HAL open science

Shape and Reflectance Recovery using Multiple Images with Known Illumination Conditions

Kuk-Jin Yoon, Emmanuel Prados, Peter Sturm, Amael Delaunoy, Pau Gargallo

▶ **To cite this version:**

Kuk-Jin Yoon, Emmanuel Prados, Peter Sturm, Amael Delaunoy, Pau Gargallo. Shape and Reflectance Recovery using Multiple Images with Known Illumination Conditions. [Research Report] RR-6309, INRIA. 2007, pp.31. inria-00175274v2

HAL Id: inria-00175274

https://inria.hal.science/inria-00175274v2

Submitted on 27 Sep 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Shape and Reflectance Recovery using Multiple Images with Known Illumination Conditions

Kuk-Jin Yoon — Emmanuel Prados — Peter Sturm — Amaël Delaunoy — Pau Gargallo

N° 6309

September 2007

Thème COG

Rapport de recherche

ISSN 0249-6399 ISRN INRIA/RR--6309--FR+ENG

Shape and Reflectance Recovery using Multiple Images with Known Illumination Conditions

Kuk-Jin Yoon*, Emmanuel Prados*, Peter Sturm*, Amaël Delaunoy*, Pau Gargallo*

Thème COG — Systèmes cognitifs
Projet PERCEPTION

Rapport de recherche n° 6309 — September 2007 — 31 pages

Abstract: We develop a variational method to recover both the shape and the reflectance of a scene surface(s) using multiple images, assuming that illumination conditions are fixed and known in advance. Scene and image formation are modeled with known information about cameras and illuminants, and scene recovery is achieved by minimizing a global cost functional with respect to both shape and reflectance. Unlike most previous methods recovering only the shape of Lambertian surfaces, the proposed method considers general dichromatic surfaces. We verify the method using synthetic data sets containing specular reflection.

Key-words: multi-view stereovision, shape and reflectance recovery, non-lambertian reflectance, variational method.

* PERCEPTION Lab., INRIA - centre de recherche Grenoble - Rhône-Alpes, France

Unité de recherche INRIA Rhône-Alpes
655, avenue de l'Europe, 38334 Montbonnot Saint Ismier (France)
Téléphone : +33 4 76 61 52 00 — Télécopie +33 4 76 61 52 52

Reconstruction de la forme et des propriétés de réflectance d'une scène tridimentionnelle à partir d'images multiples avec des conditions d'éclairage connues

Résumé : Nous développons une méthode variationnelle permettant de reconstruire conjointement la forme tridimensionnelle et les propriétés de réflectance de la surface d'une scène à partir d'images de cette scène prises de plusieurs points de vue. Nous supposons que les conditions d'illumination sont fixées et connues et que les caméras sont complètement calibrées. La méthode consiste à minimiser une fonctionnelle de coût globale par rapport, à la fois, la forme et la réflectance. Contrairement à la plupart des méthodes précédentes reconstruisant seulement la forme de surfaces Lambertiennes, la méthode proposée ici considère des surfaces dichromatiques générales.

Mots-clés : vision tridimensionnelle, reconstruction de la forme et de la réflectance d'une scène, surface non lambertienne, méthode variationnelle.

1 Introduction

Recovering the three-dimensional scene shape using multiple images is one of the major research topics in computer vision. Many methods have been proposed to solve the problem during these last two decades; refer to [15] for an evaluation of various recent methods. However, most of them pay little attention to the reflectance properties of scenes, usually assuming perfectly Lambertian surfaces.

To reduce errors due to non-Lambertian surfaces, Bhat and Nayar [1] analyzed the physics of specular reflection and the geometry of stereopsis, which leads to a relationship between stereo vergence, surface roughness, and the likelihood of a correct match. Zickler et al. [23] presented the Helmholtz stereopsis to overcome the specular reflection problem. However, these two approaches require specialized camera configurations. Some other works [11, 19, 21] also tried to handle non-Lambertian surfaces in the shape reconstruction. However, they rely on some pre-processing and/or specularity-independent photo-consistency measures that are practically hard to use in the case of multiple light sources with different colors. Moreover, they do not recover the surface reflectance.

There are a few works to recover scene geometry and reflectance from known illumination conditions and/or prior knowledge on the shape, while Yu and Malik [22] showed how to recover illumination from known scene geometry and reflectance. Samaras et al. [14] proposed a *multiple* method (succession of several independent processes) for the computation of object shape and reflectance characteristics for non-constant albedo and non-uniformly Lambertian surfaces using 3D models. Recently, Birkbeck et al. [3] proposed a method that recovers diffuse and specular reflectance with changing lights and viewpoints. Our goal is to provide a shape and reflectance estimation method that is global and completely model based as [4]. Also, we want to provide a method that improves the robustness to non-Lambertian effects by directly incorporating a physically based specular model in the mathematical formulation of the problem. By incorporating a complete photometric image formation model, we also exploit prolifically all the photometric phenomena. Also, we thus aim to provide a method that allows to naturally manage with a set of images under with several lighting conditions.

Let us note that actually there already exist recent works that provide solutions in this direction [4], [25]. In [25], Yu and Xu proposed a global and model-based method for recovering the 3D shape and the reflectance properties of a non-Lambertian object. Nevertheless, in this last paper, the authors constrain the object to be made by a *single material*; that is to say that the parameters of the reflectance (in particular the albedo) are the same for all the points of the object surface. So, the method in [25] is a "Multiview Shape From Shading" method, similarly as the one proposed by Jin et al. [10] who focuses on the Lambertian case.

Our method provides a multiview stereo/shape from shading algorithm similarly to [10, 7, 12, 18, 26] which allow to recover 3D shapes from Lambertian shading [10, 7, 12] as well as specular shading [18, 26]. Nevertheless, contrary to these previous works [10, 7, 12, 18, 26, 25], in our work, we do not want to restrain ourself to a single material: in other words, the reflectance properties of the object can spatially (strongly) change. In effect, now a day, more and more objects are now printed and so it is fundamental to be able to recover textured and patterned objects. In return, of course, we will not be able to recover lighting conditions as done [10], and we have to use a parallel

process which return them. In this work, we assume that lighting conditions are known in advance. Practically, we can use spherical objects with the reference white color to capture the directions and the colors of light sources.

In short, in this work, we develop a variational method to recover the shape and the reflectance of a scene surface using multiple images, assuming that illumination conditions are fixed and known in advance. The proposed method can handle quite general dichromatic surfaces exhibiting specular reflection, and the output of this work is the complete description of a scene surface, which can be used to synthesize novel views with arbitrary illumination conditions.

The paper is organized as follows. We first present the modeling of scene introduce the reflection model and the image formation process in Section 2.4. We then formulate the problem in Section 3 and define cost functions in Section 4. Shape evolution and reflectance estimation are described in Section 5 and experimental results for test image sets are shown in Section 7. We then summarize and conclude the work in Section 8.

2 Modeling Assumptions and Notations

We assume here that the scene can be decomposed into two entities: the foreground, which corresponds to the objects of interest, and the background; these are defined more precisely below. The foreground is composed by a set of (bounded and closed) 2D manifolds of $\mathbb{R}^3$. These surfaces are represented by S.

2.1 Lighting conditions

We assume that the scene is illuminated by a finite number of distant point light sources. We complete them by adding an ambient light term (which partially accounts for interreflections and other complex phenomena), with constant energy radiated isotropically in all directions. Note that, based on Wiener's theorems, [17] shows that such a light distribution can approximate arbitrarily well any positive distribution on the sphere. Let n_l be the number of illuminants and $\mathbf{l}_j \in \mathbb{S}^2$ and $L_j \in \mathbb{R}^c$ be the direction and the intensity[1] of the jth illuminant, respectively. $L_a \in \mathbb{R}^c$ is the intensity[1] of the ambient illumination.

