情報処理学会研究報告
IPSJ SIG Technical Report

Vol.2013-CVIM-187 No.29
2013/5/30

非負値行列因子分解を用いた画像の鏡面反射分離

明石 康宏[1]　岡谷 貴之[1]

概要：本研究は，2 色性反射モデルに基づき，単一画像から拡散反射成分と鏡面反射成分とを分離する新たな手法を提案する．このような分離を行う方法に，事前に画像中の物体表面について色空間上でいくつかの単色の領域に分割した後，各領域について 2 色性反射モデルにしたがって各成分を同定するものがある．提案手法は非負値行列因子分解を用いることで，このような領域分割を明示的に行うことなく，単色領域の同定と各反射成分の分離を同時に行う．実画像を用いて既存手法との比較実験を行ったところ，提案手法はより良好な結果を与えることが示された．

1. はじめに

本研究では，一枚の画像中の鏡面反射を特定し，その成分を拡散反射の成分と分離する問題を扱う．鏡面反射成分を分離することにはいくつかの用途がある．1 つは，陰影や色を用いた画像計測—例えば Shape from Shading[1] や Photometric Stereo[2] などの三次元形状復元—では，物体が拡散反射表面であることが一般に仮定され，したがって実画像中の鏡面反射成分を画像から除去できることは，高精度な計測の前提となる．また，画像から物体のマテリアルを認識する場合，画像中のハイライトはその特徴になり得るため，これを取り出して認識に利用することも考えられる．

このような理由により，単一画像中の鏡面反射と拡散反射の成分を，正確かつ頑健に分離する方法の研究が盛んに行われてきた [3], [4], [5], [6], [7], [8], [9], [10]．その多くは 2 色性反射モデル，すなわち物体表面での反射光（すなわち画像の各色の輝度値）が，鏡面反射成分と拡散反射成分の線形和になるというモデルに基づく [3]．2 色性反射モデルによれば，$\mathbf{i}$ をある画素の RGB 値を格納した 3 成分のベクトルとすると，照明色 $\mathbf{i}_s$ および物体色（=表面での拡散反射に起因する物体表面の色）$\mathbf{i}_d$ を用いて

$$\mathbf{i} = \alpha \mathbf{i}_s + \beta \mathbf{i}_d. \tag{1}$$

のように書ける．

照明色 $\mathbf{i}_s$ および物体色 $\mathbf{i}_d$ が複数の画素において共通だとすれば，上の式はこれらについての条件式を与える．これまでの研究の多くはこの原理に基づくもので，より正確には次を仮定する．

- 物体表面は，それぞれ異なる物体色を持つ複数の領域からなる．
- 照明色 $\mathbf{i}_s$ が既知だとし，物体色 $\mathbf{i}_d$ が未知である．上の領域ごとに $\mathbf{i}_d$ は異なり，各画素がどの領域に属するかは事前には不明である．α および β は画素ごとに異なる未知数である．

この条件の下，反射成分の分離を行う方法が研究されてきた．本研究でも，これと同じ条件を考える．

この条件の下で単一画像中の反射成分を分離する方法は，大きく 2 つに分けられる．一つは，色空間上に全画素を投影し，その分布に対しクラスタリングを行い，画像中の物体色とその領域（画素集合）を定めた後，式 (1) に基づいて未知数を定める方法である．色空間として RGB 空間を対象とする Klinker らの手法 [6] や，Hue-Saturation-Lightness 空間を用いる Bajscy らの方法 [7] がある．これらの方法は，物体表面の彩色が単純な場合には有効だが，そうでない場合には色空間で物体色を推定する部分に困難さが生じる．

そこで，このように色空間での明示的な物体色の推定を行わず，画像上の隣接画素間での物体色の関係を元に，反射成分を反復計算により分離する方法が提案された．Tan-Ikeuchi の手法 [8] や，これを単純にし高速化を図った Shen らの手法 [9] である．これらは，鏡面反射成分を含まない（がそのままでは拡散反射成分を正しく表していない）中間的な表現（Specular Free Image と呼ばれる）を経由し，成分分離を行う．この方法では，典型的にはハイライト領域の外側から内側へ向けて物体色の情報が伝搬するが，物体色の境界でこれに失敗することがよくあり，万能とはいえなかった．このような欠点を克服するため，Tan-Lin-Quan は，ハイライトが生じる領域の周囲のテク

