

GANを用いた単一画像からの ハイダイナミックレンジ画像の生成

森脇 健太^{1,a)} 吉橋 亮太¹ 川上 玲¹ 尤 少迪^{2,3,b)} 苗村 健¹

概要:

ハイダイナミックレンジ (HDR) 画像は、センサの記録範囲を超えた光量のダイナミックレンジを表現できる画像のことである。人間の視覚特性に近い、もしくはそれ以上の感度範囲を表現でき、コンピュータグラフィックスにおける照明効果のシミュレーションなどに用いられる。HDR 画像の生成には、複数の露出で撮影した画像を合成する手法が一般的であるが、撮影に手間がかかり、また HDR 撮影がされなかった画像には適用できない。本稿では、深層畳み込みニューラルネットワーク (DCNN) を用いて一枚のローダイナミックレンジ (LDR) 画像から HDR 画像を生成する手法を提案する。既存の生成モデルの一つであるエンコーダ・デコーダモデルに加えて、新たに敵対的生成ネットワーク (GAN) を導入する。また、我々は損失関数に敵対的損失および MSE 損失を定義することに加え、知覚的類似性を表すコンテンツ損失を導入することでより自然な画像を生成する。実験では、画質評価尺度である PSNR, SSIM を用いて既存の HDR 生成手法と比較を行い、我々の手法の有効性を示す。

1. はじめに

一般的な風景のダイナミックレンジは広く、しばしばコントラスト比は 100,000 : 1 を超える [1]。しかし、イメージセンサのダイナミックレンジは狭く、実世界の光量の範囲をそのまま記録できない。撮影時は主要な被写体に対して適正露出が設定されるため、たとえば太陽や影といった背景で極端に明暗差がある部分は、センサの記録可能な範囲を超えてしまう。明るすぎて白く写った部分を白飛び、暗すぎた部分を黒つぶれという。ハイダイナミックレンジ (HDR) 画像は、そのようなセンサの記録範囲を超えた幅広いダイナミックレンジを表現する画像形式である。赤、緑、青、の各色を表す仮数部に加え、指数部を導入することにより、幅広い諧調が表現される。HDR 画像を用いれば、白飛びや黒つぶれを回避できる。このため、コンピュータグラフィックスにおける照明のシミュレーションや、美しい画像編集によく利用される。

HDR 画像を得るには、露出の異なる複数枚のローダイナミックレンジ (LDR) 画像を撮影し、それらを合成する

必要がある。しかし、複数の露出で視点が完全に一致した画像を撮影するには、特殊なセンサ、もしくは三脚やカメラの遠隔操作が必要になるなど、通常よりも手間がかかる。また、複数の露出で撮影されなかった画像には適用できない。

そこで、単一の LDR 画像から HDR 画像を深層学習により推定する試みがいくつかある。雑音除去、インペインティング、色づけ、超解像、など多くの画像復元のタスクでエンコーダ・デコーダモデル [2], [3] が成功を収めており、HDR 画像生成にも応用されている。たとえば、Endo ら [4] は、エンコーダ・デコーダモデルにより複数枚の異なる露光画像を生成し、これらを従来の物理ベースの手法 [5] で合成した。Eilertsen ら [6] はエンコーダ・デコーダモデルにより直接 HDR 画像を生成した後、これを入力画像と合成する後処理を加えた。しかし、白飛びや黒つぶれにより情報が欠落している単一の LDR 画像からの HDR 画像生成は困難であり、しばしば生成画像に雑音が生じ、復元されたコントラストも完全でない。

そこで本稿では、エンコーダ・デコーダモデルに加えて、同様に画像復元で成功を収めている Generative Adversarial Network (GAN) [7] の構造を導入し、生成された HDR 画像の画質改善を試みる。白飛びの復元はインペインティングと同種のタスクであり、黒つぶれの復元は雑音除去に近い。このため、両者に有効な GAN が HDR 生成でも同

¹ 東京大学
The University of Tokyo

² Data61-CSIRO

³ オーストラリア国立大学
Australian National University

a) moriwaki,yoshi,rei,naemura@nae-lab.org

b) Shaodi.You@data61.csiro.au

様に機能するはずである。また、視覚的に自然な HDR 画像を生成するために、超解像化に利用される視覚的類似性を表す VGG19 [8] の出力による損失関数を導入する。これにより、生成画像における不自然な雑音を低減できる。我々は、この二つの要素を単一 LDR 画像からの HDR 画像生成に新たに適用し、既存手法との比較から、我々の手法の有効性を示す。

2. 関連研究

深層学習を利用し、単一の LDR 画像から HDR 画像を生成する手法は、二つに大別できる。まず複数の露出の異なる画像を生成し、これらを従来手法 [5] で合成するような二段階のステップを踏むやり方と、入力から直接 HDR 画像を推定する手法である。

前者としては、たとえば Endo ら [4] は、HDR 画像データセットから、単一の露光で撮影したような画像をシミュレートし、Grossberg と Nayar らのカメラカーブ [9] を利用して入力画像を生成した。この入力から、複数の露光を模した複数枚の LDR 画像を生成するように、エンコーダ・デコーダモデルに学習させた。同様に、Lee らは、中間的な露光の入力画像からエンコーダ・デコーダモデル、もしくは GAN を利用して、複数枚の異なる露光画像を段階的に生成する手法を提案した [10], [11]。これらの手法は、入力画像から少しずつ露出の高い/低い画像を生成する戦略をとっている。