2.2 Cameras, image data and visibility

Image data are generated by n_c pinhole cameras. The perspective projection, from world to image coordinates, performed by the ith camera, is represented by $\Pi_i : \mathbb{R}^3 \to \mathbb{R}^2$. $\pi_i \subset \mathbb{R}^2$ is the image domain of the ith camera (i.e. the area covered by the pixels). It is split into two parts: the pixels corresponding to the foreground, $\pi_{iF} = \pi_i \cap \Pi_i(S)$, and the other points $\pi_{iB} = \pi_i \setminus \pi_{iF}$ (associated to the background). $I_i : \pi_i \to \mathbb{R}^c$ is the image of the true scene, captured by the ith camera ($c = 1$ for a gray-scale image, and $c = 3$ for a color image). We denote I the set of input images: $I = \{I_1, I_2, \cdots, I_{n_c}\}$; I_{iF} and I_{iB} are the restrictions of the function I_i to π_{iF} and π_{iB},

[1] Non-normalized color vector, if $c = 3$.

respectively. In other respects, we consider the visibility function δ_{S_i} defined by: $\delta_{S_i}(\mathbf{X}) = 1$ if $\mathbf{X}$ is visible from the ith camera and $\delta_{S_i}(\mathbf{X}) = 0$ otherwise. S_i denotes the part of S that is visible from the ith camera and $\Pi_{i,S}^{-1}$ is the backprojection from the ith camera onto S: i.e. for all points $\mathbf{x} \in \pi_{iF}$, $\Pi_{i,S}^{-1}(\mathbf{x})$ is the closest point on S along the ray joining $\mathbf{X}$ to the optical center of the ith camera.

2.3 Modeling the background

As suggested by [20], to be sure that the estimated foreground surface does not shrink to an empty set (which is indeed the global optimum for most cost functionals used in other works) it is crucial to define and characterize the background. The choice of model is dictated by the scenario and the applications. For example, in [20, 7], the background is characterized by its radiance which is constrained to be constant or strongly regular. At the opposite extreme, when the background is quite irregular, one can assume that one has at his disposal the background images, i.e. the images of the scene captured by the same cameras without foreground objects. Due to lack of space, we only deal here with the latter scenario. Therefore, in addition to the images I, we assume that we detain the background images $\tilde{I} = \{\tilde{I}_1, \cdots, \tilde{I}_{n_c}\}$. Finally, we define $\tilde{I}_{iF}$ and $\tilde{I}_{iB}$ analogously to I_{iF} and I_{iB}.

2.4 Modeling the foreground surface

In this work, we model the foreground object(s) by its shape S and its reflectance R. We denote $\Omega = (S, R)$.

Contrary to most previous stereovision methods, we want to go beyond the Lambertian model. In order to get a solvable minimization problem without too many unknown variables, we chose to represent the reflectance by a parametric model. Of course the chosen model directly depends on the applications aimed at; as an example, we consider the popular Blinn-Phong shading model. In this context, and assuming that $I_i(\mathbf{x})$ is equal to the radiance of the surface S at point $\mathbf{X} = \Pi_{i,S}^{-1}(\mathbf{x})$ in the direction of the ith camera, the images I_i are then decomposed as

$$I_i = I_{id} + I_{is} + I_{ia}, \tag{1}$$

where I_{id}, I_{is}, and I_{ia} are images with the diffuse, specular, and ambient reflection component of I_i, respectively.

Diffuse reflection is caused by the subsurface scattering of light and it is independent of viewing direction. By using the cosine law, this image component is described as

$$I_{id}(\mathbf{x}) = \sum_{j=1}^{n_l} \delta_{L_j}(\mathbf{X}) \Big(\rho_d(\mathbf{X}) L_j \left(\mathbf{n}(\mathbf{X}) \cdot \mathbf{l}_j \right) \Big), \tag{2}$$

where $\rho_d(\mathbf{X}) \in \mathbb{R}^c$ is the diffusion albedo[1] at point $\mathbf{X}$, $\mathbf{n}(\mathbf{X})$ is the normal vector to the surface S at $\mathbf{X}$ and δ_{L_j} represents the light visibility function: S_{L_j} being the part of S visible from the jth illuminant, we define $\delta_{L_j}(\mathbf{X}) = 1$ if $\mathbf{X} \in S_{L_j}$, $\delta_{L_j}(\mathbf{X}) = 0$ otherwise.

Specular reflection is caused by the surface reflection, as with a mirror. This component is expressed as

$$I_{is}(\mathbf{x}) = \sum_{j=1}^{n_l} \delta_{L_j}(\mathbf{X}) \left(\rho_s(\mathbf{X}) L_j \big(\mathbf{n}(\mathbf{X}) \cdot \mathbf{h}_{ij}(\mathbf{X})\big)^{\alpha_s(\mathbf{X})} \right) \tag{3}$$

where $\mathbf{h}_{ij}(\mathbf{X})$ is the bisector of the angle between the view of the ith camera and the jth illuminant at $\mathbf{X}$, $\rho_s(\mathbf{X}) \in \mathbb{R}^c$ and $\alpha_s(\mathbf{X}) \in \mathbb{R}^+$ are the specular albedo and the shininess parameter at point $\mathbf{X}$.

The *ambient illumination* is assumed to be uniform in the scene and modeled as

$$I_{ia}(\mathbf{x}) = \rho_d(\mathbf{X}) L_a, \tag{4}$$

where L_a is defined above.

By combining the diffuse, specular, and ambient reflection, we get the image formation equation as

$$I_i(\mathbf{x}) = \sum_{j=1}^{n_l} \delta_{L_j}(\mathbf{X}) \mathbb{L}_{ij}(\mathbf{X}, \mathbf{n}(\mathbf{X})) + \rho_d(\mathbf{X}) L_a, \tag{5}$$

where

$$\begin{aligned} \mathbb{L}_{ij}(\mathbf{X}, \mathbf{n}(\mathbf{X})) &= \mathbb{L}^d_{ij}(\mathbf{X}, \mathbf{n}(\mathbf{X})) + \mathbb{L}^s_{ij}(\mathbf{X}, \mathbf{n}(\mathbf{X})) \\ &= L_j \rho_d(\mathbf{X}) \big(\mathbf{n}(\mathbf{X}) \cdot \mathbf{l}_j\big) \\ &\quad + L_j \rho_s(\mathbf{X}) \big(\mathbf{n}(\mathbf{X}) \cdot \mathbf{h}_{ij}(\mathbf{X})\big)^{\alpha_s(\mathbf{X})}. \end{aligned} \tag{6}$$

In the sequel, in order to simplify the notations, we denote $R = (R_d, R_s)$, where $R_d = \rho_d$ and $R_s = (\rho_s, \alpha_s)$.

3 Problem Formulation

From a probabilistic point of view, the goal of this work is to estimate the shape S and the reflectance R of a scene surface Ω, that maximize $P(\Omega|I)$ for given I. By Bayes' rule, the problem is then formulated as

$$\begin{aligned} P(\Omega|I) &= \frac{P(I|\Omega)\, P(\Omega)}{P(I)} \propto P(I|\Omega)\, P(\Omega) \\ &= P(I|S, R)\, P(S)\, P(R) \end{aligned} \tag{7}$$

under the assumption that S and R are independent. Here, $P(I|\Omega) = P(I|S, R)$ is a likelihood and $P(S)$ and $P(R)$ are priors on the shape and reflectance respectively.

3.1 Likelihood

If Π_i and illumination conditions are given, we can produce a synthetic image $\bar{I}_i(\Omega)$ corresponding to an input image I_i by using the current estimation of Ω. Here, the correct estimation of Ω will produce the same images as the input images under the given illumination conditions (modulo noise of course). This allows us to measure the validity of the current estimation by comparing input

images with generated ones. When assuming an independent identical distribution (i.i.d) of noise in the observations, the likelihood can be expressed as

$$P(I|\Omega) \propto \prod_{i=1}^{n_c} \exp\big(-\xi_i(\Omega)\big) = \prod_{i=1}^{n_c} \exp\big(-\xi(I_i, \bar{I}_i(\Omega))\big), \tag{8}$$

where $\xi_i(\Omega) = \xi(I_i, \bar{I}_i(\Omega))$ is a function of Ω, measuring the dissimilarity between two images I_i and $\bar{I}_i$.

3.2 Prior on surface shape S

A typical and reasonable prior for the surface shape S is about the area or about the smoothness of a surface. When using the surface area for the prior on S, it is expressed as

$$P(S) \propto \exp\big(-\psi(S)\big). \tag{9}$$

Here, $\psi(S)$ is the monotonic increasing function of the surface area $\int_S d\sigma$ where $d\sigma$ is the Euclidean surface measure.

