

画像セグメンテーションに基づいた多光源の照明方向推定の手法

23WM0217 周 海貝

1. はじめに

照明方向推定は画像処理の重要な課題としてこれまで研究され、様々な手法が提案されてきた。その中で、物体表面のシェーディング情報に基づく推定方法が有効かつ軽量であるとされているが、単一シーンに複数の光源が存在する場合への対応が難しいという欠点がある。本研究では、画像セグメンテーションに基づいた複数光源それぞれの照明方向を推定する手法を提案する。まず、各種のセグメンテーションを用いて、画像中の全画素に対してクラスラベルを予測し、物体インスタンスや照明の影響を基準に、異なる領域に分割する。次に、それぞれの領域に対して照明方向推定アルゴリズムを適用し、画素ごとに照明の三次元方向推定を行う。そして、クラスター分析を利用した複数光源に対応できる手法を検討する。最終的には、単一画像から複数光源の方向を正確に推定できる汎用的な方法の実現を目指す。

2. 先行研究

Yi らは、室外のシーンと人の顔のカラー画像における光源方向の推定を行う手法を提案した [2]。具体的には、一定サイズの YCbCr 画像の Y チャンネル画像に対してエッジ検出を行ってから、図 1 のように画像を 4×4 のブロックに 16 等分し、それぞれの領域における複雑度 [3] を計算する。そして、複雑度の低さと平均輝度値の大きさを基準に 3 つの領域を選び、誤差関数と無限光源モデルの制約関数を用いてそれぞれの領域の光源方向を推定し、それらの推定結果を合成したものを最終結果としている。このアプローチにより、画像におけるシェーディング変化が最も緩やかな領域が選択され、照明推定計算に用いられるようになる。

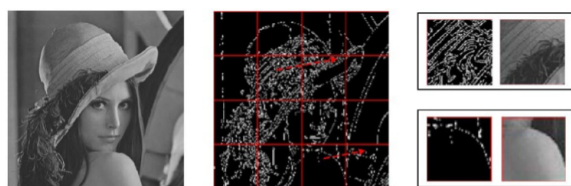


図 1 先行研究における画像分割 [2]

先行研究における実験により、この手法には複数の光源が同時に存在する場合、正しく光源方向を推定できないという課題が存在する。また、反射光源になりうる物体が多数存在する環境や、雨の日や曇りの日など太陽光が雲層で遮蔽され、光がさまざまな方向に散乱する環境において撮影された画像では、光源方向の推定が困難になる。

3. 提案手法

先行研究 [2] の問題の原因は、正方形領域の分割で不連続な三次元空間を扱うことにより、有効な解析や計算ができなくなっているためだと考えられる。そこで、本研究では、各種のセグメンテーション方法の利用により、画像における物体の領域や同一照明から影響を受けている領域を分割し、各領域ごとに照明方向推定を行う手法を検討する。その概要は図 2 のようになっている。

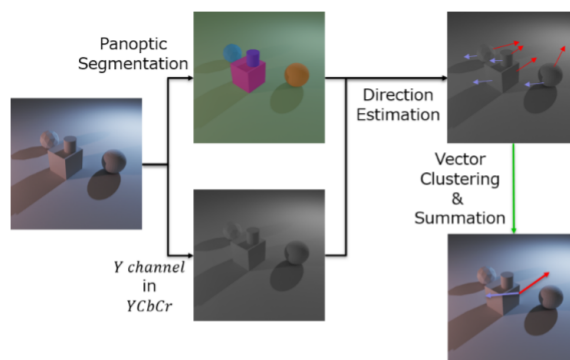


図 2 提案手法の概要図

本研究で利用する手法の一つとして、パノプティックセグメンテーション [1] がある。パノプティックセグメンテーションとは、全画素に定義済みのラベルをつけるセマンティックセグメンテーション [5] と物体の領域だけを分割するインスタンスセグメンテーション [4] を組み合わせたコンピュータビジョンタスクである。その目的は、画像中のシーンを詳しく理解することであり、結果として、画像中の全画素にはそのインスタンスに従ったラベルが付けられる。本研究では、パノプティックセグメンテーションを用いて単一画像に存在する物体と背景の領域を分割し、それぞれの領域に対して光源方向の推定を行う。そして、各領域の推定結果に対してクラスター分析を行い、クラスターごとの推定方向を合成することで複数光源の環境に対応する。

4. 実験

CG 画像を対象とし、提案手法の有効性について検証する。

4.1. 実験用画像の生成

実験に使用する画像とそのセグメンテーションマスクは、blender によって生成する。画像のサイズは 500×500 とし、色空間は RGB とする。画像のシーンには、5 個以下の幾何学物体を配置し、3 個以下の光源（スポットライト）によって照明する。また、光源の色は互いに異なり、位置はシーン内に映らないように設置する。

セグメンテーションマスクは、地面や壁などを含めた画像内のすべての物体の領域に対してラベル付けを行ったものである。それぞれのラベルは、ランダムに分配された異なる色によって表示される。

4.2. 照明方向推定のアルゴリズム

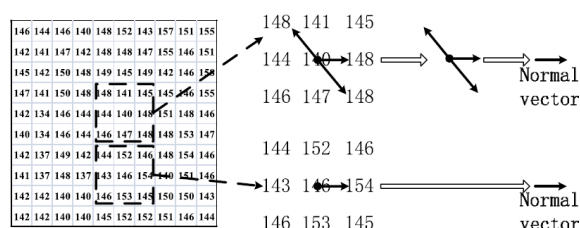


図 3 法線ベクトルの計算 [2]