直接エンド・ツー・エンドで所望の出力を学習するような後者の手法として、Marnerides ら [12] は、三つの枝分かれたニューラルネットワークを利用して HDR 画像を生成する手法を提案した。一つは、抽象的な特徴量を学習する大域的ネットワークであり、残りの二つは、高周波数およびその周辺の特徴量を学習する局所的ネットワークの構造である。Eilertsen ら [6] や、Yang ら [13] は、白飛びの度合いが同程度の露光画像に対し、エンコーダ・デコーダモデルを利用して、HDR 画像を生成した。ネットワークのエンコーダからデコーダに、U-Net [14] の構造と同様の Skip-Connections を導入することで、エンコーダの初期層で失われ得るテクスチャ情報を含んだ特徴空間の伝達を可能にし、推定画像の復元精度を向上させている。Eilertsen らと Yang らは、HDR 画像推定における課題である推定レンジの広さを画素値の対数を取ることで対処している。これにより生成される HDR 画像の値域が広がる。本稿でも、Eilertsen ら [6] のエンコーダ・デコーダモデルを利用し、また画素値の対数空間を利用する。

しかし、前者、後者の手法においても、入力画像となる LDR 画像中に過度な白飛びおよび黒つぶれが存在する際、詳細な輝度情報の復元や見た目の自然さなどの観点において結果が十分でない。そこで、画像のインペインティングや超解像化などの画像復元で良好な結果を収め注目

されている GAN [7] の利用が考えられる。GAN は生成器 (generator) と識別器 (discriminator) という二つの構造を持つ。生成器は、訓練データと同様の確率分布を生成するように学習する。識別器は、生成データと訓練データを見分けるように学習する。それぞれの最適化が完全に進めば、生成データは目標とする確率分布に必ず近づく。実際に、学習が進むと、生成器が訓練データと同等のデータを生成するようになり、識別器はこれらの識別が困難になる。

GAN を利用した画像のインペインティングの手法に [15], [16] がある。生成モデルにはエンコーダ・デコーダモデルを採用している。生成ネットワークの出力と真値との画素間の最小二乗誤差を損失関数として用いるのみでは、復元結果がぼやけてしまう一方で、GAN を導入することで、より輝度情報の詳細な復元が可能になることが示された。画像のインペインティングと同様に HDR 画像生成の問題に GAN を導入した手法に Lee ら [11] や Ning ら [17] の手法がある。GAN を導入することで欠陥した情報からの復元の精度をより高めることが可能になった。

本稿では Lee らや Ning らと同様に GAN を導入するとともに、さらに画像の超解像化 [18] で用いられた視覚的類似性を表す損失関数を HDR 画像生成のモデルに導入する。これにより、HDR 画像の画質が改善できる。

3. システムの概要

3.1 敵対的生成モデルの概要

図 1 に手法の概要を示す。ネットワークの構造は generator と discriminator の二つの構造からなる。generator には、Eilertsen ら [6] のエンコーダ・デコーダモデルを用いる。入力画像は 1 チャネルあたり 8 ビットの LDR 画像であり、ネットワークを通じて 1 チャネルあたり 32 ビットの float 型に変換させた推定画像が生成される。discriminator の構造は、画像の超解像化 [18] で用いられている構造で、8 層の畳み込み層と 2 層の全結合層からなる 10 層で構成される。discriminator には、LDR 画像生成に用いた元データで真値となる HDR 画像と generator が出力した推定 HDR 画像を入力とし、本物らしさの確率を 0 から 1 の数値で出力する。入力 LDR 画像の作成方法に関しては、4.1 章に詳細を記述する。

3.2 損失関数

我々は、GAN の generator の損失関数 L_G に三つの損失関数を用いる。画素間の距離を定義した MSE ベースの損失関数、GAN に用いられる敵対的損失関数、さらに、VGG19 のネットワークの出力を利用し、視覚的類似性を定義した損失関数であり、以下の式で定義される。

