



照度差ステレオによる三次元形状推定

松下 康之*

Photometric stereo for 3D reconstruction

Key Words : 3D imaging, computer vision

1 はじめに

人工知能ブームである。深層学習（ディープラーニング）などに代表される機械学習技術の進展に伴い、画像データからそこに写っている物体をコンピュータに「認識」させる技術が大きく進んでいる。このような画像を用いた物体認識は一見すると簡単なタスクに思えるが、これをコンピュータプログラムとして実装しようとするとなかなか一筋縄にはいかない。現状の物体認識の精度は、ある種の単純化された問題設定においては人間の性能と同等であるという報告があるが、人間による視覚情報処理は多様で複雑なタスクに対応できる。人間と同等の視覚機能を持つロボットの「眼」の実現にはより一層の高度化が待たれている。コンピュータによる視覚情報処理は主にカラー画像、あるいはカラー＋深度画像が多く用いられてきたが、これらに含まれる情報は限定的である。新しいセンシング技術によってより多くの情報を含む「視覚」データを獲得できるならば、コンピュータによる認識機能をさらに高度化させることができるのではないだろうか。本稿では、微細な3次元形状情報の獲得を可能にする「照度差ステレオ」技術について概説する。

2 照度差ステレオ

ステレオという言葉は様々なコンテキストで用い

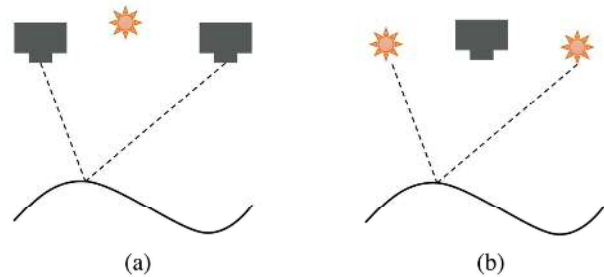


図1: (a) 二眼ステレオ：三角測量により深度を求める、
(b) 照度差ステレオ：光の当て方を変えながら撮影し陰影の変化から物体表面の法線を求める

られている。音響工学では二つのマイクで独立に収録した音声信号を二つのスピーカーを用いて再生する技術を指す。コンピュータビジョンの分野では、人間の二つの目のように、左右二つのカメラによって同一シーンを少しだけ異なる視点から撮影するステレオ撮影技術が広く用いられている。ステレオ視により得られた二枚の画像間で対応する点が求まれば、三角測量の原理によってカメラからシーンへの奥行きが求まるという仕組みである（図1(a)）。

これに対して照度差ステレオでは一つのカメラ視点において、光の方向を変えながら複数枚の画像を撮影する。いわば、二眼ステレオと対をなす仕組みである（図1(b)）。照度差ステレオでは、撮影された複数枚の画像中の陰影の変化パターンから物体表面の法線（surface normal）ベクトルを推定する。この法線ベクトルは、対象物体表面の「傾き」に関連する量である。照度差ステレオでは絶対的な深度情報は得られないものの、微細な表面形状を推定するためには非常に有効な手法として知られている。以下でまず準備として、法線ベクトル、平行光線、ランバート拡散反射モデルについて簡単に述べた上で、もっとも基本的なランバート照度差ステレオ法[4]について述べる。



* Yasuyuki MATSUSHITA

1974年9月生
東京大学大学院 工学系研究科 電子情報工学専攻（2003年）
現在、大阪大学 大学院情報科学研究科 教授 博士（工学）
コンピュータビジョン
TEL：06-6879-7825
FAX：06-6879-7827
E-mail：yasumat@ist.osaka-u.ac.jp



図2: 照度差ステレオ. ランバート拡散反射面の仮定の下では, 最低三つの異なる光源下で撮影された画像から法線ベクトルを推定できる. 左の三つの画像はこれらの入力画像, 右は推定された法線ベクトルをRGBで可視化したものである. 右下の球は, 法線ベクトルとRGBカラーコーディングの対応を示す: $(R, G, B)^T = (\mathbf{n} + \mathbf{1})/2$.

