HAL
open science

3-D Reconstruction of Shaded Objects from Multiple Images Under Unknown Illumination

Hailin Jin, Daniel Cremers, Dejun Wang, Emmanuel Prados, Anthony Yezzi, Stefano Soatto

To cite this version:

Hailin Jin, Daniel Cremers, Dejun Wang, Emmanuel Prados, Anthony Yezzi, et al.. 3-D Reconstruction of Shaded Objects from Multiple Images Under Unknown Illumination. International Journal of Computer Vision, 2008, 76 (3), pp.245-256. 10.1007/s11263-007-0055-y . inria-00260433

HAL Id: inria-00260433

https://inria.hal.science/inria-00260433

Submitted on 4 Mar 2008

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.

3-D Reconstruction of Shaded Objects from Multiple Images Under Unknown Illumination

Hailin Jin* Dejun Wang Daniel Cremers† Emmanuel Prados‡ Anthony Yezzi§ Stefano Soatto¶

Keywords: Stereoscopic segmentation, shape from shading, multi-view stereo, variational 3D reconstruction, level set methods, lighting and appearance reconstruction.

Abstract

We propose a variational algorithm to jointly estimate the shape, albedo, and light configuration of a Lambertian scene from a collection of images taken from different vantage points. Our work can be thought of as extending classical multi-view stereo to cases where point correspondence cannot be established, or extending classical shape from shading to the case of multiple views with unknown light sources. We show that a first naive formalization of this problem yields algorithms that are numerically unstable, no matter how close the initialization is to the true geometry. We then propose a computational scheme to overcome this problem, resulting in provably stable algorithms that converge to (local) minima of the cost functional. We develop a new model that explicitly enforces positivity in the light sources with the assumption that the object is Lambertian and its albedo is piecewise constant and show that the new model significantly improves the accuracy and robustness relative to existing approaches.

1 Introduction

We address the problem of recovering the three-dimensional shape of a scene or object that has diffuse, or "Lambertian," reflection, seen from multiple images taken from different vantage points. We assume that both intrinsic and extrinsic calibration parameters of the camera that took the images are known. We do not assume that we

*hljin@adobe.com, Office of Technology, Adobe Systems Incorporated, 345 Park Avenue, San Jose, CA 95110
†dcremers@cs.uni-bonn.de, Dept. of Computer Science, R"omerstr. 164, D-53117 Bonn, Germany
‡emmanuel.prados@inrialpes.fr, INRIA Rhône-Alpes, 655 avenue de l'Europe, 38330 Montbonnot, France
§ayezzi@ece.gatech.edu, School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332
¶soatto@cs.ucla.edu, Computer Science Department, University of California, Los Angeles, CA 90095

know the illumination, and explicitly estimate the ambient illumination level and the position and intensity of one or more point light sources.

This problem can be thought of as the multi-view extension of the problem of "shape from shading" [8, 19, 44, 25, 26, 43, 29, 30, 18, 6, 28] where one makes the assumption that the objects in the scene have homogeneous material, and therefore the shading effects seen on the image are due to their shape, that can be retrieved modulo a bas-relief ambiguity [2, 41]. However, in our model we allow the objects to have piecewise homogeneous materials, a common trait of most man-made objects. Also, we have multiple views available, which allow us to estimate the position and intensity of the light source along with the shape of the observed scene.

The problem also relates to "multi-view stereo", where one is given multiple views of an object and estimates its shape together with the radiance ("texture map") of the scene [22, 7, 13, 10, 33]. We refer the reader to [33] for evaluations of some of the recent algorithms. However, while in multi-view stereo one has to rely on the gradient being non-zero everywhere, and no knowledge on the reflectance properties of the scene (other than Lambertian) is enforced, we can explicitly make use of the assumption of piecewise constancy of the reflectance, and we also retrieve the lighting configuration, which is not addressed in multi-view stereo. Let us underline here that our scenario is different from other ones which deal with multiple images and varying illumination [42, 5]. More germane to our approach is the work of [31] which estimates the lighting direction and the work of [32] which has also been prior attempts to combine stereo and shape from shading, although this is different than ours because we do not attempt to establish direct correspondence.

Finally, this problem relates to segmentation, or better to multi-view, or stereoscopic, segmentation [39], in the sense that scenes for which point-to-point correspondence cannot be easily established (for instance because they do not have distinctive enough gradient profiles) are typically easy to segment, and vice-versa. For a recent reformulation of stereoscopic segmentation based on a probabilistic treatment of individual voxels we refer to [20]. Our goal is to ultimately be able to integrate all cues, and arrive at a coherent 3-D reconstruction framework that can exploit any available knowledge on the scene, and arrive at a consistent estimate of its shape.

Such an algorithm potentially has important applications in 3-D reconstruction for computer-added design, manufacturing, mapping, robotic navigation, visual recognition, and computer graphics, specifically image-based modeling and rendering.

To simplify the problem, we assume that the scene of interest is surrounded by a background with uniform radiance and has a piecewise constant albedo. We show how the problem can be formalized as an infinite-dimensional optimization task. Unfortunately the naive algorithm, based on an iterative procedure designed to have the first-order optimality conditions as its fixed point, yields a numerically unstable flow, and therefore cannot be implemented except for a few special cases. We therefore introduce a different model, based on an auxiliary

vector field. This auxiliary vector field, which is directly responsible for shading effects, is then coupled to the normal field to the shape of the scene via an energy term. This artifice, which can be interpreted as a relaxation, allows us to have a provably stable flow, that has the added benefit of not depending on any derivative of the image, resulting in increased robustness with respect to noise.

Another challenge that we had to overcome, that was a limitation in the prior art, was the enforcing of physical constraints such as the positivity of the light sources. We show that the resulting model yields successful results where prior algorithms failed, as we demonstrate experimentally on synthetic and real image datasets designed to challenge the assumptions underlying our model.

We implement our algorithm in the level set framework, that yields the added benefit of automatically allowing changes of topology both in the shape of the scene, and in the radiance profile, where the boundary between constant materials can split and merge.

This article gathers and strongly upgrades two of our previous conferences papers, [12] and [14]. [14] presented the work on recovering shape for piecewise constant radiance without considering the light effects. In [12], we recover the shape and the light configuration together and solve the variational problems via level set methods. In this new article, we combine our two preliminary works presented in [12] and [14]. We then can recover the shape, the lighting conditions and a piecewise constant approximation of the object's albedo. Moreover, we propose here a physically plausible model of illumination, that explicitly enforces positivity of the sources, without the need to revert to the concept of "negative" light sources introduced in [12]. We finally improve the algorithm by solving some intermediate steps in closed form.

2 Problem Formulation

In this section we introduce our notation and the basic formalization of the problem, and show how it yields numerically unstable algorithms. This serves as a motivation for our solution, introduced in the next section.

2.1 Notation and Modeling Assumptions

Let S be a smooth two-dimensional surface embedded in $\mathbb{R}^3$. We denote with $\mathbf{X} = [X, Y, Z]^T \in \mathbb{R}^3$ the coordinates of a generic point on S with respect to a fixed reference frame. Our goal is to reconstruct the surface S from a set of n images $I_i : \Omega_i \to \mathbb{R}^+, i = 1, ..., n$, where $\Omega_i \subset \mathbb{R}^2$. The intrinsic and extrinsic calibration parameters for each camera are assumed to be known (see e.g. Chapter 4 of [24]), so after some simple pre-processing we can model each camera as an ideal perspective projection $\pi_i : \mathbb{R}^3 \to \Omega_i; \mathbf{X} \mapsto x_i \doteq \pi_i(\mathbf{X}) = \pi(\mathbf{X}_i) = [X_i/Z_i, Y_i/Z_i]^T$,

where $\mathbf{X}_i = [X_i, Y_i, Z_i]^T$ are the coordinates of $\mathbf{X}$ in the i-th camera reference frame. $\mathbf{X}$ and $\mathbf{X}_i$ are related by a rigid body transformation, which can be represented in coordinates by a rotation matrix $R_i \in SO(3)$ and a translation vector $T_i \in \mathbb{R}^3$: $\mathbf{X}_i = R_i\mathbf{X} + T_i$. We assume that there is a background B filling the field of view of each camera, modeled as a sphere with infinite radius. For each camera we define the foreground to be the region $Q_i = \pi_i(S) \subset \Omega_i$ and denote its complement in Ω_i, the background, by Q_i^c. We also define the back-projection $\pi_i^{-1} : \Omega_i \to \mathbb{R}^3$ of x_i onto S, which is the first intersection of the ray through x_i in the i-th camera with S.

