

LDR 画像の学習に基づく CNN を用いた逆トーンマッピング法

木下 裕磨[†] 貴家 仁志[†]

[†] 首都大学東京大学院システムデザイン研究科 〒191-0065 東京都日野市旭ヶ丘 6-6
E-mail: [†]kinoshita-yuma@ed.tmu.ac.jp, ^{††}kiya@tmu.ac.jp

あらまし 本研究では、任意の単一 LDR 画像から高品質な HDR 画像の生成することを目的に、LDR 画像の学習に基づく CNN を用いた逆トーンマッピング法を提案する。CNN を用いた逆トーンマッピングにおいて、教師画像として HDR 画像を直接利用して CNN を学習することは、LDR 画像と HDR 画像におけるダイナミックレンジの違いや人間の非線形な視覚特性などの理由から、困難であることが既存の研究により指摘されている。そこで、提案する逆トーンマッピング法では、入力 LDR 画像から Reinhard のグローバルオペレータによりマッピングされた LDR 画像への変換を学習する。Reinhard のグローバルオペレータは逆変換が可能な関数である。提案法は、Reinhard のグローバルオペレータによってマッピングされたとみなすことのできる LDR 画像を入力 LDR 画像から生成し、生成された LDR 画像に対して Reinhard のグローバルオペレータの逆変換を行うことで HDR 画像を生成する方法である。HDR-VDP-2.2, および PU encoding + MS-SSIM を用いた客観評価実験により、提案法が他の手法と比較して高品質な HDR 画像を生成できることが示される。

キーワード 逆トーンマッピング, 高ダイナミックレンジ画像, 畳み込みニューラルネットワーク, 深層学習, 画像強調

An Inverse Tone Mapping Operation Using CNN with LDR Based Learning

Yuma KINOSHITA[†] and Hitoshi KIYA[†]

[†] Department of Information and Communication Systems,
Tokyo Metropolitan University, 191-0065, Tokyo, Japan
E-mail: [†]kinoshita-yuma@ed.tmu.ac.jp, ^{††}kiya@tmu.ac.jp

Abstract This paper proposes an inverse tone mapping operation using CNN with LDR based learning. In inverse tone mapping with CNNs, it is difficult to train CNNs by directly using HDR images as training data. In the proposed method, CNNs are trained with only LDR images, namely, HDR images are not used in training. The proposed CNNs learn a transformation from various input LDR images to LDR images mapped by Reinhard's global operator. Since Reinhard's global operator is invertible, HDR images can be reconstructed from LDR images mapped by the operator. For this reason, the proposed method enables us to generate high quality HDR images from various LDR images by transforming the input LDR ones via CNNs. Experimental results show that HDR images generated by the proposed method have higher quality than HDR ones generated by conventional inverse tone mapping methods, in terms of HDR-VDP-2.2 and PU encoding + MS-SSIM.

Key words Inverse Tone Mapping, High Dynamic Range Image, Convolutional Neural Network, Deep Learning, Image Enhancement

1. はじめに

撮像センサにおける輝度のダイナミックレンジは、現実シーンにおけるダイナミックレンジよりもはるかに狭い。そのため、現在のデジタルカメラによる撮影では、人間が知覚している輝度のダイナミックレンジを忠実に表現するには至って

いない。低ダイナミックレンジ (Low Dynamic Range, LDR) 画像, あるいは標準ダイナミックレンジ (Standard Dynamic Range, SDR) 画像が持つこれらの問題を解決するため、高ダイナミックレンジ (High Dynamic Range, HDR) 画像への関心がますます高まっている。

デジタルカメラにより撮影された単一 LDR 画像から HDR 画像を獲得するため、逆トーンマッピング法の研究が数多く実施されている [1, 2]. これら研究の多くは、特定の逆トーンマッピングオペレータを用いて LDR 画像のダイナミックレンジを拡張することで、HDR 画像を生成する方式の研究である. ところが、逆トーンマッピングは一般に不良設定問題であり、カメラの非線形な応答特性を表すカメラ応答関数 (Camera Response Function, CRF) などの事前情報が利用できない場合には、高品質な HDR 画像の推定が困難である. そのため、単一 LDR 画像から高品質な HDR 画像を生成するために、深層学習を用いた逆トーンマッピング法の開発が期待されている.

