

形状・反射特性・光源の同時推定法の改良と 様々な実画像への適用

槌谷 洋介¹ 右田 剛史¹ 尺長 健¹

概要: 本稿では、単一点光源の位置を多様に変化させて撮影した画像列を入力とし、撮影された物体の形状・反射特性・光源位置を推定する。従来の手法は反射モデルやカメラモデルが非線形であるため推定に時間が掛かっていた。本稿では、単純な照度差ステレオ法を用いた推定結果を初期パラメータとして与えることで、収束の高速化と安定性の向上を目指す。また、Multi-PIE Database を含む数百組の実画像を対象として推定実験を行った。その結果、良好な復元結果が得られた。また、最大で 10 倍程度の速度で復元が可能になることを確認した。

1. はじめに

画像列から物体形状・反射特性・光源位置を同時に推定する方法が右田ら [1][2][3] によって提案されている。この方法では、入力画像列を再現する形状・反射特性・光源を非線形最小 2 乗法で探索することで推定を行う。しかし、推定に用いる反射モデルが鏡面反射を含んでおり複雑であるため、収束が非常に遅いという問題があった。そこで本稿では、単純な照度差ステレオ法の結果を初期値として与えることによって収束速度の向上を目指し、様々な物体で実験を行う。

2. 形状・反射特性・光源の同時推定原理

本章では、右田ら [1] の方法 (以下、FR 法と呼ぶ) を説明する。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すると、図 1 の様な画像列が得られる。ここで推定に用いない画素は灰色で表す。FR 法では、図 1 の様な画像列から図 2 の様な形状を復元するとともに、反射特性・光源を同時に推定する。

2.1 定式化

F 枚の画像が与えられ、画像毎に光源位置 l_f が異なる。また、光源色は全画像に共通で s とする。

画素は対象物体と背景に分けられ、対象物体上の推定対象画素数を P とする。物体形状は、各画素に対応する面素までの距離 (depth) λ_p から成る距離画像で表す。面素の色 (反射率) を表す RGB ベクトルを $d_p = (w_{0p}, w_{1p}, w_{2p})^T$

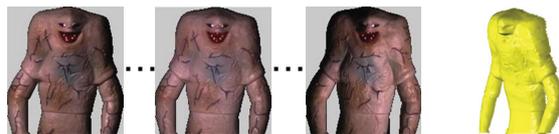


図 1 光源を変化させて撮影した画像列

図 2 復元形状

とし、鏡面反射の強度を w_{3p} とする。面素の座標 x_p は距離 λ_p とカメラモデルから計算される。

カメラは透視投影モデルであり、カメラから物体側を z 軸の正の方向とし、焦点距離を ℓ 、位置を $v = (0, 0, -\ell)^T$ とする。このカメラモデルでは、 ℓ は射影歪効果を調整するパラメータであり、像の大きさは変化しない。このため、 $\ell \rightarrow \infty$ として平行投影を扱うこともできる (ℓ は常に逆数で用いられるため、実装上は ℓ^{-1} を表す変数を用いる)。画像座標系の原点は画像中心とし、画素の座標を (x_p, y_p) とすると、画素の座標と画素からカメラへの方向 v'_p は次のように表される。

$$x_p = (x_p, y_p, 0)^T + \lambda_p u_p, \quad (1)$$

$$\text{ただし, } u_p = (x_p/\ell, y_p/\ell, 1)^T$$

$$v'_p = (v - x_p)/\ell = -(1 + \lambda_p/\ell)u_p \quad (2)$$

面素の法線ベクトル n_p は近傍の面素の座標から次のように計算される。

$$n_p = (x_R - x_L) \times (x_T - x_B) \quad (3)$$

ここで、 L, R, T, B は着目点 p の左右上下の点を表すが、物体境界外になる場合は代わりに p 自身を表す。例えば、 p の右の画素が推定対象外なら $R = p$ である。このため、幅が 2 に満たない領域は推定対象としない。なお n_p は単位

¹ 岡山大学
Okayama University

長である必要はない。

画像の輝度値 e_{fp} は Torrance-Sparrow モデルによって次のように表される。

$$e_{fp} = \eta_{fp} (d_p D_{fp} + w_{3p} s S_{fp}), \quad (4)$$

$$\text{ただし, } D_{fp} = \max\left(0, \frac{n_p^T l'_{fp}}{|n_p| |l'_{fp}|}\right), \quad (5)$$

$$S_{fp} = \frac{1}{\frac{n_p^T v'_p}{|n_p| |v'_p|}} \exp(\rho \alpha^2) \quad (6)$$

D_{fp} , S_{fp} はそれぞれ拡散反射, 鏡面反射の強度を表す係数であり, ρ は鏡面反射の広がりを表すパラメータで, 物体全体で共通である。また, 以下の定義を用いた。

$$\eta_{fp} = |l'_{fp}|^{-2}, \quad l'_{fp} = l_f - x_p, \quad (7)$$

$$\alpha_{fp} = \arccos\left(\frac{n_p^T b_{fp}}{|n_p| |b_{fp}|}\right), \quad b_{fp} = \frac{l'_{fp}}{|l'_{fp}|} + \frac{v'_p}{|v'_p|}. \quad (8)$$