We assume that the scene is made of Lambertian surfaces. For simplicity, we only deal with black and white images or with the brightness of the images (even if the extension to color images is quite direct), so the reflectance of the scene is described by positive scalar functions (called *albedos*): $\rho : S \to \mathbb{R}^+$ for the foreground, and $h : B \to \mathbb{R}^+$ for the background. In what follows we assume that the albedo of the background is constant and that the albedo of the foreground is piecewise constant. Furthermore, we assume that the background is uniformly illuminated which results in a uniform radiance. We denote the radiance value by h with an abuse of notation (for an extension to smooth backgrounds see [16]). While this may seem like a restrictive assumption, it is a rather good approximation of most man-made objects. Besides, in lack of any assumption about albedo, there is little one can say about the shape of the scene in the presence of changing illumination [5].

We assume that illumination can be well approximated by a superposition of two components: One is an *ambient term,* with constant energy E_0 radiated isotropically in all directions. This term approximates inter-reflections, diffuse illumination and short-range effects that would be too complex to model explicitly (although see [23, 34] for a way to exploit shading induced by ambient illumination). Again, it is a reasonable approximation in most well-lit environments including indoors (barring black carpet floors) and outdoors on a cloudy day. The second term consists of a number of distant *point light* sources. For simplicity we only allow one source (e.g. the sun on a clear day) but extension to any number (e.g. spotlights in a theater) presents no theoretical difficulties. Using old theorems by Wiener one can show that any positive distribution on the sphere can be approximated arbitrarily well by such a collection of sources [36].

Using Lambert's cosine law [24, 9], we can evaluate the radiance r at each $\mathbf{X} \in S$ via

$$r(\mathbf{X}) = \rho(\mathbf{X})(\langle N(\mathbf{X}), L\rangle \xi(\mathbf{X}) + E_0),$$

where $L \in \mathbb{R}^3$ is a vector pointing in the direction of the light source, with norm equal to its intensity, $N(\mathbf{X}) \in \mathbb{S}_2$ is the outward unit normal to S at $\mathbf{X}$, $\xi : S \to \{0, 1\}$ is the visibility of the light and $\rho(\mathbf{X})$ is the albedo of the surface at $\mathbf{X}$. In the case of convex objects, visibility is given by $\xi = \mathcal{H}(\langle N, L\rangle)$, where $\mathcal{H}$ denotes the Heaviside step function: $\mathcal{H}(x) = 1,\ \forall x \geq 0$ and $\mathcal{H}(x) = 0$, otherwise. More in general, under the assumptions above, we

can write the image formation model as:

$$I(\pi(\mathbf{X})) = \rho(\mathbf{X})(\langle N(\mathbf{X}), L\rangle \xi(\mathbf{X}) + E_0). \tag{1}$$

Let us remark here that we implicitly assume that the brightness of the image I at a pixel x is equal to the irradiance of the associated 3D point located on the retinal plane of the camera (our work can be easily extended to more realistic models, see for example [9]).

Since the albedo ρ is piecewise constant, for simplicity we assume that it partitions S into two regions D_1 and D_2, which can have several connected components. We denote by C the union of the curves dividing D_1 and D_2, and assume they are smooth and closed. Therefore, ρ can be described as

$$\rho(\mathbf{X}) = \rho_j \in \mathbb{R} \text{ for } \mathbf{X} \in D_j,\ j = 1, 2.$$

Note that model (1) is not minimal in the sense that there is a multiplicity ambiguity between ρ, L and E_0 as they appear together in products. We shall address this issue later. Extensions of the albedo model to more than two phases are straight forward using respective multiphase level set formulations [37, 3].

2.2 Formalization of the Problem in a Variational Framework

The scene is described by the three-dimensional (3-D) surface S, and by its reflectance, that given the assumptions above is determined by the 2-D curves C on S and the two scalars ρ_1 and ρ_2. The intensity of the background h and the light sources, L and E_0, are nuisance parameters, in the sense that they are unknown and not necessarily of direct interest, but they affect the measured images, and we will therefore recover them along with the description of the scene.

In order to recover a 3-D model of the scene, we wish to solve an optimization problem whereby we find, among all possible scenes, the one(s) that generate "virtual" images that most closely resemble the given (measured) ones. In other words, we wish to minimize the discrepancy, for instance the squared $\mathcal{L}^2$ distance, between the right-hand-side and the left-hand-side of (1). We call such a discrepancy E_{data}, in the sense that it depends on the

measured images

$$
\begin{aligned}
E_{data} &\doteq \sum_{i=1}^{n} \int_{\pi_i(D_1)} \Big(I_i(x_i) - \rho_1 \big(\langle N, L \rangle \xi(\pi_i^{-1}(x_i)) + E_0 \big) \Big)^2 d\Omega_i \\
&+ \sum_{i=1}^{n} \int_{\pi_i(D_2)} \Big(I_i(x_i) - \rho_2 \big(\langle N, L \rangle \xi(\pi_i^{-1}(x_i)) + E_0 \big) \Big)^2 d\Omega_i \\
&+ \sum_{i=1}^{n} \int_{Q_i^c} (I_i(x_i) - h)^2 d\Omega_i.
\end{aligned} \tag{2}
$$

Among the unknowns, $\rho_1, \rho_2, E_0, L, h, C, S$, the latter two are infinite-dimensional. Therefore, one needs to impose additional assumptions, for instance in the form of regularization, to arrive at a well-posed problem. The most obvious choice is to search for a solution that maximizes the regularity of C and S. That can be easily achieved by adding terms to the energy functionals, for instance

$$
E_{smooth} \doteq \int_S dA \tag{3}
$$

that measures the area of S and, when minimized, forces regularity, and similarly for the curves C

$$
E_{curv} \doteq \int_C ds, \tag{4}
$$

where dA and ds are the area form of S and the arc length of C respectively. Unfortunately, even considering the composite functional $E_{data} + \alpha E_{smooth} + \beta E_{curv}$, where α and β are positive weighting factors, is not sufficient to yield a well-posed problem, as we illustrate next.

2.3 Stable Formulation by Decoupling the Normal Field

The minimization of the energy functional just described represents an ill-posed problem. To illustrate this phenomenon, consider a simpler version of the energy above, where we neglect the background term and we consider only a constant albedo foreground, $\rho_1 = \rho_2 = \rho$. In other words, we concentrate only on the first line of (2), and for simplicity we neglect the visibility term, so $\xi(\mathbf{X}) = 1, \ \forall \ \mathbf{X} \in S$. After some calculations, the reader can verify that the curvature-dependent term of the first-order variation of E_{data} is given by [11]

$$
\sum_{i=1}^{n} 2H \left((I_i - E_0)^2 + 2\rho^2 - 3 \langle \rho L, N \rangle^2 \right) - 2\rho^2 \Pi(N \times L) \tag{5}
$$

where H denotes the mean curvature and Π the second fundamental form. In regions where the surface faces the light (and therefore $\langle L, N \rangle \to 1$) and where the modeled directional intensity ρ^2 exceeds the measured one

$(I_i - E_0)^2$, the coefficients of both H and Π are negative. This amounts to a backward heat flow, that models a physically unrealizable process and is well-known to be numerically unstable. Therefore, a local gradient-based minimization algorithm based on the first-order optimality condition is bound to be numerically unstable no matter how close the initial conditions are.