このような理由から、畳み込みニューラルネットワーク (Convolutional Neural Networks, CNN) を利用した逆トーンマッピング法がいくつか提案されている [3, 4]. しかし、CNN を用いた場合でさえ、HDR 画像を教師画像として LDR 画像から HDR 画像への変換を直接学習することには大きな困難を伴う. これは、LDR 画像と HDR 画像における画素値のレンジが大きく異なることや人間の非線形な視覚特性などの理由から、平均二乗誤差などの一般的な損失関数が HDR 画像の学習に適していないためである. そのため、明るさの異なる複数枚の LDR 画像 (多重露出画像) を用いて CNN を学習する手法 [3] や、HDR 画像を教師画像として利用するための新しい損失関数を用いて CNN を学習する手法 [4] が提案されている. しかしながら、これらの手法には、HDR 画像の生成にかかる計算コストが大きいという課題がある上に、生成される HDR 画像の品質も十分であるとはいえない.

このような背景から、本研究では、任意の入力 LDR 画像から高品質な HDR 画像を生成可能な、LDR 画像の学習に基づく CNN を用いた新しい逆トーンマッピング法を提案する. 本研究の着眼は、様々な入力 LDR 画像から Reinhard のグローバルオペレータにより生成された LDR 画像へ変換することにある. この実現のため、提案法における CNN は、HDR 画像を直接学習に使用せず、任意の入力 LDR 画像から Reinhard のグローバルオペレータにより生成された LDR 画像への変換を学習する. これらの入出力画像間にはダイナミックレンジの差がないことから、HDR 画像を直接学習する場合に比べ、容易に学習可能である. 加えて、Reinhard のグローバルオペレータは逆変換が可能な関数であり、これにより生成された LDR 画像は、HDR 画像へ高精度に変換できる. このことから、提案する CNN によって入力 LDR 画像から生成された LDR 画像に対し、Reinhard のグローバルオペレータの逆関数を適用することにより、高品質な HDR 画像を生成できる.

提案する逆トーンマッピング法の有効性を評価するため、HDR-VDP-2.2 および PU encoding + MS-SSIM という 2 つの客観評価尺度を用いる画質評価実験を実施した. 実験の結果、提案法が従来の逆トーンマッピング法に比べ高品質な HDR 画像を生成できることが、双方の評価尺度により示される.

2. 準備

ここでは、HDR 画像の生成、および畳み込みニューラルネッ

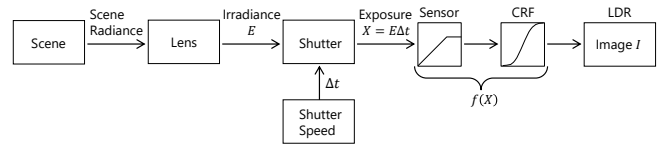


図 1 Imaging pipeline of digital camera

トワークについて簡潔に要約する. また議論の簡単化のために、特に言及しない限りグレースケール画像を仮定する.

2.1 HDR 画像の生成

HDR 画像は、デジタルカメラのセンサに照射される光の放射照度 E を直接記録した画像である. しかしながら、従来のデジタルカメラでは、センサのダイナミックレンジの制限とカメラが持つ非線形な応答 (CRF) により、放射照度 E を直接記録できない.

図 1 に、デジタルカメラを用いた代表的な画像撮影の流れを示す [5]. センサに照射される光は、シャッターが開いている時間 Δt の間累積される. この累積される光の、単位面積当たりのエネルギーを露出 X と呼ぶ. 静的なシーンであれば、露出 X は、放射照度 E とシャッタースピード Δt を用いて、以下のように表される.

$$X_{i,j} = E_{i,j} \Delta t \quad (1)$$

ここで、 $E_{i,j}$ と $X_{i,j}$ は、それぞれ画素 $(i, j) \in P$ における放射照度および露出を表し、 P は画素全体の集合を表す. 撮影される LDR 画像 I の画素値 $I_{i,j} \in [0, 1]$ は、露出 X を用いて以下のように与えられる.

$$I_{i,j} = f(X_{i,j}) \quad (2)$$

ここで、 $f: \mathbb{R} \rightarrow [0, 1]$ は、CRF とセンサの飽和を表す関数の合成関数である.

デジタルカメラにより撮影された LDR 画像 I から HDR 画像 E を獲得するため、これまでに数多くの研究が実施されている. なかでも代表的な手法は、シャッタースピード Δt を変化させながら撮影された複数枚の LDR 画像 (多重露出画像) を合成する多重露出画像合成法である. 多重露出画像合成法は、関数 f の高精度な推定が可能である. しかしながら、動く被写体を撮影する場合やカメラが固定されていない場合など、高品質な多重露出画像の撮影が困難な状況では、合成画像にゴーストと呼ばれるアーティファクトが発生してしまう問題がある. 加えて、多重露出画像合成法により生成される HDR 画像は、放射照度 E に対して線形な信号を持つものの、絶対的に E と等しいとは限らない. したがって、生成される HDR 画像はユーザによるキャリブレーションを必要とする.