この式を使って入力画像を再現するパラメータ (形状・反射特性・光源) を探索する。推定の目的関数 E は次のように表される。

$$E(\mathbf{p}) = (1/2) \sum_{(f,p)} \mathbf{r}_{fp}^T \mathbf{r}_{fp} \quad \text{where } \mathbf{r}_{fp} = e_{fp} - e'_{fp} \quad (9)$$

ここで, \mathbf{p} は全推定パラメータを含むベクトル (以後, その要素数を N とする), e'_{fp} は入力画像の輝度値 (RGB) である。なお, 総和の添字 (f,p) は, 必ずしも直積 $[0, 1, \dots, F-1] \times [0, 1, \dots, P-1]$ の全要素ではない。モデルに従わない影画素や飽和画素を閾値処理で判定し, 幾つかの (f,p) の組み合わせは除外される。 N は概ね $O(10^5)$, \mathbf{r} の次元は概ね $O(10^6)$ を想定する。

2.2 非線形最小 2 乗法の解法

関数 $E(\mathbf{p})$ を Levenberg-Marquardt 法 [4] で最小化する。この手法は, 現在の推定値 \mathbf{p} における E の勾配 \mathbf{g} とヘッセ行列 \mathbf{H} を用いて次式に従って推定値を更新する手法である。矢印 \leftarrow は変数の更新を表す。

$$\mathbf{p} \leftarrow \mathbf{p} - \alpha(\mathbf{H} + \mu \mathbf{I})^{-1} \mathbf{g} \quad (10)$$

ここで, α , μ は発散を防ぐために用いられる適当なパラメータである。なお, 式 (10) で発散を防ぐ μ を動的に調節する方法 [4] は大規模連立方程式の実行数が増えるので, 本稿では用いない。代わりに係数 α を導入し, これを推定する直線探索を行う。

関数 E が残差の 2 乗和なので, 残差ベクトル \mathbf{r} を使って $E(\mathbf{p}) = (1/2) \mathbf{r}^T \mathbf{r}$ と書くことができる。 \mathbf{r} の第 i 要素を \mathbf{p} の第 j 要素で偏微分した導関数 $(\partial r_i / \partial p_j)$ を i 行 j 列要素とするヤコビ行列 \mathbf{J} を用いて, ヘッセ行列 \mathbf{H} と勾配 \mathbf{g} は次のように表される。

$$\mathbf{H} = \mathbf{J}^T \mathbf{J}, \quad \mathbf{g} = \mathbf{J}^T \mathbf{r} \quad (11)$$



図 3 初期形状

2.3 初期値と多段階推定

文献 [1] では, 初期形状, 初期光源位置, 初期反射特性は以下のようにしている。これは複数の実験例で局所解に陥ることなく機能しているが, 極めて粗い初期値であるため収束に時間がかかる。

形状

平面を初期値とした場合, 凸物体を凹物体として復元することがある。復元対象はほとんどの場合凸状であるため, 図 3 のように少し膨らみを持った曲面を初期値とする。この曲面は, 次式を用いて計算する。

$$\lambda_p = au_p^2 + bv_p^2$$

ここで, a, b は曲面の出っぱり具合を決める係数であり, 復元対象物体に応じて適当に決める。

反射特性

拡散反射係数 d_p は物体色の RGB の比である。これは点毎に入力画像列全フレームの平均値を用いる。

$$d_p = \frac{1}{F} \sum_{f=1}^F e_{fp} \quad (12)$$

光源

光源に関するパラメータには光源色 \mathbf{s} と, 光源位置 l_f がある。このうち光源色は白色光源と仮定して, $\mathbf{s} = (1, 1, 1)^T$ を初期値とする。光源位置は, 物体からの距離 d を適当に与え, $l_f = (0, 0, -d)^T$ を初期値とする。

推定の安定性や速度のために, 文献 [1] では推定を複数の段階に分けて行っている。まず, 「Step 1」では鏡面反射項を除く拡散反射のみのモデルで推定を行う。また, 大きな μ 値を使うことで局所解を避ける。次に, 「Step 2」では鏡面反射項を含むモデルを使って推定を行う。最後に, 「Step 3」では「Step 2」から μ を 0 に近づけて推定を行う。

3. 初期値の線形推定

2.3 節で述べた初期値は粗い初期値であるため収束が非常に遅い。Step1 を Lambert 反射モデルに基づく照度差ステレオ法に置き換えることで短時間での収束を目指す。

3.1 Lambert モデル

物体表面が完全拡散反射面であり、無限遠点光源に照らされている時、観測される輝度値は次式で表される。

$$M = LN \quad (13)$$

$$M = \begin{bmatrix} e_{11}^T & \cdots & e_{1P}^T \\ \vdots & \ddots & \vdots \\ e_{F1}^T & \cdots & e_{FP}^T \end{bmatrix}, \quad (14)$$

$$L = \begin{bmatrix} \eta_1 l_1^T \\ \vdots \\ \eta_F l_F^T \end{bmatrix}, N = \begin{bmatrix} n_1 d_1^T & \cdots & n_P d_P^T \end{bmatrix}$$

