

RGB-D カメラから得られる法線マップを用いた 照明推定とリライティング

小倉 洋平^{*1}

斎藤 英雄^{*1}

Illumination Estimation and Relighting based on Normal Map from an RGB-D Camera

Yohei Ogura^{*1} and Hideo Saito^{*1}

Abstract – In this paper, we propose a relighting system which combined with an illumination estimation method using RGB-D camera. Relighting techniques can achieve the photometric consistency of composite images. They often need illumination environments of the scene which include a target object and the background scene. Some relighting method obtain the illumination environments beforehand. In our method, the illumination environment of the scene which includes the target object is estimated from pixel intensity, normal map and surface reflectance based on Inverse rendering in online processing. The background image and the illumination environment are obtained from an omni-directional image. Relighting is achieved by calculating the ratio for the estimated Illumination environment of the each scene. Thus our implementation can be used for dynamic illumination or a dynamic object.

Keywords : illumination estimation, relighting, normal map, RGB-D camera

1 はじめに

任意の物体や人物を他のシーンに合成する場合、両シーンの照明環境が異なることが多く、単純に合成させただけでは違和感が生じてしまう。この場合、幾何学的整合性の他に光学的整合性を実現することが重要である。光学的整合性の実現によって、撮影した環境とは異なるシーンや、CG で制作されたシーンに、実際の人物や物体を合成する際に違和感を軽減できるため、映像コンテンツ制作やエンタテインメント向けアプリケーションなどへの応用が考えられる。このような場合、光学的整合性を実現する手法として、リライティングがある。対象となる物体または人物への光の当たり方を変更することで、他のシーンに合成する際に、よりリアルな映像を得ることができる。リライティングには一般的に、対象となる物体の3次元形状と表面の反射特性、撮影した環境と合成先の2つの照明環境のデータが必要となる。

照明推定の既存の手法として、鏡面球を用いる手法[1]がある。鏡面球はあらゆる方向の風景を反射するので、これをキャプチャすることで、物体を照らす照明環境を入手することができる。この手法の場合、カメラの位置や照明が変化した際は鏡面球をキャプチャし直す必要がある。動的な照明下では常に鏡面球を撮影してキャプチャする必要があるため、シーンに制約が生じる。一方、Gruber ら [2] は、物体を3次元復元

することで法線マップを入手した後、インバースレンダリングに基づき、明度値、反射率、法線マップから、撮影されたシーンの照明環境を推定する手法を提案した。この手法は3次元復元をおこなうため、常に形状が変化する物体への適用は難しく、また照明は白色のみとしている。

リライティングの手法としては、Debevec ら [3] が提案した、Light Stage という大量の点光源を用いた装置を用いる手法がある。3次元形状の取得は不要であるが、対象物体を異なる光源で照らした際の画像を大量に取得する必要がある、実時間でリライティングをおこなうためには高速で光源を切り替えながら撮影をおこなう装置が必要となる。他には、Zhen ら [4] が提案した手法がある。顔のモーフィングモデルから3次元形状を取得し、あらかじめ取得した照明環境を用いてリライティングをおこなうものである。この手法は静的な対象、あるいは静止画を対象とした手法である。

そこで本稿では、RGB-D カメラを用いて照明推定とリライティングを実現する手法を提案する。はじめに、事前に求めた物体表面の反射特性を用いて、リライティング対象を含むシーンで照明推定をおこなう。合成先のシーンにおいては、全方位カメラで撮影した画像を用いて照明環境を取得する。RGB-D カメラから得られる距離画像をノイズ除去し、法線マップを求める。これらのデータを用いて人物や物体を、合成先の照明環境に合わせてリライティングをおこなうことで、光学的整合性を実現する。

^{*1}慶應義塾大学

^{*1}Keio University

2 提案手法

本手法は、照明推定とリライティングの2つのパートからなる。照明推定パートでははじめに、RGB-Dカメラから得た距離画像から、物体の法線マップを求める。そして、事前にオフラインで求めておいた物体の反射率、カラー画像から得た明度値と法線マップを用いて撮影されたシーンの照明を推定する。これをリライティング対象を含むシーンに対して実行する。合成先のシーンでは全方位カメラで撮影した全方位画像を用いて、照明環境と背景画像を取得する。リライティングパートでは、推定した照明分布、対象となる物体の法線マップと明度値を用いる。各画素に対して、法線ベクトルと照明分布の比を取り、リライティング後の明度値を得る。これを背景画像に合成することで最終的な出力画像を得る。

