

アルベドと法線ベクトルの緊密度に着目した 影と鏡面反射除去のための組み合わせ照度差ステレオ法

家永 直人^{1,a)} 斎藤 英雄^{1,b)} 清水 雅芳^{2,c)} 手塚 耕一^{2,d)} 岩村 康正^{2,e)}

概要: 本稿は影や鏡面反射を含む画像に対しても正確に法線を推定できる組み合わせ照度差ステレオ法を提案する。照度差ステレオ法は入力画像が3枚以上であれば利用できる。よって入力画像が n 枚の場合、 nC_3 通りの組み合わせで照度差ステレオ法を用いることができる。本手法では nC_3 通りの画素の組み合わせから推定されたアルベドと法線の分布を作り、影にも鏡面反射にも含まれない画素をそれ以外の画素と区別するために“緊密度”を定義する。緊密度は、アルベドと法線の集合度合いを表しており、影や鏡面反射に含まれない画素から計算されたものほど大きくなる。これを利用して、緊密度が最大のアルベドと法線の周辺から投票を行い、影と鏡面反射除去を行う。実験結果を通し、提案手法が影や鏡面反射のある画像に対しても正確に法線を推定できることだけでなく、既存研究よりも優れていることを示す。

キーワード: 照度差ステレオ法, 3次元形状復元

1. はじめに

3Dプリンタの登場に見られるように、3次元形状復元は近年注目を集めている。照度差ステレオ法は、対象物体の3次元形状を取得する手法としてよく知られている。実際には照度差ステレオ法によって得られる法線を積分することで、3次元形状を得ることができる。照度差ステレオ法は光源方向のみかわる複数枚の入力画像から、各画素のアルベド（入射光と反射光の比）と法線を推定する。だが照度差ステレオ法は、物体表面の反射がランバート反射に従うと仮定しているため、その仮定が成り立たない影や鏡面反射領域が入力画像に含まれると正しく法線を計算できない。

照度差ステレオ法は3枚以上の入力画像を必要とする。そのため入力画像が n 枚の場合、 nC_3 通りで照度差ステレオ法を適用することができる。そこで本手法では、 n 枚の入力画像の各画素の nC_3 通りの組み合わせに対し、照度差ステレオ法を適用し、各組み合わせでアルベドと法線 (p ,

q の2つの値になる。3.1参照)を算出する。そのため各組み合わせは3つの値を持つ。これを“三つ組み”と定義する。本手法では p , q (法線)の2次元空間ではなく、アルベドも含めた3次元空間で、各画素ごとに nC_3 個の三つ組みの分布を考える。影や鏡面反射に含まれない画素を選択するために、この分布において、“緊密度”を定義する。緊密度は、アルベドと法線を表す三つ組みの集合度合いである。そして緊密度が最大の三つ組みを用いて投票を行うことにより、画素を選択する。このように本稿では、緊密度を用いて、ランバート反射に従わない画素を除去する手法を提案する。

本稿は5部構成である。第2章では、関連研究について述べる。第3章では提案手法の詳細を述べる。第4章では、提案手法の有効性を確認するための実験を行い、結果とその考察を述べる。第5章では、本稿の結論を述べる。

2. 関連研究

影や鏡面反射の影響を考慮した照度差ステレオ法の研究は、長い間研究されてきた。Chungら[1]は対象物体の投影を用いて、Ward BRDFモデルのパラメータを推定し、光沢の影響を除去する手法を提案した。Hernら[2]やBarskyら[3]は入力画像が3枚、ないし4枚しか得られない場合での照度差ステレオ法を提案した。入力画像が少ないため、各画素において影や鏡面反射に含まれてもよい画素は1枚の入力画像においてのみである。これらに対して

¹ 慶應義塾大学情報工学科
Department of Information and Computer Science, Keio University

² 富士通研究所
Fujitsu Laboratories Ltd.