3.3 Prior on reflectance R

R is composed of two components, $R = (R_d, R_s)$. We express our prior as $P(R) = P(R_d)P(R_s)$ under the assumption that R_d and R_s are independent. Here, $P(R_d)$ and $P(R_s)$ can be assumed uniform in general so that $P(R)$ is constant. However, unfortunately, estimating reliable specular reflectance for all surface points with the uniform prior is very difficult unless there are enough observations exhibiting specular reflection at every surface point. For that reason, we need some specific prior on specular reflectance to be able to infer it inspite of the lack of observations[2]. It is physically valid to assume that specular reflectance varies smoothly within each homogeneous material surface patch. It is, however, also very difficult to partition Ω according to the types of materials. Instead, we use the diffuse reflectance of a surface as a soft constraint to partition Ω and define the prior on the surface reflectance as

$$P(R) \propto \exp\big(-\omega(R)\big) \tag{10}$$

where $\omega(R)$ is a function of the intrinsic gradient of the diffuse and specular reflectance of a surface. This function will be defined below.

[2]We will discuss some special cases that do not need any specific prior on the surface reflectance in Sec. 6.

4 Cost Functions

Based on the derivations in Sec. 3, the problem is formulated as

$$\begin{aligned} P(\Omega|I) &\propto P(I|\Omega)P(\Omega) \\ &\propto \prod_{i=1}^{n_c} \exp\big(-\xi_i(\Omega)\big) \times \Big(\exp\big(-\psi(S)\big)\Big) \\ &\times \Big(\exp\big(-\omega(R)\big)\Big) \end{aligned} \tag{11}$$

and it can be expressed in terms of cost functions as

$$\begin{aligned} E_{total}(\Omega) &= E_{data}(\Omega) + E_{shape}(S) + E_{refl}(R) \\ &= \sum_{i=1}^{n_c} \xi_i(\Omega) + \psi(S) + \omega(R). \end{aligned} \tag{12}$$

This shows that maximizing the probability (Eq. (11)) is equal to minimizing the total cost (Eq. (12)).

4.1 Data cost function

The current estimation of Ω gives a segmentation of the input image I_i into foreground I_{iF} and background I_{iB} and we can synthesize $\bar{I}_{iF}$ according to the above image formation model. As for $\bar{I}_{iB}$, it is generated according to the available background model. In this paper, as mentioned in Sec. 2.3, we use actual background images, i.e. $\bar{I}_{iB}$=$\tilde{I}_{iB}$. $\xi_i(\Omega) = \xi(I_i, \bar{I}_i)$ is then rewritten as

$$\begin{aligned} \xi(I_i, \bar{I}_i) &= \xi_F(I_{iF}, \bar{I}_{iF}) + \xi_B(I_{iB}, \bar{I}_{iB}) \\ &= \xi_F(I_{iF}, \bar{I}_{iF}) + \xi_B(I_{iB}, \tilde{I}_{iB}) \\ &= \hat{\xi}_F(I_{iF}, \bar{I}_{iF}) + \xi(I_i, \tilde{I}_i), \end{aligned} \tag{13}$$

where $\hat{\xi}_F(I_{iF}, \bar{I}_{iF}) = \xi_F(I_{iF}, \bar{I}_{iF}) - \xi_F(I_{iF}, \tilde{I}_{iF})$. Since $\xi(I_i, \tilde{I}_i)$ is independent of Ω, the data cost function is written as

$$E_{data}(\Omega) = \sum_{i=1}^{n_c} \hat{\xi}_F(I_{iF}, \bar{I}_{iF}) + C, \tag{14}$$

where $C = \sum_{i=1}^{n_c} C_i = \sum_{i=1}^{n_c} \xi(I_i, \tilde{I}_i)$ is constant.

4.1.1 Similarity Measure

When computing ξ, any statistical correlation among color or intensity patterns such as the sum of squared differences (SSD), cross correlation (CC), and mutual information (MI) can be used[3]. In

[3] In fact, we do not need to use any sophisticated measure because we also recover the surface reflectance. Instead, we can use the simple pixel-wise measure.

any case, ξ can be expressed as the integral over the image plane as

$$\xi(I_i, \bar{I}_i) = \int_{\pi_i} e(\mathbf{x}) d\sigma_i, \tag{15}$$

where $d\sigma_i$ is the surface measure and $e(\mathbf{x})$ is the contribution at $\mathbf{x}$ to ξ_i. The data cost function is then given as

$$E_{data}(\Omega) = \sum_{i=1}^{n_c} \int_{\pi_{iF}} \hat{e}(\mathbf{x}) d\sigma_i + C, \tag{16}$$

where $\hat{e}(\mathbf{x}) = e\big(I_i(\mathbf{x}), \bar{I}_i(\mathbf{x})\big) - e\big(I_i(\mathbf{x}), \tilde{I}_i(\mathbf{x})\big)$. We adopt the derivations proposed in [13] for ξ_i, e, and $\partial_2 e$.

4.1.2 Decoupling appearance from surface normal

As shown in Eq. (5), surface appearance (i.e., the data cost function) is dependent on both the surface normal and position, and this makes the problem hard to solve and unstable. To resolve this problem, we introduce an auxiliary unit vector field $\mathbf{v}$ satisfying $\|\mathbf{v}\| = 1$ as in [7], which is used for the computation of surface appearance. To penalize the deviation between the actual normal vector $\mathbf{n}$ and the auxiliary normal vector $\mathbf{v}$, we add a new term

$$\begin{aligned} E_{dev}(\Omega) &= \tau \int_S \chi(\mathbf{X}) d\sigma = \frac{\tau}{2} \int_S \|\mathbf{n}(\mathbf{X}) - \mathbf{v}(\mathbf{X})\|^2 d\sigma \\ &= \tau \int_S \left(1 - (\mathbf{n}(\mathbf{X}) \cdot \mathbf{v}(\mathbf{X}))\right) d\sigma, \end{aligned} \tag{17}$$

to the cost function, where τ is a control constant.

4.2 Shape area cost function

By using the area of a surface for the prior, the shape area cost function is simply defined as

$$E_{shape}(S) = \psi(S) = \lambda \int_S d\sigma, \tag{18}$$

where λ is a control constant.

4.3 Reflectance discontinuity cost function

Based on the assumption on surface reflectance in Sec. 3.3, we define a discontinuity cost function of surface reflectance as

$$E_{refl}(R) = \omega(R) = \beta \int_S f(\mathbf{X}) d\sigma, \tag{19}$$

where β is a control constant. $f(\mathbf{X})$ is defined as

$$f(\mathbf{X}) = \zeta\big(R_d(\mathbf{X})\big) \times \eta\big(R_s(\mathbf{X})\big), \tag{20}$$

where $\zeta\big(R_d(\mathbf{X})\big)$ and $\eta\big(R_s(\mathbf{X})\big)$ are defined in terms of the magnitude of the intrinsic gradients of diffuse reflectance and specular reflectance respectively as

$$\zeta\big(R_d(\mathbf{X})\big) = \left(1 - \frac{\|\nabla_S R_d(\mathbf{X})\|^2}{M}\right) \tag{21}$$

$$\eta\big(R_s(\mathbf{X})\big) = \big(\|\nabla_S \rho_s(\mathbf{X})\|^2 + \gamma\|\nabla_S \alpha_s(\mathbf{X})\|^2\big) \tag{22}$$

with a pre-defined constant M[4]. ∇_S denotes the intrinsic gradient defined on S.

The proposed discontinuity cost function of surface reflectance makes the discontinuities of specular reflectance generally coincide with the discontinuities of diffuse reflectance, which is physically valid in general. Accordingly, surface points that do not have enough specular observations get assigned specular reflectance inferred from the specular reflectance of neighboring surface points.