[1] 東北大学
Tohoku Uniersity

スチャ情報を用いて，拡散反射成分を復元する手法を提案した [10] が，画像中のハイライト部分が特定できている必要があった．

本研究では，従来同様 2 色性反射モデルに基づき，画像データについて非負値行列因子分解（NMF : Non-negative Matrix Factorization）を行うことで，1 枚の画像から鏡面反射分離を行う手法を提案する．提案手法は，どちらかと言えば色空間のクラスタリングを行って単色領域を同定する方法に近いが，これを陽に行うのではなく，単色領域の同定と反射成分の分離を一度に行う点で異なる．既存手法と異なり，常に画像全体で最適化計算を行うため，高精度かつ安定に反射成分分離を行うことが期待できる．

2. 非負値行列因子分解

2.1 概要

NMF とは，画像や音声データなどの非負値で構成されるデータを加法的な構成成分に分解することを目的とした多変量解析手法である．具体的には，非負値のデータを格納した $M \times N$ 行列 $\mathbf{V}$ を，次のように同じく非負値の $M \times R$ 行列 $\mathbf{W}$ と $R \times N$ 行列 $\mathbf{H}$ の積に分解する：

$$\mathbf{V} \simeq \mathbf{WH} \tag{2}$$

このとき，右辺の $\mathbf{W}$ および $\mathbf{H}$ も成分がすべて正であることに注意する．データの行列 $\mathbf{V}$ の j 番目の列ベクトルを $\mathbf{v}_j$ と書くと，式 (2) は $\mathbf{W} = [\mathbf{w}_1, \dots, \mathbf{w}_R]$ を用いて

$$\mathbf{v}_j \simeq \sum_{k=1}^{R} \mathbf{w}_k H_{k,j} \tag{3}$$

のように書ける．$\mathbf{W}$ は，$\mathbf{V}$ の列ベクトルをその線形和で最もよく表現できるようなベクトルを R 個，列ベクトルとして格納した行列で，$\mathbf{H}$ はその線形和の重みであると解釈できる．

このような分解は，データ行列 $\mathbf{V}$ とその再現 $\mathbf{WH}$ の乖離度 $D(\mathbf{W}, \mathbf{H})$ を目的関数として，それを最小化することで得られる．この乖離度 $D(\mathbf{W}, \mathbf{H})$ には，L2 ノルム

$$D(\mathbf{W}, \mathbf{H}) = \|\mathbf{V} - \mathbf{WH}\|_2^2 \tag{4}$$

を用いるのが最も一般的で，場合によって一般化 KL ダイバージェンス [11] や Itakura-Saito ダイバージェンス [12] を用いる．

NMF は，Lee らが効率的な反復アルゴリズムを考案した [11] ことを契機に，様々な問題に応用されるとともに，問題に応じて目的関数を修正する拡張が行われてきた [12], [13], [14], [15], [16], [17], [18].

2.2 スパース NMF

中でも重要な拡張に，スパース正則化を取り込んだスパース NMF がある [13]．これは，$\mathbf{H}$ の非零成分の数が少なくなるような分解を得ることを目的に，次の目的関数を最小化する．

$$F(\mathbf{W}, \mathbf{H}) = \frac{1}{2}\|\mathbf{V} - \mathbf{WH}\|_2^2 + \lambda \sum_{i,j} H_{i,j} \tag{5}$$

式 (5) の右辺の第 2 項は，スパースコーディング [19] などと同様に，L0 ノルムの代替として L1 ノルムを採用する緩和に基づく．この結果，スパースコーディングと同じように，各データがなるべく少ない数の基底（$\mathbf{W}$ の列ベクトル）の線形和で表現されるようになる．ただし NMF では，非負成分を持つ行列への分解を考えている点で，スパースコーディングとは異なる．

この目的関数を，$\mathbf{W}$ と $\mathbf{H}$ が非負値をとるという制約の下で最小化する数値計算の方法がにある．これは，$\mathbf{W}$ と $\mathbf{H}$ をそれぞれランダムに初期化した状態から開始し，次に示す更新式を収束するまで反復する方法である．

$$\mathbf{H} \leftarrow \mathbf{H} \odot \frac{\bar{\mathbf{W}}^T V}{\bar{\mathbf{W}}^T \bar{\mathbf{W}} \mathbf{H} + \lambda} \tag{6a}$$

$$\mathbf{W} \leftarrow \mathbf{W} \odot \frac{\mathbf{VH}^T + \bar{\mathbf{W}} \odot \mathbf{A} \bar{\mathbf{W}} \mathbf{HH}^T}{\bar{\mathbf{W}} \mathbf{HH}^T + \bar{\mathbf{W}} \odot \mathbf{AVH}^T} \tag{6b}$$

なお，式中の $\bar{\mathbf{W}}$ は列ごとに正規化された $\mathbf{W}$ を表し，$\mathbf{A}$ は要素が全て 1 である $\mathbf{M} \times \mathbf{M}$ 行列，$\odot$ の記号は行列の要素同士の乗算を表し，除算も要素ごとに行う．この反復は，局所的な最小解に必ず有限の反復で到達することが示されている [13].