一方、単一の LDR 画像のみを用いて HDR 画像生成を行う逆トーンマッピング法の研究も多く行われている. 逆トーンマッピング法は多重露出画像を必要としないことから、原理的にゴーストが発生しない. ところが、逆トーンマッピングは一般に不良設定問題であり、関数 f などの事前情報が利用できない場合には高精度な HDR 画像の推定が困難である. そのため、単一 LDR 画像から高品質な HDR 画像を生成するために、深層学習を用いた逆トーンマッピング法の開発が期待されている.

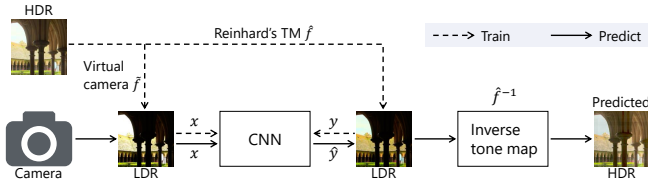


図2 Proposed inverse tone mapping scheme

2.2 畳み込みニューラルネットワーク

CNNは、ニューラルネットワークにおける各層の結合および重みに制約を加えることで畳み込み演算を実行する、畳み込み層を用いたニューラルネットワークである。第 l 層の畳み込み層における、 (i, j, c) 番目のユニットの出力 $z_{i,j,c}^{(l)}$ は、ピアス $b^{(l)}$ とカーネルサイズ $H \times W \times C$ を持つ K 個のフィルタ $\omega^{(l)}$ を用いて、以下の式で表される。

$$z_{i,j,c}^{(l)} = \sigma(u_{i,j,c}^{(l)}), \quad (3)$$

$$u_{i,j,k}^{(l)} = \sum_{c=0}^{C-1} \sum_{w=0}^{W-1} \sum_{h=0}^{H-1} z_{si+h, sj+w, c}^{(l-1)} \omega_{h,w,c,k}^{(l)} + b_{i,j,k}^{(l)} \quad (4)$$

ここで、 $z^{(l-1)}$, σ , s はそれぞれ、第 $l-1$ 層のユニットの出力、活性化関数、およびストライドを表す。フィルタのチャンネル数 C は前層のフィルタ数 K と等しいため、これ以降、カーネルサイズは $H \times W$ のように表記する。

画像生成の分野においては、CNNの中でもエンコーダ・デコーダモデルが一般的に採用される。エンコーダは、各層のユニット数が徐々に減少するように畳み込み層とプーリング層を繰り返し配置したネットワークである。代表的なプーリング層の一つである最大プーリング層は、以下の式で与えられる。

$$u_{i,j,k}^{(l)} = \max_{(i',j') \in P_{si,sj}} z_{i',j',k}^{(l-1)} \quad (5)$$

ここで、 $P_{si,sj} = \{(si+h, sj+w) | 0 \leq h \leq H-1, 0 \leq w \leq W-1\}$ である。一方デコーダは、各層のユニット数が徐々に増加するように、畳み込み層と転置畳み込み層を繰り返し配置したネットワークである。転置畳み込み層における演算は、前層の出力に対してゼロ挿入による拡大率 U のアップサンプリング処理を行い、その後、畳み込み演算を行うことに相当する。エンコーダ・デコーダモデルの中でも特に、U-Net [6]は、その有効性から様々な画像生成タスクに対して利用されている。

3. 新しい逆トーンマッピング法の提案

図2に、提案する逆トーンマッピング法の枠組みを示す。提案法におけるCNNは、HDR画像を直接学習に使用せず、入力LDR画像 x とReinhardのグローバルオペレータにより生成されたLDR画像 y との関係を学習する。