4.4 Total cost function

By combining the cost functions defined in the previous sections, the total cost function is given by

$$\begin{aligned} E_{total}(\Omega) =& E_{data}(\Omega) + E_{dev}(\Omega) \\ &+ E_{shape}(S) + E_{refl}(R). \end{aligned} \tag{23}$$

Here, it is worthy of notice that $E_{dev}(\Omega)$, $E_{shape}(S)$, and $E_{refl}(R)$ are defined over the scene surface while $E_{data}(\Omega)$ is defined as an integral over the image plane. By the change of variable

$$d\sigma_i = -\frac{\mathbf{d}_i(\mathbf{X}) \cdot \mathbf{n}(\mathbf{X})}{z_i(\mathbf{X})^3} d\sigma, \tag{24}$$

where $\mathbf{n}(\mathbf{X})$ is the outward unit surface-normal vector at $\mathbf{X}$, $\mathbf{d}_i(\mathbf{X})$ is the vector connecting the center of the i_{th} camera and $\mathbf{X}$ and $z_i(\mathbf{X})$ is the depth of $\mathbf{X}$ relative to the i_{th} camera, we can replace the integral over the image plane by an integral over the surface:

$$\begin{aligned} E_{data}(\Omega) &= C - \sum_{i=1}^{n_c} \int_{S_i} \Big(\hat{e}(\mathbf{x}) \frac{\mathbf{d}_i(\mathbf{X}) \cdot \mathbf{n}(\mathbf{X})}{z_i(\mathbf{X})^3}\Big) d\sigma \\ &= C - \int_S \left(\sum_{i=1}^{n_c} \delta_{S_i}(\mathbf{X}) \hat{e}(\mathbf{x}) \frac{\mathbf{d}_i(\mathbf{X}) \cdot \mathbf{n}(\mathbf{X})}{z_i(\mathbf{X})^3}\right) d\sigma \end{aligned} \tag{25}$$

When denoting $g(\mathbf{X}, \mathbf{n}(\mathbf{X})) : \mathbb{R}^3 \times \Omega \to \mathbb{R}$ as

$$g(\mathbf{X}, \mathbf{n}(\mathbf{X})) = \left(-\sum_{i=1}^{n_c} \Big(\delta_{S_i} \hat{e} \frac{\mathbf{d}_i \cdot \mathbf{n}}{{z_i}^3}\Big) + \tau\chi + \lambda + \beta f\right), \tag{26}$$

[4] $M \geq 3$ for gray-level images and $M \geq 9$ for color images.

Eq. (23) is simply rewritten as

$$E_{total}(\Omega) = C + \int_S g(\mathbf{X}, \mathbf{n}(\mathbf{X}))d\sigma. \tag{27}$$

Here, although the total cost function is an integral over the surface, it does not suffer from the usual minimal surface bias: most functionals used in multiple stereo have an empty set as globally optimal surface, since they do not "explain" all pixels in the input images. Our approach, like [20], takes into account all pixels in the cost function, using both the estimated foreground and the available background information.

5 Scene Recovery

Scene recovery is achieved by minimizing E_{total} while updating S and R. Unfortunately, S and R are highly coupled and, therefore, it is very complicated to estimate all unknowns simultaneously. To efficiently solve the problem, we adopt an alternating scheme, updating S for a fixed R and then R for a fixed S. This procedure is repeated until E_{total} no longer decreases and S and R no longer change.

5.1 Shape estimation – Surface evolution

When assuming that R is given, E_{total} is a function of S. In this case, the gradient of $E_{total}(S)$ is given according to the derivation in [5] and [16] as

$$\nabla_S E_{total}(S) = (\nabla_S g) \cdot \mathbf{n} + 2gH + \nabla_S \cdot g_{\mathbf{n}}, \tag{28}$$

where H is the mean curvature and $g_{\mathbf{n}}$ represents the gradient on the unit sphere.

Accordingly, the gradient descent surface evolution that minimizes the total cost function is characterized by the normal velocity of the evolution given as

$$S_t = -\left((\nabla_S g) \cdot \mathbf{n} + 2gH + \nabla_S \cdot g_{\mathbf{n}}\right). \tag{29}$$

In this work, we derive the gradient descent flows corresponding to the cost functions respectively. The final gradient descent flow is then given by

$$S_t = \left(S_t\big|_{data} + S_t\big|_{dev} + S_t\big|_{shape} + S_t\big|_{refl} \right) .\mathbf{n} \tag{30}$$

5.1.1 Gradient descent flow for the data cost

According to the form of $E_{data}(\Omega)$, $S_t\big|_{data}$ is given as

$$S_t\big|_{data} = \sum_{i=1}^{n_c} \left(\frac{\hat{e}}{z_i^3}(\nabla_S \delta_{S_i} \cdot \mathbf{d}_i) + \frac{\delta_{S_i}}{z_i^3}\left((\partial_2 \hat{e} \nabla_S \bar{I}_i) \cdot \mathbf{d}_i\right) \right) \tag{31}$$

This includes both the variation related to the camera visibility changes (the first term in Eq. (31)) and the variation related to the image changes (the second term in Eq. (31)), which also includes the variation due to the light visibility changes. By using Eq. (5), $\nabla_S \bar{I}_i$ is expressed as

$$\nabla_S \bar{I}_i = \sum_{j=1}^{n_l} \{(\nabla_S \delta_{L_j}) \mathbb{L}_{ij} + \delta_{L_j} (\nabla_S \mathbb{L}_{ij})\} + (\nabla_S \rho_a) L_a, \tag{32}$$

where $\nabla_S \mathbb{L}_{ij} = \nabla_S \mathbb{L}^d_{ij} + \nabla_S \mathbb{L}^s_{ij}$. Here, it is worthy of notice that the gradient descent flow for the data cost is not dependent on the image gradient, which is sensitive to image noise, but on the shape/reflectance estimation.

5.1.2 Gradient descent flows for the normal deviation cost and the shape area cost

$S_t\big|_{dev}$ (originating from $E_{dev}(\Omega)$) is computed as

$$S_t\big|_{dev} = (-2\tau H + \tau(\nabla_S \cdot \mathbf{v})) \tag{33}$$

and $S_t\big|_{shape}$ (from $E_{shape}(S)$) is the mean curvature flow

$$S_t\big|_{shape} = -2\lambda H. \tag{34}$$

5.1.3 Gradient descent flow for the reflectance discontinuity cost

By using the derivation in [9], we get the following equation for surface evolution.

$$\begin{aligned} S_t\big|_{refl} = -2\beta\Big(&\frac{1}{M} m(\rho_d)\eta(R_s) \\ &- (m(\rho_s) + \gamma m(\alpha_s))\, \zeta(R_d)\Big) \end{aligned} \tag{35}$$

Here,

$$m(\rho_s) = \Big(\mathrm{II}\big(\nabla_S \rho_s \times \mathbf{n}\big) + \|\nabla_S \rho_s\|^2 H\Big), \tag{36}$$

$$m(\alpha_s) = \Big(\mathrm{II}\big(\nabla_S \alpha_s \times \mathbf{n}\big) + \|\nabla_S \alpha_s\|^2 H\Big), \tag{37}$$

$$m(\rho_d) = \Big(\mathrm{II}\big(\nabla_S \rho_d \times \mathbf{n}\big) + \|\nabla_S \rho_d\|^2 H\Big), \tag{38}$$

where $\mathrm{II}(\mathbf{t})$ is the second fundamental form for a tangent vector $\mathbf{t}$ with respect to $\mathbf{n}$.