法線ベクトル 法線ベクトルとは, 物体表面の点における接平面に垂直な単位ベクトルを指す. 三次元空間中の法線ベクトルは三次元のベクトル $\mathbf{n} = [n_x, n_y, n_z]^T \in \mathbb{R}^3$ ($\|\mathbf{n}\|_2 = 1$)¹ として表現される. 法線ベクトルの要素 (n_x, n_y, n_z) は, それぞれ (x, y, z) 軸における成分を示す. 面には表裏があるので面に垂直なベクトルは二つあることになるが, 照度差ステレオのコンテキストではカメラ側から見える面を表の面として扱う.

平行光線 平行光線とは, 無限遠にある点光源からシーンに届く理論的な光線である. 平行光線で照らされた面は各点において同じ方向から同じ強度の光を受ける. 厳密な平行光線を作ることは実際には困難であるが, 十分に遠い点光源からの光は平行光線として近似することができる. 本稿では, 単位平行光線ベクトルを $\mathbf{s} \in \mathbb{R}^3$ ($\|\mathbf{s}\|_2 = 1$) と表現する.

ランバート拡散反射モデル まず反射モデルという言葉がわかりにくいものであるが, これはシーンを光で照らした時にどのような明るさの見える方をするか, を規定する関数であると考えていただきたい. ランバート拡散反射 (Lambertian diffuse reflection) は, 入射した光線をあらゆる方向に均等に反射する理想的な拡散反射モデルである. ランバート拡散反射モデルでは, 単位光強度の平行光線で照らされることを仮定すると, シーン中の物体表面が受ける光の単位面積当たりの強さは入射光 $\mathbf{s}$ と面の法線 $\mathbf{n}$ の

内積で決まる. 従って, 影の影響を無視すれば観測される明るさ $m \in \mathbb{R}_+$ は以下の式で表現される.

$$m = \rho \mathbf{s}^T \mathbf{n} \quad (1)$$

ここで非負の実数 $\rho \in \mathbb{R}_+$ は反射の度合いを表すランバート拡散反射率である. 例えば, 白い物体はより多くの光を跳ね返し (ρ が大きくなる), 黒っぽい物体ではその逆となる.

基本的な照度差ステレオ 初期の照度差ステレオ法は, 1970年代後半に Woodham によって提案された [4]. まず簡単のためにシーン内の一点に着目しこの点の法線ベクトル $\mathbf{n}$ を推定したいとしよう. このシーン内の点が平行光源 $\mathbf{s}_1$ の下で明るさ m_1 の見え方をするとする. この点の反射モデルに関して式 (1) のランバート拡散反射モデルを仮定すれば, 以下の式が一つ得られる.

$$m_1 = \rho \mathbf{s}_1^T \mathbf{n} \quad (2)$$

この平行光源をオフにして, 別の方向から平行光線を当てると似たような式が得られる. 三つの異なる平行光線下で撮影すると, 上述の式が三つ得られ, 以下のような連立線形方程式となる.

$$\begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix} = \begin{bmatrix} \mathbf{s}_1^T \\ \mathbf{s}_2^T \\ \mathbf{s}_3^T \end{bmatrix} \rho \mathbf{n} \quad (3)$$

行列 $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3] \in \mathbb{R}^{3 \times 3}$, 観測輝度ベクトル $\mathbf{m} = [m_1, m_2, m_3]^T \in \mathbb{R}^3$ とすると, $\mathbf{m} = \mathbf{S}^T (\rho \mathbf{n})$ と書ける. ここで求めたい未知数は法線ベクトル

¹ $\|\mathbf{x}\|_2$ はベクトル $\mathbf{x}$ の ℓ_2 ノルム (ユークリッドノルム) を指す.

