

2015年度 修士論文

**Minnaert 則を考慮した
小惑星の3次元形状復元**

指導教員 石川 博 教授

早稲田大学 基幹理工学研究科 情報理工・情報通信専攻

5114F017-5 岩野 俊介

2016 年 2 月 1 日 提出

目次

第1章	はじめに	5
1.1	はやぶさ2プロジェクト	5
1.2	探査機が撮影する小惑星画像の性質	6
1.3	複数枚の画像を用いた対象物体の3次元復元	7
1.4	研究目的	9
第2章	関連研究	10
2.1	複数枚の画像からの3次元復元手法	10
2.1.1	Multi-view Stereo	10
2.1.2	Photometric Stereo	12
2.1.3	Multi-view Photometric Stereo	13
2.2	物体表面の反射モデル	15
2.2.1	Lambert 則	15
2.2.2	Minnaert 則	15
第3章	提案手法	18
3.1	手法概要	18
3.2	光源方向と表面反射率の推定	20
3.3	光源方向を考慮した最適化	22
3.3.1	エネルギー最小化による各推定値の最適化	22
3.3.2	Visibility の推定と視点の選択	23
3.3.3	推定表面に沿った点の追加と外れ値除去	25
3.3.4	背景に投影される位置に移動した点の除去	28
第4章	実験	30
4.1	実装	30
4.2	合成画像を用いた光源方向の推定実験	30
4.2.1	実験目的	30

4.2.2	実験方法	30
4.2.3	実験結果と考察	31
4.3	エネルギーの閾値 ϵ とサブセットの枚数を変えた場合の比較	32
4.3.1	目的	32
4.3.2	実験方法	32
4.3.3	実験結果と考察	33
4.4	Minnaert 則の対象物体の復元	34
4.4.1	目的	34
4.4.2	実験方法	38
4.4.3	実験結果と考察	39
第 5 章 おわりに		41

図 目 次

1.1	小惑星の撮影環境の概念図.	7
1.2	はやぶさがマルチバンド分光カメラで撮影したイトカワの画像例.	8
2.1	Multi-view stereo で対象とする撮影シーンの概略図.	11
2.2	Photometric stereo の対象とする撮影シーンの概略図.	11
2.3	反射モデルによる輝度の計算に使われるパラメータの概略図.	15
2.4	平面と粗面がピクセルに観測されたときの違いの概略図.	16
2.5	粗面の観測輝度の視点による変化の概念図.	17
3.1	撮影シーンの概略図.	19
3.2	手法全体の流れの概要図.	21
3.3	Cast shadow の例.	24
3.4	Occulusion の概略図.	25
3.5	視点選択の概念図.	26
3.6	最適化手法の流れの概要図.	27
3.7	背景に投影されるエネルギーの低い点の概要図.	29
4.1	光源方向推定実験の入力画像.	31
4.2	$k = 1.0$ でレンダリングした入力画像例.	33
4.3	画像 5 枚のサブセットでの実験結果.	35
4.4	画像 6 枚のサブセットでの実験結果.	36
4.5	画像 6 枚のサブセットでの実験結果.	37
4.6	$k = 0.9$ でレンダリングした入力画像例.	38
4.7	Minnaert 則と Lambert 則での復元結果.	40

表 目 次

2.1	Multi-view photometric stereo の関連研究	13
4.1	図 4.1a を入力とした光源方向推定の実験結果	32
4.2	図 4.1b を入力とした光源方向推定の実験結果	32
4.3	エネルギーの閾値 ϵ とサブセットの枚数を変えた実験結果	34
4.4	Minnaert 則と Lambert 則での比較	39

第1章 はじめに

1.1 はやぶさ2プロジェクト

はやぶさ2プロジェクトとは、宇宙航空研究開発機構 (JAXA) による、小惑星探査機「はやぶさ2」を小惑星「Ryugu(1999JU3)」へ向けて打ち上げ、Ryugu の科学観測及びサンプルリターンを目的としたプロジェクトである [1].

はやぶさ2プロジェクトには、前身となった「はやぶさプロジェクト」がある。はやぶさプロジェクトもはやぶさ2と同様に、対象小惑星「イトカワ」の科学観測とサンプルリターンを目的としたプロジェクトである。また、サンプルリターンという技術の実証をテーマとしており、イオンエンジン、自律航法、微小重力下でのサンプル採取、採取物質の入ったカプセルの地球帰還のテストを兼ねていた [2]. 2003年に打ち上げられたはやぶさは2005年にイトカワに到着、2010年に地球に帰還し、前述のミッションはすべて実行された。はやぶさが持ち帰ったサンプルの解析により、小惑星イトカワの物質構成と形成の歴史が解明された。

はやぶさプロジェクトの経験を受け、次のミッションであるはやぶさ2プロジェクトが現在実行されている。対象小惑星である Ryugu は、表面物質に有機物や水がイトカワより多く含まれていると考えられている天体であり、これらの今でも残されている太陽系が生まれた頃の有機物や水を解析することで、地球の水や生命を構成する物質の起源を解明することがはやぶさ2プロジェクトに期待されている。

はやぶさ2は2014年12月3日に種子島宇宙センター大型ロケット発射上から H-A ロケットで打ち上げられた。2018年6月に小惑星 Ryugu に到着し、2020年12月に地球に帰還する予定である。はやぶさ2の探査機本体ははやぶさとほぼ同じ規模であり、搭載される機器ははやぶさに搭載されていたものを改良したものである。小惑星に到着後、はやぶさ2は以下のミッションを行う。

1. 多バンド可視カメラによる観測
2. 近赤外線分光器による観測

3. 中間赤外カメラによる観測
4. レーザー測距装置による観測
5. 小型ローバー，着陸機の投下
6. 小惑星表面へのタッチダウン，サンプルの採取
7. 人工クレーター作成，タッチダウン，サンプルの採取

多バンド可視カメラ (ONC; Optical Navigation Camera) は前回のはやぶさに搭載されたものとほぼ同じ装置であり，これは小惑星の写真を撮影して科学的な観測を行うだけでなく，探査機から見た小惑星の位置や距離を計測する航法用の機器でもある．ONC は望遠カメラである ONC-T と，広角カメラである ONC-W1, ONC-W2 の 3 種類で構成され，探査機と小惑星との距離が約 20km, 5km, 1km の時に撮影を行う [3]．

小型ローバーと着陸機の投下，及び探査機のタッチダウンの際，投下点を選択する必要がある．このとき，小惑星の 3 次元モデルを用いることで，画像のみの場合よりも正確な判断が行えることが期待される．地球上からの観測結果から Kawakami ら [4] により Ryugu の 3 次元形状が推定されたが，遠距離からの観測データに基づく推定であるため，小惑星に接近した探査機が得た観測データからより高精度な復元を行う必要がある．地球と探査機間の距離による通信時間が大きいいため，将来的には探査機が自動で小惑星の 3 次元形状復元から現在位置の姿勢推定までを行うことが求められている．

1.2 探査機が撮影する小惑星画像の性質

はやぶさプロジェクト及びはやぶさ 2 プロジェクトでの小惑星の撮影環境の概念図を図 1.1 に示す．探査機は小惑星に対して一定の距離を保ちながら航行する．太陽を光源として小惑星を撮影するため，探査機は太陽と小惑星の間に位置し，太陽を背にして小惑星の方を向いて撮影を行う．小惑星は自転と公転をしており，探査機に対して姿勢変化することを利用して様々な姿勢の小惑星を撮影する．

はやぶさがマルチバンド分光カメラ (AMICA; Asteroid Muliti-band Imaging Camera) によって撮影した画像の例を図 1.2 に示す．撮影された画像は 1024×1024 ピクセルの 2 次元画像であり，全長 535m の小惑星イトカワを約 7km の位置から撮

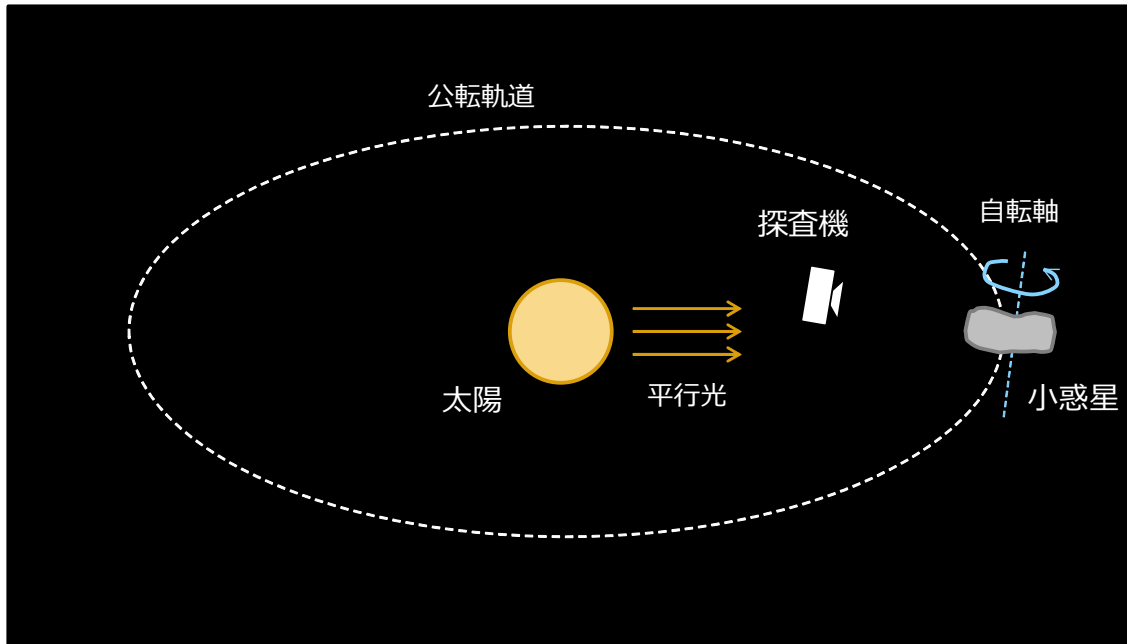


図 1.1: 小惑星の撮影環境の概念図の概念図。探査機は太陽と小惑星の間に位置し、太陽を背にして小惑星の方を向いて撮影を行う。小惑星は自転と公転をしており、探査機に対して姿勢変化することを利用して様々な姿勢の小惑星を撮影する。

影したものである。画像には小惑星のみが写っており、背景領域の画素値は 0 で黒色となる。また、光源となる太陽は十分に遠いため、平行光源とみなせる。AMICA ははやぶさ 2 に搭載されている望遠カメラ ONC-T とほぼ同じものであり、はやぶさ 2 プロジェクトでは全長 900m 程度と推定される Ryugu を 20km 離れた位置から撮影するため、撮影される画像はイトカワに近いものになる。

1.3 複数枚の画像を用いた対象物体の 3 次元復元

3 次元形状の復元は技術と計算機性能の進歩に伴って高精度化・高速化され、研究の成果はコンピュータグラフィクス (CG)、拡張現実 (AR) といった映像メディアを始め、ロボット工学、医療、教育など様々な分野に応用されている。

3 次元復元手法の一つである多視点ステレオ法では、対象となる物体やシーンを多数の視点から撮影した複数の画像から対象物体の 3 次元形状の復元を行う。現在では、画像のみから大規模な建造物や都市の形状復元が可能であり、インターネット上に存在する大量の画像から形状復元を行う研究も行われている [5]。

