Title

Region-based segmentation on evolving surfaces with application to 3D reconstruction of shape and piecewise constant radiance

Journal

Computer Vision - ECCV 2004, Pt 2, 3022

ISSN

0302-9743

Authors

Jin, H L
Yezzi, A J
Soatto, Stefano

Publication Date

2004

Peer reviewed

Region-based Segmentation on Evolving Surfaces with Application to 3D Reconstruction of Shape and Piecewise Constant Radiance

Hailin Jin[1], Anthony J. Yezzi[2], and Stefano Soatto[1]

[1] Computer Science Department, University of California, Los Angeles, CA 90095, hljin,soatto@cs.ucla.edu
[2] School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332, ayezzi@ece.gatech.edu

Abstract. We consider the problem of estimating the shape and radiance of a scene from a calibrated set of images under the assumption that the scene is Lambertian and its radiance is piecewise constant. We model the radiance segmentation explicitly using smooth curves on the surface that bound regions of constant radiance. We pose the scene reconstruction problem in a variational framework, where the unknowns are the surface, the radiance values and the segmenting curves. We propose an iterative procedure to minimize a global cost functional that combines geometric priors on both the surface and the curves with a data fitness score. We carry out the numerical implementation in the level set framework.

Keywords: variational methods, Mumford-Shah functional, image segmentation, multi-view stereo, level set methods, curve evolution on manifolds.

Fig. 1. *(COLOR) A plush model of "nemo." The object exhibits piecewise constant appearance. From a set of calibrated views, our algorithm can estimate the shape and the piecewise constant radiance.*

1 Introduction

Inferring three-dimensional shape and appearance of a scene from a collection of images has been a central problem in Computer Vision, known as multi-view stereo. Traditional approaches to this problem first match points or small

Fig. 2. *Man-made objects often exhibit piecewise constant appearance. Approximating their radiances with smooth functions would lead to gross error and "blurring" of the reconstruction. On the other hand, these objects are not textured enough to establish dense correspondence among different views. However, we can clearly see radiance boundaries that divided the objects into constant regions.*

image regions across different views and then combine the matches into a three-dimensional model[3]. Scene radiances can be reconstructed afterwards if necessary. These approaches effectively avoid directly estimating the scene radiances, which can be quite complex for real scenes. However, for these methods to work, the scene has to satisfy quite restrictive assumptions, namely the radiance has to contain "sufficient texture." When the assumptions are not fulfilled, traditional methods fail. Recently, various approaches have been proposed to fill the gaps where the assumptions underlying traditional stereo methods are violated and the scene radiance is assumed to be smooth, for instance [1, 2]. In this case, radiance is modeled explicitly rather than being annihilated through image-to-image matching. The problem of reconstructing shape and radiance is then formulated as a joint segmentation of all the images. The resulting algorithms are very robust to image noise.

However, there are certainly many scenes whose radiances are not heavily textured, but are not smooth either. For instance, man-made objects are often built with piecewise constant material properties and therefore exhibit approximately piecewise constant radiances, for instance those portrayed in Figure 2. For scenes like these, neither the assumption of having global constant or smooth radiances is satisfied, nor their radiances are textured enough to establish dense correspondence among different views.

For such scenes, one may attempt to use the algorithms designed for smooth radiances to reconstruct pieces of the surface that satisfy the assumptions and then patch them together to get the whole surface. Unfortunately, this approach does not work because patches are not closed surfaces, but even if they were, individual patches would not be able to explain the image data due to self-occlusions (see for instance Figure 4). Therefore, more complete and "global" models of the scene radiance are necessary. Our choice is to model it as a piecewise constant function. Under this choice, we can divide the scene into regions such that each region supports a uniform radiance, and the radiance is discontinuous across regions. The scene reconstruction problem now amounts to estimating the surface shape, the segmentation of the surface and the radiance value of each region.

[3] There are of course exceptions to this general approach, as we will discuss shortly.