3. NMF を用いた鏡面反射分離

3.1 定式化

画像中の物体色の数を $R-1$ とし，それぞれ $\mathbf{i}_k (k = 1, ..., R-1)$ とする．画素 p における色 $\mathbf{i}_p$ は，

$$\mathbf{i}_p = \alpha_p \mathbf{i}_s + \sum_{k=1}^{R-1} \beta_{k,p} \mathbf{i}_k \tag{7}$$

のように書ける．ただし $\beta_{k,p}$ は $k = 1, \dots, R-1$ のうち唯一の k についてのみ非零（それ以外は零）となるとする．各画素の色は，特定の物体色一つと照明色の和で与えられるからである．これを式で書くと

$$\sum_{k=1}^{R-1} \|\beta_k\|_0 = 1 \tag{8}$$

となる．また，照明色 $\mathbf{i}_s$ は正規化されている ($\|\mathbf{i}_s\|_2 = 1$) とする．なお，$\mathbf{i}_p$，$\mathbf{i}_s$，$\mathbf{i}_k$ の各成分および，α_p と $\beta_{k,p} (k = 1, \cdots, R-1)$ は全て非負であることに注意する．

以上は次のように行列表記できる．入力画像の RGB の 3 成分をそれぞれ 1 つのベクトルで表し，それら 3 つのベクトルを行ベクトルとするような $\mathbf{V}$ を定義する．つまり，画像サイズを $m \times n$ とすると，$\mathbf{V}$ は $3 \times N (N = mn)$ となる．$\mathbf{W}$ を $3 \times R$ の行列

$$\mathbf{W} = [\mathbf{i}_s, \mathbf{i}_1, ..., \mathbf{i}_{R-1}] \tag{9}$$

とし，各画素の α_p および $\beta_{k,p}$ を並べた $R \times mn$ の行列

$$\mathbf{H} = \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_N \\ \beta_{1,1} & \beta_{1,2} & \cdots & \beta_{1,N} \\ \vdots & \vdots & \vdots & \vdots \\ \beta_{R-1,1} & \beta_{R-1,2} & \cdots & \beta_{R-1,N} \end{bmatrix} \tag{10}$$

を定義すると，式 (7) および (8) にしたがって $\mathbf{i}_p$ から $\mathbf{i}_k$ および α_p,$\beta_{k,p}$ を得る問題は，次のような制約付きの非負値行列分解と再解釈できる：

$$\mathbf{V} = \mathbf{W}\mathbf{H} \tag{11a}$$

$$\text{s.t.} \quad \sum_{k=1}^{R-1} \|H_{i,j}\|_0 = 1 \tag{11b}$$

ここで $H_{i,j}$ は行列 $\mathbf{H}$ の (i,j) 成分である．$\mathbf{W}$ および $\mathbf{H}$ の成分は全て非負でなければならないことに注意する．したがって，式 (11a) の分解を式 (11b) の条件下で行う．ただし，$\mathbf{W}$ は全てが未知ではなく，$\mathbf{W}$ を照明色と物体色に分け，

$$\mathbf{W} = [\mathbf{i}_s, \mathbf{W}_d] \tag{12}$$

と書いたとき，$\mathbf{i}_s$ は既知，$\mathbf{W}_d$ のみが未知である．

3.2 問題の緩和

式 (11) は，そのままでは解くことが困難なので，次の目的関数を最小化するような非負行列 $\mathbf{W_d}$ と $\mathbf{H}$ を求めることを考える．

$$F(\mathbf{W_d}, \mathbf{H}) = \frac{1}{2}\|\mathbf{V} - \mathbf{W}\mathbf{H}\|_2^2 + \lambda_s \sum_{j=1}^{N} \|H_{1,j}\|_0 + \lambda_d \sum_{j=1}^{N}\sum_{i=2}^{R} \|H_{i,j}\|_0 \tag{13}$$

第 1 項はデータと分解結果の乖離度は小さくあるべきこと，第 2 項は画像中の鏡面反射成分を与える画素数がなるべく少なくあること，第 3 項は各画素をなるべく少数の物体色で表現することをそれぞれ要請する．第 2，3 項が与えるスパース性によって，各画素をなるべく少ない (理想的にはただ一つの) 物体色と，必要な場合にのみこれに照明色を加えて表現できるように，物体色 ($\mathbf{W_d}$) が選ばれ，この画素の色を与える線形和の重み ($\mathbf{H}$) が決まる．第 2，3 項は本来意味の異なるものだが，後に示す実験においては両者の重みを同一とし，

$$\lambda = \lambda_s = \lambda_d \tag{14}$$

としている．λ は小さいとデータとの乖離度は小さくなるが，少数の物体色と照明色の線形和に必ずしもならず，逆に大きいと乖離度が大きくなってしまう．適度に λ を選択することで目的が達成できると期待される．

