とならない. また, 提案法では高輝度よりも暗部のノイズ除去に重きを置いていることから妥当な結果である.

まとめると, 提案法は, Reinhard [18] らのトーンマッピングでは, どの従来法よりも良好な結果を得ることができ, MATLAB のトーンマッピングでは, Bilateral Filter [14] よりも良好な結果で, BM3D [15] と同等の結果が得られた. CLAHE [19] のような微細な詳細情報を過度に強調するトーンマッピングでは, 提案法では対象としていない高輝度域のノイズが強調されるため, 従来法に劣る場合もある.

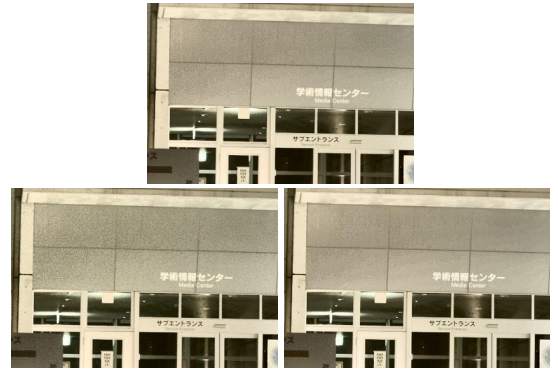


図 8 従来法 (Hat 関数) と提案法の重み関数の結果比較
Fig. 8 (upper) conventional method (Fig. 2 left: Hat function), (lower left) our method quantization noise suppression, (lower right) our method sensor noise suppression.

SSIM [17] に関しても, PSNR の結果と同様のことが言える. 提案法は, 複数の異なるシーンでも安定したノイズ除去結果を得ることができているため, ロバストな手法である.

5.4 重み関数の効果

次に, 従来法の重み付けと提案法の重み付けによる効果比較では, 高感度撮影した多重露光画像を各重み関数を用いて従来の統合手法により統合し, CLAHE [19] を基にしたトーンマッピングによりダイナミックレンジ圧縮を行った. このトーンマッピング手法は極端なコントラスト強調を行うため, センサーノイズや量子化ノイズが視認しやすくなる. 図 8 に結果を示す. 上図は従来の Hat 関数であり, 下図左, 下図右はそれぞれ, 量子化誤差, センサーノイズの除去に重きをおいた結果である. 量子化誤差を軽減する重み関数は, $b_1 = 0.001, b_2 = 0.01, b_3 = 0.95$ を, センサーノイズを軽減する重み関数は, $b_1 = 0.001, b_2 = 0.99, b_3 = 0.01$ をそれぞれ用いている. また, それぞれの

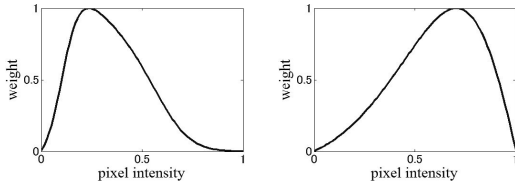


図 9 提案法の重み関数: (左図) 量子化誤差軽減, (右図) センサーノイズ軽減

Fig. 9 Proposed two weight functions: (left) quantization noise suppression, (right) sensor noise suppression.

提案法の重み関数を図 9 に示す.

図 8 より, 提案法の重み関数では, 暗部ノイズと量子化ノイズを選択的に軽減できていることが分かる. 図 8 下図左の量子化ノイズを軽減する重み関数の結果では, 白い外壁上で, 擬似エッジが見られないが, 粒子状の暗部ノイズが発生している. これに対して, 図 8 下図右の暗部ノイズを軽減する重み関数の結果では, 擬似エッジが見られるが粒子状の暗部ノイズは抑制されていることが分かる. よって, 提案法の重み関数は, 暗部ノイズと量子化ノイズを選択的に軽減が可能であることが確認できた. 二つのノイズの間には, トレードオフの関係があり, それゆえに統合時に重みの操作のみでそれらを同時に補正するのは難しい. そこで, 提案法では, 擬似エッジを好まないユーザに対しては, 量子化ノイズを軽減する重み関数を提供し, 粒子状の暗部ノイズを好まないユーザには, 暗部ノイズを軽減する重み関数を提供することで, 多くのユーザのニーズに対応することが可能と考える.

