Journal of Mathematical Imaging and Vision 18: 7–15, 2003

Segmentation with Depth but Without Detecting Junctions

SELIM ESEDOGLU*
Institute for Mathematics and its Applications, University of Minnesota, Minneapolis, MN 55455, USA
esedoglu@math.ucla.edu

RICCARDO MARCH
Istituto per le Applicazioni del Calcolo, CNR, Viale del Policlinico, 137, 00161 Rome, Italy

Abstract. Given an image that depicts a scene with several objects in it, the goal of segmentation with depth is to automatically infer the shapes of the objects and the occlusion relations between them. Nitzberg, Mumford and Shiota formulated a variational approach to this problem: in their model, the solution is obtained as the minimizer of an energy. We describe a new technique of minimizing their energy that avoids explicit detection/connection of T-junctions.

Keywords: segmentation with depth, 2.1D sketch, disocclusion, variational methods, functionals with curvature, gamma convergence, nonlinear PDEs

1. Introduction

We describe a new numerical technique for minimizing the Nitzberg-Mumford-Shiota functional [16] that appears in the variational formulation of the *segmentation with depth* problem. This is a segmentation model that allows regions to overlap in order to take into account the partial occlusion of farther objects by those that are nearer. The regions are endowed with an ordering relation indicating which are in front of each other. The minimization of the functional yields both the shapes of the objects in an image, with a reconstruction of the occluded boundaries, and the ordering of the objects in space.

In [16] the authors do not attempt to solve numerically the variational problem of minimizing the functional in general. Instead, they first compute edges and T-junctions in the image, then minimize the functional combinatorically with respect to the ordering relation using the computed T-junctions. A drawback of this approach is that the preliminary edge detection step leaves gaps at the T-junctions, so that a further junction restoration step is needed.

*To whom correspondence should be addressed.

The technique presented in this paper is novel: unlike in previous approaches to this problem, it completely avoids explicit detection and connection of T-junctions in the given image $f(x)$. Instead, we minimize the energy more directly: our algorithm is based on evolving regions (or their boundaries given by a set of curves) in the plane. In this respect, our approach is motivated by Chan and Vese's work in [5] and [17] on minimizing the Mumford-Shah segmentation functional.

Perhaps the hardest aspect of working with the Nitzberg-Mumford-Shiota energy is having to deal with the curvature of the boundaries of the regions: it contains terms of the form

$$E := \int_{\partial A} \phi(k)\, ds, \tag{1}$$

where k denotes the curvature of the boundary of the set A and ϕ is a positive, convex, even function. The presence of curvature makes it difficult both to discretize the functional E and to apply gradient descent with respect to the unknown boundary ∂A (see for instance the approach used in [4]).

De Giorgi, Bellettini and Paolini described in [2, 3, 7] techniques for approximating such energies by much more tractable ones; the approximation takes place in

the sense of Γ-convergence. Our numerical implementation utilizes an approximation to the functional E, with $\phi(k) = k^2$, by means of a sequence of elliptic functionals. The approximation is inspired by a conjecture that De Giorgi made in [7], and it was numerically experimented by one of the authors in [11] for the segmentation (without depth) problem. Also, it has been recently applied to the related problem of image inpainting in [8].

We present numerical examples on artificial images, resembling those considered in [16]. These examples show how our technique leads to interaction and connection of T-junctions automatically.

2. Background: The 2.1D Sketch Model

Given a two-dimensional grayscale image $f(x):\Omega \to [0,1]$ of a three dimensional scene with several objects in it, the goal of *segmentation with depth* is to determine the shapes of the objects that make up the scene, as well as their ordering in space.

This is an extremely difficult problem, not least because in real scenes there might not be a well defined order between the objects. Such is the case, for example, when the objects exhibit self-occlusions and entanglements. It is therefore necessary to make some simplifying assumptions.

The *2.1D sketch* refers to a theory introduced by Nitzberg, Mumford and Shiota in [15, 16] to tackle the segmentation with depth problem. Here, it is assumed that $f(x)$ is the two-dimensional image of a (simplified) three-dimensional scene in which:

- Object surfaces do not exhibit self-occlusions.
- Objects are not entangled: each object is either completely in front of or completely behind any given other.
- The gray-scale intensity of each object is approximately constant.

These assumptions allow objects in the foreground to occlude objects in the background. For simplicity, we will assume a *strict* ordering between the regions, so that any two distinct objects cannot be at the same depth: one of them is necessarily closer to the observer.