ここで、 M は全観測を含む、 $F \times 3P$ の観測行列、 L は画像毎の光源方向単位ベクトル l_f と光源強度 η_f の積を要素とする $F \times 3$ の光源特性行列、 N は点毎の単位法線ベクトル n_p と色 (反射率) d_p を要素に持つベクトルのテンソル積から成る $3 \times 3P$ の面特性行列を示す。

3.2 SVD による推定

式 (13) のように M をランク 3 に分解することで形状と光源を同時に推定できる [5]。このために、行列 M を次のように特異値分解する。

$$M = U \Sigma V^T \approx \hat{U} \hat{\Sigma} \hat{V}^T \quad (15)$$

ここで、 U 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、 V^T は $F \times 3P$ の直交行列である。実画像ではランクは 3 にならない。そこで、上位 3 つの特異値に対応する $F \times 3$ 行列 \hat{U} 、 3×3 行列 $\hat{\Sigma}$ 、 $3 \times 3P$ 行列 \hat{V}^T で近似する。

この特異値分解によって得られる直交行列 $\hat{U} \hat{\Sigma}$ を L 、 \hat{V}^T を N として用いる。

3.3 SVDMD による推定

特異値分解は全観測 M を要するため、鏡面反射や影、画像ノイズの影響を受け、光源方向や法線方向の推定がうまく出来ないことがある。この問題に対して Singular Value Decomposition with Missing Data (SVDMD) が一般に用いられている [6]。SVDMD では欠損のある行列の分解を目的としており、影や鏡面反射のある画像列で有用だと考えられる。本稿では、 M の密部分に対して SVD を行った後、 L の推定と N の推定を交互に行う交互最小 2 乗法によって SVDMD を実装する。この推定は式 (13) の 3 行、または 3 列を抜き出した次の式に基づいて行う。

$$[e_{f1}^T \cdots e_{fP}^T] = \eta_f l_f^T [n_1 d_1^T \cdots n_P d_P^T] \quad (16)$$

$$\begin{bmatrix} e_{1p}^T \\ \vdots \\ e_{Fp}^T \end{bmatrix} = \begin{bmatrix} \eta_1 l_1^T \\ \vdots \\ \eta_F l_F^T \end{bmatrix} n_p d_p^T \quad (17)$$

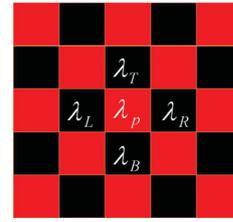


図 4 Red-Black 法の分け方

式 (16),(17) から、 $\eta_f l_f^T$ 、 $n_p d_p^T$ の更新は以下の線形最小 2 乗法によって行う。

$$\eta_f l_f^T \leftarrow [\cdots e_{fp}^T \cdots] [\cdots n_p d_p^T \cdots]^+ \quad (18)$$

$$n_p d_p^T \leftarrow \begin{bmatrix} \vdots \\ \eta_f l_f^T \\ \vdots \end{bmatrix}^+ \begin{bmatrix} \vdots \\ e_{fp}^T \\ \vdots \end{bmatrix} \quad (19)$$

ここで式 (18) の右辺の添字 p は必ずしも $[1, 2, \dots, P]$ の全要素ではない。同様に式 (19) の右辺の添字 f も必ずしも $[1, 2, \dots, F]$ の全要素ではない。これらは e_{fp} の閾値処理で影や飽和の画素を除いたものである。この計算は、以後の FR 法の計算と重複するため、完全な収束を待たずに数回で停止させる。

3.4 GBR 不定性

光源と法線を同時に推定した場合、次式のような 3×3 の正則行列 X で表される不定性が残る。

$$M = L X X^{-1} N \quad (20)$$

ただし、法線に対して積分可能条件を課すことで、3 自由度の Generalized Bas-Relief (GBR) 不定性のみが残る [7]。法線 n_p ・物体の 3 次元位置 x_p ・光源位置 l_p は GBR 変換を表す行列 G を用いて、

$$x_p^* = G x_p \quad (21)$$

$$n_p^* = G^{-T} n_p \quad (22)$$

$$l_f^* = G l_f \quad (23)$$

$$\text{ただし, } G = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ a & b & c \end{bmatrix} x_p \quad (24)$$

と変換しても同じ画像が生成される。ここで、 a, b, c は未知パラメータである。

なお、FR 法では近接点光源を用いているため、GBR 不定性を解消することが期待できる [8], [9]。

3.5 ポアソン方程式による積分

法線 n_p から画素までの距離 (depth) λ_p を復元するためにポアソン方程式を用いた。復元対象となる点数が多いと収束に時間がかかるため、Red-Black SOR (Successive

表 1 推定対象

	推定対象点数	画像数	光源配置
木目の器	24,025	36	図 6
備前焼	39,196	36	図 11
表情 1(笑顔)	63,127	24	図 17
表情 2(驚き)	68,966	24	図 17
100 人顔画像	30,564 - 48,286	24	図 17
Multi-PIE	17,281 - 40,589	18	図 25

Over-Relaxation) 法 [10] を用いた. SOR 法による λ_p 更新は次のように行う.

$$\lambda_p \leftarrow \lambda_p + \frac{\omega}{4}(\lambda_L + \lambda_R + \lambda_T + \lambda_B - 4\lambda_p - \nabla \cdot \mathbf{q}_p) \quad (25)$$

$$\mathbf{q}_p = (n_{px}/n_{pz}, n_{py}/n_{pz})$$

ここで, $\lambda_L, \lambda_R, \lambda_T, \lambda_B$ は λ_p の左右上下の距離値である.