$$L_G = L_{IR} + \alpha L_{GAN_{gen}} + \beta L_{VGG} \quad (1)$$

また、discriminator の損失関数 L_D は [7] に倣い以下の式

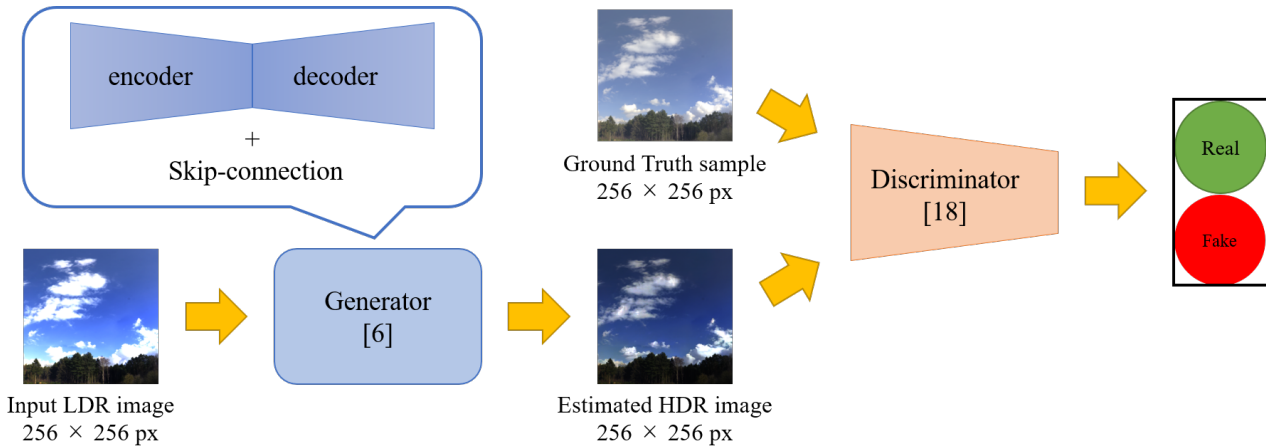


図 1 提案手法の概要

を用いる。

$$L_D = \frac{1}{N} \sum_N (-\log(1 - D(\hat{y})) + \log(D(y))) \quad (2)$$

これらはそれぞれ generator と discriminator の重みの更新に利用される。 \hat{y} はモデルによって推定された HDR 画像であり、 y は真値となる元データである。 α, β は学習における重みである。

L_{IR} は、 generator の生成画像と真値の画像間の画素ごとの最小二乗誤差を計算した損失関数である。 Eilertsen ら [6] の結果に従い、生成画像にガウシアンフィルタをかけることで低周波成分と高周波成分を分けて損失関数を計算する。

$$\log(I_i^{\hat{y}}) = (G_\sigma * \hat{y}_i) \quad (3)$$

$$\log(R_{i,c}^{\hat{y}}) = \hat{y}_{i,c} - \log(I_i^{\hat{y}}) \quad (4)$$

ここで I は低周波成分を、 R は高周波成分を、 G_σ はガウシアンローパスフィルタを、 i および c は画素とチャンネルを表す。

線形な HDR 画像の輝度値を直接用いた損失関数は、高輝度値に大きく依存し、より低い範囲の輝度値の誤差を過小評価してしまう。そこで [6] や [13] と同様に、元データおよび出力 HDR 画像を対数領域で表現し、低周波成分と高周波成分を用いて以下の式で計算する。

$$L_{IR}(\hat{y}, y) = \frac{\lambda}{N} \sum_i |m_i(\log(I_i^{\hat{y}}) - \log(I_i^y))|^2 + \frac{1-\lambda}{3N} \sum_{i,c} |m_i(\log(R_{i,c}^{\hat{y}}) - \log(R_{i,c}^y))|^2 \quad (5)$$

m_i は白飛び領域のマスクに対応し、その領域のみ差分を計算する。この m_i は以下の式で定義する。

$$m_i = \frac{\max(0, \max_c(x_{i,c}) - \tau)}{1 - \tau} \quad (6)$$

$x_{i,c}$ は [0,1] に正規化された LDR 画像であり、閾値である τ は 0.95 に設定した。

$L_{GAN_{gen}}$ は、GAN で利用される敵対的生成損失関数である。この損失関数を導入することにより元データに近い自然な生成画像が期待できる。損失関数は、以下の式で定義される。

$$L_{GAN_{gen}} = \sum_{n=1}^N -\log D(\hat{y}) \quad (7)$$

ここで \hat{y} は、generator によって生成された HDR 画像であり、 $D(\hat{y})$ は確率を表す。より良い勾配の挙動のために、 $\log(1 - D(\hat{y}))$ の代わりに $-\log D(\hat{y})$ を最小化する [19]。

L_{VGG} は、 Gatys ら [20] や Bruna ら [21], Johnson ら [22], Christian ら [18] の論文で用いられている考えであり、視覚的類似性を計算する損失関数である。我々は、Christian ら [18] と同様に Simonyan と Zisserman ら [8] の 19 層からなる学習済み VGG19 ネットワークの活性化層の出力を損失関数として定義する。 $\phi_{i,j}$ は、VGG19 における i 番目の MaxPooling 層の直前の畳み込み層の j 番目のチャンネルを表し、VGG の損失関数は次のように定義される。 $N_{i,j}$ は特徴量空間の画素の総数を表す。本稿では、 $i = 5, j = 4$ を用いた [22]。

$$L_{VGG}(\hat{y}, y) = \frac{1}{N_{i,j}} \sum_{k=1}^{N_{i,j}} (\phi_{i,j}(\hat{y})_k - \phi_{i,j}(y)_k)^2 \quad (8)$$

4. 実験

4.1 データセット

学習ベースの HDR 画像推定の重要な課題は、十分な学習データセットを得ることである。我々は Eilertsen ら [6] が使用したデータセットに加え、インターネット上から新たに HDR 画像データを収集し、合計約 1,000 枚 (69.0GB) の HDR 画像および動画を収集した。データセットは基本的に風景を撮影した画像であり、ランダムな位置で 256 x 256 ピクセルにクロッピングを行いデータの拡張を行なった。データをさらに HDR 画像から LDR 画像セットを生成する際に、HDR 画像の輝度値のサンプリング間隔を取得

