

3DCG を用いた光源推定による自然画像からの光源色の影響低減に関する研究 A Study on Reduction of Illuminations Color Influences from Natural Images by Illumination Estimation Using 3DCG Images

青沼 奏志[†] 中村 嘉彦[‡] 山本 椋太[‡]
Soshi Aonuma Yoshihiko Nakamura Ryota Yamamoto

1. はじめに

1.1 光源と光源色の影響

実世界は何らかの光源に照らされている。光源には色があり、物体の色の見え方は光源色の影響で、本来の色から変化する。人間の視覚には、色恒常性という能力がある。色恒常性とは、ある物体の色を光源色の影響にかかわらず、同じ色として認識できる能力である[1]。

デジタルカメラで撮影した画像においても、画像中の物体の色は光源色の影響によって本来の色から変化する[1]。デジタルカメラでは、計算上の色恒常性という処理により、撮影した画像から光源色の影響を除去し、白色光源の元で撮影された画像に変換する[1]。写真では、計算上の色恒常性はホワイトバランス(WB)と呼ばれることが多い[2]。計算上の色恒常性において、光源推定という手順が必要であるが、光源推定は不良設定問題である[1]。誤った WB が原因で、深層学習と画像処理を組み合わせたアプリケーションの精度が低下することがある。

1.2 深層学習による光源推定

現在、深層学習を用いた光源推定手法が提案されている。深層学習にはデータセットが必要であり、様々な光源推定用データセットが提案されている。しかし、光源推定用データセットの作成には、被写体、光源、地理的な場所、季節、時刻の多様性が要求され、それらを満足するデータセットは多くない[3]。

また、深層学習による光源推定の問題点として、任意のカメラセンサに一般化できないことが挙げられる。カメラセンサの型が異なれば、分光感度も異なり、同じシーンを撮影した画像でも色合いが異なる。そのため、学習時とは異なるカメラセンサで撮影された画像の光源推定をする場合、そのカメラセンサで撮影された画像でモデルを再学習、またはファインチューニングしないと十分な精度が出ない。

1.3 RAW 画像と sRGB 画像

光源推定の手法の多くは RAW 画像での推定を意図している[2]。RAW 画像は計算上の色恒常性や後述の非線形な操作が適用されていない画像である。しかし、深層学習と画像処理を組み合わせたアプリケーションでは、一般に sRGB 画像が使われる。sRGB 画像は着色強化、色調操作など、カメラごとに非線形な操作がなされている[2]。仮に、sRGB 画像の光源推定ができたとしても、推定できた領域のみが補正され、ほかの領域は正確に補正できない[2]。

[†] 苫小牧工業高等専門学校 National Institute of Technology, Tomakomai College

[‡] 苫小牧工業高等専門学校 National Institute of Technology, Tomakomai College

1.4 本研究の目的

本研究の目的は、自然画像から光源色の影響を低減し、深層学習と画像処理を組み合わせたアプリケーションの精度向上を実現することである。例えば、物体の色から何らかの推定を行うアプリケーションでは、学習時と実用時の光源色の違いの影響を受ける。光源色の影響を低減する上で、光源推定が重要となる。光源推定における問題は、データセットの作成が高コストであること、モデルがカメラセンサに依存すること、sRGB 画像ではなく RAW 画像を対象としていること、の 3 つであると考えられる。これら 3 つの問題を解決するため、3DCG 画像で構成される光源推定用データセットにより、実写の sRGB 画像の光源推定を行い、光源色の影響を低減する手法を提案する。

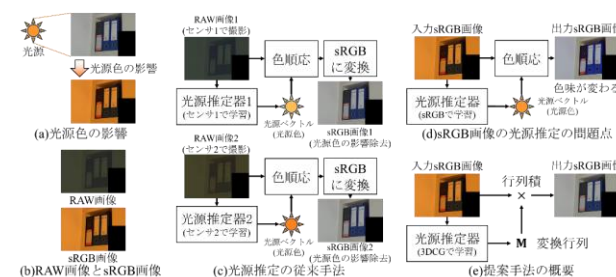


図 1 研究の背景の概要

2. 提案手法

2.1 提案手法の概要

本研究で提案する手法は、sRGB の 3DCG 画像で学習した光源推定器で sRGB の自然画像の光源推定を行い、光源色の影響を推定する。推定結果を用いて、自然画像から光源色の影響を低減する。すなわち、画像を白色光源の元で撮影された画像に近づける変換をする。