a) ienaga@hvrl.ics.keio.ac.jp

b) saito@hvrl.ics.keio.ac.jp

c) shimizu.masa@jp.fujitsu.com

d) ktezuka@jp.fujitsu.com

e) iwamura.yasumas@jp.fujitsu.com

本手法は最低 4 枚の入力画像が必要であり、各画素は最低 3 枚の画像において、影や鏡面反射の影響を受けていないことを前提とするものである。

Chandraker ら [4] は、グラフカットを用いて影の影響を受けている領域を除去し、照度差ステレオ法で形状を復元する手法を提案した。Dulac ら [5] は影だけの問題に焦点を当て、ある画素がどれかの画像において影に含まれているなら、もっとも暗い画素の画像においてであるという仮定をした。まずもっとも暗い画素、もっとも明るい画素、3 番目に明るい画素から法線を計算する。次にその法線を使い逆算し求めた明度値と観測された明度値を比較する。この差が大きければもっとも暗い画素を除く。これらの処理を差が小さくなるまで繰り返す。

宮崎ら [6], [7] もまた影や鏡面反射の問題に取り組んだ。[6] ではグラフカットを、[7] では中央値を用いた。[7] で、彼らはすべての画素において本手法と同様に nC_3 通りの法線を計算し、4 近傍の画素も含めた中央値と 4 近傍の平均値を使い各画素の法線を算出した。

3. 組み合わせ照度差ステレオ法

本手法の概要を述べる。目的は各画素において、影や鏡面反射に含まれない画素を選択することである。入力画像を n 枚とすると、本手法はまず各画素において nC_3 通りで照度差ステレオ法を適用し、 nC_3 個のアルベドと法線を得る。これらを、2 つの値で表される法線だけの 2 次元空間ではなく、アルベドも含めた 3 次元空間に分布させる。ここで緊密度を定義し、もっとも正しいと思われる値を決定するが、それはその値を最終的な値として利用するためではなく、その値を使い投票を行うことで、影や鏡面反射に含まれない画素を選択するためである。以下、上述の流れに沿って詳細を述べる。

3.1 三つ組みの推定

照度差ステレオ法は光源の位置が既知の場合と未知の場合で大別されるが、本手法は既知の場合を対象とする。各画素について、アルベドを ρ 、光源の強度を s 、入力画像の光源方向からなる n 行 3 列の行列を \mathbf{L} 、法線ベクトルを \mathbf{n} とし、対象物体表面の反射がランバート反射に従うと仮定すると、 n 枚の各画像におけるある画素の明度値を表す n 次元ベクトル \mathbf{V} は、以下のように表される。

$$\mathbf{V} = \rho s \mathbf{L} \mathbf{n} \quad (1)$$

ここで $s = 1$ とし、両辺に \mathbf{L}^{-1} をかけると、

$$\mathbf{L}^{-1} \mathbf{V} = \rho \mathbf{n} \quad (2)$$

となる。 \mathbf{n} は単位ベクトルなので、式 (2) の左辺の絶対値が ρ になる。

$$\rho = |\mathbf{L}^{-1} \mathbf{V}| \quad (3)$$

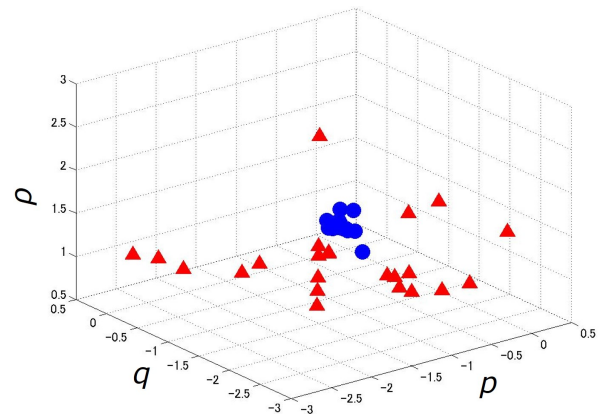


図 1 ある画素の三つ組みの分布。入力画像が 8 枚あるため $56({}_8C_3)$ 個の三つ組みがある。手動でどの三つ組みが正しく、どの三つ組みが誤っているか指定した。ほとんどの正しい三つ組みはある特定の位置に集まっている、一方誤った三つ組みの多くは広く分布している。

