

Torrance-Sparrow モデルに基づく 反射特性・形状・光源位置の同時推定の検討

荻野 真佑 右田 剛史 尺長 健

岡山大学大学院 自然科学研究科

〒700-8530 岡山市津島中 3-1-1

{ogino,migita,shaku}@chino.cs.okayama-u.ac.jp

概要: 本稿では、光源位置を多様に変化させて撮影した画像列を基にして、物体の反射特性・形状、及び光源位置を同時推定する手法を提案する。本手法では、非線形最小2乗法を用いて反射特性・形状・光源位置の同時推定を行う。このため、基底となる Torrance-Sparrow モデルをいくつか用いて、点毎に独立な物体表面の反射特性を表現する。これにより、鏡面反射を含み、テクスチャを持つ物体を扱うことができる。ここで、同時推定に必要な初期値の作成には、対象物体の反射特性・形状、及び光源位置に関する実測データは必要ない。また、カメラは未校正でよい。木製の器や人物顔を対象物体として実験を行い、手法の有効性を確認した。

Simultaneous Recovery of Reflectance Property, Shape and Light Position based on Torrance-Sparrow Model

Shinsuke Ogino, Tsuyoshi Migita and Takeshi Shakunaga

Department of Information Technology, Faculty of Engineering, Okayama University
3-1-1 Tsushima-naka, Okayama, 700-8530, JAPAN

{ogino,migita,shaku}@chino.cs.okayama-u.ac.jp

Abstract: Given a set of images, where a point light source moves around an object, we can formulate recovery of the reflectance property, texture and shape of the object, as well as the positions of the light source. We herein describe an efficient algorithm to address this large-scale nonlinear optimization problem. Various objects, including toy models and human faces, are successfully recovered without any special devices nor a precise calibration.

1 はじめに

コンピュータグラフィックスにおいて、物体の反射特性・形状、及び光源位置を与えて画像を生成することを Rendering と言う。逆に、画像で観測される輝度値に基づいて、物体の反射特性・形状、及び光源位置などを獲得する手法を Inverse Rendering と呼び、活発な研究が行われている [1][2]。

これらの処理には、物体表面における光の反射を近似する反射モデルが必要である。そのための一般的なモデルとして、双方向反射分布関数 (Bidirectional

Reflectance Distribution Function:BRDF) が知られている。BRDF は入射光強度と反射光強度の割合を表現する関数である。物体表面の模様 (テクスチャ) を表現するには、点毎に異なる BRDF を用いればよい。このためのモデルとして、一般的に双方向テクスチャ関数 (Bidirectional Texture Function:BTF) が用いられる。BTF は BRDF を要素とする 2次元配列であり、様々なテクスチャを表現することができる。BTF を効率的に表現するために、関口ら [3] は、基底となるパラメトリック反射関数モデル (本稿ではこれを基底反射特性モデル、又は単に基底モ

デルと呼ぶ)をいくつか定義し、それらの線形結合により BTF を表現する方法を提案している。また、Inverse Rendering の枠組で、静止物体に対して光源位置を多様に変化させ、1 台の固定カメラで撮影した画像列 (以後、画像集合 A と呼ぶ) を推定のための入力として与え、非線形最小 2 乗法により反射特性と形状の同時復元を行っている。しかし、関口ら [3] の研究には次のような問題があった。まず、光源位置には実測データを与えており、光源位置を計測するために、光源の移動を制御するための装置等が必要になる。また、カメラのキャリブレーション等、画像集合 A の撮影とは別に手間がかかる。さらに、初期形状の獲得には 1 台のプロジェクタと 1 台のカメラから構成されるレンジファインダを用いており、画像集合 A の他に新たな画像を大量に撮影する必要があった。

本稿では、これらの問題を解決するために、画像集合 A のみを入力とした反射特性・形状・光源位置の同時推定法を提案する。提案手法では、光源を制御するような装置や形状を計測するような装置は用いない。また、カメラも未校正で構わない。カメラが未校正の場合は、入力画像のサイズに応じて焦点距離と画像中心を設定した単純なパラメータを用いる。提案手法は、部屋を暗くし、1 つの電球と 1 台のカメラがあれば行えるため、手軽に実現できるという特長がある。