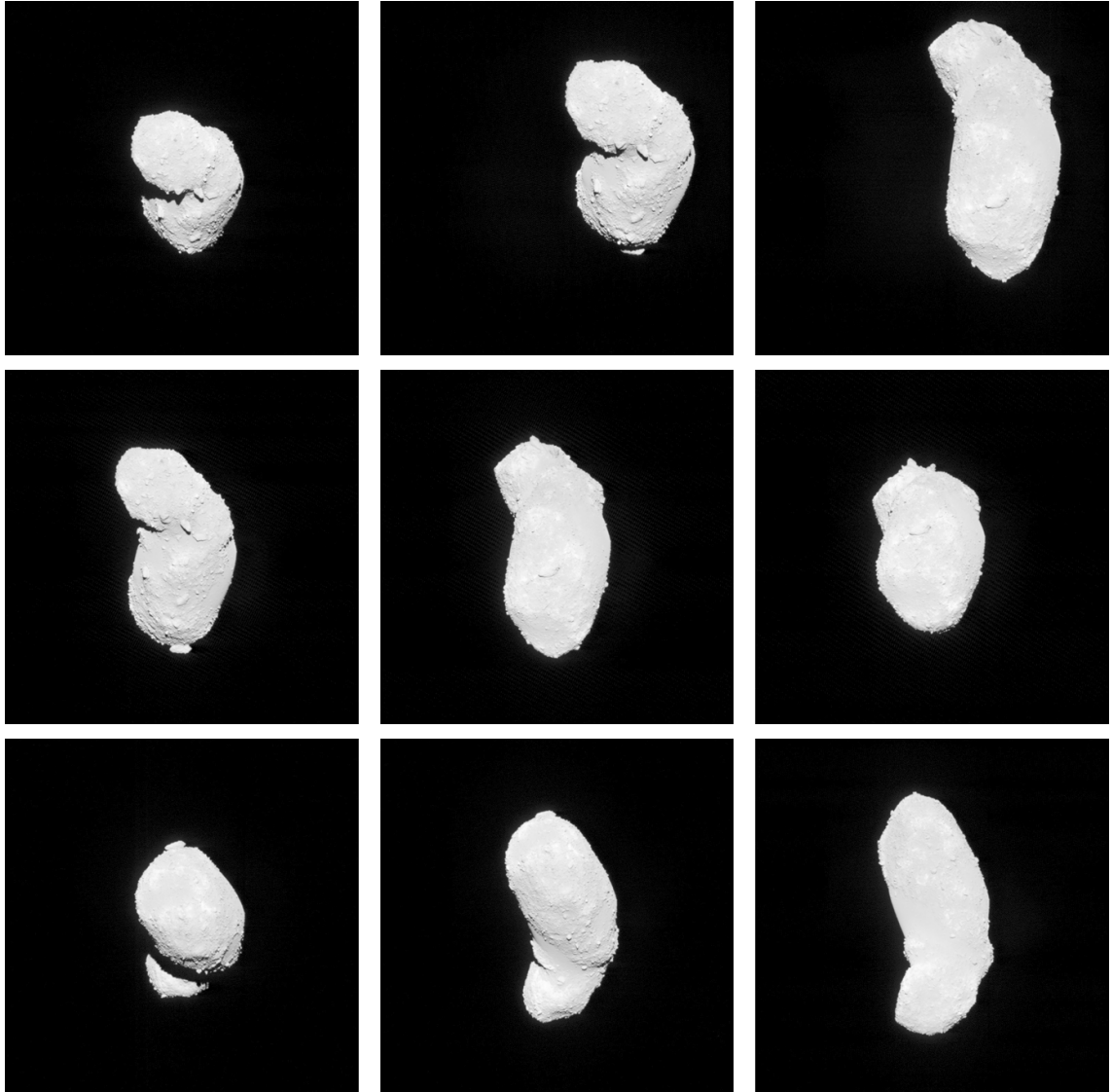


図 1.2: はやぶさがマルチバンド分光カメラで撮影したイトカワの画像例. 画像には小惑星のみが写っている.

はやぶさ 2 プロジェクトで撮影される画像は、探査衛星が光源となる太陽を背にして、移動と自転を行う小惑星を時々刻々と撮影したものとなる。小惑星は光源に対して絶えず姿勢を変化させており、それによってもって各画像に観測される小惑星の表面の明るさも変化する。既存の多視点ステレオ法で用いられる入力画像は、光源に対して静止している物体を異なる視点から撮影したものであり、はやぶさ 2 プロジェクトで撮影される画像はこの仮定を満たさない。そのため、小惑星の 3 次元復元に適した手法が必要である。

1.4 研究目的

本研究では、小惑星を撮影した画像列のみから自動で対象の 3 次元形状を復元する手法の提案を目的とする。

単一のカメラで複数枚の小惑星の画像を撮影した場合、小惑星が移動と自転を行っているため、撮影される画像は視点ごとに小惑星に対するカメラの位置・方向と光源方向が変化したものとなる。視点ごとにカメラの姿勢が異なる画像列を用いた 3 次元復元手法は「多視点ステレオ法」が代表的である。この手法は、物体表面上の点が画像上で同じ輝度で観測されることを仮定しており、視点ごとに物体に対する光源方向が変化して物体表面の見えが変化する本研究の対象シーンには適さない。そのため、視点ごとに物体に対する光源方向が変化することを考慮した 3 次元復元手法を提案する。

また、既存の 3 次元形状復元手法では物体表面の輝度が物体表面の法線方向と光源方向から計算されるランバート反射モデルを仮定しているが、画像上で観測される小惑星の物体表面は視線方向によって観測輝度が変化するため、ランバート反射モデルで正確に表現することができない。本研究では、視線方向を考慮した反射モデルを仮定した 3 次元復元手法を提案する。

提案手法の有用性を示すために、合成画像および実画像を入力に用いた実験を行い、定量的評価を行う。

第2章 関連研究

2.1 複数枚の画像からの3次元復元手法

2.1.1 Multi-view Stereo

多視点ステレオ法 (Multi-View Stereo; MVS) は、異なる視点から撮影した複数の画像を用いて対象物体上の点の3次元位置を推定する手法である [6]。Multi-view stereo で対象とする撮影シーンの概略を図 2.1 に示す。固定された光源下で撮影された複数枚の画像と、撮影したカメラの内部・外部パラメータを用いて、画像上の点の3次元位置を推定する。

物体上のある点が複数枚の画像で観測されているとき、その点の各画像上での座標を求めて対応付ける。対応を求めるためには、SIFT[8] や SURF[9] といった特徴量による特徴点の対応付けを行う。対応点の探索領域を限定するためにエピポラ拘束が用いられる。得られた対応から、三角測量 (triangulation) を用いて特徴点の各画像からの奥行きを推定し、3次元位置を決定する。

点の3次元位置 \mathbf{x} と画像上の座標 \mathbf{u} との関係は、画像を撮影したカメラの内部パラメータ \mathbf{K} と外部パラメータ $[\mathbf{R}|\mathbf{t}]$ を用いて、以下の式で表される。

$$\mathbf{u} = \mathbf{K}[\mathbf{R}|\mathbf{t}]\mathbf{x} \quad (2.1)$$

点 \mathbf{x} が画像インデックスのセット \mathcal{V} で観測され、カメラパラメータ $\mathbf{K}_i[\mathbf{R}_i|\mathbf{t}_i]$ を持つ画像 i 上の点 \mathbf{u}_i に投影されるとき、推定される3次元位置 \mathbf{x}^* は次の最小化問題で表される。

$$\mathbf{x}^* = \arg \min_{\mathbf{x}} \sum_{i \in \mathcal{V}} \|\mathbf{u}_i - \mathbf{K}_i[\mathbf{R}_i|\mathbf{t}_i]\mathbf{x}\| \quad (2.2)$$

この最小化を対応する各特徴点の組について行い、対象物体を3次元復元した点群を推定する。

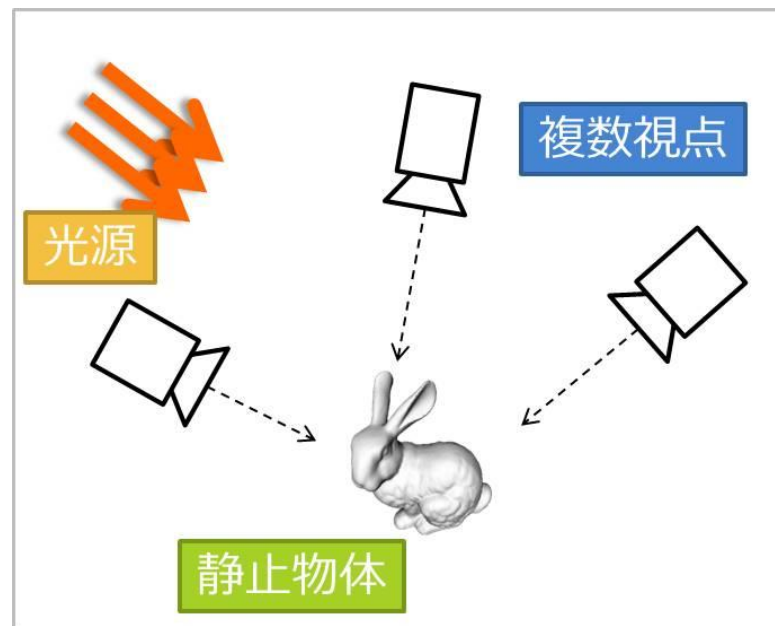


図 2.1: Multi-view stereo で対象とする撮影シーンの概略図. 固定された光源環境下で静止している対象物体を複数の視点で撮影する. 画像に用いたモデルは stanford computer graphics laboratory の “stanford bunny” [7].

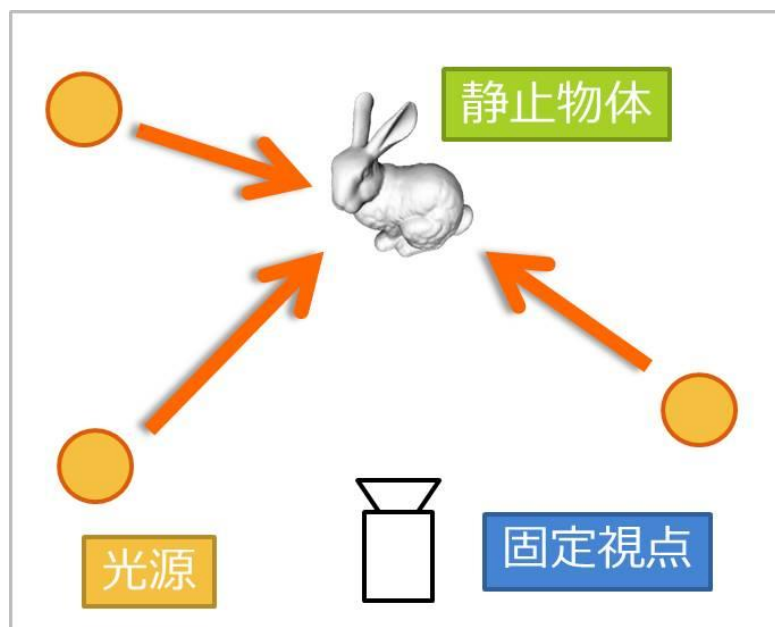


図 2.2: Photometric stereo の対象とする撮影シーンの概略図. 静止している対象物体を同じ視点から異なる照明環境下で撮影する.

Multi-view stereo は奥行き の推定に優れ、また対象物体の全体が写っている画像列を入力として与えれば、対象物体の全体の復元が可能である。しかし、物体表面の復元結果の微細さでは 2.1.2 章に述べる photometric stereo に劣る。

2.1.2 Photometric Stereo

照度差ステレオ法 (Photometric Stereo; PS) は、対象物体を異なる照明環境下で撮影した画像から、物体表面の法線を推定する手法である。Photometric stereo では、対象物体とカメラを固定し、異なる単一平行光源下で撮影した複数枚の画像を用いる。Photometric stereo の対象とする撮影シーンの概略を図 2.2 に示す。

物体の表面がランバート反射であること、光源強度が 1 であることを仮定したとき、物体上の点の輝度値 b は、表面反射率 ρ 、法線ベクトル \mathbf{n} 、光源方向ベクトル \mathbf{l} によって以下の式で表される。ここで、 \mathbf{n} と \mathbf{l} はノルムが 1 の方向ベクトルである。

$$b = \rho \mathbf{n}^\top \mathbf{l} \quad (2.3)$$

大きさが ρ で方向が \mathbf{n} のベクトル $\mathbf{N} = \rho \mathbf{n}$ を定義し、 \mathbf{N} を推定する。 \mathbf{N} の自由度は 3 であるため、 \mathbf{N} を求めるために必要な方向既知の異なる光源下で撮影された画像は 3 枚以上である。光源方向がそれぞれ $\mathbf{l}_1, \mathbf{l}_2, \mathbf{l}_3$ の 3 枚の画像の同一座標での輝度値がそれぞれ b_1, b_2, b_3 のとき、以下の関係を満たす。

$$\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} \mathbf{l}_1^\top \\ \mathbf{l}_2^\top \\ \mathbf{l}_3^\top \end{bmatrix} \mathbf{N} \quad (2.4)$$

