Segmentation with Depth: A Level Set Approach *

Wei Zhu †, Tony Chan ‡, and Selim Esedoḡlu §

Abstract

Segmentation with depth is the challenging problem of obtaining three dimensional information from a single two dimensional image. Unlike the standard segmentation problem, the goal of segmentation with depth is to determine not only the boundaries of objects that appear in the image, but also their relative distances to the observer by making use of occlusion relations. Nitzberg, Mumford, and Shiota proposed a variational formulation of this problem; according to their model, called the 2.1D Sketch model, the regions that the objects occupy and their relative depth are to be extracted from the 2D image by minimizing a complicated functional. Numerically, this is a highly non-trivial problem as the functional involves curvatures of the unknown contours. In this paper, wee develop a level set based procedure for minimizing the Nitzberg-Mumford-Shiota energy.

1 Introduction

We develop level set methods [23, 25] to numerically minimize the Nitzberg-Mumford-Shiota (NMS) functional that was proposed for the segmentation with depth problem [22, 21]. Segmentation with depth is different from standard segmentation [13, 19, 20, 17, 5, 6, 8, 9] which only depends on the gray-scale intensity distributions in image domains (see Figure 1). Segmentation with depth refers to the difficult task of extracting three dimensional information from a single two dimensional image that depicts a scene with several objects in it. Unlike the standard segmentation problem, segmentation with depth attempts to reconstruct both the regions occupied by different objects in the given two dimensional image and their relative distances (depth) to the observer. To accomplish this, the model of Nitzberg-Mumford-Shiota tries to

*This work has been supported by ONR contract N00014-03-1-0888, NSF contract DMS-9973341 and NIH contract P20 MH65166.

†Department of Mathematics, University of California, Los Angeles, 405 Hilgard Avenue, Los Angeles, CA, 90095. E-mail: wzhu@math.ucla.edu

‡Department of Mathematics, University of California, Los Angeles, 405 Hilgard Avenue, Los Angeles, CA, 90095. E-mail: TonyC@college.ucla.edu

§Department of Mathematics, University of California, Los Angeles, 405 Hilgard Avenue, Los Angeles, CA, 90095. E-mail: esedoglu@math.ucla.edu

decompose the given image into regions that are allowed to overlap, and includes terms that depend on the curvature of these regions' boundaries. The curvature dependence of the functional involved is one of the most difficult aspects of the model; specifically, the energy has terms of the form:

$$\int_{\partial R} \phi(\kappa) ds, \tag{1}$$

where κ denotes the curvature of the boundary of the region R and ϕ is a particular function.

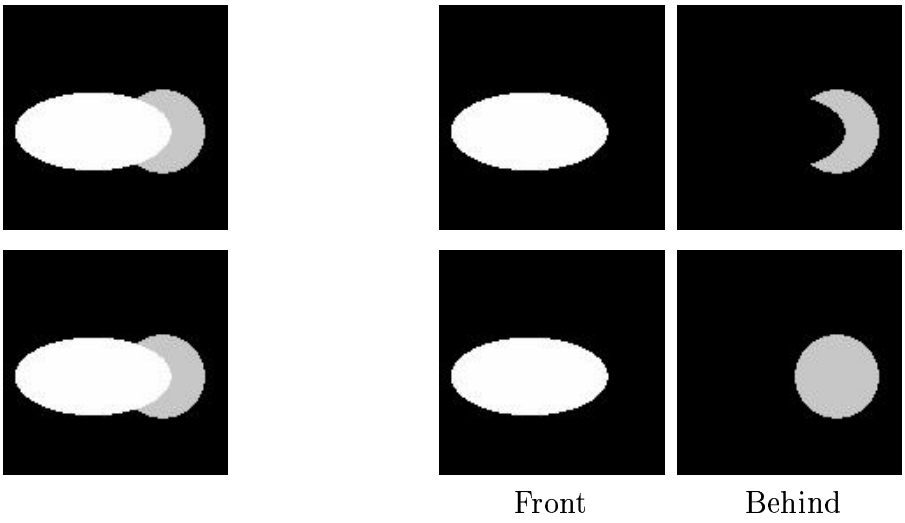


Figure 1: The first row contains the original image (Left) with two regions (middle and right) given by any multi-phase gray-scale intensity based segmentation method, and the second row illustrates the results from a segmentation with depth model. Besides the occlusion relation between the two objects ("Front" and "Behind" in this example), the missing boundary of the behind object is also restored.

In [22], the authors avoid minimizing their functional directly. Instead, they propose a combinatorial algorithm that first detects edges and T-junctions in the given 2D image and then considers all possible ways of connecting the T-junctions by new edges that is consistent with a given ordering hypothesis. The ordering hypothesis with respect to which the detected T-junctions can be connected using the smallest amount of energy is found to be the correct depth relation between the (possibly overlapping) regions. This is thus a combinatorial technique with very high complexity. Moreover, it requires explicitly detecting T-junctions in the image as a preprocessing step, which can be a significant source of errors (in the form of false positives and misses). On the other hand, the model itself makes no explicit reference to T-junctions; therefore it should be possible to minimize the energy without detecting such features first.

In [10], Esedo$\bar{g}$lu and March propose a new, direct way of minimizing the Nitzberg-Mumford-Shiota energy that is based on curve evolution and gradient descent. It

works in much the same way that the Chan-Vese algorithm ([8]) minimzes the Mumford-Shah segmentation funcitonal ([19, 20]). Their technique makes essential use of $\Gamma-$ convergence ideas ([14]) and a conjecture of E. De Giorgi ([12]) to approximate the NMS functional by elliptic ones which are numerically much more convenient (Many applications of the notion of $\Gamma-$ convergence appear in various contexts ([1, 3, 15, 16]). This approach allows them to avoid detecting and connecting T-junctions explicitly; these happen automatically as a consequence of the curve evolution (and involve changes in the topology of the curves). However, the approximation technique used by the authors force them to consider only quadratic dependence on curvature (i.e. $\phi(x) = x^2$), which is an approximation to but is different from the correct choice prescribed by NMS in their original formulation. The difference is significant as explained by the authors in [10].

In this paper, we give a level set based formulation ([23, 25, 24]) to minimize the NMS functional that is also based on curve evolution and gradient descent. However, our techniques allow us to use the more general curvature dependence directly, without using approximations to them such as in ([1, 2, 3, 4]). This way, we get an implementation that remains faithful to the model and avoids the issues caused by using quadratic dependence as in [10].

The level set method of Osher and Sethian is particularly well-suited to handling curvature dependent functionals, since the curvature of a contour has a very simple expression in terms of the level set function that represents it. For example, let ϕ be a level set function, then the curvature κ along any level curve $\phi = c$ (c is any real value) can be represented as:

$$\kappa \;=\; \nabla \cdot (\frac{\nabla \phi}{|\nabla \phi|}). \tag{2}$$

With the representation, it is also trivial to handle the absolute value of curvature that appears in the above functional. By using level set methods, the problem of minimizing the Nitzberg-Mumford-Shiota functional turns out to be solving a system of fourth order partial differential equations. This kind of equations were also solved by the first two authors in [28] for capturing illusory contours. Moreover, we employ a semi-implicit discretization scheme due to Smereka [26] to solve the resulting fourth order equations. This method can speed up computations significantly, since explicit schemes for these equations are notoriously slow to converge to a stationary state.

