

# Deep Neural Networks を用いた Inverse Tone Mapping

平尾克彦<sup>†1</sup> 竹内健<sup>†1</sup> 甲藤二郎<sup>†1</sup>

**概要** : Deep Neural Networks (DNN) を用いて, 単一の Standard Dynamic Range (SDR) 画像から High Dynamic Range (HDR) 画像を生成する inverse Tone Mapping Operator (iTMO) を提案する. 近年の撮像, 表示デバイスの進歩に伴い, 既存の SDR 画像の輝度を HDR 相当まで拡大する高精度な iTMOs 技術が求められている. 本稿では, 近年様々な分野で成功を収めている Deep learning ベースの手法を提案し, 画質評価指標である PSNR を用いて, その有効性を示す.

**キーワード** : 深層学習, HDR

## Inverse Tone Mapping using Deep Neural Networks

KATSUHIKO HIRAO<sup>†1</sup> MASARU TAKEUCHI<sup>†1</sup>  
JIRO KATTO<sup>†1</sup>

**Abstract**: We propose inverse Tone Mapping Operators (iTMOs) that generate High Dynamic Range (HDR) image from an Standard Dynamic Range (SDR) image using Deep Neural Networks (DNN). Recently, high precision iTMOs are required with the progress of imaging and display devices. Inspired by the recent success of deep learning in a variety of fields, we propose a deep learning based method and show efficiency of our method by PSNR.

**Keywords**: Deep learning, HDR

### 1. はじめに

近年, 撮像や表示デバイスの技術発展に伴い, 従来よりも広いダイナミックレンジを持つ High Dynamic Range (HDR) 画像データの普及が進んでいる. 一方, 現存する多くの画像データは Standard Dynamic Range (SDR) 画像であり, SDR 画像を HDR 画像と比較して遜色なく表示させるために, SDR 画像を HDR 相当まで拡大する高精度な inverse Tone Mapping Operator (iTMO) 技術への要求が高まっている.

本稿では, 近年超解像やカラリゼーションの分野で高い精度を示している Deep Neural Networks (DNN) を用いて, 単一の SDR 画像から HDR 画像を生成する iTMO を提案する.

### 2. 関連研究

これまでに数多くの iTMO が提案されてきた. Kuo らは [1] にて, シーン分類をアルゴリズムに含めて環境パラメータを設定し, コンテンツ毎に適切な処理を行う iTMO を提案している. また Huo らは [2] にて, 限られたパラメータ数で, 人間の視覚システムの特徴を考慮した iTMO を提案している. 更に, Kovalenski らは [3] にて, 適正露出領域だけでなく, 露出不足や露出過度な領域にも適切に機能するクロスバイラテラルフィルタベースの iTMO を提案している.

### 3. 概要

図 1 に提案手法の概要図を示す. DNN は, 4 層の畳み込み層で構築し, 畳み込み前後で入出力のサイズを変化させないように, 画像の境界値を 0 で埋める padding を適用している. また, 活性化関数として, 最終層にはシグモイド関数を使用し, それ以外の層には正規化線形関数を使用している. 出力層にシグモイド関数を使用することで, 出力値を [0,1] とする. そして, 確率的勾配効果法により, HDR 推定画像と Ground truth である HDR 画像との Mean Squared Error (MSE) loss を最小化するように学習をさせる.

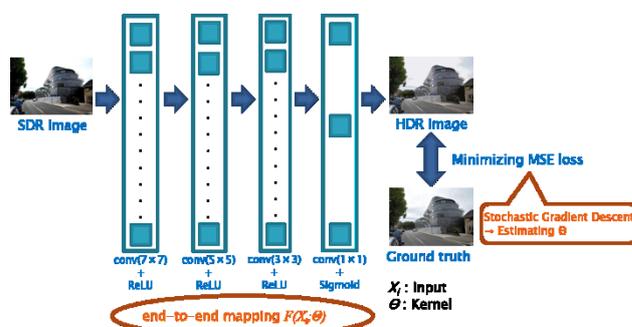


図 1 提案手法の概要図

Figure 1 Overview of our proposed method.