The above instability arises due to the strong coupling between surface appearance, as measured by the image, and its normal vector field N. In the presence of noise in the image, with fixed illumination, the surface will bend and ripple to fit the data. To circumvent this instability, we propose a relaxed cost functional in which the normal is decoupled from the surface through an auxiliary unit vector field V, defined as

$$V : S \to \mathbb{S}_2,\ \mathbf{X} \mapsto V(\mathbf{X}), \tag{6}$$

where $\mathbb{S}_2 \subset \mathbb{R}^3$ is the unit sphere of dimension 2. V will take the place of the unit normal in modeling the shading effects. We will show that the induced surface flow lacks the potentially unstable curvature-based diffusion terms. The data fitness term now becomes

$$\begin{aligned} E_{data} \doteq \sum_{i=1}^{n} \int_{\pi_i(D_1)} \Big(I_i(x_i) - \rho_1\big(\langle V(\pi_i^{-1}(x_i)), L\rangle \xi(\pi_i^{-1}(x_i)) + E_0\big)\Big)^2 d\Omega_i \\ + \sum_{i=1}^{n} \int_{\pi_i(D_2)} \Big(I_i(x_i) - \rho_2\big(\langle V(\pi_i^{-1}(x_i)), L\rangle \xi(\pi_i^{-1}(x_i)) + E_0\big)\Big)^2 d\Omega_i \\ + \sum_{i=1}^{n} \int_{Q_i^c} (I_i(x_i) - h)^2 d\Omega_i. \end{aligned} \tag{7}$$

To stay faithful to the physically motivated interaction between the surface normal and the light source direction, we will introduce an indirect coupling between the unit normal and the modeled surface radiance by adding a term to our energy which penalizes the average deviation between the true unit normal of the surface and the unit vector field V which takes its place in the new radiance model. The constraint for V is given by a penalty on the $\mathcal{L}^2$ distance between V and the unit normal field N on S:

$$E_{coupling} = \frac{1}{2} \int_S |V(\mathbf{X}) - N(\mathbf{X})|^2 dA = \int_S \big(1 - \langle V(\mathbf{X}), N(\mathbf{X})\rangle\big) dA, \tag{8}$$

where we have used the fact that the vectors V and N have unit length. The overall cost functional is simply a weighted average of the several costs:

$$E_{total} = E_{data} + \alpha E_{smooth} + \beta E_{curv} + \gamma E_{coupling}. \tag{9}$$

When minimizing E_{total}, we need to guarantee that V is always a unit vector field, i.e. $|V(\mathbf{X})|^2 = 1, \quad \forall\, \mathbf{X} \in S$, and that the radiance coefficients $\rho_1,\ \rho_2$ are positive.

Decoupling the effective normal field from the true surface normal allows us to bypass this ill-posed formulation. As the auxiliart field does not depend upon the surface derivatives (as does the true unit normal), the optimiality conditions for V (and the ensuing gradient flow) do not involve teh second-order derivatives contained in heat flows.

3 Optimal Reconstruction

In order to minimize E_{total} with respect to all the unknowns, we adopt an alternating minimization procedure, evolving one step at a time in the gradient direction of each unknown. For the simplest unknowns, we will be able to compute a closed-form solution for each iteration.

3.1 Updating of the Surface Properties: Geometry and albedos

We start with minimization with respect to S, that is the most delicate. To facilitate finding the variation of the data fitness term with respect to S, we need to introduce two more terms. Let $\chi_i : S \to \{0,1\}$ be the surface visibility function with respect to the i-th camera, i.e. $\chi_i(\mathbf{X}) = 1$ for points on S that are visible from the i-th camera and $\chi_i(\mathbf{X}) = 0$ otherwise. Let σ_i account for the change of coordinates from $d\Omega_i$ to dA, i.e, $\sigma_i = \frac{d\Omega_i}{dA} = \langle \mathbf{X}_i, N_i \rangle / Z_i^3$, where N_i the unit normal N expressed in the i-th camera reference frame, $N_i = R_i N$. For simplicity, we first illustrate the properties of the flow for the case of constant albedo foreground: The data term is

$$\begin{aligned}
&\sum_{i=1}^{n} \int_{Q_i} \Big(\big(I_i - \langle V, L\xi \rangle - E_0\big)^2 - \big(I_i - h\big)^2 \Big) d\Omega_i + \sum_{i=1}^{n} \int_{\Omega_i} \big(I_i - h\big)^2 d\Omega_i \\
= &\sum_{i=1}^{n} \int_{S} \chi_i \Big(\big(I_i - \langle V, L\xi \rangle - E_0\big)^2 - \big(I_i - h\big)^2 \Big) \sigma_i dA + \sum_{i=1}^{n} \int_{\Omega_i} \big(I_i - h\big)^2 d\Omega_i,
\end{aligned}$$

where the background integral over the complement Ω_i^c of Q_i was expressed as an the integral over the entire domain Ω_i minus an integral over Q_i.

and the gradient descent flow based on the first-order variation is given by [11]

$$S_t = \Big(\sum_{i=1}^{n} \frac{1}{Z_i^3}\big((I_i - (\langle V, L\xi\rangle + E_0))^2 - (I_i - h)^2\big) \cdot \big\langle \nabla\chi_i, R_i^T \mathbf{X}_i \big\rangle$$
$$- \sum_{i=1}^{n} 2\chi_i \Big(I_i - \langle V, L\xi\rangle - E_0 \Big) \Big(\xi L^T \nabla_S V R_i^T \mathbf{X}_i + \sum_{j=1}^{l} \langle V, L\rangle \langle \nabla\xi, R_i^T \mathbf{X}_i \rangle \Big)$$
$$+ \big(\ 2H(\alpha+\beta) - \beta \nabla_S \cdot V \ \big) \Big) N \quad (10)$$

where ∇_S is the Laplace-Beltrami operator. Note that the only second-order term (curvature term) in the flow (10) is $2H(\alpha+\beta)N$, therefore the flow is always numerically stable (with a properly chosen time step). Another advantage of flow (10) is that it depends only upon the image values, *not the image gradients*. This property greatly increases the robustness of the resulting algorithm to image noise.

For the more general case, where we have piecewise constant albedo foreground regions separated by the contours C, to facilitate the computing of the gradient descent flow we define the level set function $\phi : S \to \mathbb{R}$ to represent C via $D_1 = \{X|\phi(X) > 0\}$, $D_2 = \{X|\phi(X) < 0\}$ and $C = \{X|\phi(X) = 0\}$. $\mathcal{H}(\phi)$ is the Heaviside function of the level set function ϕ. The gradient flow of the data term with respect to the surface S is given by [11]

$$S_t = \Bigg(\sum_{i=1}^{n} \frac{\Gamma_i}{Z_i^3} \langle \chi_{iX}, R_i^T X_i \rangle$$
$$- \chi_i \frac{\delta(\phi)}{Z_i^3} \Big(\big(I_i - \rho_1(\langle V, L\rangle\xi + E_0)\big)^2 - \big(I_i - \rho_2(\langle V, L\rangle\xi + E_0)\big)^2 \Big) \langle \nabla_S\phi, R_i^T X_i \rangle$$
$$+ 2\alpha k + \beta \frac{\delta(\phi)}{|\nabla_S \phi|} \Pi(\nabla_s \phi \times N) \Bigg) N, \quad (11)$$

where we have defined

$$\Gamma_i \doteq \mathcal{H}(\phi)(I_i - \rho_1(\langle V, L\rangle\xi + E_0))^2 + (1 - \mathcal{H}(\phi))(I_i - \rho_2(\langle V, L\rangle\xi + E_0))^2 - (I_i - h)^2.$$

Similarly one can derive the component of the gradient flow due to the data term for the curves C [15]:

$$C_t = \sum_{i=1}^{n} ((I_i - \rho_2(\langle V, L\rangle\xi + E_0))^2 - (I_i - \rho_1(\langle V, L\rangle\xi + E_0))^2)\sigma_i + \beta k_g)\vec{n} \quad (12)$$

where $\vec{n}$ is the normal direction of a point in the segmenting curve between regions of different albedo, k is the curvature, k_g is the geodesic curvature. In our level set implementation, we implement the evolution of the level

set function ϕ instead of C, as follows [15]:

$$\phi_t = |\nabla_S \phi| \sum_{i=1}^{n} \Big((I_i - \rho_2(\langle V, L\rangle \xi + E_0))^2 - (I_i - \rho_1(\langle V, L\rangle \xi + E_0))^2 \Big) \sigma_i + \beta \Big(\delta_S \phi - \frac{\nabla_S^T \phi \nabla_S^2 \phi \nabla_S \phi}{|\nabla_S \phi|^2} \Big).$$