The rest of this paper is organized as follows. In the next section, we introduce the background of segmentation with depth as well as the Nitzberg-Mumford-Shiota functional. In section 3, we rewrite the functional in a level set formulation and derive the associated Euler-Lagrange equations. We then discuss numerical solution techniques including Smereka's semi-implicit scheme in the subsequent section 4. In section 5, we present some numerical experiments on synthetic images that appears in [10]. Section 6 is devoted to our conclusions.

2 Nitzberg-Mumford-Shiota functional

In this section, we sketch the problem of segmentation with depth and present the Nitzberg-Mumford-Shiota functional. Further details can be found in [22, 10].

The most important unknows of the Nitzberg-Mumford-Shiota model are the shapes of the objects that make up the scene, and their relative distances to the observer (i.e. their depth). The relative distances of the objects can be described by an ordering between them. The shapes of the objects are described by subsets of the image domain. Since the model allows for occlusions, these regions need not be disjoint: Parts of regions in the back (further from the observer) might be blocked from view by parts of regions in front (closer to the observer). The model thus inevitably involves reconstructing invisible parts of the regions' boundaries. This calls for the notion of "most likely completion" of a boundary contour. In the NMS model, the most likely completion of a curve is expressed as the minimizer of a curvature dependent functional.

The 2.1D model of NMS makes the following simplifying assumptions about the three dimensional scene:

- Objects in the scene are not entangled,
- There are no self-occlusions,
- Grayscale intensities of different objects are approximately uniform and different from each other.

The first two assumptions can be summarized as follows: the 2.1D sketch model of NMS assumes that the objects making up the three dimensional scene live in distinct planes which are parallel to each other and which are perpendicular to the line of sight of the observer. The unknowns of the problem are:

- The ordering relation between the regions (which is in front of which),
- The (un-occluded) shapes of the regions,
- The approximate grayscale intensity of the corresponding objects.

Let $f : \Omega \to R^1$ be a 2-D image defined on $\Omega \subset R^2$, and $R_1, ..., R_n$ be the regions occupied by objects inside the image. Denote the visible parts of the above objects by $R_1^{'}, ..., R_n^{'}$ respectively.

Suppose the regions $R_1, ..., R_n$ are listed in order of increasing distance to the observer (i.e. depth), so that R_1, R_n are the nearest and farthest regions respectively. Then the visible part $R_i^{'}$ of R_i is given by:

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

For simplicity, we denote by $R'_{n+1} = \Omega - \bigcup_{j<n+1} R_j$ as the background. Then the Nitzberg-Mumford-Shiota functional reads:

$$E = \sum_{i=1}^{n} \int_{\partial R_i \bigcap \Omega} [\alpha + \beta\phi(\kappa)]ds + \sum_{i=1}^{n+1} \int_{R'_i} (f(x) - c_i)^2 dx, \tag{3}$$

where α, β are two nonnegative parameters, and the unknowns $c_i \in R^1$ denote the approximate gray-scale intensities (color) of of the corresponding objects. The function ϕ determines how the curvature information will be incorporated in the functional. The authors choose ϕ as follows:

$$\phi(x) = \begin{cases} x^2, & |x| \leq 1 \\ |x|, & |x| > 1 \end{cases} \tag{4}$$

This choice of the function ϕ, as opposed to the choice made in [10], allows corners in boundaries of objects.

An important point is that the contour integrals in energy (3) are taken along the entire length of the boundaries of the regions, not just the visible parts of these boundaries. This term therefore dictates how the occluded parts of boundary curves are best reconstructed. The area integrals in the energy (3) make up the "fidelity" terms in the model. They ensure that the reconstructed configuration of regions and their ordering in three dimensional space yield a two dimensional image that agrees with the observed one, namely $f(x)$.

3 Level set method for minimizing the Nitzberg-Mumford-Shiota functional

As noted in the introduction, the level set method of Osher and Sethian is very well suited for the representation of functionals that involve geometric quantities such as curvature. In this section, we first rewrite the NMS functional in terms of level set functions that represent the unknown regions R_i. Subsequently, we derive the associated Euler-Lagrange equations in terms of the level set functions.

3.1 Level set formulation

The regions R_i are among the unknowns of the optimization problem (3) and therefore need to be updated during our iterative solution procedure. To represent them, we introduce the level set functions ϕ_i so that $R_i = \{\psi_i\}$, $i = 1, ..., n$.

Then, the first sum of (3) can be rewritten as follows:

$$\sum_{i=1}^{n} \int_{\Omega} [\alpha + \beta\phi(\kappa_i)]|\nabla\psi_i|\delta(\psi_i)dx, \tag{5}$$

where $\kappa_i = \nabla \cdot (\frac{\nabla \psi_i}{|\nabla \psi_i|})$ is the curvature of any level curve, and $\delta(x)$ is the derivative of the Heaviside function:

$$H(x) = \begin{cases} 1, & \text{x} > 0, \\ 0, & \text{x} \leq 0. \end{cases} \tag{6}$$

in the distribution sense, or the Dirac function.

On the other hand, since $R_i^{'} = R_i - \bigcup_{j<i} R_j$, and $R_i^{'} = \{\psi_i > 0\} \bigcap (\bigcup_{j<i} \{\psi_j \leq 0\})$, we may rewrite the second sum of (3) as follows:

$$\sum_{i=1}^{n} \Big\{ \int_{\Omega} (f - c_i)^2 H(\psi_i) \prod_{j=1}^{i-1} (1 - H(\psi_j)) dx \Big\} + \int_{\Omega} (f - c_{n+1})^2 \prod_{j=1}^{n} (1 - H(\psi_j)) dx, \tag{7}$$

where H is the Heaviside function:

Combining (5) and (7), we obtain the Nitzberg-Mumford-Shiota functional in a level set based formulation as follows:

$$\begin{aligned} E &= \sum_{i=1}^{n} \int_{\Omega} [\alpha + \beta \phi(\kappa_i)] |\nabla \psi_i| \delta(\psi_i) dx \\ &\quad + \sum_{i=1}^{n} \Big\{ \int_{\Omega} (f - c_i)^2 H(\psi_i) \prod_{j=1}^{i-1} (1 - H(\psi_j)) dx \Big\} \\ &\quad + \int_{\Omega} (f - c_{n+1})^2 \prod_{j=1}^{n} (1 - H(\psi_j)) dx, \end{aligned} \tag{8}$$

3.2 Derivation of the Euler-Lagrange equations

To derive the associated Euler-Lagrange equations of the functional (8), we here only consider the curvature related terms, since the fidelity terms are easy to handle. For simplicity, let's consider a functional $\widehat{E}(\psi) = \int_{\Omega} \Psi(\kappa_\psi) |\nabla \psi| \delta(\psi) dx$, where $\Psi(x) = \alpha + \beta \phi(x)$ in the functional (8). In fact, the derivation of the Euler-Lagrange equation for a similar functional can be found in [7, 11]. The functional that we consider here is an integral along a contour instead of on the whole domain Ω.