1.1 Relation to prior work and contributions

This work falls in the category of multi-view stereo reconstruction. The literature on this topic is too large to review here, so we only report on closely related work. Faugeras and Keriven [3] were the first to combine image matching and shape reconstruction in a variational framework and carry out the numerical implementation using level set methods [4]. The underlying principle of their approach is still based on image-to-image matching and therefore their algorithm works for scenes that contain significant texture. Yezzi et. al. [1] and Jin et. al. [2] approached the problem by modeling explicitly a (simplified) model of image formation, and reconstruct both shape and radiance of the scene by matching images to the underlying model, rather than to each other directly. The class of scenes they considered is Lambertian with constant or smooth radiances. In this paper, we extend their work by allowing scenes to have *discontinuous* radiances and *model explicitly the discontinuities*. In the work of shape carving by Kutulakos and Seitz [5], matching is based on the notion of photoconsistency, and the largest shape that is consistent with all the images is recovered. We use a different assumption, namely that the radiance is piecewise constant, to recover a different representation (the smoothest shape that is photometrically consistent with the data in a variational sense) as well as photometry. Since we estimate curves as radiance discontinuities, this work is related to stereo reconstruction of space curves [6, 7]. The material presented in this paper is also closely related to a wealth of contributions in the field of image segmentation, particularly region-based segmentation, starting from Mumford and Shah's pioneering work [8] and including [9, 10].

We use curves on surfaces to model the discontinuities of the radiance. We use level set methods [4] to evolve both the surface and the curve to perform optimization. Our curve evolution scheme is closely related to [11–13].

We address the problem of multi-view stereo reconstruction for Lambertian objects that have piecewise constant radiances. To the best of our knowledge we are the first to address this problem. Our solution relies *not* on matching image-to-image, but on matching all images to the underlying model of both geometry and photometry.

For scenes that satisfy the assumptions, we reconstruct (1) the shape of the scene (a collection of smooth surfaces) and the radiance of the scene, which includes (2) the segmentation of the scene radiance, defined by smooth curves, and (3) the radiance value of each region. Our implementation contains several novel aspects, including simultaneously evolving curves (radiance discontinuities) on evolving surfaces, both of which are represented by level set functions.

2 A variational formulation

We model the scene as a collection of smooth surfaces and a background. We denote collectively with $S \subset \mathbb{R}^3$ all the surface, i.e., we allow S to have multiple connected components. We denote with $\mathbf{X} = [X, Y, Z]^T$ the coordinates of a generic point on S with respect to a fixed reference frame. We assume to be able to measure n images of the scene $I_i : \Omega_i \to \mathbb{R}, i = 1, 2, \ldots, n$, where Ω_i is

the domain of each image with area element $d\Omega_i$[4]. Each image is obtained with a calibrated camera which, after pre-processing, can be modeled as an ideal perspective projection $\pi_i : \mathbb{R}^3 \to \Omega_i; \mathbf{X} \mapsto \mathbf{x}_i \doteq \pi_i(\mathbf{X}) = \pi(\mathbf{X}_i) = [X_i/Z_i, Y_i/Z_i]^T$, where $\mathbf{X}_i = [X_i, Y_i, Z_i]^T$ are the coordinates for $\mathbf{X}$ 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 $R_i \in SO(3)$[5] and a translation vector $T_i \in \mathbb{R}^3$, so that $\mathbf{X}_i = R_i\mathbf{X} + T_i$. We assume that the background, denoted with B, covers the field of view of every camera. Without loss of generality, we assume B to be a sphere with infinite radius, which can therefore be represented using angular coordinates $\Theta \in \mathbb{R}^2$. We assume that the background supports a radiance function $h : B \to \mathbb{R}$ and the surface supports another radiance function $\rho : S \to \mathbb{R}$. We define the region $Q_i \doteq \pi_i(S) \subset \Omega_i$ and denote its complement by $Q_i^c \doteq \Omega_i \backslash Q_i$.

Our assumption is that the foreground radiance ρ is a *piecewise constant function*. We refer the reader to [14] for an extension to *piecewise smooth* radiances. For simplicity, the background radiance h is assumed to be constant, although extensions to smooth, piecewise constant or piecewise smooth functions can be conceived. Furthermore, we assume that the discontinuities of ρ can be modeled as a smooth closed curve C *on the surface* S, and C partitions S into two regions D_1 and D_2 such that $D_1 \cup D_2 = S$. Note that we allow each region D_i to have multiple disconnected components. Extensions to more regions are straightforward, for instance following the work of Vese and Chan [15]. We can thus re-define ρ as follows:

$$\rho(\mathbf{X}) = \rho_i \in \mathbb{R} \qquad \text{for } \mathbf{X} \in D_i, i = 1, 2. \tag{1}$$

We denote with $\pi_i(D_1)$ and with $\pi_i(D_2)$ the projections of D_1 and D_2 in the i-th image respectively.

