LOCALIZED STATISTICAL MODELS IN COMPUTER VISION

A Thesis
Presented to
The Academic Faculty

by

Shawn M. Lankton

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Electrical and Computer Engineering

Georgia Institute of Technology
December, 2009

LOCALIZED STATISTICAL MODELS IN COMPUTER VISION

Approved by:

Professor Anthony Yezzi,
Committee Chair
School of Electrical and Computer
Engineering
Georgia Institute of Technology

Professor Allen Tannenbaum, Advisor
School of Electrical and Computer
Engineering
Georgia Institute of Technology

Professor Jeff Shamma
School of Electrical and Computer
Engineering
Georgia Institute of Technology

Professor Ghassan Al Regib
School of Electrical and Computer
Engineering
Georgia Institute of Technology

Professor Marc Niethammer
School of Electrical and Computer
Engineering
Georgia Institute of Technology

Professor Arthur Stillman
School of Electrical and Computer
Engineering
Georgia Institute of Technology

Date Approved: September 8, 2009

To my parents, Steve and Carol.

ACKNOWLEDGEMENTS

During my time as a Ph.D. student at the Georgia Institute of Technology I have had the fortune of knowing and working with amazing people. Your help, friendship, and support has made my time here truly incredible. In particular, I would like to thank:

My Advisor. Professor Allen Tannenbaum, your wisdom, humor, and kindness are truly inspiring. Thank you for your guidance and friendship over the years. You have taught me about life, business, science, friendship, and of course, computer vision. It has been an absolute honor and a pleasure to be your student.

My Thesis Committee. Professor Anthony Yezzi, who I've always looked to as a model of excellence in academics and teaching, professor Arthur Stillman, who helped direct my research toward vessel analysis, professor Marc Niethammer, who would be listed in the lab mentors section if he wasn't also a committee member, and professors Jeff Shamma, and Ghassan Al Regib whose remarks helped guide and improve my thesis.

Mentors in the Lab. Delphine Nain, Yogesh Rathi, and Eric Pichon. I'm honored to have had you as friends and lab-mates and to have learned from you during that time. I have since measured myself against each of you as models for outstanding academic and professional success.

Other Collaborators. Gozde Unal, Stephen Sprigle, Sharon Sonnenblum, Paolo Raggi, Marie Thursby, and Margi Berberi, thank you for your guidance, support, and encouragement as you helped me pursue many varied interests.

TABLE OF CONTENTS

LIST OF FIGURES

SUMMARY

Computer vision attempts to artificially approximate the capabilities of human vision using computers. Some of the many applications it has found include automation, robotics, visual tracking, and medical image processing. The research presented in this thesis is primarily in the fields of image segmentation and visual tracking. Image segmentation is a subset of computer vision that attempts to partition an image into logical parts that correspond to objects or regions of interest. Visual tracking takes that process a step further and analyzes objects of interest as they change in time-varying image sequences.

The work contained in this document explores the ramifications of one key hypothesis; by localizing the analysis of visual information, assumptions about image-makeup and object-appearances can be relaxed thereby improving the accuracy of both segmentation and tracking results.

Through investigating this assumption, a novel framework was created that allows well-known image segmentation techniques to be re-formulated with a scale parameter that controls the degree to which statistical analysis will be localized. Experiments demonstrate that the framework is useful for segmenting images that are too challenging for previously existing techniques. Accordingly, the behaviors of localized segmentation methods are studied in depth to fully understand their strengths and limitations.

Next, two important medical applications of the proposed framework are considered: segmentation of neuron bundles in the brain and analysis of vessels and coronary plaques in the heart. Local analysis is well suited to the study of these long, tubular, structures because of their fine scale, varying appearance, and statistical similarity to the surrounding image. Neuron bundle segmentation is accomplished using a novel energy created with the presented framework that is capable of operating on diffusion tensor images (DTI).

Coronary vessel analysis begins with segmenting the vessels from computed tomography angiography (CTA) imagery. Next, analysis of local regions is shown to be useful for detecting and measuring non-calcified plaques (also known as soft plaques), which are considered to be a major risk factor for cardiovascular disease. Experiments confirm that by using a combination of localized methods it is possible to detect and segment these potentially dangerous plaques despite the subtle differences that separate them from surrounding healthy tissue.