3.1 Reinhardのグローバルオペレータ

Reinhardのグローバルオペレータは、代表的なトーンマッピング法の一つであり、HDR画像 E からLDR画像 I を以下の通り生成する [7]。

$$I_{i,j} = \hat{f}(X_{i,j}), \quad (6)$$

$$\hat{f}(X_{i,j}) = \frac{X_{i,j}}{1 + X_{i,j}} \quad (7)$$

ただし、 $X_{i,j}$ は、2つのパラメータ a と $G(E)$ を用いて、

$$X_{i,j} = \frac{a}{G(E)} E_{i,j} \quad (8)$$

と与えられる。ここで、 a は出力画像の明るさを定めるパラメータであり、 $G(E)$ は E の幾何平均

$$G(E) = \exp \left(\frac{1}{|P|} \sum_{(i,j) \in P} \log(\max(E_{i,j}, \epsilon)) \right) \quad (9)$$

を表すパラメータである。また、 ϵ は、特異点を回避するための微小な正の実数である。

Reinhardのグローバルオペレータは、2つのパラメータが既知の時、逆変換が可能である。この逆変換は、

$$E'_{i,j} = \frac{G(E)}{a} X'_{i,j}, \quad (10)$$

$$X'_{i,j} = \hat{f}^{-1}(I_{i,j}) = \frac{I_{i,j}}{1 - I_{i,j}} \quad (11)$$

として与えられる。さらに、パラメータ $G(E)$ は、トーンマッピングされたLDR画像から、

$$G(E) = \exp \left(\frac{|P|}{|P_B|} \log \bar{X}' - \frac{|P_B|}{|P_B|} \log a \right) \quad (12)$$

として計算できる [2]。ここで、 $P_B = \{(i,j) \in P | I_{i,j} = 0\}$ であり、 $\bar{P}_B = \{(i,j) \in P | I_{i,j} \neq 0\}$ である。したがって、計算における有限語長の影響がない場合には、1つのパラメータ a のみを用いて、Reinhardのグローバルオペレータにより生成されたLDR画像からHDR画像を完全に再構成可能である。

しかし、一般のLDR画像は、Reinhardのグローバルオペレータにより生成されていない。本研究の着眼は、CNNにより一般のLDR画像をReinhardのグローバルオペレータにより生成されたLDR画像へ変換することにある。

3.2 提案法の枠組み

ここでは、提案法におけるCNNの学習の概略と、提案する逆トーンマッピング法を用いてLDR画像からHDR画像を生成する手順を示す (図2参照)。

学 習

CNNの学習は以下の通り実施する。

- i 仮想カメラを用いて、様々なCRFの下でHDR画像から入力LDR画像 x を生成する。このことは、CRFが未知である任意のLDR画像を入力画像として想定することに相当する。
- ii 式(6)から式(9)に従い、Reinhardのグローバルオペレータとパラメータ $a = 0.18$ を用いて、HDR画像から教師LDR画像 y を生成する。
- iii 入力LDR画像 x と教師LDR画像 y を利用して、CNNにより x から y への変換を学習する。

より詳細な学習条件は3.4節に記す。

実 行

学習済みのCNNを用いてLDR画像からHDR画像を生成する手順は以下の通りである。

- i カメラ等により撮影されたLDR画像を x とする。
- ii 学習されたCNNを用いて、入力LDR画像 x を \hat{y} に変換する。ここで、 \hat{y} は、Reinhardのグローバルオペレータにより生成されたLDR画像に相当する。

iii 式 (10) から式 (12) に従い, \hat{y} から HDR 画像を生成する. ここで, 逆トーンマッピングにおけるパラメータ a は 0.18 とする.

前節の通り, Reinhard のグローバルオペレータにより生成された LDR 画像から HDR 画像を生成するためには, パラメータ a が必要である. そのため提案法では, CNN の学習時におけるすべての教師画像 y を, 固定のパラメータ $a = 0.18$ を用いて生成する. ここで, $a = 0.18$ と設定することは, 画像の平均的な明るさを反射率 18% グレーと等しくすることを意味する [7]. これにより, CNN を用いて生成される画像 \hat{y} は, パラメータ $a = 0.18$ を用いてマッピングされた LDR 画像に相当する. したがって, 逆トーンマッピングを実行する際にも, すべての画像に対し $a = 0.18$ を利用することができる.

3.3 CNN の構成

本研究で用いる CNN は, U-Net [6] を参考に設計した (図 3 参照). この CNN は, RGB 値を画素値として持つ 512×512 画素の LDR 画像を入力とする.