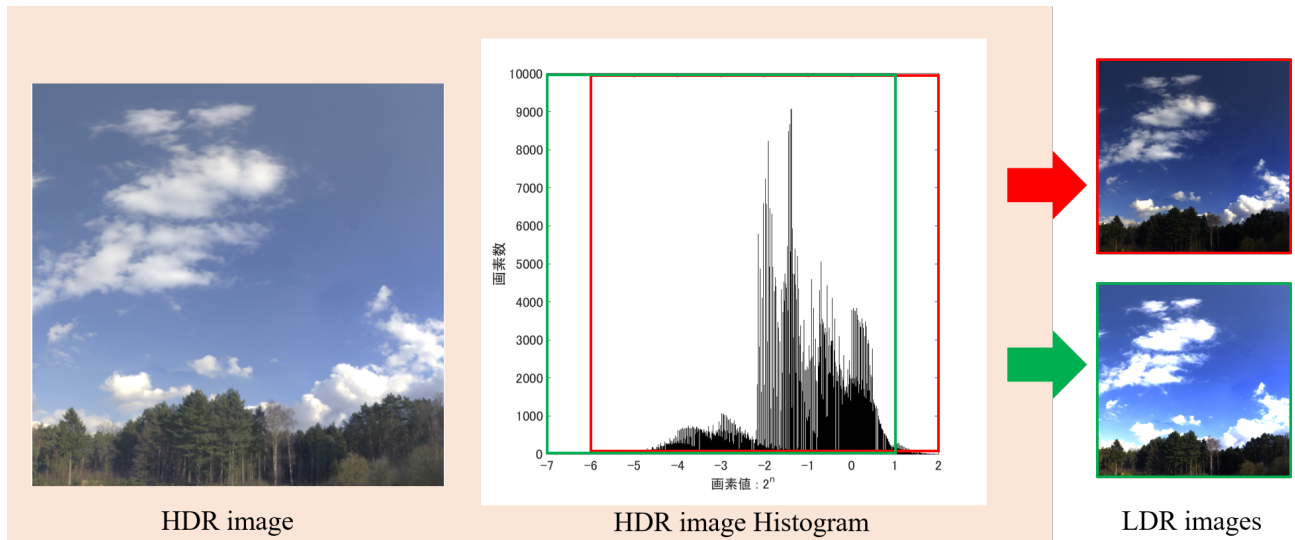


図 2 HDR 画像から LDR 画像セットを生成するまでの流れ. 右に示す例では, 赤はサンプリング間隔 2^{-6} を表し, 表 1 では d に対応する. 同様に緑は e に対応する.

表 1 HDR 画像から作成した LDR 画像セットのサンプリング間隔

Name	サンプリング間隔	輝度値の範囲
a	2^{-3}	$2^{-3} \sim 2^5$
b	2^{-4}	$2^{-4} \sim 2^4$
c	2^{-5}	$2^{-5} \sim 2^3$
d	2^{-6}	$2^{-6} \sim 2^2$
e	2^{-7}	$2^{-7} \sim 2^1$
f	2^{-8}	$2^{-8} \sim 2^0$

し, 1 チャンネルあたり 8 ビットになるようなレンジで画像を整形し 1 枚の HDR 画像から複数の露光の LDR 画像セットを拡充した. 拡充の流れは図 2 に示す通りであり, 中央値で正規化した HDR 画像からサンプリング間隔を取得し, 図 2 の中央のヒストグラムに示すように一部のレンジを抽出し, 図の右に示すような画像セットを生成した. 視覚的に十分な量のレンジを取得するため今回は, 表 1 に示す 6 つの異なる露光画像を生成し, それぞれを $a \sim f$ として定めた. 我々は最終的に HDR データセットから約 80,000 枚の学習データを作成した. データセットの一例を図 3 に示す.

4.2 学習方法

GAN の生成器の重みの初期値には, Eilertsen ら [6] の学習済みモデルを採用する. Eilertsen らは, エンコーダ・デコーダモデルの重みの初期値に, Places [23] のデータセットで事前学習を行ったものを採用している. これは, 現状入手可能な HDR データセットの量が不十分であるためであり, 大規模データセットで学習されたモデルを利用した転移学習とみなせる. この際, 白飛びや黒つぶれが生じているものは学習データとしてふさわしくないため, 白飛びや黒つぶれのないものを選択し, さらにその画像から一部を白飛びさせることで, 擬似的に LDR 画像と HDR 画像

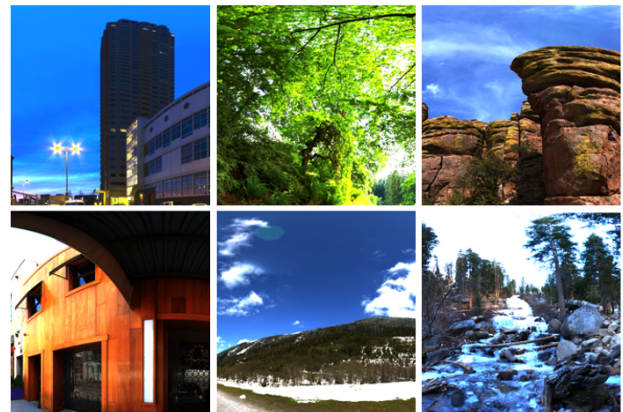


図 3 学習データの一例

の対を作成している. Places のデータは 600K 枚ほどあり, さらにクロッピングを行い, 2.5M 枚のデータを事前学習に用いた. これらの事前学習されたパラメータを初期値とし, Eilertsen ら [6] が収集した HDR 画像データセットを利用して学習されたものを提案手法における生成器の重みの初期値として用いる.