光源ベクトル行列の逆行列を計算することで次のように変形できる。

$$\mathbf{N} = \begin{bmatrix} \mathbf{l}_1^\top \\ \mathbf{l}_2^\top \\ \mathbf{l}_3^\top \end{bmatrix}^{-1} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \quad (2.5)$$

最後に、 \mathbf{N} をノルム ρ と単位ベクトル \mathbf{n} に分解し、各ピクセルにおける法線方向を求める。

$$\rho = \|\mathbf{N}\| \quad (2.6)$$

$$\mathbf{n} = \frac{\mathbf{N}}{\|\mathbf{N}\|} \quad (2.7)$$

Photometric Stereo の特徴は，物体表面の法線を各ピクセルごとに計算する点であり，点群の 3 次元位置を推定する multi-view stereo に比べて物体表面の形状をより微細に推定することができる．その反面，photometric stereo で推定するのは面の傾きであるため，multi-view stereo のように物体の奥行きを直接求めることはできない．

2.1.3 Multi-view Photometric Stereo

光源情報を取り扱った多視点ステレオ法は多視点照度差ステレオ法 (Multi-View Photometric Stereo; MVPS) と呼ばれる．MVPS に関する研究の目的は大きく 2 つに分けられる．一つは，光源やカメラパラメータが校正済みであることを想定し，多視点ステレオ法と照度差ステレオ法の双方の長所を取り入れた高精度な 3 次元形状復元を行うことを目的としたものである．もう一つは，対象シーンが異なるために既存手法では困難な入力画像からの 3 次元形状復元を目的としたものである．既存研究の相違点を表 2.1 にまとめる．

表 2.1: Multi-view photometric stereo の関連研究

第一著者	光源位置	光源数	備考
Hernandez [10]	未知	1	
Zhou [11]	未知	1	同じ視点からの照度差画像を多視点分用いる
Higo [12]	既知	1	光源付き手持ちカメラ
Yoshiyasu [13]	既知	1	光源付き手持ちカメラ
Li [14]	既知	1	非ランバート反射物体
Wu [15]	既知	多数	
Fu [16]	既知	多数	運動している対象物体

本研究の対象とするシーンでは，地上からの小惑星の観測のみで正確な光源方向と物体に対するカメラの位置姿勢を得ることは難しい．また，物体とカメラの相対位置を固定して光源方向のみを変化させた画像を得ることも困難である．

Hernández らの手法 [10] は，画像ごとに異なる方向未知の単一平行光源下で撮影した画像を対象としている．想定する画像列は回転台の上に乗った対象物体を固定

したカメラで撮影したもので、1 フレーム毎に物体の姿勢と光源方向を変化させて撮影する。対象物体の表面が Lambert 則に従うことを仮定しているが、鏡面反射成分による画像上のハイライトの影響はある程度取り除くことができる。また、画像から正確な輪郭が得られること、カメラと物体の回転軸との相対位置が固定であることを仮定している。

Hernández らの手法は、カメラパラメータの推定、光源方向の推定、初期形状の推定、3次元形状の最適化の順に行われる。最適化には2種類の単位法線方向ベクトル \mathbf{v}_f と \mathbf{n}_f を用いる。 \mathbf{v}_f はメッシュの頂点の位置から算出されるメッシュの法線ベクトルであり、 \mathbf{n}_f は Lambert 則により輝度から算出されるメッシュの法線ベクトルである。 $f \in F$ はメッシュのインデックス、 $i \in I$ は画像のインデックスである。この2つの法線方向ベクトルが一致するようにエネルギー最小化によるメッシュの頂点 $\mathbf{x}_1 \dots \mathbf{x}_J$ の位置の最適化を行う。

最適化は、以下の2つのエネルギー関数 E_v , E_m の最小化を、収束するまで交互に行う。

$$E_v(\mathbf{n}_{1,\dots,F}, \rho_{1,\dots,F}; \mathbf{x}_{1,\dots,J}) = \sum_{f=1}^F \sum_{i \in I} (\mathbf{l}_i^\top \rho_f \mathbf{n}_f - b_{f,i})^2 \quad (2.8)$$

$$E_m(\mathbf{x}_{1,\dots,J}; \mathbf{n}_{1,\dots,F}) = \sum_{f=1}^F \|\mathbf{v}_f - \mathbf{n}_f\|^2 A_f \quad (2.9)$$

A_f はメッシュ f の面積である。 E_v は、観測輝度 $b_{f,i}$ と光源方向ベクトル \mathbf{s}_i から、 i 番目の画像におけるメッシュ f の持つ法線方向ベクトル \mathbf{n} と表面反射率 ρ を最適化する。各メッシュごとに以下の式を解くことでエネルギー関数を最小化する。

$$\mathbf{n}_f, \rho_f = \arg \min_{\mathbf{n}, \rho} \sum_{i \in I} (\mathbf{l}_i^\top \rho \mathbf{n} - b_{f,i})^2 \quad (2.10)$$

E_m は、式 2.8 により求めた法線方向ベクトル \mathbf{n}_f と、頂点の位置から算出される法線方向ベクトル \mathbf{v}_f の残差が最小となるように頂点の位置 \mathbf{x}_j を最適化する。式 2.9 のエネルギー最小化は最急降下法で行う。

Hernández らの手法では、ある一つの軸に則って回転する物体を固定されたカメラで時々刻々と撮影した画像を入力として用いる。しかし本研究の対象物体である小惑星は、自転と公転によって絶えず姿勢変動と移動をしており、Hernández らの手法の仮定を完全に満たす入力画像列を得るのは困難である。また、初期形状として visual hull を用いているが、対象物体の輪郭が得られる画像を入力とする必要が

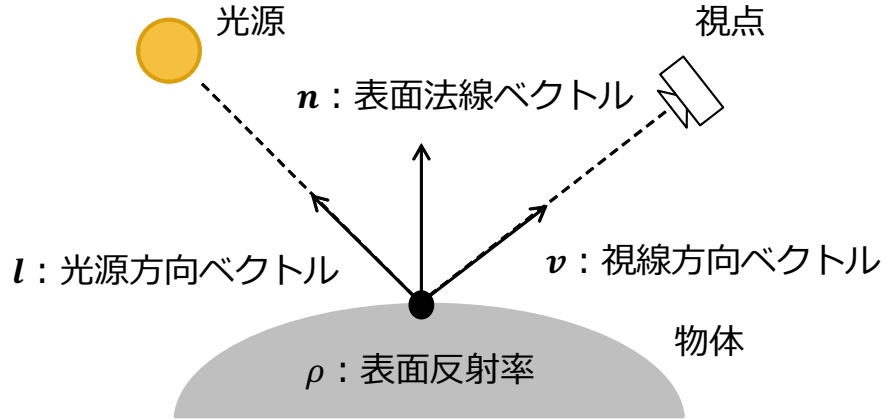


図 2.3: 反射モデルによる輝度の計算に使われるパラメータの概略図. Lambert 則は輝度 b の計算に表面反射率 ρ , 表面法線ベクトル \mathbf{n} , 光源方向ベクトル \mathbf{l} が用いられる. Minnaert 則ではこれらに加えて視線方向ベクトル \mathbf{v} を用いる.

ある. そのため, 本研究ではこの仮定を必要としない, より一般的な画像を入力とする手法を提案する.

2.2 物体表面の反射モデル

2.2.1 Lambert 則

Lambert 則による輝度の計算に使われるパラメータの概略図を図 2.3 に示す. 物体の表面が Lambert 則であること, 光源強度が 1 であることを仮定したとき, 物体上の点の輝度値 b は, 表面反射率 ρ , 法線ベクトル \mathbf{n} , 光源方向ベクトル \mathbf{l} によって以下の式で表される. \mathbf{n} と \mathbf{l} はノルムが 1 の方向ベクトルである.

$$b = \rho \mathbf{n}^T \mathbf{l} \quad (2.11)$$

この反射モデルで計算される輝度は, 物体表面の粗さや視線方向によらない.

2.2.2 Minnaert 則

本研究の対象物体である小惑星の物体表面はざらざらとした粗面である. 粗面を撮影した場合, 一つのピクセルに複数の異なる法線を持つ物体表面が観測される. 平面と粗面がピクセルに観測されたときの違いの概略図を図 2.4 に示す. このよう

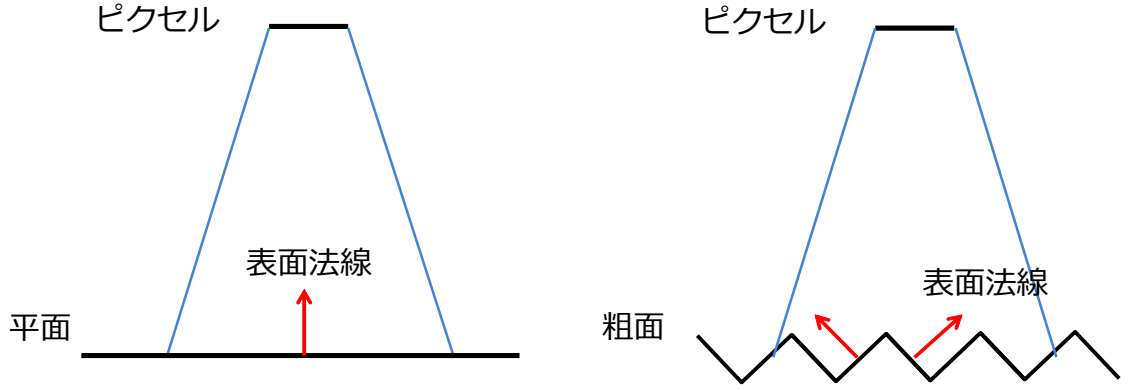


図 2.4: 平面と粗面がピクセルに観測されたときの違いの概略図. (左) 平面がピクセルに観測される場合, 観測されるのは同じ法線方向の面のみ. (右) 粗面がピクセルに観測される場合, 異なる法線方向を持つ複数の表面が一つのピクセルに観測される.

な面は視線方向によって観測輝度に変化する. 粗面の観測輝度の視点による変化の概念図を図 2.5 に示す.

多くの既存手法で仮定される Lambert 則では, 観測輝度 b は物体表面の反射率 ρ , 光源方向単位ベクトル \mathbf{l} , 法線ベクトル \mathbf{n} によって次の式で計算される.

$$b = \rho \mathbf{n}^\top \mathbf{l} \quad (2.12)$$

この反射モデルで計算される輝度は視線方向によらないため, 粗面を正確に表現することはできない. Minnaert 則による輝度の計算に使われるパラメータの概略図を図 2.3 に示す. Marcel Minnaert によって提案された Minnaert 則 [17] は, 視線方向単位ベクトル \mathbf{v} と粗さを表す係数 $0 \leq k \leq 1$ によって Lambert 則を拡張したものであり, 粗面の物体表面の輝度を次の式で表す.

$$b = \rho (\mathbf{n}^\top \mathbf{l})^k (\mathbf{n}^\top \mathbf{v})^{k-1} \quad (2.13)$$

$k = 1$ のとき, この式は Lambert 則と同じものになる. 本研究では粗面物体の高精度な 3 次元復元のために, Minnaert 則を仮定した 3 次元形状復元手法を提案する.

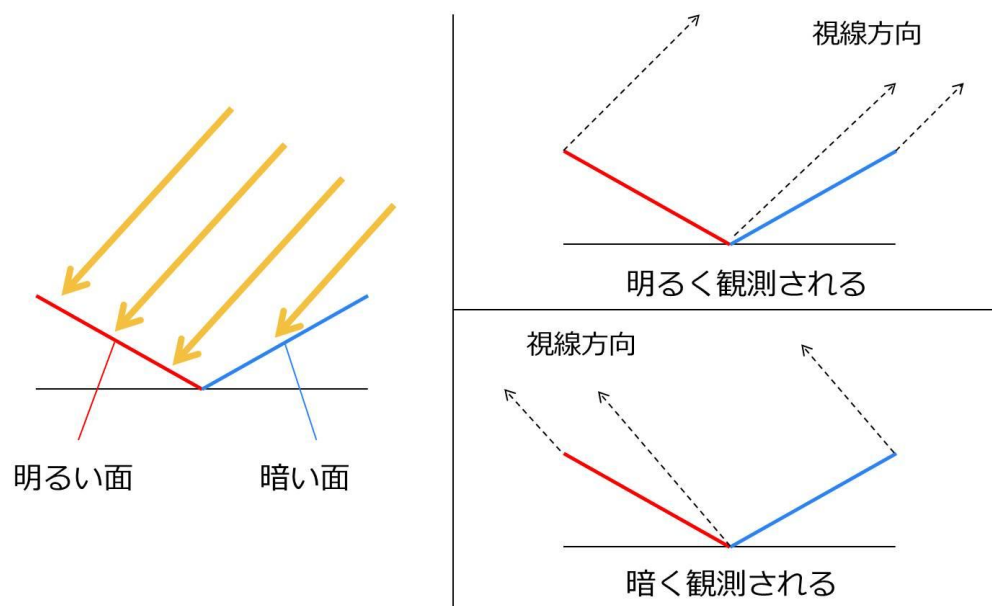


図 2.5: 粗面の観測輝度の視点による変化の概念図. 左のように明るさの異なる面が観測されるとき, 右のように視線方向により観測される面積が変化して観測輝度も変化する.