2.1 The cost functional

The task is to reconstruct S, C, ρ_1, ρ_2, and h from the data $I_i, i = 1, 2, \ldots, n$. In order to do so, we set up a cost, E_{data}, that measures the discrepancy between the prediction of the unknowns and the actual measurements. Since some unknowns, namely the surface S and the curve C, live in infinite-dimensional spaces, we need to impose regularization to make the inference problem well-posed. In particular, we assume that both the surface and the curve are smooth (geometric priors), and we leverage on our assumption that the radiance is constant within each domain. The final cost is therefore the sum of three terms:

$$E(S, C, \rho_1, \rho_2, h) = E_{data} + \alpha E_{surf} + \beta E_{curv}, \tag{2}$$

[4] More precisely, measured images are usually non-negative discrete functions defined on a discrete grid and have minimum and maximum values. For ease of notation, we will consider them to be defined on continuous domains and take values from the whole real line.

[5] $SO(3) = \{R \mid R \in \mathbb{R}^{3\times 3} \text{ s.t. } R^T R = I \text{ and } \det(R) = 1\}$.

where $\alpha, \beta \in \mathbb{R}^+$ control the relative weights among the terms. The data fitness can be measured in the sense of $\mathcal{L}^2$ as:

$$E_{data} = \sum_{i=1}^{n} \left(\int_{\pi_i(D_1)} (I_i(\mathbf{x}_i) - \rho_1)^2 d\Omega_i + \int_{\pi_i(D_2)} (I_i(\mathbf{x}_i) - \rho_2)^2 d\Omega_i \right) + \sum_{i=1}^{n} \int_{Q_i^c} (I_i(\mathbf{x}_i) - h)^2 d\Omega_i, \tag{3}$$

although other function norms would do as well. The geometric prior for S is given by the total surface area:

$$E_{surf} = \int_S dA, \tag{4}$$

and that for C is given by the total curve length:

$$E_{curv} = \int_C ds, \tag{5}$$

where dA is the Euclidean area form of S and s is the arc-length parameterization for C. Therefore, the total cost takes the expression:

$$\begin{aligned} &E_{total}(S, C, \rho_1, \rho_2, h) = E_{data} + \alpha E_{surf} + \beta E_{curv} \\ &= \sum_{i=1}^{n} \left(\int_{\pi_i(D_1)} (I_i(\mathbf{x}_i) - \rho_1)^2 d\Omega_i + \int_{\pi_i(D_2)} (I_i(\mathbf{x}_i) - \rho_2)^2 d\Omega_i \right) \\ &+ \sum_{i=1}^{n} \int_{Q_i^c} (I_i(\mathbf{x}_i) - h)^2 d\Omega_i + \alpha \int_S dA + \beta \int_C ds. \end{aligned} \tag{6}$$

This functional is in the spirit of the Mumford-Shah functional for image segmentation [8].

3 Optimization of the cost functional

In order to find the surface S, the radiances ρ_1, ρ_2, h and the curve C that minimize the cost (6), we set up an iterative procedure where we start from a generic initial condition (typically a big cube, sphere or cylinder) and update the unknowns along their gradient directions until convergence to a (necessarily local) minimum.

3.1 Updating the surface

The gradient descent flow for the surface geometric prior is given by $S_t = 2\kappa N$, where κ is the mean curvature and N is the unit normal to S. Note that we have kept 2 in the expression in order to have the weights in the final flow match the weights in the cost (2). To facilitate computing the variation of the rest terms with respect to the surface, we introduce the radiance characteristic function ϕ

to describe the location of C for a given surface S. We define $\phi : S \to \mathbb{R}$ such that

$$D_1 = \{\mathbf{X} \mid \phi(\mathbf{X}) > 0\}, \quad D_2 = \{\mathbf{X} \mid \phi(\mathbf{X}) < 0\}, \quad \text{and } C = \{\mathbf{X} \mid \phi(\mathbf{X}) = 0\}. \tag{7}$$