5.2 Updating the auxiliary vector field $\mathbf{v}$

The computed gradient descent flows minimize the total cost with respect to given reflectance and $\mathbf{v}$. We then update the auxiliary vector field $\mathbf{v}$ to minimize the total cost with respect to given shape

and reflectance. The $\mathbf{v}$ that minimizes the total cost satisfies the following equation.

$$\frac{\partial g}{\partial \mathbf{v}} = \left(-\sum_{i=1}^{n_c} \delta_{S_i} \partial_2 \hat{e} \frac{\partial \bar{I}_i}{\partial \mathbf{v}} \frac{\mathbf{d}_i \cdot \mathbf{n}}{{z_i}^3} \right) + (-\tau \mathbf{n}) = 0 \tag{39}$$

Here, $\frac{\partial \bar{I}_i}{\partial \mathbf{v}}$ is given as

$$\frac{\partial \bar{I}_i}{\partial \mathbf{v}} = \sum_{j=1}^{n_l} \delta_{L_j} L_j \left(\rho_d \mathbf{l}_j + \rho_s \alpha_s \left(\mathbf{v} \cdot \mathbf{h}_{ij} \right)^{\alpha_s - 1} \mathbf{h}_{ij} \right). \tag{40}$$

We update $\mathbf{v}$ by performing gradient descent using the following PDE, with the constraint $\|\mathbf{v}\| = 1$.

$$\frac{\partial \mathbf{v}}{\partial t} = \left(-\sum_{i=1}^{n_c} \delta_{S_i} \partial_2 \hat{e} \frac{\partial \bar{I}_i}{\partial \mathbf{v}} \frac{\mathbf{d}_i \cdot \mathbf{n}}{{z_i}^3} \right) + (-\tau \mathbf{n}) \tag{41}$$

5.3 Reflectance estimation

Here, we estimate R for a fixed S, still minimizing the total cost function. Since E_{dev} and E_{shape} do not depend on R at all, we seek an optimal R by minimizing $(E_{data}(\Omega) + E_{refl}(R))$. Here, because it is also complex to estimate diffuse and specular reflectance at the same time due to the high coupling between them, we alternatively estimate surface reflectance one by one while assuming that the rest are given. We repeat the procedure until they no longer change.

5.3.1 Estimating R_d

For given S and R_s, we estimate ρ_d that minimizes the following.

$$\begin{aligned} E_{data} + E_{refl} = \int_S \Big(\Big(-\sum_{i=1}^{n_c} \delta_{S_i} \hat{e} \frac{\mathbf{d}_i \cdot \mathbf{n}}{{z_i}^3} \Big) \\ + \beta \left(1 - \frac{\|\nabla_S \rho_d\|^2}{M} \right) \eta(R_s) \Big) d\sigma \end{aligned} \tag{42}$$

Here, ρ_d that minimizes the total cost function will satisfy the Euler-Lagrange equation given as

$$-\sum_{i=1}^{n_c} \delta_{S_i} \partial_2 \hat{e} \frac{\partial \bar{I}_i}{\partial \rho_d} \frac{\mathbf{d}_i \cdot \mathbf{n}}{{z_i}^3} + \frac{2\beta}{M} \eta(R_s) \Delta_S \rho_d = 0, \tag{43}$$

where Δ_S denotes the Laplace-Beltrami operator defined on the surface S and $\frac{\partial \bar{I}_i}{\partial \rho_d}$ is given as

$$\frac{\partial \bar{I}_i}{\partial \rho_d} = \sum_{j=1}^{n_l} \delta_{L_j} L_j \left(\mathbf{v} \cdot \mathbf{l}_j \right) + L_a. \tag{44}$$

We solve the PDE by performing gradient descent using the following PDE.

$$\frac{\partial \rho_d}{\partial t} = \left(-\sum_{i=1}^{n_c} \delta_{S_i} \partial_2 \hat{e} \frac{\partial \bar{I}_i}{\partial \rho_d} \frac{\mathbf{d}_i \cdot \mathbf{n}}{{z_i}^3} \right) + \left(\frac{2\beta}{M} \eta(R_s) \right) \Delta_S \rho_d \tag{45}$$

5.3.2 Estimating R_s

We then estimate $R_s = (\rho_s, \alpha_s)$ for given S and R_d in the same manner. ρ_s that minimizes the total cost function will satisfy the Euler-Lagrange equation given as

$$\left(-\sum_{i=1}^{n_c} \delta_{S_i} \partial_2 \hat{e} \frac{\partial \bar{I}_i}{\partial \rho_s} \frac{\mathbf{d}_i \cdot \mathbf{n}}{{z_i}^3} \right) - 2\beta(\Delta_S \rho_s)\zeta(\rho_d) = 0, \tag{46}$$

where $\frac{\partial \bar{I}_i}{\partial \rho_s}$ is given as

$$\frac{\partial \bar{I}_i}{\partial \rho_s} = \sum_{j=1}^{n_l} \delta_{L_j} L_j \left(\mathbf{v} \cdot \mathbf{h}_{ij} \right)^{\alpha_s}. \tag{47}$$

We again solve the PDE by performing gradient descent using the following PDE to get the solution of Eq. (46).

$$\frac{\partial \rho_s}{\partial t} = -\sum_{i=1}^{n_c} \left(\delta_{S_i} \partial_2 \hat{e} \frac{\partial \bar{I}_i}{\partial \rho_s} \frac{\mathbf{d}_i \cdot \mathbf{n}}{{z_i}^3} \right) - 2\beta(\Delta_S \rho_s)\zeta(\rho_d) \tag{48}$$

α_s is also estimated in the same manner by solving the PDE as

$$\frac{\partial \alpha_s}{\partial t} = -\sum_{i=1}^{n_c} \left(\delta_{S_i} \partial_2 \hat{e} \frac{\partial \bar{I}_i}{\partial \alpha_s} \frac{\mathbf{d}_i \cdot \mathbf{n}}{{z_i}^3} \right) - 2\beta\gamma(\Delta_S \alpha_s)\zeta(\rho_d), \tag{49}$$

where $\frac{\partial \bar{I}_i}{\partial \alpha_s}$ is given as

$$\frac{\partial \bar{I}_i}{\partial \alpha_s} = \sum_{j=1}^{n_l} \delta_{L_j} L_j \rho_s \left(\mathbf{v} \cdot \mathbf{h}_{ij} \right)^{\alpha_s} \ln \left(\mathbf{v} \cdot \mathbf{h}_{ij} \right). \tag{50}$$

6 Single-material surface case

When dealing with a single-material surface, it is possible to set $\rho_s(\mathbf{X}) = \rho_s$ and $\alpha_s(\mathbf{X}) = \alpha_s$ for all surface points. In this case, the discontinuity cost function of surface reflectance, $E_{refl}(R)$, can be excluded because $f(\mathbf{X})$ in Eq. (20) is zero everywhere on the surface. Hence the gradient descent flow is then given by

$$S_t = \left(S_t\big|_{data} + S_t\big|_{dev} + S_t\big|_{shape} \right) \mathbf{n} \tag{51}$$

and the PDE used for the estimation of ρ_d , Eq. (45), is simplified as

$$\frac{\partial \rho_d}{\partial t} = -\sum_{i=1}^{n_c} \delta_{S_i} \partial_2 \hat{e} \frac{\partial \bar{I}_i}{\partial \rho_d} \frac{\mathbf{d}_i \cdot \mathbf{n}}{{z_i}^3}. \tag{52}$$

In addition, ρ_s and α_s are computed by performing gradient descent using the following PDEs.

$$\frac{\partial \rho_s}{\partial t} = \int_S \left(-\sum_{i=1}^{n_c} \delta_{S_i} \partial_2 \hat{e} \frac{\partial \bar{I}_i}{\partial \rho_s} \frac{\mathbf{d}_i \cdot \mathbf{n}}{{z_i}^3} \right) d\sigma \tag{53}$$

$$\frac{\partial \alpha_s}{\partial t} = \int_S \left(-\sum_{i=1}^{n_c} \delta_{S_i} \partial_2 \hat{e} \frac{\partial \bar{I}_i}{\partial \alpha_s} \frac{\mathbf{d}_i \cdot \mathbf{n}}{{z_i}^3} \right) d\sigma \tag{54}$$

Table 1: Performance of the proposed method

	accuracy (mm - 90%)	completeness (% - 5.0mm)	e_{imgae}
dragon	shape = mm, ρ_{dr}= , ρ_{dg}=, ρ_{db}=, ρ_{dr}=, ρ_{dr}=,		
bunny	0.004765	89.8748	1.24