ポアソン方程式は近接する 4 点に対して依存性があり, 互いに独立な要素を並列に処理するために Red-Black 法を用いた. Red-Black 法は全点を図 4 の市松模様のように 2 つの集団に分け, 片方ずつ処理することで依存性を回避する方法であり, 並列処理が可能になる.

6 万点程度の顔形状が収束するのに $\omega = 1.0$ の SOR 法 (Gauss-Seidel 法) で 80 秒以上要する処理を $\omega = 1.99$ とすることで 1.7 秒, さらに Red-Black SOR 法を用いることで 0.5 秒程度で収束した.

なお, この段階ではカメラモデルは平行投影を仮定しており, 後の FR 法で透視投影での形状に修正する.

3.6 その他のパラメータの初期推定

式 (13) に基づいて推定した形状・光源方向を式 (4) のモデルに適応させるために物体色, 光源強度 (光源距離) η_f を推定する. また, η_f を光源までの距離に変換する. 更に, 物体色は, 法線方向と光源方向, 画素値から最小 2 乗法を用いることで再推定する.

推定対象約 6 万点, 24 枚の顔画像列から形状・反射特性・光源を推定し, FR 法の初期値として与えるのに, 1.2 秒程度必要とする.

4. 実験

物体形状・反射特性・光源の同時推定実験を行った. 推定法には 2.2 節で述べた非線形最小 2 乗法を用い, FR 法の初期値は 2.3 節で述べたものを用いる. さらに 3.2, 3.3 節で述べた方法による初期値を用いた場合の結果を載せる. 実画像実験では多種の条件下における推定を行った. 実験に用いる画像の各画素値は 8bit 整数で表され 0-255 の範囲を取る. 各対象に関する推定対象点数と画像のフレーム数を表 1 に載せる. 実験に用いた計算機の CPU は Core i7 3930K, GPU は GTX580 である.

4.1 実物体を用いた復元

木目の器

図 5 に入力画像例を示す. これは図 6 のような円筒状に光源を配置して撮影した画像である. 表 2 に各方法における FR 法



図 5 木目の器の入力画像例

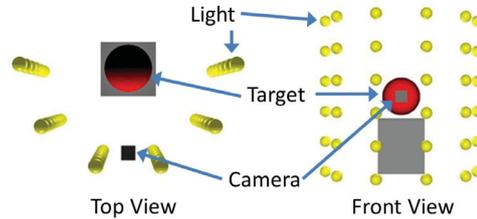


図 6 撮影環境 1

表 2 各 Step の反復回数 (木目の器)

	Step1	Step2	Step3
FR	20	20	20
SVD+FR	-	10	10
SVDMD+FR	-	-	7

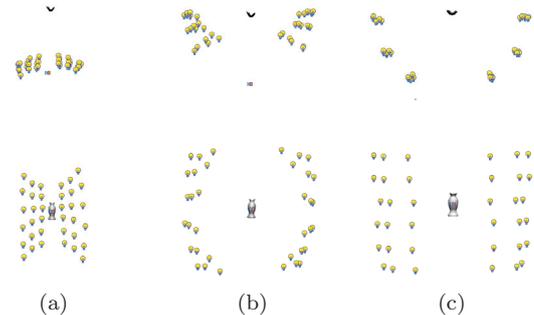


図 7 木目の器の光源配置 ((a):SVD 後, (b):SVDMD 後, (c):SVDMD+FR 後)

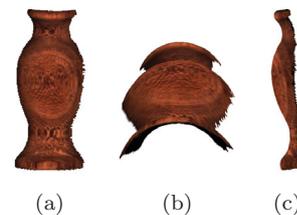


図 8 木目の器の推定形状 ((a):正面,(b):斜め下,(c):横)

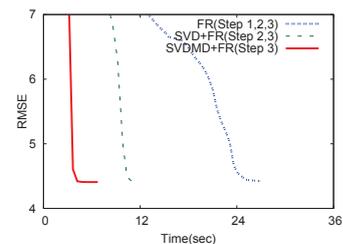


図 9 木目の器の処理時間と残差

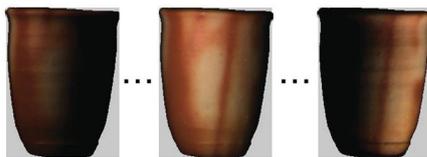


図 10 備前焼の入力画像例

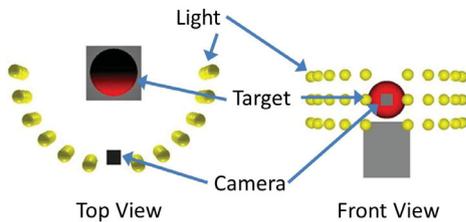


図 11 撮影環境 2

表 3 各 Step の反復回数 (備前焼)

	Step1	Step2	Step3
FR	10	25	130
SVD+FR	-	10	90
SVDMD+FR	-	-	10

の各 Step の反復回数を示す。これはそれぞれ反復回数を変えて実験を行った結果から最適な回数を選んだ。SVD による初期値を用いた場合、Step 2,3 を 10 回ずつ反復推定することで収束する。SVDMD による初期値を用いた場合 Step 2,3 を用いるよりも、Step 3 のみを用いることで速く収束し、7 回の反復で収束した。