2.1 法線マップの取得

本手法では、物体の3次元復元などはおこなわず、RGB-Dカメラから得られる距離画像から法線マップを取得する。法線マップの精度は照明推定、リライティングの両方に影響を及ぼすので重要である。得られた距離画像から直接法線マップを求めると、ノイズの影響によって満足な精度を得られない場合がある。そこで、距離画像に対してバイラテラルフィルタと時間フィルタを使用してノイズを軽減する。その後、カメラパラメータが既知として、距離画像をカメラ座標系で表される3次元点群マップ V に変換し、Holzerら[5]の手法を用いて法線マップを取得する。この手法により、スムーズな法線マップを得ることができるが、物体の境界付近など、距離値の差が大きい部分には適用することが出来ない。これらの領域内の点 $\mathbf{u} = (u, v)$ の法線ベクトル $N(\mathbf{u})$ は以下の式によって求められる。

$$N(\mathbf{u}) = (V(u+1, v) - V(u, v)) \times (V(u, v+1) - V(u, v)) \quad (1)$$

この結果とHolzerらの手法によって得られた結果を組み合わせることで、最終的な法線マップとする。これを各フレームにおいて計算する。

2.2 反射率推定と照明推定

対象物体を含むシーンにおける照明推定は、Gruberら[2]の手法にもとづいている。撮影された物体の明度値、反射率、法線マップからシーンの照明環境を推定する。物体の表面は鏡面反射せず、拡散反射のみとし、光源は無限遠に存在するものと仮定する。この時、物体上の点 $\mathbf{x}$ で観測された画像中の明度値 $I(\mathbf{x})$ は、以下のように表すことができる。

$$I(\mathbf{x}) = R_d \int_{\Omega} L(\omega) \max((\omega \cdot \mathbf{n}(\mathbf{x})), 0) d\omega \quad (2)$$

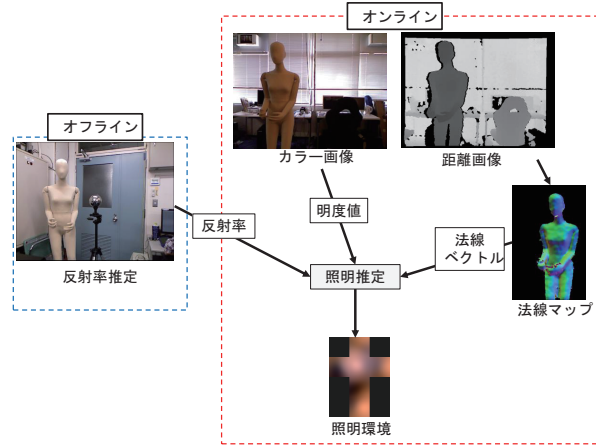


図1 照明推定の流れ
Fig. 1 Flow of illumination estimation

$L(\omega)$ は ω 方向の入射光の強さ、 $\mathbf{n}(\mathbf{x})$ は点 $\mathbf{x}$ における法線ベクトル、 R_d は物体の反射率を示す。 $\max((\omega \cdot \mathbf{n}(\mathbf{x})), 0)$ は入射光の方向ベクトルと法線ベクトルの内積を表し、方向が一致する場合は入射光の強度を全て考慮し、90度以上の場合は0となり、その方向の入射光は考慮しないこととする。

これにより、照明環境を推定することができるが、推定すべき変数の数が膨大になり、処理に時間がかかってしまう。そのため、照明環境をSpherical Harmonicsで近似することで、計算量を削減している。

$$I(\mathbf{x}) = R_d \sum_{l=0}^2 \sum_{m=-l}^l A_l(\theta) L_{l,m} Y_{l,m}(\omega) \quad (3)$$

$I(\mathbf{x})$ は明度値、 $Y_{l,m}$ はSHの基底関数、 $L_{l,m}, A_l(\theta)$ はそれぞれSHに近似した入射光と、法線ベクトルと入射光の方向ベクトルの内積である。