S and C being now updated, we then update ρ_1 and ρ_2 as follows:

$$\rho_1 = \frac{\sum_{i=1}^{n} \int_{\pi_i(D_1)} I_i(x_i) d\Omega_i}{\sum_{i=1}^{n} \int_{\pi_i(D_1)} (\langle V, L\rangle \xi + E_0) d\Omega_i}, \tag{13}$$

$$\rho_2 = \frac{\sum_{i=1}^{n} \int_{\pi_i(D_2)} I_i(x_i) d\Omega_i}{\sum_{i=1}^{n} \int_{\pi_i(D_2)} (\langle V, L\rangle \xi + E_0) d\Omega_i}. \tag{14}$$

The next step consists in fixing S, C, ρ_1, ρ_2 and in minimizing the energy with respect to $V,\ L,\ E_0,\ h$ while enforcing normalization and positivity constraints.

3.2 Updating of Radiance Parameters L, E_0 and h

In this step, we fix S, C, V, ρ_1 and ρ_2 and we minimize E_{total} with respect to L, E_0 and h.

Let us start with the updating of h. One can easily verify that the optimal value for h is the mean, on all the images (I_i, Ω_i), of the intensity of the pixels located in the background parts Q_i^c (let us remind that Q_i^c represents the complement of the region $Q_i = \pi_i(S) \subset \Omega_i$); i.e.

$$h = \frac{\sum_{i=1}^{n} \int_{Q_i^c} I_i d\Omega_i}{\sum_{i=1}^{n} \int_{Q_i^c} d\Omega_i}. \tag{15}$$

Now, let us consider the updating of E_0 and L. This step is quite more complicated than the previous one. In effect, a basic optimization with respect to E_0 would clearly involve some changes of its sign. In order to ensure the physical plausibility of this variable at any iteration, we have then to constrain E_0 to be positive

$$E_0 \geq 0$$

during the minimization process. Furthermore in order to increase the consistency and the efficiency of the algorithm, we also minimize our energy with respect to E_0 and L *simultaneously*.

The minimization with respect to (E_0, L) then consists in a constrained optimization problem. The Kuhn-Tucker conditions [21] provide the appropriate tools for dealing with this kind of problems. These tools provide necessary conditions for a solution to be optimal; we then use these conditions for finding a solution as one usually does in standard optimization. Being given the energy we are minimizing and the enforced constraint, the associated

Lagrangian [21] is:

$$\sum_{i=1}^{n}\int_{\pi_i(D_1)}(\ I_i(x_i)-\rho_1(\langle V(\pi_i^{-1}(x_i)),L\rangle\xi(\pi_i^{-1}(x_i))+E_0)\ )^2 d\Omega_i$$
$$+\sum_{i=1}^{n}\int_{\pi_i(D_2)}(\ I_i(x_i)-\rho_2(\langle V(\pi_i^{-1}(x_i)),L\rangle\xi(\pi_i^{-1}(x_i))+E_0)\ )^2 d\Omega_i$$
$$+\sum_{i=1}^{n}\int_{Q_i^c}(I_i(x_i)-h)^2 d\Omega_i+\alpha\int_S dA+\beta\int_C ds$$
$$+\gamma\int_S(1-\langle V(\mathbf{X}),N(\mathbf{X})\rangle)dA\ \ +\ \ \lambda E_0\ \ =0 \quad (16)$$

constrained by $\lambda \geq 0$. The Kuhn-Tucker [21] conditions are: If (E_0, L) is a local minimum, then there exists a non-negative constant λ s.t.

$$\begin{cases} 1)\ \sum_{i=1}^{n}\int_{\pi_i(D_1)}(I_i(x_i)-\rho_1(\langle V,L\rangle\xi+E_0))\ \rho_1 V\xi\ d\Omega_i \\ \qquad +\sum_{i=1}^{n}\int_{\pi_i(D_2)}(I_i(x_i)-\rho_2(\langle V,L\rangle\xi+E_0))\ \rho_2 V\xi\ d\Omega_i=0, \\ 2)\ -\sum_{i=1}^{n}\int_{\pi_i(D_1)}(I_i(x_i)-\rho_1(\langle V,L\rangle\xi+E_0))\ \rho_1\ d\Omega_i \\ \qquad -\sum_{i=1}^{n}\int_{\pi_i(D_2)}(I_i(x_i)-\rho_2(\langle V,L\rangle\xi+E_0))\ \rho_2\ d\Omega_i\ +\ \lambda=0, \\ 3)\ \lambda E_0=0. \end{cases} \quad (17)$$

In the equations above, 1) is derived by differentiating Eq. (16) with respect to L and 2) is derived by differentiating (16) with respect to E_0.

We now focus on the domain of definition of the above equations. If a solution is retrieved within the domain $\{\ E_0 \mid E_0 > 0\ \}$, then $\lambda = 0$. Thus according to the first and second equations of (17), we have

$$\begin{cases} L=\left[\rho_1^2\sum_{i=1}^{n}\int_{\pi_i(D_1)}VV^T\xi d\Omega_i+\rho_2^2\sum_{i=1}^{n}\int_{\pi_i(D_2)}VV^T\xi d\Omega_i\right]^{-1} \\ \qquad\cdot\ \left(\rho_1\sum_{i=1}^{n}\int_{\pi_i(D_1)}(I_i(x_i)-\rho_1E_0)V\xi d\Omega_i\right. \\ \qquad\qquad \left.+\rho_2\sum_{i=1}^{n}\int_{\pi_i(D_2)}(I_i(x_i)-\rho_2E_0)V\xi d\Omega_i\right), \\ E_0=\left(\sum_{i=1}^{n}\int_{\pi_i(D_1)}(I_i(x_i)-\rho_1\langle V,L\rangle\xi)\rho_1 d\Omega_i\right. \\ \qquad \left.+\sum_{i=1}^{n}\int_{\pi_i(D_2)}(I_i(x_i)-\rho_2\langle V,L\rangle\xi)\rho_2 d\Omega_i\right) \\ \qquad\qquad\cdot\ 1/(\sum_{i=1}^{n}\int_{\pi_i(D_1)}\rho_1^2 d\Omega_i+\sum_{i=1}^{n}\int_{\pi_i(D_2)}\rho_2^2 d\Omega_i), \end{cases} \quad (18)$$

where $[.]^{-1}$ denotes the matrix inverse operator.

In order to simplify the above equations, we let

$$\begin{cases} A &= \rho_1 \sum_{i=1}^n \int_{\pi_i(D_1)} I_i(x_i) d\Omega_i \quad + \quad \rho_2 \sum_{i=1}^n \int_{\pi_i(D_2)} I_i(x_i) d\Omega_i, \\ V^* &= \rho_1^2 \sum_{i=1}^n \int_{\pi_i(D_1)} V \xi d\Omega_i \quad + \quad \rho_2^2 \sum_{i=1}^n \int_{\pi_i(D_2)} V \xi d\Omega_i, \\ M &= \rho_1^2 \sum_{i=1}^n \int_{\pi_i(D_1)} V V^T \xi d\Omega_i \quad + \quad \rho_2^2 \sum_{i=1}^n \int_{\pi_i(D_2)} V V^T \xi d\Omega_i, \\ p^* &= \rho_1 \sum_{i=1}^n \int_{\pi_i(D_1)} I_i(x_i) V \xi d\Omega_i + \rho_2 \sum_{i=1}^n \int_{\pi_i(D_2)} I_i(x_i) V \xi d\Omega_i, \\ B &= \rho_1^2 \sum_{i=1}^n \int_{\pi_i(D_1)} d\Omega_i \quad + \quad \rho_2^2 \sum_{\pi_i(D_2)} d\Omega_i. \end{cases} \tag{19}$$

where $A \in \mathbb{R}$, $V^* \in \mathbb{R}^3$, $M \in M_{3\times 3}(\mathbb{R})$, $p^* \in \mathbb{R}^3$ and $B \in \mathbb{R}^+$. Let us note here that, even if for all $\mathbf{X} \in S$ the matrix $V(\mathbf{X})V^T(\mathbf{X})$ is not invertible (its rank is 1), it is reasonable to assume that M is. Moreover, as soon as the surface S is visible from any camera, then $B > 0$. By using the above abbreviations, which are all known quantitites in this phase of the optimization, we simplify the equations as follows:

$$\begin{cases} L = M^{-1}(p^* - E_0 V^*), \\ E_0 = (A - \langle V^*, L \rangle)/B. \end{cases} \tag{20}$$