5.5 ノイズ耐性の評価

ノイズ耐性を評価するため, Ground truth の生成に用いたノイズレスな多重露光画像に対してガウス雑音を加え, この劣化した多重露光画像を入力とし各統合手法で HDR 画像を生成した. そして, Ground truth との誤差を測定した. ただし, センサーノイズよりも強く劣化するように標準偏差 0.05 のガウス雑音を用いている. このとき, 各統合手法ともに, 4.2. のセンサーノイズを軽減するパラメータ設定の提案法の重み関数を用いている. ここでは, NonlinearSNR (NSNR) と HDR-VDP-2 [20] により HDR 画像の誤差の評価を行う. NSNR は, 次式により定義される非線形の関数を各 HDR 画像に施し, 非線形化された HDR 画像の SNR を算出する手法である.

$$s(x) = L_{max}^2 \frac{x}{x + 12.6x^{0.63}} \quad (21)$$

表 3 ガウス雑音を加えた多重露光画像を用いた定量評価
Table 3 Quantitative evaluation using multiple exposure image with additional Gaussian noises (BRF: Bilateral Filter).

		Noisy	our	BRF	BM3D
NSNR	(scene1)	20.99	25.68	23.61	25.21
	(scene2)	23.72	30.41	27.80	30.09
HDR-VDP-2	(scene1)	-1.17	-1.29	-1.21	-1.29
	(scene2)	-0.84	-0.97	-0.88	-0.99

ここで, s は非線形関数であり, x は入力の画素値, L_{max} は表示するディスプレイの最大輝度である. 式 (21) は Daly が提案した人間の視覚特性をモデル化した関数 [21] である. また, HDR-VDP-2 [20] は Rafal らによって提案された HDR 画像を対象とした誤差検出手法である. この手法は人間の視覚特性を考慮した誤差検出が可能となっている.

表 3 に各評価手法の結果を示す. ここで, HDR-VDP-2 [20] では数値が低いと類似度が高いことを意味する. 表 3 より, 提案法は Bilateral Filter [14] よりも良好で, BM3D と同等の結果となった.

図 10 に各手法の HDR 画像の結果を示す. ここでは, Reinhard 等のグローバルトーンマッピング [18] によりトーンマッピング処理を施した LDR 画像を示している. 図 10 より, Bilateral Filter [14] では, ノイズを除去できているものの木の輪郭等の強いエッジを保存できているものの, 草などの細かいディテールが失われている. また, BM3D [15] では, ディテールを保ちつつノイズ除去を行えていることが確認できる. しかし, 草の細かい詳細情報は平滑化により失われている. BM3D [15] は, パッチ処理によりノイズ除去を行うため, 図 10 の草むらのようにエッジの変化が激しい領域では, パッチのオーバーラップによる平滑化によって細かなエッジの再現ができない. 従来法に対して提案法は, 草や木の表面等の細かな詳細をよく保存できているが, 高輝度領域で色むらが見られる. 劣化が大きい場合は, 暗部においては従来法よりも詳細を復元することが可能であるが, 高輝度領域においては若干ではあるが色むらが残る結果となった.

5.6 処理速度

本手法は MATLAB で実装されており, サイズ (1016 × 1592 × 3) の画像に対して処理時間は 1sec 以下である. 一方, BM3D [15] の実行時間は平均 36sec 程度である. 結果, 本手法はノイズ除去性能の非常に高い BM3D [15] と視覚的に同等のノイズ除去効果がより少ない計算時間で得られることがわかる. また,



図 10 Reinhard [18] らのグローバルトーンマッピングによりトーンマッピングされた HDR 画像: (左から) Ground truth, ノイズのある画像, 提案法, Bilateral Filter, BM3D

Fig. 10 Tonemapped HDR images by Reinhard Global Tonemap: (from left to right) Ground truth, Noisy image, our method, Bilateral Filter, BM3D.