本研究ではまず、対象画像のモノクロ画像とセグメンテーションマスクに基づいて、各物体領域に対して法線ベクトルに基づく照明方向推定のアルゴリズムを使用する。具体的には、図3のように、対象のピクセルに対して、隣接する8個の周辺ピクセルから最大の画素値を持つものを見つけ、その最大値と対象画像の画素値との差を計算する。その差が正の値の場合は、最大値を持つピクセルの方向を法線ベクトルとし、0以下の値の場合は、零ベクトルを法線ベクトルとする。また、最大画素値を持つピクセルが複数存在する場合は、すべての方向ベクトルを合成したものを法線ベクトルとする。

以上の処理を分割された領域ごとに行い、すべての領域に対して計算することで、画像中の全画素には一つだけの法線ベクトルが存在することになる。

4.3. クラスタ分析

クラスタ分析はデータ解析手法の一種であり、類似する特徴を持つ多次元データを同一のクラスターに分類することができる。本研究では推定された法線ベクトルに対してクラスタ分析を行い、同一の光源からの影響を受けたベクトルを同じクラスターに分類することで、複数の光源が存在する環境に対応する。なお、本実験では階層型手法の一種であるDBSCAN[6]を用いる。DBSCANは、階層的なクラスタ分析手法の一種であり、事前にクラスターの数指定せず分類ができるという特徴がある。

具体的な処理の流れを説明する。まず照明方向推定を終えたそれぞれの領域に対して、一回目のクラスタ分析を行う。半径 ϵ と最小サンプル個数 minPts は実験から得た最適解である(5, 51)とし、対象データの構造は (h, w, y, x, v) とする。ここで、 (h, w) は対象ピクセルの座標を表し、 (y, x, v) は法線ベクトルを表す。

次に、各領域に含まれる法線ベクトルをクラスター単位で合成し、領域数 \times クラスター数の方向ベクトルを生成する。そして、この結果に二回目のクラスタリングと合成を行い、結果として総クラスター数に相当するベクトルを生成する。この時のベクトルの数は設置された光源の数と一致することが望まれる。

5. 結果と考察

図4の対象画像に対する実験結果は図5のようになった。対象画像には二個の光源がシーンの左右にそれぞれ存在しているのに対して、結果は全く異なる方向のベクトルを示した。

この原因として、各物体の領域に存在する他の物体の影が、クラスタ分析ではその物体で有効なクラスタとして分類され、その後の合成計算に用いられたため、全体の推定結果に悪影響を与えた可能性が高い。また、地面領域内のシェーディング情報が統合されず、規則性に欠けていることや、クラスタ分析に使用されるデータの構造が不適切であるなどの要因も考えられる。

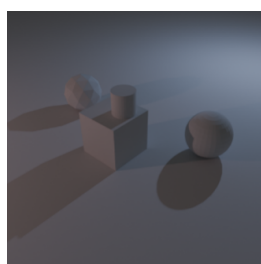


図4 対象画像

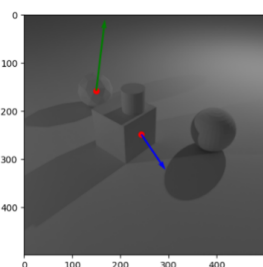


図5 推定結果

6. 今後の課題

前節で述べた実験結果を踏まえ、今後はまずセグメンテーションの手法を改善し、シェーディング情報に基づいた領域分割の手法による実験を検討する。具体的には、輝度値の変化の幅を基準に領域分割をし、それぞれの領域に対して一つだけのメインクラスターを算出する。そして、各領域のメインクラスターを合成したベクトルに対して再度クラスタリングと合成を行い、その出力を最終結果とする方法が考えられる。

その他に、クラスタ分析に用いられるベクトルの構造を5次元から3次元に下げることや、先行研究のように局所的な複雑度と平均輝度値によって適切な領域を選び出すこと、そして、事前に想定した光源の数を指定して非階層的なクラスタ分析手法を用いることも、有効な改善策だと考えられる。

7. おわりに

本研究では、複数光源が存在する画像に対して三次元的に光源方向を推定する手法を提案し、その提案手法の有効性を検証する初期的な実験を行った。今後の予定として、輝度値の変化幅をもとにするセグメンテーション手法を導入し、シェーディング情報や輝度値の変化幅を基準に領域分割をしたうえで、様々なクラスタ分析手法を適用する実験を行う。

参考文献

- [1] Alexander Kirillov, Kaiming He, Ross Girshick, Carsten Rother, Piotr Dollár: "Panoptic Segmentation", CVPR(2019)
- [2] Jizheng Yi, Xia Mao, Lijiang Chen, Yuli Xue, Angelo Compare: "Illuminant direction estimation for a single image based on local region complexity analysis and average gray value", Applied Optics 53(2):226-36(2014)
- [3] M.Chacon, L.E.Aguilar, and A.Delgado: "Fuzzy adaptive edge definition based on the complexity of the image", in Proceedings of the 10th IEEE International Conference on Fuzzy Systems, pp.675-678(2001)
- [4] Kaiming He, Georgia Gkioxari, Piotr Dollár, Ross Girshick: "Mask R-CNN", arXiv:1703.06870 [cs.CV]
- [5] Jonathan Long, Evan Shelhamer, Trevor Darrell: "Fully Convolutional Networks for Semantic Segmentation", CVPR(2015)
- [6] Martin Ester, Hans-Peter Kriegel, Jörg Sander, Xiaowei Xu: "A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise", KDD, (1996)