このように書き換えても，式 (13) をこのまま解くことは難しいため，L0 ノルムをその代替の L1 ノルムで置き換えた

$$F(\mathbf{W_d}, \mathbf{H}) = \frac{1}{2}\|\mathbf{V} - \mathbf{W}\mathbf{H}\|_2^2 + \lambda \sum_{j=1}^{N}\sum_{i=1}^{R} \|H_{i,j}\|_1 \tag{15}$$

を目的関数とし，最小化する．すると，式 (15) は，式 (5) と同じ目的関数となり，スパース NMF の最適化アルゴリズムをそのまま用いることができる．ただし，今 $\mathbf{W}$ の照明色については既知としてあるので，この成分については更新を行わず，$\mathbf{W_d}$ についてのみ更新を行うようにアルゴリズムを修正する．具体的にいえば，式 (6b) を次のように変更する．

$$\mathbf{W_d} \leftarrow \mathbf{W_d} \odot \frac{\mathbf{V}'\mathbf{H_d}^T + \bar{\mathbf{W}}_\mathbf{d} \odot \mathbf{A}\bar{\mathbf{W}}_\mathbf{d}\mathbf{H_d}\mathbf{H_d}^T}{\bar{\mathbf{W}}_\mathbf{d}\mathbf{H_d}\mathbf{H_d}^T + \bar{\mathbf{W}}_\mathbf{d} \odot \mathbf{A}\mathbf{V}'\mathbf{H_d}^T} \tag{16}$$

ここで

$$\mathbf{V}' = \mathbf{V} - \mathbf{i}_s\mathbf{h}_s \tag{17}$$

$$\mathbf{h}_s = \begin{bmatrix} \alpha_1 & \cdots & \alpha_N \end{bmatrix} \tag{18}$$

$$\mathbf{H_d} = \begin{bmatrix} \beta_{1,1} & \cdots & \beta_{1,N} \\ \vdots & \vdots & \vdots \\ \beta_{R-1,1} & \cdots & \beta_{R-1,N} \end{bmatrix} \tag{19}$$

であり，$\mathbf{A}$ は先ほどと同じく要素が全て 1 である $M \times M$ 行列である．これにより，$\mathbf{i}_s$ は初期値のまま更新されることは無い．

4. 実験

4.1 実験全体における NMF の設定

以上の方法を実行するには，$\mathbf{W}$ の列数 R を設定する必要がある．画像中に含まれる色（単色領域）の数を正確に言い当てられるのであれば，その値を用いるのが良い（その数が $R-1$ になる）．現実には，正確にそれを知るのは難しいと考えられるが，後の実験に示すように，それほど正確でなくても十分良い結果が得られる．

照明色のベクトル $\mathbf{i}_s$ は $\frac{1}{\sqrt{3}}[1,1,1]^T$ とし，式 (6a) と 式 (16) に基づいて更新を行い，$\mathbf{W_d}$ と $\mathbf{H}$ を求める．更新は収束するまで行う．収束判定は，ある時刻 t における目的関数の値を F_t とすれば，ひとつ前の時刻における F_{t-1} を用いて

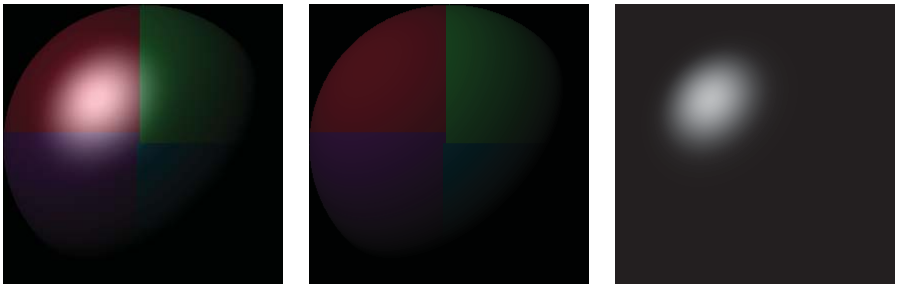
$$|F_t - F_{t-1}| < \epsilon |F_t| \tag{20}$$

が満たされたとき収束したとみなす．ϵ は $\exp(-15)$ とした．そして最適化された $\mathbf{W_d}$ と $\mathbf{H}$ について，鏡面反射成分を

$$\mathbf{I}_s = \mathbf{i}_s\mathbf{h}_s \tag{21}$$

また拡散反射成分を

$$\mathbf{I}_d = \mathbf{W}_d\mathbf{H}_d \tag{22}$$



(a) (b) (c)

