

疎な画像列からの光源状況と反射特性の推定

西野 恒[†] 池内 克史^{††} 張 正友^{†††}

本研究では、少ない枚数の画像列と物体の形状モデルを入力とし、光源状況と物体の反射特性を推定することにより仮想物体画像の合成を行う手法を提案する。従来の inverse rendering の手法とは異なり、光源状況、反射特性、テクスチャのすべてを未知とし、それらを推定するのが特徴である。これは光源状況を半球上の点光源集合として近似し、鏡面反射をこの半球表面上における畳み込み積分として定式化したうえで、逆畳み込み問題を解くことにより行う。この結果、非常に簡潔な表現により、任意視点、任意光源状況下における仮想物体の見えの生成を行うことが可能となる。

Light Sources and Reflectance from a Sparse Set of Images

KO NISHINO,[†] KATSUSHI IKEUCHI^{††} and ZHENGYOU ZHANG^{†††}

In this paper, we propose a new framework for view-dependent rendering from arbitrary viewpoints and relighting under novel illumination conditions of a real object, from a sparse set of images and a pre-acquired geometric model of the object. Unlike previous inverse rendering approaches, we recover all three unknowns, namely diffuse texture, parametric BRDF and lighting distribution, from the observation of real objects. This is realized by representing the illumination environment as a set of point light sources on the surface of a hemisphere and formalizing the problem of simultaneously estimating the reflectance property and true illumination distribution as a 2D blind deconvolution on the surface of the hemisphere, and then solving this deconvolution with Alternating Minimization algorithm. As a result, we are able to accomplish rendering of real objects from arbitrary viewpoints and under novel lighting conditions with a very compact representation.

1. はじめに

現実世界に存在する物体をコンピュータ上に電子的に表示する際、その見えをより現実感高く生成するための研究は、コンピュータビジョンとコンピュータグラフィックスの両研究分野にまたがって現在までさかに行われてきている。コンピュータ上の電子的な仮想空間における物体の見えを、まるで現実の物体の写真であるかのように表示することが目的であるため、近年では自然な流れとして、現実物体の観察に基づき仮想物体画像を合成する手法の研究が主流となっている。これらの手法のうち代表的な共通アプローチとして、image-based rendering, 3D photography と inverse rendering の 3 つがあげられる。

Image-based rendering と 3D photography による

手法は、基本的に見えの補間をその根幹としているため、必然的にさまざまな環境下(視点, 光源状況)における実物体の見えの変化を非常に密に観察したものを入力とする必要がある。その結果、典型的には 100 枚オーダーから 1,000 枚オーダー以上の入力画像を扱う必要があり、しかも、それらのサンプル点から大きく逸脱した視点や光源状況下の画像合成が困難であるという問題点を持つ^{1),3),4),8),9),12),13),24)}。

一方、inverse rendering と総称される手法は、新しい視点や光源状況下における画像を合成するための情報を入力画像列から推定するというアプローチをとる。任意視点、任意光源状況下における仮想物体画像の合成には、テクスチャ、双方向反射分布関数(BRDF)、光源状況、対象シーンの幾何形状を知る必要がある。これらのうち、幾何形状は既知とし、残りの 3 つの要素のうち 1 つか 2 つを未知として推定する手法が現在まで数多く提案されている^{2),11),23),28)}。これらの多くは、光源状況を既知とし、反射特性の推定に重点を

[†] コロンビア大学コンピュータサイエンス科

Department of Computer Science, Columbia University

^{††} 東京大学生産技術研究所

Institute of Industrial Science, The University of Tokyo

^{†††} マイクロソフトリサーチ, レドモンド

Microsoft Research, Redmond

光源位置および明るさ

ピクセルごとの拡散反射成分パラメータを求めれば、それはテクスチャと同等である。

おいたものである．光源推定に関しては，Marschnerらによる inverse lighting と呼ばれる手法があげられる¹⁴⁾．光源状況を物体を覆う仮定の球面上の領域に対応する基本光源領域の線形和として近似し，各基本光源領域に対応する基本画像を一連の線型方程式を解くことにより推定している．この際，反射特性に関しては明示的な推定を行わないため，拡散反射成分のみの物体に関してのみ新たな光源状況下での仮想物体画像合成が可能となっている．Satoらは，物体の落とす影領域に着目し，同様の手法を用いて光源状況を推定している²¹⁾．

最近，Ramamoorthiらにより，信号処理的アプローチにより inverse rendering を行う枠組みが提案されている¹⁹⁾．無限遠光源を仮定することにより，物体表面における光の反射を極座標系における畳み込み積分として解析し，球面調和関数表現を用いた逆畳み込みにより，光源状況既知の場合のテクスチャと BRDF の推定と，テクスチャ既知の場合の BRDF と光源状況推定を実現している．

本研究では，特に，対象物体の画像列と 3 次元モデルから任意視点，任意光源状況下における対象物体の画像を合成する手法を提案する．従来手法が少なくとも数十枚オーダーの画像列を入力とするのに対し，本提案手法はより少ない枚数の画像列（のちに示す実験では 6 枚）から任意視点/光源下の画像合成を実現する．従来の inverse rendering の手法とは異なり，光源状況，反射特性，テクスチャのすべてを未知とし，それらを推定するのが特徴である．これは，まず入力画像列において反射成分の分離を行い，光源状況を半球上の点光源集合として近似し，鏡面反射をこの半球表面上における畳み込み積分として定式化したうえで，逆畳み込み問題を解くことにより行う．Nishinoら¹⁸⁾は，本研究で提案するものと同様の枠組みを用いて，少ない枚数の画像列からの任意視点画像の合成を行っている．しかし，光源状況と反射特性の同時推定が不定な問題である点に留意しておらず，結果も安定したものがない．我々は，この同時推定問題を半球表面上の逆畳み込み問題として定式化し，安定に解く枠組みを提案する．さらに，任意視点画像ばかりではなく，任意光源状況下の画像合成も行えることを示す．

提案する手法では，光源状況と物体が固定された状態で，視点を変えながら得られた少数枚の画像列（図 1）と，あらかじめ取得した距離画像列から構築した対象物体の 3 次元モデルを入力とする．撮像に用いたカメラはあらかじめ校正用物体等を用いてそ



図 1 入力画像（6 枚中の 3 枚）
Fig. 1 Three out of six input images.