また \mathbf{n} は次の通りである。

$$\mathbf{n} = \frac{\mathbf{L}^{-1} \mathbf{V}}{|\mathbf{L}^{-1} \mathbf{V}|} = (n_x, n_y, n_z)^T \quad (4)$$

本手法は法線の冗長さを取り除くために、物体表面に平行な x 方向ベクトル $\mathbf{r}_x = (1, 0, p)^T$ と y 方向ベクトル $\mathbf{r}_y = (0, 1, q)^T$ を用いた。 p と q はそれぞれ x 方向と y 方向における表面の勾配と呼ばれる。法線はこれらのベクトルの外積により計算される。 p 、 q と式 (4) の関係は以下のようになる。

$$\mathbf{n} = \mathbf{r}_x \times \mathbf{r}_y = \begin{pmatrix} -p \\ -q \\ 1 \end{pmatrix} = \begin{pmatrix} n_x/n_z \\ n_y/n_z \\ 1 \end{pmatrix} \quad (5)$$

本手法は入力画像 n 枚の各画素 nC_3 通りの組み合わせに対し、 ρ と p 、 q をそれぞれ式 (3)、(5) によって求める。そしてこれらの ρ 、 p 、 q 、3 つの値を“三つ組み”と定義する。

3.2 3次元分布の作成

図 1 は、 p 、 q 、 ρ の 3 次元空間上で、図 7 (4.2 章に記載) のある画素における nC_3 個の三つ組みの分布を示している。円は影にも鏡面反射にも含まれない画素（以下、正しい画素と呼ぶ）のみから計算された三つ組み（以下、正しい三つ組みと呼ぶ）、三角は影や鏡面反射に含まれる画素（以下、誤った画素と呼ぶ）を 1 つ以上使い計算された三つ組み（以下、誤った三つ組みと呼ぶ）である。

求めた三つ組みを法線、つまり p 、 q の 2 次元空間に分布させるのではなく、 p 、 q 、 ρ の 3 次元空間に分布させたのは、より正しい三つ組みと誤った三つ組みを区別しやすくするためである。

本手法の目的である、正しい画素（影や鏡面反射に含まれない画素）の選択をするには、正しい三つ組みがどこに分布しているかがわかりさえすればよい。なぜなら正しい三つ組みがどれかさえわかれば、あとはそれらがどの画素から計算されたかを確認すればよいだけだからである。正しい三つ組みの構成要素である画素は、正しい画素である（これは定義である）。しかし、正しい三つ組みがどれかを推測するために、三つ組みの正しさを示す値が必要である。

3.3 緊密度の計算

正しい三つ組みは近い値をもっている（すべての画素がランバート反射に従いかつ誤差がなければ、本来完全に一致する）ため、図1のように集まることがわかる。この性質は正しい三つ組みがどれかを推測するために有用である。この性質を利用するためには、各三つ組みの周囲にくつ三つ組みがあるかを数えればよい。周囲により多くの三つ組みがある三つ組みほど、正しい三つ組みと言える。これを“緊密度”として定義する。

$$\text{緊密度} = \text{ある三つ組みの周囲の三つ組みの数} \quad (6)$$

具体的には、注目している三つ組みのしきい値 th_d 以内にある三つ組みの数が緊密度になる。正しい三つ組みは集まりやすく、周囲に三つ組みが多いので、正しい三つ組みほど緊密度が大きくなる。

だが、緊密度が最大の三つ組みの法線を、単純にその画素における最終的な法線として利用しない。なぜなら緊密度が最大とは言え、 n 個画素があるのに常に3つの画素しか利用しないよりも、影や鏡面反射の影響のない画素をできる限り多く利用したほうが正確な法線を算出できるからである。そこで次に述べるように、影や鏡面反射の影響がないと判断されるすべての画素を選択する。

