# Abstruct

我々は，灰色のピクセルを見つけるための新しいグレーネスインデックスを提案し，照明推定におけるその有効性と効率性を実証する．このグレーネス・インデックス（GI）は，Dichromatic Reflection Modelを用いて導出され，学習不要である．GIは、色に偏りのある画像において、1つまたは複数の照明源を推定することができます。GIは、標準的な単一照明および複数照明の推定ベンチマークにおいて、最先端の統計的手法や、最近の多くの深層心理学的手法よりも優れています。GIはシンプルかつ高速で、数十行のコードで書かれており、最適化されていないMatlabのコードでも1080pの画像を約0.4秒で処理することができます。

# 1.Introduction

人間の目には、撮影条件や照明の変化に適応する能力があります。コンピュータビジョンで定評のある「色の恒常性」（CC）は、民生用デジタルカメラにも同様の能力を持たせようとするものです。完全な色の恒常性があれば、RGBの値が等しいかどうかを確認するだけで、グレーのピクセルを見つけることはまったく問題ありません。しかし、色に偏りのない画像の場合、グレーの画素、つまり無彩色の表面を観察している画素を検出することは、難しく、非課徴的な問題です。一方、画像内のグレーのピクセルが「完全」であれば、色の恒常性が満たされていることになります。そこで、ここからは、グレーの画素を見つけることと、色の恒常性を保つことを同等の問題として扱います（図1参照）。色の恒常性の問題は，計算写真，本質的な画像分解，セマンティックセグメンテーション，シーンレンダリング，オブジェクトトラッキングなど，多くのコンピュータビジョンや画像処理のアプリケーションで発生する．[18].

図1：画像のグレーと非グレーのピクセル（左）。グレーネス・インデックス（GI）マップ（中央、青は高いグレーネス値を示す）。GIマップから推定されたグローバルな照明色（右上）と空間的に変化する照明色（右）。

何十年もの間，色の恒常性に対する古典的なアプローチである学習不要法は，照明の色がシーン全体で一定であり，したがって大域的な処理によって推定できるという仮定に基づいていた [6, 2, 38, 17, 19, 41, 12]．このアプローチは，照明特性が画像ごとに推定されるため，撮影デバイスに依存しないという利点がある．最近では，畳み込みニューラルネットワーク（CNN）をはじめとする最新の学習ベースの手法が，特定のデータセットで検証された場合，一貫して統計的手法を上回っている[9, 25, 22, 24, 29]．我々は，学習ベースの手法は，照明やシーンコンテンツの統計的分布がトレーニング画像とテスト画像で類似しているという仮定に依存していると主張する．言い換えれば，学習ベースの手法は，与えられた画像の撮像および照明条件が過去の学習例から推測できることを前提としており，したがって，学習データに大きく依存することになる[21]。

本稿では、学習不要のアプローチに注目します。具体的な例として、ユーザーがWeb上でリニアRGB（ガンマ補正）画像を取得し、その色を補正したい場合を考えます。このシナリオでは、使用されているCC法がそのカメラからの画像を見たことがない場合、撮像装置や撮影されたシーンに関する強い仮定なしに、照明推定と色補正を行う必要があります。このような環境では、学習不要の手法は、学習ベースの手法と比較して、より有望でロバストな結果を示すことを実験的に示しました。そのため、撮影した画像のカメラや撮像プロセスなどのパラメータに依存しない学習不要な手法が求められています。

ほとんどのカメラセンサでは，標準的なニュートラル照明下のリニアRGB画像ではグレーピクセルがグレーにレンダリングされるため，グレーネスは入射照明の色を推定するための潜在的な尺度となる．我々は，Shaferの二色性反射モデル（DRM）[33]を採用して，新しいグレーネス・インデックス（GI）を開発し，グレーネスに応じてすべての画像ピクセルをランク付けすることができる．魅力的な点は (GIは、単純かつ高速に計算できること、②物理的な意味が明確であること、③鏡面反射のハイライトをある程度扱うことができること（定性的な比較から）、④ピクセルレベルの照明を推定できること、⑤異なるカメラでも一貫した予測が可能であること。単一照明および複数照明の色の恒常性データセットを用いた包括的な結果から、GIは最先端の学習不要の手法を上回り、クロスデータセットの設定では最先端を達成していることが示された。

# 2. Related Work

リニアデジタルカメラのセンサーで撮影した画像を、黒レベルを補正し、彩度をかけずに撮影したものを考えてみましょう。二色性反射モデルでは，1つのグローバル照明源の下での(x, y)におけるピクセル値は，次のようにモデル化される[33]。

文字が書かれている

低い精度で自動的に生成された説明

ここで，I (x,y) iは(x, y)における画素値，L(λ)はグローバルな光のスペクトル分布，Fi(λ)はセンサー感度，3色カメラの場合はi = {R, G, B}，λは波長である。彩色項のRb(λ)とRs(λ)は、それぞれ身体反射と表面反射を表し、無彩色項のγbとγsは、上記2種類の反射の強さを表しています。