の内部パラメータは求めているものとし，さらに各画像に対応するカメラ視点，向き等の外部パラメータも structure-from-motion 等の手法を用いて事前に求めたものであるとする．

2. 反射成分分離

本研究で対象とする物体は，その物体表面における光の反射が二色性反射モデルにより近似でき，それぞれの反射成分である拡散反射成分が Lambertian 反射モデル，鏡面反射成分が Torrance-Sparrow 反射モデルにより表現できるものとする．実世界に存在する多くの物体は，塗料等が表面に塗られているため，このような反射モデルの組合せによる近似が可能である．

ここで，画像輝度（ピクセル値）は対象シーン輝度と正比例関係にあることが分かっているため¹⁰⁾，以下の議論を簡単にするため両者の間の比例定数を無視し（どちらかに埋め込んだと仮定して），ピクセル値はそのままシーン輝度値を表すものとする．このとき，二色性反射モデルにより，物体上のある 1 点における反射光 \mathbf{I} は，式 (1) のように拡散反射成分 \mathbf{I}_D と鏡面反射成分 \mathbf{I}_S の線形和で表される．

$$\mathbf{I} = \mathbf{I}_D + \mathbf{I}_S \quad (1)$$

ただし，ボールド体は RGB の 3 次元色ベクトルを表す ($\mathbf{I} = [I_R \ I_G \ I_B]^T$)．

さらに，上述のようにそれぞれの反射成分は以下に示す式 (2)，(3) のようにそれぞれ Lambertian 反射モデル，Torrance-Sparrow 反射モデルにより近似する．

$$\mathbf{I}_D = \max \left[0, \mathbf{K}_D \int_{\Omega} L_i(\theta_i, \phi_i) \cos \theta_i d\omega_i \right] \quad (2)$$

ただし， \mathbf{K}_D は物体表面自体の色と光源色により決まる 3 次元色ベクトルであり， $L_i(\theta_i, \phi_i)$ は対象となっている物体表面上の点を中心とする極座標系における光源の放射輝度， ω_i は光源の占める立体角である．

$$\mathbf{I}_S = \int_{\Omega} \frac{\mathbf{K}_S F G}{\cos \theta_r} L_i(\theta_i, \phi_i) \exp \left[-\frac{\alpha^2}{2\sigma^2} \right] d\omega_i \quad (3)$$

ただし， \mathbf{K}_S は鏡面反射光の色ベクトルであり，指数

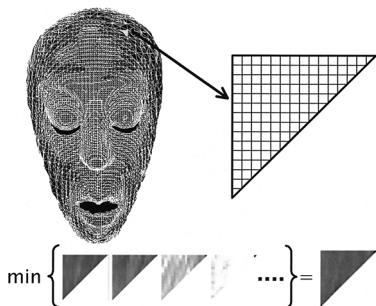


図 2 拡散反射テクスチャ
Fig. 2 Diffuse texture.

部の正規化と鏡面反射強度も含む値である。また、 F はフレネル係数、 G は幾何的減衰係数であり、 σ は物体表面の粗さを表す係数である。それぞれの角度は拡散反射光の場合と同様の極座標系であり、 θ_r は法線方向と視点方向間の角度、 α は視点方向と光源方向間の二等分角と法線方向の間の角度である。

2.1 拡散テクスチャマップ

物体表面における光の反射が Lambert 則に従う拡散反射と Torrance-Sparrow 反射モデルにより表される鏡面反射の線形和で表されると考えると、物体固定、光源状況固定で視点のみが移動している状況では、拡散反射成分はつねに一定の値をとる。したがって、物体表面上の各点における輝度値 ($|I| = \sqrt{I_R^2 + I_G^2 + I_B^2}$) の変化を各画像 (各視点) について並べ、その中で輝度値が最小値をとる色ベクトル I_{min} をもって拡散反射成分 I_D とする。

物体表面上の各点において輝度値の変化を観察し、その最小値を求めることは、物体表面が連続であるため不可能である。そのため、本手法ではこの拡散反射成分を 1 枚のテクスチャマップとして保持する。これは図 2 に示すように、物体の 3 次元モデルの各三角パッチ内に 3 次元の格子を考え、各格子点が各画像において射影される 2 次元点のピクセル値の変化を追うことにより行う。すなわち、各三角パッチについて各画像から対応する画像の領域を切り出し、その小三角画像列において各格子点について最小輝度値をとる色ベクトルを並べたものが、その三角パッチの拡散反射成分テクスチャとなる。

2.2 光源半球

前節の拡散反射テクスチャマップを各視点において 3 次元モデルにテクスチャマッピングすることにより、各視点における拡散反射成分画像を得ることができる。次に、これらの拡散反射成分画像を元の入力画像からそれぞれピクセルごとに差し引くことにより、鏡面反

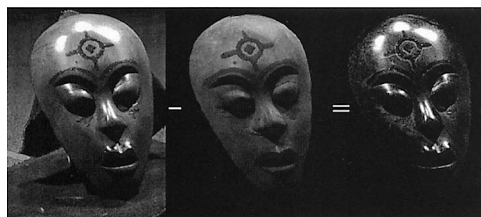


図 3 鏡面反射成分画像
Fig. 3 Specular image.