2 基底反射特性モデルの線形結合による BTF の近似表現

静止物体に対して光源位置を多様に変化させ、1 台の固定カメラで撮影した画像列を考える。関口ら [3] は基底となる Torrance-Sparrow モデル [4] をいくつか定義し、それらの線形結合によって BTF を効率的に表現している。このとき、画像 f 枚目における点 p の画素値 i_{fp} は次式で表される。

$$i_{fp} = \sum_{m=1}^M w_{mp} M_{mfp} \quad (1)$$

ここで、

$$M_{mfp} = d_m \cos \beta_{fp} + s_m \frac{\exp(\rho_m \alpha_{fp}^2)}{\cos \gamma_p} \quad (2)$$

$$\cos \beta_{fp} = \mathbf{n}_p^T \mathcal{N}[\mathbf{l}_f - \mathbf{x}_p]$$

$$\cos \gamma_p = \mathbf{n}_p^T \mathcal{N}[\mathbf{v} - \mathbf{x}_p]$$

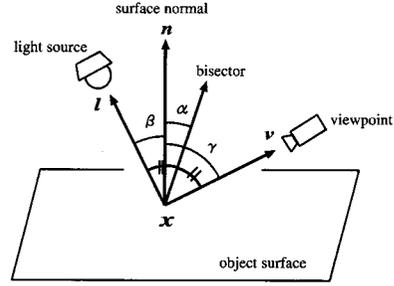


図 1: Torrance-Sparrow モデル

$$\cos \alpha_{fp} = \mathbf{n}_p^T \mathcal{N}[\mathcal{N}[\mathbf{l}_f - \mathbf{x}_p] + \mathcal{N}[\mathbf{v} - \mathbf{x}_p]]$$

であり、 w_{mp} は基底モデル M_{mfp} の結合重み、 M は用いる基底モデル数、 i_{fp} は RGB 成分からなる 3 次元ベクトル、 d_m は各要素に RGB 成分の拡散反射率を持つ 3 次元ベクトル、 s_m は各要素に鏡面反射率を持つ 3 次元ベクトル、 ρ_m は物体表面の粗さを表す係数、 $\mathbf{x}_p, \mathbf{n}_p$ はそれぞれ物体表面の 3 次元座標並びにその法線方向単位ベクトル、 \mathbf{l}_f は光源位置、 \mathbf{v} は視点位置、 $\mathcal{N}[\cdot]$ は $[\cdot]$ 内のベクトルのノルムを 1 に正規化する演算子を示す (図 1 参照)。

基底モデルに関して、まず、拡散反射成分の反射特性を表現するため、次の 3 つの基底モデルを導入する。ここでは、表記の煩雑さを避けるため、添字 f, p を省略する。

$$M_1 = \begin{pmatrix} 100 \\ 0 \\ 0 \end{pmatrix} \cos \beta, \quad M_2 = \begin{pmatrix} 0 \\ 100 \\ 0 \end{pmatrix} \cos \beta, \quad M_3 = \begin{pmatrix} 0 \\ 0 \\ 100 \end{pmatrix} \cos \beta \quad (3)$$

次に、鏡面反射成分の反射特性を表現する基底モデルを考慮する。鏡面反射は、光源色と物体表面の粗さが表現できればよい。このことから、以下のモデルを追加する。

$$M_4 = s_4 \frac{\exp(\rho_4 \alpha^2)}{\cos \gamma}, \quad M_5 = \dots \quad (4)$$

鏡面反射の強さは物体を構成する材質毎に異なるため、この基底モデルは材質の数に応じて追加すればよい。本稿では以後、 d_m, s_m, ρ_m をそれぞれ拡散反射係数、鏡面反射係数、粗さ係数と呼ぶ。

3 反射特性・形状・光源位置の同時推定

3.1 推定対象パラメータ

2章で述べた反射モデルを構成するパラメータを推定することで、反射特性・形状・光源位置の同時推定を行う。このとき、推定すべきパラメータは以下の通りである。

反射特性を表すパラメータには、拡散反射係数 d_m 、鏡面反射係数 s_m 、粗さ係数 ρ_m 、基底モデルの結合重み w_{mp} がある。このうち、拡散反射係数は式 (3) のように固定するため推定しなくてよい。したがって、反射特性に関する推定対象パラメータは鏡面反射パラメータ ($s_4, \rho_4, \dots, s_M, \rho_M$) と結合重み (w_{11}, \dots, w_{MP}) となる。ここで、 P は推定対象点数を示す。