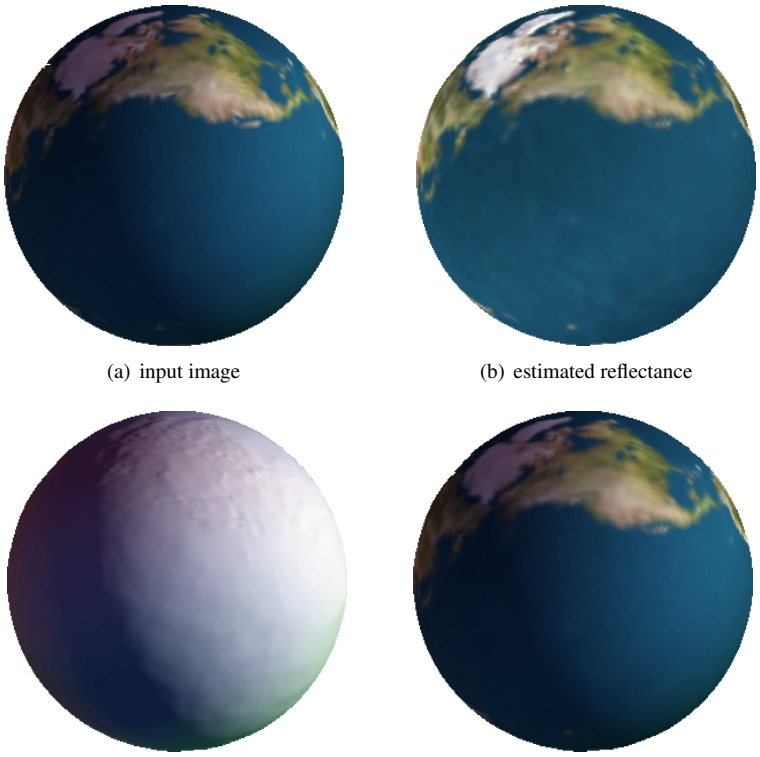
7 Experiments

Instead of implementing the surface evolution directly, we have implemented the gradient descent surface evolution in the level set framework, in which the topological changes of surfaces are handled automatically. The camera and light visibility are computed by using the OpenGL. To verify the proposed method, we generated synthetic images by specifying illumination conditions and surface reflectance for 3D models with various geometries. The algorithm starts from the visual hull obtained by rough silhouette images to reduce the computational time and to avoid local minima. We used the simple L^2-norm to compute the image similarity, e.

7.1 Results for the Lambertian case

In these experiments, $(128 \times 128 \times 128)$ grids were used except the "dragon" image set[5]. Figure 1 shows one of 32 input images and the synthesized image generated by using the estimated shape (and shading) and reflectance. We can generate the images of the scene with different lighting conditions as shown in Fig. 2 by using the estimated shape and reflectance. The results for the "torus" image set are also shown in Fig. 3 and the results for more complex object is shown in Fig. 4. We can see that the images synthesized by using the estimation closely resemble the input images while the shading and the reflectance are successfully separated. The image synthesized by using the estimated shape and reflectance with different lighting conditions and a different viewpoint is shown in Fig. 5.

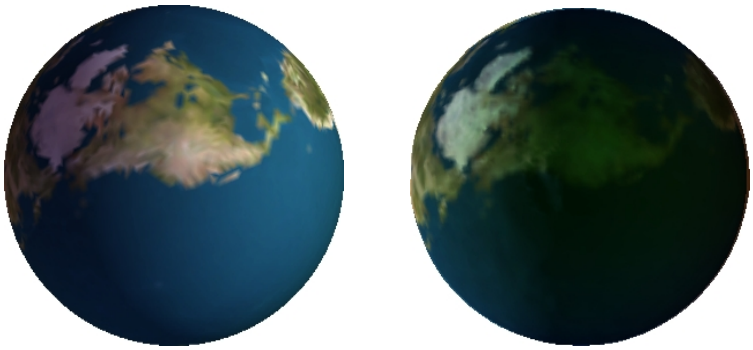
[5] $(160 \times 80 \times 128)$ grids was used for the "dragon" image set.



(a) input image (b) estimated reflectance

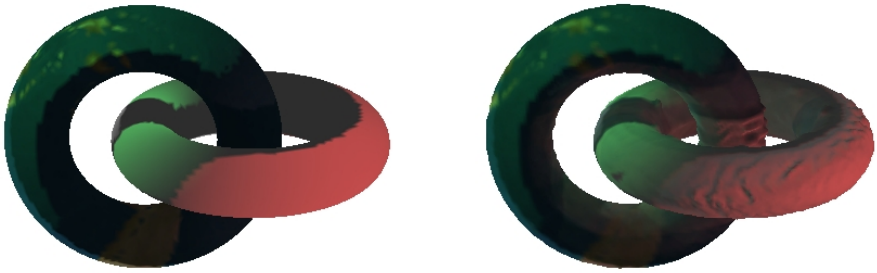
(c) estimated shading (d) synthesized image

Figure 1: Estimation results for a “sphere” image set



(a) original input image (b) synthesized image

Figure 2: Scene synthesis under different lighting conditions



(a) input image (b) synthesized image

Figure 3: Estimation results for a "torus" image set

The estimated shape is quantitatively evaluated in terms of accuracy and completeness as in [15]. We used 95% for accuracy and the 10.0mm error for completeness. Here, beside the shape evaluation, we also evaluated the estimated reflectance in the same manner. For each point on a estimated surface, we found the nearest point on the surface and compute the distance and reflectance differences and vice versa. In addition, for the more quantitative evaluation of surface shape and reflectance, we computed the average of the differences between input images and synthesized images using the L^2-norm as

$$e_{image} = \frac{1}{n_c} \sum_{i=1}^{n_c} \frac{1}{A} \int_{\pi_i} \| \big(I_i(\mathbf{x}) - \bar{I}_i(\mathbf{x}) \big) \| d\sigma_i, \tag{55}$$

where $A = \int_{\pi_i} d\sigma_i$. The performance of the proposed method is summarized in Table 2.

Table 2: Performance of the proposed method

	accuracy (95%) (shape, ρ_{dr}, ρ_{dg}, ρ_{db})	completeness (10.0mm) (shape, ρ_{dr}, ρ_{dg}, ρ_{db})	e_{imgae}
sphere	14.04mm, 0.0254, 0.0189, 0.0167	97.17%, 0.0228, 0.0175, 0.0161	0.6026
dragon	2.63mm, 0.0897, 0.0734, 0.0655	99.88%, 0.0658, 0.0575, 0.0543	5.4812

7.2 Experiments in the general case

To verify the proposed method in the general case, we assigned the uniform reflectance for the 'dragon' image set for the simple evaluation of results. However, we estimated reflectance at every point on a surface without using the fact of uniform reflectance. The 'dragon' and the 'bunny' models used for experiments and ground truth light shading and surface reflectance maps are shown in Fig. 6. Three point light sources with different colors were placed. We took 32 images for the 'dragon' model and 16 images for the 'bunny' model as inputs for recovery.

(a) input image (b) synthesized image

(c) true reflectance of (a) (d) estimated reflectance of (a)

(e) true shading of (a) (f) estimated shading of (a)

Figure 4: Estimation results for a “dragon” image set

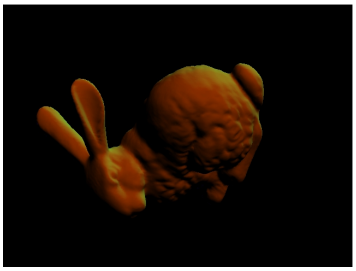
(a) input image (b) synthesized image

(c) true reflectance of (a) (d) estimated reflectance of (a)

(e) true shading of (a) (f) estimated shading of (a)

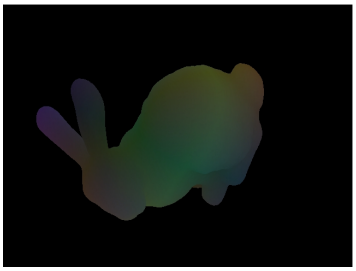
Figure 5: Synthesized image with different lighting conditions and a different viewpoint

(a) input model



(b) diffuse shading

(c) specular shading

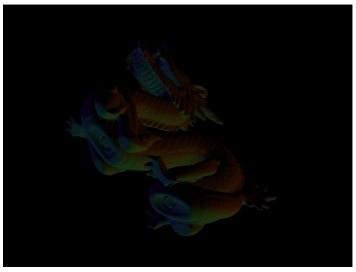


(d) diffuse reflectance

(e) specular reflectance

(f) shininess

(g) ambient refl. comp. of an input image



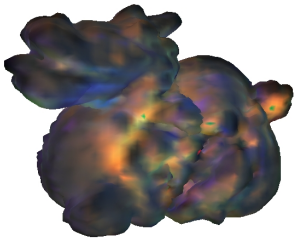
(h) diffuse refl. comp. of an input image

(i) specular refl. comp. of an input image

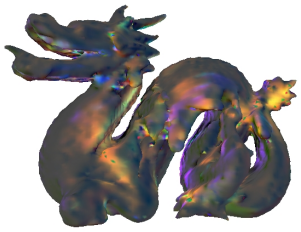
Figure 6: Input models and ground truth shading and reflectance maps

The initial estimations (shape + reflectance) and the refined outputs are shown in Fig. 7 and in Fig. 8. Input images and corresponding synthesized images are also shown for comparison. We can see that the images synthesized by using the estimation closely resemble input images.

