

# 面光源と CNN に基づく少数画像からの再照明

平尾 寿希<sup>1,a)</sup> 王 超<sup>1,b)</sup> 川原 僚<sup>1,c)</sup> 岡部 孝弘<sup>1,d)</sup>

**概要:** 様々な照明環境で撮影した実画像を組合わせて任意照明環境下の画像を生成する技術は再照明と呼ばれる。再照明において鏡面反射成分などの高周波数成分を生成するためには、一般に、密に配置した光源下で撮影した大量の画像を補間する必要がある。本稿では、少数の照明環境下で撮影された画像を用いた再照明法を提案する。提案手法では、少数の照明環境で鏡面反射成分を効率よく捉えるために、点光源だけでなく様々な大きさの面光源やそれらの組合せも利用する。また、畳み込みカーネルを用いて照明環境を表現できることに着目して、照明環境と補間処理の両方を、畳み込みニューラルネットワークの枠組みで同時に最適化する。実画像を用いた実験を行い、提案手法の有効性を示す。

## 1. はじめに

被写体の見えは照明環境に依存することから、任意照明環境下画像の生成はコンピュータグラフィックス (CG) やコンピュータビジョン (CV) 分野における重要な課題である。様々な照明環境下で撮影された実画像を用いて任意照明環境下の画像を生成する技術は再照明と呼ばれる。

一般に、画像は拡散反射や鏡面反射など様々な成分で構成される。拡散反射成分については、異なる光源方向下で撮影した 3 枚の基底画像の線形結合により、任意光源方向下の画像を生成できることが知られている [1]。一方、鏡面反射成分のような高周波数成分については、密な光源方向下で撮影した大量の画像や、局所的に物体が一様な反射特性および滑らかな形状を持つことを仮定したうえで非線形な補間 [3] が必要になることが知られている。

本稿では、少数の照明環境下で撮影された画像を用いた再照明手法を提案する。提案手法では、少数の照明環境で鏡面反射成分を効率よくとらえるために、点光源だけでなく様々な大きさの面光源やそれらの組み合わせも利用する。重ね合わせの原理から、複数光源下の画像と単一光源下の画像の関係は、 $1 \times 1$  の畳み込みカーネルを用いて表現できることに着目し、照明環境と補間処理の両方を畳み込みニューラルネットワーク (CNN) の枠組みで同時に最適化する。さらに、実画像を用いた実験を行い、提案手法の有効性を示す。

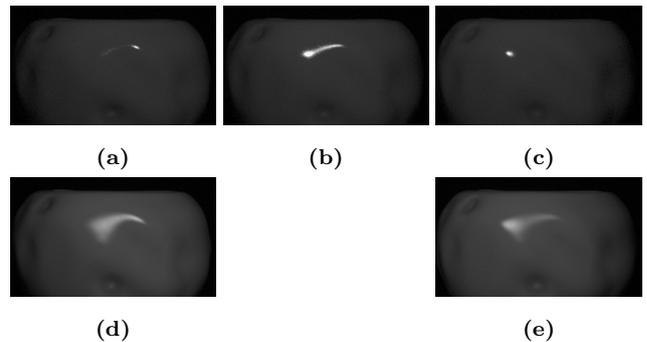


図 1: 面光源利用の有効性

近年のコンピュータビジョン分野におけるディープニューラルネットワーク (DNN) を用いた研究では、Chakrabarti ら [4] の手法のように、撮影済みの画像の処理方法のみならず、撮影方法 (カメラ) までも最適化するディープセンシングと呼ばれる手法が提案されている。本稿では、畳み込みカーネルを用いて照明環境を表現できることに着目し、CNN の枠組みにおいて、再照明を行うために照明環境も同時に最適化するディープライティングという手法を提案する。

## 2. 関連研究

### 2.1 拡散反射成分・鏡面反射成分の補間

低周波数成分である拡散反射成分については、線形補間により、任意光源方向下の拡散反射成分を生成できることが知られている。具体的には、Lambert モデルを仮定すると、任意光源方向下の画像の拡散反射成分は、異なる光源方向下で撮影された 3 枚の基底画像の線形結合により補間できる [1]。

一方、高周波数成分である鏡面反射成分については、拡

<sup>1</sup> 九州工業大学  
a) hirao.toshiki929@mail.kyutech.jp  
b) c\_wang@pluto.ai.kyutech.ac.jp  
c) rkawahara@ai.kyutech.ac.jp  
d) okabe@ai.kyutech.ac.jp