3.4 投票による画素の選択

計算された緊密度を使い、正しい画素を選択するために緊密度がもっとも大きかった三つ組みを使い投票を行う。具体的には緊密度が最大の三つ組みの th_d 以内にあるすべての三つ組みを構成する画素を数える。図2の例では、画素1は2票、画素2は3票、画素3は0票・・・といった具合である。緊密度が最大の三つ組みが複数ある場合でも同じ表に投票する。緊密度が最大の三つ組みとして正しい三つ組みが選ばれることと、その周囲には正しい三つ組みが多いことが期待できるので、正しい画素の票が多くなるはずである（正しい三つ組みは正しい画素のみから計算されるから）。そこで票数が全体の平均 - 標準偏差 以上の画素を選択する。なお、この投票の際のしきい値 th_d は、緊密度の計算の際のそれよりも大きくする。正しい三つ組みを取りこぼさないためである。

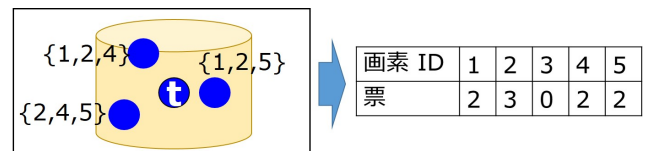


図2 投票例。円柱は“ t ”と書かれた注目している三つ組みから、しきい値 th_d の範囲を表している。周囲に3つの三つ組みがある。それぞれの三つ組みを構成する3つの画素に投票した結果が右の表である。

3.5 しきい値の決定

前述のように本手法では th_d を2回設定しておく必要がある。加えて、 ρ と p , q のスケールが異なる（このため図2に円柱が描かれている）ため、 th_d に関して4つの値を決定しなければならない。これらのしきい値を簡単に決定するために以下のアルゴリズムを用いる。まず緊密度計算の際には、初期値ですべての三つ組みの th_d 以内に、他の三つ組みがしきい値 th_f 個以上ない場合、互いに th_d 外にある三つ組み間で最小の距離を、 th_d の足し再度計算を行う。また画素選択の際は、緊密度が最大のすべての三つ組みの th_d 以内に他の三つ組みがないとき、緊密度が最大の三つ組みとそれ以外の三つ組みとの間で最小の距離を初期値に足し、再度投票を行う。このアルゴリズムにより特に緊密度計算の際、非常に簡単に th_d の初期値を決定できる。なぜなら緊密度計算のときには緊密度が最大の三つ組みを選ぶことが目的であり、 th_d は小さい値であるべきだが、上記のアルゴリズムにより th_d を小さくし過ぎてしまうことはない。

3.6 3次元形状復元

画素を選択した後、それらの画素を使い法線を計算する。そして法線を積分して対象物体の3次元形状を復元する。この工程では Xu ら [8] の手法を用いた。

4. 実験と結果

4.1 影と鏡面反射に対する本手法の効果

この実験では通常の照度差ステレオ法（すべての画素を法線の計算に用いる）と比較し、影と鏡面反射に対する本手法の効果を示す。図3はPOV-Ray[10]と[9]から入手したobjファイル*1を使い作成した8枚の入力画像のうち1枚である。影と鏡面反射を含んでいる。図4は通常の照度差ステレオ法（左）と本手法（右）のそれぞれで得た法線をもとに復元された3次元形状である。

図5は本手法の鏡面反射に対する効果を示している。通常の照度差ステレオ法（左）では、図3にあるように鏡面反射の影響により、形状の中央が突起している。一方、本手法（右）ではより滑らかな形状になっている。

*1 Josea, “Clay vase garden pottery,” (http://artist-3d.com/free.3d_models/dnm/model_disp.php?uid=3792)



図 3 入力画像 図 4 復元された形状. 左:通常の照度差ステレオ法. 右:本手法.

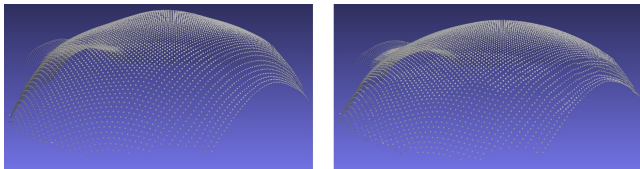


図 5 図 4 を左下から見た 3 次元点群. 左:通常の照度差ステレオ法. 右:本手法.