第3章 提案手法

3.1 手法概要

本研究では、対象物体に対する光源環境の異なる多視点画像から、対象物体の3次元形状と対象シーンの光源方向を推定する手法を提案する。図 3.1 に対象とする撮影環境を示す。固定された単一平行光源によって照らされながら運動している物体を、固定された視点から時々刻々と撮影した画像を入力とする。対象物体の表面反射特性は Minnaert 則に従うと仮定し、Minnaert 係数 k は既知とする。また、対象物体の表面反射率は一定であると仮定する。光源は単一の無限遠平行光源を仮定する。本手法の最終的な出力は対象物体の表面の3次元点群である。

本研究の対象シーンでは、カメラと光源環境が固定されているため、対象物体の姿勢が変化すると物体に対する光源環境は変化する。そのため、物体に対する光源環境が固定であることを仮定している既存の multi-view stereo による3次元復元は精度が落ちる。そこで本研究の提案手法ではまず、画像ごとに変化する物体に対する光源環境を考慮するために光源方向の推定を行う。その後、multi-view stereo の出力の3次元形状にたいして光源方向を考慮した最適化を行い、より精度の高い復元結果を得る。

本手法は次の流れで行う。

1. Structure-from-motion によるカメラパラメータの推定
2. Multi-view stereo による3次元点の位置と法線の初期推定
3. シーン中の光源方向と物体の表面反射率の推定
4. 3次元点の位置と法線の最適化

この内、1. Structure-from-motion によるカメラパラメータの推定と 2. Multi-view stereo による3次元点の位置と法線の初期推定については既存手法を用いる。本研究では、入力画像と推定されたカメラパラメータ、3次元点の位置と法線から、カメ

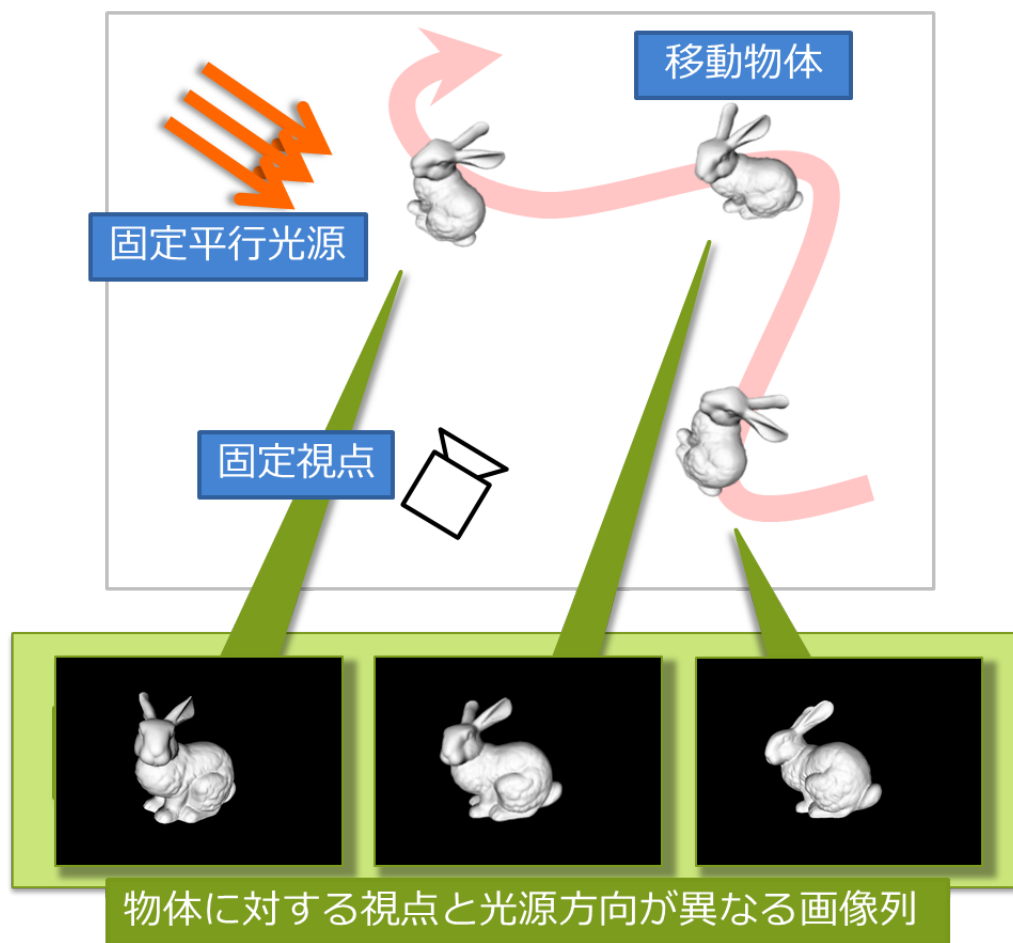


図 3.1: 撮影シーンの概略図. 運動する対象物体を固定されたカメラで撮影する.

ラに対する光源方向と物体の表面反射率を推定し、3次元点の位置と法線を最適化する手法を提案する。画像を入力とし、3次元形状の最適化を行う提案手法全体の流れを図 3.2 に示す。

入力画像の番号の集合を \mathcal{I} とし、入力画像の番号を $i \in \mathcal{I}$ とする。入力画像は $I_i, i \in \mathcal{I}$ である。カメラの内部パラメータは各画像間で共通であり、 3×3 行列 \mathbf{K} で表す。各画像における物体の姿勢はカメラの外部パラメータとみなすことができ、向きを 3×3 行列 \mathbf{R}_i 、位置を 3次元ベクトル \mathbf{t}_i で表す。光源方向はカメラに対して固定であるため、カメラ座標系における光源方向を 3次元ベクトル \mathbf{l} で表す。3次元点の番号の集合を \mathcal{P} とし、3次元点の番号を $p \in \mathcal{P}$ とする。また、最終的な復元点群の集合を \mathcal{P}^* とする。各 3次元点 p は 3次元位置 \mathbf{x}_p と法線 \mathbf{n}_p を持つ。物体の表面反射率は一様であることを仮定し、 ρ で表す。 ρ は $0 \leq \rho \leq 1$ の範囲を取り、入射光に対する反射光の強さの割合を示す。画像 \mathcal{I} の内、ある点 p が観測されるものの集合を \mathcal{V}_p とする。 i 番目の画像上での 3次元点 p の観測輝度を $b_{p,i}$ で表す。

まずはじめに 3.2 節で光源方向推定方法について述べ、次に 3.3 節で 3次元形状の最適化について述べる。

3.2 光源方向と表面反射率の推定

提案する光源方向推定法は、物体上の点 p の 3次元位置 \mathbf{x}_p 、表面法線ベクトル \mathbf{n}_p 、カメラの視線を表す外部パラメータ \mathbf{R}_i 及び入力画像 I_i 上で観測される輝度値 b_i を用いて、Minnaert 則を元に光源方向 \mathbf{l} と表面反射率 ρ を推定する。推定する光源方向はカメラ座標系におけるものであり、本手法ではカメラに対する光源環境が固定であることを仮定しているため、カメラ座標系における光源方向は各視点において同じとみなせる。

対象物体上の点 \mathbf{x}_p が画像 i 上において観測される輝度 $b_{p,i}$ は

$$b_{p,i} = \rho [(\mathbf{R}_i \mathbf{n}_p)^\top \mathbf{l}]^k [(\mathbf{R}_i \mathbf{n}_p)^\top \mathbf{v}_{p,i}]^{k-1} \quad (3.1)$$

となる。ここで、 $\mathbf{R}_i \mathbf{n}_p$ はカメラ座標系における \mathbf{n}_p 、 $\mathbf{v}_{p,i}$ は点 p から見た視点 i の方向ベクトルである。

$\mathbf{s} = \rho^{1/k} \mathbf{l}$ と置くと、上述の式は次のように変形できる。

$$(\mathbf{R}_i \mathbf{n}_p)^\top \mathbf{s} - \left[\frac{b_{p,i}}{(\mathbf{R}_i \mathbf{n}_p)^\top \mathbf{v}_{p,i}} \right]^{\frac{1}{k}} \quad (3.2)$$

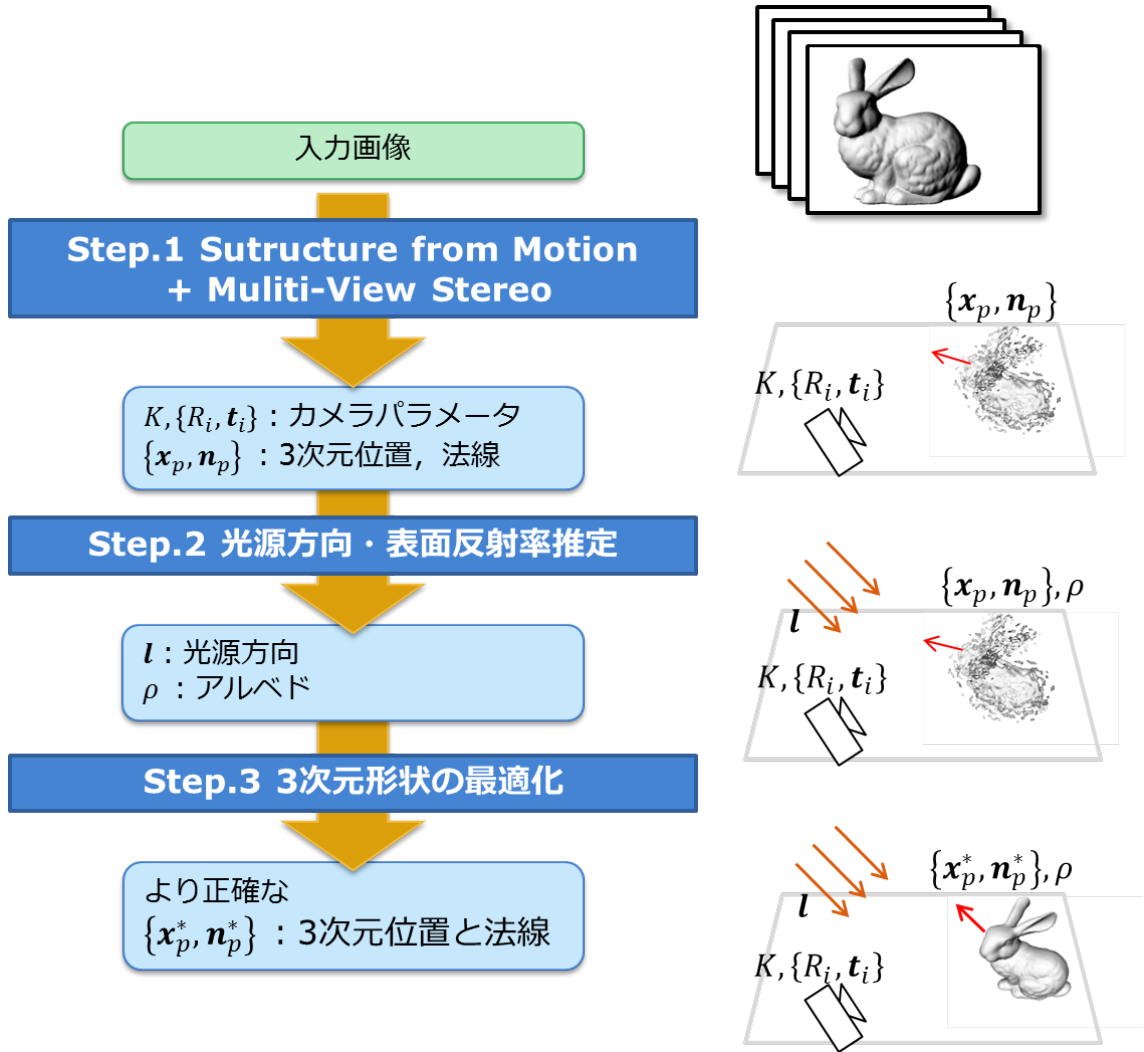


図 3.2: 手法全体の流れの概要図. Step.1 既存手法である Structure from Motion と多視点ステレオ法から, カメラパラメータと対象の3次元点の位置と法線方向が得られる. Step.2 提案する光源方向推定により, 光源方向と表面反射率を得る. Step.3 提案する最適化手法により, 3次元点の位置と法線方向を最適化する.