Bilateral Filter [14] に関しては提案法よりも更に高速に実装できるが、ノイズ除去性能は提案法よりも劣る。

6. む す び

提案法では、極端な暗所をもつシーンから詳細を残したままノイズを大幅に軽減した HDR 画像を得ることが可能である。つまり、多重露光画像間でのノイズの軽減により各画像間の揺らぎを取り除き、また、提案法の重み関数によりセンサーノイズが生じている画素値に小さい重みを付け統合によるセンサーノイズの軽減効果を高め、そして、統合のための wavelet-shrinkage によりノイズの除去と統合を行うことで、本論文の目的であるノイズレスな HDR 画像を生成する多重露光画像統合法を示すことができた。ここで、提案法においてノイズ除去効果が最も高いのは wavelet-shrinkage による処理であるが、画像間のノイズ除去とセンサーノイズを軽減する重み付けを組み合わせることで、暗部の詳細を残しつつノイズ除去を実現することが可能となっている。

また、視覚評価では、従来法の BM3D [15] とほぼ同等のノイズ除去性能があると言える。処理時間に関しては、提案法は、複雑なマッチング処理を行っていないので、従来法 (BM3D [15]) に比べ高速な処理が可能である。

謝辞 本研究は科研費若手 (B)(研究課題番号: 20760244), JST・A-step 探索タイプ, KDDI 財団の助成を得た。ここに謝意を示す。

文 献

- [1] E. Reinhard, S. Pattanaik, G. Ward, and P. Debevec, High Dynamic Range Imaging: Acquisition, Display, and Image-Based Lighting (Morgan Kaufmann Series in Computer Graphics and Geometric Modeling), Morgan Kaufmann Publisher, 2005.
- [2] P.E. Debevec and J. Malik, “Recovering high dynamic range radiance maps from photographs,” Proc. SIGGRAPH 97, Computer Graphics Proceedings, pp.369–378, 1997.
- [3] S. Mann and R. Picard, “On being ‘undigital’ with digital cameras: Extending dynamic range by combining differently exposed pictures,” Proc. IS&T 46th Annual Conference (May 1995), pp.422–428, 1995.
- [4] T. Mitsunaga and S.K. Nayer, “Radiometric self calibration,” IEEE Conference on Computer Vision and Pattern Recognition (CVPR), vol.1, pp.374–380, June 1999.
- [5] S.W. Hasinoff, F. Durand, and W.T. Freeman, “Noise-optimal capture for high dynamic range photography,” 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.553–560, June 2010.
- [6] A. Akyuz and E. Reinhard, “Noise reduction in high dynamic range imaging,” J. Visual Communication and Image Representation, vol.18, pp.366–376, 2007.
- [7] M. Granados, B. Ajdin, M. Wand, C. Theobalt, H.-P. Seidel, and H. Lensch, “Optimal HDR reconstruction with linear digital cameras,” 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.215–222, June 2010.
- [8] T. Jinno and M. Okuda, “Motion blur free HDR image acquisition using multiple exposures,” IEEE International Conference on Image Processing, pp.1304–1307, Oct. 2008.
- [9] R.H. Chan, R.F. Chan, L. Shen, and Z. Shen, “Wavelet algorithms for high-resolution image reconstruction,” SIAM Journal on Scientific Computing, vol.24, pp.1408–1432, 2002.
- [10] T. Saito, N. Fujii, and T. Komatsu, “Iterative soft color-shrinkage for color-image denoising,” IEEE International Conference on Image Processing, pp.3837–3840, Nov. 2009.
- [11] M. Elad, “Why simple shrinkage is still relevant for redundant representations?,” IEEE Trans. Inf. Theory, vol.52, no.12, pp.5559–5569, Dec. 2006.
- [12] D.L. Donoho, “Denoising by soft thresholding,”