射成分画像を得ることができる。実際には図 3 の右端のような残余画像が得られ、その各ピクセル値は、鏡面反射成分、相互反射成分、拡散反射テクスチャマップを最小値フィルタにより求めた際の拡散反射成分の残余の合計となる。しかし、相互反射は本手法では扱わず、これらの残余画像列は鏡面反射成分が主であるため、残りの成分はこのあとの議論では単純にノイズとして扱うことにする。そのため、以後これらの残余画像列を鏡面反射成分画像列と呼ぶ。

次に、これらの鏡面反射成分画像から光源状況と鏡面反射特性の推定を行う。式 (3) において明らかなように、物体表面における鏡面反射は、完全鏡面反射方向を中心として (厳密には $\frac{1}{\cos \theta_r}$ の項のため、完全鏡面反射方向の少しずれた方向を中心に) 分布する。そのため、上述の方法により得た鏡面反射成分画像の各ピクセル値を、各ピクセルに対応する物体表面上の 3 次元点をもとに完全鏡面反射方向に打ち返すことにより、おおよその光源状況が推定できる。そこで、光源状況を物体を覆う半球を考え、その半球の表面上を一樣に分割した格子点を点光源とする点光源集合として近似する。点光源となりうる格子点を半球上に一樣に分布させるため、複数回分割した (後の実験では 70 回、約 25,000 格子点) 正 20 面体を用いる。

各鏡面反射成分画像の各ピクセルにつき、(1) 対応する 3 次元点を求め、(2) その法線ベクトルを用いて完全鏡面反射ベクトルを求め、(3) その完全鏡面反射ベクトルと光源半球の交点の最近傍格子点にそのピクセル値を割り振る作業を全ピクセルについて行くと、図 4 の左 2 つの図に示すように、各視点における光源半球が求まる。これらの各視点に対応する光源半球は、各視点における物体表面の放射輝度に関与しない光源も存在するため、本来の光源状況を完全に表現しているとは限らない。そのため、これらの各視点に対応する光源半球 (以後、部分光源半球) を統合した光

ただし、図は光源半球を真上から真下に見下ろした状態を表示している。

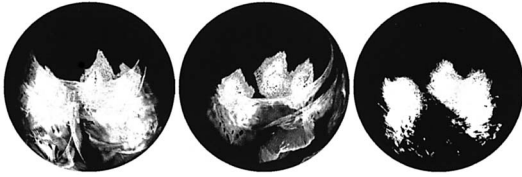


図4 左, 中: 部分光源半球, 右: 初期光源半球
Fig.4 Left, Middle: Partial illumination hemispheres.
Right: Initial illumination hemisphere.

源半球を構築する。さらに、各部分光源半球にはわずかな相互反射やノイズの影響がそのまま打ち返されて割り付けられているため、これらの影響も考慮する必要がある。このような相互反射やノイズを光源としてとらえた場合、静止せずに移動する光源と考えられる。そのため、各部分光源半球の各格子点において、ピクセル値が割り付けられたかどうかの論理積をとれば、これらの影響を除外することができる。そこで(全画像数 n) 回各部分光源半球においてピクセル値の割り付けられたものを有効な点光源として残すという戦略により光源半球を構築する。各画像においてすべての光源が必ず物体表面反射に寄与している場合には、上記の閾値は全画像数回でよいが($n = 0$)、光源の配置によっては特定の画像において物体の見えに寄与しない光源が存在することがある。このような場合も考慮し、上記の閾値を多少緩めることにより($n = 2$ 等)、すべての光源が含まれる光源半球を構築する。その結果、図4の右の図のように、おおまかに光源状況をとらえた点光源集合を得ることができる。

3. 光源状況と鏡面反射特性の同時推定

前章において構築した光源半球は、物体表面が完全鏡面であるという仮定に基づいて得られた光源状況を表している。実際には式(3)にあるように、物体表面における鏡面反射成分は粗さ係数 σ で表現される正規分布に従い、完全鏡面反射方向を中心に分布している。そこで、前章で得た光源半球を、光源状況の初期推定値とし、この初期光源半球と鏡面反射成分画像列を用いて、鏡面反射成分のパラメータと真の光源状況を光源半球として推定する。

3.1 定式化

Torrance-Sparrow 反射モデルでその鏡面反射成分が近似できる現実物体の多くにおいて、幾何減衰定数 G は 1.0 であり、フレネル反射係数 F は一定であることが分かっている²⁵⁾。そのため、物体表面上のある 1 点 v における鏡面反射成分は、式(3)を簡略化し、

$$\mathbf{I}_S(v) = \frac{\mathbf{K}_{S,v}}{\cos \theta_r} \int_{\Omega} L_i(\theta_i, \phi_i) \exp \left[-\frac{\alpha^2}{2\sigma^2} \right] d\omega_i \quad (4)$$

で与えられる。さらに、光源状況を離散化された点光源集合として表現するため、この光源半球の各点光源を極座標系で表すと、この光源半球下における鏡面反射成分は、

$$\mathbf{I}_S(v) = \frac{\mathbf{K}_{S,v}}{\cos \theta_r} \sum_l^{N_L} \omega_l L_l(\theta_l, \phi_l) \exp \left[-\frac{\alpha^2}{2\sigma^2} \right] \quad (5)$$

となる。ただし、 $L_l(\theta_l, \phi_l)$ は光源半球上の座標 (θ_l, ϕ_l) における点光源 l の放射輝度である。

ここで、 N_L 個の格子点が存在する光源半球では $\omega_l = \frac{2\pi}{N_L}$ となる。さらに、鏡面反射色と光源色は同一の色ベクトルとなるため、対象シーンにおける光源はすべて同一の色ベクトル \mathbf{L} であるとし、光源半球のすべての点光源の正規化された平均色ベクトルを \mathbf{L} とする。この仮定により、後に示すように異なる光源色下での画像合成を行えなくなるが、本研究では光源の放射輝度の空間的分布を推定することに集中する。また、対象物体表面上の鏡面反射特性は均一であると仮定すると、ある 1 点における鏡面反射成分は、

$$\mathbf{I}_S(v) = I_S(v) \mathbf{L} \quad (6)$$

$$I_S(v) = \frac{2\pi}{N_L} \frac{K_S}{\cos \theta_r} \sum_l^{N_L} L_l \exp \left[-\frac{\alpha^2}{2\sigma^2} \right] \quad (7)$$

と表されることになる。

この結果、式(7)のすべての変数がスカラー量となり、光源状況と鏡面反射特性の同時推定は、次式に示すように、式(7)と光源半球を用いた最小自乗問題として定式化できる。

$$\arg \min_{\sigma, K_S, L_l(l=0, \dots, N_L)} \sum_{k=0}^{N_K} \sum_{s,t=0,0}^{N_S, N_T} |I_{(s,t,k)} - I_{S,(s,t,k)}|^2 \quad (8)$$