According to the Nitzberg, Mumford and Shiota theory such a simplified three-dimensional scene is described by the following, which are the *unknowns* of the 2.1D sketch problem:

- The number n of objects in the scene.
- The regions $R_1, \ldots, R_n$ that these objects occupy in the image plane (i.e. their *shapes*).
- The approximate gray-scale intensities $c_1, \ldots, c_n$ of the objects.
- The order relation between the regions (out of $n!$ possibilities).

Here, $R_1, \ldots, R_n$ are subsets of the image domain $\Omega \subset \mathbf{R}^2$, and the constants $c_1, \ldots, c_n$ are in $[0, 1]$. Since occlusions are allowed in this model, the regions R_j need not be disjoint. Furthermore, since we are assuming a strict ordering between the regions, we may adopt the following convention:

$$j < i \Rightarrow R_j \text{ is in front of } R_i.$$

In the two-dimensional image, a region R_i might not be completely visible; parts of it might be occluded by objects that are in front of it. The **visible** part R_i' of the i-th region R_i is given by:

$$R_1' := R_1, \quad R_i' := R_i - \bigcup_{j<i} R_j, \quad \text{for } i = 2, \ldots, n.$$

Moreover we denote by $R_{n+1}' := \Omega - \bigcup_{j=1}^{n} R_j$ the background region, and by c_{n+1} the corresponding gray-scale intensity.

The two-dimensional observed image $u(x)$ corresponding to such a three–dimensional scene can be expressed in terms of these unknowns (the regions R_i, the constants c_i, and the order relation) as follows:

$$u(x) = \sum_{i=1}^{n+1} c_i \chi_{R_i'}(x).$$

The goal of the 2.1D sketch problem is the inverse process of obtaining the regions, the order relation, and the approximate gray-scale intensities from the two-dimensional image.

In [15, 16] Nitzberg, Mumford and Shiota gave a variational formulation to this problem by looking for a minimizer of the following energy:

$$E_{2.1} := \sum_{i=1}^{n} \int_{\partial R_i \cap \Omega} [\alpha + \beta\phi(k)]\, ds + \sum_{i=1}^{n+1} \int_{R_i'} (f(x) - c_i)^2\, dx, \qquad (2)$$

which is to be minimized over the number of objects n, the regions R_i, and the constants c_i. The function $\phi(x)$ is to be a positive, convex, even function such that $\phi(x) = x^2$ when $|x|$ is small, but with linear growth

at infinity, in order to allow corners along ∂R_i. The (open) set Ω is considered as a window on the image plane, so that the contour integrals exclude portions of the boundary ∂R_i that lie outside Ω.

In [16] an additional term is included in the energy to distinguish between certain ambiguous scenes; for the present we will ignore it: it is neither a difficult term to deal with, nor essential for the examples we will consider.

3. The Approximation Method

Our strategy for minimizing (2) is first to write down a sequence of approximating functionals that are numerically more tractable. Our inspiration comes from the notion of Γ-convergence and the many subsequent applications of this notion to the approximation of functionals that appear in various contexts [1, 2, 10, 11, 13]. We use an approach proposed by one of the authors in [11] to construct our approximation. The Γ-convergence, introduced by De Giorgi, makes it possible to go to the limit in the corresponding minimization problem: the minimizers of the approximating functionals converge, in an appropriate metric, to the minimizers of the limit functional. For the properties of Γ-convergence see [6], and [10] for a simple explanation.

In this paper, we will concentrate on the case

$$\phi(x) = x^2,$$

thereby deviating from the original formulation proposed by Nitzberg, Mumford, and Shiota. The motivation for doing so is to be able to use an approximation conjectured by De Giorgi in [7] for the term (1) with $\phi(x) = x^2$. Some disadvantages of this are indicated in the conclusions section.

Particularly relevant to the present work is the paper [2] by Bellettini, where an approximation for functionals of type (1) is proposed, and the Γ-convergence of the approximations is rigorously established. We opted for the approximation method of De Giorgi mainly because its numerical treatment seems to be simpler.

We will keep track of the regions $R_1, \ldots, R_n$ via approximations $u_1(x), \ldots, u_n(x)$ to their characteristic functions:

$$u_i(x) : \Omega \to \mathbf{R}, \qquad \text{with } u_i(x) \approx \chi_{R_i}(x),$$

where the approximation takes place in the L^1-topology. Moreover we define $R_{n+1} := \Omega$ and $u_{n+1} := 1$. We now explain how the various terms that appear in (2) are approximated in terms of the functions $u_i(x)$. Let $W(t) := t^2(1-t)^2$.