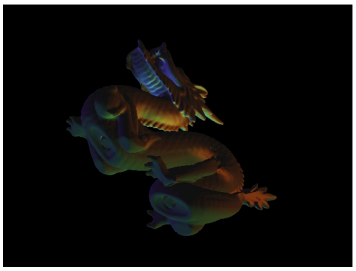
The computational time depends on the number of images and grids. For instance, the 'dragon' image set took about 90 minutes when using 32 images, $(96 \times 96 \times 96)$ grids, and the linux machine with 2.66GHz CPU and 2G memory.



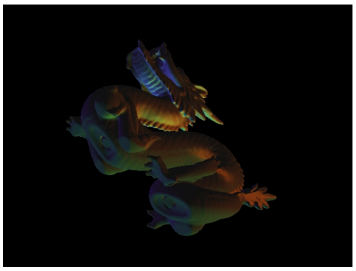
(a) initial shape - ‘dragon’



(b) final shape - ‘dragon’

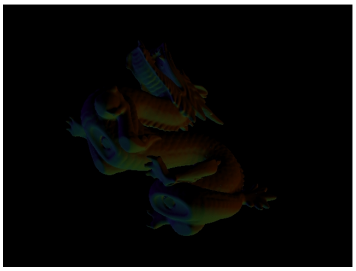


(c) input image



(d) synthesized image

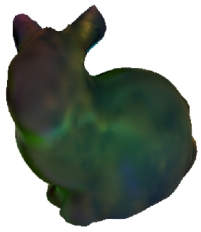
(e) ambient refl. comp.



(f) diffuse refl. comp.

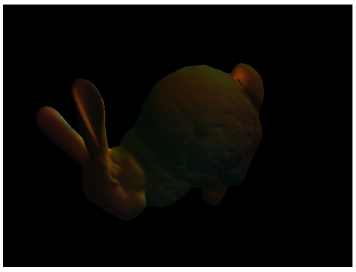
(g) specular refl. comp.

Figure 7: Results for the ‘dragon’ model



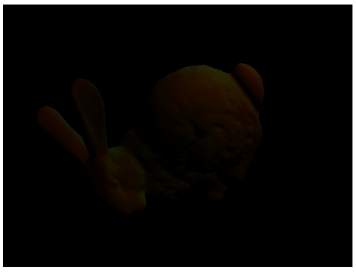
(a) initial shape - ‘bunny’

(b) final shape - ‘bunny’



(c) input image

(d) ambient refl. comp.



(e) diffuse refl. comp.

(f) specular refl. comp.

Figure 8: Results for the ‘bunny’ model

8 Conclusion

In this paper, we have presented a variational method that recovers both the shape and the reflectance of scene surfaces using multiple images. Scene recovery was achieved by minimizing the global cost functional alternatively. As a result, the proposed method produced the complete description of a scene surface.

The main contribution of this paper lies in proposing a solution for dealing with general dichromatic surfaces by utilizing illuminant conditions. We modeled the scene and image formation using know information about cameras and illuminants. We then formulated the problem via Bayes' rule and defined global cost functional in terms of data, shape, and reflectance cost functions. Especially, we presented an efficient reflectance discontinuity cost function to make the problem tractable in spite of the lack of specular reflection observation. In addition, we derived analytic formula for surface evolution and reflectance estimation.

A Appendices

A.1 Notations

- Σ : a set of surfaces in a scene, $\Sigma = \{\Omega, \Omega^C\}$
- Ω : foreground surface in Σ, described by its shape and reflectance as $\Omega = (S, R)$
- Ω^C : background surface in Σ, described by the shape and reflectance as $\Omega^C = (S^C, R^C)$
- S : the shape of Ω
- R : the reflectance of Ω, $R = (R_d, R_s)$
- R_d : the diffuse reflectance (albedo) of Ω, $R_d = \rho_d$[6], $0 \leqq \rho_d \leqq 1$
- R_s : the specular reflectance of Ω, $R_s = (\rho_s, \alpha_s)$
- ρ_s : the coefficient of the specular reflection of Ω, $0 \leqq \rho_s \leqq 1$
- α_s : the shininess constant of the specular reflection of Ω, $\alpha_s \geqq 0$
- n_c : the number of cameras (= the number of images)
- $\mathbf{o}_i$: the view direction of the i_{th} camera
- Π_i : the perspective projection performed by the i_{th} camera, $\Pi_i : \mathbb{R}^3 \rightarrow \mathbb{R}^2$
- π_i : the image plane of the i_{th} camera
- π_{iF} : a region in π_i corresponding to S, $\pi_{iF} = \pi_i \cap \Pi_i(S)$
- π_{iB} : a region in π_i as $(\pi_i - \pi_{iF})$[7]
- I_i : an image captured by the i_{th} camera with foreground surfaces, $I_i : \pi_i \subset \mathbb{R}^2 \rightarrow \mathbb{R}^d$[8]
- I : a set of input images with foreground surfaces, $I = \{I_1, I_2, \cdots, I_{n_c}\}$
- $\tilde{I}_i$: an image captured by the i_{th} camera without foreground surfaces (i.e., backgound image), $\tilde{I}_i : \pi_i \subset \mathbb{R}^2 \rightarrow \mathbb{R}^d$
- $\tilde{I}$: a set of input images without foreground surfaces, $I = \{\tilde{I}_1, \tilde{I}_2, \cdots, \tilde{I}_{n_c}\}$
- I_{iF} : an image region in I_i corresponding to π_{iF}, $I_{iF} : \pi_{iF} \subset \mathbb{R}^2 \rightarrow \mathbb{R}^d$
- I_{iB} : an image region in I_i corresponding to π_{iB}, $I_{iB} : \pi_{iB} \subset \mathbb{R}^2 \rightarrow \mathbb{R}^d$[9]
- $\tilde{I}_{iF}$: an image region in $\tilde{I}_i$ corresponding to π_{iF}, $\tilde{I}_{iF} : \pi_{iF} \subset \mathbb{R}^2 \rightarrow \mathbb{R}^d$
- $\tilde{I}_{iB}$: an image region in $\tilde{I}_i$ corresponding to π_{iB}, $\tilde{I}_{iB} : \pi_{iB} \subset \mathbb{R}^2 \rightarrow \mathbb{R}^d$
- S_i : the part of S visible from the i_{th} camera
- $\Pi_{i,S}^{-1}$: the inverse projection from the i_{th} camera onto S, $\Pi_{i,S}^{-1} : \pi_{iF} \rightarrow S_i$
- $\mathbf{X}$: a point on S
- $\mathbf{x}_i$: a point in the i_{th} image plane corresponding to $\mathbf{X}$, $\mathbf{x}_i \in \pi_i = \Pi_i(\mathbf{X} \in S_i)$

[6] For color images, $\rho_d = (\rho_{dr}, \rho_{dg}, \rho_{db})$. We use R_d and ρ_d interchangeably in this work.
[7] $\pi_i = \pi_{iF} \cup \pi_{iB}, \pi_{iF} \cap \pi_{iB} = \emptyset$
[8] $d = 1$ for a gray image and $d = 3$ for a color image
[9] The mapping from a surface point to an image point is characterized by Π_i and I_i: Π_i and I_i represent the geometric mapping (related to S) and the photometric mapping (related to R) between points respectively.