Finally, a visual tracking algorithm is proposed which makes use of the ideas of localization temporally rather than spatially. The presented tracking method uses localized time-windows to maintain appearance models of the tracked object even as its appearance changes throughout a time-varying image sequence. Using this technique it is possible to track objects that change drastically in scale, shape, and visual characteristics over time.

The research in this thesis confirms that localized analysis of images can produce better results than global analysis in many cases. Throughout this thesis experiments demonstrate the utility and effectiveness of localization in a wide range of applications. For completeness, the presented techniques are explained in detail so that readers can use and build upon the ideas presented herein.

CHAPTER I

INTRODUCTION
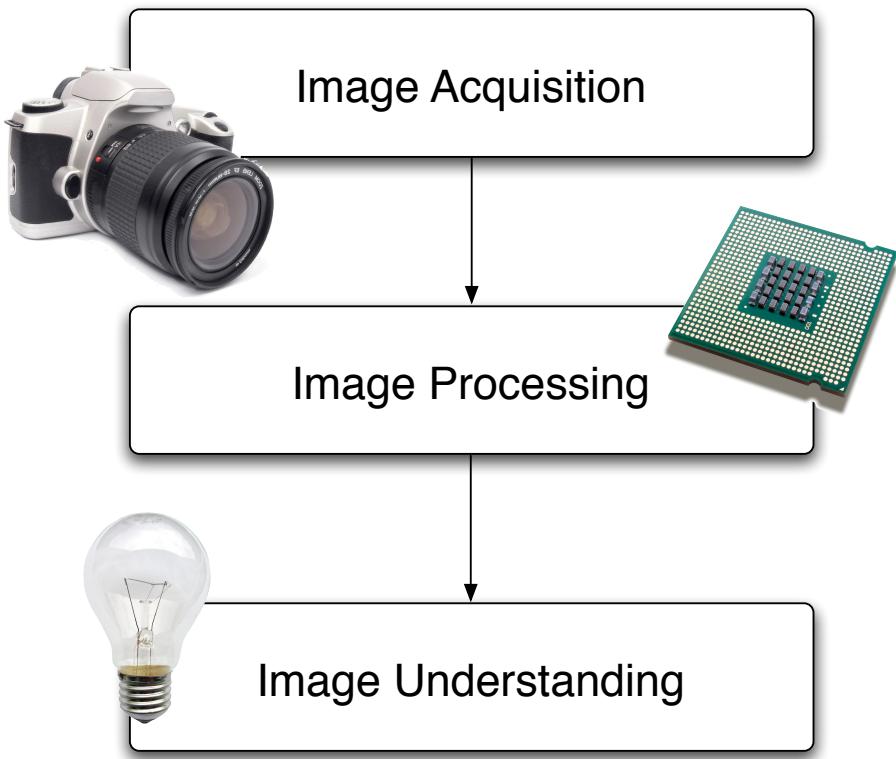
This chapter introduces the field of computer vision and the specific topics of segmentation, medical image processing, and visual tracking. This introduction will allow the reader to understand the contributions of this thesis, which are described in broad terms at the end of the chapter. Finally, the organization of the thesis is presented giving the reader a context for understanding and navigating the included research.

1.1 Computer Vision

Vision is perhaps the most important sense used by humans to understand the world around them. By analyzing light reflected from the world around us, our brains are able to draw conclusions about our surroundings and help us function in our ever-changing environment. The field of computer vision emerged in the 1970s with the goal of approximating the abilities of human vision and perception artificially [30]. More specifically, computer vision includes both the ability to acquire images and the ability to analyze and make sense of them.

Although the goal of computer vision is similar to that of biological vision the methodology is quite different. In biological vision our brains interpret signals from our eyes using neural pathways adapted over millions of years and learning acquired over a lifetime of experience [15]. In computer vision, images are acquired from a variety imaging devices, stored digitally, and processed by a multitude of algorithms designed for specific tasks [12, 34, 75]. By combining these algorithms in an intelligent way it is possible to create systems that simulate the high-level processing that is completed so effortlessly by the human brain.