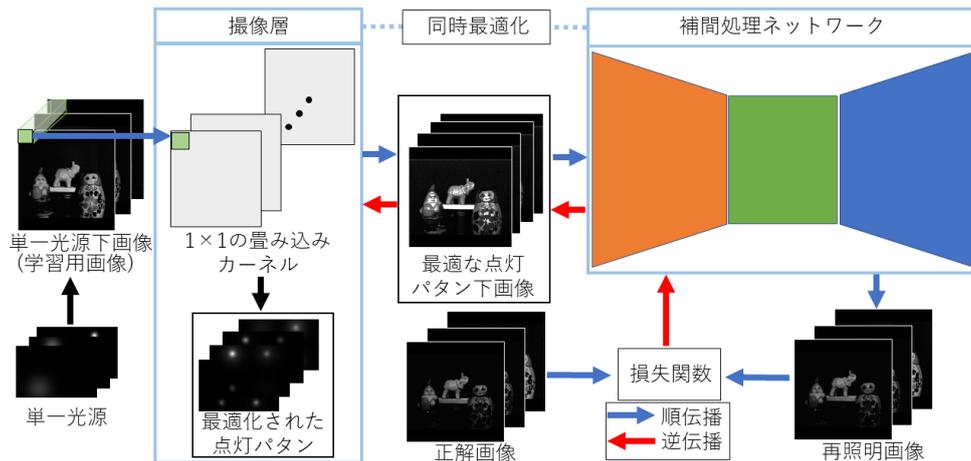


図 2: 点灯パターンと画像補間の同時最適化ネットワーク

散反射成分とは異なり、単純な基底画像の線形結合では任意光源方向下の画像を生成できないことが知られている。そこで Debevec ら [2] は、被写体を顔に限定したうえで、密な光源方向下で撮影した数百枚の画像を使用してライトトランスポートを密に計測することで、鏡面反射などの高周波数成分を生成する手法を提案した。しかしながら、上記の手法では再照明に大量の画像を撮影する必要があり、撮影コストが高い。

一方、Fuchs ら [3] は、比較的少数の異なる照明環境下の画像から任意照明環境下の画像の補間を行う手法を提案した。具体的には、撮影画像から、画素値をもとに鏡面反射領域を分離し、物体表面において鏡面反射が滑らかに移動することを仮定して、オプティカルフローに基づいた非線形補間を行った。しかしながら、実シーンにおいては、曲率の変化や材質の変化のために、鏡面反射が滑らかに移動するという仮定は成り立たない。

そこで提案手法では、光源として点光源だけでなく面光源も利用することでこの問題を解決する。図 1 に示すように左右の光源方向の鏡面反射から、中央の光源方向の鏡面反射の予測を行う。この際、図 1(a), (c) のように左右の光源として点光源を用いた場合は、鏡面反射を観測できる領域が狭く、図 1(b) に示す中央方向での鏡面反射光の振る舞いを予測することが困難である。一方、図 1(d), (e) のように左右の光源として面光源を用いた場合は、中央方向での鏡面反射光を観測できるため、点光源だけでなく面光源を組み合わせることで、図 1(b) に示す中央方向での鏡面反射光の振る舞いの予測が可能であると考えられる。さらに、提案手法では、面光源を含む複数の光源の組み合わせも利用することで、鏡面反射成分を効率よく捉え、少数の照明環境下の画像を用いた再照明を実現する。

## 2.2 センサと画像処理の同時最適化

CV 分野において深層学習を応用した研究が多く行われている。従来の応用としては、撮影済みの画像を用いて、

所望の画像処理タスクの処理方法のみを学習により最適化することで、様々な画像処理タスクにおいて性能の向上を実現してきた。しかし、学習に使われる画像を撮影する際に利用するカメラや照明の設計まで工夫することは少なく、工夫を施す場合も信号処理理論などの解析的な手法で行われることが一般的である。

これに対し、近年、撮影済みの画像の処理方法のみならず、深層学習の枠組みにおいて、カメラのレンズやフィルタ等のセンサ部分の設計までも同時に最適化するディープセンシング [4], [5], [6] と呼ばれる手法が提案されている。Chakrabarti ら [4] は、カラーカメラに使用されているカラーフィルタアレイ、および、それを用いて撮影された画像からフルカラー画像の生成を行う手法を DNN の枠組みで同時に最適化する手法を提案した。しかし、ディープセンシングではセンシング部分の最適化は行っているものの、照明環境の最適化は行っていない。

そこで、提案手法では、撮影済みの画像の処理方法のみならず、照明環境も最適化するディープライティングという手法を提案する。具体的には、提案手法では、照明環境の最適化として特定の大きさ・方向の光源の組み合わせ方の最適化を行う。光源には点光源だけでなく面光源も利用し、さらに照明環境として複数光源の同時点灯を許容することで、再照明のために照明環境までも CNN の枠組みで同時に最適化する。