また，狭いスペクトル応答Fi(λ)を仮定すると，式1はさらに簡略化されて[3]のようになります。

ここで、◦はハダマード積を表し、◦は

テキスト, 手紙

自動的に生成された説明

　ここで，{R,G,B}の添え字はFiと交差するスペクトルの対応する部分を表しています。式2は，カメラが捉えたグローバルライトLの下で，体Wと表面反射Vを示すシーンの位置に対応する画像Iの画素値の成り立ちを示している。

　CCの目的は，Iが与えられたときに，Wを復元するためにLを推定することである。この問題を解決するために使用される戦略に基づいて，色の恒常性の手法を，学習ベースの手法と学習なしの手法の2つのカテゴリーに分けている。

## Learning-based Methods

[9,25,22,24,29,31,32]では，豊富な学習データから，撮影された画像Iと求められる照明Lを関連付けるモデルを構築することを目指している．最先端のアプローチの中でも特に性能が高いのは，CCC法[3]であり，2次元のログクロマ空間で畳み込みフィルタを識別的に学習している．このフレームワークは，彩度トーラス上の高速フーリエ変換を用いて高速化されています[4]．Chakrabartiら[8]は、条件付きの彩度分布を学習することで、照明予測に正規化された輝度を活用しています。DS-Net [35]とFC4 Net [28]は2つの深層学習手法であり，前者は2分岐CNNアーキテクチャを用いて複数の照明推定値から推定値を選択し，後者はセグメンテーションのようなフレームワークを用いてパッチの局所的な推定の曖昧さに対処する．学習ベースの手法は、事前に記録された「グランドトゥルース」の照明色をかなり正確に予測することに成功しているが、トレーニング画像とテスト画像の両方に同じカメラやシーンがあるかどうかに大きく依存している（項3および項4.2参照）。Corrected-Moment法[14]は、データセットごとに補正行列を学習する必要があるため、学習ベースの手法と考えることもできます。

# 3. Grayness Index

　まず最初に、二色性反射モデル（DRM）の文脈で、（Lambertianモデルから導かれた）以前のGray Pixel [41]をレビューします。

## 3.1. Gray Pixel in [41]

Yangら[41]は、灰色のピクセルは、一連の制約によって求められると主張している。しかし、彼らの定式化は、明らかにカラーピクセルであるグレーピクセルを識別することが多い。この現象は注目されていたが、適切に分析されていなかった。ここでは、DRMを用いてGPを解析し、オリジナルの定式化の潜在的な失敗例を指摘する。ナローバンドセンサーを仮定すると、式1は次のように単純化される。

概略図

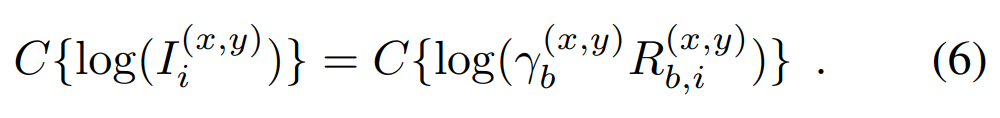
低い精度で自動的に生成された説明

次に，Yangら[41]に倣って，log(-)と局所的なコントラスト演算子C{-}(5\*5のLoGフィルタ)を双方に適用する．(Laplacian of Gaussian, 詳細は項4を参照)を両辺に適用し、次のように求めます。

テキスト, 手紙

自動的に生成された説明

γs＝0（表面反射がないことを意味する）とすると、次のようになります。



γs 6=0の場合、式（5）のγbとγs Rs,i Rb,iの相互作用により、これらの色付きの画素は灰色の画素として誤って認識されてしまう。GPでは、γs=0の場合、均一なグレーの表面に不均一な強度の光を投射すると、各チャンネルに同じ量の「コントラスト」が発生するということが重要です。光の強さの変化は、表面と照明の間の幾何学的な関係（シェーディング）や、異なる表面間の幾何学的な関係（オクルージョン）に起因します。この問題を解決するために、我々は二色性反射モデルを採用し、より複雑な環境下でグレーのピクセルを識別するための別の方法を模索しています。

## 3.2. Grayness Index using Dichromatic Reflection Model

　すべての演算は（x, y）を中心とした局所的な近傍に適用されるため，簡単のために，以降では上付き文字（x, y）を削除する．まず、赤チャンネルと輝度の残差を対数空間で計算し、局所的なコントラスト演算子C{-}をEq.5に次のように適用します。

テキスト, 手紙

自動的に生成された説明

図2：グレーのピクセルを見つける。(a)入力画像。(b) 計算されたグレーネスインデックスGI．濃い青はグレーネスの度合いが高いことを示す．(c) (a)の対応するピクセルの色(緑がかった色)を用いてレンダリングされた，N%の最も灰色のピクセル．(d) 推定された照明色．(e) グランドトゥルースカラー。(f) (d)で補正した画像。

ここで、｜I｜は輝度の大きさ（IR +IG +IB）を表す。

　この場合，中性界面反射（NIR）の仮定により，灰色の画素については，Rj,R = Rj,G = Rj,B = R¯ j with j∈{s, b} [30]が成立することになる。この場合、式7は次のように簡略化される。

