

影に基づく光源分布とカメラ応答関数の同時推定

中島 拓人^{1,a)} 岡部 孝弘^{2,b)}

概要：本研究では、単一画像から撮影シーンの光源分布を推定する手法を提案する。従来の光源推定では、カメラ応答関数と呼ばれる物理的な明るさを画素値に変換する関数が線形であることを仮定している。しかし、一般に民生用カメラは非線形の応答関数を持つことが知られているため、従来手法をそのまま適用することができない。カメラ応答関数による変換を考慮した推定手法も提案されているが、被写体を凸物体と仮定しており、推定可能なシーンが限られてしまう。そこで提案手法では、影に基づいて、光源分布とカメラの逆応答関数を同時に推定する。具体的には、光源分布と応答関数を基底関数の線形結合で表現し、その結合係数を求めることによりこれを実現する。

キーワード：光源分布推定、カメラ応答関数、影、球面調和関数

1. はじめに

形状・反射特性・光源環境などの情報を利用して画像を生成する技術を一般的にレンダリングと呼ぶ。それとは逆に、1枚の画像から形状・反射特性・光源環境などのシーンの記述を復元する技術はインバースレンダリングと呼ばれ、コンピュータビジョンの分野において中心的な研究課題の一つである。その中でも、インバースライティングと呼ばれる光源環境推定は、拡張現実や複合現実において光学的整合性を実現するために重要である。これにより、CG物体と現実シーンの違和感のない合成を可能になる。

従来研究では、カメラの特性についてあまり考慮されておらず、シーンの明るさを画素値に変換するカメラ応答関数が線形であることを仮定していた。しかしながら、民生用カメラは、非線形のカメラ応答関数を持つことが知られており、応答関数が線形である特殊なカメラで撮影された画像でなければ光源環境を正確に推定することができない。

そこで本研究では、ある物体が他の物体に落とす影であるキャストシャドウなどを手掛かりとして、未知のカメラ応答関数を持つカメラで撮影された単一画像から、シーンの光源分布とカメラ応答関数を同時に推定する手法を提案する。これにより、インターネット画像等のカメラ応答関数未知の画像においても、光源分布を推定し、光学的整合性を実現することが可能になる。

具体的には、実際の明るさである輝度値と光源分布、画素値のそれぞれの関係から、光源分布とカメラ応答関数を基底関数の線形結合で表現し、その結合係数を求めることで、同時推定を実現する。さらに、キャストシャドウなどの輝度変化の大きい影などを考慮し、光源分布をより多くの基底を用いて表現する。

2. 関連研究

2.1 光源分布の推定

画像から光源環境の推定を行うインバースライティングについて述べる。光源が2つあるようなシーンを考える。このとき、2つの光源で照らされたシーンの輝度値は、それぞれ1つの光源で照らされたシーンの輝度値の和で表される。これは輝度値について、重ね合わせの原理が成り立つからである。Marschnerらは、この原理を利用して光源分布の推定を行っている[4]。

光源分布を求めたい入力画像の被写体の形状と反射特性、カメラの位置情報が既知であると仮定する。基底となる光源分布を複数用意し、被写体の形状と反射特性、カメラの位置情報が既知であることを利用して、それぞれの基底光源によって照らされた被写体を入力画像と同じカメラ位置で撮影した基底画像を作成する。このとき、重ね合わせの原理により、入力画像は作成した基底画像の線形結合で近似することができる。この線形結合の結合係数と基底光源分布から入力画像の光源分布を求めることができる。

他の光源環境の推定手法として、Satoらは物体表面の輝度情報のみでなく、入射光によるシーン中の遮蔽情報も手掛かりにして、物体表面上の観測される画像輝度からシー

¹ 九州工業大学 大学院情報工学府 先端情報工学専攻

² 九州工業大学 大学院情報工学研究院 知能情報工学研究系

^{a)} takuto_nakashima@pluto.ai.kyutech.ac.jp

^{b)} okabe@ai.kyutech.ac.jp

ンの光源分布を復元する手法を提案している [8].