## 3. 提案手法

### 3.1 学習とテスト

少数の照明環境下の画像から再照明を行うために、テストの際は、図 2 のように、入力として最適な照明環境で撮影した少数の画像（最適な点灯パターン下画像）を使用し、特定の単一光源下の再照明画像を出力とするネットワークを考える。学習の際は、様々な大きさ・方向の単一光源下画像を入力とし、CNN の枠組みにおいて、図 2 の撮像層で照明環境を、補間処理ネットワークで非線形な補間処理

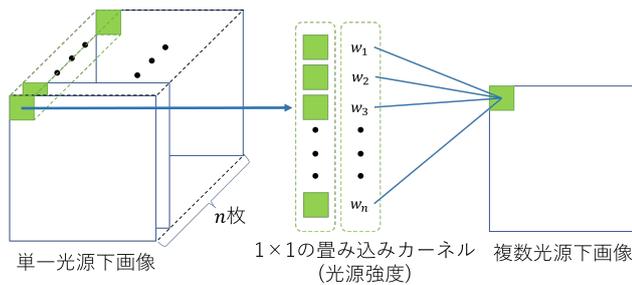


図 3: 1 × 1 の畳み込み

手法を学習することで、照明環境および補間処理の同時最適化を行う。

### 3.2 光源点灯パターン

最適な照明環境として、様々な大きさ・方向の光源の組み合わせ(点灯パターン)を考える。複数の光源を同時点灯して撮影した画像を縦ベクトルで表現した  $i_{mul}$  は、重ね合わせの原理により、単一光源下で撮影された画像を横に並べて行列で表現した  $I$  を用いて、

$$i_{mul} = Iw \quad (1)$$

のように線形結合で表現でき、各光源の明るさ  $w$  が結合係数となる。このような線形結合は CNN の枠組みにおいて 1 × 1 の畳み込みで、結合係数である各光源の明るさ  $w$  は 1 × 1 の畳み込みカーネルで表現される(図 3)。このことから、最適な点灯パターンの学習は、図 2 の撮像層における 1 × 1 の畳み込みカーネルの学習に帰着する。

### 3.3 同時最適化ネットワーク

再照明画像の生成において、前述のように、任意光源方向下の拡散反射成分は、異なる光源方向下で撮影した 3 枚の基底画像の線形結合により補間できるのに対して、鏡面反射成分の生成には非線形の補間が必要である。そこで、少数の最適な照明環境下の画像を入力とし、特定の単一光源下画像を生成する非線形補間を図 2 の補間処理ネットワークで学習する。具体的には、補間処理ネットワークとして、U-Net[7] のデコーダ部分を 2 つに増やしたネットワークを用いる(図 4)。デコーダ部分の層を増やすことでより多くの特徴マップを用いた学習を行うとともに、スキップ接続により大域的な特徴と局所的な特徴をとらえることで、拡散反射などの低周波数成分および鏡面反射などの高周波数成分の両方の補間が可能であると考えられる。

ネットワークの前半部分では、256 × 256 サイズの画像を入力とし、カーネルサイズ 3 × 3 の畳み込み処理後、バッチ正規化 [8] を行い、活性化関数 ELU による非線形変換を行う処理を 2 回経て、2 × 2 サイズの max pooling を行う操作を 4 回繰り返す。活性化関数として ELU を使用することで、活性化関数として ReLU を使用する場合に入力値が