図 7 に初期値の光源配置と最終的な光源配置の推定結果を示す。この物体は図 6 の光源配置を用いて撮影したため、光源位置は円筒状に並ぶはずである。(a)(b)を見ると SVD を用いた推定では光源が物体正面に集まり、うまく推定できていないが、SVDMD を用いた推定では対象物体の周囲に推定されており、比較的良く推定できている。これは画像列が影を多く含み、全観測を必要とする SVD ではうまく推定できないことが要因だと考えられる。(c)を見ると、光源位置が物体にそってきれいな円柱状に並んでおり、良好な推定が行えているのが分かる。

図 8 に最終結果の形状を示す。どの初期値でも目視で分からないほどの GBR 不定性の除いて最終結果は変わらないため、最も速く収束する SVDMD を初期値としたものを示す。木目の器は回転体であるので、(a)を半分にした形と(c)は同形になり。(b)の底の形も半円となるはずである。GBR 変換によって多少の傾きと奥行きが潰れが見られるが、その通りに推定できている。

図 9 に計算時間と残差の関係を示す。SVD による初期値を用いることで 1/2 程度、SVDMD による初期値を用いることで 1/5 程度の時間で収束していることが分かる。

備前焼

図 10 に入力画像例を示す。これは図 11 のような光源配置で撮影した画像列である。表 3 に各方法における FR 法の各 Step の反復回数を示す。これはそれぞれ反復回数を変えて実験を

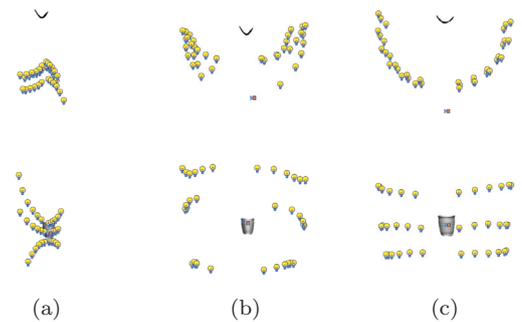


図 12 備前焼の光源配置 ((a):SVD 後, (b):SVDMD 後, (c):SVDMD+FR 後)

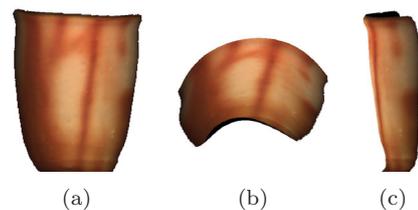


図 13 備前焼の推定形状 ((a):正面,(b):斜め下,(c):横)

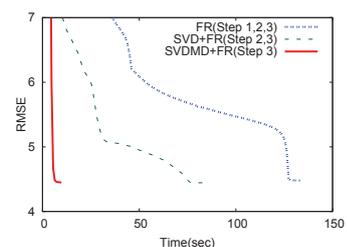


図 14 備前焼の処理時間と残差

行った結果から最適な回数を選んだ。SVDMD による初期値を用いた場合 Step 2,3 を用いるよりも、Step 3 のみを用いることで速く収束した。

図 12 に初期値の光源配置と最終的な光源配置の推定結果を示す。光源が真横までくるような非常に影の多い画像列なので、全観測を必要とする SVD では推定は難しいため、(a)の SVD を用いた推定では光源が物体側面に回り込まず失敗しているが、(b)の SVDMD を用いると概ね推定できている。(c)を見ると、光源位置が物体にそって 3 段の円周状に並んでおり、良好な推定が行えているのが分かる。

図 13 に最終結果の形状を示す。どの初期値でも、目視で分からないほどの GBR 不定性の除いて最終結果は変わらないため、最も速く収束する SVDMD を初期値としたものを示す。備前焼は木目の器と同じく回転体である。(b)の底部分や(c)の側面図を見ると GBR 変換によるものと思われる多少の潰れが見られるものの、回転体として復元できている。図 14 に計算時間と残差の関係を示す。SVD による初期値を用いることで 2/3 程度、SVDMD による初期値を用いることで 1/10 以下の時間で収束していることが分かる。