2.2 カメラ応答関数

一般に, 民生用のカメラで撮影した画像の明るさと物理的なシーンの明るさは異なる. これは輝度値を画像に保存する際, 人が明るさの差異等を知覚しやすいようにするため非線形の変換が施されるためである. この変換関数をカメラ応答関数と呼ぶ.

画像ベースの研究では, カメラ応答関数を線形と仮定することがほとんどである. しかし, 民生用のカメラはさまざまな非線形の応答関数を持つことが知られている. そのため, 応答関数を線形とした手法を行う場合, カメラ応答関数が線形である特殊なカメラで画像の撮影を行うか, 撮影を行うカメラの応答関数を事前に測定し, 測定した関数の逆変換により画像を較正しておかなければならない.

Grossberg らは, 実際のカメラ応答関数のデータを収集して, そのデータベースからカメラ応答関数の統計モデル (Empirical Model of Response: EMoR) を作成し, そのモデルを用いて単一画像からカメラ応答関数の推定を行っている [2].

2.3 カメラ応答関数との同時推定手法

従来の光源分布推定手法では, カメラ応答関数が線形であることを仮定しているため, 特殊なカメラで撮影された画像, もしくは, 撮影したカメラの応答関数を測定して較正した画像でなければ推定を行うことができない.

その問題を解決するために, Ohta らはカメラ応答関数を光源分布と同時に推定する手法を提案している [5]. インバースライティングとカメラ応答関数の統計モデルによる表現により, 光源分布と応答関数を同時推定している. この手法を用いることで, 撮影したカメラ応答関数が未知の画像であっても, 正しい光源分布の推定が可能になる.

この手法では, 画像中の被写体が凸物体であることを仮定している. しかし, 画像中の物体が凸である場合は多くなく, キャストシャドウなどの輝度変化の大きい影が現れてしまうため, 推定できるシーンが限られてしまう.

そこで本研究では, キャストシャドウなどの影を手掛かりとして, 未知のカメラ応答関数を持つカメラで撮影された単一画像から, シーンの光源分布とカメラ応答関数を同時に推定する手法を提案する. これにより, 非線形のカメラ応答関数を持つ民生用のカメラで撮影された画像でも, 被写体が凸物体でない場合でも, 光源分布の推定が可能になる.

3. 提案手法

3.1 光源分布と輝度値の関係

被写体の形状が既知であり, 反射特性が Lambert モデ

ルと呼ばれる理想的な拡散反射を行うモデルに従うと仮定する. このとき, 入力画像は, 凸物体であれば 2 次までの球面調和関数による基底光源で照らされた 9 枚の画像の線形結合で近似的に表現できる [6], [7]. 非凸物体の場合は, 影を考慮するため, 3 次以上の球面調和関数も用いて, 10 枚以上の基底画像の線形結合で入力画像を表す.

ある光源分布 $L(\theta, \phi)$ は, 基底の光源分布を $Y_n(\theta, \phi)$, 光源分布の結合係数を α_n とすると,

$$L(\theta, \phi) = \sum_{n=1}^N \alpha_n Y_n(\theta, \phi) \quad (1)$$

で表せる.

カメラ応答関数が線形のとき, p 番目の画素について, n 個目の基底光源 $Y_n(\theta, \phi)$ で照らされた画像の画素値を $R_p(Y_n(\theta, \phi))$ とすると, 放射輝度値 I_p は,

$$I_p = \sum_{n=1}^N \alpha_n R_p(Y_n(\theta, \phi)) \quad (2)$$

で与えられる.

3.2 画素値と輝度値の関係

カメラ応答関数にはガンマ補正による表現や統計モデルによる表現など様々な表現方法があるが, 本手法では多項式表現を用いる [3]. カメラ応答関数はシーンの明るさを画素値へ変換するものであるが, 画素値から実際の明るさを知りたいため, 逆の変換を行うカメラの逆応答関数 g について考える. p 番目の画素の画素値を I'_p , 逆応答関数の結合係数を β_m として, 画像の放射輝度値 I_p を

$$I_p = g(I'_p) = I'_p + I'_p(I'_p - 1) \sum_{m=1}^M \beta_m I'^{M-m}_p \quad (3)$$

のように近似的に表現する [3]. ただし, 放射輝度値と画素値は, それぞれ $0 < I < 1$, $0 < I' < 1$ の範囲に正規化されているものとする. 物理的な明るさとカメラ応答関数によって変換された画素値の明るさの大小関係は変わらないため, カメラ応答関数は単調増加である. よって, 逆応答関数も単調増加であるので, 任意の画素値 I'_p について逆応答関数は,

$$g(I'_p) < g\left(I'_p + \frac{1}{255}\right) \quad (4)$$

という制約を満たす.

