Local or Global Minima: Flexible Dual-Front Active Contours

Hua Li, *Member*, *IEEE*, and Anthony Yezzi, *Senior Member*, *IEEE*

Abstract—Most variational active contour models are designed to find local minima of data-dependent energy functionals with the hope that reasonable initial placement of the active contour will drive it toward a "desirable" local minimum as opposed to an undesirable configuration due to noise or complex image structure. As such, there has been much research into the design of complex region-based energy functionals that are less likely to yield undesirable local minima when compared to simpler edge-based energy functionals whose sensitivity to noise and texture is significantly worse. Unfortunately, most of these more "robust" region-based energy functionals are applicable to a much narrower class of imagery compared to typical edge-based energies due to stronger global assumptions about the underlying image data. Devising new implementation algorithms for active contours that attempt to capture more global minimizers of already proposed image-based energies would allow us to choose an energy that makes sense for a particular class of energy without concern over its sensitivity to local minima. Such implementations have been proposed for capturing global minima. However, sometimes the completely-global minimum is just as undesirable as a minimum that is too local. In this paper, we propose a novel, fast, and flexible dual front implementation of active contours, motivated by minimal path techniques and utilizing fast sweeping algorithms, which is easily manipulated to yield minima with variable "degrees" of localness and globalness. By simply adjusting the size of active regions, the ability to gracefully move from capturing minima that are more local (according to the initial placement of the active contour/surface) to minima that are more global allows this model to more easily obtain "desirable" minimizers (which often are neither the most local nor the most global). Experiments on various 2D and 3D images and comparisons with some active contour models and region-growing methods are also given to illustrate the properties of this model and its performance in a variety of segmentation applications.

Index Terms—Active contours, curve evolution, dual front evolution, morphological dilation, local minima, global minima, minimal path technique, level set methods, fast sweeping methods, image segmentation.

1 Introduction

Image segmentation is one of the first and most important tasks in image analysis and computer vision. In the computer vision literature, various methods dealing with object segmentation and feature extraction are discussed [1]. However, due to the variety and complexity of images, the design of robust and efficient segmentation algorithms is still a very challenging research topic.

Since the introduction of snakes [2], active contours have been applied to various problems in image processing and computer vision such as segmentation and feature extraction, image registration, shape analysis and modeling, and visual tracking. The original snake model was formulated to minimize the energy functional

$$E(C) = \alpha \int_0^1 |C'(s)|^2 ds + \beta \int_0^1 |C''(s)|^2 ds + \lambda \int_0^1 P(C(s)) ds, \quad (1)$$

where α, β, and λ are real positive weighting constants, $C : [0,1] \rightarrow \mathbb{R}^2$ is a parameterized curve, and $P(C)$ is a potential which depends upon some desirable image feature. In (1), the first two terms are internal forces, which control the regularity on curve C while the potential P attracts the curve C toward the desired boundary. Normally, the potential P depends on the gradient of the processed image.

As is well-known, the original snake models were dependent upon an arbitrary parameterization of the curve and had difficulty dealing with topology changes. A geometric active contour models, such as [3], [4], where introduced shortly afterward based on curve evolution theory, yielding a parameterization independent model which could also handle topology changes very naturally when implemented using level set methods proposed by Osher and Sethian [5].

The geometric active contour model most closely related to the original snake model is probably the *geodesic active contour* model [6], [7] which has the form

$$\frac{\partial C(s,t)}{\partial t} = P(C)(w+k)\vec{N} - (\nabla P \cdot \vec{N})\vec{N}, \quad (2)$$

where k is the curvature, $\vec{N}$ is the inward unit normal to the curve, and w is an optional constant *inflationary force* [8] (which parallels the *balloon force* introduced by [9] in the context of the original snake formulation).

Both the original snakes as well as geodesic active contours, however, are prone to getting "trapped" by extraneous edges due to image noise or texture, yielding many undesirable local minima of their corresponding energy functionals. As a result, initializations must be chosen carefully. This is a common trait of variational active contour models which typically designed to find local minima of data-dependent energy functionals with the hope that reasonable initial placement of the active contour will drive it toward a "desirable" local minimum rather

• *The authors are with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332.*
E-mail: {hua.li, ayezzi}@ece.gatech.edu.

Manuscript received 1 Sept. 2005; revised 20 Mar. 2006; accepted 19 May 2006; published online 13 Nov. 2006.
Recommended for acceptance by G. Sapiro.
For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number TPAMI-0472-0905.

than an undesirable configuration due to the noise or complex image structure. Devising new implementation algorithms for active contours that attempt to capture more global minimizers of already proposed image-based energies would allow us to choose an energy that makes sense for a particular image feature without concern over its sensitivity to local minima.

The minimal path technique proposed by Cohen et al. [10], [11], [12] is one such implementation which captures the global minimum of a contour energy between two fixed user-defined end points. In their technique, an image is defined as an oriented graph characterized by its cost function (or potential), thus the boundary segmentation problem becomes an optimal path search problem between two user-defined points in the graph. Their technique leads to a global minimum of an snake-like energy, thereby avoiding "intervening" local minima. However, this technique requires the user supplied end-points to be located precisely on the desired 2D boundary. A topology-based saddle search routine is needed to extended their technique to closed curve extraction. The original minimal path technique can be used for 3D tree-structured object extraction [13], but not for general 3D surface extraction. Ardon and Cohen [14] proposed a more general scheme for 3D surface extraction between user supplied end-curves, but this scheme also requires the user supplied curves to be located precisely on the desired 3D boundary (making it impractical for segmenting complex 3D structures without simple end-curves, such as the cortex in MRI brain imagery).

Other implementations have also been proposed for capturing more global minimizers by restricting the search space. One method with restricted search space was proposed by Gunn and Nixon via dual snakes [15], [16]. In their method, one snake is set inside the desired object and expands, whereas the other one is set outside the desired object and contracts. The two snakes are interlinked by arc-length and reach the inner and outer boundaries of the desired object, respectively. Similar methods were also proposed by Giraldi et al. [17] and Georgoulas et al. [18]. Aboutanos et al. [19] and Erdem et al. [20] restrict their search spaces by considering normals' lengths to an initial contour. Dawood [21], instead, proposed a dual-band with predefined width chosen by the Euclidean distance transform of the initial contour as a means to restrict the search space. While these methods may find more desirable minima for some images, they have several disadvantages. One is about the choice of search space, i.e., the desired boundary should be included in the search space. Second, these methods are restricted to detection of objects with simple topologies. Finally, not all of these methods are easy to extend to three dimensions.

In [22], Xu et al. combined active contours with the optimization tool of graph-cuts. They used morphological dilation to restrict the search space for graph-cuts segmentation. Their method may provide more global results and lead to smooth contours. However, some drawbacks exist. First, the computational cost is higher because a graph with appropriate pixel connectivity and edge weights needs to be prebuilt. Second, if an initial boundary is far from the actual object boundary, it may not find the actual boundary. This is similar to the local minimum problem of classical snakes and geodesic active contours. Finally, their method cannot be used for segmentation of multiple objects simultaneously.

All of the methods discussed so far have come to be known collectively as "edge-based" models. In many important applications, however, strong edge information is not always present along the entire boundary of the objects to be segmented. Therefore, the performance of purely edge-based models is often inadequate. As such, there has been much research into the design of complex region-based energy functionals that are less likely to yield undesirable local minima when compared to simpler edge-based energy functionals. In general, region-based models [23], [24], [25], [26], [27], [28], [29], [30] utilize image information not only near the evolving contour, but image statistics inside and outside the contour as well in order to yield more robust performance. Many of these methods were inspired by the "Region Competition" algorithm presented by Zhu and Yuille [31]. Unfortunately, most of these more "robust" region-based energy functionals assume highly constrained models for pixel intensities within each region and require a priori knowledge of the number of region types. These functionals are applicable to a much narrower class of imagery compared to typical edge-based energies due to stronger global assumptions about the underlying image data.

Whether by design of more robust energy functionals or by strategic implementation techniques, there has clearly been much research in efforts to yield active contour models which capture more global minimizers. It should be pointed out, however, that sometimes a minimum that is too global may be just as undesirable as a minimum that is too local. One example is illustrated in Fig. 6.

Finally, there is the issue of computational cost. Geometric active contour models implemented via level set methods [5] incur a heavy computational cost, even when using narrow band techniques [32]. Fast marching methods were proposed for monotonically evolving fronts [33], [34], [35], which are much faster. However, since the front can only move monotonically, it is prone to passing over the true boundary. These methods are also unable to incorporate curvature-based terms to control the smoothness of the evolving front, and it is difficult to design automated stopping criteria to end the front evolutions. Research on novel evolution schemes which combine the advantages of level set methods and fast marching methods is an very interesting topic. How to initialize freely, formulate potentials, and determine stopping criteria are also very important for the design of fast, yet powerful, curve evolution methods [36].

In this paper, we propose a novel, fast and flexible dual front implementation of active contours, motivated by minimal path techniques [10] and utilizing fast sweeping algorithms [37], [38]. This model is easily manipulated to yield minima with variable "degrees" of localness and globalness. The degree of global or local minima can be controlled in a graceful manner by adjusting the width of the active regions used to propagate the contour. This ability to gracefully move from capturing minima that are more local (according to the initial placement of the active contour/surface) to minima that are more global makes it much easier to obtain "desirable" minimizers (which often are neither the most local nor the most global).

This model is an iterative process of alternating dual front evolution and active region relocation which is robust to noise and poor initialization. Initializations can be chosen freely, thereby requiring limited or no user interaction. The dual front evolution forms a new global minimal partition curve within a narrow active region, which also guarantees the smoothness of evolving curves with the capability to handle topology changes. The whole iterative process convergences to the desired object automatically. The algorithm implementation

combines advantages of level-set methods and fast marching methods, while avoiding some of their disadvantages. The computational complexity, which is reduced significantly, is $\mathcal{O}(N)$, where N is the number of grid points involved in evolutions. Multiple objects can be segmented simultaneously just by one initial curve, and the model is easily extended to 3D and higher dimensions.

2 Dual-Front Active Contours

2.1 Background—Minimal Path Technique

In this section, we briefly review the minimal path technique proposed by Cohen et al. [10], [11], [12]. Their technique is a boundary extraction approach which detects the global minimum of a contour energy between two user-supplied points located on the desired boundary, thereby avoiding local minima arising from the sensitivity to initializations in snakes or geodesic active contours. Contrary to the energy functional (1) for traditional snakes [2], they proposed a simplified energy minimization model,

$$E(C) = \int_\Omega \{w + P(C(s))\} ds = \int_\Omega \widetilde{P}(C) ds, \tag{3}$$

without the second derivative term, where s represents the arc-length parameter on a defined domain $\Omega = [0, 1]$, $C(s) \in \mathbb{R}^n$ represents a curve, $\|C'(S)\| = 1$, $E(C)$ represents the energy along curve C, P is the potential associated to image features, w is a real positive constant, and $\widetilde{P} = P + w$. In this model, energy $E(C)$ includes the internal regularization energy (smoothing terms) in potential P, and controls the smoothness of curve C using P and $w > 0$.