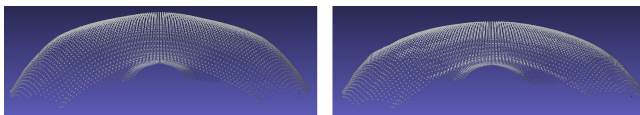


図 6 図 4 を下から見た 3 次元点群. 左:通常の照度差ステレオ法. 右:本手法.



図 7 入力画像. 鏡面反射はほとんどなく, 影を多く含む.

また図 6 は本手法の影に対する効果を示している. 影の影響により, 通常の照度差ステレオ法 (左) は台形に近い形状になってしまっていることがわかる. しかし, 本手法 (右) はより自然な丸みを帯びた形である.

4.2 実画像を用いた関連研究との比較

この実験は, Xiong ら [11] のデータ *2 の図 7 に示した画像を用いた. この実験では通常の照度差ステレオ法だけでなく, 宮崎ら [7] の手法, Dulac ら [5] の手法とも本手法を比較した. 図 7 は 8 枚の入力画像のうちの 3 枚である. これらはあまり鏡面反射を含んでおらず, 影を多く含んでいる. これらの画像を選んだのは, 影の問題のみを扱っている Dulac らの手法と比較するためである.

それぞれの手法で得られた法線と, 法線の真値との成す角を比較した. 図 8 で, 黒い画素は成す角が 15 度以上, 白い画素は成す角が 0 度 (真値と等しい) であることを示している. 左上は通常の照度差ステレオ法である. 影のある領

*2 ダウンロードについては以下の URL を参照のこと.
(<http://vision.seas.harvard.edu/qsfs/Data.html>)

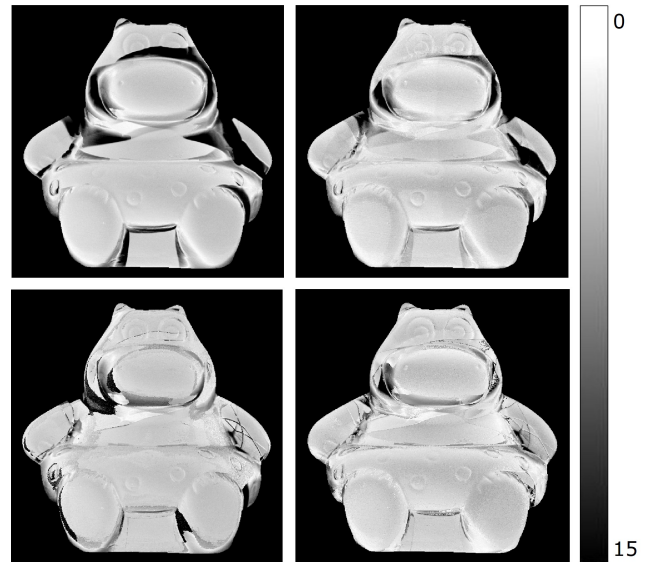


図 8 グレースケール画像. 真値との成す角 0-15 度が明度値 255-0 に対応している. 左上:通常の照度差ステレオ法. 右上:宮崎ら [7] の手法. 左下:Dulac ら [5] の手法. 右下:本手法.

表 1 平均二乗誤差 (ここで誤差は角度のこと).

通常の 照度差ステレオ法	宮崎ら [7]	Dulac ら [5]	本手法
7.76	4.97	4.93	3.95

表 2 アルベドと投票の効果.

	投票あり	投票なし
p, q, ρ 3 次元空間	3.95	4.43
p, q 2 次元空間	4.26	5.23

域はほとんど黒い. 右上と左下はそれぞれ宮崎らと Dulac らの手法である. 通常の照度差ステレオ法に比べ, 黒い領域が減っている. だが, まだ黒い領域が目立っており, 十分に影の影響を除去できたとは言えない. 右下が本手法の結果である. もっとも黒い領域が少ない.

表 1 は真値との成す角を誤差と捉え, 物体領域全体から算出したそれぞれの手法の平均二乗誤差を示している. 定量的に評価しても本手法が他の手法よりも効果的であることがわかる.