3.3 光源分布とカメラ応答関数の同時推定

インバースライティングによる光源分布の推定方法とカメラ応答関数の多項式表現を用いて同時推定を行う. 式

(2) と式 (3) の右边が等しいことから,

$$\sum_{n=1}^N \alpha_n R_p(Y_n(\theta, \phi)) = I'_p + I'_p(I'_p - 1) \sum_{m=1}^M \beta_m I'_p{}^{M-m} \quad (5)$$

が成り立つ. 各画素について式 (5) が成り立つため, 画像の画素数を P とすると, 方程式の数が P 個で未知数が $(M + N)$ 個の 1 次の連立方程式が得られる. したがって, $P > (M + N)$ であれば,

$$\{\alpha, \beta\} = \arg \min_{\{\alpha_n, \beta_m\}} \sum_{p=1}^P \left[\sum_{n=1}^N \alpha_n R_p(Y_n(\theta, \phi)) - I'_p - I'_p(I'_p - 1) \sum_{m=1}^M \beta_m I'_p{}^{M-m} \right]^2 \quad (6)$$

という最適化問題を, 式 (4) の逆応答関数が単調増加であるという制約の下で解くことで, 光源分布とカメラの逆応答関数の結合係数 α_n と β_m を求めることができる.

しかし, 実際は逆応答関数が単調増加という制約だけでは推定が不安定になってしまうため, カメラ応答関数についての平滑化項

$$c \sum_{l=1}^{255} \left[\frac{\partial^2 g(I'_l)}{\partial I'^2_l} \right]^2 \quad (7)$$

を加える. ここで, I'_l は 8bit の画素値が 0 から 1 の範囲で正規化されているもの, つまり, $I'_l = l/255$ であり, c は平滑化項の重みを表す. 式 (7) を式 (6) に追加して, 光源分布とカメラ応答関数の同時推定を行う.

4. 実験

実際に非線形のカメラ応答関数により変換された CG 画像について推定を行った. 図 1(a) のような HDR の光源データ [1] を用いて影が含まれるような画像 (b) を作成し, 非線形のカメラ応答関数を適用することで, 入力画像 (c) を作成した. 入力画像 (c) に対して, カメラ応答関数が線形であると仮定して推定を行う従来手法と提案手法を用いて, それぞれ推定を行った. 従来手法と提案手法で推定した光源分布をそれぞれ (d) と (e) に示す. (f) は入力画像を作成する際に適用した逆応答関数の真値と提案手法の推定結果をグラフで表したものである. このシーンにおける推定では, 3 次までの球面調和関数による基底光源を用い, 逆応答関数を 5 次の多項式で近似した. カメラ応答関数は, $\gamma = 1.5$ として, 入力画像を作成した. 図 1 より, 光源分布, カメラの逆応答関数共に推定結果が良好であることがわかる.

さらに光源分布とカメラの逆応答関数の推定結果について, RMSE(Root Mean Squared Error) による定量的評価を行った. 評価結果を表に示す. 表 1 より, 定量的にみて