ϕ can be viewed as the level set function of C. However, one has to keep in mind that ϕ is defined on S. We can then express the curve length as $\int_C ds = \int_S \|\nabla_S \mathcal{H}(\phi)\| dA$ where $\mathcal{H}$ is the Heaviside function. We prove in the technical report [14] that the gradient descent flow for the curve smoothness term has the following expression

$$S_t = \frac{\delta(\phi)}{\|\nabla_S \phi\|} \mathbf{II}(\nabla_S \phi \times N) N, \tag{8}$$

where $\mathbf{II}(\mathbf{t})$ denotes the second fundamental form of a vector $\mathbf{t} \in T_P(S)$, i.e. the normal curvature along $\mathbf{t}$ for $\|\mathbf{t}\| = 1$ and δ denotes the one-dimensional Dirac distribution: $\delta = \dot{\mathcal{H}}$. $T_P(S)$ is the tangent space for S at P. Note that $\nabla_S \phi \times N \perp N$ and therefore $\nabla_S \phi \times N \in T_P(S)$. Since flow (8) involves $\delta(\phi)$, it only acts on the places where ϕ is zero, i.e., the curve C. To find the variation of the data fitness term with respect to S, we need to introduce two more terms. Let $\chi_i : S \to \mathbb{R}$ 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 be the change of coordinates from $d\Omega_i$ to dA, i.e, $\sigma_i \doteq \frac{d\Omega_i}{dA} = \langle \mathbf{X}_i, N_i \rangle / Z_i^3$, where N_i the unit normal N expressed in the i-th camera reference frame. We now can express the data term as follows (see the technical report [14] for more details):

$$\begin{aligned} &\sum_{i=1}^n \left(\int_{\pi_i(D_1)} (I_i(\mathbf{x}_i) - \rho_1)^2 d\Omega_i + \int_{\pi_i(D_2)} (I_i(\mathbf{x}_i) - \rho_2)^2 d\Omega_i + \int_{Q_i^c} (I_i(\mathbf{x}_i) - h)^2 d\Omega_i \right) \\ &= \sum_{i=1}^n \int_S \chi_i \Gamma_i \sigma_i dA + \sum_{i=1}^n \int_{\Omega_i} (I_i(\mathbf{x}_i) - h)^2 d\Omega_i, \end{aligned} \tag{9}$$

where $\Gamma_i \doteq \mathcal{H}(\phi)(I_i - \rho_1)^2 + (1 - \mathcal{H}(\phi))(I_i - \rho_2)^2 - (I_i - h)^2$. Note that we have dropped the arguments for I_i and ϕ for ease of notation. Since, for a fixed h, $\sum_{i=1}^n \int_{\Omega_i} (I_i(\mathbf{x}_i) - h)^2 d\Omega_i$ does not depend upon the unknown surface, we only need to compute the variation of the first term $\sum_{i=1}^n \int_S \chi_i \Gamma_i \sigma_i dA$ with respect to S. We prove in the technical report [14] that the gradient descent flow for minimizing cost functionals of a general type $\sum_{i=1}^n \int_S \chi_i \Gamma_i \sigma_i dA$ takes the form:

$$S_t = \sum_{i=1}^n \frac{1}{Z_i^3} \Big(\Gamma_i \left\langle \chi_{i\mathbf{X}}, R_i^T \mathbf{X}_i \right\rangle - \chi_i \left\langle \Gamma_{i\mathbf{X}}, R_i^T \mathbf{X}_i \right\rangle \Big) N, \tag{10}$$

where $\chi_{i\mathbf{X}}$ and $\Gamma_{i\mathbf{X}}$ denote the derivatives of χ_i and Γ_i with respect to $\mathbf{X}$ respectively. We further note that $\left\langle I_{i\mathbf{X}}, R_i^T \mathbf{X}_i \right\rangle = 0$ [16] and obtain

$$\left\langle \Gamma_{i\mathbf{X}}, R_i^T \mathbf{X}_i \right\rangle = \delta(\phi) \big((I_i - \rho_1)^2 - (I_i - \rho_2)^2 \big) \left\langle \nabla_S \phi, R_i^T \mathbf{X}_i \right\rangle. \tag{11}$$