畳み込みブロック (Conv. block) は, カーネルサイズが 3×3 のフィルタを持つ 2 つの畳み込み層からなる. これら 2 つの畳み込み層におけるフィルタ数 K は等しい. 各畳み込みブロックにおけるフィルタ数 K は入力側からそれぞれ, 32, 64, 128, 256, 512, 1024, 512, 256, 128, 64, 32 とする. 一方, 転置畳み込み層におけるカーネルサイズは 4×4 であり, これら各層のカーネル数 K は, 入力側から 512, 256, 128, 64, 32 である. さらに, 各転置畳み込み層の拡大率 U はすべて 2 である. 畳み込みブロックにおける各層, および各転置畳み込み層のストライド s は 1 であり, 活性化関数 σ には次式で与えられる Rectified Linear Unit (ReLU) [8] を用いる.

$$\sigma(u) = \max(u, 0) \quad (13)$$

加えて, これら各層では ReLU を適用したのち, Batch normalization [9] を適用する.

最大プーリング層は, カーネルサイズ 2×2 およびストライド s を 2 とする. 出力層直前に配置された畳み込み層は, カーネルサイズ 3×3 のフィルタ 3 つからなり, 活性化関数としてシグモイド関数

$$\sigma(u) = \frac{1}{1 + e^{-u}} \quad (14)$$

を利用する.

3.4 CNN の学習

提案法における CNN の学習のためには, 様々な条件で撮影された LDR 画像と Reinhard のグローバルオペレータによりマッピングされた LDR 画像の組が必要である. 一方, 実際のカメラを用いて十分な量の学習用画像を収集することは難しい. この理由から, 仮想カメラ [4] を用いて入力 LDR 画像 x を, Reinhard のグローバルオペレータを用いて教師 LDR 画像 y を HDR 画像 E からそれぞれ生成し, これら x, y を CNN の学習に利用する. 学習用画像として, オンライン上で利用可能なデータベース [10–15] から 978 枚の HDR 画像を収集した.

CNN の学習手順は以下の通りである.

- i 978 枚の HDR 画像の中からランダムに 8 枚の画像を選択する.
- ii 以下の手順に従い, 各 HDR 画像から入力画像 x および教師画像 y を 1 枚ずつ生成し, 入力画像と教師画像の組を計 8 組得る.
 - (a) HDR 画像 E から, $N \times N$ 画素の領域 \tilde{E} を切り出す. ここで, N は, 区間 $[0.2, 0.6]$ 上の一様乱数と, 各画像の短辺の長さの積として与える. また, 領域を切り出す位置は一様な乱数に従い決定する.
 - (b) \tilde{E} を 512×512 画素にリサイズする.
 - (c) 手順 (b) で得られた \tilde{E} を, 50% の確率で左右反転させる.
 - (d) 手順 (c) で得られた \tilde{E} を, 50% の確率で上下反転させる.
 - (e) 手順 (d) で得られた \tilde{E} から, 露出 X を計算する. ただし, シャッタースピード Δt は, 区間 $[-4, 4]$ の一様乱数 v を用いて $\Delta t = 0.18 \cdot 2^v / G(\tilde{E})$ として計算する.
 - (f) 計算された露出 X を用いて, 次式の関数 f を持つ仮想カメラにより, 入力画像 x を生成する.

$$\tilde{f}(X) = \max\left((1 + \eta) \frac{X^\gamma}{X^\gamma + \eta}, 1\right) \quad (15)$$
 ただし, η, γ は, 平均 0.6 分散 0.1 の正規分布および平均 0.9 分散 0.1 の正規分布にそれぞれ従う.
 - (g) Reinhard のグローバルオペレータ (式 (6) から式 (9) 参照) を用いて, 手順 (d) で得られた \tilde{E} から教師画像 y を生成する. ただし, パラメータ a として 0.18 を用いる.
- iii CNN により, 8 枚の入力画像 x から 8 枚の予測画像 \hat{y} をそれぞれ得る.
- iv 平均二乗誤差を用いて, 予測画像 \hat{y} と教師画像 y 間の損失を計算する.
- v 誤差逆伝播法によりフィルタ ω およびバイアス b を更新する.

本研究においては, 上記手順を 122 回繰り返したものを 1 エポックとし, 5000 エポックの学習を行った. ただし, 手順 i において, 各エポック中で同一の HDR 画像が複数回選択されることはない.

CNN の初期化には He らの手法 [16] を用いた. また, 誤差逆伝播法における最適化アルゴリズムには, Adam (Adaptive Moment Estimation) [17] を利用した. ただし, Adam におけるパラメータは $\alpha = 0.002, \beta_1 = 0.5, \beta_2 = 0.999$ とした.

4. 実験

提案する逆トーンマッピング法の有効性を評価するため, 客観評価尺度を用いた画質評価実験を実施した.