物体の形状は3次元点 \mathbf{x}_p の集合で表される。 \mathbf{x}_p は、カメラの投影中心と画像平面上の点 (u_p, v_p) を通る直線上にあるので、カメラ位置からの距離によって表現できる。これにより、1点当たりの形状に関する推定パラメータ数を3から1に削減することができる。 \mathbf{x}_p は次式によって表される。

$$\mathbf{x}_p = \mathbf{v} + \lambda_p \mathbf{t}_p \quad (5)$$

ここで、 λ_p はカメラから物体までの距離、 \mathbf{t}_p はカメラ中心から点 p へ向かう単位ベクトルを示す。そのため、形状に関する推定対象パラメータは ($\lambda_1, \dots, \lambda_P$) となる。

光源位置は画像毎に異なるため、全画像に対してその3次元座標を推定する必要がある。そのため、光源位置に関する推定対象パラメータは (l_1, \dots, l_F) となる。ここで、 F は入力画像枚数を示す。

以上より、推定すべきパラメータベクトル \mathbf{p} は次のようになる。

$$\mathbf{p} = (\quad s_4^T, \rho_4, \dots, s_M^T, \rho_M, \quad w_{11}, \dots, w_{MP}, \quad \lambda_1, \dots, \lambda_P, \quad l_1^T, \dots, l_F^T)^T \quad (6)$$

例えば、入力画像枚数 40、推定対象点数 20000、基底モデル数 4 の場合、推定パラメータ総数は 100124 となり、10 万次元を超える。

3.2 評価関数

次の評価関数を最小化するようにパラメータを選ぶことで、反射特性・形状・光源位置の同時推定を行う。

$$E(\mathbf{p}) = \frac{1}{2} \sum_{f=1}^F \sum_{p=1}^P \mathbf{r}_{fp}^T \mathbf{r}_{fp} \quad (7)$$

$$\text{ここで、} \mathbf{r}_{fp} = \left(\sum_{m=1}^M w_{mp} \mathbf{M}_{mfp} \right) - \mathbf{i}_{fp}$$

この評価関数を最小化するため、次節に示す手法を用いる。

3.3 BDCG 法

3.3.1 パラメータ更新法

右田ら [5] は、共役勾配法に基づく BDCG 法 (Block Diagonal matrix preconditioned Conjugate Gradient method: ブロック対角行列による前処理付き共役勾配法) を提案している。大規模な非線形最適化問題において、ヘッセ行列が特殊な疎構造である場合、BDCG 法によりメモリの消費や計算量を大幅に抑えることができる。BDCG 法におけるパラメータの更新法を以下に示す¹。

$$\mathbf{p}_k = \begin{cases} \text{初期パラメータ} & (k = 0) \\ \mathbf{p}_{k-1} - \alpha_k \mathbf{d}_{k-1} & (k > 0) \end{cases} \quad (8)$$

$$\mathbf{d}_k = \begin{cases} \mathbf{C}_0^{-1} \mathbf{g}_0 & (k = 0) \\ \mathbf{C}_k^{-1} \mathbf{g}_k + \beta_k \mathbf{d}_{k-1} & (k > 0) \end{cases}$$

ここで、

$$\alpha_k = \arg \min_{\alpha} E(\mathbf{p}_k - \alpha \mathbf{d}_k)$$

$$\mathbf{g}_k = \nabla E(\mathbf{p}_k)$$

$$\beta_k = \frac{\mathbf{g}_k^T \mathbf{C}_k^{-1} \mathbf{g}_k}{\mathbf{g}_{k-1}^T \mathbf{C}_{k-1}^{-1} \mathbf{g}_{k-1}}$$

であり、 \mathbf{d}_k は反復 k 回目における探索方向ベクトル、 \mathbf{g}_k は勾配ベクトルを示す。本稿では、反復 5 回毎に現在の推定値を新たに初期値として再出発 ($\beta_k = 0$) を行う。また、 \mathbf{C}_k は探索方向に前処理を施す前処理行列を示し、これを適切に選ぶことで、反復回数を少なくすることができる。前処理行列を単位行列