Before deriving the Euler-Lagrange equation of ψ according to the minimization of $\widehat{E}(\psi)$, let's notice two simple facts.

Lemma 1. *Let $\overrightarrow{n} \in R^2$ be a unit vector, and $I, P_{\overrightarrow{n}} : R^2 \to R^2$ be two operators with $I(\overrightarrow{v}) = \overrightarrow{v}$, $P_{\overrightarrow{n}}(\overrightarrow{v}) = (\overrightarrow{v} \cdot \overrightarrow{n}) \overrightarrow{n}$, for any $\overrightarrow{v} \in R^2$. Then for the curvature $\kappa = \nabla \cdot (\frac{\nabla \psi}{|\nabla \psi|})$, we have $\frac{\partial \kappa}{\partial t} = \nabla \cdot ((I - P_{\overrightarrow{n}}) \frac{\nabla \psi_t}{|\nabla \psi|})$, where $\overrightarrow{n} = \frac{\nabla \psi}{|\nabla \psi|}$.*

Proof. Simple calculation shows the conclusion. □

Lemma 2. *With the same notation as in lemma 1, we have $(I - P_{\overrightarrow{n}}) \overrightarrow{v} \cdot \overrightarrow{w} = (I - P_{\overrightarrow{n}}) \overrightarrow{w} \cdot \overrightarrow{v}$, for any $\overrightarrow{v}, \overrightarrow{w} \in R^2$.*

We may thus obtain the following theorem.

Theorem 1. *Let* $\widehat{E}(\psi) = \int_\Omega \Psi(\kappa)|\nabla\psi|\delta(\psi)dx$, *and* $\Psi : R^1 \to R^1$ *is continuous and piecewise differentiable with* $\Psi'(0) = 0$, *then the associated Euler-Lagrange equation reads: (*$P = P_{\frac{\nabla\psi}{|\nabla\psi|}}$*.)*

$$\frac{\partial\psi}{\partial t} = |\nabla\psi|\nabla\cdot[\frac{\nabla\psi}{|\nabla\psi|}\Psi(\kappa) - \frac{1}{|\nabla\psi|}(I-P)(\nabla(\Psi'(\kappa)|\nabla\psi|))], \tag{9}$$

with the boundary condition: $\frac{\partial\psi}{\partial\overrightarrow{\nu}}|_{\partial\Omega} = 0$, $\kappa_\psi|_{\partial\Omega} = 0$, *where* $\overrightarrow{\nu}$ *is the unit vector outward normal to* $\partial\Omega$.

Proof. To get the Euler-Lagrange equation, let's take the derivative of $\widehat{E}$ according to time t, then determine the derivative of ψ to t by decreasing the functional $\widehat{E}$.

In fact,

$$\begin{aligned}
\frac{d\widehat{E}}{dt} &= \int_\Omega \Psi'(\kappa)\kappa_t|\nabla\psi|\delta(\psi) + \Psi(\kappa)|\nabla(\psi)|_t\delta(\psi) + \Psi(\kappa)|\nabla(\psi)|\delta'(\psi)\psi_t \\
&= \int_\Omega \Psi'(\kappa)\nabla\cdot[(I-P)(\frac{\nabla\psi_t}{|\nabla\psi|})]|\nabla\psi|\delta(\psi) + \Psi(\kappa)\nabla\psi_t\cdot\frac{\nabla\psi}{|\nabla\psi|}\delta(\psi) \\
&\quad + \int_\Omega \Psi(\kappa)|\nabla(\psi)|\delta'(\psi)\psi_t \\
&= \int_\Omega \nabla\cdot[(I-P)(\frac{\nabla\psi_t}{|\nabla\psi|})\Psi'(\kappa)|\nabla\psi|\delta(\psi)] + \int_\Omega \nabla\cdot[\psi_t\frac{\nabla\psi}{|\nabla\psi|}\Psi(\kappa)\delta(\psi)] \\
&\quad - \int_\Omega (I-P)(\frac{\nabla\psi_t}{|\nabla\psi|})\cdot\nabla(\Psi'(\kappa)|\nabla\psi|\delta(\psi)) - \psi_t\nabla\cdot[\frac{\nabla\psi}{|\nabla\psi|}\Psi(\kappa)\delta(\psi)] \\
&\quad + \int_\Omega \Psi(\kappa)|\nabla(\psi)|\delta'(\psi)\psi_t,
\end{aligned} \tag{10}$$

by the Neunman boundary condition, i.e. $\frac{\partial\psi}{\partial\overrightarrow{\nu}} = \nabla\psi\cdot\overrightarrow{\nu} = 0$, we see that $\nabla\psi_t\cdot\overrightarrow{\nu} = (\nabla\psi\cdot\overrightarrow{\nu})_t - \nabla\psi\cdot\overrightarrow{\nu}_t = 0$, therefore by applying the Green-Gauss Theorem, we may drop the two divergence terms, and get

$$\begin{aligned}
\frac{d\widehat{E}}{dt} &= \int_\Omega -(I-P)(\frac{\nabla\psi_t}{|\nabla\psi|})\cdot\nabla(\Psi'(\kappa)|\nabla\psi|\delta(\psi)) \\
&\quad - \int_\Omega \psi_t\nabla\cdot[\frac{\nabla\psi}{|\nabla\psi|}\Psi(\kappa)\delta(\psi)] + \Psi(\kappa)|\nabla(\psi)|\delta'(\psi)\psi_t,
\end{aligned} \tag{11}$$

by lemma 2, $(I-P)\overrightarrow{v}\cdot\overrightarrow{w} = \overrightarrow{v}\cdot(I-P)\overrightarrow{w}$, we have

$$\begin{aligned}
\frac{d\widehat{E}}{dt} &= \int_\Omega -\frac{\nabla\psi_t}{|\nabla\psi|}\cdot(I-P)[\nabla(\Psi'(\kappa)|\nabla\psi|\delta(\psi)] \\
&\quad -\psi_t\nabla\cdot[\frac{\nabla\psi}{|\nabla\psi|}\Psi(\kappa)\delta(\psi)] + \int_\Omega \Psi(\kappa)|\nabla(\psi)|\delta'(\psi)\psi_t
\end{aligned}$$

$$
\begin{aligned}
&= \int_\Omega -\nabla\cdot[\psi_t\frac{1}{|\nabla\psi|}(I-P)(\nabla(\Psi^{'}(\kappa)|\nabla\psi|\delta(\psi))] \\
&\quad +\psi_t\nabla\cdot[\frac{1}{|\nabla\psi|}(I-P)(\nabla(\Psi^{'}(\kappa)|\nabla\psi|\delta(\psi))] \\
&\quad -\int_\Omega \psi_t[\nabla\cdot(\frac{\nabla\psi}{|\nabla\psi|}\Psi(\kappa))\delta(\psi)+\frac{\nabla\psi}{|\nabla\psi|}\Psi(\kappa)\cdot\delta^{'}(\psi)\nabla(\psi)] \\
&\quad +\int_\Omega \Psi(\kappa)|\nabla(\psi)|\delta^{'}(\psi)\psi_t. \qquad (12)
\end{aligned}
$$