Length Term: Part of the boundary integral in energy (2) involves the length of the boundaries of the regions R_i. A theorem of Modica and Mortola [12, 13] tells us how these terms can be approximated:

$$\int_\Omega \left(\varepsilon |\nabla u_i|^2 + \frac{1}{\varepsilon} W(u_i) \right) dx \xrightarrow[\Gamma]{\varepsilon \to 0^+} \int_{\partial R_i \cap \Omega} ds,$$

where convergence takes place in the sense of Γ-convergence, up to a (multiplicative) constant that depends on the function $W(t)$. The approximate functionals remain bounded as $\varepsilon \to 0^+$ only if u_i converge, in L^1, to the characteristic function of a set R_i; the meaning of Γ-convergence is that the minimum possible value of the approximate functionals along such a sequence tends to the length of $\partial R_i \cap \Omega$.

Curvature Term: based on a conjecture by De Giorgi in [7], we propose to approximate the boundary integrals in (2), including the curvature dependent part, as follows:

$$\begin{aligned} \alpha \int_\Omega \left(\varepsilon |\nabla u_i|^2 + \frac{1}{\varepsilon} W(u_i) \right) dx \\ + \frac{\beta}{\varepsilon} \int_\Omega \left(2\varepsilon \Delta u_i - \frac{1}{\varepsilon} W'(u_i) \right)^2 dx \\ \xrightarrow{\varepsilon \to 0^+} \int_{\partial R_i \cap \Omega} (\alpha + \beta k^2)\, ds, \end{aligned}$$

where all multiplicative constants of the approximation are absorbed into the weights α and β of the functional $E_{2.1}$. This convergence has not been proved completely; an informal explanation can be found in [11], and some partial results can be found in [3, 9]. The meaning of the approximation is the following: since the Modica-Mortola functionals approximate the length of ∂R_i, and the curvature is the first variation of the length functional, then the contour integral of the square of curvature is approximated by the domain integral of the square first variation of the Modica-Mortola functional; the factor $1/\varepsilon$ multiplying β is necessary to ensure the convergence to the desired limit functional.

Fidelity Terms: These are the easiest. Since $u_i(x)$ is an approximation to the characteristic function of R_i, we have:

$$\begin{aligned} 1 - u_i(x) &\approx \chi_{R_i^c}(x), \\ u_i(x) u_j(x) &\approx \chi_{R_i \cap R_j}(x), \end{aligned}$$

and so on. These observations lead to:

$$\int_{R_i'} (f(x) - c_i)^2 \, dx \approx \int_\Omega (f(x) - c_i)^2 \, \eta(u_i) \prod_{j<i} \eta(1 - u_j) \, dx,$$

where $\eta(x)$ is a positive function increasing on $x \in [0, 1]$ with $\eta(0) = 0$ and $\eta(1) = 1$. In principle, η can be chosen to be as simple as $\eta(x) = x^2$.

Putting together the various terms discussed above, we arrive at the following approximation to the Nitzberg-Mumford-Shiota functional:

$$\begin{aligned} E_\varepsilon := \sum_{i=1}^{n} \int_\Omega \Bigg[& \alpha \left(\varepsilon |\nabla u_i|^2 + \frac{1}{\varepsilon} W(u_i) \right) \\ & + \frac{\beta}{\varepsilon} \left(2\varepsilon \Delta u_i - \frac{1}{\varepsilon} W'(u_i) \right)^2 \Bigg] dx \\ & + \sum_{i=1}^{n+1} \int_\Omega \eta(u_i)(c_i - f)^2 \prod_{j<i} \eta(1 - u_j) \, dx, \end{aligned} \tag{3}$$

which is to be minimized over the functions $u_1, \ldots, u_n$, and the constants $c_1, \ldots, c_{n+1}$ (assuming for the moment that we know the number of regions n). To get the Euler-Lagrange equations, we take first variations with respect to the functions u_k; that yields:

$$\begin{aligned} & 4\beta \Delta v_k - \left(\alpha + \frac{2\beta}{\varepsilon^2} W''(u_k) \right) v_k + \eta'(u_k)(c_k - f)^2 \\ & \quad \times \prod_{j<k} \eta(1 - u_j) - \sum_{i>k} \eta(u_i)(c_i - f)^2 \eta'(1 - u_k) \\ & \quad \times \prod_{\substack{j<i \\ j \neq k}} \eta(1 - u_j) = 0, \end{aligned} \tag{4}$$

for $k = 1, \ldots, n$, where

$$v_k := 2\varepsilon \Delta u_k - \frac{1}{\varepsilon} W'(u_k). \tag{5}$$

Equation (4) is fourth order in u_k. Variations with respect to the constants c_k give:

$$\begin{aligned} c_k = & \left(\int_\Omega f \, \eta(u_k) \prod_{j<k} \eta(1 - u_j) \, dx \right) \\ & \times \left(\int_\Omega \eta(u_k) \prod_{j<k} \eta(1 - u_j) \, dx \right)^{-1}, \end{aligned} \tag{6}$$

for $k = 1, \ldots, n + 1$. Hence, as $\varepsilon \to 0^+$, c_k approximates the average of the original image $f(x)$ over the *visible* part of the k-th region R_k, which is just the value of the constant gray-scale intensity that minimizes the functional $E_{2.1}$.

We will explain in some detail our numerical treatment of Eqs. (4)–(6) in the following section.

4. Numerical Implementation

Our method for solving Eqs. (4)–(6) is based on an *iteration* (or time marching) that requires suitable *intial guesses* for the regions R_i and the constants c_i.

4.1. Choosing an Initial Guess

A preliminary segmentation without depth (obtained for instance by minimizing the Mumford-Shah functional [14]) is a natural way to obtain initial guesses for the regions and the constants, as well as an estimate for the number of objects in the scene. Indeed, it is reasonable to assume that each region found by a successful segmentation (without depth) algorithm should be part of a *single* (but possibly larger) region found by the segmentation with depth algorithm. In other words, when compared to standard segmentation, in most situations segmentation with depth would be expected to yield fewer and bigger regions.

More precisely, if $A_1, \ldots, A_{n+1}$ are all the regions found by a segmentation (without depth) algorithm, so that:

$$\Omega \approx \bigcup_{i=1}^{n+1} A_i,$$

then to minimize (2) we need to start with n functions $u_1, \ldots, u_n$ and $n + 1$ constants $c_1, \ldots, c_{n+1}$ that are initialized as:

$$\begin{aligned} u_j(x, t = 0) &= \chi_{A_{\sigma(j)}}(x) \quad \text{for } j = 1, \ldots, n, \\ c_j(t = 0) &= \frac{1}{|A_{\sigma(j)}|} \int_{A_{\sigma(j)}} f(x) \, dx, \end{aligned} \tag{7}$$

where σ is a permutation on n objects, and A_{n+1} is assumed to be the initial guess for the background region. There are thus $n!$ ways to initialize the segmentation with depth algorithm based on the n regions found by a presegmentation. Our technique requires minimizing (2) separately for each possible initialization, because (based on numerical experiments) different initial configurations often lie in the basin of attraction of different local minimizers of (2).

4.2. Treatment of the PDE System

Recall that the Euler-Lagrange equation for the function u_k turns out to be a fourth order, non-linear elliptic equation, the highest order term of which is given by the biharmonic operator acting on u_k. There are many ways to approach its numerical solution; for instance, its parabolic version (i.e. gradient descent for the associated energy) enjoys a certain similarity to the Cahn-Hilliard equation, whose numerical solution has been extensively studied. We found it convenient to replace each fourth order equation by the two coupled *second order* equations expressed in (4) and (5), and compute the solution to the following evolutionary equations up to a large value of time t:

$$\partial_t u_k = 2\varepsilon \Delta u_k - \frac{1}{\varepsilon} W'(u_k) - v_k, \tag{8}$$

and

$$\begin{aligned}\partial_t v_k = {} & 4\beta \Delta v_k - \left(\alpha + \frac{2\beta}{\varepsilon^2} W''(u_k)\right) v_k \\ & + \eta'(u_k)(c_k - f)^2 \prod_{j<k} \eta(1-u_j) \\ & - \sum_{i>k} \eta(u_i)(c_i - f)^2 \eta'(1-u_k) \prod_{\substack{j<i \\ j\neq k}} \eta(1-u_j).\end{aligned} \tag{9}$$