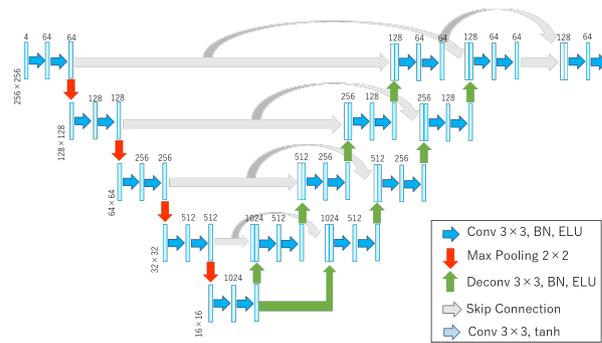


図 4: 画像補間ネットワーク

0 以下で勾配が消失する問題を解消でき、より学習が進みやすくなると考えられる。その後、同様に畳み込み、バッチ正規化、活性化関数 ELU による非線形変換処理を 2 回行う。ネットワークの後半部分では、カーネルサイズ 3 × 3 の逆畳み込み操作によって特徴量の拡大を行い、前半同様バッチ正規化、活性化関数 ELU による非線形変換を行う。この出力と、前半の畳み込み層を経た出力をスキップ接続により連結した特徴マップを入力とし、前半と同様の畳み込み、バッチ正規化、活性化関数 ELU による非線形変換の操作を 2 回繰り返す。この操作を 2 つのデコーダ部分で 4 回繰り返すことで、入力と同解像度の出力を得る。最後に 2 つのデコーダの出力を連結した特徴マップを入力とし、畳み込み、バッチ正規化、活性化関数 ELU による非線形変換を 2 回行い、最後にカーネルサイズ 3 × 3、活性化関数 tanh の畳み込み処理によって再照明画像の生成を行う。

## 4. 実験

提案手法の有効性を示すために実画像を用いた実験を行った。実験では光源として液晶ディスプレイを利用した。今回の実験では、図 6(a) に示す 4 方向の光源下の画像から、図 6(d) に示す 5 方向の光源をそれぞれ点灯した単一光源下の画像を生成する。学習の際は、学習用画像として図 6(a), (b), (c) の 4 方向・3 サイズの光源をそれぞれ順番に点灯した単一光源下画像 12 枚、正解画像として図 6(d) の 5 方向の光源をそれぞれ順番に点灯した単一光源下画像 5 枚を用いた。今回の実験では学習データとして 8 シーン、評価データとして 1 シーン、テストデータとして 5 シーンの画像を準備して実験を行った。限られた学習データから効率的に学習するために、画像データを 256 × 256 のサイズの小画像としてランダムに切り出して使用した。

画像補間ネットワークは、最適化手法として Adam オプティマイザー [9] を用いて学習を行い、学習率は  $3.0 \times 10^{-4}$  とした。損失関数としては MSE (最小二乗誤差)、評価関数は MSE, SSIM を用いて学習を行った。また、全ての重みの初期値は He の初期値 [10] を利用した。

図 5 に、(a) 正解画像、(b) 図 6(a) の 4 方向の単一点光源下画像 4 枚を用いて単純な線形補間により生成した画像、

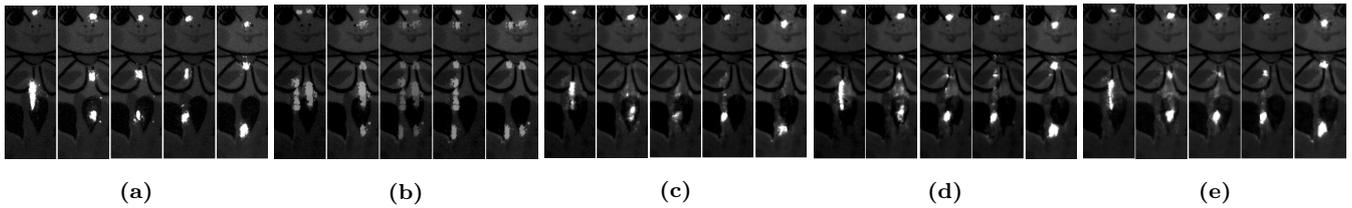


図 5: 正解画像と再照明結果 (鏡面反射の周囲を切り出し)

(c) 図 6(a) の 4 方向の単一点光源下画像 4 枚を用いて非線形補間により生成した画像, (d) 最適化された照明環境下画像 4 枚を用いて非線形補間により生成した画像 (提案手法), (e) 図 6(a), (b), (c) の 4 方向・3 サイズの単一光源下画像 12 枚を用いて非線形補間により生成した画像を示す. なお, 比較のため生成結果の一部を抜粋したものを示している. 各画像の順番としては, 左から順に図 6(d) に示す上方向・右方向, 中央方向, 左方向, 下方向の単一光源下の生成結果となっている.

まず, 図 5(b) の単純な線形補間により生成した画像では, 鏡面反射は入力画像の鏡面反射が移動しておらず, 図 5(a) の正解画像と比較すると鏡面反射の位置および外形が大きく異なるため, 補間が不十分であることが確認できる. 次に, 図 5(c) の 4 つの点光源のみを用いて非線形補間により生成された画像は鏡面反射の位置および外形ともに概ね正しい. しかし, 図 5(d) の提案手法と比較すると提案手法の方が良好な結果が得られており, 照明環境の最適化が再照明において有効であることが確認できる. また, 図 5(e) の 12 枚の画像を用いて生成された画像と図 5(d) の提案手法とを比較すると, 図 5(d) は図 5(e) よりも画像数を減らしたにもかかわらず, 鏡面反射の位置および外形が類似しており, 定性的に良好な結果が得られている.