図 15 笑顔の入力画像例

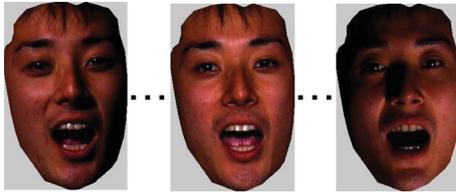


図 16 驚き顔の入力画像例

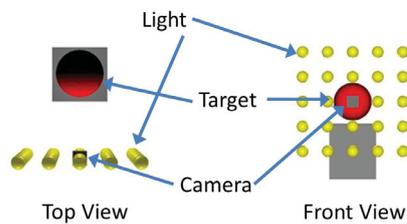


図 17 撮影環境 3

表 4 各 Step の反復回数
(a) 笑顔

	Step1	Step2	Step3
FR	10	10	100
SVD+FR	-	20	50
SVDMD+FR	-	-	50

(b) 驚き顔

	Step1	Step2	Step3
FR	10	10	100
SVD+FR	-	-	50
SVDMD+FR	-	-	50

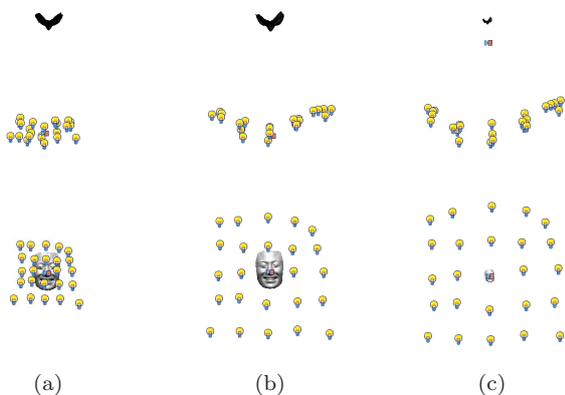


図 18 笑顔の光源配置 ((a):SVD 後, (b):SVDMD 後, (c):SVDMD+FR 後)

4.2 人物の顔を用いた復元 表情のある顔

図 15,16 のような画像列を入力として推定を行った。これは

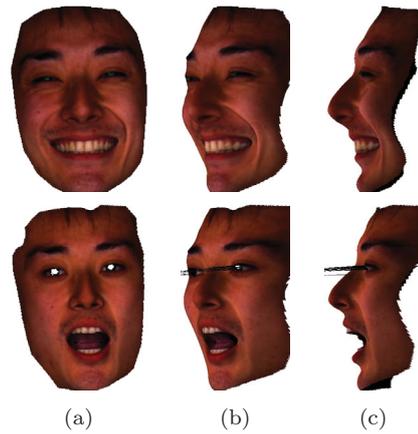


図 19 表情のある顔の推定形状 ((a):正面,(b):斜め横,(c):横)

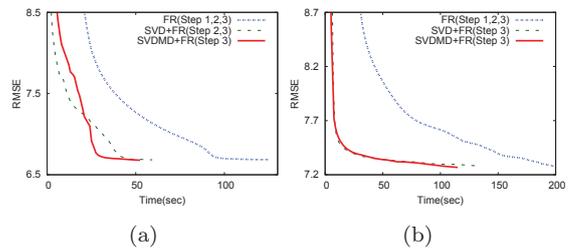


図 20 処理時間と残差 ((a):笑顔, (b):驚き顔)

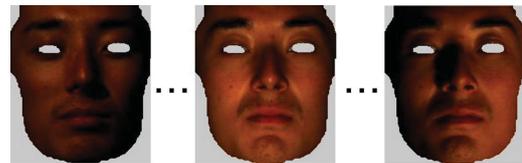


図 21 人物顔の入力画像例

図 17 のような格子状に光源を配置した環境で撮影した。表 4 に各方法における FR 法の各 Step の反復回数を示す。図 18 は推定光源である。光源は 5×5 の格子状に並んでおり、良好に推定されている。図 19 に推定形状を示す。唇や鼻など細かな部分でも形がしっかりと復元されており、良好な推定結果が得られている。入力画像の中には鼻周辺では cast shadow が起こっているものも含まれるが、図 19 ではそれに起因する形状の歪みは見られない。両目が飛び出したように見えるのは、目に生じる光沢が式 (4) の反射モデルに従わないためと考えられる。

図 20(a) に笑顔の画像列を対象とした推定実験の処理時間と残差の関係を示す。SVD, SVDMD による初期値を用いることで $1/3$ 程度の時間で収束していることが分かる。図 20(b) は驚き顔の場合であり、SVD, SVDMD による初期値を用いることで $1/2$ 程度の時間で収束していることが分かる。収束にかかる時間を見ると大きな差はないが、初期値を与えることで短時間で残差が小さくなっていることが分かる。

100 人の顔

図 21 の様な画像列を入力として実験を行う。これは図 17 と同じ環境で 100 人の人物顔を対象として撮影を行ったうちの 1 例である。表 5 に各方法における FR 法の各 Step の反復回数

表 5 各 Step の反復回数 (100 人共通)