¹ここで用いる $\alpha_k, \beta_k, \mathbf{d}_k$ は他の節に現れる反射モデルのパラメータとは異なる変数である

光源に関する推定パラメータには、光源位置と光源強度がある。物体表面が完全拡散反射面であり、無限遠光源の場合、Lambert モデルによれば、複数画像で観測される輝度値は次式で表される。

$$\begin{bmatrix} i_{11}^T & \cdots & i_{1P}^T \\ \vdots & \ddots & \vdots \\ i_{F1}^T & \cdots & i_{FP}^T \end{bmatrix} = \mathbf{L}\mathbf{N} \quad (16)$$

$$\mathbf{L} = \begin{bmatrix} \eta_1 \mathbf{l}_1^T \\ \vdots \\ \eta_F \mathbf{l}_F^T \end{bmatrix}, \quad \mathbf{N} = \begin{bmatrix} \mathbf{n}_1 \mathbf{d}_1^T & \cdots & \mathbf{n}_P \mathbf{d}_P^T \end{bmatrix}$$

ここで、 \mathbf{L} はフレーム毎の光源方向単位ベクトル \mathbf{l} と光源強度との積を要素に持つ $F \times 3$ の光源特性行列、 \mathbf{N} は点毎の物体表面の法線方向単位ベクトルと拡散反射率とのテンソル積から成る $3 \times 3P$ の面特性行列を示す。また、左辺の観測輝度値を表す行列を輝度行列 \mathbf{A} とし、 \mathbf{A} を特異値分解することにより次のように分解できる。

$$\mathbf{A} = \mathbf{V}(\Sigma \mathbf{U}^T) \quad (17)$$

なお、 \mathbf{V} は $F \times F$ の直交行列、 Σ は対角要素に特異値を持つ $F \times F$ の特異値行列、 \mathbf{U}^T は $F \times 3P$ の直交行列である。物体表面が完全拡散反射面の場合、 \mathbf{A} はランク 3 になる。式 (16),(17) より、特異値分解から光源特性行列を求め、光源特性行列から光源方向を得る方法が考えられる。しかし、 \mathbf{A} は次のようにも分解できる。

$$\mathbf{A} = (\mathbf{L}\mathbf{X})(\mathbf{X}^{-1}\mathbf{N}) \quad (18)$$

ここで、 \mathbf{X} は任意の正則行列である。このように、 \mathbf{A} の分解には不定性が残るため、正確な光源方向は分からない。したがって、この段階では \mathbf{X} は単位行列と仮定する。また、無限遠光源の仮定より光源位置を知ることはできない。よって、位置ではなく方向に着目して、特異値分解によって得られる \mathbf{V} の上位 3 次元を光源位置の初期値とした。また、本稿では鏡面反射や影も扱うため \mathbf{A} はランク 3 にならない。これも、光源特性行列から得られる光源方向に誤差が生じる原因となるが、これらは全て同時推定により修正する。また、光源強度は経験的に $\eta_1 = \cdots = \eta_F = 1$ を初期値とした。

形状に関して、式 (17) の特異値分解で得られる面特性行列から法線を求め、法線を用いて形状を求め、

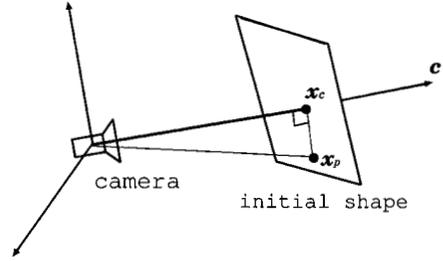


図 2: 初期形状

これを初期値に用いる方法が考えられるが、光源の場合と同じように不定性 \mathbf{X} がある。そこで、このような情報は用いず、カメラの向きに対して垂直な平面を初期値とする (図 2 参照)。本稿で行った実験では、平面を初期値としても十分に収束する。

3.6 推定手順

初めから複雑な反射モデルを用いて推定を行うと、あまり良い推定結果が得られないことが多い。そこで、単純な反射モデルから複雑な反射モデルへと段階を踏んで推定を行う。以下にその推定手順を示す。