対象物体上の点 $p \in \mathcal{P}$ が画像 $i \in \mathcal{V}_p$ 上で観測される時、次の最小二乗問題の解として光源ベクトル \mathbf{s} を推定する.

$$\mathbf{s}^* = \arg \min_{\mathbf{s}} \sum_{p \in \mathcal{P}} \sum_{i \in \mathcal{V}_p} \left\| (\mathbf{R}_i \mathbf{n}_p)^\top \mathbf{s} - \left[\frac{b_{p,i}}{(\mathbf{R}_i \mathbf{n}_p)^\top \mathbf{v}_{p,i}} \right]^{\frac{1}{k}} \right\|^2 \quad (3.3)$$

光源方向 \mathbf{l} と表面反射率 ρ は, \mathbf{s} のノルムが 1 であることを利用し,

$$\rho^* = \|\mathbf{s}^*\|^k, \quad (3.4)$$

$$\mathbf{l}^* = \frac{\mathbf{s}^*}{\|\mathbf{s}^*\|} \quad (3.5)$$

と推定できる.

3.3 光源方向を考慮した最適化

3.3.1 エネルギー最小化による各推定値の最適化

このステップでは, エネルギー関数 E を最小化することで, これまでに推定したパラメータを最適化し, 最終的な出力となる 3 次元形状を得る.

エネルギー関数は輝度の再投影誤差にエネルギーを与え, 復元点の位置と法線 $\{\mathbf{x}_p, \mathbf{n}_p\}$ について最適化を行う.

$$E(\mathbf{x}_p, \mathbf{n}_p) = \sum_{i \in \mathcal{V}_p} \left\| I_{p,i}^{\text{geo}} - I_{p,i}^{\text{photo}} \right\| \quad (3.6)$$

ここで, $I_{p,i}^{\text{geo}}$ は画像上で観測される 3 次元点の輝度, $I_{p,i}^{\text{photo}}$ は法線と光源方向, 表面反射率から計算される 3 次元点の輝度である.

$$I_{p,i}^{\text{geo}} = I_i(\mathbf{K}[\mathbf{R}_i | \mathbf{t}_i] \mathbf{x}_p) \quad (3.7)$$

$$I_{p,i}^{\text{photo}} = \rho \left[(\mathbf{R}_i \mathbf{n}_p)^\top \mathbf{l} \right]^k \left[(\mathbf{R}_i \mathbf{n}_p)^\top \mathbf{v} \right]^{k-1} \quad (3.8)$$

3.3.2 Visibility の推定と視点の選択

Visibility の推定

本項では、ある物体表面上の点 p が、実際に観測されている画像群 $\mathcal{V}_p \in \mathcal{I}$ を推定する手法について述べる。visibility とは、ある点がある画像において観測されているかどうかの判定であり、 \mathcal{V}_p は点 p について visibility が真となる画像群である。

点 p の画像 i における visibility の推定には、点 p の法線 \mathbf{n}_p と点 p からの視点方向 $\mathbf{n}_{p,i}$ を利用する手法と、点 p が画像 i に投影された点の座標と輝度を利用する手法の 2 通りを組み合わせる。次の 2 つの条件を両方満たした画像の visibility を真とする。

まず、法線 \mathbf{n}_p と視点方向 $\mathbf{v}_{p,i}$ を利用する手法について述べる。法線 \mathbf{n}_p と視点方向 $\mathbf{v}_{p,i}$ の成す角を θ とする。物体表面上の点の法線 \mathbf{n}_p が視点側を向いているときに点 p は観測され、視点側を向いていないときは点 p は観測されない。そのため、 $\theta < 90^\circ$ を満たすとき、visibility を真とする。

次に、点 p が画像 i に投影された点の座標と輝度を利用する手法について述べる。投影点の座標が画像内に収まっていなければ、その画像には観測されないため、visibility を偽とする。また、投影点の輝度を $b_{p,i}$ としたとき、 $b_{p,i}$ が 0 であるときは、その点は画像の背景に投影されているか、物体が自身に落とした影 (Cast shadow) に投影されていると考えられるため、visibility を偽とする。Cast Shadow の例を図 3.3 に示す。

画像の選択

前項の手法によって得られた visibility が真の画像群は、occlusion を考慮していない。Occlusion とは、ある物体表面上の点が、別の物体表面によって遮られ、画像に観測されないことである。Occlusion の概略図を図 3.4 に示す。本来観測される点とは別の輝度を持つ点が観測されるため、最適化手法に悪影響を及ぼす。本項では、occlusion のある画像を取り除き、より良い \mathcal{V}_p を求める手法を提案する。

Visibility の推定によって visibility が真と判定された画像群を \mathcal{V}_p^* と定義する。 \mathcal{V}_p^* は occlusion のある画像を含んでいる。Occlusion のある画像を含んだ画像群を用いて式 (3.6) による最適化を行った場合、式 (3.6) は輝度の残差をエネルギーとしているため、エネルギーは大きくなる。これを利用し、 \mathcal{V}_p^* のサブセットを複数作成し、それぞれのサブセットについて式 (3.6) による最適化を行い、最もエネルギーが

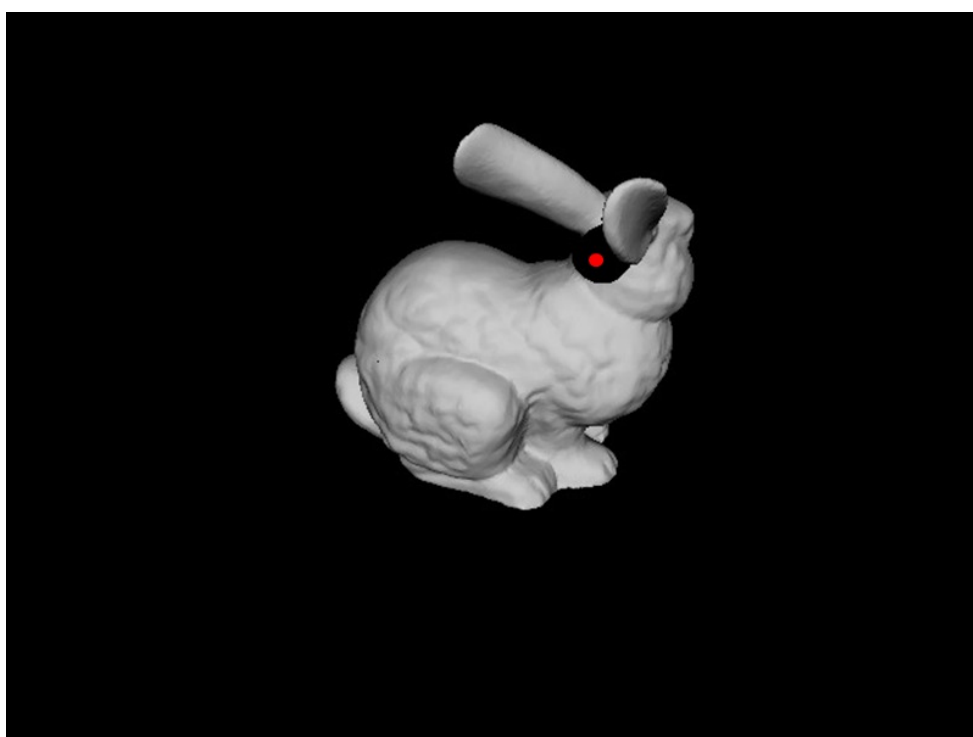


図 3.3: Cast shadow の例. 赤で示した点は物体表面を観測しているが, 物体が自身に落とした影によって輝度が本来より低く観測されている.

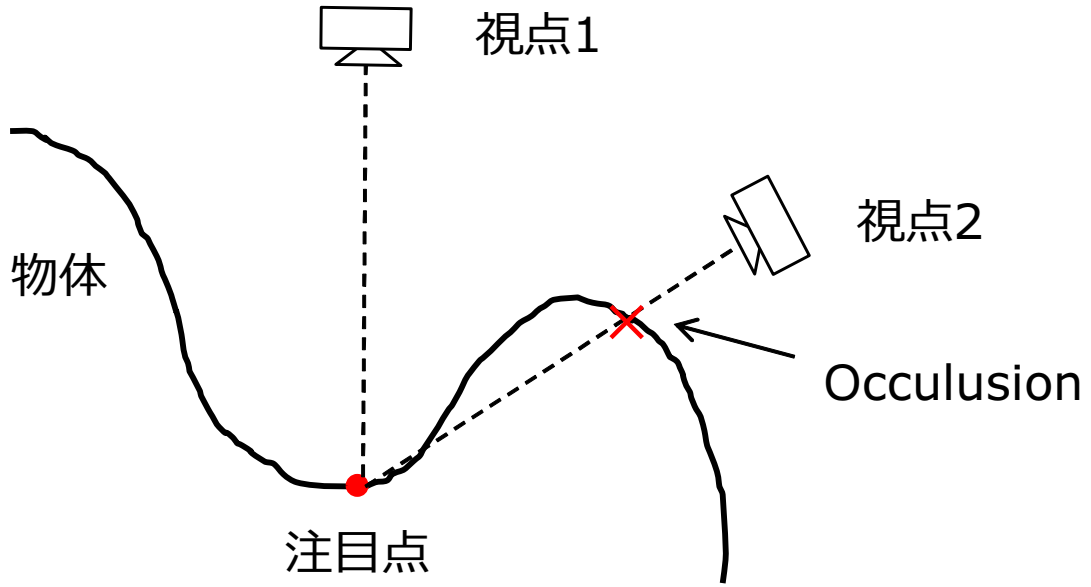


図 3.4: Occlusion の概略図. 視点 1 は正しく注目点を観測しているが, 視点 2 は別の物体表面によって視線が遮られ, 注目点を観測していない.

低くなったサブセットを occlusion のある画像を含まない \mathcal{V}_p として採用する. 視点選択の概念図を図 3.5 に示す.

3.3.3 推定表面に沿った点の追加と外れ値除去

復元された 3 次元点の数が多いほど, 物体表面を詳細に復元できると考えられる. そのため, 最適化後のエネルギーが十分低くなった状態で収束した 3 次元点に対し, 周りに 3 次元点を増やすことで, より密な 3 次元復元結果を得る. 増やした点についても同様の最適化を行う.

また, 最適化によってエネルギーが低い状態で収束したとしても, 3 次元点が物体表面から離れた位置でローカルミニマムに収束してしまい, 復元結果に悪影響を与える場合がある. この問題を解決するために, 増やした点の収束後のエネルギーをもとに, 増やす元となったローカルミニマムの点を取り除く手法を提案する. 点の追加による最適化手法の概要をに示す.