For convenience, we imposed periodic boundary conditions (which we expect is not a significant matter for the sample scenes we considered, where all objects are well separated from the boundary of the domain). We discretized Eqs. (8) and (9) in space and time. Let δt be the time step size, and let u_k^l and v_k^l denote the discrete time approximation to u_k and v_k at lth time step (i.e. at $t = l\delta t$). Our time discretization scheme is as follows:

$$\frac{u_k^{l+1} - u_k^l}{\delta t} - 2\varepsilon \Delta u_k^{l+1} = -\frac{1}{\varepsilon} W'\big(u_k^l\big) - v_k^l,$$

and

$$\begin{aligned}&\frac{v_k^{l+1} - v_k^l}{\delta t} - 4\beta \Delta v_k^{l+1} + \alpha v_k^{l+1} \\ &\quad = -\frac{2\beta}{\varepsilon^2} W''\big(u_k^l\big) v_k^l + \eta'\big(u_k^l\big)(c_k - f)^2 \prod_{j<k} \eta\big(1-u_j^l\big) \\ &\qquad - \sum_{i>k} \eta\big(u_i^l\big)(c_i - f)^2 \eta'\big(1-u_k^l\big) \prod_{\substack{j<i \\ j\neq k}} \eta\big(1-u_j^l\big).\end{aligned}$$

At each time step, $u_k^{l+1}(x)$ and $v_k^{l+1}(x)$ are obtained from $u_k^l(x)$ and $v_k^l(x)$ by solving the two equations above. Their right hand sides depend only on u_k^l and v_k^l, while their left hand sides are constant coefficient linear elliptic operators on u_k^{l+1} and v_k^{l+1}. At each time step, we discretize these elliptic equations in space by means of a finite difference approximation; then we use the fast Fourier transform to solve them. The small parameter ε, which has to be of the same order of the space grid size [11], constitutes the main restriction on how large the time steps can be chosen.

4.3. Treatment of the Constants

In general, the constants c_k need to be updated at each time step according to formula (6). However, in many cases it might be reasonable to assume that distinct regions obtained via (pre)segmentation (as initial data) will remain distinct even after the segmentation with depth procedure. This would be an appropriate assumption, for example, when the initial guesses for the approximate intensities c_k are well separated from each other, and all the initial regions have large area. Then the *initial* approximate intensities $c_1, \ldots, c_{n+1}$ would also be the *final* approximate intensities. In that case, $c_1, \ldots, c_{n+1}$ need not be updated in computations, once they are obtained from the (pre)segmentation. For the numerical examples we present in this paper, we tried it both ways; it did not make much of a difference, confirming this discussion.

5. Sample Computations

In this section, we present computational results that use our approach on two synthetic images.

5.1. Two Regions and a Background

We start with an image, the leftmost in Fig. 1, that clearly has two distinct regions on top of a background. The image is simple enough that any reasonable segmentation algorithm would decompose it as:

- The bright vertical bar with approximate intensity 1.
- The broken fork, with approximate intensity 0.48.

They appear on a dark background of approximate intensity 0. For the purposes of this paper, we can assume

Figure 1. From left to right, these figures are the original image used in the first numerical example and the two regions obtained from it by the preliminary segmentation (without depth) procedure. The resolution of the images is 128×128.

that this presegmentation step has already been accomplished successfully. The second and third images in Fig. 1 show the initial regions which we assume have been obtained. Let us refer to these two regions as A and B, in that order.

The central task is to minimize (3). As explained in Section 4, we have to consider separately the two different ways of assigning A and B as initial guesses. The first option, which we denote AB, is:

- $u_1(x, t = 0) =$ Characteristic function of the bar, and $c_1 = 1$.
- $u_2(x, t = 0) =$ Characteristic function of the fork, and $c_2 = 0.48$.

whereas the second option, denoted BA, is to take:

- $u_1(x, t = 0) =$ Characteristic function of the fork, and $c_1 = 0.48$.
- $u_2(x, t = 0) =$ Characteristic function of the bar, and $c_2 = 1$.

In both cases, $c_3 = 0$.

The top row of Fig. 2 shows the two regions found as the minimizer under *Hypothesis AB*. Since in this case the vertical bar A is assumed to be in the foreground, the fidelity terms in the energy do not prevent the second region (which is initially the broken fork B) to spread underneath A. And in fact, as shown in the figures, the two teeth of the fork propagate under the bar, and form what is in some sense the most likely completion. The energy of the final configuration (which *seems* to be a local minimizer) turns out to be about 3.48.