In most cases human perception is far superior to solutions provided by computer vision. In some cases however, computer vision does out-perform human vision. For instance, computers are able to perform tasks quickly and with level of detail that would be impossible, or at least terribly meticulous, for a human. Furthermore, computer vision systems have

Figure 1.1: The three high-level operations of computer vision: Image acquisition, image processing, and image understanding.
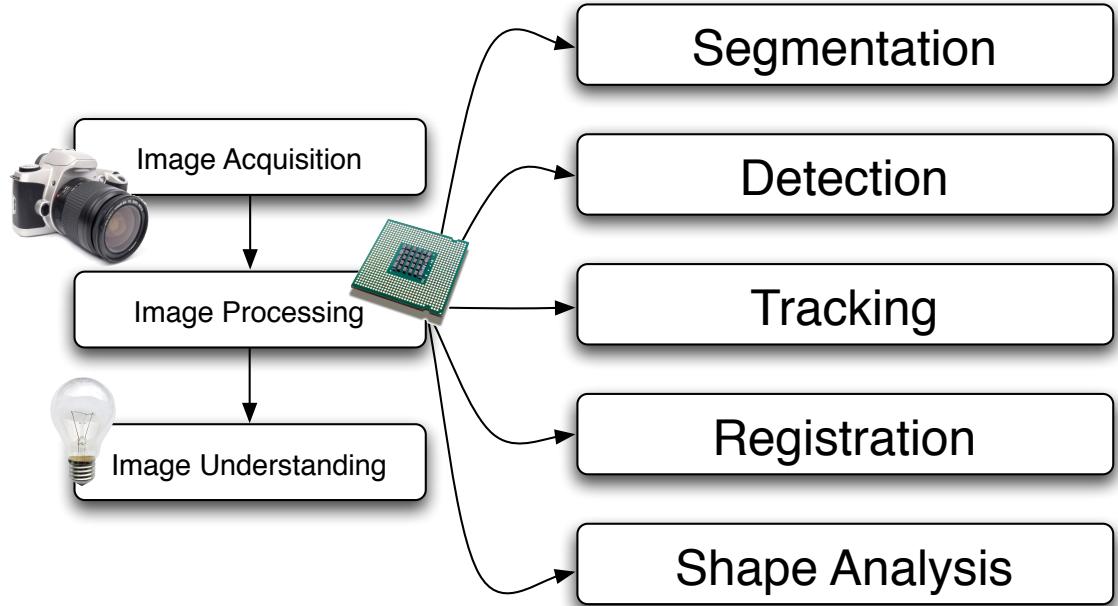
the ability to remove inter-operator variation that is often present when people perform image analysis tasks. Consider three different people analyzing the same image: They will likely return three different, although similar answers [58]. A deterministic computer vision system, on the other hand, will return the exact same answer every time. Finally, computers are well suited to analyzing image data that is extremely large, or of a dimensionality, that humans cannot readily comprehend. For example, the medical imaging systems discussed in Section 1.1.2 return three dimensional (3-D) volumetric data that is cumbersome for a human to interact with, but which a computer can easily interpret.

While human vision remains far more complex than man-made approximations, research in this field has resulted in huge leaps in technological capabilities. Computer vision is now used extensively in surveillance, robotics, defense, and medicine, just to name a few applications [3, 25, 35, 106].

Within computer vision, one can consider three high-level operations: image acquisition, image processing, and image understanding. In image acquisition, data is acquired with various sensors and then filtered until a usable image is created. Research in this area has produced a wide array of imaging modalities allowing computer vision to "see" far beyond human vision. Technologies such as 3-D laser imaging, long-range tactical imaging, and in-vivo 3-D imaging inside the human body are just a few examples. The next task, image processing, deals with analyzing the acquired images to locate, identify, and analyze salient objects or patterns. This crucial step allows the image to be represented symbolically rather than as a collection of data points representing individual picture elements (pixels). Image processing is still a topic of intense research and remains largely unsolved although significant progress has been made. The ability to derive rich and accurate symbolic representations from image data is a necessary step before image understanding can be accomplished. Image understanding bridges the gap between computer vision and artificial intelligence and is often performed by humans or is inherent in a particular application.