Notice that $\nabla(\Psi^{'}(\kappa)|\nabla\psi|\delta(\psi)) = \nabla(\Psi^{'}(\kappa)|\nabla\psi|)\delta(\psi)+\Psi^{'}(\kappa)|\nabla\psi|\delta^{'}(\psi)\nabla\psi$, then $(I-P)\nabla(\Psi^{'}(\kappa)|\nabla\psi|\delta(\psi)) = \delta(\psi)(I-P)\nabla(\Psi^{'}(\kappa)|\nabla\psi|)$, since $(I-P)\nabla\psi = 0$. Therefore, by applying Green-Gauss Theorem again and note that $\Psi^{'}(\kappa)|_{\partial\Omega} = 0$, we drop the divergence part in (12) and obtain

$$
\begin{aligned}
\frac{d\widehat{E}}{dt} &= \int_\Omega \psi_t\nabla\cdot[\frac{1}{|\nabla\psi|}(I-P)(\nabla(\Psi^{'}(\kappa)|\nabla\psi|)\delta(\psi)] - \int_\Omega \psi_t\nabla\cdot[\frac{\nabla\psi}{|\nabla\psi|}\Psi(\kappa)]\delta(\psi) \\
&= \int_\Omega \psi_t\Big\{\nabla\cdot[\frac{1}{|\nabla\psi|}(I-P)(\nabla(\Psi^{'}(\kappa)|\nabla\psi|)]\delta(\psi) \\
&\quad +\frac{1}{|\nabla\psi|}(I-P)(\nabla(\Psi^{'}(\kappa)|\nabla\psi|)\cdot\delta^{'}(\psi)\nabla\psi\Big\} - \int_\Omega \psi_t\nabla\cdot[\frac{\nabla\psi}{|\nabla\psi|}\Psi(\kappa)]\delta(\psi) \\
&= \int_\Omega \psi_t\delta(\psi)\nabla\cdot[\frac{1}{|\nabla\psi|}(I-P)(\nabla(\Psi^{'}(\kappa)|\nabla\psi|) - \frac{\nabla\psi}{|\nabla\psi|}\Psi(\kappa)] \qquad (13)
\end{aligned}
$$

Therefore, if we choose $\psi_t = -\delta(\psi)\nabla\cdot[\frac{1}{|\nabla\psi|}(I-P)(\nabla(\Psi^{'}(\kappa)|\nabla\psi|)) - \frac{\nabla\psi}{|\nabla\psi|}\Psi(\kappa)]$, we see $\frac{d\widehat{E}}{dt} \le 0$. Moreover, by employing the standard technique which replaces $\delta(\psi)$ by $|\nabla\psi|$ in [27] (the factor $|\nabla\psi|$ also helps to accelerate the evolution process as discussed in [18]), we then have the associated Euler-Lagrange equation:

$$
\frac{\partial\psi}{\partial t} = |\nabla\psi|\nabla\cdot[\frac{\nabla\psi}{|\nabla\psi|}\Psi(\kappa) - \frac{1}{|\nabla\psi|}(I-P)(\nabla(\Psi^{'}(\kappa)|\nabla\psi|))]. \qquad (14)
$$

□

Let's turn to the original Nitzberg-Mumford-Shiota functional (8). From Theorem 1, the system of Euler-Lagrange equations reads:

$$
\begin{aligned}
\frac{\partial\psi_i}{\partial t} &= |\nabla\psi_i|\nabla\cdot[\frac{\nabla\psi_i}{|\nabla\psi_i|}\Psi(\kappa_i) - \frac{1}{|\nabla\psi_i|}(I-P_{\frac{\nabla\psi_i}{|\nabla\psi_i|}})(\nabla(\Psi^{'}(\kappa_i)|\nabla\psi_i|))] \\
&\quad -|\nabla\psi_i|(f-c_i)^2\prod_{j=1}^{i-1}(1-H(\psi_j)) \\
&\quad +|\nabla\psi_i|\sum_{s=i+1}^{n+1}\Big\{(f-c_s)^2H(\psi_s)\prod_{j=1}^{i-1}(1-H(\psi_j))\prod_{j=i+1}^{s-1}(1-H(\psi_j))\Big\}, \qquad (15)
\end{aligned}
$$

with the boundary conditions: $\frac{\partial\psi_i}{\partial\vec{\nu}}|_{\partial\Omega} = 0$, $\kappa_i|_{\partial\Omega} = 0$, $i = 1, ..., n$, where $H(\psi_{n+1}) = 1$.

4 Numerical Implementation

In this section, we show how to discretize each term in the equations (15) and explain how to apply Smereka's Semi-Implicit method to solve the equations.

4.1 Discretization of the Euler-Lagrange equations

We first discuss the schemes for the fourth order terms in (15). The discretization is essentially the same as the one in [7]. We include its details for the sake of completeness of the present discussion.

As in [7], we rewrite them as follow:

$$\frac{\partial \phi}{\partial t} = |\nabla \phi| \nabla \cdot \vec{V}, \tag{16}$$

where $\vec{V} = (V_1, V_2) = \frac{\nabla \phi}{|\nabla \phi|}\Psi(\kappa) - \frac{1}{|\nabla \phi|}(I - P)(\nabla(\Psi'(\kappa)|\nabla \phi|))$.

Then, denote $\Psi'(\kappa)|\nabla \psi|$ simply by A, we get:

$$V_1 = \Psi(\kappa)\frac{\psi_x}{|\nabla \psi|} - \frac{\psi_y}{|\nabla \psi|^3}[A_x \psi_y - A_y \psi_x], \tag{17}$$

and

$$V_2 = \Psi(\kappa)\frac{\psi_y}{|\nabla \psi|} - \frac{\psi_x}{|\nabla \psi|^3}[A_y \psi_x - A_x \psi_y], \tag{18}$$

where $A_x = \partial A/\partial x, A_y = \partial A/\partial y$.

Let (i, j) denote the pixel locations, h be the spatial step size and the time be discretized to be $\{ndt\}$, where $n = 0, 1, 2, ...$, and dt is the time step. Then $\psi^n_{i,j}$ refers to the value of ψ at pixel (i, j) at time ndt.

Let's discuss the disretization of each term in (16$\sim$ 18). In (16), we approximate $|\nabla \psi|$ by central difference as:

$$|\nabla \psi|_{i,j} = \sqrt{[\frac{\psi_{(i+1,j)} - \psi_{(i-1,j)}}{2h}]^2 + [\frac{\psi_{(i,j+1)} - \psi_{(i,j-1)}}{2h}]^2}, \tag{19}$$

As for the terms in (17, 18), since

$$\begin{aligned} \nabla \cdot \vec{V}_{(i,j)} &= V_{1x(i,j)} + V_{2y(i,j)} \\ &= \frac{V_{1(i+1/2,j)} - V_{1(i-1/2,j)}}{h} + \frac{V_{2(i,j+1/2)} - V_{2(i,j-1/2)}}{h}, \end{aligned} \tag{20}$$

we need to calculate the values of V_1 at the x-half-pixel $(i + 1/2, j)$, while the values of V_2 at the y-half-pixel $(i, j + 1/2)$.