<sup>†1</sup> 早稲田大学  
Waseda University

## 4. データセット

データセットとして、株式会社 IMAGICA が販売している UHD/HDR の評価画像セット LUCORE [4] を使用する。LUCORE のデータセットに含まれている画像は ITU-R BT.2020 に準拠しており、HDR 画像は SMPTE ST 2084 に準拠している。LUCORE の SDR (100nits) バージョンは、ガンマ値 2.2 で最適になるようにグレーディングをしたもので、1000nits バージョンは、SONY の BVM-X300 で、BT.2020, ST 2084 の設定でグレーディングしたものである。また、10000nits バージョンは IMAGICA 社にて組み上げたリアプロジェクションシステムで最高輝度 10000nits に設定してグレーディングを行ったものである。以下の 3 パターンを訓練データとして、学習させる。

### (1) Pattern A

SDR バージョンと HDR 1000nits バージョン  
 映像信号レベルでの学習

### (2) Pattern B

SDR バージョンと HDR 10000nits バージョン  
 映像信号レベルでの学習

### (3) Pattern C

SDR バージョンと HDR 10000nits バージョン  
 ディスプレイ絶対輝度での学習

つまり、SDR 画像に対しては、逆ガンマ補正 ( $\gamma=2.2$ ) を適用し、HDR 画像に対しては、PQ (Perceptual Quantization) 方式の EOTF (Electro-Optical Transfer Function) を適用した後学習させる。そしてテスト時に、ガンマ補正後の SDR 画像を入力とし、DNN の出力に対して Inverse EOTF を適用することで HDR 推定画像とする。

各シーンそれぞれ 1~2 枚を使用して、パッチ生成 (40×40×3) を行い、約 30 万枚の訓練データを生成した。

## 5. 評価実験

200 万イテレーション学習させて HDR 推定画像を生成し、Ground truth である HDR 画像と PSNR による比較をする。約 40 のテストシーンを用意し、PSNR の平均を取った各 Pattern の結果を表 1 に示す。表 1 の結果を見ると、Pattern A と Pattern B の比較から、10000nits のデータを使用した結果よりも、1000nits のデータを使用した結果のほうが良い結果となった。これは、1000nits のデータのほうが SDR のデータに類似しているために、学習が容易になったためであると推測される。また、Pattern B と Pattern C の比較から、

表 1 PSNR の平均値の取得結果

Table 1 Results of mean PSNR.

Pattern	Mean PSNR
A	35.64
B	31.71
C	24.13

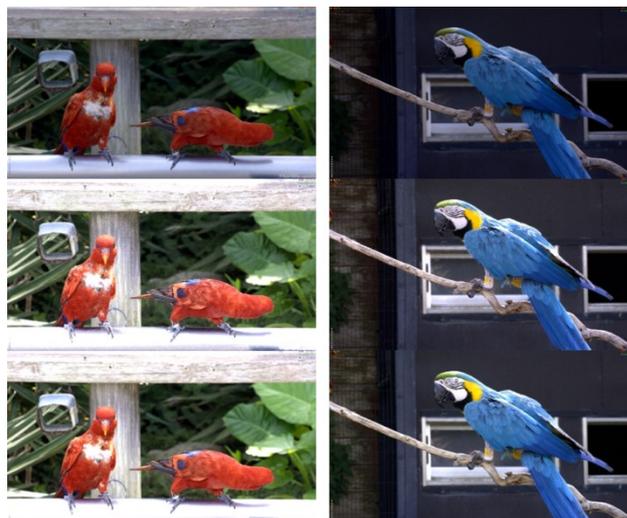


図 2 実験結果 (上から SDR 画像, HDR 推定画像, Ground truth の順)

Figure 2 Experimental results (top row: SDR images, middle row: estimated HDR images, and bottom row: Ground truth images)

データを映像信号レベルで学習させるほうが、ディスプレイ絶対輝度で学習させるよりも良い結果が得られることがわかった。また、Pattern A における、入力である SDR 画像と出力結果である HDR 推定画像、Ground truth の例を図 2 に示す。図 2 の結果は、データをディスプレイ絶対輝度に変換後、Radiance 形式にして保存し、NVIDIA の HDR 画像を表示させるための SDK [5] を使用することで表示している。

## 6. まとめと今後の課題

本稿では、DNN を用いた iTMO を提案し、PSNR によりその有効性を示した。今後の課題として、精度改善のために Residual Network [6] を導入することを検討している。[6] により、更に DNN を多層化し、高レベルな特徴量を抽出することで、より良い結果が得られると推測される。

## 参考文献

- [1] Kuo, Pin-Hung et al. "Content-adaptive inverse tone mapping," IEEE VCIP 2012, Nov.2012.
- [2] Huo, Yongqing, et al.: "Physiological inverse tone mapping based on retina response," The Visual Computer, Vol.30, Issue 5, pp.507-517, May.2014.
- [3] Kovalski, Rafael P., and Manuel M. Oliveira.: "High-quality reverse tone mapping for a wide range of exposures," SIBGRAPH 2014, Aug.2014.
- [4] <https://www.imagica.com/news/lucore/>.
- [5] <https://developer.nvidia.com/high-dynamic-range-display-development>
- [6] He, Kaiming, et al. "Deep residual learning for image recognition," IEEE CVPR 2016, June 2016.