- step1

前節で述べた方法で初期値を設定し、まず基底モデル数 3 で拡散反射のみを表現するパラメータの推定を行う。

- step2

step1 で得られた推定結果を初期値として、次に基底モデル数 4 で拡散反射と鏡面反射を表現するパラメータの推定を行う。

- step3

step2 で得られた推定結果を初期値として、再び基底モデル数 4 で推定を行う。このとき、光源強度の推定を含める。これにより、全体のパラメータの最適化を行う。

4 実験

4.1 実験条件

本稿では、実物体から反射特性・形状・光源位置の同時推定を行うために、静止物体に対して光源位置を多様に変化させ、1 台の固定カメラで撮影を行っ

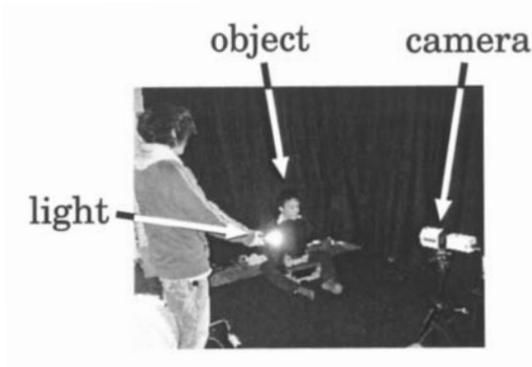


図 3: 物体を撮影した環境



図 4: 「木目の器」の入力画像（一部）
及び初期結合重み（右端）

た（図 3 参照）．この部屋は外からの光が入射しない暗室であり，撮影時には部屋の照明を消した．対象物体には，木目の器，人形，人物顔の 3 種類を用いた．なお，BCDG 法の反復回数は step1 では 100 回，step2，step3 では 200 回とした．また，推定精度の評価には次式で表される RMS 誤差を用いた．

$$\text{RMS Error} := \sqrt{\frac{\sum_{f=1}^F \sum_{p=1}^P \mathbf{r}_{fp}^T \mathbf{r}_{fp}}{3FP}} \quad (19)$$

4.2 実画像を用いた反射特性・形状・光源位置の同時推定実験

4.2.1 木目の器

木目の器を対象として，反射特性・形状・光源位置の同時推定実験を行った．入力画像は，光源を変化させ，カメラ（本実験では，SONY XC-003 を用いた）で撮影した計 36 枚の画像を用いた．画像サイズは 128×296 であり，このうち物体領域内の推定対象点数は 25480 であった．また，カメラは未校正であり，画像サイズに応じて焦点距離と画像中心を設定した単純なパラメータを用いた．本実験で用

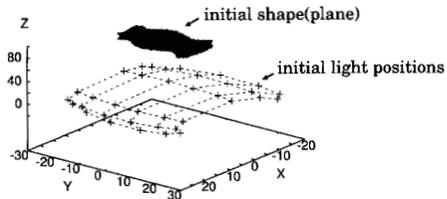


図 5: 「木目の器」の初期形状・初期光源位置



図 6: 「木目の器」の推定形状

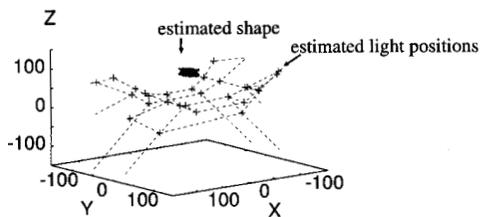


図 7: 「木目の器」の推定光源位置と推定形状の位置関係

いたカメラの射影行列 P を次式に示す．

$$P = \begin{bmatrix} 1000 & 0 & 64 & 0 \\ 0 & 1000 & 148 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

入力画像の一部と初期結合重みを図 4 に，初期形状・初期光源位置を図 5 に示す．なお，推定すべきパラメータ数は 127512 であった．また，撮影した画像は RGB それぞれ 8 ビットのカラー画像で，画素は 0 から 255 の値を取る．観測輝度が飽和している画素については，その画素に対応する \mathbf{r}_{fp} の項を評価関数から外した．