By combining the two equations above, we have

$$L = M^{-1}\Big(p^* - (A - \langle V^*, L \rangle) V^*/B\Big),$$

so

$$L \;=\; \langle V^*, L \rangle (M^{-1} V^*)/B \quad + \quad M^{-1}(p^* - AV^*/B). \tag{21}$$

In order to simplify this last equation, we introduce two new vector variables, $\mathbf{a}$ and $\mathbf{b}$:

$$\mathbf{a} = (M^{-1} V^*)/B \qquad \in \mathbb{R}^3, \tag{22}$$

and

$$\mathbf{b} = M^{-1}(p^* - AV^*/B) \qquad \in \mathbb{R}^3. \tag{23}$$

We have then

$$L \;=\; \mathbf{a} V^{*T} L \;+\; \mathbf{b}. \tag{24}$$

L is then the solution of the linear system

$$WL = \mathbf{b}, \tag{25}$$

where $W = Id_{3\times 3} - \mathbf{a}V^{*T}$.

It is therefore a simple matter of linear algebra to retrieve L from the equation above, since all other terms are known.

E_0 can now be derived from L using equation (20). If it is greater than zero, then it is what we are searching for and so L and E_0 are obtained. Otherwise, we have to look for the minimum such that $E_0 = 0$. This problem is significantly easier. We can directly derive L from equation (20):

$$L = M^{-1}p^*. \tag{26}$$

3.3 Updating of the Auxiliary Vector Field V

Having fixed $S,\ C,\ \rho_1,\ \rho_2,\ L,\ E_0$ and h, we now minimize the energy with respect to V. Recall that we have to guarantee

$$|V| = 1,$$

that is to say that V has to stay on the manifold $\mathbb{S}_2$ (the unit sphere). In our previous conference paper [12], we proposed to update the auxiliary field V by an iterated process. Here we propose a "closed-form" solution.

For any Riemannian manifold $\mathcal{M} \subset \mathbb{R}^m$ and any differentiable energy $E : \mathbb{R}^m \to \mathbb{R}$, we have the following (necessary) optimality condition: if $x \in \mathcal{M}$ is a local minima of $E_{|\mathcal{M}}$, then $\nabla E(x)$ is orthogonal to the tangent plane $T\mathcal{M}(x)$, i.e.

$$\nabla E(x) \in \big(T\mathcal{M}(x)\big)^{\perp}.$$

Here, $\nabla E(x)$ denotes the classical gradient of $E : \mathbb{R}^m \to \mathbb{R}$ at point x defined in the framework of the differential calculus. Also, the above statement directly results from the chain rule and the classical optimality condition in the differential calculus framework; see for example [35, 1, 4].

Now, let us remark that for any *point* V on the unit sphere $\mathbb{S}_2$, the *vector* V is orthogonal to the tangent plane to the sphere at point V, i.e.

$$\forall V \in \mathbb{S}_2,\ V \in (\ T\mathbb{S}_2(V)\ )^{\perp}.$$

In the case where $\mathcal{M} = \mathbb{S}_2$, since the dimension of $(\ T\mathbb{S}_2(V)\ )^{\perp}$ is one, then the above optimality condition becomes:

$$\text{if } V \in \mathbb{S}_2 \text{ is a local minima of } E_{|\mathbb{S}_2}, \text{ then } \nabla E(V) \propto V. \tag{27}$$

Beyond this, the terms of E_{total} which depend on V are E_{data} and $E_{coupling}$. Let us stress that, even if V is indirectly regularized by the coupling term $E_{coupling}$ (which forces V to be close to N which is smooth because of the regularity constraint on the surface S), our energy E_{total} does not directly contain a regularization term for the auxiliary vector field V. The values $V(\mathbf{X})$ of V at different points $\mathbf{X}$ on the surface S are then decorrelated. The optimal vector field $V(.)$ minimizing $E_{data} + \gamma E_{coupling}$ (and then E_{total}) is then the vector field V such that for all $\mathbf{X}$ in S, $V(\mathbf{X})$ is given by:

$$V(\mathbf{X}) \quad = \quad \arg\min_{\tilde{V} \in \mathbb{S}_2} \gamma\big(1 \ - \ \langle \tilde{V}, N(\mathbf{X})\rangle\big) \ + \ \sum_{i=1}^{n} \chi_i(\mathbf{X})\Big(I_i(\pi_i(\mathbf{X})) \ - \ \rho(\mathbf{X})\big(\langle \tilde{V}, L\xi(\mathbf{X})\rangle \ - \ E_0\big)\Big)^2 \sigma_i(\mathbf{X}). \tag{28}$$

Accordingly to (27) and (28), the vector field $V(.)$ which minimizes E_{total} verifies: for all fixed $\mathbf{X} \in S$, there exists a real scalar ν (depending on $\mathbf{X}$) such that

$$\nu V(\mathbf{X}) \quad = \quad \sum_{i=1}^{n} \chi_i(\mathbf{X}) \ \ (\ \ I_i(\pi_i(\mathbf{X})) \ - \ \rho(\mathbf{X})(\langle V(\mathbf{X}), L\rangle\, \xi(\mathbf{X}) \ + \ E_0) \ \ ) \ \ \rho(\mathbf{X}) L\xi(\mathbf{X}) \ \ \sigma_i(\mathbf{X}) \ + \ \frac{\gamma}{2} N(\mathbf{X}). \tag{29}$$

If $\xi(\mathbf{X}) = 0$, obviously the solution is $V(\mathbf{X}) = \frac{N(\mathbf{X})}{|N(\mathbf{X})|}$, else we have

$$\begin{aligned} \nu V(\mathbf{X}) = -\left(\sum_{i=1}^{n} \chi_i(\mathbf{X})\sigma_i(\mathbf{X})\right) \ \rho^2(\mathbf{X}) \ \langle V(\mathbf{X}), L\rangle \ \ L \\ + \sum_{i=1}^{n} \chi_i(\mathbf{X}) \ \Big(\ I_i(\pi_i(\mathbf{X})) - \rho(\mathbf{X}) E_0 \ \Big) \ \rho(\mathbf{X}) L\sigma_i(\mathbf{X}) + \frac{\gamma}{2} N(\mathbf{X}). \end{aligned} \tag{30}$$

If we denote

$$H = \left(\sum_{i=1}^{n} \chi_i \sigma_i\right) \ \rho^2 \ L \ L^T \quad \in M_{3\times 3}(\mathbb{R}) \tag{31}$$

and

$$B = \sum_{i=1}^{n} \chi_i \ (\ I_i(x_i) - \rho E_0 \ ) \ \rho L \sigma_i + \frac{\gamma}{2} N \quad \in \mathbb{R}^3 \tag{32}$$

(for simplicity, we have removed the $\mathbf{X}$ in the equations; x_i denotes $\pi_i(\mathbf{X})$)

then (30) becomes

$$(\nu Id_{3\times 3} + H)\ V = B. \tag{33}$$

Thus V has to verify: there exists ν in $\mathbb{R}$ such that

$$\begin{cases} \left|\ \left[\nu\ Id_{3\times 3} + H\right]^{-1} B\ \right| \ = \ 1, \\ V \ = \ \left[\nu\ Id_{3\times 3} + H\right]^{-1} B, \end{cases} \tag{34}$$

where $[.]^{-1}$ is the matrix inverse operator (as in section 3.2).