Given a potential $P > 0$ that takes lower values near the desired boundary, the objective of the minimal path technique is to look for a path (connecting two user-supplied end points) along which the integral of $\widetilde{P} = P + w$ is minimal. A minimal action map $U_{p_0}(p)$ is defined as the minimal energy integrated along a path between a starting point p_0 and any point p,

$$U_{p_0}(p) = \inf_{A_{p_0,p}} \left\{ \int_\Omega \widetilde{P}(C(s)) ds \right\} = \inf_{A_{p_0,p}} \{E(C)\}, \tag{4}$$

where $A_{p_0,p}$ is defined as the set of all paths between p_0 and p. The value of each point p in the minimal action map $U_{p_0}(p)$ corresponds to the minimal energy integrated along a path starting from point p_0 to point p. So, the minimal path between point p_0 and point p can be easily deduced by calculating the action map $U_{p_0}(P)$ and then sliding back from point p to point p_0 on this action map U_{p_0} according to the gradient descent.

They also proposed that, if given a minimal action map U_{p_0} to point p_0 and a minimal action map U_{p_1} to point p_1, the minimal path between points p_0 and p_1 is exactly the set of points p_g which satisfy

$$U_{p_0}(p_g) + U_{p_1}(p_g) = \inf_p \{U_{p_0}(p) + U_{p_1}(p)\}. \tag{5}$$

They defined a saddle point p' as the first point that U_{p_0} and U_{p_1} meet each other, which means that p' satisfies $U_{p_0}(p) = U_{p_1}(p)$ and (5) simultaneously. The minimal path between points p_0 and p_1 may also be determined by calculating U_{p_0} and U_{p_1} and then, respectively, sliding back from the saddle point p' on U_{p_0} to point p_0 and from the saddle point p' on U_{p_1} to point p_1 according to the gradient descent. This idea was used in [12] for finding closed contours. In order to compute $U_{p_0}(p)$, they formulated a PDE equation,

$$\frac{\partial L(v,t)}{\partial t} = \frac{1}{\widetilde{P}} \vec{n}(v,t), \tag{6}$$

to describe the set of contours L in "time" t, where t represent heights of the level sets of U_{p_0} and values of energy $E(C)$, $v \in \mathbb{S}^1$ is an arbitrary parameter, and $\vec{n}(v,t)$ is the normal to the closed curve $L(v,t)$. These curves $L(v,t)$ correspond to the set of points p and the values of $U_{p_0}(p)$ on these points are equal to t. Equation (6) represents that a front starting from an infinitesimal circle around p_0, evolves with velocity $1/\widetilde{P}$ until each point inside the image domain is assigned a value for U_{p_0}. In fact, the minimal action map is also the potential weighted distance map. Because the action map U_{p_0} has only one minimum value at the starting point p_0 and increases from p_0 outward, it can be easily determined by solving the Eikonal equation

$$||\nabla U_{p_0}|| = \widetilde{P} \qquad \text{with} \qquad U_{p_0}(p_0) = 0. \tag{7}$$

They described three algorithms in [10], [11] to compute this map U_{p_0}, which are all consistent with the continuous propagation rule while implemented on a rectangular grid. These three algorithms utilize level set methods [5], [32], shape from shading methods [39], [40], and fast marching methods [35]. They favored fast marching methods to calculate U_{p_0} because of their lower complexity compared to the other two algorithms.

2.2 Principle of Dual-Front Active Contours

In this section, we propose the dual-front active contour model. Without loss of generality, we assume that an original image I has two regions R_0 and R_1, and B is the boundary of R_0 and R_1. We choose one point p_0 from R_0 and another point p_1 from R_1. Then, we define a velocity $1/\widetilde{P}$ taking lower values near the boundary B and define two minimal action maps $U_{p_0}(p)$ and $U_{p_1}(p)$ according to (4). Contrary to just considering the saddle point p' which satisfies $U_{p_0}(p') = U_{p_1}(p')$ and (5) simultaneously, we consider the set of points p_e which satisfy $U_{p_0}(p_e) = U_{p_1}(p_e)$. At these points p_e, the level sets of U_{p_0} meet the level sets of U_{p_1}. These points p_e form a partition curve B' which divides I into two regions. This partition is also a velocity (or potential) weighted Voronoi diagram. The region containing p_0 will be referred to as R'_0, while the other region containing p_1 will be referred to as R'_1. All points in R'_0 are closer to p_0 than to p_1 and all points in R'_1 are closer to p_1 than to p_0 in terms of U_{p_0} and U_{p_1}. Because the action maps are potential weighted distance maps, B' is called *the potential weighted global minimal partition curve*.

The level sets of U_{p_0} and U_{p_1} represent the evolving fronts, and the front evolving velocity $1/\widetilde{P}$ takes lower values near B. When an evolving front arrives at the actual boundary B, it evolves very slowly and, therefore, takes a long time to across B. By choosing appropriate potentials when defining U_{p_0} and U_{p_1}, we may cause the partition curve B' (the potential weighted global minimal partition curve) formed by the meeting points of the level sets of U_{p_0} and U_{p_1} to correspond with the desired boundary B.

Now, let us consider minimal action maps having a set of starting points. Similar to the definitions in [12], we let X be a set of points in image I (for example, X is a 2D curve or a 3D surface), and define a minimal action map $U_X(p)$ as the

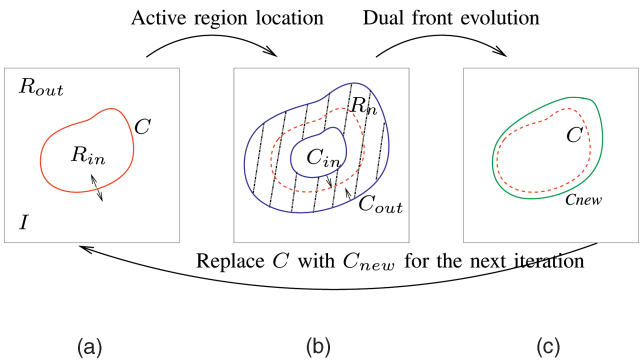

Fig. 1. Iteration process of the dual-front active contour model.

minimal energy integrated along a path between a starting point $p_0 \in X$ and any point $p \notin X$,

$$U_X(p) = \min_{p_0 \in X} \left(\inf_{A_{p_0,p}} \left\{ \int_\Omega \widetilde{P}(C(s))ds \right\} \right). \quad (8)$$

We choose a set of points X_i from R_0 and another set of points X_j from R_1 and define two minimal action maps $U_{X_i}(p)$ and $U_{X_j}(p)$ according to (8). All points satisfying $U_{X_i}(p) = U_{X_j}(p)$, form a partition boundary B'' and divide I into two regions. One region contains X_i and the other region contains X_j. Because $U_{X_i}(p)$ and $U_{X_j}(p)$ are the potential weighted distance maps, B'' is a potential weighted minimal partition of I. With appropriate potentials, it is also possible that B'' is exactly the actual boundary B of R_0 and R_1.

Therefore, we propose a dual front evolution principle to find a potential weighted global minimum partition curve within an active region. This principle is shown in Fig. 1. In Fig. 1a, an initial contour C separates image I to two regions R_{in} and R_{out}. In Fig. 1b, a narrow active region R_n is formed by extending the initial curve C. For example, it may be formed by dilating C with morphological dilation. R_n has an inner boundary C_{in} and an outer boundary C_{out}. As shown in Fig. 1c, the inner and outer boundaries C_{in} and C_{out} of R_n are set as the initializations of two minimal action maps $U_{C_{in}}$ and $U_{C_{out}}$. These minimal action maps $U_{C_{in}}$ and $U_{C_{out}}$ are defined by different potentials $\widetilde{P}_{in}$ and $\widetilde{P}_{out}$, respectively, based on (8). When the level sets of $U_{C_{in}}$ and $U_{C_{out}}$ meet each other, the meeting points form a potential weighted minimal partition curve C_{new} in active region R_n. The evolution of curves C_{in} and C_{out} and their meeting locations p_g can also be obtained using the "time of arrival" functions which satisfy the Eikonal equations

$$\begin{cases} ||\nabla U_{C_{in}}|| = \widetilde{P}_{in} & with \quad U_{C_{in}}(C_{in}) = 0 \\ ||\nabla U_{C_{out}}|| = \widetilde{P}_{out} & with \quad U_{C_{out}}(C_{out}) = 0 \\ U_{C_{in}}(p_g) = U_{C_{out}}(p_g) & on\ C_{new}. \end{cases} \quad (9)$$

Since the dual front evolution is to find the global minimal partition curve only within an active region, not in the whole image, the degree of this globalness may be changed flexibly by adjusting the size of active regions.

We now outline the complete dual-front active contour model. It is an iterative process including the dual front evolution and the active region relocation. It may be summarized in following steps:

1. Choose an initial contour and extend a narrow active region around this contour and extract its inner and outer boundaries.
2. Identify separated boundaries with different labels according to certain conditions and define potentials for the labeled contours.
3. Use the dual front evolution to propagate the labeled contours. The meeting points of the evolving contours separate the current active region in what we call the potential weighted minimal partition curve.
4. Use the new partition curve as an initial contour and repeat Steps 1 through 3 to find new global minimal partition curves again until convergence (i.e., when the new minimal partition curve is almost identical to the previous curve).

In dual-front active contours, the segmentation objective of finding a minimum with a certain degree of localness/globalness in a certain region is translated into finding a global minimum partition contour within a narrow active region extended from the initial curve, and then iteratively changing this active region and reminimizing until convergence. After each iteration (before convergence) the contour moves dramatically in comparison with standard level set PDE iterations. The whole evolution process of dual-front active contours looks like a sequence of big "jumps" in the curve before it arrives at the final desired boundary.

The iterative process may obviously be terminated manually, however, some natural automatic methods may be employed as well. The most natural automatic method, and the one that we employ in our experiments, is to halt the iterative procedure as soon as difference between two consecutive contours is insignificant (or even zero). This is effective and also quite natural given that large jumps are expected each iteration as a global minimization occurs each iteration (within the active region).

Dual-front active contours can handle topology changes just like standard level set PDE evolutions. There are two ways to decide the labels of the separated boundaries in Step 2. First, the labels may be reset in each iteration loop. The number of labels is decided by the number of separated boundaries. In this case, the number of labels may be changed and different with the number of the user-defined initial curves. So, we can detect several disjoint objects just by one initial curve. One example is shown in Fig. 10.