損失関数は 3.2 章で定義した (1) および (2) を使用する. この二つの損失関数を用い, generator と discriminator の重みを交互に更新する. ここではイテレーション数が偶数の時は generator の重みを更新し, 奇数の時は discriminator の重みを更新する. 最適化手法として Adam [24] を用い, 学習率は 2.0×10^{-5} , バッチサイズは 16 枚, エポック数を 60 として, 合計 30 万回のイテレーション数で学習を行った. また (1) 中の α および β は学習の重みを表し, 今回は, $\alpha = 1.0 \times 10^{-3}$, $\beta = 5.0 \times 10^{-4}$ としている.

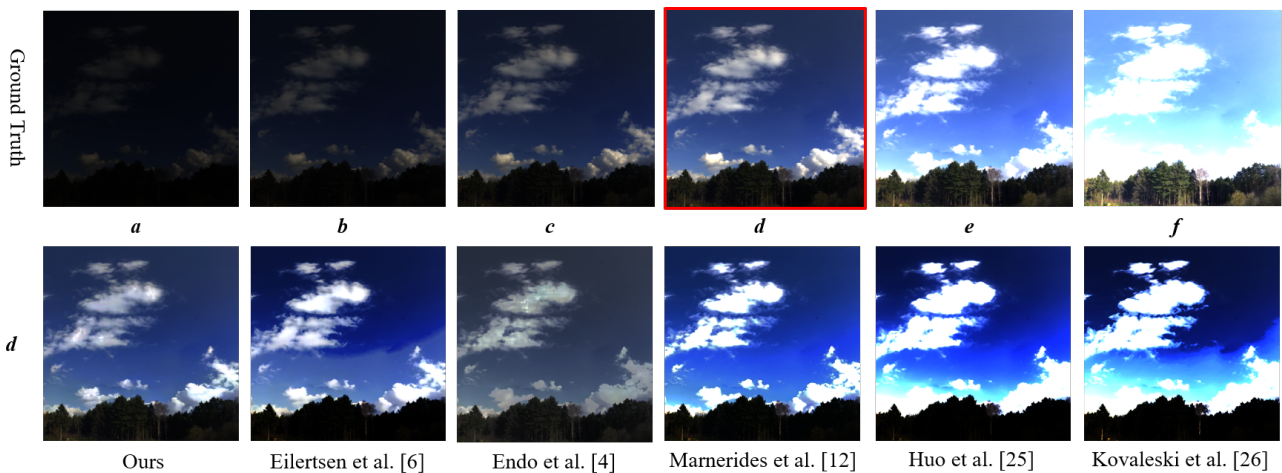


図 4 比較結果 1. 他手法との比較を図 2 に示す方法で LDR 画像セットを生成し、定性的な比較を可視化したものである。今回は入力に e の領域を使用し d の領域における手法の比較結果を図示した。



図 5 入力 LDR 画像の例。表 1 の e の領域に対応する画像を実験の入力画像としている。

表 2 PSNR(dB) : Comparison of the inferred LDR image stack

Method	a	b	c	d	e	f	Mean
Ours	34.57	30.66	27.67	27.82	32.44	32.16	30.88
[6]	27.66	23.20	19.44	17.67	18.90	16.39	20.54
[4]	29.41	24.90	20.97	19.49	21.76	21.30	22.97
[12]	25.44	20.89	17.22	16.15	19.95	18.84	19.74
[25]	19.09	14.13	11.35	12.74	16.48	14.50	14.71
[26]	17.13	12.26	11.26	13.50	16.97	15.21	14.38

4.3 実験結果

本章では、いくつかの HDR 画像生成の結果を示しながら、我々の手法の有用性を示す。本研究では、用意したデータセットのうち、4.1 章に示す LDR 画像セットの中から $d \sim f$ に対応するレンジの LDR 画像のみを学習データとし

表 3 SSIM : Comparison of the inferred LDR image stacks

Method	a	b	c	d	e	f	Mean
Ours	0.937	0.931	0.931	0.951	0.978	0.964	0.948
[6]	0.833	0.784	0.752	0.753	0.782	0.793	0.782
[4]	0.831	0.782	0.741	0.748	0.824	0.839	0.794
[12]	0.786	0.738	0.718	0.756	0.841	0.877	0.786
[25]	0.597	0.534	0.516	0.581	0.681	0.753	0.610
[26]	0.567	0.515	0.534	0.616	0.706	0.762	0.616