$\mathbf{n}$ である。光線行列 $\mathbf{S}$ が逆行列を持てば、反射率 ρ でスケールされた法線ベクトル $\rho \mathbf{n}$ は $\rho \mathbf{n} = (\mathbf{S}^\top)^{-1} \mathbf{m}$ と解析的に決まる。さらに法線ベクトルは単位ベクトル ($\|\mathbf{n}\|_2 = 1$) であることを考えると、以下の手順で拡散反射率 ρ と法線ベクトル $\mathbf{n}$ を決定できる。

$$\begin{cases} \rho \leftarrow \|\rho \mathbf{n}\|_2 \\ \mathbf{n} \leftarrow \rho \mathbf{n} / \rho \end{cases} \quad (4)$$

ここまではシーン中の一点に着目してきた。実際にはカメラを用いて複数の点を同時に観測することができるので、上記の結果を複数点に拡張してみよう。 p 個の点が観測されるとすると (つまり p ピクセルの画像)、式 (3) を拡張して

$$\begin{bmatrix} m_{11} & m_{1p} \\ m_{21} & \cdots & m_{2p} \\ m_{31} & m_{3p} \end{bmatrix} = \begin{bmatrix} \mathbf{s}_1^\top \\ \mathbf{s}_2^\top \\ \mathbf{s}_3^\top \end{bmatrix} \begin{bmatrix} \rho_1 \mathbf{n}_1 & \cdots & \rho_p \mathbf{n}_p \end{bmatrix} \quad (5)$$

となる。左辺を $\mathbf{M} \in \mathbb{R}^{3 \times p}$ 、右辺の法線ベクトルの行列を $\mathbf{N} \in \mathbb{R}^{3 \times p}$ とすれば、上述の手順と同様に $\mathbf{N} = (\mathbf{S}^\top)^{-1} \mathbf{M}$ のように各点における拡散反射率 ρ でスケールされた法線ベクトルマップが求まる。最終的に、各々の点において式 (4) のように拡散反射率を決定することで、単位法線ベクトルから成る法線ベクトルマップが得られる (図2)。

3 照度差ステレオの実験

前節で述べた照度差ステレオを実環境において適用する際には幾つかの問題に直面するだろう。まず、影の影響を無視してきたために、式 (1) による観測輝度 m を負にしてしまうような法線 $\mathbf{n}$ が求まるケースがある。実際には、負の輝度値というものは観測されることはなく、影の領域の輝度値はゼロとなるため $m = \max(\rho \mathbf{s}^\top \mathbf{n}, 0)$ のように表現する方が適切となる。画像センサのノイズも完全には無視できない。さらには、ランバート反射モデルによって表現されない、物体間での相互反射や鏡面反射などの影響も現実には存在するであろう。これらすべての要因を数理モデルとして表現することは困難であるため、まず単純な拡張としてより多くの (>3) 画像を撮影することによってこれらのモデル化されていない要因を軽減しようというアプローチが取られてきた。つまりランバート反射モデルに沿わない観測値を所謂「ノイズ」として扱おうというわけである。

記述の簡略化のためにシーン中の一点に着目しているセッティングに戻る。式 (3) では三つの光源下における連立方程式を扱ったが、これを $f(>3)$ 個の異なる光源下での式に拡張する。

$$\begin{bmatrix} m_1 \\ \vdots \\ m_f \end{bmatrix} = \begin{bmatrix} \mathbf{s}_1^\top \\ \vdots \\ \mathbf{s}_f^\top \end{bmatrix} \rho \mathbf{n} \quad (6)$$

この左辺を $\mathbf{m} \in \mathbb{R}^f$ 、光源行列を $\mathbf{S} \in \mathbb{R}^{3 \times f}$ 、スケールされた法線ベクトルを $\mathbf{b}(=\rho \mathbf{n}) \in \mathbb{R}^3$ とすると、上式は $\mathbf{m} = \mathbf{S}^\top \mathbf{b}$ と書ける。これは優決定系の線形システムとなり、ノイズの影響を想定しているので等号を用いずに $\mathbf{m} \simeq \mathbf{S}^\top \mathbf{b}$ とするのが適当である。つまり、観測輝度ベクトル $\mathbf{m}$ と既知の光源行列 $\mathbf{S}$ が与えられた状況で、この近似式を (完全にではないにせよ) 最もよく満たす法線行列の近似解 $\mathbf{b}^*$ を推定したい、という問題に帰着する。