Second, the labels may be decided just according to the number of the user-defined initial curves. In each iteration loop, the labels of the separated boundaries of the new active region are decided by the result from the previous iteration. For example, if one point doesn't belong to the current active region, we choose this point's label in the previous iteration as this point's current label. So, the labels' number remains constant during the whole iterative process even when the evolving curve's topology structure changes. This restriction is suitable for detecting objects with tiny protuberant parts, such as the surface of gray matter or white matter of human brains. In this manner, tiny parts of an object are always connected with other big parts of the same object by keeping the same label, and the segmentation result can be more accurate. We use this restriction in 3D brain cortex segmentation and one experiment is shown in Fig. 13.

3 Properties of Dual-Front Active Contours

In this section, we analyze a number of properties of dual-front active contours. These include flexible local-to-global

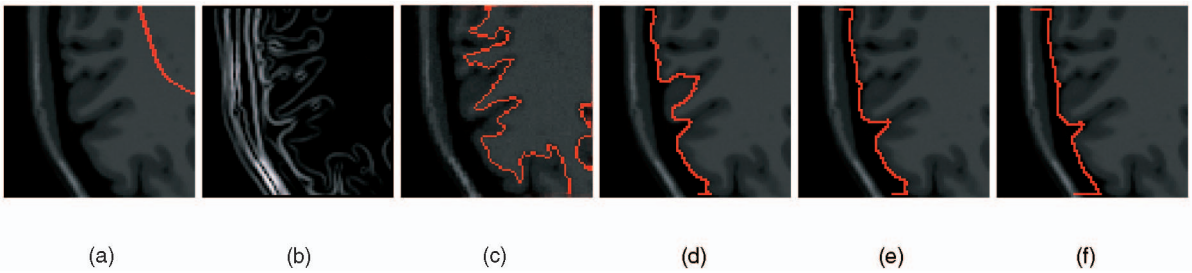

Fig. 2. By choosing active regions with different sizes, the dual-front active contour model captures minima with varying degrees of globalness/localness.

minima, evolution schemes and computational efficiency, evolution potentials, regularization terms, and evolution convergence.

3.1 Flexible Local or Global Minima

As discussed in Section 1, problems with overly global minima and problems with overly local minima problems are common in minimal path methods and active contour methods, respectively. But, in dual-front active contours, the degree of global or local minima can be controlled in a graceful manner by adjusting the width of the narrow active region and the weight of region-based and/or edge-based information combined in potentials for the dual front evolution. This ability to gracefully move from capturing minima that are more local (according to the initial placement of the active contour/surface) to minima that are more global makes this model much easier to obtain "desirable" minimizers (which often are neither the most local nor the most global).

The result of the dual front evolution is a potential weighted global minimum partition curve within an active region. So, the size and shape of active regions affects final segmentation results. If the active region's size is extremely small, it may lead to local minima problems because of the effect of noise and local image structure. If the active region's size is extremely large, the desired boundary may be missed as an overly global minimizer may be favored instead. The active region's size should be selected based on the shape and size of detected objects, the image's noise level, the structure of the background, etc. In dual-front active contours, we provide a very flexible way to define active regions. We may treat an active region as a restricted search space and many methods can be used to decide active regions. For example, as introduced in Section 1, active regions (or restricted search spaces) can be decided by calculating the Euclidean distance [21] to an initial curve or the normal's length [19], [20] of an initial curve, or by using morphological dilation [22], [41] to form a narrow region around an initial curve, or even by using simple thresholding operators [42]. In our experiments, we generally use morphological dilation and erosion to generate an active region around the current curve. In this way, the size and shape of the active regions can be controlled easily by adjusting the associated structuring elements. It is even useful sometimes (for additional computational speed) to dynamically change the size of the active region during the iteration processes. For example, when an initial curve is far from the desired object, we may first use wider active regions to extend the searching scale for each iteration, speed up the computational time and avoid the effect from noise. After a few iterations, as the curve nears the desired boundary, we may use narrower active regions to speed up the remaining iterations.

In Fig. 2, we demonstrate that by choosing different active regions with different sizes, dual-front active contours may achieve different degrees' global minima in the whole image. The potential of point (x, y) in image I was chosen as $\widetilde{P}(x, y) = |I(x, y) - \mu_l| + (1 + |\nabla I|)^2/10 + 0.1$, where μ_l is the mean value of points having the same label l as the point (x, y). Fig. 2a shows the original image with the initialization, Fig. 2b shows the corresponding gradient information, Figs. 2c, 2d, 2e, and 2f show the different segmentation results using 5×5, 7×7, 11×11, and 15×15 pixels circle structuring elements for morphological dilation after 15 iterations, respectively.

3.2 Numerical Implementation

In this section, we analyze the numerical implementation of dual-front active contours. As shown in (9), the dual front evolution detects the meeting points of the level sets of two minimal action maps $U_{C_{in}}$ and $U_{C_{out}}$. These two minimal action maps may be computed by solving Eikonal equations $\|\nabla U\| = \widetilde{P}$. In the minimal path technique proposed in [10], they used fast marching methods described in [35] to solve Eikonal equations. Fast marching methods are computationally efficient tools to solve Eikonal equations, in which upwind difference schemes and heap-sort algorithms are used for guaranteeing the solution is strictly increasing or decreasing on grid points. The computational complexity of fast marching methods is $\mathcal{O}(N \log N)$, where N is the number of grid points, and $\log N$ comes from the heap-sort algorithm. Tsitsiklis [34] first used heap-sort structures to solve Eikonal equations, Sethian [35] and Helmsen et al. [33] also reported similar approaches lately. Recently, Yatziv et al. [43] proposed a new implementation of fast marching methods which reduces the computational cost of fast marching methods to linear complexity $\mathcal{O}(N)$ using a data structure "untidy priority queue" instead of a heap-sort structure.

Another algorithm for solving Eikonal equations is the fast sweeping method [37], [38] based on iteration strategies. It is suited for computing the solution of Eikonal equations on a rectangular grid. Its main idea is to combine nonlinear up-wind differences and Gauss-Seidel iterations with alternating sweeping orders so that the causality along characteristics of all directions is followed in an optimal way. The idea of alternating sweeping order was first proposed by Danielesson [44] to compute distance mapping. But, Danielsson's algorithm cannot be used for calculating the distance to a curve or a surface. Another discrete approach based on the idea of fast sweeping was also proposed by Tsai [45] to compute distance functions, but their work does not apply to general Eikonal equations

either. In fast sweeping methods [37], [38], the characteristics are divided into a finite number of groups according to their directions and each sweep of Gauss-Seidel iterations with a specific order covers a group of characteristics simultaneously. 2^n Gauss-Seidel iterations with alternating sweeping order are used to compute a first order accurate numerical solution for the distance function in n dimensions. Fast sweeping methods have an optimal complexity of $\mathcal{O}(N)$ for N grid points, are extremely simple to implement in any dimension, and give similar results as fast marching methods. The details of fast sweeping methods may be seen in [38], and their extension to more general Hamilton-Jacobi equations are discussed [46], [47].

Both fast marching methods and fast sweeping methods can be used in the dual front evolution for finding the minimal partition curve in an active region. When using fast marching methods to implement the dual front evolution, there is no need to solve Eikonal equations shown in (9) on the whole active region R_n and then look for the set of p_g. The dual front evolution may be implemented by labeling initial curves with different labels and evolving the labeled curves with different potentials simultaneously until each point inside the active region is assigned a label. The computational complexity of the dual front evolution is $\mathcal{O}(N \log N)$, where N is the number of grid points in R_n.

In this paper, in contrast with more classical minimal path techniques [10], the dual front evolution scheme utilizes fast sweeping methods because of their lower complexity $O(N)$, where N is the number of grid points in R_n. Dual-front active contours includes the dual front evolution and morphological dilation. Because the low-computational cost of fast sweeping methods is maintained and the calculation of all minimal action maps can be finished simultaneously, the complexity of the dual front evolution is still $\mathcal{O}(N)$. Furthermore, the complexity of morphological dilation is lower than $\mathcal{O}(N)$, and the boundary tracking process can be finished in finite iterations. So, the total complexity of dual-front active contours is still $\mathcal{O}(N)$, where N is the number of grid points (on average) in an active region.

3.3 Evolution Potentials

Generally, active contour models can be divided into two categories, edge-based approaches and region-based approaches. Defining appropriate potentials is an essential task of dual-front active contours. Here, we consider potentials combining both region and edge information.

Without loss of generality, we assume an image domain Ω_I includes an object Ω_O with boundary Γ_O and background Ω_B. We also choose an initial contour C_0. In the tth iteration of dual-front active contours, a closed curve $C(t-1)$ is dilated/eroded to form an active region $R_n(t)$ with two boundaries $C_{in}(t)$ and $C_{out}(t)$, and Ω_I is divided into five subsets: three open subsets $R_{in}(t)$, $R_{out}(t)$, $R_n(t)$ and their common boundaries $C_{in}(t)$, $C_{out}(t)$. This situation just looks like that shown in Fig. 1b. Similar to the definitions in [30], we also define two region descriptors $k^{in}(t)$ and $k^{out}(t)$, which are globally attached to their respective regions $R_{in}(t)$ and $R_{out}(t)$ and, therefore, depend on them. For example, statistical features of regions like means or variances fall in this class of region descriptors. We also define boundary descriptors $k^b(t)$, such as edge strength indicators used in many edge-based active contour models [6], [7] which are typically functions of the image gradient ∇I.

For dual-front active contours, therefore, let us consider energy functionals

$$\begin{cases} E_{in}(C) = \omega_{in}^r \int\int_{R_{in}} k^{in}(x,y,R_{in})dxdy \\ \qquad\qquad +\omega_{in}^b \int_\Omega (k^b(x,y)+\omega_{in})ds \\ E_{out}(C) = \omega_{out}^r \int\int_{R_{out}} k^{out}(x,y,R_{out})dxdy \\ \qquad\qquad +\omega_{out}^b \int_\Omega (k^b(x,y)+\omega_{out})ds, \end{cases} \tag{10}$$

where ω_{in}^r, ω_{out}^r, ω_{in}^b, ω_{out}^b, ω_{in}, and ω_{out} are real positive weighting constants. The integrands contain all of the information to be extracted from the image. The objective of dual-front active contours is to make an initial contour $C(0)$ evolve toward a final minimal partition $(\Omega_O, \Omega_B, \Gamma_O)$ of image Ω_I by minimizing $E_{in}(C)$ and $E_{out}(C)$.