続いて, 図 5(b)-(e) の画像間の定量比較を行った結果を表 1 に示す. 表 1 の結果からわかるように, PSNR, SSIM の両指標において, 図 5(d) の提案手法が図 5(b), (c) の点光源のみを用いて生成された画像よりも良好な結果が得られている. また, 図 5(e) の 12 枚の画像を用いた結果と比較しても大きな差はなく, 複数光源の組み合わせの利用が再照明において有効に機能することが分かる. 併せて, 図 6(e)-(h) に今回の実験で得られた最適な点灯パターンを示す. 最適な点灯パターンでは, 点光源のみならず, 面光源も利用しているため, 再照明において面光源の利用が有効であることが確認できる.

## 5. むすび

本稿では, 再照明のための照明環境と画像補間の同時最適化手法を提案した. 提案手法では, 少数の照明環境下の画像から効率的に鏡面反射成分をとらえるために, 点光源だけでなく様々な大きさの面光源やそれらの組み合わせを利用した. また, 照明環境が CNN の枠組みにおいて  $1 \times 1$

表 1: 定量比較結果

図 5	PSNR $\uparrow$	SSIM $\uparrow$
(b) 線形補間 (単一サイズ光源下 4 枚)	36.85	0.932
(c) 非線形補間 (単一サイズ光源下 4 枚)	39.94	0.948
(d) 提案手法: 非線形補間 (最適照明下 4 枚)	40.49	0.949
(e) 非線形補間 (複数サイズ光源下 12 枚)	40.53	0.948

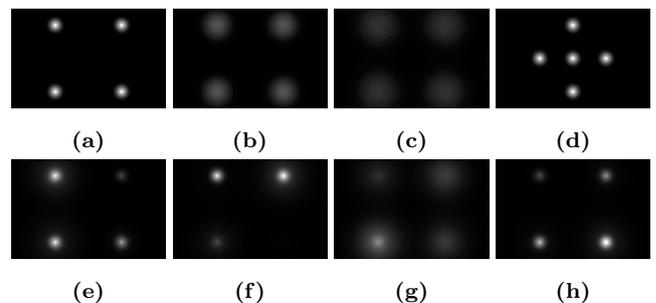


図 6: 光源点灯パターン

の畳み込みカーネルを用いて表現できることに着目し, 再照明のために照明環境と補間処理の両方を同時に最適化する新たなアプローチを提案した. 実画像を用いた実験により, 照明環境および補間処理の同時最適化が, 再照明において有効であることを確認した. 影などの他の画像中の成分に対する本手法の拡張が今後の課題である.

**謝辞** 本研究の一部は, JSPS 科研費 JP20H00612 の助成を受けた.

## 参考文献

- [1] A. Shashua, "On photometric issues in 3D visual recognition from a single image", IJCV, Vol.21, No.1-2, pp.99-122, 1997.
- [2] P. Debevec, T. Hawkins, C. Tchou, H. Duiker, W. Sarokin, and M. Sagar, "Acquiring the reflectance field of human face", In Proc. SIGGRAPH2000, pp.145-156, 2000.
- [3] M. Fuchs, H. Lensch, V. Blanz, and H. Seidel, "Super-resolution Reflectance Fields: Synthesizing images for intermediate light directions", In Proc. EGSR2007, pp.447-456, 2007.
- [4] A. Chakrabarti, "Learning sensor multiplexing design through back-propagation", In Proc. NIPS2016, pp.3081-3089, 2016.

- [5] S. Nie, L. Gu, Y. Zheng, A. Lam, N. Ono, and I. Sato, “Deeply learned filter response functions for hyperspectral reconstruction”, In Proc. IEEE CVPR 2018, pp.4767–4776, 2018.
- [6] C. A. Metzler, H. Ikoma, Y. Peng, and G. Wetzstein, “Deep optics for single-shot high-dynamic-range imaging”, In Proc. IEEE/CVF CVPR 2020, pp.1375–1385, 2020.
- [7] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation.”, In Proc. MICCAI 2015, pp.234–241, 2015.
- [8] S. Loffe, and C. Szegedy, “Batch normalization: Accelerating deep network training by reducing internal covariate shift.”, In: Proc. ICML 2015, pp.448–456, 2015.
- [9] D. P. Kingma, and J. Ba, “Adam: A method for stochastic optimization”, In: Proc. ICLR2015, 2015.
- [10] K. He, X. Zhang, S. Ren, and J. Sun, “Delvin Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification”, In Proc. IEEE ICCV 2015, pp.1026–1034, 2015.