In the above equations, H and B are known. We get ν via Newton methods applied to the first equation of (34). The initial value ν_0 used in this iterative process is determined from the previous value of V (we denote $\hat{V}$) and other parameters of equation (29). More precisely, we get ν_0 by solving the following linear minimization problem

$$\nu_0 = \arg\min_{\nu} \Big| \sum_{i=1}^{n} \chi_i (I_i(x_i) - \rho\big(\langle \hat{V}, L\rangle + E_0\big))\ L\sigma_i\rho + \frac{\gamma}{2} N - \nu\hat{V} \Big|^2. \tag{35}$$

After calculating ν, we get V from the second equation of (34). Thus the computed V fulfills the requirement of $|V| = 1$ and it minimizes E_{total} on the manifold $\mathbb{S}_2$. While this is not strictly speaking a closed-form solution, it is simple and efficient to implement.

3.4 Ambiguities in illumination

As we previously mentioned in Section 2.1, model (1) is not minimal and there is a multiplicity ambiguity. In particular, when minimizing the energy with respect to ρ_1, ρ_2, L, E_0, we have that if $(\rho_1, \rho_2, E_0, L, C, S)$ is a solution, then so is $(\delta\rho_1, \delta\rho_2, E_0/\delta, L/\delta,\ C, S)$ with $\delta \in \mathbb{R}^+$.

Although the ambiguity makes it impossible to extract the albedo and illumination exactly, we could still recover $\rho E_0; |L|/E_0; L/|L|$, i.e., reconstruct the parameters up to a scalar transformation, or recover ρ_1/ρ_2. When we know the maximum albedo on the object, we would still expect to generate the exact albedos and the amplitude of ambient term and point light source as well. In practice, in each iterative step, we fix E_0 to a constant.

4 Experiments

In the first set of experiments, we took 28 calibrated images of a doll figure of approximately uniform albedo standing on a table. The background is dark, and the doll is illuminated both by standard fluorescent overhead lamps and by an additional strong spotlight. Note that the actual environment only approximately satisfies the

Figure 1: *Example views of the input data set consisting 28 images of a dancer statuette.*

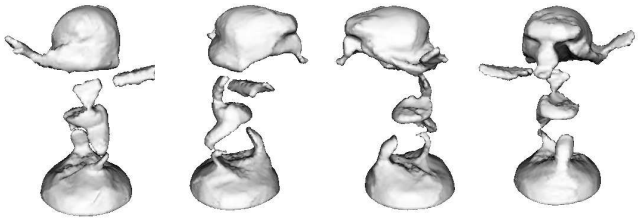
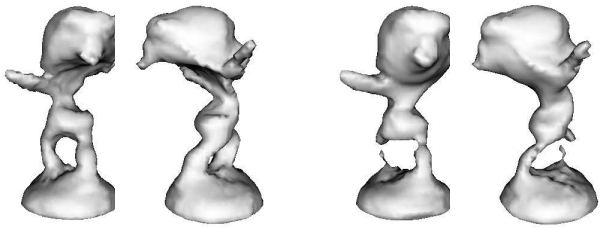


Figure 2: *Final shape estimated using [38]. The algorithm fails to reconstruct the doll, notably the legs and the back, because the assumption of constant radiance of the object is violated.*



Positive spot light Negative spot light

Figure 3: *Shape estimated by using the algorithm [12] with either a positive (left) or a negative (right) directional light. Compared with Figure 2, we obtained improved reconstructions in both light configurations over that of [38]. However, the results are still not satisfactory.*

conditions for which the algorithm is designed, but this is on purpose. In the next experiment we will illustrate the behaviour of the algorithm quantitatively on synthetic sequences generated to satisfy the assumptions precisely. Figure 1 shows 4 representative views that show how the light modulates the image intensity from light (front of the head) to dark (upper back). Notice also the non-trivial topology of the object, which is handled automatically by our algorithm as a side benefit of implementing the gradient flow within the level set framework [27].

The first test consisted in testing the stereoscopic segmentation algorithm [38] on this dataset. This corresponds to iterating the surface evolution in Section 3, without directional lighting, i.e. $L = 0$. The underlying assumption is that radiance, not just albedo, is constant, which is patently not the case for the dataset in question. As expected, the algorithm fails to converge to a viable estimate of the shape of the statuette (see Figure 2). In particular, the darker parts of the legs and back are ascribed to the background.

In the second test, we introduce one light source into the reconstruction problem. Figure 3 shows the results obtained from the algorithm presented in [12] with *one* punctual light source. In [12], the positivity of the light source was not enforced, and the underlying image-formation model was

$$I(\pi(\mathbf{X})) = \rho(\mathbf{X}) \left(\sum_i \lambda_i \langle N(\mathbf{X}), L_i \rangle \xi_i(\mathbf{X}) + E_0 \right)$$

where the λ_i and E_0 could be negative. The signs of these variables could change during the optimization process (see section 4.3 of [12]) and the final sign of the lights (returned after the convergence) depends on the initialization. Figure 3 shows two typical examples of reconstruction obtained from the algorithm of [12]; according to the initialization of the position of the light source, the light converged towards a positive (left image) or negative sign (right image). In either case the results improve stereoscopic segmentation [38], where light is not explicitly modeled. In order to improve the model [12], one needs two light sources; by using a positive and a negative light, one can obtain reasonable results such as those displayed in Figure 4. However, since the algorithm in [12] does not impose positivity, the resulting lighting configuration is physically impossible.

Next we tested the full-fledged algorithm described in the previous sections. Contrary to [12], here we enforce the positivity of the light sources. Figure 6 shows the results obtained with a single light source; the improvement over [12] are evident from Fig. 3, including the version with two light sources illustrated in Fig. 4. The improvement is particularly clear on details of the surface reconstruction, especially on the hands and lower back (Figure 7, compare with Figure 5).

In order to gain a more quantitative understanding of the operation of the algorithm, we tested it on various synthetic sequences, one of which we report in Figure 8 as an example. It is a sphere with constant albedo illuminated by an ambient light and a point light source. The results of the algorithm are shown in Figure 9, and the table in Figure 10 summarizes the numerical results.

In a final experiment, we tested the algorithm on a dataset that contained piecewise constant albedo, shown in Figure 11. In order to avoid the scale ambiguity, we rescale ρ_1, ρ_2, E_0, L in each evolution to $\delta = E_0/100$; $E_0 = E_0/\delta; |L| = |L|/\delta; \rho_1 = \rho_1\delta; \rho_2 = \rho_2\delta;$. Here 100 is just an arbitrary scale. Figure 12 shows the reconstructed shape and partitioning curve for the synthesized painted ball illuminated by directional light. Figure 13 shows the reconstructed shape for the Fish model. We also show the evolution process of the shape and partitioning curve in Figure 14.

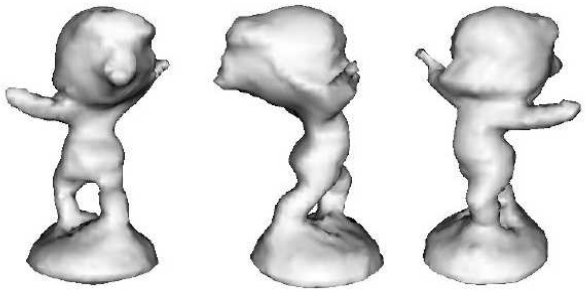


Figure 4: *Final shape estimated using the algorithm [12] with a combination of one positive light and one negative light. The algorithm reconstructs the 3D object more accurately (compared with Figure 3), nevertheless the lighting configuration if not physically plausible.*

5 Conclusions

We proposed an image formation model for Lambertian scenes with piecewise constant albedos illuminated by ambient and point light sources. We designed an algorithm to simultaneously infer all model parameters from a collection of calibrated images by minimizing a single energy. We solved the resulting infinite-dimensional optimization problem by numerically integrating partial differential equations that converge in steady states to a local minimum of the first variation of the functional. Our work extends the prior art on Stereoscopic Segmentation [38] by explicitly modeling and exploiting the light and the reflectance laws. It does so while enforcing physical constraints, such as the positivity of the energy distribution of light sources.