In tth iteration, the minimal partition curve $C(t)$ is the global minimum of the energy

$$\inf_{C \in R_n(t)} \{E_{out}(C) + E_{in}(C)\} \approx E_{out}(C(t)) + E_{in}(C(t)) \tag{11}$$

within the narrow active region $R_n(t)$ enclosed by $C_{in}(t)$ and $C_{out}(t)$. According to (4), (7), and (10), we may calculate (9) according to potentials

$$\begin{cases} \widetilde{P}_{in}(x,y) = w_{in}^r \times f\Big(|I(x,y)-\mu_{in}|, \sigma_{in}^2\Big) \\ \qquad\qquad +w_{in}^b \times g(\nabla I(x,y)) + w_{in} \\ \widetilde{P}_{out}(x,y) = w_{out}^r \times f\Big(|I(x,y)-\mu_{out}|, \sigma_{out}^2\Big) \\ \qquad\qquad +w_{out}^b \times g(\nabla I(x,y)) + w_{out}, \end{cases} \tag{12}$$

where $I(x,y)$ is the image intensity, $g(\nabla I(x,y))$ is a function of the image gradient, μ_{in} and μ_{out} are mean values of I, and σ_{in}^2 and σ_{out}^2 are variances within the regions R_{in} and R_{out}. In (12), region and edge information is combined in the potentials for the dual front evolution. In dual-front active contours, different functions f and g and different weights for the components of the potentials should be chosen for different segmentation objectives. For example, if the desired object has strong and reliable edge information, we should increase the weight of the edge descriptors. Otherwise, if the desired objects have weak edges or the image is very noisy, we should increase the weight on the region descriptors. As with any segmentation algorithm, the optimal set of parameters is very application dependent.

3.4 Simple Regularization Terms

In classical active contour models, regularization forces arise from internal energy terms in the energy that penalize length or elasticity of the evolving contour. The resulting forces typically take the form of mean curvature flows (or derivatives of mean curvature in the case of elasticity). While curvature-based terms do not explicitly appear in our schemes, we may still obtain regularity in the design of the potential functions. We propose three solutions to control the regularity of the evolving fronts.

The dual-front active contour model is motivated by the minimal path technique [11], in which they introduced an upper bound on the curvature along an extracted contour $C(s)$ by controlling potentials P. They proved that given a potential $P > 0$ defined on an image domain D, and $\widetilde{P} = \omega + P$, the curvature magnitude $|k| = \|C_{ss}\|$ along the geodesic minimizing $\int_\Omega (\omega + P(C(s)))ds$ is bounded by

$$|k| \le \sup_D \frac{\|\nabla P\|}{\omega}. \tag{13}$$

According to this relationship, we propose two solutions for ensuring smooth contours in the dual front evolution algorithm. The first solution is to decrease $\sup_D$ by increasing the constant ω added to P. Fig. 3a shows an

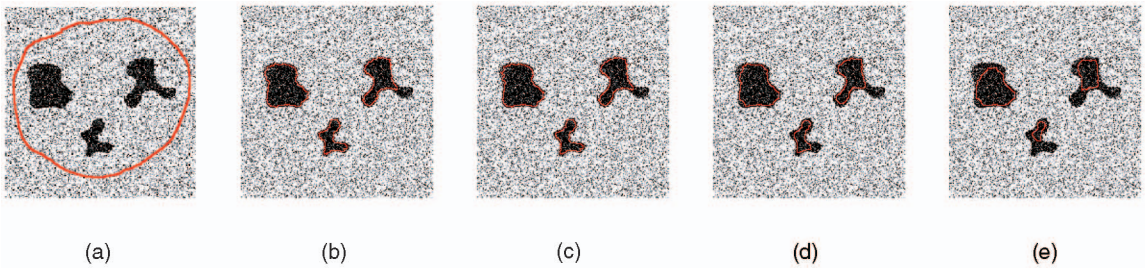

Fig. 3. The smoothness of contours is adjusted by changing the constant w added to the potential P.

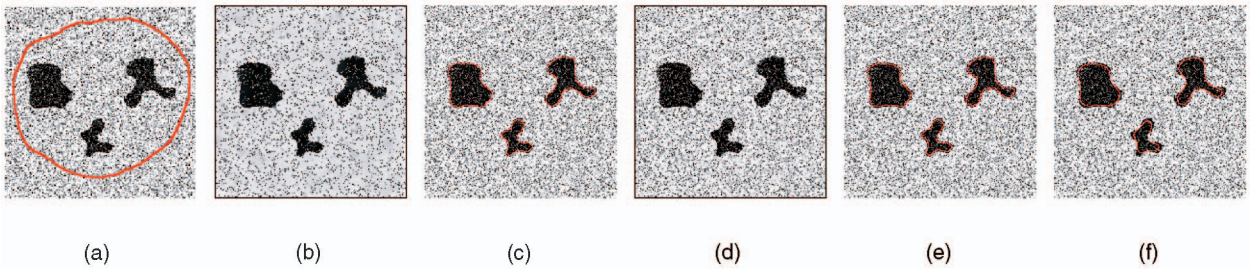

Fig. 4. The extracted boundary also can be smoothed using the last two solutions proposed in Section 3.4.

original image with an initial contour, Figs. 3b, 3c, 3d, and 3e show the different smoothing effects in dual-front active contours by changing the coefficient w in the potential $\widetilde{P}(x,y) = (|I(x,y) - \mu_l| + \omega)$ from 0.1 to 10, 100, 200, where μ_l is the mean value of points having the same label l as the point (x,y). This Fig. 3 illustrates that by increasing the constant w added to the potential P, the smoothness of contours is increased. In this figure, the structuring element for the morphological dilation is a 5×5 circle mask and all results are obtained after 20 iterations.

The second solution is to decrease $\sup_D$ by smoothing the potentials (or original images) instead of increasing the constant ω. Gibou and Fedkiw [48] developed a hybrid numerical technique for image segmentation that draws on the speed and simplicity of k-means procedures and the robustness of level set algorithms. They suggested that a diffusion-like operation can be used to replace mean curvatures to regularize the level set evolution. Motivated by their method, we take the isotropic nonlinear diffusion operator proposed by Perona and Malik [49] for denoising an image while still keeping the image edges. This nonlinear equation is

$$\frac{\partial I(\vec{x},t)}{\partial t} = \nabla \cdot (g(|\nabla I|)\nabla I), \tag{14}$$

where I defines the image intensity map at the voxel location x and fictitious time t, g is an edge-stopping function, and $\lim_{|\nabla I| \to \infty} g(|\nabla I|) = 0$ so that diffusion stops at the location of large gradients. Based on the original function proposed by Perona and Malik [49], we take

$$g(s) = \nu/(1 + |\nabla I|^2/K^2) \tag{15}$$

to smooth image I, where K is a threshold parameter tuning the edge-stopping sensitivity on the image gradient and ν is a parameter controlling length scales.

The third solution is to use postprocessing operators to smooth results. Level set evolutions with curvature terms are good choices for smoothing curves (or surfaces). In [50], [51], they combined fast matching evolutions with level set evolutions to extract desired objects. Level set evolutions with mean curvature terms start from the segmentation result achieved by fast marching evolutions. Refined final results are received after a just a few iterations. In dual-front active contours, we also can use level set evolutions to refine the obtained curves or surfaces.

In Fig. 4, we give two examples to illustrate how the last two solutions control the smoothness of segmentation results. The potential at a point (x,y) was chosen as $\widetilde{P}(x,y) = (|I(x,y) - \mu_l| + 0.1)$, where μ_l is the mean value of points having the same label l as the point (x,y). Fig. 4a shows an original image with initializations. Fig. 4b shows the smoothed original image using isotropic nonlinear diffusion operator shown in (15) with $K = 20$ and $\nu = 1$ after 120 iterations. Fig. 4c shows the segmentation result from dual-front active contours. We also test the smoothness effect from the third solution by using postprocessing operator. Fig. 4d shows the smoothed original image using isotropic nonlinear diffusion operator shown in (15) with $K = 20$ and $\nu = 1$ after 20 iterations. Fig. 4e shows the segmentation result using dual-front active contours. Fig. 4f shows the refined result after using 10 iterations mean curvature flow evolution proposed in [50] on the result in Fig. 4e.

3.5 Automatic Evolution Convergence

The stopping criterion in active contour models is also an important issue whether using fast marching methods or level set methods. Generally, fast marching methods do not give explicit stopping criteria. Level set methods reach the desired boundary and keep an equilibrium at the boundary only with appropriate energy functionals. Now, we analyze the evolution convergence properties of dual-front active contours.

First, the dual front evolution provides an automatic stopping criterion in each iteration. Since all initial contours are classified into two (or multiple) groups, all contours evolute simultaneously but based on different potentials. Whenever two contours from the same group meet, they merge into a single contour. On the other hand, if two contours from different groups meet, both contours stop evolving and a common boundary is formed by the meeting points automatically. This automatic stopping criterion is similar to that in region-growing methods in [52] and

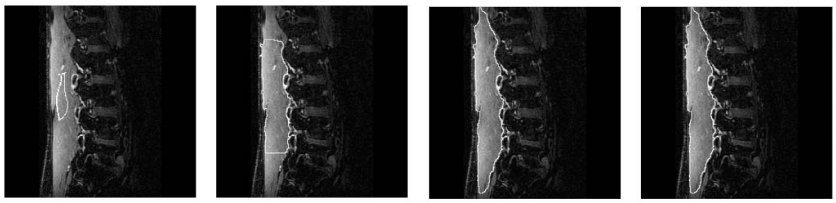

Fig. 5. The convergence property of dual-front active contours. The structuring element for morphological dilation is a 7×7 circle mask. (a) The original image and the initialization. (b), (c), and (d) The segmentation results after 15, 30, and 60 iterations, respectively.

multilabel fast marching methods [36], [51], [53]. The comparison of those methods with dual-front active contours is introduced in Section 4.2. The dual front evolution also guarantees the continuity and smoothness of curves with the capability to handle topology changes.

Second, the iteration process of dual-front active contours may be stopped automatically by comparing results from consecutive iterations. In each iteration, the result of the dual front evolution is a global minimum partition curve within the current narrow active region. After a finite number iterations, by taking significant jumps each time, the evolving curve reaches the desired boundary. When current global minimum partition curve is the same as that of last iteration or the difference between them is less than a predefined tolerance, the procedure may be stopped. Because in each iteration, the global minimum partition curve is confined to the active region, the size of active regions decides the degree of globalness/localness of the minimizer. The result using a narrow active region may eventually to the result in a large active region as long as there are no intervening local minimizers.

Fig. 5 gives an example to show this property. In this figure, the segmentation result after 30 iterations is the same as that after 60 iterations from dual-front active contours. We set the stopping criterion such that the algorithm terminates when a curve reoccurs exactly in consecutive iterations. The potential at a point (x, y) was chosen as $\widetilde{P}(x, y) = (|I(x, y) - \mu_l| + (1 + |\nabla I|)^2/10 + 0.1)$, where μ_l is the mean value of points having the same label l as the point (x, y).

4 Comparison with Other Boundary Extraction Methods

In this section, we compare and contrast dual-front active contours with some other active contours for boundary extraction. Since dual-front active contours combine region and boundary constraints as well as a number of properties from both level set methods and minimal path based fast marching methods, we compare our method to such edge-based approaches as geodesic active contours [6], [7] and the minimal path technique [10] as well as to region-based approaches such as Chan-Vese's model [24] and the more general Mumford-Shah model [54]. Finally, because of their evolution properties, we also compare dual-front active contours with region-growing methods such as watershed algorithms [52], and multilabel fast marching methods [51], [53].