照明分布と反射率を同時に推定することは困難であるので、本手法では事前に既知の照明下で反射率を求めておき、その反射率を用いて、未知の照明環境下において照明推定をおこなう。反射率の推定と照明推定ではどちらも(3)式を用いる。

反射率推定の場合は、明度値、照明環境、法線マップを入力として、物体の反射率を求める。この時、物体の反射率は同じ素材、色であれば場所によらず一定とし、シーン中のある1つの領域の反射率を求める。

照明推定の場合は、画像中から選択された複数のサンプル点の明度値、反射率、法線マップを入力として、(3)式を複数立てる。これを最小二乗法を用いて解くことで、シーンの照明環境を求める。本手法は、照明推定を各フレームにおいて独立におこなうため、照明が動的に変化する場合であっても適用できる。

合成先のシーンの照明環境の取得には、全方位画像を用いる。全方位画像を用いることで、背景画像を取得する際に、カメラ背面に位置するような照明の分

複合・現実：RGB-D カメラから得られる法線マップを用いた照明推定とライティング

布も取得することができる。全方位画像をシーン全体を覆うような無限遠の立方体に投影し、キューブマップ画像を得る。これを球面調和関数で近似し、照明環境を取得する。全方位画像を用いることで、合成先のシーンにおいて反射率が既知の物体が不要となる。

2.3 リライティング

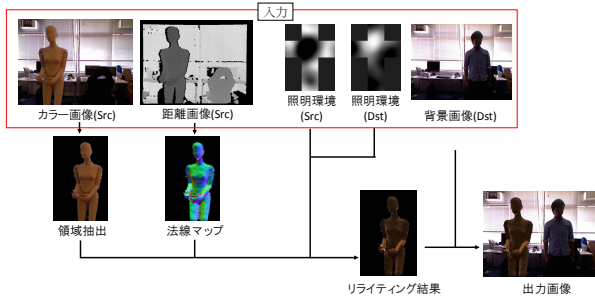


図2 リライティングの流れ
Fig. 2 Flow of relighting

リライティングは、Zhen ら [4] の手法にもとづいている。リライティング対象の物体を含むシーンを Src 、合成先のシーンを Dst とすると、求めたい物は対象物体を Dst の照明にあわせてリライティングした明度値 $I_{Dst}(\mathbf{x})$ である。 Src 、 Dst の各シーンについて (3) 式を立て、その比を取ることで、以下の式を導くことができる。

$$I_{Dst}(\mathbf{x}) = I_{Src}(\mathbf{x}) \frac{\sum_{l=0}^2 \sum_{m=-l}^l A_l(\theta) L_{l,m}^{Dst} Y_{l,m}(\omega)}{\sum_{l=0}^2 \sum_{m=-l}^l A_l(\theta) L_{l,m}^{Src} Y_{l,m}(\omega)} \quad (4)$$

これにより、 Src の物体の法線マップと明度値、そして Src 、 Dst 両方の照明環境を入力として、リライティング結果 $I_{Dst}(\mathbf{x})$ を得ることができる。 Src と Dst の比を取ることで、対象物体の反射率をキャンセルできている。つまり、リライティング処理において物体の反射率を新たに求める必要はない。リライティングした対象物体を Dst の背景に合成することで、最終的な結果画像とする。

3 実験

提案手法の有用性を示すために実験をおこなった。 Src に含まれる人物を、 Dst の照明環境にあわせてリライティングをおこなうことを目的とする。既知の照明下で事前に求めた反射率を用いて、 Src において照明推定をおこなった。 Dst では全方位カメラで撮影された全方位画像を用いて、照明環境と背景画像を取得した。 Src では青いシャツの反射率を使用した。その後、推定した照明環境データを用いて、 Src の人物に対してリライティングをおこない、 Dst のシーンに合成したものを結果画像とした。 Src は動的な対象物体

を含み、 Dst は日中の屋外で撮影したパターン 1 と、夜景を撮影したパターン 2 の 2 つの画像を用いる。(図 2)