ただし、 $I_{(s,t,k)}$ は k 番目の鏡面反射成分画像の座標値 (s, t) におけるピクセルの輝度値であり、 (N_S, N_T) は鏡面反射成分画像の幅と高さである。各鏡面反射成分画像に対応するカメラの視点は分かっているため、 (s, t, k) より対応する物体表面上の 3 次元点 v を求めることができ、式(7)により $I_{S,(s,t,k)}$ を計算できる。

3.2 畳み込み積分としての鏡面反射

前節に述べたように、式(8)の最小自乗問題を鏡面反射パラメータ (K_S, σ) と光源半球の各点光源の放射輝度 L_l について解けばよいが、実はこの最小自乗問

題は安定に解くことができない．これは鏡面反射が畳み込み積分になるためであるが，なぜ畳み込み積分になるのか，式 (7) に陰に含まれている幾何変換を追いながら見てゆく．

まず，対象物体の 3 次元モデルを持っているため，鏡面反射成分画像内の 1 つのピクセルに対応する物体表面上の 3 次元点を求めることができる．ここで，ある関数 M が鏡面反射成分画像の画像番号と 2 次元座標値をとり，対応する 3 次元点を返すとする．すなわち，

$$M : (s, t, k) \mapsto (x, y, z). \quad (9)$$

また，ある 3 次元点の法線ベクトルを返す関数を N とする．

$$N : (x, y, z) \mapsto (n_x, n_y, n_z). \quad (10)$$

光源状況は，光源半球上の点光源集合として表し，極座標系 (θ, ϕ, r) を用いる．本研究では光源半球の半径はあらかじめ与えることとし (R) ，光源半球上の光源は 2 次元座標系 (θ, ϕ) により表す．

上記の関数群を用い，ある鏡面反射成分画像のある座標値 (s, t, k) から，その点に対応する 3 次元点の法線ベクトルの光源半球表面との交点を求める関数 C を考える．

$$C : (s, t, k) \mapsto (\theta_C, \phi_C) \quad (11)$$

$$\begin{aligned} C(s, t, k) &= M(s, t, k) + \zeta \cdot N(M(s, t, k)) \\ &= (x + \zeta n_x, y + \zeta n_y, z + \zeta n_z) \\ &= (\theta_C, \phi_C, R) \end{aligned}$$

ただし， ζ は $|C(x, y, z)| = R$ を満たすスカラー値とする．同様に，鏡面反射成分画像のある座標値とともに，光源半球上の 1 つの点光源の座標値 (θ, ϕ, R) を与えると，視点方向 (v_x^k, v_y^k, v_z^k) と光源方向の間の 2 等分ベクトルと光源半球の交点を求める関数を考える．

$$BS : (s, t, k, \theta, \phi) \mapsto (\theta_{BS}, \phi_{BS}) \quad (12)$$

以上の関数を用いると，各鏡面反射成分画像におけるピクセルの輝度は簡略化された Torrance-Sparrow 反射モデルを用いて以下のように求められる．

$$\begin{aligned} I_S(s, t, k) &= \frac{K_s \omega}{\cos \theta_r} \\ &\cdot \sum_l^{N_L} L_l(\theta_l, \phi_l) g_{\frac{1}{2\sigma^2}}(|BS(s, t, k, \theta_l, \phi_l) - C(s, t, k)|) \end{aligned} \quad (13)$$

ただし， $\omega = \frac{2\pi}{N_L}$ かつ， g_γ はガウス分布関数である．

$$g_\gamma(\phi) = \exp[-\gamma\phi^2].$$

式 (13) から，鏡面反射は光源半球の表面上における 2 次元畳み込み積分としてとらえられる．すなわち，鏡面反射成分画像上のある 1 つの 2 次元点 (s, t, k) よ

り，対応する光源半球表面上の 1 点 (θ_C, ϕ_C) が定まり，その点を中心とするガウス分布フィルタに基づいて近傍の点光源の放射輝度が積分される．

3.3 逆畳み込みによるパラメータ推定

ここで，以後の議論において用いる記号を参考とする文献と同じにすると，元の画像（真の光源半球表面） $u(x, y)$ が畳み込みフィルタ（ガウシアンフィルタ） $h(x, y)$ により畳み込み積分されたものが観察される画像（鏡面反射成分画像） $z(x, y)$ となる．したがって，鏡面反射成分画像から光源状況と鏡面反射パラメータを推定する問題は，次式を最もよく説明する u, h を求めることにほかならない．

$$z = h * u + \eta \quad (14)$$

ただし， η はノイズであり，我々の場合には，拡散反射テクスチャマップを求めた際のノイズや相互反射，光学系によるノイズがこれに含まれる．

上記の問題は，隠れ逆畳み込み問題（blind deconvolution）と呼ばれ，画像処理や信号処理の分野で深く研究されている．この隠れ逆畳み込み問題の困難な点は，式 (14) が元の画像 u （真の光源状況）と畳み込みフィルタ h （ガウシアンフィルタ）の両者に関して不定（ill-posed）である点である．そこで，何かしらの制約を加えることにより，この逆畳み込みを安定に解く手法がいくつか提案されている．特に最近，Youらにより²⁷⁾，各ソース u, h のそれぞれに正規化制約を加えることにより逆畳み込みを解く手法が提案されている．すなわち，

$$\begin{aligned} \min_{u, h} f(u, h) &= \min_{u, h} \|h * u - z\|_{L^2(\Omega)}^2 \\ &\quad + \alpha_1 TV(u) + \alpha_2 TV(h) \end{aligned} \quad (15)$$