注目点 p について考える. p は初期位置 \mathbf{x}_p と初期法線 \mathbf{n}_p を持ち, 式 (3.6) のエネルギー最小化によって位置 \mathbf{x}_p^* と法線 \mathbf{n}_p^* に最適化されたとする. このとき, p が最適化後に物体表面にあればエネルギーが比較的低い状態で収束し, 物体表面から離

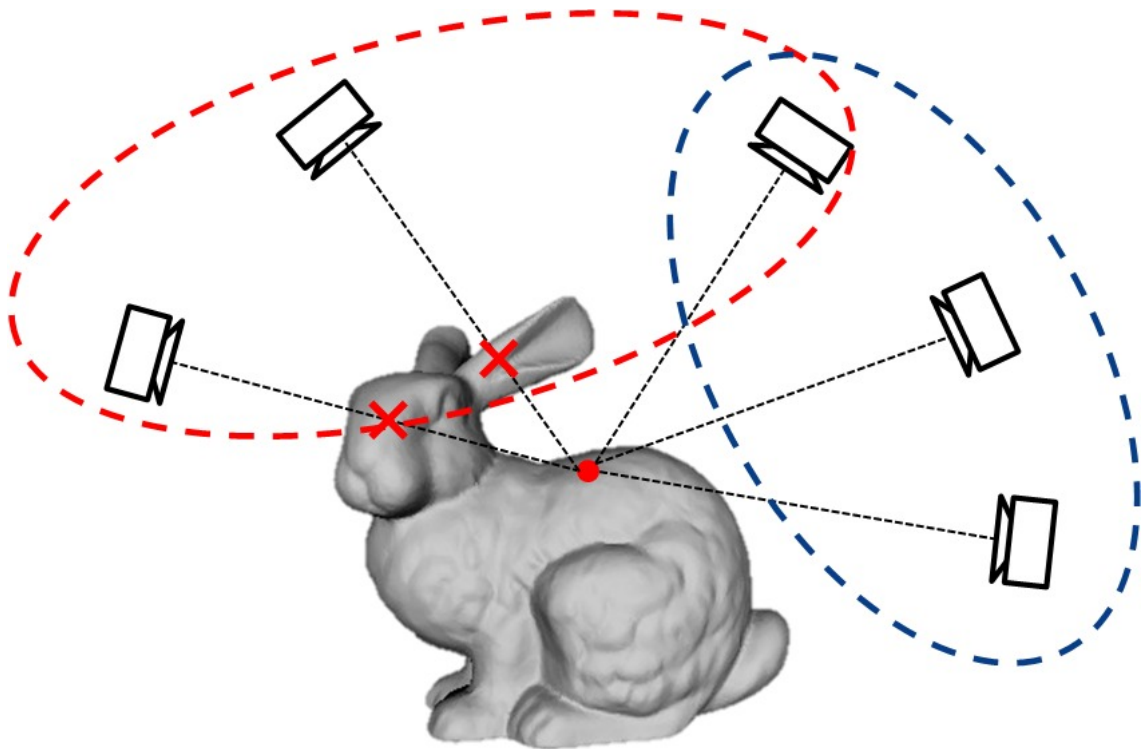


図 3.5: 視点選択の概念図. 図中の 5 つの視点のうち, 赤で囲まれたサブセットは occlusion のある画像を含み, 青で囲まれたサブセットは occlusion のある画像を含まない. 赤のサブセットはエネルギーが高くなり, 青のサブセットは低くなることが期待される.

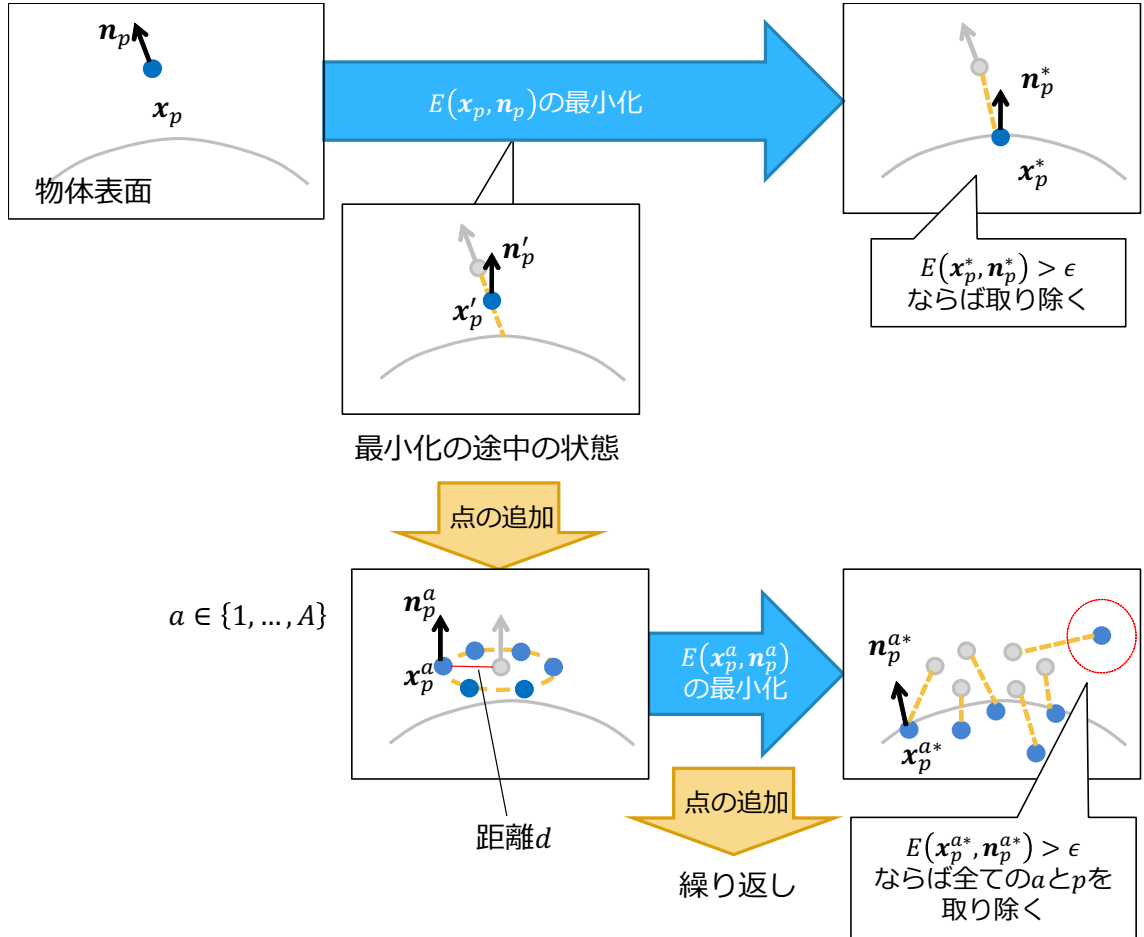


図 3.6: 最適化手法の流れの概要図. まず点 p について $E(x_p, n_p)$ の最小化を行う. 最小化後のエネルギー $E(x_p^*, n_p^*)$ が閾値 ϵ 以上なら取り除き, 以下ならその点周りに点 $a \in \{1, \dots, A\}$ を追加する. 追加した点 a についても最小化を行う. 全ての $a \in \{1, \dots, A\}$ について $E(x_p^{a*}, n_p^{a*}) < \epsilon$ ならば p を復元結果とし, そうでなければ p を取り除く.

れていれば高い状態で収束すると考えられる．ここで，エネルギーの閾値 ϵ を定め，収束後の p のエネルギー $E(\mathbf{x}_p^*, \mathbf{n}_p^*)$ が $E(\mathbf{x}_p^*, \mathbf{n}_p^*) < \epsilon$ を満たさなければ，外れ値として復元点群 \mathcal{P} から取り除く．

$E(\mathbf{x}_p^*, \mathbf{n}_p^*) < \epsilon$ を満たした場合， p を元に点を複数増やす．このとき，最適化後の $\mathbf{x}_p^*, \mathbf{n}_p^*$ ではなく，最適化途中の状態の位置と法線 $\mathbf{x}'_p, \mathbf{n}'_p$ を利用する．これは，まだ物体表面にない点を元に点を増やすことで，増やした点を最適化した際にそれぞれが p から離れた位置に収束することと，本項で後述するローカルミニマムを取り除くステップで最適化した際に大きく移動した点を取り除かれやすくなることを期待している． p について A 個の点を増やすとき，増やした点のインデックスを $a \in \{1, \dots, A\}$ とし， a の初期位置と初期法線を $\mathbf{x}_p^a, \mathbf{n}_p^a$ とする． \mathbf{x}_p^a は， \mathbf{n}'_p が定義する平面上の， \mathbf{x}'_p から距離 d だけ離れた位置に， A 個の追加点が等間隔になるように決定する． \mathbf{n}_p^a は \mathbf{n}'_p と同じ方向ベクトルに設定する．

次に，追加した点 $a \in \{1, \dots, A\}$ のそれぞれの位置と法線 $\mathbf{x}_p^a, \mathbf{n}_p^a$ について式 (3.6) のエネルギー最小化を行う．最適化後の $\mathbf{x}_p^a, \mathbf{n}_p^a$ を $\mathbf{x}_p^{a*}, \mathbf{n}_p^{a*}$ とする．このとき，元の点 p がローカルミニマムで収束していた場合，周りに増やした点 a はローカルミニマムに収束せず，エネルギー $E(\mathbf{x}_p^*, \mathbf{n}_p^*)$ が比較的高い状態で収束することが期待される．そのため， $a \in \{1, \dots, A\}$ のうち一つでも $E(\mathbf{x}_p^*, \mathbf{n}_p^*) < \epsilon$ を満たさなければ，元の点 p をローカルミニマムに収束していた外れ値として復元点群 \mathcal{P} から取り除く．

全ての $a \in \{1, \dots, A\}$ について $E(\mathbf{x}_p^*, \mathbf{n}_p^*) < \epsilon$ が満たされていれば，元となった点 p を最終的な復元点の集合 \mathcal{P}^* に追加する．これを全ての点 $p \in \mathcal{P}$ について行い，残った $p \in \mathcal{P}$ の $a \in \{1, \dots, A\}$ の集合を新たな \mathcal{P} として，これまでの処理をあらかじめ決めた回数 M 回だけ繰り返す．

3.3.4 背景に投影される位置に移動した点の除去

視線方向と光源方向が物体の回転軸に直交している場合，物体表面から離れた位置に低いエネルギーで収束する場合がある．エネルギー最小化によって，ある点 p が，画像 $i \in \mathcal{V}_p$ 上に投影されたときの輝度がすべて 0 となる位置に移動し，物体の回転軸と平行な法線になった場合， $I_{p,i}^{\text{geo}}$ と $I_{p,i}^{\text{photo}}$ の両方が 0 になり，その残差を評価するエネルギー関数式 (3.6) の値も 0 となるためである．この場合，点 p の近傍にも条件を満たす位置と法線の組が存在するため，3.3.3 節の手法で取り除くことはできない．この状態の点の概念図を図 3.7 に示す．

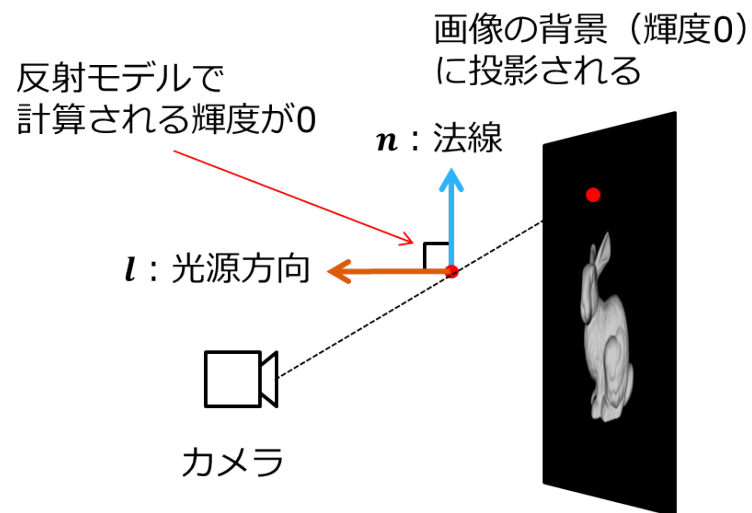


図 3.7: 背景に投影されるエネルギーの低い点の概要図. 画像上の投影点の輝度も, 反射モデルで計算される輝度も 0 となる.

このような点を取り除くために, 最適化後に画像上の輝度 0 の点に投影される位置にある 3 次元点を取り除く.