	Step1	Step2	Step3
FR	20	20	150
SVD+FR	-	10	100
SVDMD+FR	-	-	50

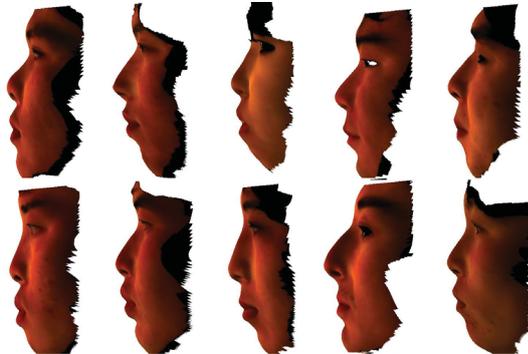


図 22 人物顔の推定形状例

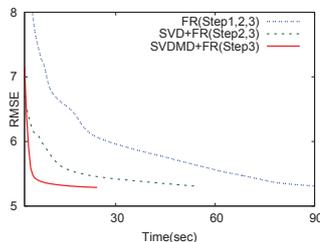


図 23 人物顔 100 人の処理時間と残差平均

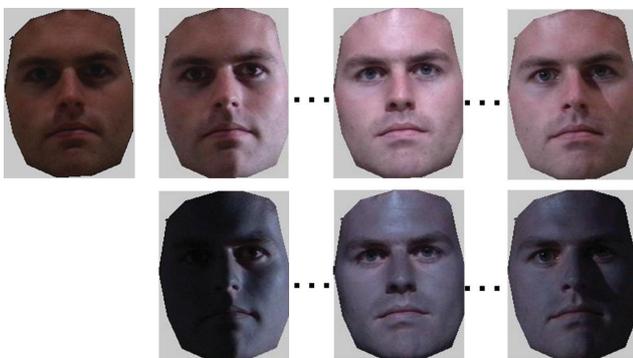


図 24 Multi-PIE の入力画像例 (左上:環境光画像, 右上:照度差画像, 下:照度差画像と環境光画像の差分)

を示す。ここで各 Step の反復回数は 100 人全員に対して共通である。

100 人全員に対して GBR 不定性の影響が多少あるものの良好な顔形状を得られている。図 22 に推定形状例として 10 人の横顔を示す。10 人全員に対して、全体的に良好な推定結果が得られていることが分かる。図 23 は横軸は処理時間、縦軸は 100 人の残差平均を表す。SVD による初期値で 2/3, SVDMD による初期値を用いることで 1/3 以下の時間で残差が収束していることが分かる。

The CMU Multi-PIE Face Database[11]

The CMU Multi-PIE Face Database は 19 照明のうち 1 つが

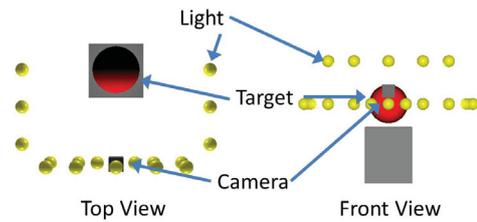


図 25 Multi-PIE の撮影環境 (文献 [11] を参考に図示したもの)

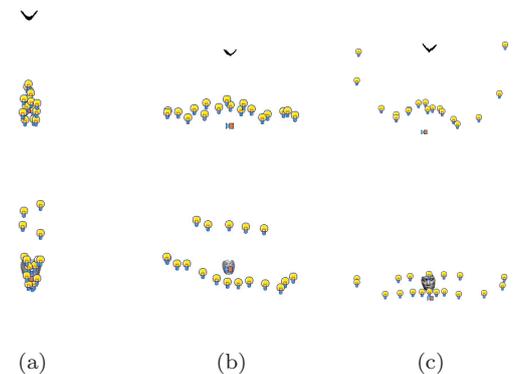


図 26 Multi-PIE の光源配置 ((a):SVD 後, (b):SVDMD 後, (c):SVDMD+FR 後)

表 6 各 Step の反復回数 (Multi-PIE 共通)

	Step1	Step2	Step3
FR	200	200	200
SVD+FR	-	100	100
SVDMD+FR	-	100	100

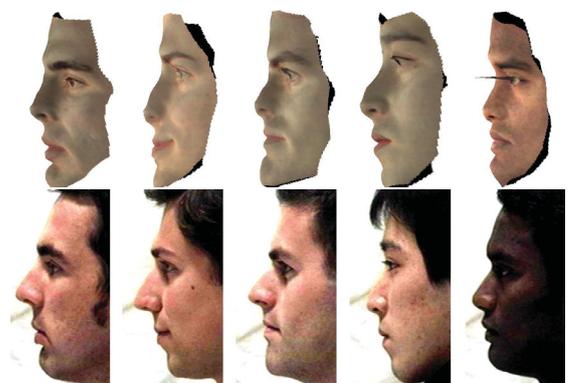


図 27 Multi-PIE の推定形状 (上段) と比較画像 (下段)

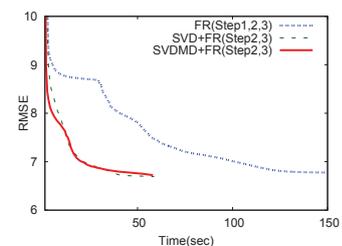


図 28 Multi-PIE 5 人 平均

環境光画像 (図 24 左上) であり、他の 18 照明はそこに単一照明を追加した画像 (図 24 右上) となっている。文献 [11] による

表 7 各 Step の反復回数
対象:木目の器

対象:木目の器			対象:備前焼		
Step1	Step2	Step3	Step1	Step2	Step3
-	50	50	-	100	100
-	25	50	-	50	100
-	-	50	-	-	100

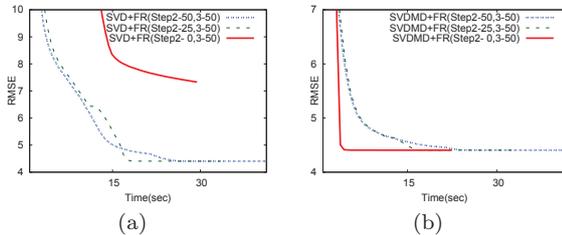


図 29 各 Step の反復回数の変化による処理時間と残差の比較
対象: 木目の器 ((a):SVD+FR, (b):SVDMD+FR)