の最小化問題を解くことにより逆畳み込みを行う．ただし， Ω は u, h の領域であり， TV 正規化項は以下のように定義される⁵⁾．

$$TV(u) \equiv \int_{\Omega} |\nabla u| dx dy \quad (16)$$

また， α_1 と α_2 は正規化具合とパラメータのフィット具合を調整する正の重み付けスカラー値である．

式 (15) の最小化問題を解くわけだが，目的関数 $f(u, h)$ は 2 変数関数として凸になっていない．しかし， u と h の片方を固定し定数と扱えば，もう一方の変数に関してこの目的関数は凸となる．この事実に基づき，Youらは，片方の変数を固定し，もう片方の変数に対して目的関数の最小化を行うことを，固定する変数を交換しながら繰り返し実行する AM (Alternating Minimization) アルゴリズムを提案した．この AM

アルゴリズムには、固定する変数の順番により、以下にあげる2つの派生法が存在し、いずれも収束が証明されている⁶⁾。

AMHU

u^0 を初期値とし：

1. $\nabla_h f(u, h^{k-1}) = 0$ の解 h^k を求める。
2. $\nabla_u f(u, h^{k-1}) = 0$ の解 u^k を求める。
3. 収束判定．収束していなければ1に戻る。

AMUH

h^0 を初期値とし：

1. $\nabla_u f(u, h^{k-1}) = 0$ の解 u^{k-1} を求める。
2. $\nabla_h f(u^{k-1}, h) = 0$ の解 h^k を求める。
3. 収束判定．収束していなければ1に戻る。

そこで、光源半球と鏡面反射成分画像を用いた真の光源半球と鏡面反射パラメータの同時推定をTV正規化を用いたAMアルゴリズムによって求めることを考える。目的関数を考えた場合、畳み込みフィルタに関しては1次元のガウス関数であり、その変数に関して滑らかであるため、正規化項を考慮する必要がない。また、式(13)を見ると明らかのように、 K_S と各点光源の放射輝度 L を分離することは不可能である。したがって、 $K_S = 1.0$ とし、光源状況を K_S によって定数スケールされたものとして求めることにする。以上より、光源状況に対する正規化項を加えた解くべき最小化問題は以下ようになる。

$$\begin{aligned} \min_{L, \sigma} f(L, \sigma) \equiv & \\ \min_{L, \sigma} \sum_k^{N_K} \sum_l^{N_L} & \|I_S(s, t, k) - I(s, t, k)\|_{L_2}^2 \\ & + \rho \sum_l^{N_L} |\nabla L(\theta_l, \phi_l)| \end{aligned} \quad (17)$$

ここで、 ρ は正規化項に対する重み付けパラメータであり、光源状況半球上における局所分布における変化の滑らかさを規定する値となる。後の実験では非常に小さい値($\rho = 1.0e-6$)を用い、なるべく強い拘束を与えないように配慮した。これを上記のAMアルゴ

リズムで解くわけであるが、元のアルゴリズムと異なり、以下の2つの点に特に留意する。

M推定量を用いた共役勾配法

鏡面反射成分画像には、拡散反射テクスチャマップを作成した際の誤差、本手法では扱わない相互反射成分、そして画像光学系によるノイズが鏡面反射成分に加重されている。そのため、これらをノイズとして扱い式(17)を最小化する必要がある。そこで、M推定量によりこれらを外乱として扱いAMアルゴリズムにおける各最小化を行う。また、効率的に最小化を行うため、共役勾配法を用いる。

初期推定値

文献5)にも実験データとともに述べられているように、AMアルゴリズムによってエッジのない滑らかな畳み込みフィルタを推定する場合、レンズボケ等による畳み込みフィルタを推定する場合に比べ、収束が非常に遅いことが分かっている。収束を早める手段として良い初期値を与えることが有効である⁵⁾。そこで、Torrance-Sparrow反射モデルにより近似することのできる鏡面反射を行う物体に関しては、 σ の値が大体0.1のオーダーになることが経験的に分かっているので、 $\sigma_0 = 0.1$ として、AMアルゴリズムをAMUHの順に実行する。

ただし、面光源等の拡散光源に対する反射パラメータと光源状況の同時推定は理論的に不可能であることが分かっている¹⁹⁾。そのため、後述の実験では点光源を用いた光源状況下で撮像された画像列を入力とする。

4. 仮想物体画像の合成

任意視点画像の合成

推定された光源半球、鏡面反射パラメータ、拡散反射テクスチャマップを用いて任意視点における仮想物体画像を合成できる。これは、与えられた視点に対して以下の手順の作業を行うことにより実現する。1) 2.1節において作成した拡散反射テクスチャマップを3次元モデルにテクスチャマッピングすることにより拡散反射成分画像を合成し、2) shadow z-buffer法²⁶⁾を用いて、物体表面の各点についてどの光源が見えているかを表す影マップを作成し、3) この影マップと推定された光源半球、鏡面反射パラメータを用いて式(7)と式(6)を用いて、鏡面反射成分画像を合成し、4) 最終的に、拡散反射成分画像と鏡面反射成分画像を足し合わせる。

任意光源下画像の合成

推定された結果を用いて、任意視点画像に加え任意光源状況下における画像も合成できる。

実際には、画像処理におけるぼけ除去等の逆畳み込み問題を解く際には、この ρ の値を順次変化させ、視覚的に最も良い結果を得られる値を用いるというアプローチがとられている。本論文における実験では、安定に収束させるために正規化項を導入しているため、アルゴリズムが収束する範囲内でできるだけ小さい値を用いることを目的とし、上記の値を用いた。

ここで、元の光源状況の放射輝度分布を光源半球上の点光源集合 $\{L_l : l = 0, \dots, N_L\}$ とする。ただし、3.3 に触れたように、これらは真の放射輝度が鏡面反射強度によりスケールされたものである ($K_S L_l^{true}$)。しかし、以下の議論では、結局この K_S は消去されるので影響はない。この元の光源半球下において、物体表面上のある点 v における拡散反射成分の色ベクトルは、

$$\mathbf{I}_{D,v} = \mathbf{K}_{D,v} \sum_l^{N_L} M_{l,v} L_l \cos \theta_{l,v} \quad (18)$$