第4章 実験

4.1 実装

対象物体の3次元点位置と表面法線方向を得るために、structure from motion のソフトウェアである Bundler [18] と、multi-view stereo のソフトウェアである Patch-based Multi-View Stereo version2 (PMVS2) [19] を用いる。Bundler は画像から抽出した SIFT 特徴量を用いてカメラの内部・外部パラメータを推定するソフトウェアである。PMVS2 は画像とそのカメラの内部・外部パラメータから物体の3次元形状を推定するソフトウェアである。PMVS2 は3次元位置と表面法線方向を持つパッチを用いて推定を行い、通常が多視点ステレオ法で推定される3次元位置に加えて、その点の表面法線方向も推定する。エネルギー最小化には、レーベンバーグ・マーカート法を用いた。

4.2 合成画像を用いた光源方向の推定実験

4.2.1 実験目的

光源方向と物体の表面反射率が既知の合成画像を用いて、光源方向の推定結果を ground truth と比較する。合成画像は Minnaert 則でレンダリングし、これに対して Lambert 則を仮定した推定と Minnaert 則を仮定した推定を行い、Minnaert 則の場合に精度が上がることを確認する。

4.2.2 実験方法

3次元モデルをレンダリングした合成画像1枚を入力として与え、提案手法による光源方向推定を行う。入力画像を図4.1に示す。カメラの視線方向と光源方向が

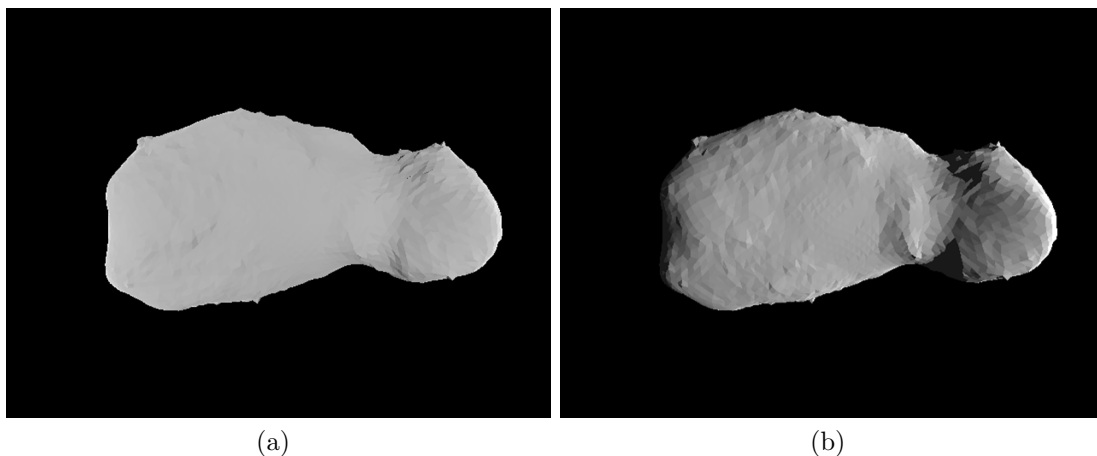


図 4.1: 光源方向推定実験の入力画像. (a) カメラの視線方向と光源方向が同じ場合. (b) 視線方向に対して 45° の角度の光源方向の場合.

同じ場合 (図 4.1a) と, 視線方向に対して 45° の角度の光源方向の場合 (図 4.1b) の 2 種類で行った. 光源方向推定に必要な初期形状として, 正解形状を与える. 入力画像をレンダリングする際の Minnaert 係数は $k = 0.7$ とした. Ground truth の光源方向と推定した光源方向のなす角度で誤差を計測する.

4.2.3 実験結果と考察

図 4.1a を入力とした光源方向推定の実験結果を表 4.1 に, 図 4.1b を入力とした光源方向推定の実験結果を表 4.2 に示す.

正面から光が当たっている場合は, $k = 0.7$ でレンダリングした画像を用いたにもかかわらず, $k = 1.0$ で推定した時が最も誤差が小さくなった. これは, 物体の法線が全方向に満遍なく分布していることと, 複数の法線から一つの光源方向を求めることにより, 各法線における誤差の合計が正面方向に寄るからだと考えられる. 視線方向に対して角度のある光源方向の画像を用いた表 4.2 の場合は, 期待通りに $k = 0.7$ で推定したときの誤差が最も小さくなった.

表 4.1: 図 4.1a を入力とした光源方向推定の実験結果

推定に用いた Minnaert 係数 k	誤差 (degree)
1.0(Lambert)	1.716313315
0.9	2.115396212
0.8	2.745992293
0.7	3.844033651
0.6	6.136557543
0.5	12.90418114

表 4.2: 図 4.1b を入力とした光源方向推定の実験結果

推定に用いた Minnaert 係数 k	誤差 (degree)
1.0(Lambert)	10.4839556
0.9	7.426857619
0.8	3.748863813
0.7	1.106146295
0.6	6.882915639
0.5	14.84275541

4.3 エネルギーの閾値 ϵ とサブセットの枚数を変えた場合の比較

4.3.1 目的

Ground Truth の得られる 3 次元モデルをレンダリングした合成画像を入力として用い、提案手法による復元結果の定量的評価を行う。また、提案手法で用いられるエネルギーの閾値 ϵ と、正しい visibility を求めるために作る画像のサブセット一つ当たりの枚数を変えて実験を行い、この二つの復元結果への影響を確認する。

4.3.2 実験方法

3 次元モデルをレンダリングした画像 60 枚を入力画像として用いた。入力画像例を図 4.2 に示す。モデルは Stanford Bunny を使用した。視線方向と同じ光源方向、表面反射率 $\rho = 1.0$ 、Minnaert 係数 $k = 1.0$ でレンダリングしたものである。対象

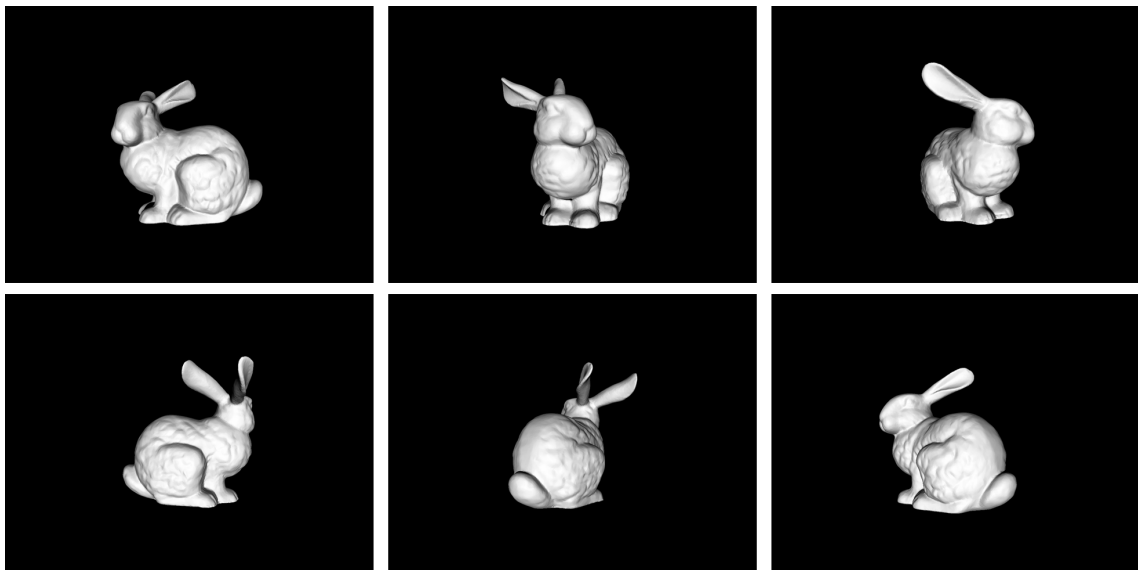


図 4.2: $k = 1.0$ でレンダリングした入力画像例.

物体は一つの軸に沿って姿勢変化している. Ground truth は, 入力画像のレンダリングに用いた 3 次元モデルとする. カメラパラメータは structure from motion での推定結果を用い, 初期形状には multi-view stereo の推定結果を用いた. 光源方向と表面反射率は提案手法による推定結果を用いた.

エネルギーの閾値が $\epsilon = 0, 0001, 0.00001, 0.000001$ の 3 種類, サブセットの枚数が 5 枚, 6 枚, 7 枚の 3 種類の 9 通りの組み合わせで実験を行った. 点の追加は 2 回行った. 追加点と元の点の距離 d を初期点群のスケールから $d = 0.1$ とし, 元の 1 点につき 6 点を等間隔に追加した.

評価方法について述べる. 各手法の復元結果と ground truth を Iterative Closest Point (ICP) によって位置合わせし, その後復元結果の各点から ground truth のメッシュまでのユークリッド距離を求め, その平均を評価する.

4.3.3 実験結果と考察

各実験結果の ground truth までの距離の平均と復元点数を表 4.3 に示す. サブセットの画像枚数が 5, 6, 7 のときの復元結果の画像をそれぞれ図 4.3, 図 4.4, 図 4.5 に示す. 入力画像 60 枚のうち, structure from motion によってカメラパラメータの推定ができたものは 56 枚であった.

初期形状の復元点数は 23171, ground truth までの距離の平均は 0.126195 である.

表 4.3: エネルギーの閾値 ϵ とサブセットの枚数を変えた実験結果

サブセットの枚数	閾値 ϵ	Ground truth までの距離の平均	復元点の数
5	0.0001	0.159024	144743
5	0.00001	0.156506	110733
5	0.000001	0.152997	66209
6	0.0001	0.145082	141774
6	0.00001	0.143076	86456
6	0.000001	0.124688	21408
7	0.0001	0.136057	129438
7	0.00001	0.132758	61815
7	0.000001	0.108579	3698

Ground truth までの距離の平均については提案手法によって大きな改善は見られなかったが、復元点数を増やせている場合が多く、精度を変化させずにより密度の高い復元点群が得られたといえる。

サブセットの枚数を増やすほど ground truth までの距離の平均は小さくなり、物体表面から遠い点を取り除きやすくなることがわかる。しかし、復元点数が少なくなり、復元点の密度も偏っている。これは足の付け根や耳などの凹んでいる部分は観測している画像が少なく、サブセットの枚数に満たないためだと考えられる。サブセット枚数が5枚のときは、他と比べて凹んだ物体表面上の点も残っている。今回の実験ではサブセットの枚数は固定であったが、異なる枚数のサブセットを混合して用いるなど、サブセットの枚数が少ない時と多いときの双方の利点を取り入れることができれば精度の向上が見込める。

4.4 Minnaert 則の対象物体の復元

4.4.1 目的

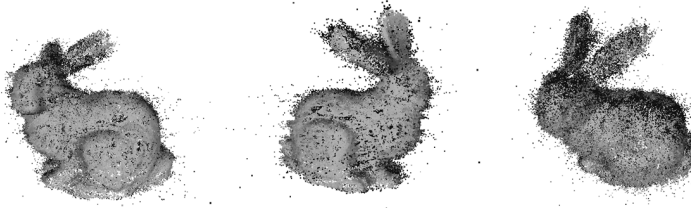
Minnaert 則でレンダリングした合成画像を入力として用い、Minnaert 則を仮定した復元と Lambert 則を仮定した復元を行い、その結果を比較することで Minnaert 則を仮定することの影響を確認する。



(a) Ground truth



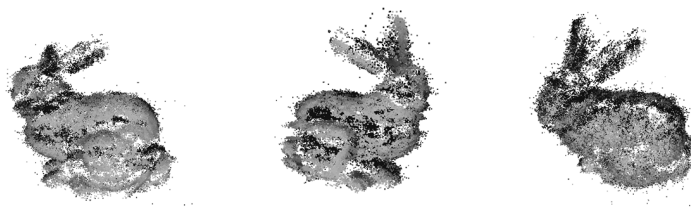
(b) 初期形状 (multi-view stereo)



(c) 提案手法 $\epsilon = 0.0001$



(d) 提案手法 $\epsilon = 0.00001$

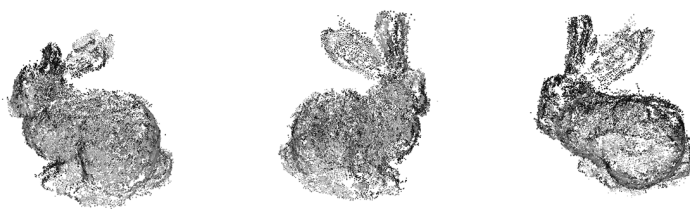


(e) 提案手法 $\epsilon = 0.000001$

図 4.3: 画像 5 枚のサブセットでの実験結果.



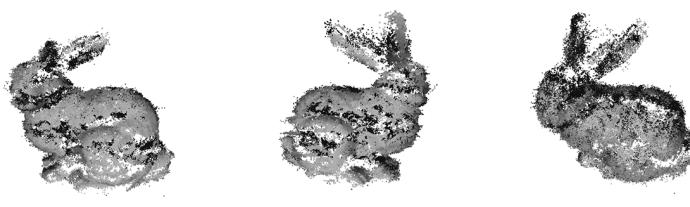
(a) Ground truth



(b) 初期形状 (multi-view stereo)



(c) 提案手法 $\epsilon = 0.0001$



(d) 提案手法 $\epsilon = 0.00001$



(e) 提案手法 $\epsilon = 0.000001$

図 4.4: 画像 6 枚のサブセットでの実験結果.