We approximate the terms in (17) at $(i+1/2, j)$ as follows:

$$\begin{aligned} \kappa &= \text{minmod}(\kappa_{(i,j)}, \kappa_{(i,j)}), \\ \psi_x &= \frac{\psi_{(i+1,j)} - \psi_{(i,j)}}{h}, \\ \psi_y &= \text{minmod}(\frac{\psi_{(i+1,j+1)-\psi_{(i+1,j-1)}}}{2h}, \frac{\psi_{(i,j+1)} - \psi_{i,j-1}}{2h}), \\ |\nabla\psi| &= \sqrt{\psi_x^2 + \psi_y^2}, \end{aligned} \tag{21}$$

$$\begin{aligned} A_x &= \frac{A_{(i+1,j)} - A_{(i,j)}}{h}, \\ &= \frac{\Psi^{'}(\kappa_{(i+1,j)})|\nabla\psi|_{(i+1,j)} - \Psi^{'}(\kappa_{(i,j)})|\nabla\psi|_{(i,j)}}{h}, \\ A_y &= \text{minmod}(\frac{\Psi^{'}(\kappa_{(i+1,j+1)})|\nabla\psi|_{(i+1,j+1)} - \Psi^{'}(\kappa_{(i+1,j-1)})|\nabla\psi|_{(i+1,j-1)}}{2h}, \\ &\quad \frac{\Psi^{'}(\kappa_{(i,j+1)})|\nabla\psi|_{(i,j+1)} - \Psi^{'}(\kappa_{(i,j+1)})|\nabla\psi|_{(i,j-1)}}{2h}), \end{aligned} \tag{22}$$

where

$$\text{minmod}(x, y) = \frac{\text{sgn}(x) + \text{sgn}(y)}{2}\min(|x|, |y|).$$

Similarly, we may approximate the terms in (18).

It is quite subtle to discretize the remain terms in (15). Specifically, it should be careful to approximate $|\nabla\psi|$. In this paper, we use the technique that discussed in [25] (see Chapter 6). The detail can be found in our previous work [28]. We omit them here.

Moreover, we approximate the Heaviside function $H(x)$ by a standard regularized function as in [8], which reads:

$$H_\epsilon(x) = \frac{1}{2}(1 + \frac{2}{\pi}\arctan(\frac{x}{\epsilon})),$$

where ϵ is a fixed small number. In the experiments, we choose ϵ as small as h^2.

4.2 Smereka's Semi-Implicit method

Since equations (15) are fourth order and parabolic, we would expect the CLF condition for an explicit scheme to require the time step size to be proportional to the fourth power of spatial step size, i.e., $dt \sim dx^4$. This restriction would result in an intolerably long time for the minimization process. To ease the time step restriction, we will apply a semi-implicit discretization method due to Smereka ([26]), which significantly accelerated the convergence of the gradient descent.

For simplicity, as in [26], let's denote the Euler-Lagrange equations as follows:

$$\psi_t = S(\psi), \tag{24}$$

then, we can rewrite it as:

$$\psi_t = -\lambda\Delta^2\psi + \lambda\Delta^2\psi + S(\psi), \tag{25}$$

where Δ^2 is a bi-Laplacian (since equations (15) are fourth order equations) and λ is a positive parameter. Then we may discretize the equation (25) in time as:

$$\psi^{n+1} - \psi^n = dt(-\lambda\Delta^2\psi^{n+1} + \lambda\Delta^2\psi^n + S(\psi^n)),$$

and then

$$\psi^{n+1} = \psi^n + dt(1 + dt\lambda\Delta^2)^{-1}S(\psi^n). \tag{27}$$

The operator $(1 + dt\lambda\Delta^2)$ is positive definite, and it can be efficiently inverted by the fast Fourier transform (FFT).

Therefore, in our case, we only need to calculate the values on the right side of (15) at time nh as $S(\psi^n)$ for each function ψ_i, then we may get the values of ψ_i at the new time $(n+1)h$ by using FFT.

Smereka's semi-implicit method will relax the CFL condition considerably, since the operator Δ^2 helps to smooth possible high oscillating waves. Numerical experiments in the following section also show that we may choose a much larger time step size than what is required by CFL condition.

4.3 Procedure of minimizing NMS functional

During the evolution of the variables ψ_i according to equation (15), the unknown intensities c_i also need to be updated. Variations of the objective energy E with respect to the c_i yield the following optimality conditions:

$$c_i = \frac{\int_\Omega f H(\psi_i)\prod_{j=1}^{i-1}(1 - H(\psi_j))}{\int_\Omega H(\psi_i)\prod_{j=1}^{i-1}(1 - H(\psi_j))}, \quad i = 1, ..., n. \tag{28}$$

and

$$c_{n+1} = \frac{\int_\Omega f \prod_{j=1}^{n}(1 - H(\psi_j))}{\int_\Omega \prod_{j=1}^{n}(1 - H(\psi_j))}. \tag{29}$$

Then, given the initial guesses for the level set functions ψ_i and the intensity constants c_i, we iteratively solve the equations (15) and (28), (29).

Now, let's discuss how to choose the initial guesses. We first employ any multiphase segmentation method, for instance Chan-Vese's method [9], to segment the image domain Ω into several regions in each of which the intensity is homogenous, and thus obtain the associated level set function and intensity constant for each region. Then, as in [10], we use these results as the initial guesses for solving (15) and (28), (29).

Moreover, the re-initializations of ψ_i are necessary during the above evolutions. Otherwise, the terms derived from fidelity ones in (15) will make ψ_i keep increasing or decreasing to infinity. On the other hand, the process of re-initializations can not be taken more frequently since it will prevent the zero level sets from growing.

5 Numerical Experiments

In this section, we present numerical results of applying our technique on some synthetic images, including the typical example (Bar and Fork) that appears in [10].

5.1 Two regions and a background

We here show two examples: Bar and Fork, Ellipse and Disk. For each example, the intensities for the two regions and the background are 1.0, 0.5, and 0 respectively.

The first example is an image with a bar and a broken fork inside a background (Figure 2). For this case, there are only two possible orderings: the bar is in front of the fork, or vice versa. We minimize the Nitzberg-Mumford-Shiota functional under both of these order hypothesis separately. The way that we build the hypothesis into the computation is by the assignment of initial conditions for the regions, i.e., in one of the computations, we take the bar to be the initial guess for the region that is in front and the broken fork to be the initial guess for the region that is behind, and we switch them for the other case. The values are listed as follow:

- Bar front, Fork behind, the NMS functional = 0.1762.
- Fork front, Bar behind, the NMS functional = 0.2718.

This shows that the assumption that the bar is in front of the fork yields a smaller value of the NMS functional, the scene is made up of a fork that is behind and occluded by a vertical bar.