Therefore, the whole gradient descent flow for the cost (6) is given by

$$S_t = \Big(\sum_{i=1}^{n} \frac{\Gamma_i}{Z_i^3} \left\langle \chi_i \mathbf{X}, R_i^T \mathbf{X}_i \right\rangle - \chi_i \frac{\delta(\phi)}{Z_i^3} \big((I_i - \rho_1)^2 - (I_i - \rho_2)^2 \big) \left\langle \nabla_S \phi, R_i^T \mathbf{X}_i \right\rangle + 2\alpha\kappa + \beta \frac{\delta(\phi)}{\|\nabla_S \phi\|} \mathbf{II}(\nabla_S \phi \times N) \Big) N. \tag{12}$$

Note that flow (12) depends only upon the image values, *not the image gradients*. This property greatly improves the robustness of the resulting algorithm to image noise when compared to other variational approaches [3] to stereo based on image-to-image matching (i.e. less prone to become "trapped" in local minima).

3.2 Updating the curve

We show in the technical report [14] that the gradient descent flow for C related to the smoothness of the curve is given by

$$C_t = \Big(\sum_{i=1}^{n} \big((I_i - \rho_2)^2 - (I_i - \rho_1)^2 \big) \sigma_i + \beta \kappa_g \Big) \mathbf{n}, \tag{13}$$

where κ_g is the geodesic curvature and $\mathbf{n}$ is the normal to the curve in $T_P(S)$ (commonly referred to as the *intrinsic normal* to the curve C). Since $\mathbf{n} \in T_P(S)$, C stays in S as it evolves according to equation (13).

3.3 Updating the radiances

Finally, the optimization with respect to the radiances can be solved in closed forms as:

$$\begin{cases} \rho_1 = \frac{\sum_{i=1}^{n} \int_{\pi_i(D_1)} I_i(\mathbf{x}_i) d\Omega_i}{\sum_{i=1}^{n} \int_{\pi_i(D_1)} d\Omega_i} \\ \rho_2 = \frac{\sum_{i=1}^{n} \int_{\pi_i(D_2)} I_i(\mathbf{x}_i) d\Omega_i}{\sum_{i=1}^{n} \int_{\pi_i(D_2)} d\Omega_i} \\ h = \frac{\sum_{i=1}^{n} \int_{Q_i^c} I_i(\mathbf{x}_i) d\Omega_i}{\sum_{i=1}^{n} \int_{Q_i^c} d\Omega_i}, \end{cases} \tag{14}$$

i.e., the optimal values are the sample averages of the intensity values in corresponding regions.

4 A few words on the numerical implementation

In this section, we report some details on implementing the proposed algorithm. Both the surface and curve evolutions are carried out in the level set framework [4]. Since there has been a lot of work on shape reconstruction using level set methods and the space is limited, we refer the interested reader to [2, 3, 17] for general issues. We would like to point out that we do not include the term (8) in the implementation of the surface evolution because experimental testing has empirically shown that convinging results can be obtained even neglecting this term given its localized influence only near the segmenting curve. The numerical implementation of equations (14) should be also straightforward, since one only

needs to compute the sample average of the intensities in the regions $\pi_i(D_1)$, $\pi_i(D_2)$ and Q_i^c. Therefore, we will devote the rest of this section to issues related to the implementation of the flow (13).

Note that flow (13) is not a simple planar curve evolution. The curve is defined on the unknown surface, and therefore its motion has to be constrained on the surface. (It does not make sense to move the curve freely in $\mathbb{R}^3$, which would lead the curve out of the surface.) The way we approach the problem is to exploit the radiance characteristic function ϕ. Our approach is similar to the one considered by [11, 12]. Recall that C is the zeros of ϕ. We can express all the terms related to C using ϕ. In particular, the geodesic curvature is given by (we refer the reader to [14] for details on deriving this expression and the rest equations in this section):

$$\begin{aligned}\kappa_g &= \nabla_S \cdot \left(\frac{\nabla_S \phi}{\|\nabla_S \phi\|}\right) = \frac{\nabla_S \cdot \nabla_S \phi}{\|\nabla_S \phi\|} - \left\langle \nabla_S \phi, \nabla_S \left(\frac{1}{\|\nabla_S \phi\|}\right)\right\rangle \\ &= \frac{\Delta_S \phi}{\|\nabla_S \phi\|} - \frac{\nabla_S^T \phi \nabla_S^2 \phi \nabla_S \phi}{\|\nabla_S \phi\|^3},\end{aligned} \tag{15}$$