推定形状を図 6 に，推定光源位置と推定形状の位置関係を図 7 に，左から入力画像・復元画像・拡散反射成分のみを表示した画像・鏡面反射成分のみを表示した画像，及び推定結合重みを図 8 に示す．反射特性に関して，図 8 の右端下に示す物体上端における数点の鏡面反射の結合重みに対しては十分な結果を得ることができなかったが，復元画像・拡散反射画像・鏡面反射画像から拡散反射成分と鏡面反射

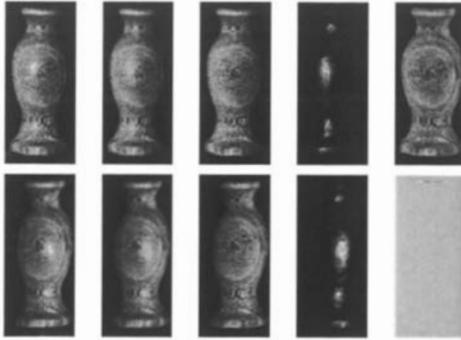


図 8: 「木目の器」の入力と推定結果: 左から入力画像・復元画像・拡散反射画像・鏡面反射画像, 及び推定結合重み (上 拡散成分, 下 鏡面成分)

成分の分離は良好に行われたことが分かる。また、RMS 誤差は約 5 であり、これは入力画像に対する復元画像の誤差が画素値の範囲に比べて約 2%であることを示す。このことから、反射特性は概ね良好に推定できていると言える。形状も、図 6 に示す通り丸みのある滑らかな結果が得られ、概ね良好に推定できていることが分かる。一方、図 7 に示す光源位置に関して 36 個のうち丸枠の外にある 5 つの光源が大きく離れた。この原因は次のように考えられる。すなわち、鏡面反射の強さは、物体から光源までの距離、結合重み、粗さ係数によって決まる。ここで、光源までの距離に関する情報が他のパラメータに吸収されたことにより、光源位置を正しく求められなかったと考えられる。一方、カメラは未校正であっても同時推定を行えることが確認できた。

4.2.2 人形

人形 (材質: 非フタル酸系塩化ビニール) を対象として、反射特性・形状・光源位置の同時推定実験を行った。入力画像は、手で持った電球を移動させながら、前節と同じカメラで撮影した計 30 枚の画像を用いた。

対象物体と推定領域、及び推定形状を図 9 に、左から入力画像・復元画像・拡散反射成分のみを表示した画像・鏡面反射成分のみを表示した画像を図 10 示す。図 9 (右) に示す通り、形状は概ね良好に推定できていることが分かる。一方、図 10 の復元画像・拡散反射画像・鏡面反射画像に示す通り、鏡面反射成分の多くが拡散反射成分と解釈され、反射特性

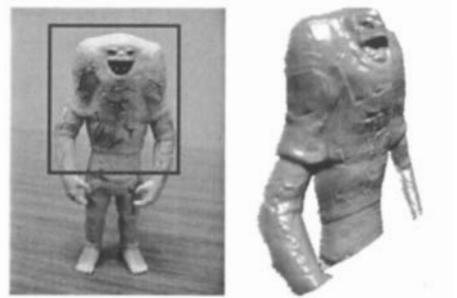


図 9: 「人形」の推定領域 (左), 推定形状 (右)

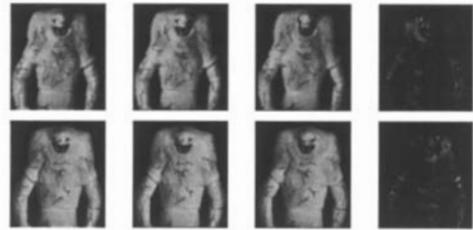


図 10: 「人形」の入力と推定結果: 左から入力画像・復元画像・拡散反射画像・鏡面反射画像

の推定結果は不十分であった。光源位置も木目の器と同じ原因から、推定結果は不十分であった。また、その他の人形に対する実験結果を図 11,12 に示す。

4.2.3 人物顔

人物顔を対象として、反射特性・形状・光源位置の同時推定実験を行った。入力画像には、前節と同様に光源を移動させながら、デジタルカメラ (本実験では、Canon IXY DIGITAL 600 を用いた) で撮影した計 30 枚の画像を用いた。入力画像の一部を図 13 に示す。