4.1 Comparison with Edge-Based and Region-Based Approaches

As mentioned in Section 1, geodesic active contours [6], [7] are based on the planar evolution (2) to capture localized image features, most notably edges. But, spurious edges generated by noise may stop the evolution of curves, yield undesirable local minimizers. Therefore, initializations must be chosen very carefully. The minimal path technique proposed by Cohen and Kimmel [10] instead captures global minimizers of the same energy between two user-defined points. However, these two initial points must be located precisely on the desired boundary.

Region-based forces have already been introduced in many active contour models [23], [24], [29], [31], [54], [55]. Considering a possibly noisy 2D image u_0 with image domain Ω_I and segmenting curve C. Mumford and Shah [54] proposed to decompose an image into piecewise-smooth functions by minimizing the following energy:

$$E^{MS}(u, C) = \mu L(C) + \lambda \int_{\Omega_I} (u_0(x, y) - u(x, y))^2 dxdy + \nu \int_{\Omega_I \setminus C} |\nabla u(x, y)|^2 dxdy, \tag{16}$$

where $L(C)$ denotes the length of C, μ, ν, and λ are positive parameters.

Chan and Vese [24] considered the limiting case (as did Mumford and Shah earlier) of infinite penalty on the final term, leading to piecewise constant approximations of the image by minimizing the following energy

$$E(c_1, c_2, C) = \mu L(C) + \lambda_1 \int_{in(C)} (u_0(x, y) - c_1)^2 dxdy + \lambda_2 \int_{out(C)} (u_0(x, y) - c_2)^2 dxdy, \tag{17}$$

where c_1 and c_2 are constants and $in(C)$ and $out(C)$ denote the interior and exterior of C, respectively. μ, λ_1, and λ_2 are positive parameters.

In Fig. 6, we compare geodesic active contours [6] (shown in Fig. 6a), the minimal path technique [10] (shown in Fig. 6b), Chan-Vese's method [24] (shown in Fig. 6c), and Mumford-Shah method [54] (shown in Fig. 6d), with dual-front active contours (shown in Fig. 6e). The test image is one 2D slice of a human brain MRI image, and the objective is to find the interface of gray matter and white matter. The image size is 80×80 pixels. The structuring element for the morphological dilation step of the dual-front active contour model was chosen to be a 5×5 circle mask. The gradient information used in panel Figs. 6a, 6b, and 6e is shown in Fig. 2. The top row shows the original image and the

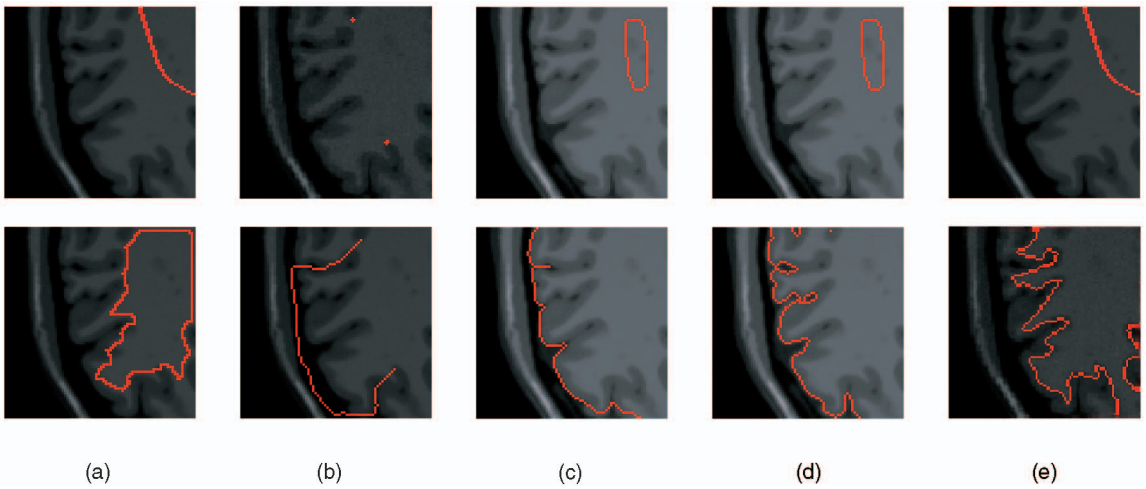

Fig. 6. Comparison of different segmentation results of the interface of white matter and gray matter via different active contour models with different degrees of local or global minima.

initializations for the curve evolutions. The bottom row shows corresponding edge segmentation results. As this figure indicates, the neither completely local nor completely global minimum found by dual-front active contours yields much more desirable boundaries compared to the local minimum of the geodesic active contour and the global minimum of the minimal path technique.

In Fig. 7, we give two examples which compare Chan and Vese's method [24], and the Mumford-Shah method [54] with dual-front active contours. Figs. 7a and 7e are two 2D medical images with the initializations, Figs. 7b and 7f are the results from Chan-Vese's model after 100 iterations, Figs. 7c and 7g are the results from Mumford-Shah model after 200 iterations, Figs. 7d and 7h are the results using dual-front active contours with a 7×7 circular structuring element and after 20 iterations. Because Chan and Vese's method and Mumford's method use global terms in the evolution equation, sometimes these methods cannot find the correct boundaries. But, for dual-front active contours, the degree of localness/globalness can be controlled by the size of the active regions, thereby yielding more flexibility to cope with images with more complicated structure while still avoiding purely local minimizers due to noise. The potential at a point (x, y) for the dual-front active contour model was chosen as $\widetilde{P}(x, y) = (|I(x, y) - \mu_l| + 0.1)$, where μ_l is the mean value of points having the same label l as the point (x, y). The parameters in Chan and Vese's method and in Mumford and Shah's method were chosen to be $\lambda = \lambda_1 = \lambda_2 = 1$, $\Delta t = 0.5$, and $\mu = 0$.

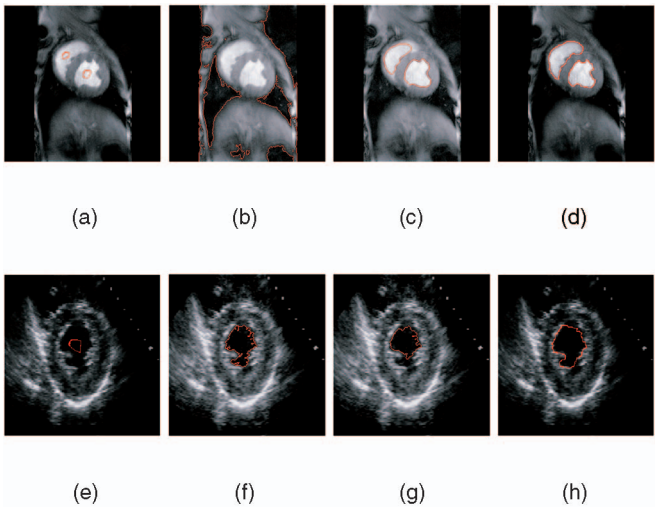

Fig. 7. Comparison of Chan-Vese's methods, Mumford-Shah's methods, and dual-front active contours.

4.2 Comparison with Morphological Watershed Transform and Multilabel Fast Marching Method

The watershed transform proposed by Vincent and Soille [52] is a well-known segmentation technique, which is based on immersion simulation, and allows the partitioning of an image into regions. This technique is based on the assumption that image contours correspond to the crest lines of the gradient magnitude which can be detected via watershed tracing. The strength of watershed segmentation is that it produces a unique solution for a particular image and can be easily adapted to any kind of digital grid or extended to n-dimensional images and graphs. However, the watershed transform typically leads to an over-segmentation because the flowing process strongly relies on the quality of the gradient and the choice of seed points. Normally, the seed point is chosen according to local minima and at least one seed must be initialized within each interesting object. But, very often, local minima are extremely numerous for noisy images, and it is a hard task to choose appropriate local minima seeds.

Another interesting region-growing approach is the multilabel fast marching method presented by Sifakis and Tziritas [53] for motion analysis in video processing. It is an extension of fast marching methods. The contour of each region is propagated according to a motion field which depends on labels and the absolute interframe difference. The propagation speeds of labeled contours are based on statistical descriptions of the propagated classes. Deschamps and coauthors [51] also proposed a similar multilabel fast marching method, in which the speed function is derived from local information (not just gradients) together with each region's statistical information.

The evolution process in multilabel fast marching methods is analogous to the immersion process used in watershed methods [52]. Both of them need a sufficient number of appropriate initial seed points for accurate

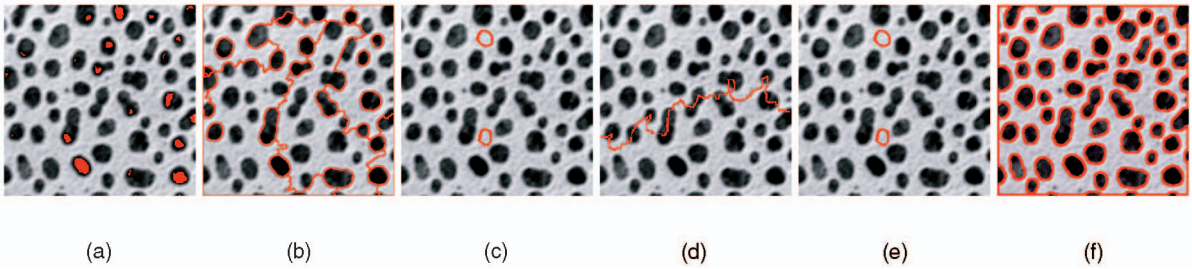

Fig. 8. Comparison of watershed algorithms, multilabel fast marching methods, and dual-front active contours.

segmentation. The number of seed points determines the number of final segmented regions. These methods are typically used as initial segmentations which requiring following refinement.

The difference between dual-front active contours and the above two methods (the multilabel fast marching methods and the watershed methods) is that the dual-front active contour method is an iterative method, which allows topology changes with simple and free initializations. In addition, dual front evolutions also allow similar groups of contours to merge, thereby avoiding oversegmentation problems common to watershed immersion methods. Dual-front active contours consider not only image gradients, but also the important region information, which is not considered in watershed transform methods.

In Fig. 8, we compare morphological watershed algorithms [52] (shown in Fig. 8b), multilabel fast marching algorithms [51], [53] (shown in Fig. 8d), and dual-front active contours (shown in Fig. 8f). Figs. 8a, 8c, and 8e show the original images with initializations, and Figs. 8b, 8d, and 8f show the corresponding segmentation results. Dual-front active contours only need very simple initializations to segment multiple cell blobs and obtain accurate results. But, for watershed algorithms and multilabel fast marching algorithms, each blob and the background should be assigned a distinctly labeled seed point, and no two regions are allowed to merge during the region-growing process. The number and the location of the initial seed points also affect the final result dramatically. The potential used both for the multilabel fast marching method and for the dual-front active contour was chosen as $\tilde{P}(x,y) = (|I(x,y) - \mu_l| + 0.1)$, where μ_l is the mean value of points having the same label l as the point (x,y). The structuring element for the dual-front active contour was chosen to be a 7×7 circle mask.