where $\nabla_S^2 \phi$ and $\Delta_S \phi$ denote the intrinsic Hessian and the intrinsic Laplacian of ϕ respectively. After representing the curve C with ϕ, we can implement the curve evolution by evolving the function ϕ on the surface. We further relax[6] ϕ from being a function defined on S to being a function defined on $\mathbb{R}^3$. This is related to the work of [13] for smoothing functions on surfaces. We denote with φ the extended function. We can then express the intrinsic gradient as follows:

$$\nabla_S \phi = \nabla \varphi - \langle \nabla \varphi, N \rangle N, \tag{16}$$

and the intrinsic Hessian as follows:

$$\nabla_S^2 \phi = (I - NN^T)\nabla^2 \varphi (I - NN^T) - (N^T \nabla \varphi) \frac{(I - NN^T)\nabla^2 \psi (I - NN^T)}{\|\nabla \psi\|}, \tag{17}$$

where ∇^2 stands for the standard Hessian in space and ψ is the level set function for S. $\Delta_S \phi$ can be computed as

$$\Delta_S \phi = \mathrm{trace}(\nabla_S^2 \phi) = \Delta \varphi - 2\kappa N^T \nabla \varphi - N^T \nabla^2 \varphi N. \tag{18}$$

Finally the curve evolution (13) is given by updating the following partial differential equation

$$\phi_t = \|\nabla_S \phi\| \sum_{i=1}^{n} \big((I_i - \rho_2)^2 - (I_i - \rho_1)^2\big)\sigma_i + \beta\Big(\Delta_S \phi - \frac{\nabla_S^T \phi \nabla_S^2 \phi \nabla_S \phi}{\|\nabla_S \phi\|^2}\Big), \tag{19}$$

with ϕ replaced by φ and $\nabla_S \phi$, $\nabla_S^2 \phi$ and $\Delta_S \phi$ replaced by the corresponding terms of φ according to equations (16), (17) and (18).

[6] This relaxation does not necessarily have to cover the entire $\mathbb{R}^3$. It only needs to cover the regions where the numerical computation operates.

Fig. 3. *(COLOR) The left* 2 *images are* 2 *out of* 26 *views of a synthetic scene. The scene consists of two spheres, each of which is painted in black with the word "ECCV". The rest of the spheres is white and the background is gray. Each image is of size* 257×257*. The right* 2 *images are* 2 *out of* 31 *images from the "nemo" dataset. Each image is of size* 335×315 *and calibrated manually using a calibration rig.*

5 Experiments

In Figure 3 (left 2 images) we show 2 out of 26 views of a synthetic scene, which consists of two spheres. Each image is of size 257×257. Each sphere is painted in black with the word "ECCV" and the rest is white. The background is gray. Clearly modeling this scene with one single constant radiance would lead to gross errors. One cannot even reconstruct either white or black part using the smooth radiance model in [1, 2] due to occlusions. For comparison purpose, we report the results of our implementation of [1] in Figure 4 (the right 2 images). The left 4 images in Figure 4 show the final reconstructed shape using the proposed algorithm. The red curve is where the discontinuities of the radiance are. The explicit modeling of radiance discontinuities may enable further applications. For instance, one can flatten the surface and the curve and perform character recognition of the letters. The numerical grid used in both algorithms is the same and of size $128 \times 128 \times 128$. In Figure 5 we show the surface evolving from a large cylinder to a final solid model. The foreground in all the images is rendered with its estimated radiance values (ρ_1 and ρ_2) and the segmenting curve is rendered in red. In Figure 6 we show the images reconstructed using the estimated surface, radiances and segmenting curve compared with an actual image from the original dataset.