また表 2 は, p, q の 2 次元空間ではなく, ρ も加えた 3 次元空間で三つ組みの緊密度を考えたこと, そして, 緊密度が最大の三つ組みの法線をそのまま利用するのではなく, 投票し画素の選択をしたこと, それら 2 つの効果を示している. 3 次元空間を利用し, かつ投票も行った結果が一番よい. 一方, 2 次元空間を利用し, 投票も行わない場合は既存手法よりも悪い結果となっている.

図 9 は通常の照度差ステレオ法 (左) と本手法 (右) によって得られた 3 次元形状である. 通常の照度差ステレオ

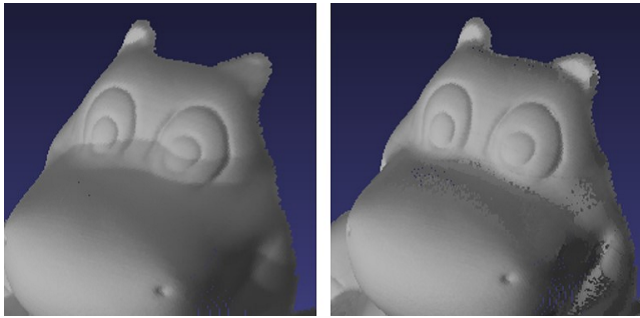


図 9 3次元形状. 左:通常の照度差ステレオ法. 右:本手法.

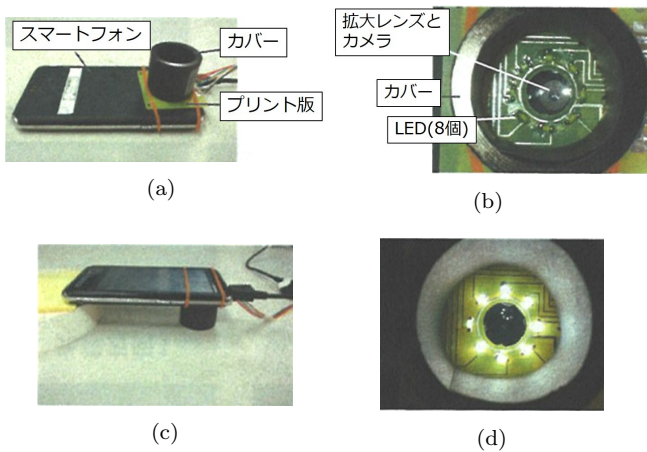


図 10 開発した装置. (a)(b):装置の詳細. 8個のLEDが順番に点灯しカメラが8枚の画像を撮影する. (c):撮影時の様子. (d):すべてのLEDを点灯させたとき.



図 11 撮影したシリコンレプリカの画像

法では影の影響により、法線を正確に計算できなかった。結果として目と鼻の間の窪みが膨れている。このように不正確な法線を積分すると、復元される形状は膨らむ傾向があると言える。

4.3 実利用の実験

最後に著者が撮影した実画像を用いて、本手法の実利用の実験を示す。図 10 に示した小型装置を開発した。これはスマートフォンに取り付けることができ、点灯している LED をかえながら 8 枚の画像を撮影する装置である。図 11 は 50 円硬貨のシリコンレプリカを装置で撮影した、8 枚のうちの 2 枚の画像である。鏡面反射を多く含んでいる。これらを入力画像とし、通常の照度差ステレオ法と本手法によって得られた 3 次元形状が図 12 や 13 である。

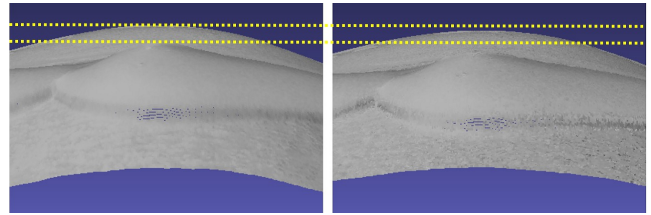


図 12 3次元形状. 左:通常の照度差ステレオ法. 右:本手法. 2本の点線は画像に平行である. 上の線は、通常の照度差ステレオ法の形状の中央と同じ高さで、下の線は、通常の照度差ステレオ法の形状の縁の膨らみと同じ高さである. しかし、2本の線は、本手法のそれらよりも高い.