4.1 実験条件

本実験では, 3.3 に示す手順 ii(a) から ii(f) に従い, HDR 画像 \tilde{E} から生成された 512×512 画素の LDR 画像 x を入力とし, 提案法により HDR 画像 \tilde{E}' を生成した. ここで, 手順 ii(e)

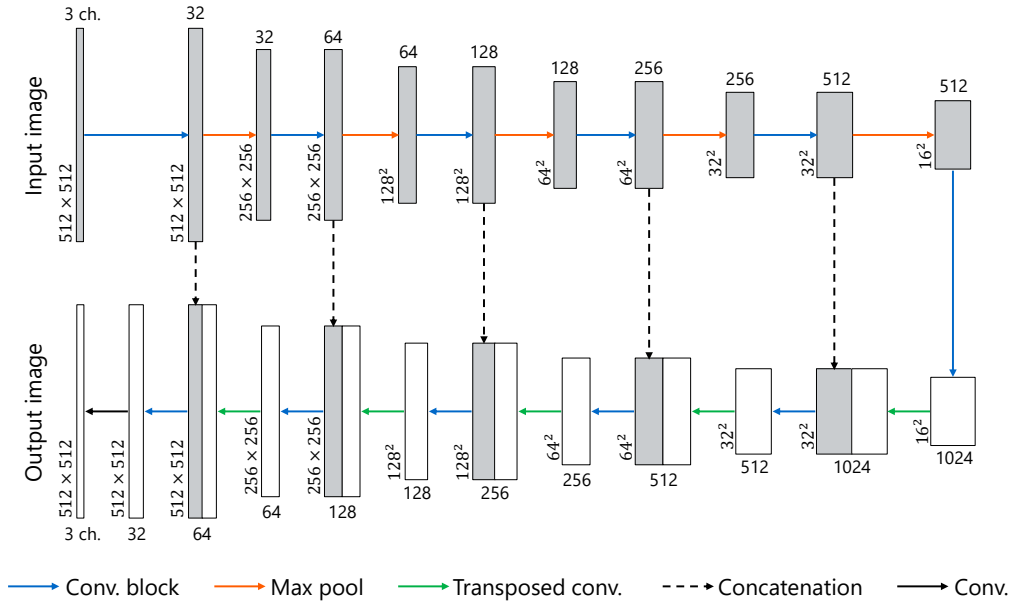


図3 Network architecture

および ii(f) の処理は画像の輝度に対して行い、HDR 画像と生成される LDR 画像間の RGB 値の比が等しくなるように変換した。実験で用いた HDR 画像は、CNN の学習に利用されていない 7 枚の画像を選択した。

生成された HDR 画像 \tilde{E}' の品質は、原 HDR 画像 \tilde{E} を参照画像として、HDR 画像の品質評価のための客観評価尺度である HDR-VDP-2.2 [18]、および PU encoding [19] と MS-SSIM [20] を組み合わせた尺度により評価した。これらの尺度が HDR 画像品質の評価に適していることは、文献 [21] により確かめられている。

また、比較手法として、式 (10) から (12) に従い入力画像 x を直接逆トーンマッピングしたもの (Direct ITMO) [2]、機械学習を利用しない従来の逆トーンマッピング法である Pseudo-Multi-Exposure-Based Tone Fusion (PMET) [1]、CNN を利用した逆トーンマッピング法である Deep Reverse Tone Mapping (DrTMO) [3] の 3 つを用いた。

4.2 実験結果

図 4 および 5 に、各手法により生成された HDR 画像の例を示す。ただし、HDR 画像は通常の LDR 機器に直接表示できないため、トーンマッピングを施している。

図 4 より、提案する逆トーンマッピング法は他の手法と比較して、原 HDR 画像 \tilde{E} と近い高品質な画像を生成できていることが確認できる。一方、DrTMO により生成された画像には、不自然な色歪みが発生している。これは、DrTMO が極端に明るい多重露出画像を生成したことに起因する。一方、図 5 では、提案法および DrTMO により生成された画像が低輝度領域をよく表現できているのに対して、Direct ITMO および PMET により生成された画像は低輝度領域が不明瞭である。したがって、提案法は他の手法と比較して高品質に HDR 画像を生成できるといえる。