5 Experimental Results

In addition to the experiments shown in Section 4, we give more experiments on 2D and 3D real images to illustrate the properties of dual-front active contours in this section.

5.1 Experimental Results on 2D Images

In dual-front active contours, one very attractive feature is that this model automatically proceeds in the correct directions without relying upon additional inflationary terms commonly employed by many active contour models. To illustrate this property, we compare the segmentation results with those in [28] using a similar noisy synthetic hand image in Fig. 9. An initial contour completely contained within the object flows outward to the boundary (shown in the first row). An initial contour partially inside and partially outside the interesting object flows in both directions toward the boundary (shown in the second row). An initial contour encircling the interesting object flows inward to the boundary (shown in the third row). And, an initial contour located outside the interesting object flows outward and wraps around the boundary (shown in the last row). The first column shows the original image and the initializing contour. The second and the third column show two intermediate steps of the algorithm. The last column shows the final segmentation result. The results illustrate that the shape of the hand can be segmented automatically and naturally. In this example, the potential at a point (x,y) was chosen as $\tilde{P}(x,y) = (|I(x,y) - \mu_l| + 0.1)$, where μ_l is the mean value of points having the same label l as the point (x,y). The structuring element for dilation was chosen as a 7×7 circle mask.

In their experiment on a similar image in [28], when an initial contour is outside the object, in addition to the final curve that outlines the boundary of the hand, there are extraneous curves around four corners of the image which do not correspond to the actual image edges due to an undesired local minimum. However, in our case, there are no extraneous curves near the corners in Fig. 9. This is because in each dilation/evolution loop of dual-front active contours, the detected curve is limited in the current active region and

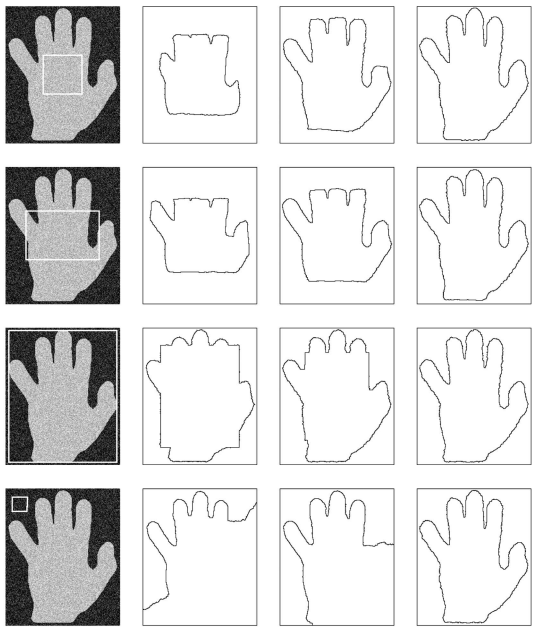

Fig. 9. The segmentation results on a hand image with different initializations.

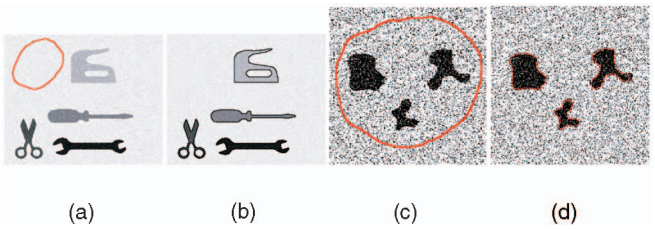

Fig. 10. Detection of multiple objects with different intensities in two noisy images with a 5×5 structuring element. (a) and (c) Two original images with initializations. (b) and (d) The segmentation results after 20 iterations.

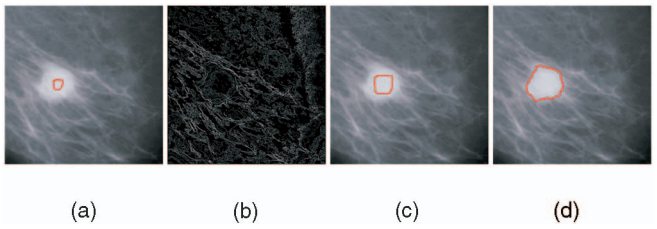

Fig. 11. The segmentation result on a 2D cyst image without gradient information.

converges to the desired edges. This difference shows that dual-front active contours can achieve better results.

Moreover, the dual front evolution is much faster than level set methods used in other active contour algorithms since fast marching or fast sweeping schemes are used in the numerical implementation of the dual front evolution technique. Since dual front evolution can handle topology changes, we can detect several objects with only one initial curve. The class of imagery that dual-front active contours can handle is not restricted to images with only two distinct means but includes images with multiple nonoverlapping regions of different means. Additionally, we do not need to know in advance the number of regions. But, such prior information is very important for Chan and Vese's method [24]. An example is shown in Fig. 10.

In Fig. 10, the segmentation is performed on a noisy synthetic image with four foreground regions having different means situated on a spatially varying background region. In this example, the potential at a point (x, y) was $\widetilde{P}(x, y) = (|I(x, y) - \mu_l| + 0.1)$, where μ_l is the mean value of points having the same label l as the point (x, y). This example illustrates automatic topology changes as well as the ability of dual-front active contours to detect different objects with different intensities and blurred boundaries. This is due to the fact that the potentials combine global dependencies and the evolving curve is automatically attracted toward the desired objects. The initial curve does not necessarily surround the objects and we use the first label assignment solution introduced in Section 2.2.

However, there are two things that need to be mentioned. The first is that, if we want to segment multiple objects in one image using only one initial contour, we must be careful to place an initial contour to enclose all objects or none of them. Otherwise, the initial contour will evolve to extract only some objects, not all objects. This property also provides a flexible way for segmenting multiple objects with different objectives. The second is that we cannot detect the inner and outer boundaries of an object with holes simultaneously just using one single initial curve. For example, in order to detect the holes of the object shown in Fig. 10a, we should set an initial seed inside the hole and another outside, and then detect the border of the hole only using the dual front evolution.

In Fig. 11, we give another example to show that dual-front active contours can be used for extracting objects without strong gradient information. Fig. 11a shows the original image with initialization, Fig. 11b shows the corresponding gradient information, Fig. 11c shows an intermediate step of the segmentation process after five iterations, and Fig. 11d shows the final segmentation after 20 iterations. In this example, the potential used was $\widetilde{P}(x, y) = (|I(x, y) - I_{mean}| + 0.1)$, and the structuring element was a 5×5 circle mask. This test shows that the segmentation of the cyst is refined even with high noise level.

5.2 Experimental Results on 3D Images

We now demonstrate dual-front active contours on several 3D images to show that this model is easy to extend to the 3D case. In Fig. 12, we present another 3D result of tumor detection in a real 3D MRI T2-weighted brain image. The 3D image size is $256 \times 256 \times 256$, the initialization is a sphere mask centered at (180, 100, 120) with size $30 \times 30 \times 30$, the structuring element is a $11 \times 11 \times 11$ sphere mask, the result is received after 20 iterations, and the potential was chosen as $\widetilde{P}(x, y, z) = (|I(x, y, z) - \mu_l| + 0.1)$, where μ_l is the mean value of points having the same label l as the point (x, y).

In recent years, various active contour models have been proposed for segmenting complex brain cortical surfaces. These methods either used active contour models as the final step for cortex segmentation or applied geometric and anatomical constraints and/or utilized significant preprocessing of original data sets to obtain desirable final segmentations.

Davatzikos and Bryan [56] used the homogeneity of intensity levels within the gray matter region to introduce a force that would drive a deformable surface toward the center of the gray matter. Teo et al. [57] proposed a four-step segmentation method which includes WM/CSF segmentation, desired WM component selection, verification/corrections, and GM representation. MacDonald et al. [58] proposed an iterative algorithm for simultaneous deformation of multiple surfaces with intersurface proximity constraints and self-intersection avoidance. Xu et al. [59] proposed a method using an external force model, called gradient vector flow [60], for cortex surface deformation. Zeng et al. [61] used the fact that the cortical layer has a nearly constant thickness to design a coupled surface model, in which two embedded surfaces evolve simultaneously, each driven by its own image-dependent forces so long as the intersurface distance remained within a predefined range. Goldenberg et al. [62] proposed a similar coupled surfaces principle and developed a model using a variational geometric framework. In their method, the surface propagation equations are derived from a minimization problem and the implementation is based on a fast geodesic active contours approach [63] for surface evolution that yields a geometrically consistent technique for improving the computation speed.

We want to clarify that the dual-front active contour model is totally different with dual snakes [16], [17], [18] or coupled surfaces [58], [61], [62]. Dual snakes or coupled surfaces were proposed to evolve coupled curves together to find two different contours simultaneously and some constraints between coupled curves (or snakes) are used to guide each curve's evolution. But, the dual front evolution is designed to find a single potential weighted minimal partition curve

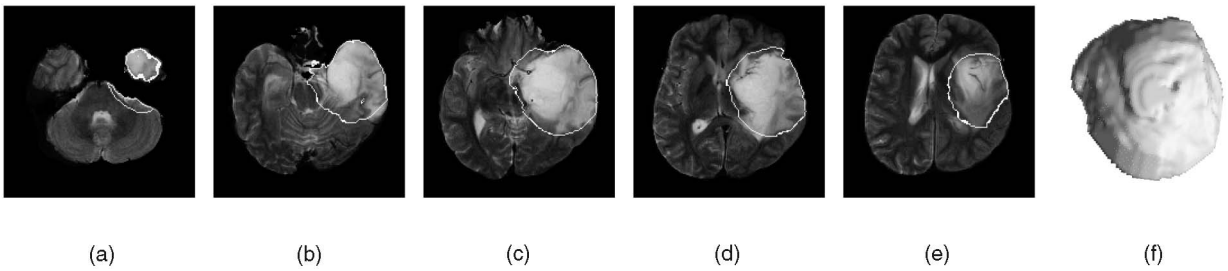

Fig. 12. The segmentation result on a 3D MRI T2-weighted brain tumor image. (a), (b), (c), (d), and (e) The segmentation results on five different slices. (f) The corresponding 3D segmentation result.