図 1 実験に用いた CG 画像と，分離の正解画像．(a) が元画像で，(b) が拡散反射成分のみの画像，(c) が鏡面反射成分のみの画像である．

(a) (b) (c) (d)

図 2 各手法での分離結果． (a)，(b) は提案手法による分離結果で，(c)，(d) は Tan-Ikeuchi[8] による分離結果である．

表 1 RMSE 値の比較

比較成分	提案手法	Tan-Ikeuchi の手法
拡散反射成分	0.916	7.73
鏡面反射成分	0.891	7.74

と復元する．スパース正則化の λ は 3 とした．$\mathbf{H}$ と $\mathbf{W_d}$ の初期値は一様分布に基づく 1 から 255 までの乱数によって与え，$\mathbf{W_d}$ については大きさが 1 となるように正規化しておく．

こうして再構成された各画像について Tan-Ikeuchi の手法 [8] と比較を行った．Tan-Ikeuchi[8] については彼らが公開しているソースコードを用いた（`http://www.staff.science.uu.nl/~tan00109/code.html`）．

4.2 合成画像を用いた実験

まず，合成した画像を用いて実験を行った，合成画像および拡散反射と鏡面反射の成分を図 1 に示す．この画像に対する，提案手法と Tan-Ikeuchi[8] の結果を図 2 に示す．提案手法では $R = 5$，$\lambda = 3$ とした．分離した各反射成分について，平均二乗誤差（RMSE）を算出したものを表 1 に示す．

図 2 と表 1 から，Tan-Ikeuchi[8] はアーティファクトが生じるとともに，RMSE 値からの精度が低いことが分かる．一方提案手法は良好な分離結果を与え，数値的にも高い精度を達成できている．

(a) (b) (c)