となる。ただし、 $\mathbf{K}_{D,v}$ は拡散反射色であり、 $M_{l,v}$ は前節に説明した 0 か 1 を値としてとる影マップである。

さて、新しい光源状況を光源半球上の点光源集合 $\{\tilde{L}_l : l = 0, \dots, \tilde{N}_L\}$ として与えられたとする。このときの拡散反射成分の放射輝度は、

$$\widetilde{\mathbf{I}}_{D,v} = \mathbf{K}_{D,v} \sum_l^{\tilde{N}_L} \widetilde{M}_{l,v} \tilde{L}_l \cos \theta_{l,v}. \quad (19)$$

となる。このとき、式 (18) と式 (19) の比をとると、

$$\frac{\mathbf{I}_{D,v}}{\widetilde{\mathbf{I}}_{D,v}} = \frac{\sum_l^{N_L} M_{l,v} L_l \cos \theta_{l,v}}{\sum_l^{\tilde{N}_L} \widetilde{M}_{l,v} \tilde{L}_l \cos \theta_{l,v}}. \quad (20)$$

となる。ここで、元の光源半球における各点の照度 (右辺分子)、新しい光源半球における各点の照度 (右辺分母) は計算でき、かつ、元の光源半球における各点の放射輝度 (左辺分子) は拡散反射成分テクスチャマップをテクスチャマッピングすることにより得られるので、 \mathbf{K}_D を知らなくても、新しい光源半球における拡散反射成分の放射輝度を計算できる。

したがって、任意視点において以下の手順に従い、任意光源状況下における仮想物体画像を合成できる。

1) 前節と同様に影マップ M と \tilde{M} を求め、2) 拡散反射テクスチャマップをテクスチャマッピングすることにより元の拡散反射成分画像を合成し、3) この拡散反射成分画像のピクセル値を対応する各 3 次元点における照度 $\sum_l^{N_L} M_{l,v} L_l \cos \theta_{l,v}$ で割ることにより、 $\frac{1}{K_S} \mathbf{K}_D$ の分布である、擬似アルビド画像を合成し、4) 新しい光源半球をもとに、各 3 次元点における照度 $\sum_l^{\tilde{N}_L} \widetilde{M}_{l,v} \tilde{L}_l \cos \theta_{l,v}$ を求めたうえで、これを擬似アルビド画像の各 2 次元座標に対応する 3 次元点について掛けることにより、新しい光源状況下での拡散反射成分画像を合成し、5) 与えられた視点のもとで、前節と同様に新しい光源半球を元に鏡面反射成分画像を

合成し、これを上述の新拡散反射成分画像に加算する。

5. 実験

図 1 を入力画像列とし、本手法を用い光源状況と反射パラメータの解析、推定を行った。各視点においてシャッタースピードを変えつつ複数枚の画像を取得し、それらの画像を組み合わせて^{7),16)} 各ピクセル値が 3 バンドの浮動小数点で表される HDR 画像列を生成した。図 1 はシャッタースピードが約 $\frac{1}{30}$ のときの画像を表示している。

また、各視点において明るさ画像と同時に、光線パターンを投影することにより距離画像を取得した²²⁾。これとは別に Minolta VIVID 900¹⁵⁾ を用いてあらかじめ取得しておいた距離画像列から位置合わせ¹⁷⁾ と統合²⁰⁾ により図 2 の左上にある 3 次元モデルを生成した。視点の推定は、この 3 次元モデルに対し、前述の各視点における距離画像を位置合わせ¹⁷⁾ することにより行った。

5.1 推定結果

2.1 節に述べた手順により得られた拡散反射テクスチャマップを用いて図 3 の真中の図のような拡散反射成分画像を合成でき、これらを各視点について作成したうえで各入力画像から引き、図 3 の右の図のような鏡面反射成分画像列を得られる。それらの鏡面反射成分画像列から得られた部分光源半球を統合した初期光源半球 (図 4 右) をもとに、真の光源半球と鏡面反射成分パラメータの推定を行った結果、図 5 の左に示す結果を得た。図 5 の右は、実験に用いた実光源環境を魚眼レンズを通して撮像した写真であるが、推定された光源状況はこれに合致していることが分かる。この際、光源半球上の最も明るい点光源は 1370.83 であり、最も暗い点光源は 0.65 という値であった。これらは 3.3 節に述べたように、光源の放射輝度を K_S でスケールした値である。また、このとき推定された鏡面反射パラメータ σ は 0.075 であった。

入力画像の撮影の際に用いた光源は、図 5 右に示すように、内側に金属塗料の施された電球であったため、厳密には点光源ではなく、一定の面積を持った光源である。図 5 左に見られるように、本手法による推定結果はこの面積に対応した広がりを持った点光源集合として光源を表している。ただし、解を安定に求めるために光源状況に対して正規化項を加えているため、実光源では一定の領域外では明るさがゼロになるものが、推定結果はなだらかに減衰する分布として求まる。理論的には面光源等の低周波成分しか存在しない拡散光源状況下での反射パラメータと光源状況の同

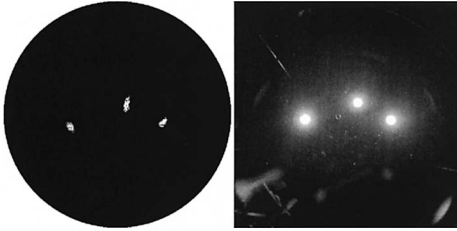


図 5 左：推定された光源半球，右：魚眼レンズを用いて撮像された実光源環境

Fig. 5 Left: Estimated illumination hemisphere, Right: Ground truth lighting environment captured with fish-eye lens.

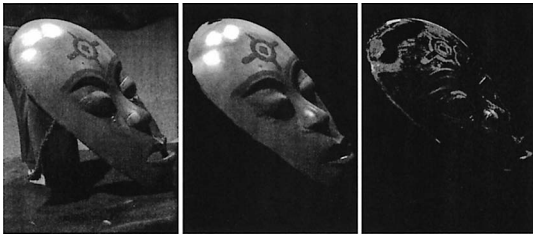


図 6 左：実画像，中：仮想物体画像，右：差分画像

Fig. 6 Left: Real photograph, Middle: Synthetic image, Right: Difference image.