Image processing is an intermediate, but crucial step in computer vision, and is the main topic of discussion in this thesis. Within image processing there are five main components: segmentation, detection, tracking, registration, and shape analysis. Segmentation, which is discussed further in Section 1.1.1, consists of partitioning an image into parts that each correspond to distinct objects or regions within the image. This is often a first step in image processing because it gives important clues about the nature of the image data. Next, detection is the process of automatically finding the location of objects or regions within an image with minimal input from a user. Continuously detecting, and possibly segmenting, moving objects in time-varying images is the third goal, tracking. Tracking is introduced in Section 1.1.3 and typically involves finding an object's location or outline at each frame in a video sequence.

Registration and shape analysis are not discussed in detail in this thesis but are important tasks within computer vision. Registration is the process of aligning images or shapes automatically, and shape analysis provides a way to compare the shape of objects that have been previously segmented from an image to determine similarities and differences that

Figure 1.2: The critical tasks within image processing: Segmentation, detection, tracking, registration, and shape analysis.

may exist between them. See [29, 32, 70, 109] and the references within for information regarding registration and shape analysis.

1.1.1 Segmentation

Segmentation is the process of partitioning an image into meaningful parts. Typically, the output will be a label map that clearly marks each pixel according to the part of the image it represents. This is a necessary initial step for subsequent image processing tasks such as tracking, detection, object recognition, and shape analysis. In medical imaging in particular, segmentation is valuable on its own as discussed in Section 1.1.2.
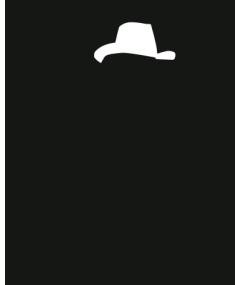
Two major challenges exist in computer-assisted segmentation: The first is defining which objects are "meaningful." This task is dependent on the particular application at hand [58]. Consider the image of a man in Figure 1.3. In some scenarios, the entire man may be the meaningful object. However, the man is made up of other objects (hat, face, jacket, hands, pants, shoes, etc.) that may be "meaningful" on their own in other scenarios. This example demonstrates the complex nature of image segmentation. In most cases the goal will be to identify the boundary of an object of interest given some initial input. The

4
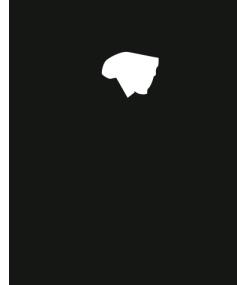
(a) Original Image

(b) Entire Man
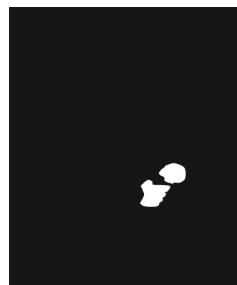
(c) Hat

(d) Face
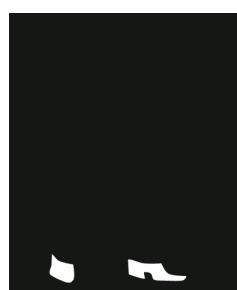
(e) Jacket

(f) Hands

(g) Pants

(h) Shoes

Figure 1.3: (a) An image of a man and multiple "correct" segmentations of the image: (b) the entire man, (c) hat only, (d) face only, (e) jacket only, (f) hands only, (g) pants only, and (f) shoes only.

input, which can be provided by a user or generated automatically, should give clues as to what constitutes "meaningful" in a particular scenario so that the segmentation may proceed.

The second major challenge concerns the fact that the object of interest may have a complex appearance that makes it difficult to distinguish from other surrounding objects or the image background. In these cases it may be difficult to devise an algorithm that is capable of automatically finding the object's boundaries. These two challenges: segmenting the correct object, and segmenting an object correctly are the basis for the majority of on-going research in the field of image segmentation.

Typical approaches to segmentation begin with observations about images and their makeup. Based on these observations, assumptions are made about the image, the object (or objects) of interest, and the type of initialization. From this a well-designed algorithm will determine a partitioning of the image domain that is consistent with these assumptions and is (hopefully) an accurate segmentation. Therefore, the design of the algorithm and the choice of appropriate assumptions are of great importance. In Chapter 2 a detailed background on the segmentation methods used in this thesis is provided. Then, in Chapters 3, 4, and 5 specific contributions are made to the field of image segmentation.