主な従来手法では、学習・テスト両方で RAW 画像が用いられている[1]。提案手法では、学習に sRGB の 3DCG 画像、テストに sRGB の自然画像を用いる。

提案手法で用いるネットワークを図 2 に示す。sRGB 画像での光源推定と、光源色の影響の低減を行うために、

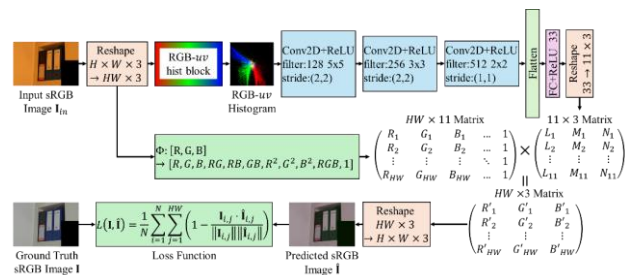


図 2 提案手法のネットワーク

RGB-uvヒストグラムを用いる。RGB-uvヒストグラムの色分布を用いて光源推定を行い、 11×3 の変換行列を推定し、入力画像の光源の影響を低減する。

2.2 光源推定

光源推定とは、計算上の色恒常性の 1 番目の手順である。計算上の色恒常性とは、画像シーン中の物体の色から、光源色の影響を除去することである[1]。計算上の色恒常性は 2 つの手順からなる。1 番目の手順は光源推定であり、画像から光源ベクトル(光源色)を推定する。光源ベクトルは 3 成分ベクトルであり、それぞれの要素は光の三原色(赤, 緑, 青)に対応する[1]。2 番目の手順は色順応と呼ばれ、画像中の各ピクセルの値を、推定した光源ベクトルの対応する要素の値で除算する。色順応が適用されることで、画像中の物体の色から光源色の影響が除去され、白色光源の元で撮影された画像に変換される[1]。

計算上の色恒常性における画像形成モデルとして、Lambertian 仮定を用いたものが主に使われる[1]。

$$f_c(\mathbf{x}) = \int_{\omega} I(\lambda, \mathbf{x}) R(\lambda, \mathbf{x}) \rho_c(\lambda) d\lambda \quad (1)$$

ここで、 $c \in \{R, G, B\}$ は色チャンネル、 \mathbf{x} はピクセル座標、 f はピクセル値、 $I(\lambda, \mathbf{x})$ は光源の分光分布、 $R(\lambda, \mathbf{x})$ は表面反射率、 $\rho_c(\lambda)$ は色チャンネル c のカメラセンサの分光感度、 λ は可視光スペクトル ω の波長である。画像シーン中の光源が 1 つかつ、光源ベクトル \mathbf{e} がピクセル座標 \mathbf{x} に依存せず均一であると仮定すると、光源ベクトル \mathbf{e} は次で表される[1]。

$$\mathbf{e} = \int_{\omega} I(\lambda) \boldsymbol{\rho}(\lambda) d\lambda \quad (2)$$

ほとんどの場合、 $I(\lambda)$ と $\boldsymbol{\rho}(\lambda)$ は不明であり、画像のピクセル値 f しか分かっていない。加えて、ピクセル値 f に対応する光源と表面反射率の組み合わせは無限に存在する。このため、光源推定は不良設定問題である[1]。

2.3 3DCG 画像

3DCG の物体色は Phong の反射モデルで計算する。これは、実世界の光源と反射現象を単純化したものであり、拡散反射、鏡面反射、環境光という 3 つの要素の加算によって決定される。拡散反射とは、光が入射点から全方位に反射する現象である。鏡面反射とは、光が正反射方向に反射する現象である。環境光とは、別の物体表面での反射を経て、物体表面に届く光である。