図 3 Tan-Ikeuchi が公開している実画像データセット [8]．(a)fish，(b)head，(c)toys．

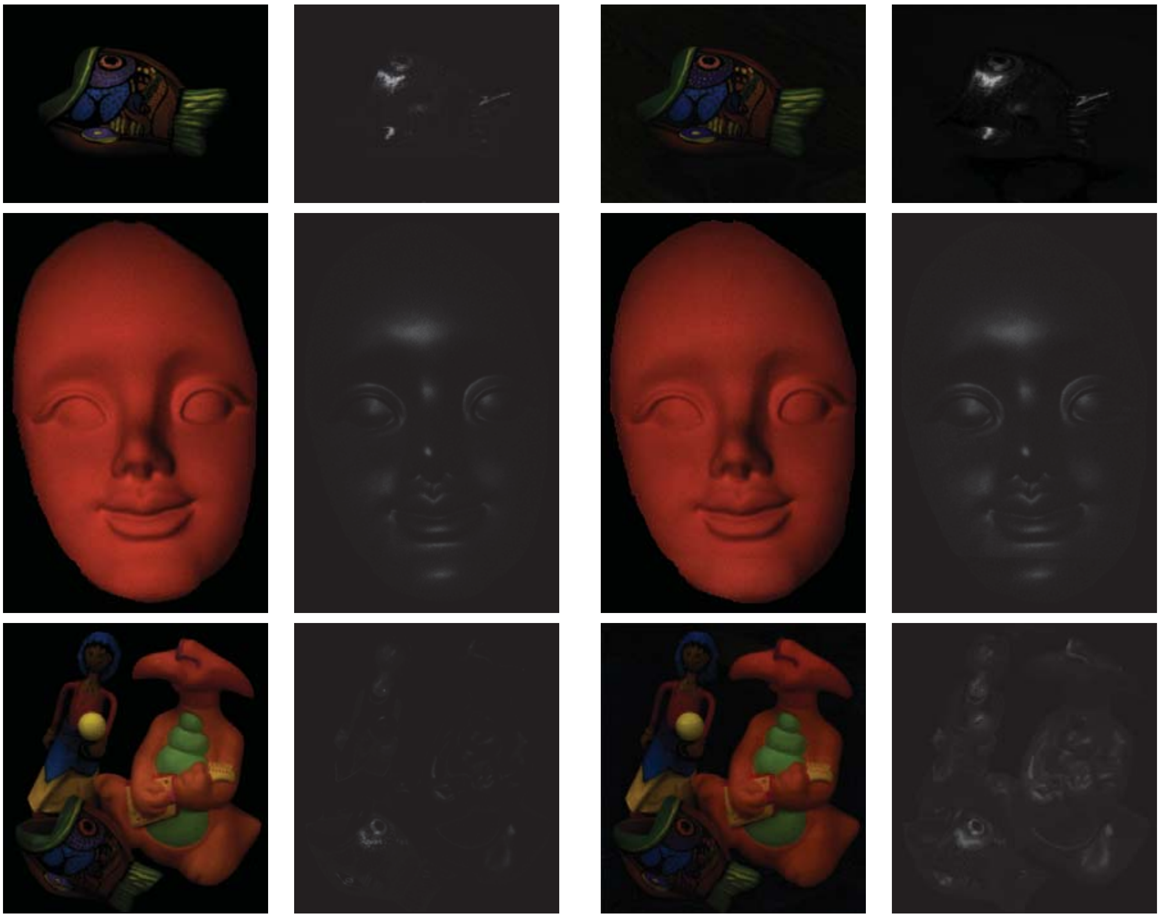
4.3 実画像を用いた実験

次に，図 3 に示す公開されている実画像データセット (`http://www.staff.science.uu.nl/~tan00109/code.html`) を用いて，実験を行った．各画像に対する両手法の結果を図 4 に示す．なお提案手法では $R = 5$，$\lambda = 3$ とした．

図 4 を見ると，提案手法と Tan-Ikeuchi の結果はかなり似通っているのが分かる．しかし詳細にみると，図 5 に示すように，Tan-Ikeuchi[8] では合成画像同様，アーティファクトが一部生じてしまっているが，提案手法の結果は良好である．また，提案手法で指定する必要のある R を，いくつか変えたときの結果を図 6 に示す．この図から分かるように，R は真の色数と一致する必要はなく，ある程度正しければ似たような結果が得られた．

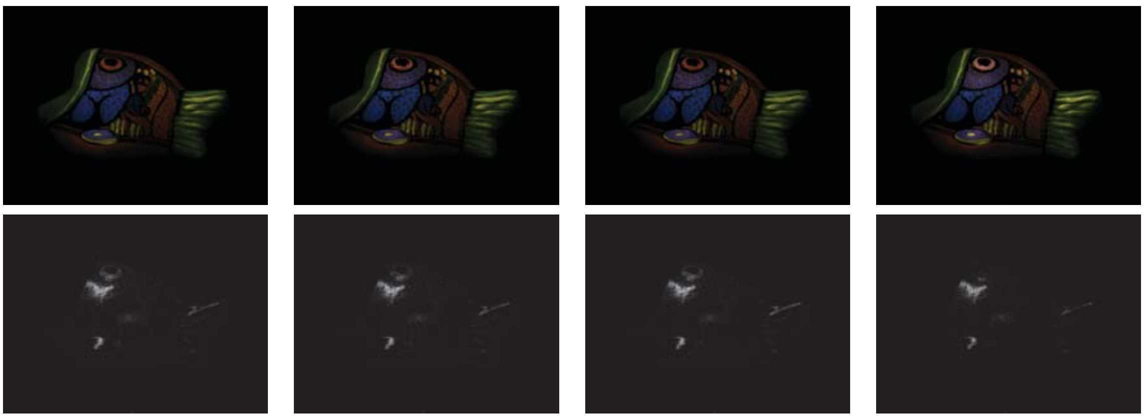
4.4 カメラの応答関数が未知である画像を用いた実験

多くの既存の鏡面反射分離手法では，カメラに到達する光の輝度と画像の輝度値の関係が線形であることを前提と



(a) (b) (c) (d)

図 4 各手法での分離結果．(a)，(b) は提案手法，(c)，(d) は Tan-Ikeuchi[8] による結果である．



(a) (b) (c) (d)