We also addressed a structural problem related to shape estimation under the standard Lambertian shading model, that results in unstable flows when the standard cost functional is minimized in a naive way. We introduced a smooth auxiliary vector field, that is coupled to the normal field via an energy term, and show that this model is not prone to instability like the naive one.

Naturally, the model we proposed is limited by the assumptions of Lambertian reflection and point/diffuse illumination. The assumption of piecewise constant albedo is not very restrictive in theory, as many man-made objects can be well approximated by this model, but in practice when objects become complex the topology and geometry of the albedo boundaries become so difficult that our algorithm may fail to capture subtle variations. One of the potential advantages of our approach is its potential to be integrated with multi-view stereo algorithms, by providing additional constraints in cases where objects with simple albedo do not provide enough constraints to establish point-to-point correspondence and therefore local feature-based approaches fail to provide a dense reconstruction.

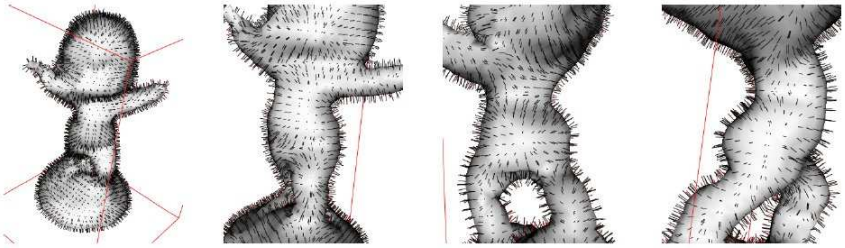


Figure 5: *Visualization of the auxiliary vector field V estimated in the same time as the surface displayed in Figure 4 via the algorithm [12] (with the combination of one positive light and one negative light).*

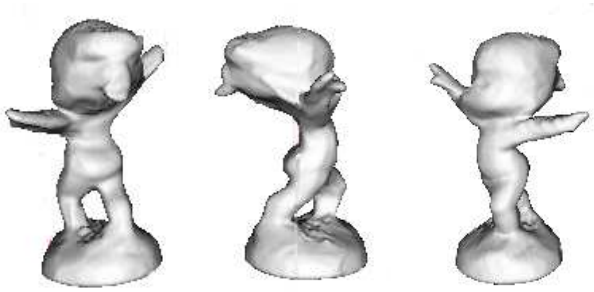


Figure 6: *Final shape estimated via the proposed algorithm containing the various optimization improvements (in particular the closed-form estimation of the auxiliary vector field) and the addition of the positivity constraint. The improvement on 3D shape details, especially around the hands and lower back, is visible.*

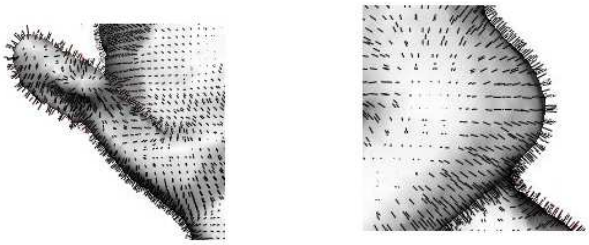


Figure 7: *Visualization of the auxiliary vector field via the proposed algorithm when we add the positivity constraint for the light and when we estimate the auxiliary vector field in closed-form.*

Acknowledgment

This work is supported by ONR N00014-03-1-0850:P0001 and AFOSR E-16-V91-G2.

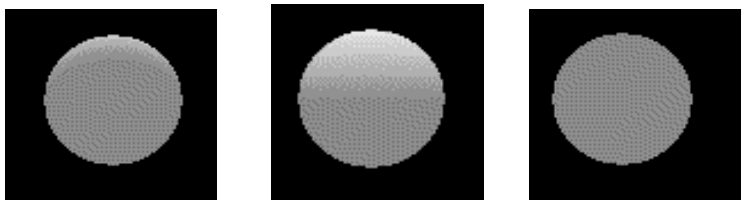


Figure 8: *Three views of the synthesized constant albedo sphere illuminated by a point light source of intensity 100 × located at (0,0,1), and an ambient light of intensity 100.*

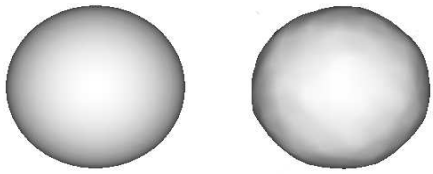


Figure 9: *Final shape estimated by the algorithm. The ground truth shape is shown on the left.*

References

[1] R. Abraham, J.E. Marsden, and T. Ratiu. *Manifolds, Tensor Analysis, and Applications*, volume 75 of *Applied Mathematical Sciences*. Springer, 2nd ed. edition, 1993.

[2] P. Belhumeur, D. Kriegman, and A. L. Yuille. The generalized bas relief ambiguity. *Intl. J. of Comp. Vis.*, 35:33–44, 1999.

[3] T. Brox and J. Weickert. Level set segmentation with multiple regions. *IEEE Transactions on Image Processing*, 15(10), October 2006.

[4] C. Chefd'hotel, D. Tschumperlé, R. Deriche, and O. Faugeras. Regularizing flows for constrained matrix-valued images. *J. Math. Imaging Vis.*, 20(1-2):147–162, 2004.

[5] H. F. Chen, P. N. Belhumeur, and D. W. Jacobs. In search of illumination invariants. In *Proc. IEEE Conf. on Comp. Vision and Pattern Recogn.*, 2000.

Reconstruction result	using positive light	ground truth.
L	(0.000,0.001,0.999)	(0.000,0.000,1.000)
E_0 Error	1.6%	100.00
Error of Intensity of point light	1.5 %	100.00
Shape Error	9 % %	0 (Note: Radius of ball is 10)

Figure 10: The table lists the results of the reconstruction results on synthesized data: A ball with radius 10 illuminated by a point light source and ambient light.

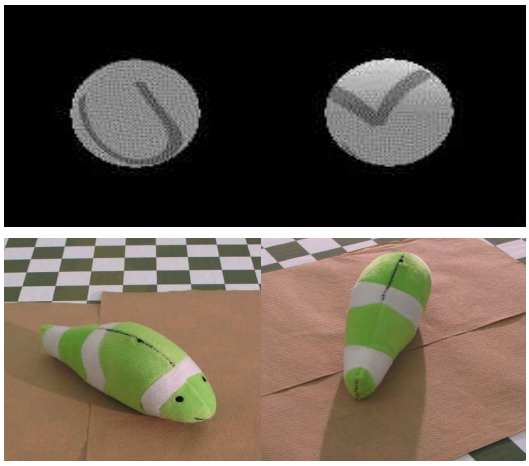


Figure 11: *The upper two images are 2 out of 26 views of a synthetic scene. The scene consists of one sphere painted with a word 'CV' in grey and the rest is white. The sphere is illuminated by ambient light of intensity* 100 *and a point light source of* $100 \times (0, 0.717, 0.717)^T$*; The lower two images are 2 out of 30 views of a real scene of a toy fish. Each image size is* 390×400 *and calibrated with a rig. The fish is green with white stripes. It is illuminated by a point light source from the head of the fish.*

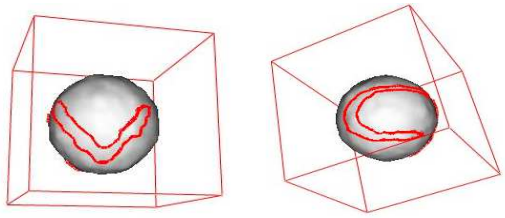


Figure 12: *The reconstructed shape and partitioning curve between regions with different albedo for the synthesized data in Figure 11. The two images show the final shape result of the proposed algorithm together with the red segmenting curve. The red straight lines show the coordinate frame for the spheres.*

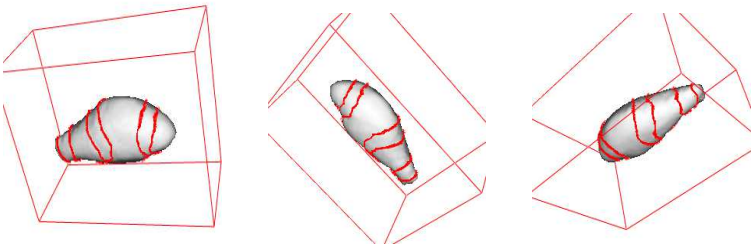


Figure 13: *Recovered shape via the algorithm proposed for the real scene shown in Figure 11. The red curves represents the segmenting curves for different albedo.*

[6] J-D. Durou, M. Falcone, and M. Sagona. A survey of numerical methods for shape from shading. Research report 2004-2-R, IRIT, January 2004.