- $\mathbf{n}(\mathbf{X})$: an outward unit surface-normal vector at $\mathbf{X}$
- n_l : the number of illuminants
- $\mathbf{l}_j$: the direction of the j_{th} illuminant
- L_j : the intensity of the j_{th} illuminant
- L_a : the intensity of the ambient illumination
- ρ_a : the coefficient of the ambient reflection of Ω, $0 \leqq \rho_a \leqq 1$

A.2 Intrinsic gradient on the manifold S

The intrinsic gradient of a function $f : S \rightarrow \mathbb{R}$ on the $(n-1)$-dimensional manifold S embedded in $\mathbb{R}^n$ can be simply defined as the projection of the gradient of a function onto the manifold S as in [2]. When $\mathbf{n}$ denotes a unit normal vector of S, $\nabla_S f$ can be computed simply as

$$\nabla_S f = \nabla \tilde{f} - \left(\mathbf{n}^{\mathrm{T}} \nabla \tilde{f}\right) \mathbf{n} \tag{56}$$

where $\tilde{f} : \mathbb{R}^n \rightarrow \mathbb{R}$ represent a differentiable function whose restriction to S is f.

On the other hand, when S is parameterized using u and v in $\mathbb{R}^3$, the intrinsic gradient on S, ∇_S , can be expresses as

$$\nabla_S f = [S_u, S_v] \begin{bmatrix} S_u \cdot S_u & S_u \cdot S_v \\ S_u \cdot S_v & S_v \cdot S_v \end{bmatrix}^{-1} \begin{bmatrix} f_u \\ f_v \end{bmatrix} \tag{57}$$

A.3 Laplace-Beltrami operator

The Laplace operator is a second order differential operator in the n-dimensional Euclidean space, defined as the divergence of the gradient. The Laplacian can be extended to functions defined on Riemannian and pseudo-Riemannian manifolds. This is named the Laplace-Beltrami operator. According to the definition, the Laplace-Beltrami operator is expressed as

$$\Delta_S f = \nabla_S \cdot (\nabla_S f) \tag{58}$$

B Acknowledgments

This work has been financially supported by the Flamenco project ANR-06-MDCA-007.

References

[1] D.N. Bhat, S.K. Nayar. Stereo and specular reflection. *IJCV*, 26(2):91–106, 1998.

[2] M. Bertalmío and L-T. Cheng and S. Osher, G. Sapiro. Variational Problems and Partial Differential Equations on Implicit Surfaces. *Journal of Computational Physics*, 174 (2):759–780,2001.

[3] N. Birkbeck, D. Cobzas, P. Sturm, M. Jägersand. Variational shape and reflectance estimation under changing light and viewpoints. *ECCV*, 1:536–549, 2006.

[4] A. S. Georghiades. Incorporating the torrance and sparrow model of reflectance in uncalibrated photometric stereo. In *IEEE International Conference on Computer Vision*, volume 02, pages 816–823, 2003.

[5] B. Goldluecke, M. Magnor. Weighted minimal hypersurfaces and their applications in computer vision. *ECCV*, 2:366–378, 2004.

[6] H. Jin, D. Cremers, A.J. Yezzi, S. Soatto. Shedding light on stereoscopic segmentation. *CVPR*, 1:36–42, 2004.

[7] H. Jin, D. Cremers, A. J. Yezzi, and S. Soatto. Shedding light on stereoscopic segmentation. *IEEE Conference on Computer Vision and Pattern Recognition*, 01:36–42, 2004.

[8] H. Jin, S. Soatto, A.J. Yezzi. Multiple stereo reconstruction of dense shape and complex appearance. *IJCV*, 63(3):175–189, 2005.

[9] H. Jin, A.J. Yezzi, Y.-H. Tsai, L.-T. Cheng, S. Soatto. Estimation of 3D surface shape and smooth radiance from 2D images: A level set approach. *Journal of Scientific Computing*, 19(1-3):267–292, 2003.

[10] H. Jin, D. Cremers, D. Wang, E. Prados, A. Yezzi, and S. Soatto. 3-d reconstruction of shaded objects from multiple images under unknown illumination. *To appear in the International Journal of Computer Vision*, 2007.

[11] J. Kim, V. Kolmogorov, R.R. Zabih. Visual correspondence using energy minimization and mutual information. *ICCV*, 1033–1040, 2003.

[12] J. Lu and J. Little. Reflectance function estimation and shape recovery from image sequence of a rotating object. In *IEEE International Conference on Computer Vision*, pages 80–86, 1995.

[13] J.-P. Pons, R. Keriven, O. Faugeras. Modelling dynamic scenes by registering multiple image sequences. *CVPR*, 2:822–827, 2005.

[14] D. Samaras, D. Metaxas, P. Fua, Y.G. Leclerc. Variable albedo surface reconstruction from stereo and shape from shading. *CVPR*, 480–487, 2000.

[15] S.M. Seitz, B. Curless, J. Diebel, D. Scharstein, R. Szeliski. A comparison and evaluation of multiple stereo reconstruction algorithms. *CVPR*, 519–528, 2006.

[16] J.E. Solem, N.C. Overgaard. A geometric formulation of gradient descent for variational problems with moving surfaces. *Scale-Space*, 419–430, 2005.

[17] A. Vedaldi, S. Soatto. Viewpoint invariance for non-planar scenes. Technical Report TR-050012, UCLA CSD, 2006.

[18] G. Vogiatzis, P. Favaro, and R. Cipolla. Using frontier points to recover shape, reflectance and illumunation. In *IEEE International Conference on Computer Vision*, volume 1, pages 228–235, 2005.

[19] R. Yang, M. Pollefeys, G. Welch. Dealing with textureless regions and specular highlights – A progressive space carving scheme using a novel photo-consistency measure. *ICCV*, 576–583, 2003.

[20] A. Yezzi, S. Soatto. Stereoscopic segmentation. *IJCV*, 53(1):31–43, 2003.

[21] K.-J. Yoon, I.-S. Kweon. Correspondence search in the presence of specular highlights using specular-free two-band images. *ACCV*, 761–770, 2006.

[22] Y. Yu, J. Malik. Recovering photometric properties of architectural scenes from photographs. *SIGGRAPH*, 207–217, 1998.

[23] T.E. Zickler, P.N. Belhumeur, D.J. Kriegman. Helmholtz stereopsis: Exploiting reciprocity for surface reconstruction. *IJCV*, 49(2-3):215–227, 2002.

[24] T. Zickler. Reciprocal image features for uncalibrated helmholtz stereopsis. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1801–1808, 2006.

[25] T. Yu, N. Xu, and N. Ahuja. Shape and view independent reflectance map from multiple views. *International Journal of Computer Vision*, 73(2):123–138, 2007.

[26] T. Yu, N. Xu, and N. Ahuja. Recovering shape and reflectance model of non-lambertian objects from multiple views. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 226–233, 2004.

Contents

1 Introduction **3**

2 Modeling Assumptions and Notations **4**
2.1 Lighting conditions 4
2.2 Cameras, image data and visibility 4
2.3 Modeling the background 5
2.4 Modeling the foreground surface 5

3 Problem Formulation **6**
3.1 Likelihood 6
3.2 Prior on surface shape S 7
3.3 Prior on reflectance R 7

4 Cost Functions **8**
4.1 Data cost function 8
4.1.1 Similarity Measure 8
4.1.2 Decoupling appearance from surface normal 9
4.2 Shape area cost function 9
4.3 Reflectance discontinuity cost function 9
4.4 Total cost function 10

5 Scene Recovery **11**
5.1 Shape estimation – Surface evolution 11
5.1.1 Gradient descent flow for the data cost 11
5.1.2 Gradient descent flows for the normal deviation cost and the shape area cost . 12
5.1.3 Gradient descent flow for the reflectance discontinuity cost 12
5.2 Updating the auxiliary vector field $\mathbf{v}$ 12
5.3 Reflectance estimation 13
5.3.1 Estimating R_d 13
5.3.2 Estimating R_s 14

6 Single-material surface case **14**

7 Experiments **15**
7.1 Results for the Lambertian case 15
7.2 Experiments in the general case 17

8 Conclusion **24**

A Appendices **25**
A.1 Notations 25
A.2 Intrinsic gradient on the manifold S 26
A.3 Laplace-Beltrami operator 26

B Acknowledgments **27**

Unité de recherche INRIA Rhône-Alpes
655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)

Unité de recherche INRIA Futurs : Parc Club Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 ORSAY Cedex (France)
Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)
Unité de recherche INRIA Sophia Antipolis : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399