図3 Src 環境, Dst 環境
Fig. 3 Src and Dst environment

3.1 リライティングの実施例

推定された照明環境を用いてリライティングをおこない、 Dst のパターン 1 の背景に合成した結果を図 5 に、パターン 2 の背景に合成した結果を図 6 に示す。 Src の照明に比べて、 Dst の照明環境はそれぞれ大きく異なる。提案手法を用いてリライティングをおこなうことで、人物への光の当たり方を変更できている。照明環境が大きく異なる背景に合成した場合でも、光学的整合性を実現できている。また、提案手法では照明推定およびリライティングに用いる法線マップを、距離画像から毎フレーム独立に求めている。 Src において対象が動的な場合であっても、法線マップを取得するため、照明推定およびリライティングを実行することができる。

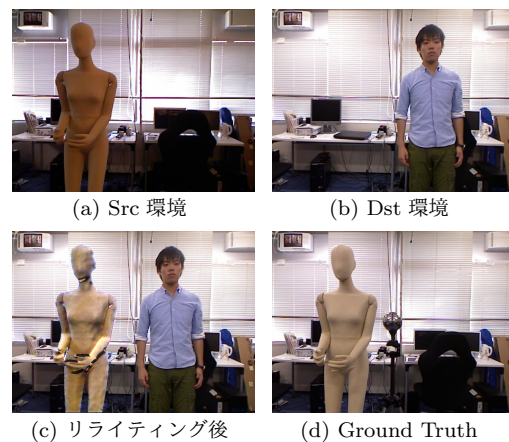


図4 Ground Truth との比較
Fig. 4 Comparison with ground truth

3.2 結果の評価および考察

実験結果の評価として、 Src 環境にマネキンを設置しリライティングをおこない、その結果を Ground



図5 リライティング結果: パターン 1
Fig. 5 Relighting result: pattern 1



図6 リライティング結果: パターン 2
Fig. 6 Relighting result: pattern 2

Truth との RMS(Root Mean Square) Error を全ピクセルに対して計算し、その平均値を各フレームで求めた。この時、*Dst* の照明分布は *Src* と同じ手法で推定した。*Src*, *Dst* 環境, リライティング結果画像, Ground Truth を図 4 に示し、この時の誤差を表 1 に示す。

表 1 Ground Truth との RMS Error(pixel value)
Table 1 RMS error between result and ground truth

	frame: 001	frame:350	frame:700
リライティングあり	17.04	16.87	19.11
リライティングなし	25.01	21.46	24.48

リライティング後の結果は各フレームにおいて Ground Truth との誤差を小さくできていることが分かる。*Src* における照明が変化していても、各フレームで照明を推定しているので、リライティング後の結果への影響を軽減できていると言える。

4 結論

本稿では、RGB-D カメラを用いて照明推定をおこない、それを元に対象となる物体、または人物をリライティングする手法を提案した。これにより、撮影された対象を他の背景に合成する際に光学的整合性を実現することができた。各フレームにおいて独立に照明推定をおこなうため、照明が動的に変化する場合や、対象が静的でない場合であっても適用できる手法を示した。また、シーン中の物体から推定した照明を用い

てリライティングをおこない、リライティングをおこなわなかった場合の画像と比較することで、本手法の有用性を示した。

今後の課題として、法線マップや照明推定の精度向上や、セグメンテーションの改善、鏡面反射成分を持つ物体への対応などがあげられる。

謝辞

本研究の一部は、科学研究費 基盤研究 (S) 24220004 の補助により行われた。

参考文献

- [1] D. Nowrouzezahrai, S. Geiger, K. Mitchell, R. Sumner, W. Jarosz and M. Gross “Light factorization for mixed-frequency shadows in augmented reality,” ISMAR, pp. 173–179, 2011.
- [2] L. Gruber, T. Richter-Trummer, and D. Schmalstieg, “Real-time photometric registration from arbitrary geometry,” ISMAR, pp. 119–128, 2012.
- [3] P. Debevec, T. Hawkins, C. Tchou H. P. Duiker, W. Sarokin and M. Sagar, “Acquiring the reflectance field of a human face,” SIGGRAPH, pp. 145–156, 2000.
- [4] Z. Wen, Z. Liu and TS. Huang, “Face Relighting with Radiance Environment Maps,” IEEE Computer Society CVPR, pp. 158–165, 2003.
- [5] S. Holzer, R. Bogdan Rusu, M. Dixon, S. Gedikli and N. Navab, “Adaptive neighborhood selection for real-time surface normal estimation from organized point cloud data using integral images,” IROS, pp. 2684–2689, 2012.