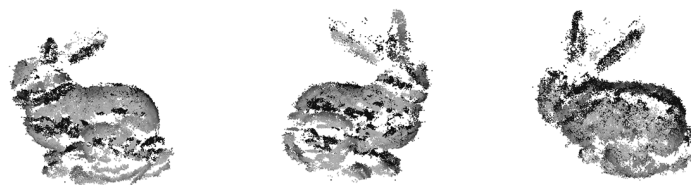
(a) Ground truth



(b) 初期形状 (multi-view stereo)



(c) 提案手法 $\epsilon = 0.0001$



(d) 提案手法 $\epsilon = 0.00001$



(e) 提案手法 $\epsilon = 0.000001$

図 4.5: 画像 7 枚のサブセットでの実験結果.

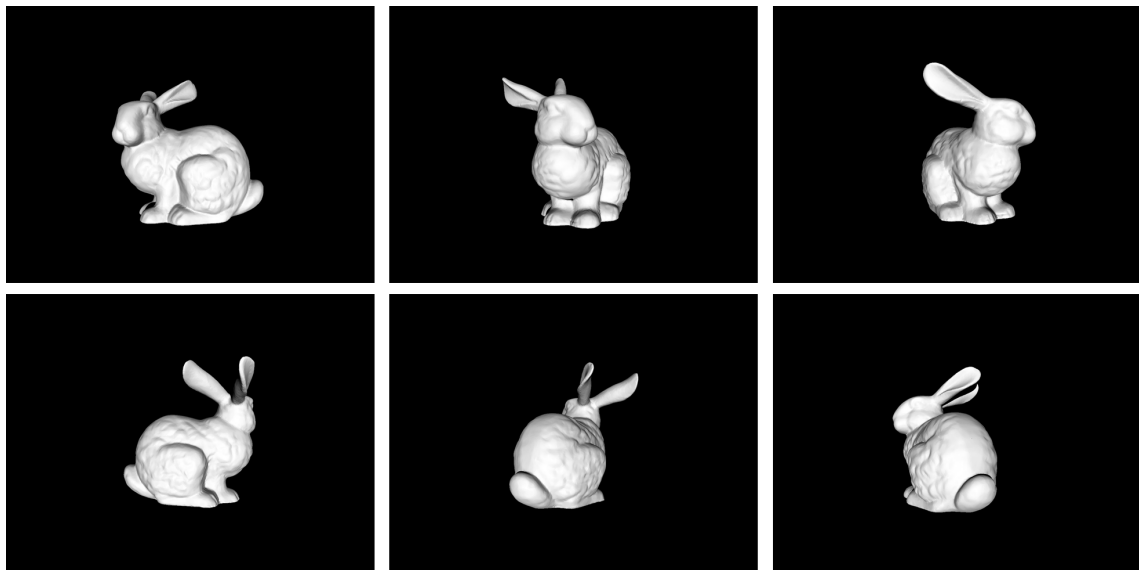


図 4.6: $k = 0.9$ でレンダリングした入力画像例.

4.4.2 実験方法

3次元モデルをレンダリングした画像 60 枚を入力画像として用いた. 入力画像例を図 4.6 に示す. モデルは Stanford Bunny を使用した. 視線方向と同じ光源方向, 表面反射率 $\rho = 1.0$, Minnaert 係数 $k = 0.9$ でレンダリングしたものである. 対象物体は一つの軸に沿って姿勢変化している. Ground truth は, 入力画像のレンダリングに用いた 3次元モデルとする. カメラパラメータは structure from motion での推定結果を用い, 初期形状には multi-view stereo の推定結果を用いた. 光源方向と表面反射率は提案手法による推定結果を用いた.

エネルギーの閾値を $\epsilon = 0.00001$, サブセットの枚数を 6 枚に設定した. 点の追加は 2 回行い, 追加点と元の点の距離 d を初期点群のスケールから $d = 0.1$ とし, 元の 1 点につき 6 点を等間隔に追加した.

光源方向・表面反射率推定とエネルギー最小化に使用する Minnaert 係数 k を $k = 1.0$ (Lambert), $k = 0.9$ の 2 通りで実験を行った.

評価方法は 4.3 節と同様に, 各復元結果と ground truth を ICP によって位置合わせし, その後復元結果の各点から ground truth のメッシュまでのユークリッド距離を求め, その平均を評価する.

表 4.4: Minnaert 則と Lambert 則での比較

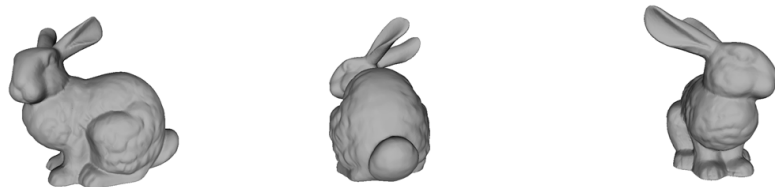
反射モデル	Ground truth までの距離の平均	復元点の数
Minnaert	0.127087	70700
Lambert	0.118318	73916

4.4.3 実験結果と考察

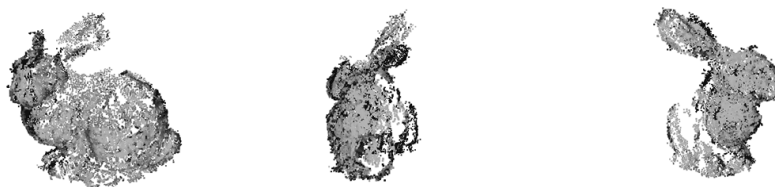
各実験結果の ground truth までの距離の平均と復元点数を表 4.4 に示し、復元結果の画像を図 4.7 に示す。入力画像 60 枚のうち、structure from motion によってカメラパラメータの推定ができたものは 27 枚であった。そのため、復元結果は全周ではなくなっている。

初期形状の復元点数は 15607、ground truth までの距離の平均は 0.133499 である。Minnaert の場合も Lambert の場合もこれを上回っている。

Minnaert 則でレンダリングした画像を入力に用いたが ground truth までの距離の平均と復元点数の両方で Lambert の方が良い結果となった。また、Minnaert 則での結果では縦方向に不自然に復元できていない部分が見られる。理由としては、仮定する反射モデルが複雑になったことで、エネルギー最小化に用いるレベンバーグ・マーカート法での勾配の計算が難しくなっていることが挙げられる。精度の改善のためには、エネルギー関数の設計の見直し、より良い最小化手法の選択が必要である。



(a) Ground truth



(b) 初期形状 (multi-view stereo)



(c) 提案手法 (Minnaert)



(d) 提案手法 (Lambert)

図 4.7: Minnaert 則と Lambert 則での復元結果.

第5章 おわりに

本論文では，multi-view stereo の復元結果に対して光源方向と反射モデルを考慮した最適化を行うことで，視点ごとに光源環境が異なる画像からより精度の高い3次元復元を行う手法を提案した．また，Minnaert 則を仮定することで，小惑星のような粗い表面を持つ物体の復元精度を向上できるか検証を行った．

提案手法では3次元点の位置と法線について最適化を行ったが，4.4節での実験結果が全周の復元にならなかったように，カメラパラメータの推定結果に大きく依存している．より精度の高い復元を行うためには，カメラパラメータについても最適化を行う必要がある．カメラパラメータ，光源方向，3次元点の順に繰り返し最適化を行うことで，3次元復元にかかわる全てのパラメータを調整していき，最終的な復元結果の精度を向上させることが，今後の展望として挙げられる．

謝辞

本研究に際して，指導教官の石川博教授にはご多忙の中様々なご指導を頂きました．また，望月義彦助教，飯塚里誌研究員助教，シモセラ・エドガー研究員助教，鳥取大学の小山田雄仁助教，国立情報学研究所の杉本晃教授本には，研究を進めるにあたり日頃よりご指導，ご指摘を頂きました．会津大学平田成准教授にはデータセット作成に際してご支援頂きました．石川研究室の皆様もデータセット作成への協力を快く引き受けてくださいました．協力して頂いた皆様に感謝の気持ちと御礼を申し上げます．謝辞にかえさせていただきます．

2016 年 2 月 岩野 俊介

参考文献

- [1] はやぶさ2プロジェクト. <http://www.jspec.jaxa.jp/hayabusa2/>.
- [2] 吉川真, 南野浩之, 中澤暁, 安部正真, 津田雄一. 小惑星サンプルリターンミッション「はやぶさ2」の進捗状況(宇宙応用シンポジウム). 電子情報通信学会技術研究報告. SANE, 宇宙・航行エレクトロニクス, Vol. 112, No. 107, pp. 19–22, 2012.
- [3] S Sugita, S Kameda, H Sawada, H Suzuki, M Yamada, C Honda, K Ogawa, K Shirai, Y Iijima, Y Cho, et al. Calibration and first light of optical navigation camera (onc) of hayabusa 2. In *Lunar and Planetary Science Conference*, Vol. 46, p. 2169, 2015.
- [4] 川上恭子, 安部正真, 長谷川直, 春日敏測. はやぶさ後継機に向けた小惑星(162173) 1999ju 3 の観測. 遊星人(日本惑星科学会誌), Vol. 19, pp. 4–11, 2010.
- [5] Yasutaka Furukawa, Brian Curless, Steven M Seitz, and Richard Szeliski. Towards internet-scale multi-view stereo. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pp. 1434–1441. IEEE, 2010.
- [6] Steven M Seitz, Brian Curless, James Diebel, Daniel Scharstein, and Richard Szeliski. A comparison and evaluation of multi-view stereo reconstruction algorithms. In *Computer vision and pattern recognition, 2006 IEEE Computer Society Conference on*, Vol. 1, pp. 519–528. IEEE, 2006.
- [7] Greg Turk and Marc Levoy. The stanford bunny, 2005.
- [8] David G Lowe. Object recognition from local scale-invariant features. In *Computer vision, 1999. The proceedings of the seventh IEEE international conference on*, Vol. 2, pp. 1150–1157. Ieee, 1999.

- [9] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. Surf: Speeded up robust features. In *Computer vision–ECCV 2006*, pp. 404–417. Springer, 2006.
- [10] C. Hernandez, G. Vogiatzis, and R. Cipolla. Multiview photometric stereo. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, Vol. 30, No. 3, pp. 548–554, 2008.
- [11] Zhenglong Zhou, Zhe Wu, and Ping Tan. Multi-view photometric stereo with spatially varying isotropic materials. In *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pp. 1482–1489. IEEE, 2013.
- [12] Tomoaki Higo, Yasuyuki Matsushita, Neel Joshi, and Katsushi Ikeuchi. A hand-held photometric stereo camera for 3-d modeling. In *Computer Vision, 2009 IEEE 12th International Conference on*, pp. 1234–1241. IEEE, 2009.
- [13] Yusuke Yoshiyasu and Nobutoshi Yamazaki. Topology-adaptive multi-view photometric stereo. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pp. 1001–1008. IEEE, 2011.
- [14] Guannan Li, Yebin Liu, and Qionghai Dai. Multi-view photometric stereo of non-lambertian surface under general illuminations. In *3D Imaging (IC3D), 2011 International Conference on*, pp. 1–6. IEEE, 2011.
- [15] Chenglei Wu, Yebin Liu, Xiangyang Ji, and Qionghai Dai. Multi-view reconstruction under varying illumination conditions. In *Multimedia and Expo, 2009. ICME 2009. IEEE International Conference on*, pp. 930–933. IEEE, 2009.
- [16] Daniel Vlasic, Pieter Peers, Ilya Baran, Paul Debevec, Jovan Popović, Szymon Rusinkiewicz, and Wojciech Matusik. Dynamic shape capture using multi-view photometric stereo. In *ACM Transactions on Graphics (TOG)*, Vol. 28, p. 174. ACM, 2009.
- [17] Marcel Minnaert. The reciprocity principle in lunar photometry. *The Astrophysical Journal*, Vol. 93, pp. 403–410, 1941.
- [18] Noah Snavely. Bundler-structure from motion software. *Web site:*) <http://phototour.cs.washington.edu/bundler>.

- [19] Yasutaka Furukawa and Jean Ponce. Accurate, dense, and robust multi-view stereopsis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, Vol. 32, No. 8, pp. 1362–1376, 2010.