[7] O. Faugeras and R. Keriven. Variational principles, surface evolution, pdes, level set methods, and the stereo

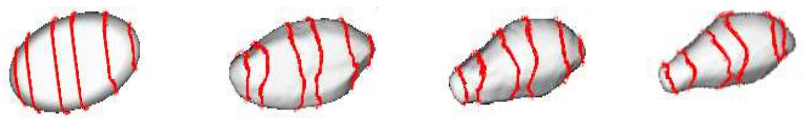


Figure 14: *The evolution process of shape and the segmenting curve for different albedo. The initial surface is an ellipsoid.*

problem. *IEEE Trans. on Image Processing*, 7(3):336–344, March 1998.

[8] B. Horn and M. Brooks (eds.). *Shape from Shading.* MIT Press, 1989.

[9] B. K. P. Horn. *Robot vision.* MIT press, 1986.

[10] R. Keriven J.-P. Pons and O. Faugeras. Multi-view stereo reconstruction and scene flow estimation with a global image-based matching score. to appear in the international journal of computer vision. *Intl. J. of Comp. Vis.*, 2006.

[11] H. Jin. Variational Methods for Shape Reconstruction in Computer Vision. Ph.D. thesis, Electrical Engineering Department, Washington University, August 2003.

[12] H. Jin, D. Cremers, A. Yezzi, and S. Soatto. Shedding light on stereoscopic segmentation. In *Proc. IEEE Conf. on Comp. Vision and Pattern Recogn.*, pages 36–42, 2004.

[13] H. Jin, S. Soatto, and A. J. Yezzi. Multi-view stereo beyond lambert. In *Proc. IEEE Conf. on Comp. Vision and Pattern Recogn.*, pages I–171–178, June 2003.

[14] H. Jin, A. J. Yezzi, and S. Soatto. Region-based segmentation on evolving surfaces with application to 3d shape and radiance estimation. In *Proc. of the Eur. Conf. on Comp. Vision*, pages 114–125, May 2004.

[15] H. Jin, A. Yezzi, and S. Soatto. Mumford-Shah on the Move: Region-based Segmentation on Deforming Manifolds with Application to 3-D Reconstruction of Shape and Appearance from Video. Submitted to *J. of Mathematical Imaging and Vision*, September 2006.

[16] H. Jin, A. J. Yezzi, Y.-H. Tsai, L.-T. Cheng, and 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, December 2003.

[17] C-Y. Kim, A.P. Petrov, H-K. Choh, Y-S. Seo, and I-S. Kweon. Illuminant direction and shape of a bump. *J. Opt. Soc. Am. A*, 15(9), pp.2341-2350, 1998.

[18] R. Klette, R. Kozera, and K. Schlüns. Shape from shading and photometric stereo methods. Technical Report CITR-TR-20, University of Auckland, New Zealand, 1998.

[19] J. Koenderink and A. van Doorn. Photometric invariants related to solid shape. *Optica Acta*, 27(7):981–996, 1980.

[20] K. Kolev, T. Brox, and D. Cremers. Robust variational segmentation of 3D objects from multiple views. In K. Franke et al., editor, *Pattern Recognition (Proc. DAGM)*, volume 4174 of *LNCS*, pages 688–697, Berlin, Germany, September 2006. Springer.

[21] H.W. Kuhn and A.W. Tucker. Nonlinear programming. In J. Neyman, editor, *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, pages 481–492. University of California Press, Berkeley, California, 1951.

[22] K. N. Kutulakos and S. M. Seitz. A theory of shape by space carving. *Intl. J. of Comp. Vis.*, 38(3):199–218, July 2000.

[23] M. S. Langer and S. W. Zucker. Shape from shading on a cloudy day. *Journal of Optical Society of America*, 11:467–478, 1994.

[24] Y. Ma, S. Soatto, J. Kosecka, and S. Sastry. *An invitation to 3D vision, from images to models.* Springer Verlag, 2003.

[25] S. Nayar, K. Ikeuchi, and T. Kanade. Surface reflection: physical and geometrical perspectives. *IEEE Trans. Pattern Anal. Mach. Intell.*, 13(7):611–634, 1991.

[26] J. Oliensis and P. Dupuis. A global algorithm for shape from shading. In *Proc. of the Intl. Conf. on Computer Vision*, pages 692–710, 1993.

[27] S. Osher and J. Sethian. Fronts propagating with curvature-dependent speed: algorithms based on hamilton-jacobi equations. *J. of Comp. Physics*, 79:12–49, 1988.

[28] E. Prados. *Application of the theory of the viscosity solutions to the Shape From Shading problem.* PhD thesis, Univ. of Nice-Sophia Antipolis, 2004.

[29] A. Robles-Kelly and E.R.Hancock. Estimating the surface radiance function from single images. *3DPVT*, 00:494–501, 2004.

[30] A. Robles-Kelly and E.R.Hancock. Surface radiance correction for shape-from-shading. *Pattern Recognition*, 38(10):1574–1595, 2005.

[31] D. Samaras. Illumination constraints in deformable models for shape and light direction estimation. *IEEE Trans. Pattern Anal. Mach. Intell.*, pages 247–264, 2003.

[32] D. Samaras, D. Metaxas, P.V. Fua, and Y.G. Leclerc. Variable albedo surface reconstruction from stereo and shape from shading. In *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*, pages I: 480–487, 2000.

[33] S. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski. A comparison and evaluation of multi-view stereo reconstruction algorithms. In *Proc. IEEE Conf. on Comp. Vision and Pattern Recogn.*, pages 519–526, 2006.

[34] A.J. Stewart and M. S. Langer. Towards accurate recovery of shape from shading under diffuse lighting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(9):1020–1025, 1997.

[35] H. Urakawa. *Calculus of Variations and Harmonic Maps.* American Mathematical Society, 1993.

[36] A. Vedaldi and S. Soatto. Viewpoint invariance for non-planar scenes. Technical Report TR050012, UCLA CSD, 2006.

[37] L.A. Vese and T.F. Chan. A multiphase level set framework for image processing using the Mumford–Shah functional. *Intl. J. of Comp. Vis.*, 50(3):271–293, 2002.

[38] A. Yezzi and S. Soatto. Stereoscopic segmentation. In *Proc. of the Intl. Conf. on Computer Vision*, pages 59–66, 2001.

[39] A. Yezzi and S. Soatto. Stereoscopic segmentation. *Intl. J. of Computer Vision*, 53(1):31–43, 2003.

[40] A. Yilmaz and M. Shah. Estimation of Arbitrary Albedo and Shape from Shading for Symmetric Objects. *In Proc. of British Machine Vision Conference*, 728–736, 2002.

[41] A. Yuille, J. M. Coughlan, and S. Konishi. Kgbr viewpoint-lighting ambiguity. *J. Opt. Soc. Am. A*, 20(1):24–31, 2003.

[42] A. L. Yuille, D. Snow, R. Epstein, and P. Belhumeur. Determining generative models of objects under varying illumination: shape and albedo from multiple images using svd and integrability. *Intl. J. of Comp. Vis.*, 35:203–222, 1999.

[43] R. Zhang, P.-S. Tsai, J Cryer, and M. Shah. Shape from shading: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(8):690–706, 1999.

[44] Q. Zheng and R. Chellappa. Estimation of illuminant direction, albedo and shape from shading. *IEEE Trans. Pattern Anal. Mach. Intell.*, 13(7):680–702, 1991.