2.4 RGB-uvヒストグラム

RGB-uvヒストグラムは、画像の色分布を uv でパラメータ化した $m \times m \times 3$ テンソルとして表現したものである[2]。RGB-uvヒストグラムは以下の関数 $\mathbf{h}(\mathbf{I})$ で生成される[2]。

$$\begin{aligned} \mathbf{I}_{y(i)} &= \sqrt{\mathbf{I}_{R(i)}^2 + \mathbf{I}_{G(i)}^2 + \mathbf{I}_{B(i)}^2}, \\ \mathbf{I}_{u1(i)} &= \log(\mathbf{I}_{R(i)}) - \log(\mathbf{I}_{G(i)}), \\ \mathbf{I}_{v1(i)} &= \log(\mathbf{I}_{R(i)}) - \log(\mathbf{I}_{B(i)}), \end{aligned} \quad (3)$$

$$\begin{aligned} \mathbf{I}_{u2} &= -\mathbf{I}_{u1}, \mathbf{I}_{v2} = -\mathbf{I}_{u1} + \mathbf{I}_{v1}, \\ \mathbf{I}_{u3} &= -\mathbf{I}_{v1}, \mathbf{I}_{v3} = -\mathbf{I}_{v1} + \mathbf{I}_{u1} \end{aligned}$$

$$\mathbf{H}(\mathbf{I})_{(u,v,c)} = \sum_i \mathbf{I}_{y(i)} \left[\left| \mathbf{I}_{uc(i)} - u \right| \leq \frac{\varepsilon}{2} \wedge \left| \mathbf{I}_{vc(i)} - v \right| \leq \frac{\varepsilon}{2} \right] \quad (4)$$

$$h(\mathbf{I})_{(u,v,c)} = \sqrt{\frac{\mathbf{H}(\mathbf{I})_{(u,v,c)}}{\sum_{u'} \sum_{v'} \mathbf{H}(\mathbf{I})_{(u',v',c)}}} \quad (5)$$

ここで、 $i = \{1, \dots, N\}$ 、 R, G, B は画像 \mathbf{I} 中の色チャンネル、 N は画像のピクセル数、 $C \in \{1, 2, 3\}$ はヒストグラム中の色チャンネル、 ε はヒストグラムのビンの幅である[2]。

3. 結果と考察

sRGB の 3DCG 画像で学習する前に、sRGB の自然画像で学習して光源推定と影響低減ができるか実験した。従来手法[1]と提案手法の誤差を表 1 に示す。従来手法は評価に角度誤差を用いており、推定光源ベクトルと正解光源ベクトルのなす角を求めている。提案手法は出力画像とグラントゥールス画像の各ピクセル間で角度誤差を求め、その平均値を評価に用いている。

表 1 従来手法と提案手法の誤差の比較

手法	平均値	中央値	Avg.
従来手法[1]	1.86	1.27	1.43
提案手法	4.17	3.73	3.87

提案手法で光源色の影響を低減した結果を図 3 に示す。図 3(a)は sRGB の入力画像であり、光源色の影響を除去すると、図 3(c)の画像となる。図 3(b)は sRGB の自然画像で学習したネットワークで光源色の影響を低減した結果である。図 3(d)は図 3(b)と(c)のピクセルごとの角度誤差を可視化した結果であり、誤差が大きくなるほど明るく表される。



図 3 提案手法で光源色の影響を低減した結果

図 3(d)より、図 3(a)の暗い部分は誤差が大きいことがわかる。暗い部分は影になっている、または黒色の物体であり、光源色の影響を受けにくいと考える。そのため、暗い部分を除去することで、誤差が減少するかどうか、検証する必要があると考えた。

また、今回は学習に sRGB の自然画像を用いたが、sRGB の 3DCG 画像を用いることで、どの程度の精度が得られるか検証・評価する必要がある。

4. おわりに

本研究の目的は、sRGB の自然画像から光源色の影響を低減し、深層学習と画像処理を組み合わせたアプリケーションの精度向上を実現することである。そのために sRGB の自然画像から光源色の影響を低減する手法を提案した。

今後の課題として、3DCG 画像で学習した場合、どの程度の精度が得られるか検証・評価する必要がある。

参考文献

- [1] Karlo Koščević, Marko Subašić, Sven Lončarić, “Deep Learning-Based Illumination Estimation Using Light Source Classification”, IEEE Access, Vol.8 (2020).
- [2] Mahmoud Afifi, Brian Price, Scott Cohen, Michael S. Brown, “When Color Constancy Goes Wrong: Correcting Improperly White-Balanced Images”, 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2019).
- [3] Egor Ershov, Alexey Savchik, Illya Semenov, Nikola Banić, Alexander Belokopytov, Daria Senshina, Karlo Koščević, Marko Subašić, Sven Lončarić, “The Cube++ Illumination Estimation Dataset”, IEEE Access, Vol.8 (2020).