In Figure 3, we list the evolution process of the two regions under the first assumption of ordering. From these groups of figures, one can see how the broken fork is connecting while the bar is keeping stable; note especially the behavior of its corners. It is the choice of $\phi(x) = |x|$ when $|x|$ big in (3) that preserves the corners. The choice $\phi(x) = x^2$ would inevitably smear those corners.

We omit the evolution process under the second assumption since the two regions are keeping stable.

Moreover, in the experiments, due to the use of Smereka's Semi-Implicit method, we choose the time step size $dt = 5 \times 10^{-4}$, which is far larger than $dt = 10^{-5}$ which is required by the strict CFL condition.

The second example is an image with an ellipse and a crescent on top of a dark background (Figure 4). We calculate the values of the Nitzberg-Mumford-Shiota functional for the two ordering as follow:

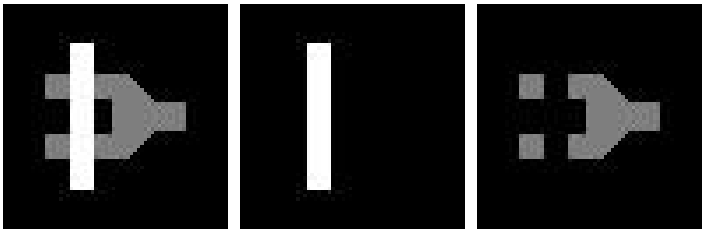


Figure 2: Bar and Fork. From left to right, these figures are the original image and the two regions obtained from standard segmentation model.

- Ellipse front, Crescent behind, the NMS functional = 0.1404.
- Crescent front, Ellipse behind, the NMS functional = 0.2812.

This shows that the ellipse is in front of the disk (instead of crescent!). In Figure 5, we list the evolution process of the two regions under the first assumption of ordering and we still omit the process under the second assumption. From these groups of figures, one can see how the crescent is growing to be a moon while the front ellipse is keeping stable.

5.2 Three regions and a background

We present here an example with an annulus and two bars inside a background (Figure 6). The intensities for them are 1.0, 0.9, 0.8, and 0. respectively.

Since there are three objects inside the image, $3! = 6$ cases of the ordering should be considered. Here we list the evolution process under the ordering assumption that the Annulus is in front of the vertical Bar while the horizontal Bar is behind the vertical Bar, which yields the smallest value of the Nitzberg-Mumford-Shiota's functional (Figure 7).

6 Conclusion

In this paper, we develop a level set based technique to numerically minimize the Nitzberg-Mumford-Shiota functional that appears in the variational formulation of the segmentation with depth problem. Our technique allows us to directly handle general curvature related terms. The numerical experiments listed in this paper demonstrate that the technique yields desirable results.

References

[1] G. Bellettini, *Variational approximation of functionals with curvatures and related properties,* J. Convex Anal., Vol. 4(1), pp. 91-108, 1997.

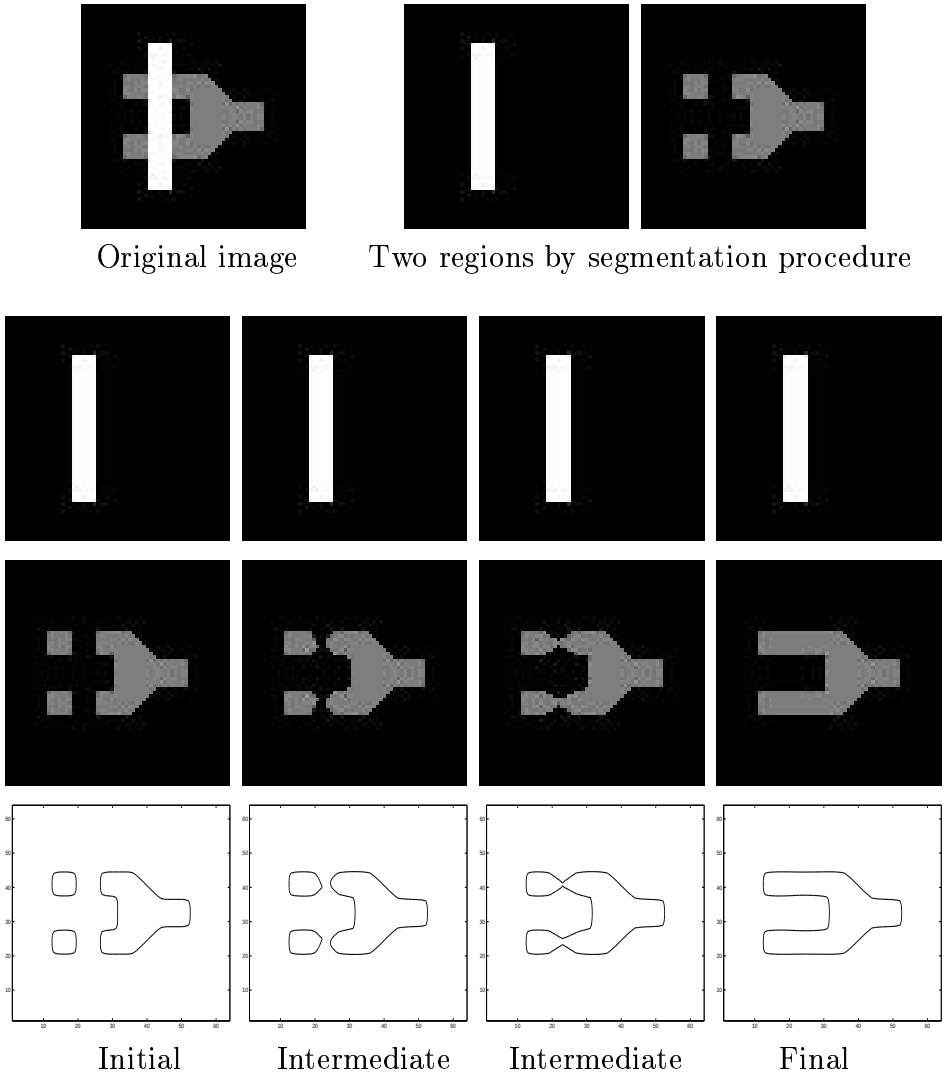


Figure 3: The evolution process under the assumption that the Bar is in front of the Fork. In the second and third row, the evolutions of the two regions by solving (15) are listed. The fourth row shows how the boundary of the Fork is growing. During the process, the front region occupied by the Bar is keeping stable while the region by the Fork is connecting behind the Bar. In this experiment, the parameters are: $\alpha = 0.00005$, $\beta = 0.005$, and $\lambda = 0.1$ in Smereka's semi-implicit method.

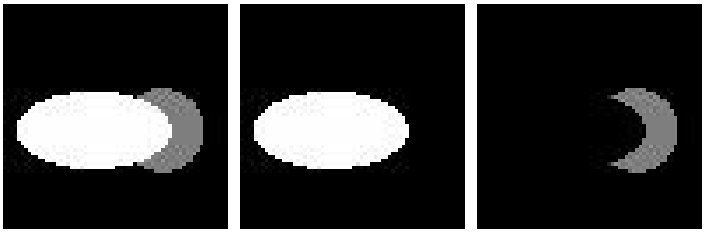


Figure 4: Ellipse and Crescent. From left to right, these figures are the original image, and the two regions obtained from standard segmentation model.