時推定は不定性をともなうが¹⁹⁾，本手法では光源半球表面上での逆置み込みとして定式化しているため，与えられた入力画像列の鏡面反射成分の変化を十分に説明するだけの推定結果を得ることができることが分かる．完全に拡散光源である場合は当然非常に不安定な推定結果が求まることになる．しかし，ある程度高周波成分の存在する光源状況である限り，少なくとも入力画像列と同じ視点における画像列を再合成結果が互いに等しくなるような各パラメータが求まるはずであり，入力画像列を増やせば増やすほどより真値に近づくものと考えられる．この点に関しては，今後解析を行いたい．

5.2 任意視点画像

図 6 に一例を示す．比較対象となっている実画像は，入力画像としては用いていないものである．図 6 より，ほぼ実画像と相違ない見えを生成できることが分かる．頭頂部のテクスチャの欠落や鏡面反射成分の微小な形の違いは，主に 3 次元モデルの精度によるものであり，誤差は形状の誤差に沿って一貫したものと現れるため，視覚的には現実感を損なわない．

5.3 任意光源状況下画像

推定された光源半球には図 5 に示すように，3 つの点光源に対応する点光源集合がクラスタとして現れている．これらの各実点光源に対応する点光源集合部分



図 7 左：実画像，中：仮想物体画像，右：差分画像

Fig. 7 Left: Real photograph, Middle: Synthetic image, Right: Difference image.

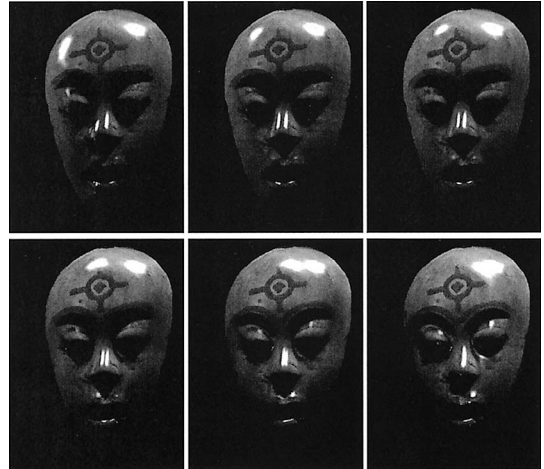


図 8 動的な光源環境下の仮想物体画像列

Fig. 8 Synthetic images rendered under a dynamic lighting condition.

をそれぞれマスキングすることにより，図 7 に一例を示すように，各実点光源を別個に点灯した画像を合成できる．

さらに，任意の光源状況下の画像合成の一例として，移動する光源状況下での画像列合成を行った．図 8 に，2 つの点光源の片方が移動する光源環境下での物体の見えの変化を示したムービーからの数フレームを示す．この例より，鏡面反射成分だけでなく，拡散反射成分も正しく relighting できることが分かる（面の右半面の明るさの変化等）．

6. おわりに

本論文では，固定された光源状況下で，固定された物体の見えの変化を疎に観察した画像列と，対象物体の幾何形状モデルを入力とし，任意視点，任意光源状況下における仮想物体画像の合成手法を提案した．従来の手法と大きく異なる特徴として，反射特性だけでなく光源状況も未知とし，入力画像列から反射特性は拡散反射テクスチャマップと鏡面反射モデルのパラメータとして，光源状況は半球上の点光源分布として

推定する。これは、はじめに反射成分を分離することにより拡散反射成分を1枚のテクスチャマップとして求め、次に鏡面反射成分画像列から光源状況と鏡面反射特性の同時推定を行うことにより実現される。この同時推定は、鏡面反射を半球の表面における2次元畳み込み積分としてとらえ、逆畳み込みを行うことにより実現されている。我々の知る限り、forward renderingに必要な情報を幾何形状をのぞいてすべて未知とし、それらをすべて安定に推定する初めての inverse rendering 手法である。