文字が書かれている

中程度の精度で自動的に生成された説明

　局所的な小さな領域では，鋳造物の照度とセンサの応答が一定であると仮定することができ[41]，C{log(FRLR)}=0，C{log((FRLR + FGLG + FBLB)}=0となり，以下のようになる．

テキスト, ホワイトボード

自動的に生成された説明

　式(9)はグレーピクセルの必要条件ではありますが、十分条件ではありません。グレーの画素を検出するためのより限定的な条件は、式（9）をさらに1つのカラーチャンネルに拡張することで与えられます（すべてのチャンネルを冗長に使用すると、センサーではRとBのスペクトル応答がほとんど重なりません）。

テキスト

自動的に生成された説明

式(7)から、I(x, y)に対するGIを次のように定義します。

文字が書かれている

中程度の精度で自動的に生成された説明

ここで||・||はℓ2ノルム(俗に言うユークリッド距離)を意味します。GIが小さければ小さいほど、対応するピクセルがグレーである可能性が高い。

さらに、「小さな」GI値が、平坦なカラーパッチ（空間的な手がかりがない）ではなく、光の強さが変化するグレーのピクセルから得られることを保証するために、局所的なコントラストに制限を加え、次のように記述しました。



ここで、ǫは小さなコントラストの閾値です。

GIを算出するプロセスは、2つのステップで構成されています。

1. 式（11）を用いて予備的なGIマップを計算する

2. 式12を用いて、空間的な手がかりを持たないGIのピクセルを破棄する。主にカメラノイズに起因する孤立したグレーピクセルの影響を弱めるために、GIマップは7×7のウィンドウで平均化されます。

図2は、GIの計算とその予測照明のフローチャートを示しています。提案されたGIは、2つの重要な点でGPと異なります。

第一に，提案GIは，より完全な画像形成モデルに基づいてグレーピクセルを検出する新しいメカニズムを利用しており，これにより，異なるフォーミュレーションを実現している．第二に、提案するGIは、明るい画素と暗い画素を、その輝度に応じて選択的に強調することなく動作する。言い換えれば、提案するGIは、暗いピクセルの影響を弱めることはありません。

3.3. GI Application in Color Constancy

カラーコンスタンシーは、グレイピクセルをそのまま応用したものです。ここでは、グレイピクセルから照明色を計算する2つのパイプライン、シングルイルミネーションパイプラインとマルチイルミネーションパイプラインについて説明します。

1つのシーンに1つのグローバルイルミネーションしかない場合、パイプラインは簡単です。図2に示すように、画像の全ピクセルをGI値でランク付けした後、上位N％のピクセルの平均値としてグローバルイルミネーションを算出する。複数の光源で撮影されたシーンでは，ピクセル単位の照明マップが必要となります．41]と同様に，まずGIマップを計算し，次に上位N%のピクセルをK-meansクラスタリングして，あらかじめ設定された数のM個のクラスタに分類します．ここで，クラスタベースで平均化を行い，クラスタmの照明ベクトルLmを得る．最終的な空間照明マップは以下の方法で計算される．

式13

ここで、ωmは、画素I(x, y)とクラスタmとの間の接続を制御するもので、次のように書かれます。

式14

ここでDmは、ピクセルからクラスタmのセントロイドまでのユークリッド距離です。式14は、近くのピクセルが同じような照明を共有することを促します。

4. Evaluation

GIは，2つの色の恒常性の設定で評価しました。(1）シーン全体の照明が，赤，緑，青の各チャンネルの単一のクロマベクトルで記述される「単一照明推定」と，（2）シーンごとに2つ以上の有効な照明が存在する「複数照明推定」である．さらに，学習ベースの手法では困難なクロスデータセットでの実験も行いました。

Datasets

* Gehler-Shi Dataset [34, 22]：単一照明、568枚のハイダイナミックリニア画像、2台のカメラ 1 .
* The NUS 8-Camera Dataset [12]：単一照明，1,736枚のハイダイナミックリニア画像，8台のカメラ（カメラリストは表2参照
* - MIMO Dataset [5]: マルチイルミネーション、78枚のリニア画像、58枚の実験室画像、20枚のハードワイルド画像

Single-illumination Experiment Settings

* 式11の局所的なコントラスト演算子は，サイズが5ピクセルのガウスフィルタのラプラシアンである。
* 色の推定に使用される最良のグレーピクセルの割合は、N = 0.1%に設定されています。
* コントラストの閾値は、ǫ = 1e-4に設定されています。

これらのパラメータは，予備的なグリッド探索（セクション4.3参照）に基づいて選択され，両データセットを用いたすべての実験において固定されたままであった．

Multi-illumination Experiment Settings

* 局所的なコントラスト演算子とコントラスト閾値は、単眼鏡実験と同じです。
* 選択されたピクセルの割合は、より多くのイルミナントが含まれるため、N=10.0%に設定されています。
* テストしたクラスターMの数は2、4、6でした。