て実験を行った。 d は平均として 8%、 e は平均として 20%、 f は平均として 50% の白飛び領域が存在する。実際に訓練画像には、約 80,000 枚の画像を使用し、テストデータには、約 100 枚の画像を使用し、性能評価を行った。評価対象は、我々と同様に深層学習を利用した最新の手法である Endo ら [4] や Eilertsen ら [6]、Marnerides ら [12] の手法に加え、従来の物理ベースの逆トーンマッピングの手法である Huo ら [25] と Kovaleski と Oliveira ら [26] の手法との比較を行った。

生成した HDR 画像から 4.1 章に記述した手法で LDR 画像セットを生成し、真値となる LDR 画像セットと PSNR および SSIM を利用して評価した。PSNR の評価結果を表 2 に、SSIM の評価結果を表 3 に示す。また生成された LDR 画像セットおよび真値となる LDR 画像セットの一例を図 4 に示す。表中の二重線で区切られているものは、上部が我々のデータで再学習したものである。下部の手法は公開されている学習済みモデルを使用して評価したものであり、学習データが異なる可能性がある。表 2、3 に示す PSNR および SSIM のスコアは、生成結果がどれほど真値に近いものなのかを示しており、既存の他手法と比較してみても、提案手法が優れていることがわかる。図 4 に示しているように、従来手法では、図 5 に示す画像のような白飛び領域を含む画像を入力した際、LDR 画像から HDR 画像への推定が困難であることから、一部のレンジ画像におい

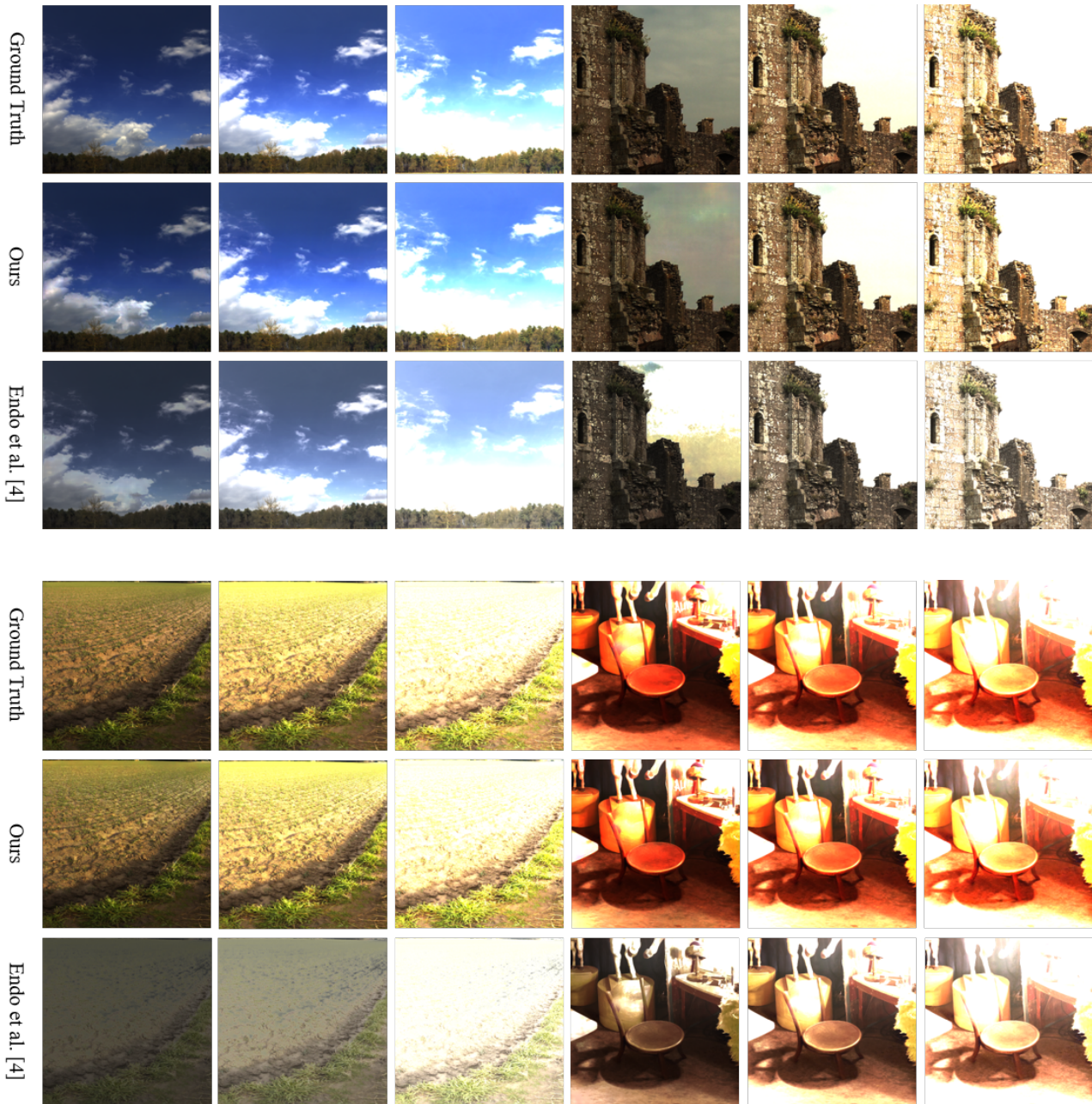


図 6 比較結果 2. 表 2, 3 の結果を参考に他の既存手法のうち最も復元精度の高い Endo ら [4] の手法と我々の手法の復元結果を比較したものを.