最小二乗法による近似解 優決定系線形システムの近似解法として最も広く用いられているものは最小二乗法であろう。線形システムの最小二乗近似解法はガウスとルジャンドルにより独立に導かれ、最適な近似解が解析的に求まることが知られている。問題設定は以下の通りである。

$$\mathbf{b}^* = \operatorname{argmin}_{\mathbf{b}} \|\mathbf{m} - \mathbf{S}^\top \mathbf{b}\|_2^2 \quad (7)$$

線形システムの二次式であるので未知数 $\mathbf{b}$ に関して微分を取り、その極値点が近似解 $\mathbf{b}^*$ となる。

$$\mathbf{b}^* = (\mathbf{S}\mathbf{S}^\top)^{-1} \mathbf{S}\mathbf{m} \quad (8)$$

線形システムの最小二乗解法は多くの数値計算ライブラリやソフトウェアに実装されており、非常に手軽に実装することができる。例えば、matlab では、 $\mathbf{b} = \mathbf{S}' \backslash \mathbf{m}$ といったコマンドインターフェースにより式 (7) の近似解を得ることができ大変便利である。

外れ値にロバストな推定 最小二乗法は残差の分布が正規分布に従う時に最適な近似解を与えることが知られている。しかしながら、前述した影や鏡面反射 (図3) によるランバート拡散反射モデルからの逸脱は正規分布に従うと期待されるものだろうか。より最近の研究で、最小二乗法ではなく残差ベクトル $\mathbf{r} = \mathbf{m} - \mathbf{S}^\top \mathbf{b}$ の ℓ_1 ノルム ($\|\mathbf{r}\|_1$) の最小化によ



図3: 実際に観測されるシーンには鏡面反射や影などの非ランバート拡散反射成分が含まれる。

る法線ベクトルの推定が効果的であることが報告されている [5, 2]。残差の ℓ_1 ノルム最小化問題は以下のように記述される。

$$\mathbf{b}^* = \operatorname{argmin}_{\mathbf{b}} \|\mathbf{m} - \mathbf{S}^T \mathbf{b}\|_1 \quad (9)$$

一見すると式 (7) とほぼ同じであるが、ノルムを示す右下の添字が2から1に変わり、さらに二乗の記号が落ちている。式 (9) のセッティングは外れ値に対してロバストな、もしくは残差を「疎 (スパース)」にする回帰をすることで知られており、他の ℓ_1 ノルムを扱う圧縮センシング、スパース回帰などの問題とともに広く注目を集めている。残念ながら式 (9) は最小二乗法とは異なり、解析解を導くことはできないが、簡単な数値計算アルゴリズムが幾つか知られている。その一つに Iteratively Reweighted Least Squares (IRLS) 法というものがある。このアルゴリズムでは重み付き最小二乗問題をその重みの更新をしながら繰り返し解くことにより、徐々に ℓ_1 残差の最小化を行うというものである (アルゴリズム 1) [1]。アルゴリズム 1 中の (A, b, x) をそれぞれ

式 (9) 中の $(\mathbf{S}^T, \mathbf{b}, \mathbf{m})$ と読みかえれば、式 (9) のコンテキストから利用できることがわかる。

照度差ステレオにおける外れ値 (影・鏡面反射など) は、非常に大きな誤差を持つ一方でそれらが観測されるピクセル数は全体に比して十分に小さい。言い換えれば、これらの誤差は大きいながらも「疎」であることが期待できる。従ってこのようなケースにおいては、大きい疎な外れ値に対してロバストな ℓ_1 最小化が有効に働く。図4に、40枚の画像からなるシーンに対して最小二乗法、残差の ℓ_1 最小化法をそれぞれ適用した結果を示す。このシーンは影や鏡面反射の成分を含むために、 ℓ_1 最小化が最小二乗法に比して有効であることを示している。