The bottom row of Fig. 2 illustrates the minimizer found this time under *Hypothesis BA*. The energy of the final configuration in this case (which again *seems* to be a local minimizer) turns out to be around 4.23.

The result of the computation is summarized in the following table:

Order	Minimum energy
AB	3.48
BA	4.23

Based on these minimal values of the energy in the two possible scenarios, the first hypothesis wins. *The conclusion is that the scene consists of a bright vertical bar partially occluding a fork on a dark background.*

Notice how under hypothesis AB the four T-junctions formed by occlusion on the left side of the bar connected up with the four other T-junctions on the right side of the bar. This all happened automatically, without explicitly detecting the junctions and trying out different ways of connecting them.

5.2. Three Regions and a Background

Figures 3 and 4 illustrate the results of a numerical experiment with a scene that is composed of 3 regions of approximately constant intensity on a dark background.

Once again, we assume that a successful (pre)segmentation algorithm is available that decomposes the original image shown in the upper left hand position in Fig. 3 into the three regions (plus the background) also shown in the same figure. Let A denote the ellipse, B the half bottle shaped region, and C the crescent shape. Note that in general the presegmentation need not be of particularly high quality, since these regions will merely be used as initial guesses.

As there are now three regions in the presegmentation, it is possible to assign them in $3! = 6$ different ways as initial guess. As before, we minimized our approximation to the Nitzberg-Mumford-Shiota energy

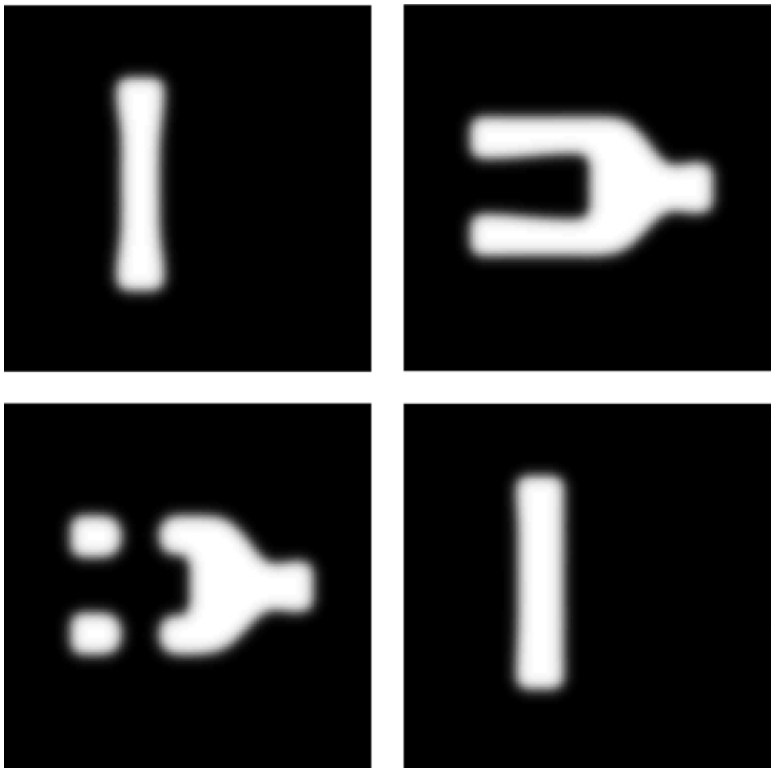

Figure 2. The top row of figures represents the minimizer obtained by our approach under order hypothesis *AB*, where the left figure is the front and the right figure is the back region found. The minimum energy reached in this case was: 3.48. The bottom row of images represents the minimizer obtained under the competing order hypothesis *BA*. The minimum energy in that case turned out to be: 4.23. The constants used in the computation were: $\alpha = 10^{-4}$, $\beta = 5 \times 10^{-7}$, and $\varepsilon = 0.01$. We took 500 time steps of size $\delta t = 16$ for each hypothesis.

starting with each. The results of the experiment are summarized in the table below:

Order	Minimum energy
ABC	4.21
ACB	4.21
BAC	6.08
BCA	6.53
CAB	4.71
CBA	6.53

The two orderings *ABC* and *ACB* come out on top with the least energy among all possible 6 orderings; the remaining four lead to significantly higher minimum energies. The appropriate conclusion seems to be:

The scene consists of a bright ellipse that partially occludes the bottle shaped object and a round object. The relative depths of the bottle shaped object and the round object cannot be determined.