HDR-VDP および PU encoding + MS-SSIM を用いて各手法により生成された HDR 画像を評価した結果を、表 1 および 2

表 1 HDR-VDP-2.2 scores

	Direct ITMO [2]	PMET [1]	DrTMO [3]	Proposed
Image 1	32.89	32.89	47.92	50.94
Image 2	31.19	31.19	31.41	35.17
Image 3	28.57	28.57	45.83	44.86
Image 4	56.93	49.94	48.47	56.39
Image 5	39.38	39.38	34.33	48.24
Image 6	33.06	33.06	48.26	57.61
Image 7	44.77	44.77	43.82	61.49

表 2 PU encoding + MS-SSIM scores

	Direct ITMO [2]	PMET [1]	DrTMO [3]	Proposed
Image 1	0.6399	0.7750	0.4912	0.8828
Image 2	0.0599	0.0641	0.2382	0.8059
Image 3	0.7444	0.7393	0.5566	0.8502
Image 4	0.8445	0.5893	0.6624	0.9143
Image 5	0.4358	0.5734	0.1729	0.6641
Image 6	0.9156	0.7982	0.4083	0.9696
Image 7	0.5906	0.6997	0.6042	0.9804

にそれぞれ示す。HDR-VDP による評価では、提案法が 7 枚中 5 枚の画像で最も高いスコアを示した。さらに、PU encoding + MS-SSIM を用いた評価では、提案法が全画像で最も高いスコアを示した。すなわち、CNN を用いた最新の手法を含む他の手法と比較して、提案法は高精度に HDR 画像を復元できるといえる。

以上の結果より、提案法は、様々な単一 LDR 画像から高品質な HDR 画像を生成できることが示された。

5. おわりに

本研究では、LDR 画像の学習に基づく CNN を用いた新しい逆トーンマッピング法を提案した。提案法における CNN は、学習に直接 HDR 画像を利用せず、任意の LDR 画像と Reinhard のグローバルオペレータによりマッピングされた LDR 画像間の変換を学習する。さらに、Reinhard のグローバルオペレータの逆変換により、提案する CNN を用いて変換された LDR

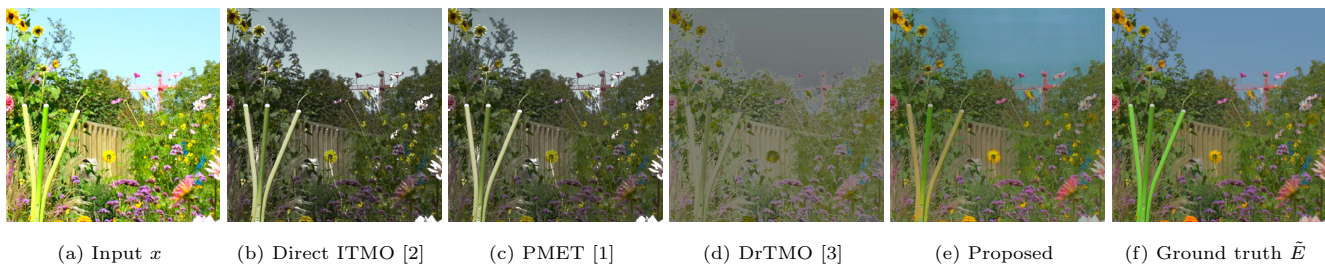


図 4 Experimental Results (Image 1)

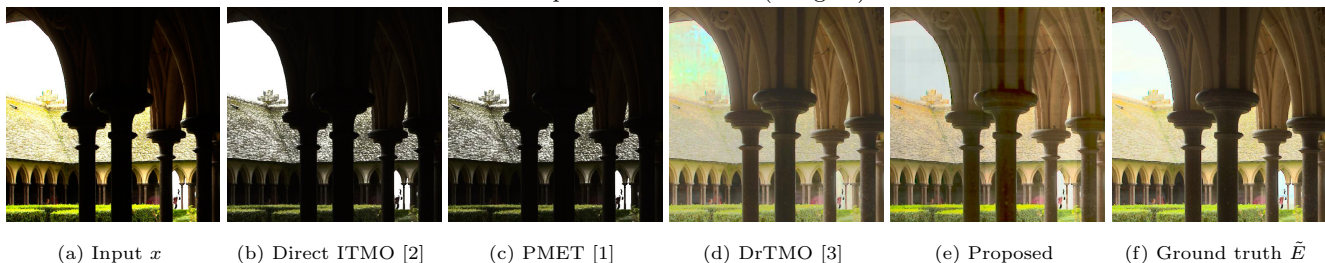


図 5 Experimental Results (Image 5)

画像から HDR 画像を生成できる。提案する逆トーンマッピング法は、学習された CNN 以外の事前情報無しに、単一 LDR 画像から高品質な HDR 画像の生成を可能とする。