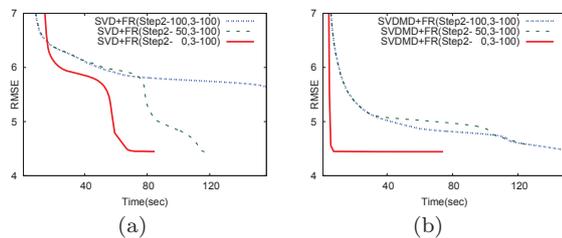


図 30 各 Step の反復回数の変化による処理時間と残差の比較
対象: 備前焼 ((a):SVD+FR, (b):SVDMD+FR)

と, 18 照明は図 25 のように, コの字形に配置された 13 照明と上段に 5 照明が配置されている. 式 (4) のモデルは環境光を扱えないが, 18 照明の画像から環境光画像の輝度値を引くことで今回の手法を適用することができる (図 24 下). ここで, 原画像に飽和を起こした画素がある場合や減算後に負となる画素がある場合, 処理の対象から外す. 表 6 に各方法における FR 法の各 Step の反復回数を示す. 249 人の実験を行い, 比較的良好な結果を得られた 5 人の結果を示す.

図 27 は正面から撮影した画像列から復元した形状を横から見たものと, 真横のカメラから撮影した画像を比較したものである. 一部は鼻のつぶれたような形となっており GBR 変換の影響を受けているように考えられるが 5 人全員, 良好な推定結果が得られている. 推定光源を図 26 に示す. 顔の横の光源が良好に推定できている. 図 28 に 5 人の処理時間と残差の平均を示す, SVD, SVDMD による初期値を用いることで 1/3 程度の時間で残差が収束していることが分かる.

各 Step の反復回数による残差の変化

実物体実験で用いた画像について, 各 Step の反復回数を変えて実験を行った. 各 Step の反復回数を表 7 に示す. 図 29 は木目の器を対象とした実験であり, SVD による初期値推定後に Step3 のみ行うと, 残差が極めて大きな残差で停滞するが, SVDMD による初期値推定後に Step3 のみ行うと, 残差が最も速く収束することが分かる.

図 30 は備前焼を対象とした実験であり, SVD や SVDMD による初期値推定後に Step3 のみ行うことで, 残差が比較的速く収束する. 特に SVDMD による初期値推定を用いると, 非常に速く収束することが分かる. これらの結果から, SVDMD によって精度のよい初期値を推定できた場合, μ を非常に小さい値から始めることで高速な推定が可能なが分かった.

5. まとめ

本稿では, 物体形状・反射特性・光源の同時推定法の初期値やパラメータ調整による高速化について述べた. 多様な画像列を用いた実験で初期値を用いることによる処理の高速化と良好な推定結果を確認し, 最大 1/10 程度の時間で推定可能なことを示した. The CMU Multi-PIE Face Database を対象とする実験も行い, 提案法の安定性を確認した.

本研究の一部は文部科学省科学研究費若手研究 (B 22700181) の助成による.

参考文献

- [1] 右田剛史, 樋谷洋介, 尺長健, “物体形状・反射特性・光源の同時推定の実装法と評価,” 情報処理学会研究報告, CVIM-181-19, 2012.
- [2] T. Migita, S. Ogino, and T. Shakunaga, “Direct bundle estimation for recovery of shape, reflectance property and light position,” Proc. European Conference on Computer Vision ECCV2008, Vol.5304, pp.412-425, 2008.
- [3] 澤見幸枝, 右田剛史, 尺長健, “形状・反射特性・光源位置の同時推定に基づく顔形状の復元,” 電子情報通信学会研究報告, PRMU2009-265, vol.109, no.470, pp.187-192, 2009.
- [4] D.W. Marquardt, “An algorithm for least-squares estimation of nonlinear parameters,” Journal of the Society for Industrial and Applied Mathematics, vol.11, no.2, pp.431-441, 1963.
- [5] H. Hayakawa, “Photometric stereo under a light source with arbitrary motion,” Journal of the Optical Society of America A: Optics, Image Science, and Vision, vol.11, no.11, pp.3079-3089, 1994.
- [6] H.Y. Shum, K. Ikeuchi, and R. Reddy, “Principal component analysis with missing data and its application to polyhedral object modeling,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.17, no.9, pp.854-867, 1995.
- [7] A. Yuille, D. Snow, R. Epstein, and P. Belhumeur, “Determining generative models of objects under varying illumination: Shape and albedo from multiple images using svd and integrability,” International Journal of Computer Vision, vol.35, no.3, pp.203-222, 1999.
- [8] 岡部孝弘, 佐藤洋一, “近接点光源は未較正照度差ステレオにおける形状復元の不定性を解決するか?,” 情報処理学会研究報告, CVIM-157-21, 2007.
- [9] 祖川和弘, 右田剛史, 尺長健, “光源色の補空間を用いた物体形状・反射特性・光源の同時推定,” 情報処理学会研究報告, CVIM-176-23, 2011.
- [10] 金田康正編, 並列化数値処理 -高速化と性能向上のために-, コロナ社, 2010.
- [11] R. Gross, I. Matthews, J. Cohn, T. Kanade, and S. Baker, “Multi-pie,” Proceedings of the Eighth IEEE International Conference on Automatic Face and Gesture Recognition, Vol.25 No.5, 2008.