In Figure 3 (right 2 images) we show 2 out of 31 views of a real scene, which contains a plush model of "nemo". The intrinsics and extrinsics of all the images are calibrated off-line. Each image is of size 335×315. Nemo is red with white stripes. For the proposed algorithm to work with color images, we have extended the model (6) as follows: we consider images to take vector values (RGB color in our case) and modify the square error between scalars in equation (6) to the simple square of Euclidean vector norm. In Figure 7 we show several shaded views of the final reconstructed shape using the proposed algorithm. The radiance discontinuities are rendered as green curves. The numerical grid used here is of size $128 \times 60 \times 100$. In Figure 8 we show the surface evolving from an initial shape that neither contains nor is contained in the shape of the scene, to a final solid model. The foreground in all the images is rendered with its

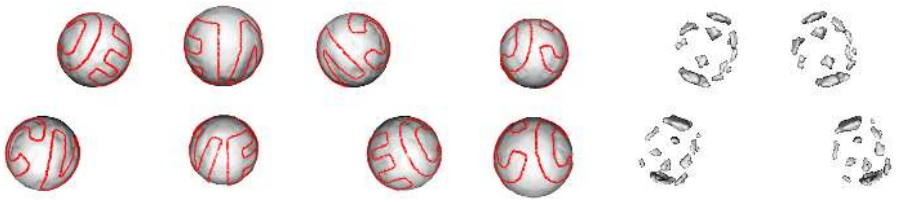

Fig. 4. *(COLOR) The first* 4 *images are shaded views of the final shape estimated using the proposed algorithm. Radiance discontinuities have been rendered as red curves. The location of the radiance discontinuities can be exploited for further purposes, for instance character recognition. The last* 2 *images are the results of assuming that the foreground has constant radiance, as in [1]. Note that the algorithm of [1] cannot capture all the white parts or all the black parts of the spheres, because that is not consistent with the input images due to occlusions.*

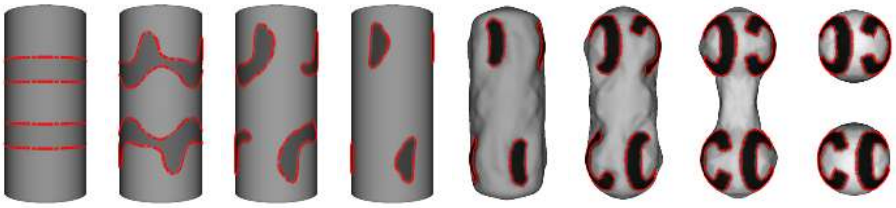

Fig. 5. *(COLOR) Rendered surface during evolution. The foreground in all the images is rendered with the current estimate of the radiance (ρ_1 and ρ_2) plus some shading effects for ease of visualization.*

Fig. 6. *The first image is just one view from the original dataset. The remaining* 6 *images are rendered using estimates from different stages of the estimation process. In particular, the second image is rendered using the initial data and the last image is rendered using the final estimates.*

estimated radiance values (ρ_1 and ρ_2) and the segmenting curve is rendered in green. In Figure 9 we show the images reconstructed using the estimated surface, radiances and segmenting curve compared with one actual image in the original dataset.

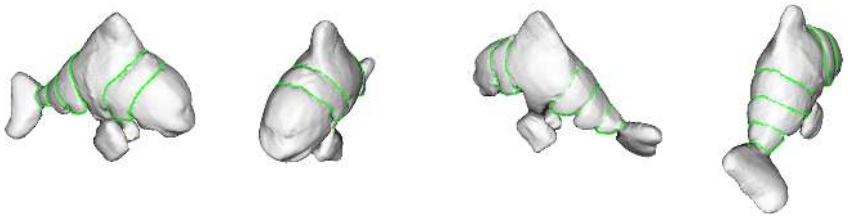

Fig. 7. *(COLOR) Several shaded views of the final reconstructed surface. The radiance discontinuities have been highlighted in green.*

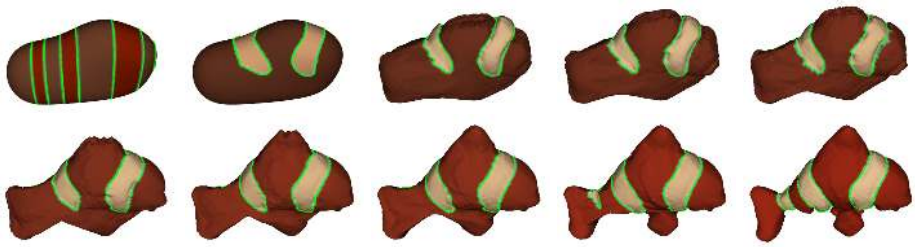

Fig. 8. *(COLOR) Rendered surface during evolution. Notice that the initial surface is neither contained nor contains the actual object. The foreground in all the images are rendered with the current estimate of the radiance values (ρ_1 and ρ_2) plus some shading effects for ease of visualization.*