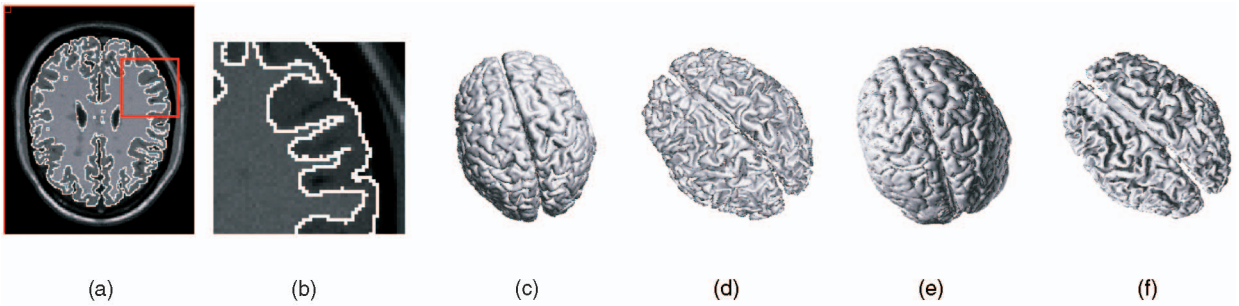

Fig. 13. Comparison of 3D models of the outer and inner cortical surfaces from our method, and the ground truth data.

within an active region, which is formed by the meeting points of the dual evolving curves. By iteratively forming a new narrow active region based on current partition curve and then using the dual front evolution to find a new global partition curve within the narrow active region, dual-front active contours can find the boundary of the single desired object. Furthermore, the principle of "dual front evolution" can be easily extended to multiple front evolution. Generally, any number of independent initial contours may be used to initialize the same number of action maps, each action map is defined by potentially different potentials and assigned a different label. Whenever the level sets of two (or more) active maps meet each other, both (or more) curve evolutions stop at the point of contact and determine the new boundary automatically. The whole process stops when each point in an active region is assigned a final label.

Finally, to demonstrate the ability of dual-front active contours to handle complex structures and topology changes, we test this model on a simulated MRI 3D brain image data set to extract the complex boundaries between three tissues: GM (gray matter), WM (white matter), and CSF (cerebral spinal fluid). After skull stripping and nonbrain tissue removal, we confine ourselves to the remaining brain region and use a dual-front active contour to capture the CSF boundary. Then, we use a second dual-front active contour to capture the WM/GM boundary. In each stage of this process, we use the second label assignment method introduced in Section 2.2 for the dual front evolution.

In this experiment, we use $T1$-weighted images because they provide better GM/WM contrast. The experimental result is obtained by processing the original 3D volume directly. The test image is available from BrainWeb [64] and was generated from the MS Lesion brain database using $T1$ modality, $1mm$ slice thickness, 3 percent noise level and 20 percent intensity nonuniformity setting. The image size is $181 \times 217 \times 217$. The initialization for the hierarchical segmentation is a sphere mask centered at $(100, 100, 95)$ with size $75 \times 75 \times 150$. The potential used was $\widetilde{P}(x, y, z) = (|I(x, y, z) - \mu_l| + 0.1)$, where μ_l is the mean value of points having the same label l as the point (x, y). The structuring element was a $5 \times 5 \times 5$ sphere mask. In Figs. 13a and 13b, we present the segmented outer (CSF-GM interface) and inner (GM-WM interface) cortical surfaces in one slice of the 3D simulated brain image and a zoom-in of the extracted boundaries for this slice. We also compare 3D models of the outer and inner cortical surfaces from our method on the same 3D image, and the ground truth data provided by BrainWeb. Figs. 13c and 13d show 3D models of the outer and inner cortical surfaces obtained from our method while Fig. 13e and 13f show 3D models using the corresponding ground truth data.

The experiments in this section were chosen primarily for their ability to illustrate a number of properties of the dual-front active contour model: robustness to both local and global image artifacts, topology changes, ease of initialization, ability to capture complex structures, etc. As is well known, the segmentation of most medical images, especially the complex 3D brain cortex, remains a challenging problem because of the variety and complexity of anatomical structures. As such, the best results are typically obtained on an application dependent basis by using additional preprocessing, postprocessing, and by applying sensible constraints (and, in some cases, by using known shape priors). In most of these applications, active contours only constitute a part of the entire algorithm. Further, the parameters used for this particular part of the algorithm also vary from application to application if one wishes to obtain the best possible results. In Section 3.1, we gave some basic principles about how to choose appropriate potentials and active regions, as it is not advisable to use any single fixed choice to process all classes of imagery.

One of our ongoing research projects focuses on the quantitative analysis of 3D brain cortex segmentation using a more complete, application-specific algorithm based on dual-front active contours. We hope to report the results of this work in the near future after extensive quantitative

comparison with many other state-of-the-art brain segmentation models.

6 Conclusions and Future Work

In this paper, we have presented a novel dual-front active contour model for 2D and 3D image segmentation to iteratively evolve an initial contour toward a desired boundary via alternating dual front evolution and morphological dilation. A key attraction of dual-front active contours is their ability to gracefully move from capturing minima that are more local to minima that are more global in nature, making it much easier to obtain "desirable" minimizers (which often are neither the most local nor the most global). Furthermore, this model combines advantages of level-set methods and fast marching methods, avoids some of their disadvantages, and has low-computational complexity $\mathcal{O}(N)$. It can detect object contours with or without gradients, and is easy to extend to 3D case. Comparison with other active contour models and segmentation results on various 2D and 3D real images illustrate that this novel model is a fast yet powerful technique for unsupervised image segmentation.

The fact that the dual-front approach may be customed tailored to capture minimizers that are flexible in their degrees of localness and globalness allows us to construct around this basic building block an algorithm that may be controlled and adapted in ways that other active contour models cannot. This key point greatly extends the usefulness of this model to many important applications in computer vision, especially medical imaging, where user control and interaction is highly desirable.

Future research work on dual-front active contours includes combining this model with other powerful image smoothing, denoising methods or other image segmentation methods for more challenging segmentation objectives in medical imaging, searching for new methods of defining more appropriate active regions for improving the accuracy of the segmentation results, and working on quantitative analysis of 3D medical image segmentation.

Acknowledgments

The authors would like to thank the anonymous reviewers for valuable comments and suggestions. They would also like to thank Dr. Sebastien Fourey and Dr. Regis Clouard of GREYC-ENSICAEN, France for providing MVox 3D model visualization and PANDORE image processing platform, Professor Laurent Cohen of CEREMADE, Universite Paris Dauphine, and Ganesh Sundaramoorthi of School of ECE, Georgia Institute of Technology for their helpful suggestions which have improved the quality of this paper. Support for this research was made possible from NSF grant CCR-0133736 and NIH/NINDS grant R01NS037747.

References

[1] X. Munoz et al., "Strategies for Image Segmentation Combining Region and Boundary Information," *Pattern Recognition Letters,* vol. 24, no. 1, pp. 375-392, 2003.

[2] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active Contour Models," *Int'l J. Computer Vision,* vol. 1, no. 4, pp. 321-332, 1988.

[3] V. Caselles et al., "A Geometric Model for Active Contours in Image Processing," *Numerical Math.,* vol. 66, no. 1, pp. 1-31, 1993.

[4] R. Malladi, J. Sethian, and B. Vemuri, "Shape Modeling with Front Propagation: A Level Set Approach," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 17, no. 1, pp. 158-175, Jan. 1995.

[5] S. Osher and J. Sethian, "Fronts Propagating with Curvature Dependent Speed: Algorithms Based on the Hamilton-Jacobi Formulation," *J. Computational Physics,* vol. 79, no. 1, pp. 12-49, 1988.

[6] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic Active Contours," *Int'l J. Computer Vision,* vol. 22, no. 1, pp. 61-79, 1997.

[7] A. Yezzi, S. Kichenassamy, A. Kumar, P. Olver, and A. Tannenbaum, "A Geometric Snake Model for Segmentation of Medical Imagery," *IEEE Trans. Medical Imaging,* vol. 16, no. 2, pp. 199-209, 1997.

[8] H. Tek and B. Kimia, "Image Segmentation by Reaction Diffusion Bubbles," *Proc. Int'l Conf. Computer Vision,* pp. 156-162, 1995.

[9] L. Cohen and I. Cohen, "Finite-Element Methods for Active Contour Models and Balloons for 2D and 3D Images," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 15, no. 11, pp. 1131-1147, Nov. 1993.

[10] L. Cohen and R. Kimmel, "Global Minimum for Active Contour Models: A Minimal Path Approach," *Proc. IEEE Int'l Conf. Computer Vision and Pattern Recognition,* pp. 666-673, 1996.

[11] L. Cohen and R. Kimmel, "Global Minimum for Active Contour Models: A Minimal Path Approach," *Int'l J. Computer Vision,* vol. 24, no. 1, pp. 57-78, 1997.

[12] L. Cohen, "Multiple Contour Finding and Perceptual Grouping Using Minimal Paths," *J. Math. Imaging and Vision,* vol. 14, no. 3, pp. 225-236, 2001.

[13] T. Deschamps and L. Cohen, "Fast Extraction of Minimal Paths in 3D Images and Applications to Virtual Endoscopy," *Medical Image Analysis,* vol. 5, no. 4, pp. 281-299, 2001.

[14] R. Ardon and L. Cohen, "Fast Constrained Surface Extraction by Minimal Paths," *Proc. Int'l Conf. Computer Vision-Workshop Variational, Geometric, and Level Set Methods,* pp. 233-244, 2003.

[15] S. Gunn and M. Nixon, "A Model Based Dual Active Contour," *Proc. British Machine Vision Conf.,* pp. 305-314, 1994.

[16] S. Gunn and M. Nixon, "A Robust Snake Implementation: A Dual Active Contour," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 19, no. 1, pp. 63-68, Jan. 1997.

[17] G. Giraldi, L. Goncalves, and A. Oliveira, "Dual Topologically Adaptable Snakes," *Proc. Int'l Conf. Computer Vision, Pattern Recognition, and Image Processing,* pp. 103-107, Feb. 2000.

[18] G. Georgoulas, G. Nikolakopoulos, Y. Koutroulis, A. Tzes, and P. Groumpos, "An Intelligent Visual-Based System for Object Inspection and Welding, Relying on Active Contour Models-Algorithms," *Proc. Second Hellenic Conf. Artificial Intelligence,* pp. 399-410, Apr. 2002.

[19] G. Aboutanos, J. Nikanne, N. Watkins, and B. Dawant, "Model Creation and Deformation for the Automatic Segmentation of the Brain in MR Images," *IEEE Trans. Biomedical Eng.,* vol. 46, no. 11, pp. 1346-1356, 1999.

[20] C. Erdem, A. Tekalp, and B. Sankur, "Video Object Tracking with Feedback of Performance Measures," *IEEE Trans. Circuits and Systems for Video Technology,* vol. 13, no. 4, pp. 310-324, 2003.

[21] M. Dawood, X. Jiang, and K. Schafers, "Reliable Dual-Band Based Contour Detection: A Double Dynamic Programming Based Approach," *Lecture Notes in Computer Science,* vol. 3212, pp. 544-551, 2004.

[22] N. Xu, R. Bansal, and N. Ahuja, "Object Segmentation Using Graph Cuts Based Active Contours," *Proc. IEEE Int'l Conf. Computer Vision and Pattern Recognition,* vol. 2, pp. 46-53, 2003.