て輝度情報の復元ができていないことや, 人工的な不自然な境界が現れてしまう一方で, 我々の手法は真値と比較した際に, 画像のコントラストおよび, 輝度情報の復元の精度がよいことがわかる. また図 4 および表 2,3 において結果の優れていた Endo ら [4] の手法とさらに複数の復元結果を比較したものを図 6 に示す. この結果から定性的に判断しても我々の手法が優れていることがわかる. 一方で図 7 に示すようにより過度な露光画像においては, 輝度情報が十分に復元されず, 生成時に格子状の人工的な雑音が生じる画像の存在も確認しており, Endo らの手法より真値に近い画像を生成している一方で, 不自然さが目立っていると

いう課題も存在する.

5. おわりに

本稿において, 我々は深層畳み込みニューラルネットワークを利用し, 1 枚の LDR 画像から直接的に HDR 画像を生成する手法を提案した. 我々の手法は, 従来手法と比較して白飛び領域が多少ある画像においても, より真値に近い画像を生成することが可能であることを示した.

一方で図 7 に示すように過度な露光画像においては, 4.3 章で述べた課題が存在する. この結果から GAN の導入効果を十分に発揮できていない可能性も考えられる. 今後の

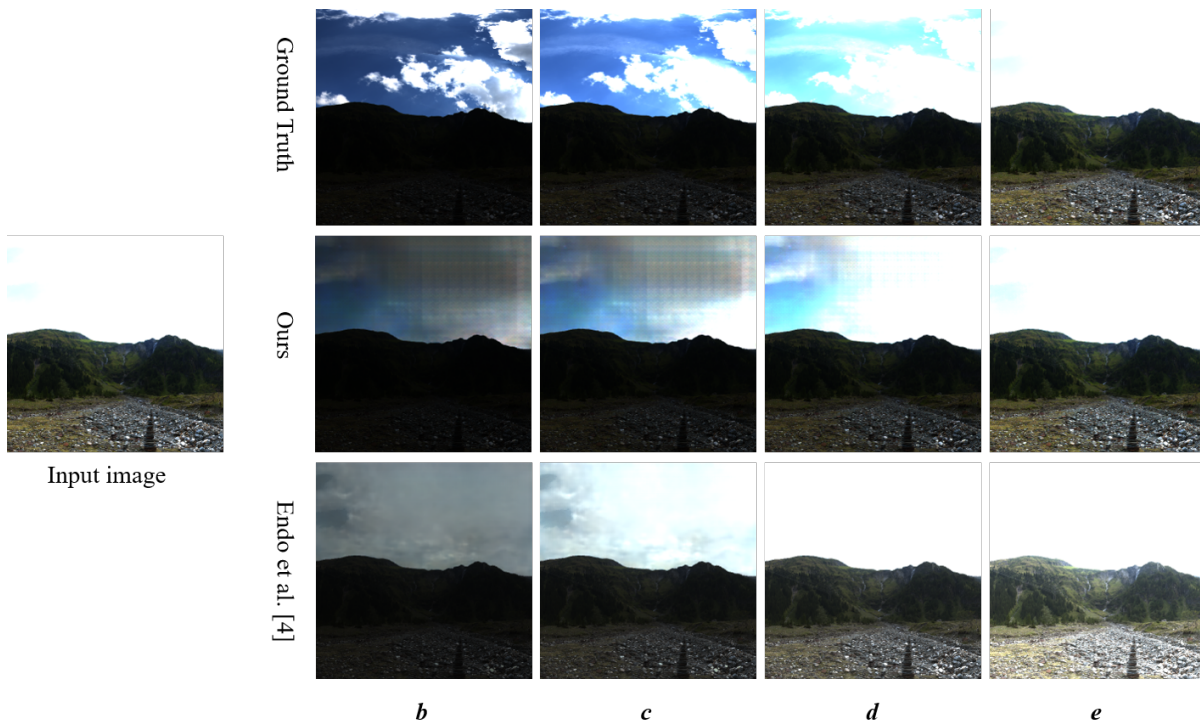


図 7 過度な露光画像に対しては十分な復元結果が得られない場合も存在する. Endo ら [4] の結果と比較しても真値に近いという観点では優れているが, 人工的なノイズが生じている.

研究では, 結果を踏まえ, ネットワーク構造の改善などを行う予定である.

謝辞

本研究は, JSPS 科研費 JP18K11348 と JP16J04552 の助成を受けたものです.