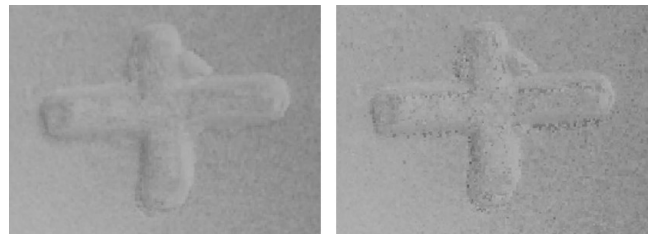


図 13 十の字の3次元形状. 左:通常の照度差ステレオ法. 右:本手法.

図 12 は 3 次元形状を下から見た図である。左が通常の照度差ステレオ法で、右が本手法である。通常の照度差ステレオ法により得られた形状は、図 9 のように、本手法よりも膨らんでいる。さらに図 13 では、通常の照度差ステレオ法の形状が膨らんでいるため十の字が、本手法よりも広がって見える。ゆえに、図 9 の結論から、本手法は通常の照度差ステレオ法よりも正確に法線を計算したと言える。このように本手法を用いて影や鏡面反射の影響を除きつつ、スマートフォンに装着可能な小型装置で手軽に任意の物体の 3 次元形状を取得することができた。

5. 結論

本稿では照度差ステレオ法において、影と鏡面反射の影響を除去する手法を提案した。入力画像 n 枚に対し、各画素で nC_3 通りの組み合わせからアルベドと法線を計算し、影にも鏡面反射にも含まれない画素のみから計算されたアルベドや法線の考察に基づく値、“緊密度”を定義した。この緊密度を使い、影や鏡面反射に含まれている画素とそうでない画素を識別した。実験の結果を通して、本手法が影と鏡面反射の影響を除去したことだけでなく、既存手法よりもよい結果を得たこと、さらに実利用の可能性を示した。

参考文献

- [1] H. Chung and J. Jia, “Efficient photometric stereo on glossy surface with wide specular lobes,” in *Proc. 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2008)*, June, 2008.

- [2] C. Hernández, G. Vogiatzis and R. Cipolla, “Overcoming shadows in 3-source photometric stereo,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 33, no. 2, pp. 419-426, Feb. 2011.
- [3] S. Barsky and M. Petrou, “The 4-source photometric stereo technique for three-dimensional surface in the presence of highlights and shadows,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 25, no. 10, pp. 1239-1252, Oct. 2003.
- [4] M. Chandraker, S. Agarwal, and D. Kriegman, “Shadow-cuts: photometric stereo with shadows,” in *Proc. 2007 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2007)*, June 2007.
- [5] A. Dulac, S. Martedi, H. Saito, K. Tezuka, and M. Shimizu, “Combination method: photometric stereo with shadows,” in *Proc. 2014 Irish Machine Vision and Image Processing (IMVIP2014)*, pp. 137-142, Aug. 2014.
- [6] D. Miyazaki and K. Ikeuchi, “Photometric stereo using graph cut and m-estimation for a virtual tumulus in the presence of highlights and shadows,” in *Proc. Workshop on Applications of Computer Vision in Archaeology*, June 2010.
- [7] D. Miyazaki, K. Hara and K. Ikeuchi, “Median photometric stereo as applied to the segonko tumulus and museum objects,” *International Journal of Computer Vision*, vol. 86, no. 2-3, pp. 229-242, Jan. 2010.
- [8] J. Xu and D. Pachauri, “Project 3: photometric stereo,” (<http://pages.cs.wisc.edu/~jiayu/projects/Stereo/>) (2015/01/17)
- [9] ARTIST-3D.COM (<http://artist-3d.com/>) (2015/01/21)
- [10] POV-Ray (<http://www.povray.org/>) (2015/01/21)
- [11] Y. Xiong, A. Chakrabarti, R. Basri, S. J. Gortler, D. W. Jacobs and T. Zickler, “From shading to local shape,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 37, no. 1, pp. 67-79, Jan. 2015.
- [12] B. K. P. Horn, *Robot Vision*. Massachusetts: The MIT Press, 1986.