客観評価尺度を用いた画質評価実験の結果、提案法は他の手法と比較して高品質な HDR 画像を生成できることが、HDR-VDP-2.2, および PU encoding と MS-SSIM を組み合わせた尺度の双方により確認された。

謝辞 本研究は JSPS 科研費 JP18J20326 の助成を受けたものである。

文 献

- [1] T.-H. Wang, C.-W. Chiu, W.-C. Wu, J.-W. Wang, C.-Y. Lin, C.-T. Chiu, and J.-J. Liou, "Pseudo-multiple-exposure-based tone fusion with local region adjustment," *Multimedia, IEEE Transactions on*, vol.17, no.4, pp.470–484, 2015.
- [2] Y. Kinoshita, S. Shiota, and H. Kiya, "Fast inverse tone mapping based on reinhard's global operator with estimated parameters," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol.100, no.11, pp.2248–2255, 2017.
- [3] Y. Endo, Y. Kanamori, and J. Mitani, "Deep reverse tone mapping," *ACM Transactions on Graphics (Proc. of SIGGRAPH ASIA 2017)*, vol.36, no.6, pp.177:1–177:10, Nov. 2017.
- [4] G. Eilertsen, J. Kronander, G. Denes, R. Mantiuk, and J. Unger, "HDR image reconstruction from a single exposure using deep CNNs," *ACM Transactions on Graphics (TOG)*, vol.36, no.6, pp.178:1–178:15, 2017.
- [5] F. Dufaux, P.L. Callet, R. Mantiuk, and M. Mrak, *High Dynamic Range Video, From Acquisition, to Display and Applications*, Elsevier Ltd., 2016.
- [6] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, vol.9351, pp.234–241, LNCS, Springer, 2015. (available on arXiv:1505.04597 [cs.CV]). <http://lmb.informatik.uni-freiburg.de/Publications/2015/RFB15a>
- [7] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, "Photographic tone reproduction for digital images," *ACM Transactions on Graphics (TOG)*, vol.21, no.3, pp.267–276, 2002.
- [8] X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural networks," *Proceedings of the fourteenth international conference on artificial intelligence and statistics*, pp.315–323, 2011.
- [9] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *arXiv preprint arXiv:1502.03167*, pp.1–11, 2015.
- [10] "Github - openexr". <https://github.com/openexr/>
- [11] "High dynamic range image examples". <http://www.anyhere.com/gward/hdrenc/pages/originals.html>
- [12] "The HDR photographic survey". <http://rit-mcsl.org/fairchild/HDRPS/HDRthumbs.html>
- [13] "EMPA HDR images dataset". This dataset is unavailable now. <http://empamedia.ethz.ch/hdrdatabase/index.php>
- [14] H. Nemoto, P. Korshunov, P. Hanhart, and T. Ebrahimi, "Visual attention in LDR and HDR images," 9th International Workshop on Video Processing and Quality Metrics for Consumer Electronics (VPQM), pp.1–6, 2015. <http://mmspg.epfl.ch/hdr-eye>
- [15] "Max planck institut informatik". <http://resources.mpi-inf.mpg.de/hdr/gallery.html>
- [16] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," *Proceedings of the IEEE international conference on computer vision*, pp.1026–1034, 2015.
- [17] D.P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, pp.1–15, 2014.
- [18] M. Narwaria, R.K. Mantiuk, M.P. Da Silva, and P. Le Callet, "Hdr-vdp-2.2: a calibrated method for objective quality prediction of high-dynamic range and standard images," *Journal of Electronic Imaging*, vol.24, no.1, pp.010501–010501, 2015.
- [19] T.O. Aydin, R. Mantiuk, and H.-P. Seidel, "Extending quality metrics to full luminance range images," *Electronic Imaging 2008*, pp.68060B–68060B, International Society for Optics and Photonics, 2008.
- [20] Z. Wang, E.P. Simoncelli, and A.C. Bovik, "Multiscale structural similarity for image quality assessment," *Signals, Systems and Computers*, 2004. Conference Record of the Thirty-Seventh Asilomar Conference on, vol.2, pp.1398–1402, IEEE, 2003.
- [21] P. Hanhart, M.V. Bernardo, M. Pereira, A.M. Pinheiro, and T. Ebrahimi, "Benchmarking of objective quality metrics for hdr image quality assessment," *EURASIP Journal on Image and Video Processing*, vol.2015, no.1, pp.1–18, 2015.