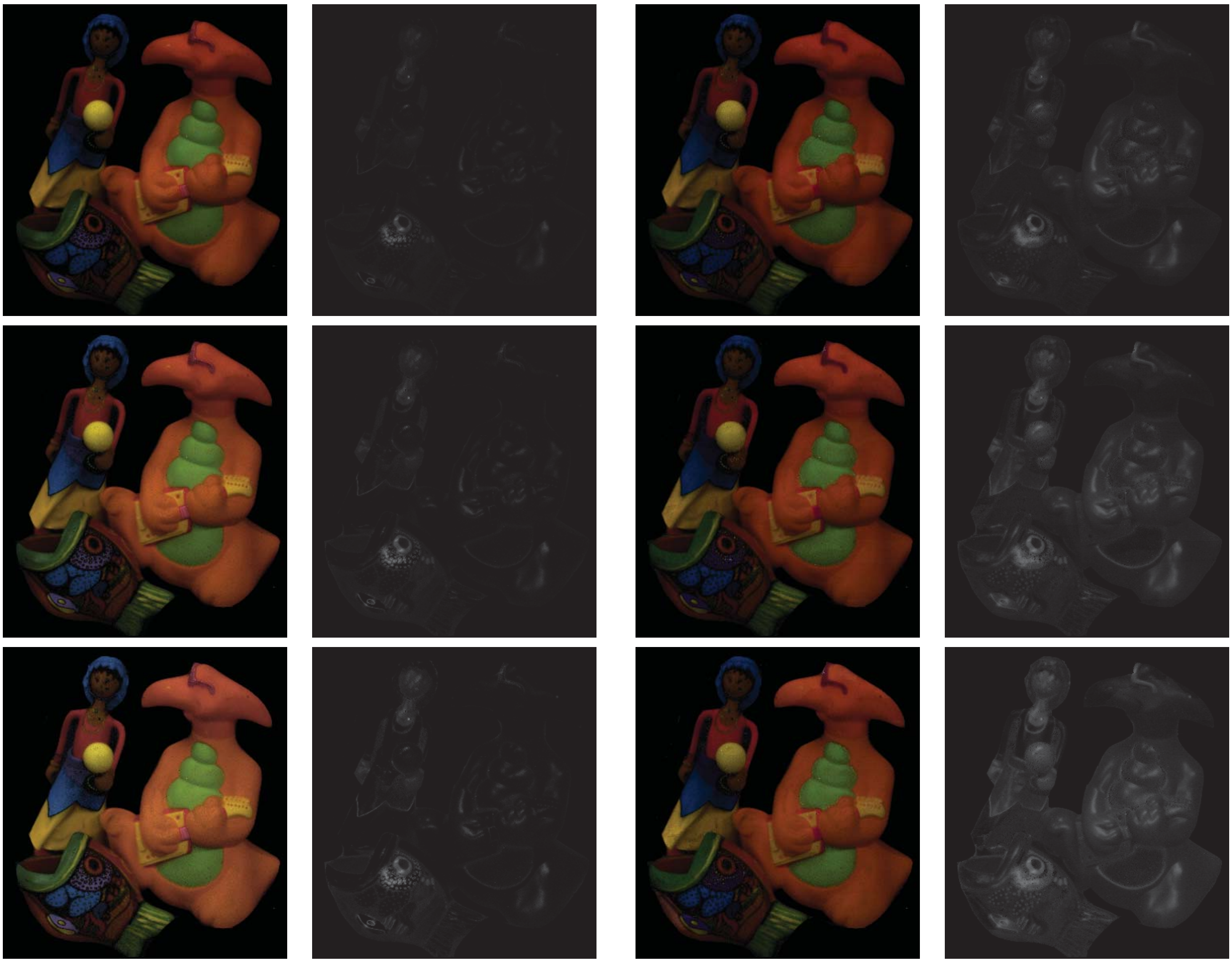
図 6 提案手法による各 R に対する分離結果．上段が拡散反射成分，下段が鏡面反射成分で，(a) は $R=4$，(b) は $R=5$，(c) は $R=6$，(d) は $R=7$ である．

する．入力画像を撮影したカメラの応答関数が既知であれば，このようなことが可能となる．しかしそのためには，カメラの応答関数を校正によって定める必要があるが，一般にそこにはに誤差が含まれる．そのような応答関数の校正誤差をシミュレートするため，画像の輝度値 x を故意に $x^{\frac{1}{\gamma}}$ と変換し，変換後の画像を入力として反射成分分離を行った．

このように輝度値を変換した画像を図 7 に示す．またそれらに対する提案手法および Tan-Ikeuchi[8] の結果を，図 8 に示す．図 8 より，Tan-Ikeuchi は γ の値が大きくなるほど分離された鏡面反射成分が大きくなり，逆に拡散成分が小さくなっている．一方，提案手法はその変化がずっと少なく，明らかにより良好な結果を与えている．以上より提案手法は，カメラの応答関数の校正誤差に対し，より頑

(a) (b) (c)

図 7 4.4 節の実験に用いた画像．(a) は $\gamma = 1.1$，(b) は $\gamma = 1.3$，(c) は $\gamma = 1.5$ である．



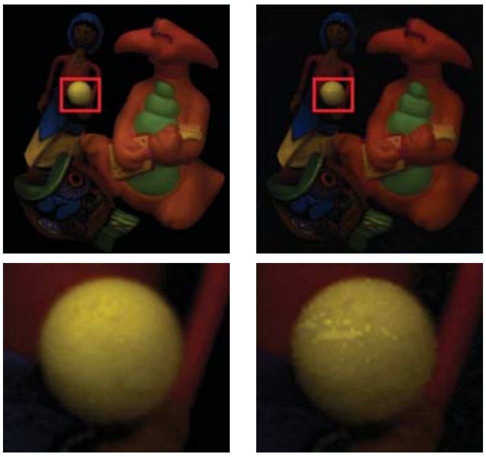
(a) (b) (c) (d)