[23] A. Chakraborty, L. Staib, and J. Duncan, "Deformable Boundary Finding in Medical Images by Integrating Gradient and Region Information," *IEEE Trans. Medical Imaging,* vol. 15, no. 6, pp. 859-870, 1996.

[24] T. Chan and L. Vese, "Active Contours without Edges," *IEEE Trans. Image Processing,* vol. 10, no. 2, pp. 266-277, 2001.

[25] N. Paragios and R. Deriche, "Geodesic Active Regions: A New Framework to Deal with Frame Partition Problems in Computer Vision," *J. Visual Comm. and Image Presentation,* vol. 13, no. 1, pp. 249-268, 2002.

[26] C. Samson et al., "A Level Set Method for Image Classification," *Proc. Int'l Conf. Scale-Space Theories in Computer Vision,* pp. 306-317, 1999.

[27] A. Yezzi, A. Tsai, and A. Willsky, "A Statistical Approach to Image Segmentation for Bimodal and Trimodal Imagery," *Proc. Int'l Conf. Computer Vision,* vol. 2, pp. 1-5, 1999.

[28] A. Tsai, A. Yezzi, and A. Willsky, "Curve Evolution Implementation of the Mumford-Shah Functional for Image Segmentation, Denoising, Interpolation, and Magnification," *IEEE Trans. Image Processing,* vol. 10, no. 8, pp. 1169-1186, 2001.

[29] A. Yezzi, A. Tsai, and A. Willsky, "A Fully Global Approach to Image Segmentation via Coupled Curve Evolution Equations," *J. Visual Comm. Image Representation,* vol. 13, no. 1, pp. 195-216, 2002.

[30] S. Jehan-Besson, M. Barlaud, and G. Aubert, "DREAM^{2}s: Deformable Regions Driven by an Eulerian Accurate Minimization Method for Image and Video Segmentation," *Int'l J. Computer Vision,* vol. 53, no. 1, pp. 45-70, 2003.

[31] S. Zhu and A. Yuille, "Region Competition: Unifying Snakes, Region Growing, and Bayes/MDL for Multiband Image Segmentation," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 18, no. 9, pp. 884-900, Sept. 1996.

[32] D. Adalsteinsson and J. Sethian, "A Fast Level Set Method for Propagating Interfaces," *J. Computational Physics,* vol. 118, no. 2, pp. 269-277, 1995.

[33] J. Helmsen, E. Puckett, P. Colella, and M. Dorr, "Two New Methods for Simulating Photolithography Development in 3D," *Proc. SPIE,* vol. 2726, pp. 253-261, 1996.

[34] J. Tsitsiklis, "Efficient Algorithms for Globally Optimal Trajectories," *IEEE Trans. Automatic Control,* vol. 40, no. 9, pp. 1528-1538, 1995.

[35] J. Sethian, "A Fast Marching Level Set Method for Monotonically Advancing Fronts," *Proc. Nat'l Academy of Science of USA,* vol. 93, no. 4, pp. 1591-1595, 1996.

[36] C. Bajaj, Z. Yu, and M. Auer, "Volumetric Feature Extraction and Visualization of Tomographic Molecular Imaging," *J. Structural Biology,* vol. 144, no. 1-2, pp. 132-143, 2003.

[37] M. Boué and P. Dupuis, "Markov Chain Approximations for Deterministic Control Problems with Affine Dynamics and Quadratic Cost in the Control," *SIAM J. Numerical Analysis,* vol. 36, no. 3, pp. 667-695, 1999.

[38] H. Zhao, "Fast Sweeping Method for Eikonal Equations," *Math. Computation,* vol. 74, pp. 603-627, 2004.

[39] E. Rouy and A. Tourin, "A Viscosity Solutions Approach to Shape-from-Shading," *SIAM J. Numerical Analysis,* vol. 29, no. 3, pp. 867-884, 1992.

[40] P. Dupuis and J. Oliensis, "An Optimal Control Formulation and Related Numerical Methods for a Problem in Shape Reconstruction," *Annals of Applied Probability,* vol. 4, no. 2, pp. 287-346, 1994.

[41] H. Li, A. Elmoataz, J. Fadili, and S. Ruan, "Dual Front Evolution Model and Its Application in Medical Imaging," *Proc. Int'l Conf. Medical Image Computing and Computer-Assisted Intervention,* pp. 103-110, Sept. 2004.

[42] H. Li, A. Yezzi, and L. Cohen, "Fast 3D Brain Segmentation Using Dual-Front Active Contours with Optional User-Interaction," *Proc. Int'l Conf. Computer Vision Workshop-Computer Vision for Biomedical Image Applications,* pp. 335-345, Oct. 212005.

[43] L. Yatziv, A. Bartesaghi, and G. Sapiro, "O(N) Implementation of the Fast Marching Algorithm," *J. Computational Physics,* vol. 212, no. 2, pp. 393-399, 2006.

[44] P. Danielsson, "Euclidean Distance Mapping," *Computer Graphics and Image Processing,* vol. 14, pp. 227-248, 1980.

[45] Y. Tsai, "Rapid and Accurate Computation of the Distance Function Using Grids," *J. Computational Physics,* vol. 178, no. 1, pp. 175-195, 2002.

[46] Y. Tsai, L. Cheng, S. Osher, and H. Zhao, "Fast Sweeping Algorithms for a Class of Hamilton-Jacobi Equations," *SIAM J. Numerical Analysis,* vol. 41, no. 2, pp. 673-694, 2003.

[47] C. Kao, S. Osher, and J. Qian, "Lax-Friedrichs Sweeping Scheme for Static Hamilton-Jacobi Equations," *J. Computational Physics,* vol. 196, no. 1, pp. 367-391, 2004.

[48] F. Gibou and R. Fedkiw, "A Fast Hybrid *k*-Means Level Set Algorithm for Segmentation," *Proc. Fourth Ann. Hawaii Int'l Conf. Statistics and Math.,* pp. 281-291, Jan. 2005.

[49] P. Perona and J. Malik, "Scale-Space and Edge Detection Using Anisotropic Diffusion," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 12, no. 7, pp. 629-639, July 1990.

[50] R. Malladi and J. Sethian, "A Unified Approach to Noise Removal, Image Enhancement, and Shape Recovery," *IEEE Trans. Image Processing,* vol. 5, no. 11, pp. 1554-1568, 1996.

[51] T. Deschamps, "Curve and Shape Extraction with Minimal Path and Level-Sets Techniques: Applications to 3D Medical Imaging," PhD dissertation, Universite de Paris-Dauphine, Dec. 2001.

[52] L. Vincent and P. Soille, "Watersheds in Digital Spaces: An Efficient Algorithm Based on Immersion Simulation," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 13, no. 6, pp. 583-598, June 1991.

[53] E. Sifakis and G. Tziritas, "Moving Object Localization Using a Multi-Label Fast Marching Algorithm," *Signal Processing: Image Comm.,* vol. 16, no. 10, pp. 963-976, 2001.

[54] D. Mumford and J. Shah, "Optimal Approximation by Piecewise Smooth Functions and Associated Variational Problems," *Comm. Pure Applied Math.,* vol. 42, pp. 577-685, 1989.

[55] R. Ronfard, "Region-Based Strategies for Active Contour Models," *Int'l J. Computer Vision,* vol. 3, no. 2, pp. 229-251, 1994.

[56] C. Davatzikos and R. Bryan, "Using a Deformable Surface Model to Obtain a Shape Representation of the Cortex," *IEEE Trans. Medical Imaging,* vol. 15, no. 6, pp. 785-795, 1996.

[57] P. Teo, G. Sapiro, and B. Wandell, "Creating Connected Representations of Cortical Gray Matter for Functional MRI Visualization," *IEEE Trans. Medical Imaging,* vol. 16, no. 6, pp. 852-863, 1997.

[58] D. MacDonald, D. Avis, and A. Evans, "Proximity Constraints in Deformable Models for Cortical Surface Identification," *Proc. Int'l Conf. Medical Image Computing and Computer-Assisted Intervention,* pp. 650-659, 1998.

[59] C. Xu, D. Pham, J. Prince, M. Etemad, and D. Yu, "Reconstruction of the Central Layer of the Human Cerebral Cortex from MR Images," *Proc. Int'l Conf. Medical Image Computing and Computer-Assisted Intervention,* pp. 481-488, 1998.

[60] C. Xu and J. Prince, "Snakes, Shapes, and Gradient Vector Flow," *IEEE Trans. Image Processing,* vol. 7, no. 3, pp. 359-369, 1998.

[61] X. Zeng, L. Staib, R. Schultz, and J. Duncan, "Segmentation and Measurement of the Cortex from 3D MR Images Using Coupled Surfaces Propagation," *IEEE Trans. Medical Imaging,* vol. 18, no. 10, pp. 100-111, 1999.

[62] R. Goldenberg, R. Kimmel, E. Rivlin, and M. Rudzsky, "Cortex Segmentation: A Fast Variational Geometric Approach," *IEEE Trans. Medical Imaging,* vol. 21, no. 2, pp. 1544-1551, 2002.

[63] R. Goldenberg, R. Kimmel, E. Rivlin, and M. Rudzsky, "Fast Geodesic Active Contours," *IEEE Trans. Image Processing,* vol. 10, no. 10, pp. 1467-1475, 2001.

[64] C. Cocosco, V. Kollokian, R. Kwan, and A. Evans, "Brainweb: Online Interface to a 3D MRI Simulated Brain Database," *NeuroImage,* vol. 5, no. 4, part 2/4, S425, http://www.bic.mni.mcgill.ca/brainweb/, 1997.

Hua Li received the PhD degree in 2001 through the Department of Electronics and Information Engineering at Huazhong University of Science and Technology, China. Currently, she is a postdoctoral research fellow at the School of Electrical and Computer Engineering, Georgia Institute of Technology. Her research interests are in the fields of image processing and computer vision, in particular PDEs, active contour models, curve and surface evolution theory, and mathematical morphology for 3D image segmentation and analysis with applications to medical imaging. She is a member of the IEEE.

Anthony Yezzi received the PhD degree in 1997 through the Department of Electrical Engineering at the University of Minnesota. After completing a postdoctoral research position in the Laboratory for Information and Decision Systems (LIDS) at Massachusetts Institute of Technology, he joined the faculty of the School of Electrical and Computer Engineering at Georgia Institute of Technology in 1999, where he currently holds the position of associate professor. Professor Yezzi has also consulted for a number of medical imaging companies including GE, Picker, and VTI, and has been a senior member of the IEEE since 1999. His research lies primarily within the fields of image processing and computer vision. He has worked on a variety of problems including image denoising, edge-detection, segmentation and grouping, shape analysis, multiframe stereo reconstruction, tracking, and registration. Some central themes of his research include curve and surface evolution theory, differential geometry, and partial differential equations.