[2] G. Bellettini and R. March, *Variational properties of a model for image segmentation with overlapping regions,* Variational Methods for Discontinuous Structures, G. Dal Maso and F. Tomarelli editors. Progress in Nonlinear Differential Equations and Their Applications. 51, pp. 9-17, 2002.

[3] G. Bellettini and R. March, *An image segmentation variational model with free discontinuities and contour curvature,* Mathematical Models and Methods in Applied Sciences. 14, pp. 1-45, 2004.

[4] A. Braides, R. March, *Approximation by Gamma convergence of a curvature dependent functional in visual reconstructions,* Preprint, July 2004.

[5] V. Caselles, R. Kimmel, and G. Sapiro, *Geodesic active contour,* International Journal of Computer Vision, Vol. 22(1), pp. 61-79, 1997.

[6] V. Caselles, F. Catté, T. Coll, and F. Dibos *A geometric model for active contours in image processing,* Numerische Mathematik, Vol 66(1), pp. 1-32, 1993.

[7] T. Chan, S. H. Kang, and J.H. Shen, *Euler's Elastica and Curvature Based Inpaintings,* SIAM J. Appl. Math, Vol. 63(2), pp. 564-592, 2002.

[8] T. Chan and L.A. Vese, *Active contours without edges,* IEEE Transaction on Image Processing, Vol. 10(2), pp. 266-277, 2001.

[9] T. Chan and L.A. Vese, *A level set algorithm for minimizing the Mumford-Shah functional in image processing,* In IEEE Workshop on Variational and Level Set Methods, pp. 161-168, Vancouver, CA, 2001.

[10] S. Esedoḡlu and R. March, *Segmentation with depth but without detecting junctions,* Journal of Mathematical Imaging and Vision, Vol. 18, pp. 7-15, 2003.

[11] S. Esedoḡlu and J. Shen, *Image inpaiting by the Mumford-Shah-Euler model,* European J. Appl. Math., Vol. 11, pp. 203-213, 2000.

[12] E. De Giorgi, *Some remarks on convergence and least squares methods,* Composite Media and Homogenization Theory, G Dal Maso and G.F. Dell's Antonio (Eds.), pp. 135-142, Birlhauser, Boston, 1991.

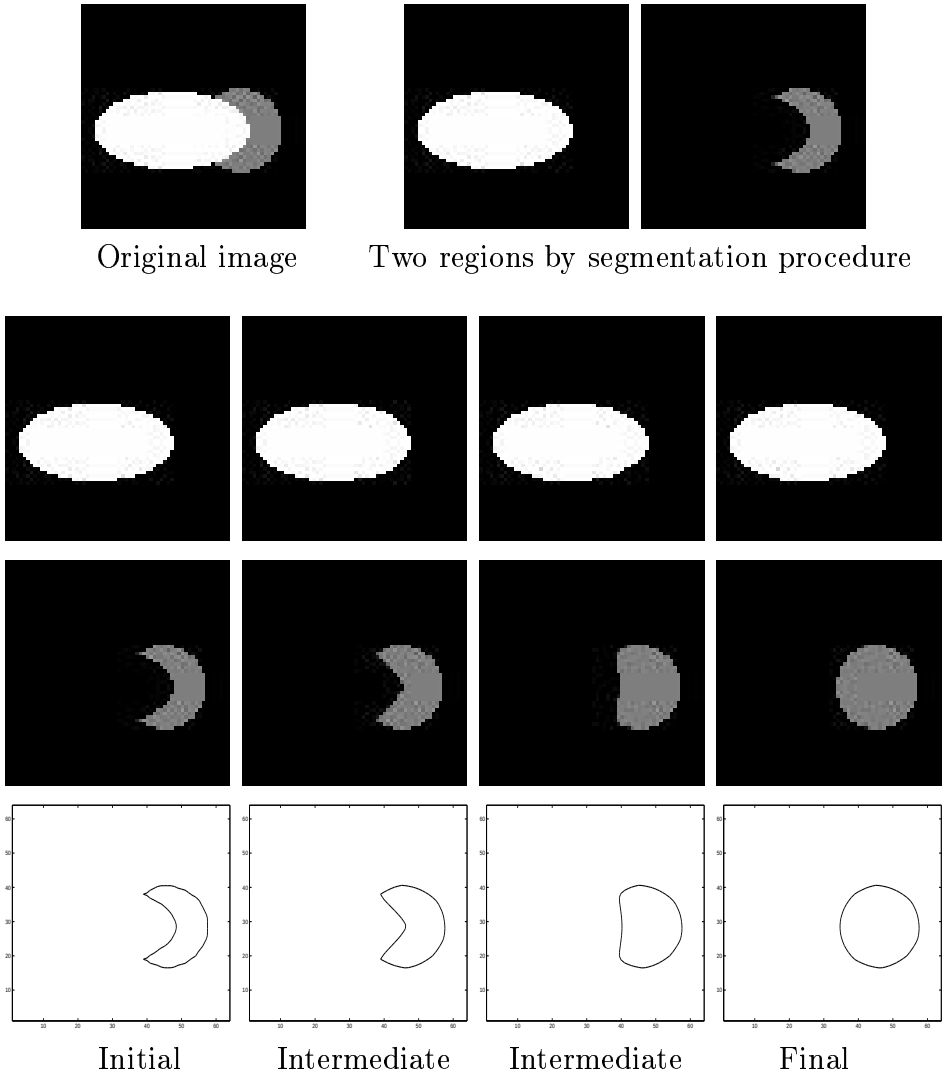


Figure 5: The evolution process under the assumption that the Ellipse is in front of the Crescent. In the second and third row, the evolutions of the two regions by solving (15) are listed. The fourth row shows how the boundary of the Crescent is growing. During the process, the front region occupied by the Ellipse is keeping stable while the region by the Crescent is connecting behind the Ellipse. In this experiment, the parameters are: $\alpha = 0.00005$, $\beta = 0.01$, and $\lambda = 0.1$ in Smereka's semi-implicit method.

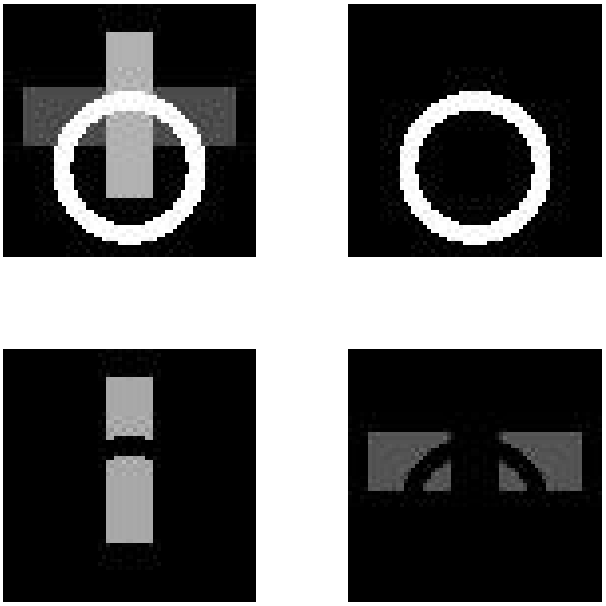


Figure 6: An Annulus and two Bars. The top left is the original image, and the other three figures show the regions obtained from the original image by standard segmentation model.