図 8 輝度値を故意に変換したときの分離結果．上段から順に $\gamma = 1.1, 1.3, 1.5$ である．(a), (b) は提案手法による分離結果で，(c)，(d) は Tan-Ikeuchi[8] による分離結果である．

健であると言える．

5. まとめ

本研究では，NMF を用いて単一画像の反射成分の分離を行う方法を述べた．提案手法は，単色領域の同定と反射成分の分離を一度に，スパース NMF の最適化計算によって行うのが特徴である．これにより，色空間でのクラスタリングを最初に行なって単色領域を同定する既存手法や，隣接画素間での物体色の関係に基づいて反射成分を分離する既存手法と比べて，より安定かつ高精度な成分分離ができると期待できる．後者の既存手法である Tan-Ikeuchi との比較を行ったところ，合成画像，実画像を通じてより良好な分離が達成できていることがわかった．特に，カメラの応答関数の校正誤差により頑健であることが分かった．



(a) (b)

図 5 (a) は提案手法，(b) は Tan-Ikeuchi．上段は toys の拡散反射成分画像で，赤色の枠で囲われた部分を拡大した画像を下段に示す．

参考文献

[1] E. Prados and O. Faugeras: Shape from shading, *Handbook of mathematical models in computer vision*, pp. 1–17 , (2006).

[2] R. Woodham: Photometric method for determining surface orientation from multiple images, *Optical engineering*, Vol. 19, pp. 139–144, (1980).

[3] S. Shafer: Using color to separate reflection components, *Color Research & Application*, Vol. 10, pp. 43–51 (online), (1985).

[4] R. Swaminathan, S. Kang and R. Szeliski: On the motion and appearance of specularities in image sequences, *ECCV 2002*, pp. 508–523, (2002).

[5] R. Feris, R. Raskar and M. Turk: Specular reflection reduction with multi-flash imaging, *Proceedings. 17th Brazilian Symposium on Computer Graphics and Image Processing*, pp. 316–321, (2004).

[6] G. Klinker, S. Shafer and T. Kanade: The measurement of highlights in color images, *IJCV*, Vol. 32 , (1988).

[7] R. Bajcsy, S. Lee and A. Leonardis: Detection of diffuse and specular interface reflections and inter-reflections by color image segmentation, *IJCV*, Vol. 17, No. 3, pp. 241–272, (1996).

[8] R. T. Tan and K. Ikeuchi: Separating reflection components of textured surfaces using a single image., *IEEE PAMI*, Vol. 27, No. 2, pp. 178–93, (2005).

[9] H. L. Shen and Q. Y. Cai: Simple and efficient method for specularity removal in an image., *Applied optics*, Vol. 48, No. 14, (2009).

[10] P. Tan, S. Lin and L. Quan: Separation of Highlight Reflections on Textured Surfaces, *2006 IEEE CVPR*, Vol. 2, pp. 1855–1860 , (2006).

[11] D. Lee and H. Seung: Learning the parts of objects by non-negative matrix factorization, *Nature*, Vol. 401, No. October 1999, pp. 788–791, (1999).

[12] C. Févotte, N. Bertin and J. Durrieu: Nonnegative matrix factorization with the itakura-saito divergence: With application to music analysis, *Neural Computation*, Vol. 21, No. No. 3, pp. 793–830, (2009).

[13] J. Eggert and E. Korner: Sparse coding and NMF, *Neural Networks, 2004. Proceedings. 2004*, Vol. 2, No. 4, pp. 2529–2533, (2004).

[14] P. Hoyer: Non-negative matrix factorization with sparseness constraints, *The Journal of Machine Learning Research*, pp. 1–13, (2004).

[15] T. Virtanen: Monaural sound source separation by nonnegative matrix factorization with temporal continuity and sparseness criteria, *Audio, Speech, and Language Processing, IEEE*, Vol. 15, No. 3, pp. 1066–1074, (2007).

[16] M. Schmidt: Speech separation using non-negative features and sparse non-negative matrix factorization, *Computer Speech and Language 2008*, No. June 2008 , (2008).

[17] S. Choi: Algorithms for orthogonal nonnegative matrix factorization, *2008 IEEE International Joint Conference on Neural Networks* , pp. 1828–1832, (2008).

[18] N. Bertin, R. Badeau and E. Vincent: Enforcing Harmonicity and Smoothness in Bayesian Non-Negative Matrix Factorization Applied to Polyphonic Music Transcription, *IEEE Transactions on Audio, Speech, and Language Processing*, Vol. 18, No. 3, pp. 538–549, (2010).

[19] B. A. Olshausen and D. J. Field: Sparse coding of sensory inputs., *Current opinion in neurobiology*, Vol. 14, No. 4, pp. 481–7, (2004).