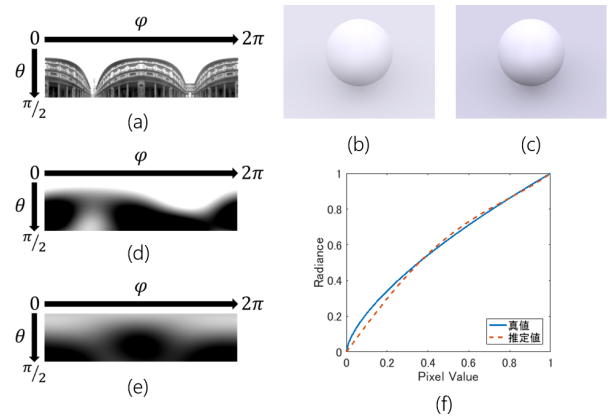


図 1 推定結果 1

表 1 定量的評価結果 1

光源分布 (従来手法)	光源分布 (提案手法)	逆応答関数 (提案手法)
4.10	3.12	0.0292

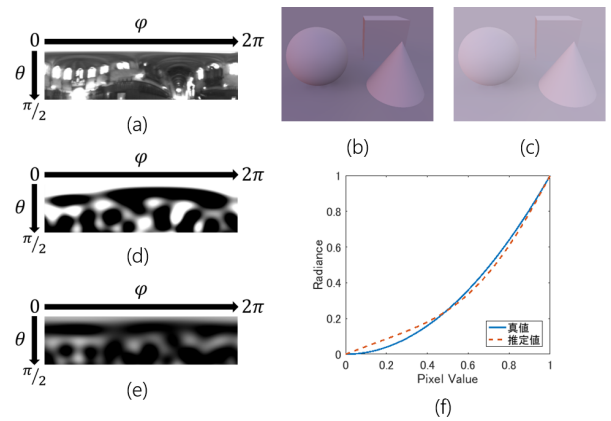


図 2 推定結果 2

表 2 定量的評価結果 2

光源分布 (従来手法)	光源分布 (提案手法)	逆応答関数 (提案手法)
5.90	5.27	0.0285

も提案手法で良好な結果が得られていることがわかる.

別のシーンにおいても同様の推定を行った. こちらの推定では, 10 次までの球面調和関数による基底光源を用い, 逆応答関数を 5 次の多項式で近似した. カメラ応答関数は, $\gamma = 0.5$ として, 入力画像を作成した. 図 2, 表 2 より, このシーンについても光源分布, カメラの逆応答関数共に推定結果が良好であることがわかる.

5. むすび

本稿では, シーン中の影に基づいた, カメラ応答関数未知の画像における光源分布の推定手法を提案した. CG 画像によるシミュレーション実験による, 従来のカメラ応答関数を線形と仮定したインバースライティングと比較し, 定性的, 定量的に本手法の有効性を示した.

今後は、実画像実験を行うと共に、球面調和関数の最適な次数や別の基底光源など光源分布の表現方法について考えていきたい。

謝辞 本研究の一部は、JSPS 科研費 JP17H00744 の助成を受けた。

参考文献

- [1] P. Debevec, “Rendering synthetic objects into real scenes: bridging traditional and image-based graphics with global illumination and high dynamic range photography,” ACM SIGGRAPH 1998, pp.189–198, 1998.
- [2] M. Grossberg and S. Nayar, “What is space of camera response functions?,” CVPR 2003, pp.602–609, 2003.
- [3] J.-Y. Lee, Y. Matsushita, B. Shi, I. S. Kweon, and K. Ikeuchi, “Radiometric calibration by rank minimization,” IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.35, No.1, pp.144–156, 2013.
- [4] S. Marschner and D. Greenberg, “Inverse lighting for photography,” IS&T/SID Fifth Color Imaging Conference, pp.262–265, 1997.
- [5] S. Ohta and T. Okabe, “Does inverse lighting work well under unknown under response function?,” VISAPP 2015, pp.652–657, 2015.
- [6] R. Ramamoorthi and P. Hanrahan, “An efficient representation for irradiance environment maps,” In Proc. ACM SIGGRAPH 2001, pp.497–500, 2001.
- [7] R. Ramamoorthi and P. Hanrahan, “A signal-processing framework for inverse rendering,” ACM ToG (In Proc. ACM SIGGRAPH 2001), Vol.20, No.3, pp.117–128, 2001.
- [8] I. Sato, Y. Sato, and K. Ikeuchi, “Illumination from shadows,” IEEE Trans. PAMI, Vol.25, No.3, pp.290–300, 2003.