1.1.2 Medical Image Processing

Two applications of image segmentation, discussed in Chapters 3 and 4, are from the field of medical image processing. Specifically, the segmentation and detection of structures within the brain and heart are considered. Current medical imaging technology is capable of recording 3-D volumetric data about a patient's body with minimally invasive protocols. The data obtained by these scans can be paramount in understanding and diagnosing pathologies within the body. Computer vision is particularly well suited for medical image processing because manually analyzing the large 3-D data sets can be time consuming and cumbersome for doctors and radiologists. Furthermore, diagnostic speed and accuracy can be improved by producing 3-D visualizations based on segmentations and detecting certain pathologies automatically [3].

Imaging modalities discussed in this thesis, particularly those used to study the brain and heart, will now be briefly introduced. For imaging the brain, the focus will be on magnetic resonance imaging (MRI) and diffusion tensor imagery (DTI) [5]. MRI acquires images in which each volume element (voxel) has an intensity that depends on the type of tissue at the corresponding point within the body. This type of data is useful for studying the makeup of organs and has been used extensively to study almost every part of the human anatomy. In this thesis, the focus will be on MRI scans of the brain. In contrast to MRI, DTI makes multiple measurements at each voxel where each measurement records the amount of water diffusion along a particular direction. DTI data is therefore most useful for analyzing fine structures in the body that restrict the diffusion of water non-uniformly [67]. This thesis will examine a particular application: segmenting groups of neuron fibers that are bundled together within the brain.

For analysis of the heart, the focus will be on computed tomography (CT) imagery, which produces 3-D volumes of data where voxel intensity is based on the X-ray attenuation of the corresponding tissue in the body. In particular, the focus will be on CT imagery taken with a CT angiography (CTA) protocol. In CTA, a contrast agent is introduced intravenously to increase the brightness of the blood in the resulting CT images. Whereas MRI and DTI are non-invasive procedures, CTA is considered minimally invasive due to the introduction of the contrast agent and the radiation dose associated with CT imagery.

1.1.3 Visual Tracking

Visual tracking is the process of finding the position of an object of interest at each frame within a time-varying sequence. In many cases it is also desirable to find the shape or pose of the object. Hence, one can consider tracking to be the process of locating an object continuously as it changes over time. Just as segmentation is a necessary first step for image processing tasks, tracking is a necessary first step for video processing tasks. Tracking lends insight into behavior patterns and object interactions within a video sequence.

The same problems that make segmentation difficult also exist within tracking. Defining what constitutes the "object of interest" and compensating for ambiguous, incomplete,

or misleading image information are frequent obstacles. In addition, new challenges can emerge such as handling occlusions of the object of interest by other objects in the scene, rapid or erratic object movement, and changes in the objects size or appearance. These challenges often require various filtering and prediction schemes such as the Kalman filter [40] or particle filter [26]. These filtering schemes utilize information about the object's dynamics that are learned a priori or collected on-line from previously tracked frames to aid in subsequent tracking.

The final challenge of visual tracking systems is that it is often desirable for them to be used in real-time systems. This means that any tracking system that will be used in real-world applications must be computationally efficient enough so that it can process incoming video frames as they arrive. This typically limits allowable processing time to at most one tenth of a second and often much less.

Real-time tracking is used in robotics, surveillance, and military applications to help locate objects as they move about in the world. Once located, a system can use the output of a visual tracking algorithm to "close the loop" on a controller that is used to navigate toward or away from the object of interest. Specific applications can include tracking missiles [33, 100], unmanned aerial vehicles [10, 31, 37], people [35], and cars [106]. This thesis describes a contribution to the field of visual tracking in Chapter 6.

1.2 Contributions and Organization of this Thesis

In this thesis, four major contributions are presented that fall into the fields of segmentation, detection, and tracking as described above.

1. **A novel framework for image segmentation that uses localized statistical models to drive an active contour segmentation toward object boundaries.** The presented framework extends the ideas of active contours presented in Chapter 2 by allowing any active contour segmentation to be re-formulated in a localized manner so that assumptions placed on the image data can be relaxed. This reduces the statistical complexity needed to accurately model the image and leads to significant improvement of segmentation performance in most cases. Several localized segmentation methods are studied in depth to understand the role of localization, the benefits of this method, and potential drawbacks. Results are presented across a wide variety of challenging natural and medical images.