- IEEE Trans. Inf. Theory, vol.41, no.3, pp.613–627, May 1995.
- [13] T. Buades, Y. Lou, J.M. Morel, and Z. Tang, “A note on multi-image denoising,” International Workshop on Local and Non-local Approximation in Image Processing, pp.1–15, Aug. 2009.
- [14] C. Tomasi and R. Manduchi, “Bilateral filtering for gray and color images,” Proc. IEEE 6th Inter. Conf. Computer Vision, pp.839–846, 1988.
- [15] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, “Image denoising by 3D transform-domain collaborative filtering,” IEEE Trans. Image Process., vol.16, no.8, pp.2080–2095, Aug. 2007.
- [16] G.J. Ward, H. Rushmeier, and C. Piatko, “A visibility matching tone reproduction operator for high dynamic range scenes,” IEEE Trans. Vis. Comput. Graphics, vol.3, no.4, pp.291–306, Dec. 1997.
- [17] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli, “Image quality assessment: From error visibility to structural similarity,” IEEE Trans. Image Process., vol.13, no.4, pp.600–612, 2004.
- [18] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, “Photographic tone reproduction for digital images,” ACM Trans. Graphics, vol.21, no.3, pp.267–276, July 2002 (Proc. SIGGRAPH 2002).
- [19] K. Zuiderveld, “Contrast limited adaptive histogram equalization,” Graphic Gems IV, pp.474–485, Academic Press Professional, San Diego, 1994.
- [20] R. Mantiuk K.J. Kim, A.G. Rempel, and W. Heidrich, “HDR-VDP-2: A calibrated visual metric for visibility and quality predictions in all luminance conditions,” ACM Trans. Graphics, vol.30, pp.1–14, July 2011.
- [21] S. Daly, “The visible differences predictor: An algorithm for the assessment of image fidelity,” in Digital Image and Human Vision, ed. A.B. Watson, pp.179–206, MIT Press, Cambridge, MA, 1993.

付 録

ここでは、本文式 (14) の導出について述べる。

まず、式 (13) のコスト関数 $E(h)$ は $h = 0$ で不連続である。 $h \neq 0$ のとき、

$$E(h) = 1 + \frac{\lambda}{N} \sum_{n=1}^N (h - l(n))^2 \quad (\text{A} \cdot 1)$$

となり、解は $h^* = \frac{1}{N} \sum_{n=1}^N l(n)$ となる。

(i) $h = 0$ のとき、

$$E(0) = \frac{\lambda}{N} \sum_{n=1}^N l(n)^2 \quad (\text{A} \cdot 2)$$

となり、

(ii) $h = \frac{1}{N} \sum_{n=1}^N l(n)$ のとき、

$$\begin{aligned} E(h) &= 1 + \frac{\lambda}{N} \sum_{n=1}^N \left(\left(\frac{1}{N} \sum_{k=1}^N l(k) \right)^2 \right. \\ &\quad \left. - 2 \left(\frac{1}{N} \sum_{k=1}^N l(k) \right) l(n) + l(n)^2 \right) \\ &= 1 - \lambda \left(\frac{1}{N} \sum_{n=1}^N l(n) \right)^2 + \frac{\lambda}{N} \sum_{n=1}^N l(n)^2 \end{aligned} \quad (\text{A} \cdot 3)$$

となる。よって、最適解 h^* は次式によって与えられる。

$$h^* = \begin{cases} 0, & \text{if } 1 - \lambda \left(\frac{1}{N} \sum_{n=1}^N l(n) \right)^2 > 0 \\ \frac{1}{N} \sum_{n=1}^N l(n), & \text{otherwise} \end{cases} \quad (\text{A} \cdot 4)$$

(平成 25 年 5 月 17 日受付, 9 月 30 日再受付)



松岡 諒 (学生会員)

2013 北九州市立大学大学院国際環境工学研究科情報工学専攻博士前期課程了。現在同大学院博士後期課程に在学中。



神納 貴生 (正員)

2011 北九州市立大学大学院国際環境工学研究科情報工学専攻博士後期課程了。工博。現在、豊橋技術科学大学大学院工学研究科助教。



奥田 正浩 (正員)

平 5 慶大・理工・電気卒。平 10 同大学院博士課程了。工博。平 13 北九州市立大学・助教授, 平 23 同教授。信号処理, 画像処理に関する研究に従事。