推定形状を図 14 に、左から入力画像・復元画像・拡散反射成分のみを表示した画像・鏡面反射成分のみを表示した画像を図 15 に示す。図 14 に示す通り、形状は概ね良好に推定できていることが分かる。一方、図 15 の復元画像・拡散反射画像・鏡面反射画像から拡散反射成分と鏡面反射成分について目や鼻まわりに推定が不十分な点がいくつかあることが分かる。これは、鼻による cast shadow が生じていること等が原因と考えられる。光源位置も木目の器と同じ原因から、推定結果は不十分であった。

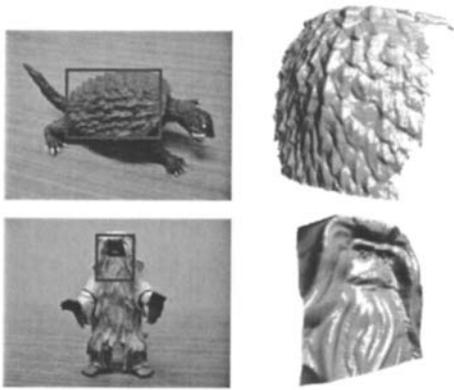


図 11: その他の人形と推定領域 (左), 推定形状 (右)

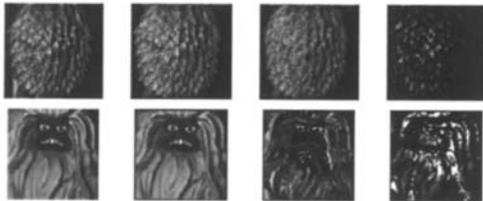


図 12: その他の人形の入力画像・復元画像・拡散反射画像・鏡面反射画像



図 13: 「人物顔」の入力画像 (一部)

5 おわりに

本稿では、実物体を撮影した画像列のみからの対象物体の反射特性・形状、及び光源位置の同時推定を行った。反射モデルを構成するパラメータを推定することで、反射特性・形状・光源位置の同時推定が行える。反射モデルには基底反射特性モデルの線形結合による反射特性の表現法を用い、パラメータ推定には BDCG 法を用いた。入力画像は、静止物体に対し光源位置を多様に変化させ、1 台の固定カメラで撮影したものを用いた。また、入力画像列から得られる輝度値を利用した初期値設定法や 3 段階構成の推定手順を導入した。

本手法の有効性を確認するために、木目の器と人形数体と人物顔を対象として、反射特性・形状・光源位置の同時推定実験を行った。全てにおいて、形



図 14: 「人物顔」の推定形状



図 15: 「人物顔」の入力と推定結果: 左から入力画像・復元画像・拡散反射画像・鏡面反射画像

状は十分な推定結果が得られたが、光源位置の推定結果は不十分であった。反射特性について、木目の器は良好に推定できた。しかし、cast shadow が生じているものは、推定結果が不十分であった。これは、用いた反射モデルでは cast shadow を扱うことができないためである。また、カメラが未校正でも同時推定が行えることを確認した。

今後の課題として、光源位置ではなく方向に着目し無限遠光源を仮定したパラメータ推定、及び扱える光学現象や材質の幅を広げた反射モデルに基づく同時推定などが挙げられる。

参考文献

- [1] 佐藤洋一, “イメージベースドモデリング,” 計測自動制御学会誌計測と制御, vol.41, no.9, pp.633-638, 9 2002.
- [2] 佐藤洋一, 向川康博, “インバースレンダリング,” 情報処理学会研究報告 CVIM 2004-145-9, pp.65-76, 9 2004.
- [3] T.Migita, S.Sekiguchi, and T.Shakunaga, “Estimation of bidirectional texture function from bizenware images,” 11th Korea-Japan Joint Workshop on Frontiers of Computer Vision, pp.171-176, 1 2005.
- [4] K.E. Torrance, and E.M. Sparrow, “Theory for off-specular reflection from roughened surfaces,” Journal of Optical Society of America, vol.57, pp.1105-1114, 1967.
- [5] 右田剛史, 天野晃, 浅田尚紀, “3 次元形状・運動復元のための高速非線形最適化計算法,” 情報処理学会論文誌, vol.44, no.11, 11 2003.