Fig. 9. *(COLOR) The first image is just one view from the original dataset. The remaining* 6 *images are rendered using estimates from different stages of the estimation process. In particular, the second image is rendered using the initial data and the last image is rendered using the final estimates.*

6 Conclusions

We have presented an algorithm to reconstruct the shape and radiance of a Lambertian scene with piecewise constant radiance from a collection of calibrated views. We set the problem in a variational framework and minimize a cost functional with respect to the unknown shape, unknown radiance values in each region, and unknown radiance discontinuities. We use gradient-descent partial differential equations to simultaneously evolve a surface in space (shape), a curve defined on the surface (radiance discontinuities) and radiance values of each region, which are implemented numerically using level set methods.

Acknowledgements

This work is supported by AFOSR grant F49620-03-1-0095, NSF grants IIS-0208197, CCR-0121778 and CCR-0133736, and ONR grant N00014-03-1-0850. We would like to thank Daniel Cremers for stimulating discussions and Li-Tien Cheng for suggestions on implementing curve evolution on surfaces.

References

1. Yezzi, A.J., Soatto, S.: Stereoscopic segmentation. In: Proc. of Intl. Conf. on Computer Vision. Volume 1. (2001) 59–66
2. Jin, H., Yezzi, A.J., Tsai, Y.H., Cheng, L.T., Soatto, S.: Estimation of 3d surface shape and smooth radiance from 2d images: A level set approach. J. Scientific Computing **19** (2003) 267–292
3. Faugeras, O., Keriven, R.: Variational principles, surface evolution, pdes, level set methods, and the stereo problem. IEEE Trans. on Image Processing **7** (1998) 336–344
4. Osher, S.J., Sethian, J.A.: Fronts propagating with curvature dependent speed: Algorithms based on hamilton-jacobi formulations. J. Comput. Phys. **79** (1988) 12–49
5. Kutulakos, K.N., Seitz, S.M.: A theory of shape by space carving. Int. J. of Computer Vision **38** (2000) 199–218
6. Kahl, F., August, J.: Multiview reconstruction of space curves. In: Proc. of Intl. Conf. on Computer Vision. Volume 2. (2003) 1017–1024
7. Schmid, C., Zisserman, A.: The geometry and matching of lines and curves over multiple views. Int. J. of Computer Vision **40** (2000) 199–233
8. Mumford, D., Shah, J.: Optimal approximations by piecewise smooth functions and associated variational problems. Comm. Pure Appl. Math. **42** (1989) 577–685
9. Chan, T.F., Vese, L.A.: Active contours without edges. IEEE Trans. on Image Processing **10** (2001) 266–277
10. Yezzi, A.J., Tsai, A., Willsky, A.: A statistical approach to snakes for bimodal and trimodal imagery. In: Proc. of Intl. Conf. on Computer Vision. Volume 2. (1999) 898–903
11. Cheng, L.T., Burchard, P., Merriman, B., Osher, S.J.: Motion of curves constrained on surfaces using a level-set approach. J. Comput. Phys. **175** (2002) 604–644
12. Kimmel, R.: Intrinsic scale space for images on surfaces: the geodesic curvature flow. Graphical Models and Image Processing **59** (1997) 365–372
13. Bertalmio, M., Cheng, L., Osher, S.J., Sapiro, G.: Variational problems and partial differential equations on implicit surfaces. J. Comput. Phys. **174** (2001) 759–780
14. Jin, H., Yezzi, A.J., Soatto, S.: Region-based segmentation on evolving surfaces with application to 3d reconstruction of shape and piecewise smooth radiance. Technical Report CSD-TR04-0004, University of California at Los Angeles (2004)
15. Vese, L.A., Chan, T.F.: A multiphase level set framework for image segmentation using the mumford and shah model. Int. J. of Computer Vision **50** (2002) 271–293
16. Soatto, S., Yezzi, A.J., Jin, H.: Tales of shape and radiance in multi-view stereo. In: Proc. of Intl. Conf. on Computer Vision. Volume 1. (2003) 171–178
17. Jin, H., Soatto, S., Yezzi, A.J.: Multi-view stereo beyond lambert. In: Proc. of IEEE Conf. on Computer Vision and Pattern Recognition. Volume 1. (2003) 171–178