It is hardly surprising that the algorithm fails to distinguish between *ABC* and *ACB*: the original image in Fig. 3 gives no clues about the relative depths of *B* (the half bottle) and *C* (the crescent).

6. Conclusion

In this paper we described a technique for minimizing the Nitzberg-Mumford-Shiota energy that avoids detecting and subsequently connecting T-junctions in the given image explicitly. One of the main drawbacks of our current implementation is that it deviates from the original model of Nitzberg, Mumford, and Shiota where the function $\phi(x)$ is supposed to have linear growth at infinity. This difference is important. Indeed, let us evaluate the boundary integral in (2) with our choice of $\phi(x) = x^2$ on a disk D_r of radius r:

$$\int_{\partial D_r} (\alpha + \beta k^2)\, ds = 2\pi\left(\alpha r + \frac{\beta}{r}\right) \xrightarrow{r \to 0^+} \infty, \quad (10)$$

as long as $\beta > 0$. Consequently, under an iterative minimization technique small disks would tend to grow,

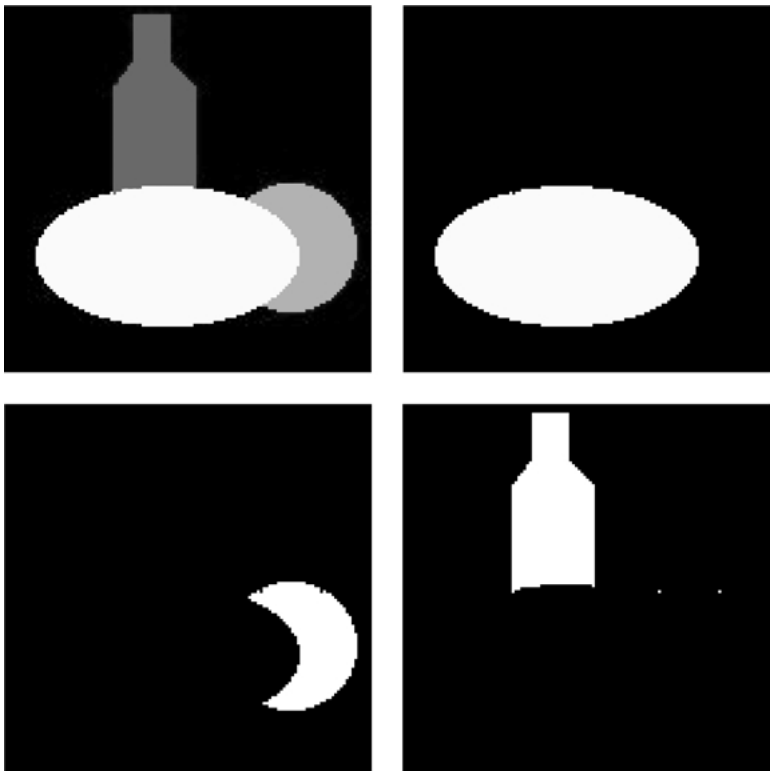

Figure 3. The upper left hand image shows the original image used in the sample calculation with three regions on a background, and the remaining three show the regions obtained from it by the preliminary segmentation (without depth) step. The resolution is 128×128.

Figure 4. The front, middle, and back figures, in that order, obtained as the minimizer in the computation with three regions on a background, corresponding to the order hypothesis *ACB*. The minimum energy in that case turned out to be: 4.21. The constants used in the computation were: $\alpha = 2 \times 10^{-4}$, $\beta = 4 \times 10^{-6}$, and $\varepsilon = 0.01$. We took 2000 time steps of size $\delta t = 2$ for each hypothesis (the remaining five of which are not shown).

raising a stability issue. Secondly, the number of regions in the segmentation cannot decrease in the course of the minimization, because this can happen only if at least one of the regions shrinks down to a point, which is prevented from (10). Note that to decrease energy (2) it might be advantageous to merge two regions whose grayscale intensities are approximately equal into a single region. Because of the said reasons, our current implementation does not allow such a possibility to be explored automatically in the course of a minimization. In order to overcome these drawbacks we are currently working on the problem of finding a numerically feasible approximation to (2) with $\phi(x)$ having linear growth at infinity.