[13] M. Kass, A. Witkin, and D. Terzopoulos, *Snakes: Active contour models,* International journal of computer vision, Vol. 1, pp. 321-331, 1988.

[14] G. Dal Maso, *An introduction to Gamma-Convergence,* Birkhauser, Boston, 1993.

[15] R. March, *Visual reconstruction with discontinuities using variational methods,* Image and Vision Computing, Vol. 10, pp. 30-38, 1992.

[16] R. March and M. Dozio, *A variational method for the recovery of smooth boundaries,* Image and Vision Computing, Vol. 15, pp. 705-712, 1997.

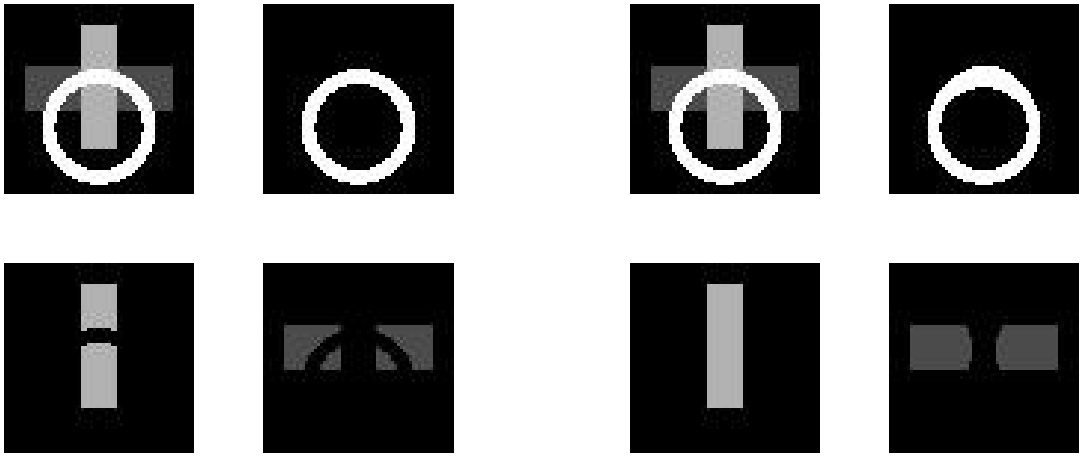
[17] R. Malladi, J.A. Sethian, and B.C. Vemuri, *Shape modeling with front propagation: A level set approach,* IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 17(2), pp. 158-175, 1995.

[18] A. Marquina and S. Osher, *Explicit algorithms for a new time dependent model based on level set motion for nonlinear deblurring and noise removal,* SIAM J. Sci. Comput., Vol. 22(2), pp. 387-405, 2000.

[19] D. Mumford and J. Shah *Optimal approximation by piecewise smooth functions and associated variational problems,* Comm. Pure Appl. Math., Vol. 42 pp. 577-685, 1989.

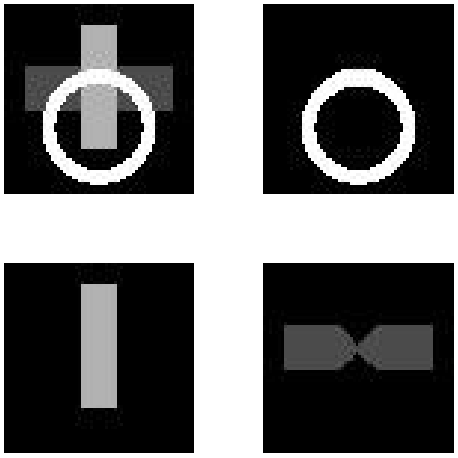
[20] J.M. Morel and S. Solimini, *Variational methods in image segmentation,* Birkhauser, Boston, 1995.

[21] M. Nitzberg, D. Mumford, and T. Shiota *The 2.1D Sketch,* Proceedings of the Third International Conference on Computer Vision, Osaka, 1990.

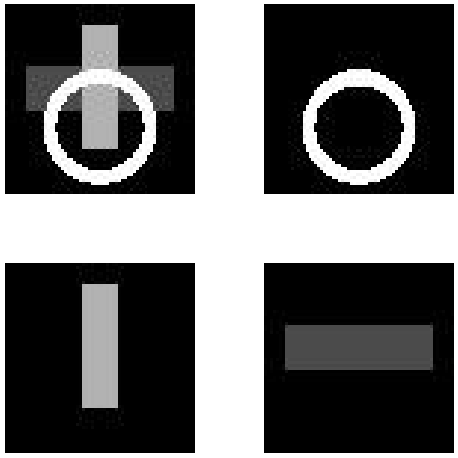


Step 1 (Initial Step)

Step 2 (Intermediate Step)



Step 3 (Intermediate Step)



Step 4 (Final Step)

Figure 7: The evolution process under the ordering assumption that the Annulus, the vertical Bar, and the horizontal Bar are listed from the nearest to the farthest. In each group of figures, the left top is the original image while the others show the evolution of each regions. During this process, the Annulus is almost keeping stable while the vertical Bar is connecting then the horizontal Bar. In this experiment, we choose the parameters are: $\alpha = 0.00005$, $\beta = 0.03$, and $\lambda = 0.1$ in Smereka's semi-implicit method.

[22] M. Nitzberg, D. Mumford, and T. Shiota *Filering, segmentation, and depth,* Lecture Notes in COmputer Science, Vol. 662, Springer Verlag, Berlin, 1993.

[23] S. Osher and J.A. Sethian, *Fronts Propagating with Curvature-Dependent Speed-Algorithm Based on Hamilton-Jacobi Formulations,* J. Comput. Phys, Vol. 79, pp. 12-49, 1988.

[24] S. Osher and R. Fedkiw, *Level set methods: an overview and some recent results,* J. Comput. Phys, Vol. 169(2), pp. 463-502, 2001.

[25] S. Osher and R. Fedkiw, *Level Set Methods and Dynamic Implicit Surfaces,* Springer-Verlag, 2002.

[26] P. Smereka, *Semi-implicit level set methods for curvature and surface diffusion motion,* Journal of Scientific Computing, Vol. 19, Nos 1-3, pp. 439-456, 2003.

[27] H.K. Zhao, T. Chan, B. Merriman and S. Osher, *A variational level set approach to multiphase motion,* J.Comput.Phys., Vol. 127, pp. 179-195, 1996.

[28] W. Zhu and T. Chan, *Capturing illusory contours: a level set approach,* UCLA CAM Report 03-65, 2003.