本論文では、この手法の理論とともに、実物体による実験結果を示し、その有効性を示した。

参 考 文 献

- 1) Adelson, E.H and Bergen, J.R.: The Plenoptic Function and the Elements of Early Vision, *Computational Models of Visual Processing*, Landy, M. and Movshon, J.A.(Eds), pp.3-20, MIT Press (1991).
- 2) Boivin, S. and Gagalowicz, A.: Image-based rendering of diffuse, specular and glossy surfaces from a single image, *Computer Graphics Proceedings, ACM SIGGRAPH 01*, pp.197-116 (Aug. 2001).
- 3) Buehler, C., Bosse, M., McMillan, L., Gortler, S. and Cohen, M.: Unstructured Lumigraph Rendering, *Computer Graphics Proceedings, ACM SIGGRAPH 01*, pp.425-432 (Aug. 2001).
- 4) Chai, J., Kang, S.B. and Shum, H.: Rendering with Non-uniform Concentric Mosaics and Invariable Input Sampling, *Proc. ECCV 2000 Workshop SMILE 2000*, pp.94-108 (Jul. 2000).
- 5) Chan, T. and Wong, C.: Total Variation Blind Deconvolution, *IEEE Trans. Image Processing*, Vol.7, pp.370-375 (Mar. 1998).
- 6) Chan, T.F. and Wong, C.K.: Convergence of the Alternating Minimization Algorithm for Blind Deconvolution, *Linear Algebra and its Applications*, Vol.316, No.1-3, pp.259-285 (2000).
- 7) Debevec, P.E. and Malik, J.: Recovering High Dynamic Range Radiance Maps from Photographs, *Computer Graphics Proceedings, ACM SIGGRAPH 97*, pp.369-378 (Aug. 1997).
- 8) Gortler, S.J., Grzeszczuk, R., Szeliski, R. and Cohen, M.F.: The Lumigraph, *Computer Graphics Proceedings, ACM SIGGRAPH 96*, pp.43-54 (Aug. 1996).
- 9) Heigl, B., Koch, R., Pollefeys, M., Denzler, J. and Van Gool, L.: Plenoptic modeling and rendering from image sequences taken by handheld camera, *Proc. DAGM '99*, pp.94-101 (1999).
- 10) Horn, B.K.P.: *Robot Vision*, McGraw-Hill Book Company (1986).
- 11) Ikeuchi, K. and Sato, K.: Determining Reflectance Properties of an Object Using Range and Brightness Images, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.13, No.11, pp.1139-1153 (Nov. 1991).
- 12) Isaksen, A., McMillan, L. and Gortler, S.: Dynamically Reparameterized Light Fields, *Computer Graphics Proceedings, ACM SIGGRAPH 00*, pp.297-306 (Jul. 2000).
- 13) Levoy, M. and Hanrahan, P.: Light Field Rendering, *Computer Graphics Proceedings, ACM SIGGRAPH 96*, pp.31-42 (Aug. 1996).
- 14) Marschner, S.R. and Greenberg, D.P.: Inverse Lighting for Photography, *Proc. IS&T/SID Fifth Color Imaging Conference*, pp.262-265 (Nov. 1997).
- 15) Minolta. *Vivid 900* (2001).
<http://www.minolta-rio.com/vivid/>
- 16) Mitsunaga, T. and Nayar, S.K.: Radiometric Self Calibration, *Proc. Computer Vision and Pattern Recognition '99*, pp.374-380 (Jun. 1999).
- 17) Nishino, K. and Ikeuchi, K.: Robust Simultaneous Registration of Multiple Range Images, *Proc. 5th Asian Conference on Computer Vision, ACCV 02*, pp.454-461 (Jan. 2002).
- 18) Nishino, K., Zhang, Z. and Ikeuchi, K.: Determining Reflectance Parameters and Illumination Distribution from a Sparse Set of Images for View-dependent Image Synthesis, *Proc. 8th IEEE International Conference on Computer Vision ICCV '01*, Vol.1, pp.599-606 (Jul.2001).
- 19) Ramamoorthi, R. and Hanrahan, P.: A Signal-Processing Framework for Inverse Rendering, *Computer Graphics Proceedings, ACM SIGGRAPH 01*, pp.117-128 (Aug. 2001).
- 20) Sagawa, R., Nishino, K. and Ikeuchi, K.: Robust and Adaptive Integration of Multiple Range Images with Photometric Attributes, *Proc. IEEE Computer Vision and Pattern Recognition, CVPR '01*, pp.172-179 (Dec. 2001).
- 21) Sato, I., Sato, Y. and Ikeuchi, K.: Illumination distribution from brightness in shadows: adaptive estimation of illumination distribution with unknown reflectance properties in shadow regions, *Proc. 7th International Conference on Computer Vision ICCV '99*, Vol.1, pp.875-882 (Sep. 1999).
- 22) Sato, K. and Inokuchi, S.: Range-imaging system utilizing nematic liquid crystal mask, *1st*

International Conference on Computer Vision, pp.657-661 (1987).

- 23) Sato, Y., Wheeler, M.D. and Ikeuchi, K.: Object shape and reflectance modeling from observation, *Computer Graphics Proceedings, ACM SIGGRAPH 97*, pp.379-387 (Aug. 1997).
- 24) Shum, H-Y. and He, L-W.: Rendering with Concentric Mosaics, *Computer Graphics Proceedings, ACM SIGGRAPH 99*, pp.299-306 (Aug. 1999).
- 25) Solomon, F. and Ikeuchi, K.: Extracting the shape and roughness of specular lobe objects using four light photometric stereo, *Proc. Computer Vision and Pattern Recognition '92*, pp.466-471 (1992).
- 26) Williams, L.: Casting curved shadows on curved surfaces, *Computer Graphics Proceedings, ACM SIGGRAPH 78*, pp.270-274 (1978).
- 27) You, Y. and Kaveh, M.: A regularization approach to joint blur identification and image restoration, *IEEE Trans. Image Processing*, Vol.5, pp.416-428 (Mar. 1996).
- 28) Yu, Y., Debevec, P., Malik, J. and Hawkins, T.: Inverse Global Illumination: Recovering Reflectance Models of Real Scenes From Photographs, *Computer Graphics Proceedings, ACM SIGGRAPH 99*, pp.215-224 (Aug. 1999).

(平成 14 年 7 月 11 日受付)

(平成 15 年 1 月 16 日採録)

(担当編集委員 横矢 直和)



西野 恒

1997 年東京大学工学部電子情報工学科卒業。1999 年東京大学大学院工学系研究科電子情報工学専攻修士課程修了。2002 年東京大学大学院理学系研究科情報科学専攻博士課程修了。理学博士。現在、米国コロンビア大学コンピュータサイエンス科 Research Scientist。コンピュータビジョン、コンピュータグラフィックの分野における研究に従事し、特に physics-based vision, image-based rendering 等に関する研究を行っている。最優秀論文賞(日本バーチャルリアリティ学会:1999)、最優秀論文賞(VSMM:2000)等を受賞。IEEE, ACM 各会員。



池内 克史(正会員)

1973 年京都大学工学部機械工学科卒業。1978 年東京大学大学院工学系研究科情報工学科博士課程修了。工学博士。MIT 人工知能研究所、電総研、CMU 計算機科学およびロボティクス学科を経て、1996 年より東京大学生産技術研究所教授。人間の視覚機能、明るさ解析、物体認識、人間による組み立て作業の自動認識等の研究に従事。D. Marr 賞(ICCV:1990)、IEEE 優秀論文賞(CVPR:1991)、最多引用論文賞(AI Journal:1992)、最優秀論文賞(日本バーチャルリアリティ学会:1999)、最優秀論文賞(VSMM:2000)等受賞。IEEE Fellow。電子情報通信学会、日本ロボット学会各会員。



張 正友

1985 年精華大学工学部電気工学科卒業。1987 年パリ XI 大学 CS 博士課程了。Ph.D., D.Sc. フランス・INRIA を経て、現在マイクロソフトリサーチ Senior Researcher。カメラキャリブレーション、三次元形状復元等の low-level vision から視線検出、顔や表情のモデリングに至る応用分野まで幅広く研究を行っている。IEEE Senior Member, IJPRAI Associate Editor, IEEE TPAMI Associate Editor。