参考文献

- [1] Myszkowski, K., Mantiuk, R. and Krawczyk, G.: High dynamic range video, *Synthesis Lectures on Computer Graphics and Animation*, Vol. 1, No. 1, pp. 1–158 (2008).
- [2] Hinton, G. E. and Salakhutdinov, R. R.: Reducing the dimensionality of data with neural networks, *Science*, Vol. 313, No. 5786, pp. 504–507 (2006).
- [3] Vincent, P., Larochelle, H., Bengio, Y. and Manzagol, P.-A.: Extracting and Composing Robust Features with Denoising Autoencoders, pp. 1096–1103 (2008).
- [4] Endo, Y., Kanamori, Y. and Mitani, J.: Deep Reverse Tone Mapping, *ACM Transactions on Graphics (Proc. of SIGGRAPH ASIA 2017)*, Vol. 36, No. 6 (2017).
- [5] Debevec, P. E. and Malik, J.: Recovering high dynamic range radiance maps from photographs., *SIGGRAPH (Owen, G. S., Whitted, T. and Mones-Hattal, B., eds.)*, ACM, pp. 369–378 (1997).
- [6] Eilertsen, G., Kronander, J., Denes, G., Mantiuk, R. and Unger, J.: HDR image reconstruction from a single exposure using deep CNNs, *ACM Transactions on Graphics (TOG)*, Vol. 36, No. 6 (2017).
- [7] Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y.: Generative Adversarial Nets, *Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2*, NIPS'14, pp. 2672–2680 (2014).
- [8] Simonyan, K. and Zisserman, A.: Very Deep Convolutional Networks for Large-Scale Image Recognition, *In International Conference on Learning Recognition (ICLR)* (2015).
- [9] Grossberg, M. D. and Nayar, S. K.: What is the space of camera response functions?, *2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003.Proceedings.*, Vol. 2, pp. 602–612 (2003).
- [10] Lee, S., An, G. H. and Kang, S.-J.: Deep Chain HDRI: Reconstructing a High Dynamic Range Image from a Single Low Dynamic Range Image, *IEEE Access*, Vol. 6, pp. 49913–49924 (2018).
- [11] Lee, S., Hwan An, G. and Kang, S.-J.: Deep Recursive HDRI: Inverse Tone Mapping using Generative Adversarial Networks, *The European Conference on Computer Vision (ECCV)* (2018).
- [12] Marnerides, D., Bashford-Rogers, T., Hatchett, J. and DeBattista, K.: ExpandNet: A Deep Convolutional Neural Network for High Dynamic Range Expansion from Low Dynamic Range Content, *Comput. Graph. Forum*, Vol. 37, No. 2, pp. 37–49 (2018).
- [13] Yang, X., Xu, K., Song, Y., Zhang, Q., Wei, X. and Lau, R. W.: Image Correction via Deep Reciprocating HDR Transformation, *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2018).
- [14] Ronneberger, O., P.Fischer and Brox, T.: U-Net: Convolutional Networks for Biomedical Image Segmentation, *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, LNCS, Vol. 9351, Springer, pp. 234–241 (2015).
- [15] Pathak, D., Krähenbühl, P., Donahue, J., Darrell, T. and Efros, A.: Context Encoders: Feature Learning by

- Inpainting, *CVPR* (2016).
- [16] Yang, C., Lu, X., Lin, Z., Shechtman, E., Wang, O. and Li, H.: High-Resolution Image Inpainting Using Multi-Scale Neural Patch Synthesis, *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2017).
 - [17] Ning, S., Xu, H., Song, L., Xie, R. and Zhang, W.: Learning an Inverse Tone Mapping Network with a Generative Adversarial Regularizer (2018).
 - [18] Ledig, C., Theis, L., Huszar, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A. P., Tejani, A., Totz, J., Wang, Z. and Shi, W.: Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network, *2017 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, Honolulu, HI, USA, July 21-26, 2017*, pp. 105–114 (2017).
 - [19] Arjovsky, M. and Bottou, L.: Towards Principled Methods for Training Generative Adversarial Networks, *International Conference on Learning Representations (ICLR)* (2017).
 - [20] Gatys, L. A., Ecker, A. S. and Bethge, M.: Texture Synthesis Using Convolutional Neural Networks, *NIPS* (2015).
 - [21] Bruna, J., Sprechmann, P. and LeCun, Y.: Super-Resolution with Deep Convolutional Sufficient Statistics, *International Conference on Learning Representations (ICLR)* (2015).
 - [22] Johnson, J., Alahi, A. and Fei-Fei, L.: Perceptual losses for real-time style transfer and super-resolution, *European Conference on Computer Vision* (2016).
 - [23] Zhou, B., Lapedriza, A., Xiao, J., Torralba, A. and Oliva, A.: Learning Deep Features for Scene Recognition using Places Database, *NIPS*.
 - [24] Kingma, D. P. and Ba, J.: Adam: A method for stochastic optimization, *International Conference on Learning Representations (ICLR)* (2015).
 - [25] Huo, Y., Yang, F., Dong, L. and Brost, V.: Physiological inverse tone mapping based on retina response, *The Visual Computer*, Vol. 30, No. 5, pp. 507–517 (2014).
 - [26] Kovaleski, R. and Oliveira, M. M.: High-Quality Reverse Tone Mapping for a Wide Range of Exposures, *Graphics, Patterns and Images (SIBGRAPI), 2014 27th SIBGRAPI Conference on*, pp. 49–56 (online), DOI: 10.1109/SIBGRAPI.2014.29 (2014).