アルゴリズム 1 IRLS による ℓ_1 残差の最小化

入力: $A \in \mathbb{R}^{m \times n}$ と $b \in \mathbb{R}^m$

出力: $Ax \simeq b$ の ℓ_1 近似解 x

// 重み対角行列 W の初期化

$W \leftarrow I$ // I は単位行列

while not converged do

// 重み付き最小二乗

$x \leftarrow \operatorname{argmin}_x (Ax - b)^T W (Ax - b)$

// 残差の計算

$r \leftarrow Ax - b$

// 重み対角行列の更新 (i は要素の index)

$w_i \leftarrow \frac{1}{|r_i| + \epsilon}$ // ϵ は非常に小さな正の数

end while

4 むすび

駆け足ではあったが、物体表面の微細な形状情報を推定するために有効な照度差ステレオ法について概説した。照度差ステレオ法は本稿で紹介した以外にも様々な方向に発展している。例えば、平行光源の仮定を緩和したり、あるいはランバート拡散反射

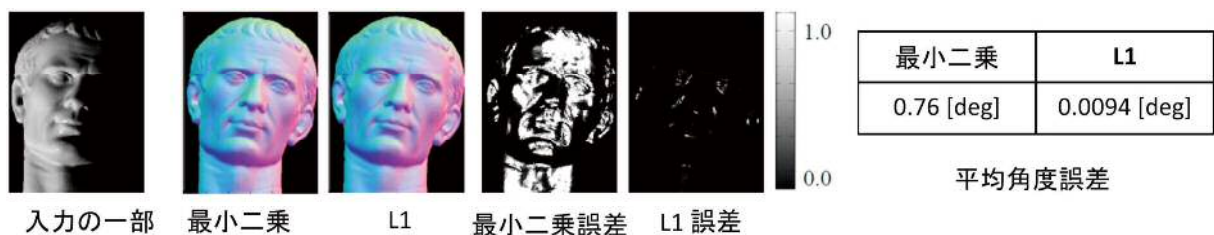


図4: 40枚の画像を用いた照度差ステレオの結果。最小二乗法と ℓ_1 残差最小化に基づく法線ベクトルの推定結果とそれらの誤差マップ。右表はそれぞれの解法による平均角度誤差を示す。

モデルよりもより表現力の高い反射モデルを導入したりなどすることで、実世界の様々なアプリケーションへの適用を目指している。実際、非ランバート拡散面への照度差ステレオの適用は、1980年代初頭から検討されている [3]。最近では、凸最適化の進歩に伴い、これまでに数値的に解くことが困難であったような種の設定（例えばロバスト主成分分析による方法 [5]）でも現実的な計算時間で法線ベクトルの頑健な推定が可能になってきている。これらの進展により、照度差ステレオ法もそろそろ実用に耐えるものになってきたと言っても過言ではないだろう。今後はこのようなセンシング技術を用いた物体認識等のアプリケーションの高度化に期待したい。

参考文献

- [1] I. F. Gorodnitsky and B. D. Rao. Sparse signal reconstruction from limited data using focuss: A re-weighted minimum norm algorithm. *IEEE Transactions on Signal Processing*, 45(3):600 – 616, March 1997.
- [2] S. Ikehata, D. Wipf, Y. Matsushita, and K. Aizawa. Robust photometric stereo using sparse regression. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 318 – 325, 2012.
- [3] K. Ikeuchi. Determining Surface Orientations of Specular Surfaces by Using the Photometric Stereo Method. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 3(6):661 – 669, 1981.
- [4] R. J. Woodham. Reflectance map techniques for analyzing surface defects in metal castings. Technical report, TR-457, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, 1978.
- [5] L. Wu, A. Ganesh, B. Shi, Y. Matsushita, Y. Wang, and Y. Ma. Robust photometric stereo via low-rank matrix completion and recovery. In *Proceedings of Asian Conference on Computer Vision (ACCV)*, volume 3, pages 703 – 717, 2010.