Another major issue is having to carry out $n!$ minimizations when there are n regions in the presegmentation. This issue is shared by other approaches to the solution of the 2.1D sketch problem. Indeed, in Nitzberg, Mumford, and Shiota's implementation, as described in [16], the number of minimizations required is on the order of 2^{n^4}, where n is the number of *contours* detected in the original image (which can be a lot larger than the number of *regions*). An improvement might be achieved by carrying out the procedures described in this paper locally (over subsections of the image) where there might be only a few objects present so that $n!$ is not too large; the question is then how to gracefully combine results obtained from various subsections to reach a global conclusion.

Acknowledgments

The authors would like to thank G. Bellettini, F. Santosa and J. Shen for many useful discussions.

References

1. L. Ambrosio and V.M. Tortorelli, "Approximation of functionals depending on jumps by elliptic functionals via Γ-convergence," *Comm. Pure Appl. Math.*, Vol. 43, No. 8, pp. 999–1036, 1990.
2. G. Bellettini, "Variational approximation of functionals with curvatures and related properties," *J. Convex Anal.*, Vol. 4, No. 1, pp. 91–108, 1997.
3. G. Bellettini and M. Paolini, "Approssimazione variazionale di funzionali con curvatura. Seminario di Analisi Matematica," Dipartimento do Matematica dell' Universita' di Bologna, 1993.
4. T.F. Chan, S.H. Kang, and J. Shen, "Euler's elastica and curvature based inpaintings," UCLA Preprint, April 2001.
5. T. Chan and L. Vese, "A level set algorithm for minimizing the Mumford-Shah functional in image procesing," in *Proceedings of the 1st IEEE Workshop on "Variational and Level Set Methods in Computer Vision,"* 2001, pp. 161–168.
6. G. Dal Maso, *An Introduction to* Γ*-Convergence*, Birkhäuser, Boston, 1993.
7. E. De Giorgi, "Some remarks on Γ-convergence and least squares methods," in *Composite Media and Homogenization Theory*, G. Dal Maso and G.F. Dell'Antonio (Eds.), Birkhauser: Boston, 1991, pp. 135–142.
8. S. Esedoglu and J. Shen, "Image inpainting by the Mumford–Shah–Euler model," *European J. Appl. Math.*, Vol. 13, pp. 353–370, 2002.
9. P. Loreti and R. March, "Propagation of fronts in a nonlinear fourth order equation," *European J. Appl. Math.*, Vol. 11, pp. 203–213, 2000.
10. R. March, "Visual reconstruction with discontinuities using variational methods," *Image and Vision Computing*, Vol. 10, pp. 30–38, 1992.
11. R. March and M. Dozio, "A variational method for the recovery of smooth boundaries," *Image and Vision Computing*, Vol. 15, pp. 705–712, 1997.
12. L. Modica, "The gradient theory of phase transitions and the minimal interface criterion," *Arch. Rational Mech. Anal.*, Vol. 98, pp. 123–142, 1987.
13. L. Modica and S. Mortola, "Un esempio di Γ-convergenza," *Boll. Un. Mat. Ital.* B(5) Vol. 14, No. 1, pp. 285–299, 1977.
14. D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational problems," *Comm. Pure Appl. Math.*, Vol. 42, pp. 577–685, 1989.
15. M. Nitzberg and D. Mumford, "The 2.1-D sketch," in *Proceedings of the Third International Conference on Computer Vision.* Osaka, 1990.
16. M. Nitzberg, D. Mumford, and T. Shiota, "Filtering, segmentation, and depth," in *Lecture Notes in Computer Science*, Vol. 662, Springer Verlag: Berlin, 1993.
17. L. Vese and T. Chan, "A multiphase level set framework for image segmentation using the Mumford and Shah model," UCLA CAM Report 01-25, 2001. Also in *International Journal of Computer Vision*, Vol. 50, No. 3, pp. 271–293, 2002.

Selim Esedoglu received his BSc. in Mathematics from Brown University in 1996, and his Ph.D. in Mathematics from the Courant Institute of Mathematical Sciences, New York University, in 2000.

He is currently a CAM Assistant Professor at UCLA Department of Mathematics. His research interest is the mathematical problems of computer vision.

Riccardo March was born in Livorno, Italy, in 1956. He received the Laurea degree in Electronic Engineering cum laude from the University of Pisa in 1982. He joined the Italian National Research Council in 1984 where he is a senior researcher. His research interests are variational methods in computer vision, free discontinuity problems and the variational approximation of functionals. He is also interested in the application of the theory of Pade' approximation to superresolution in signal processing.