2. **An application of localized segmentations to analyze connectivity in the brain by providing volumetric quantification of neural pathways.** It is shown that by using the proposed localization framework, pathways in the brain responsible for connecting different functional areas can be segmented from DTI data using simple initializations. The specific contribution is a novel localized segmentation method adapted to operate on tensor information in DTI data, which contains significantly more information than standard scalar-valued images. Results from this technique are shown on an important and challenging neural pathway known as the cingulum bundle.

3. **A set of novel algorithms to analyze CTA data, segment coronary vessels, and detect potentially dangerous plaques automatically.** Specifically, localized segmentation is applied to robustly extract coronary vessels from CTA data. Then by performing a coupled segmentation of the interior and exterior of the vessel walls using a different localized segmentation protocol coronary plaques are identified. Vessel segmentation and plaque detection represent the major challenges in computational

9

cardiology and may help to improve the detection and treatment of cardiovascular disease (CVD). Results are shown on numerous CTA datasets that have been validated by human experts in order to demonstrate the reliability and validity of the algorithms.

4. **A novel visual tracking system that uses temporal localization to adapt to objects that change in size or appearance throughout a video sequence.** By applying the ideas of localization temporally rather than spatially, a system is created that is capable of tracking challenging sequences in which an object's scale within the image changes quickly. Specifically, the system employs a fast, two-part, template-matching algorithm coupled with periodic template updates to ensure that the position of the object is found at each frame. Additionally, the algorithm delivers real-time results even if objects grow to fill the field of view. The tracking system is demonstrated on several sequences that show large changes in the objects' size, shape, and appearance.

The contributions are organized as follows:

- **Chapter 2:** Gives definitions and a detailed description of active contour image segmentation techniques and level set methods. This provides a mathematical foundation on which the remainder of the thesis will be based.

- **Chapter 3:** Presents the localized active contour framework and describes the behavior of this new class of active contours. The following chapters make use of this framework as they explore potential applications for localized active contours in medical image processing [47].

- **Chapter 4:** Describes how localized contours can be used for segmenting neuron bundles in DTI data. New localized segmentation energies are given as well as experiments validating the technique [48].

- **Chapter 5:** Shows how localized segmentation can be applied to coronary vessels analysis in CTA imagery. This chapter includes methods to automatically segment vessels and detect potentially dangerous plaques. The specific algorithms are described in detail and results are compared to those produced by human experts [51].

- **Chapter 6:** Presents a visual tracking algorithm based on applying the ideas of localization temporally to track objects with rapidly changing appearances. This chapter gives a short background on tracking, describes the proposed tracking system, and shows experiments that demonstrate the capabilities of the tracker [50].

- **Chapter 7:** Offers a summary of the presented work, draws conclusions from the thesis, and discusses possible directions for future work.

- **Appendix A and B:** Give details on derivations and numerical implementations used within the thesis.

CHAPTER II

BACKGROUND

Segmentation as described in Chapter 1 can be performed in a number of ways. What follows provides a mathematical background for the class of segmentation algorithms known as *active contours* and a particular implementation strategy known as *level set methods*. This foundation will facilitate the reader's understanding of the research and contributions put forth in the following chapters.

2.1 Variational Active Contours

Variational methods seek to solve complex problems via optimization. In order to solve a problem with a variational approach, one must define a quantity or "energy" that will be at an optimum (maxima or minima) when a desirable solution is reached. Typically an energy is minimized by finding its first variation with respect to key arguments and iteratively reducing the energy by manipulating those arguments through gradient descent [65].

Many computer vision techniques can be formulated as optimization problems [3], and doing so yields several benefits. For instance, a problem that is described variationally can be understood by analyzing the underlying energy functional. This compact, continuous, mathematical representation is simple to analyze and easy to amend if needed. Additionally, any valid energy minimization scheme will produce a similar final result thus decoupling the choice of energy functional from any specific numerical implementation. This means that the key to a variational solution is the energy itself, not the minutiae and heuristics related to the implementation.

Active contours are one class of variational methods that have been used extensively within computer vision to address the problem of image segmentation and tracking [12, 76]. In active contour segmentation, an energy functional is defined whose arguments include the image in question and a closed curve that partitions the image. The energy is then

minimized by iteratively deforming the contour via gradient descent. If the energy is well-chosen the curve will come to rest when a meaningful segmentation has been achieved.

First, a formal definition of curves is given. For simplicity, assume that curves exist in the domain $\Omega \in \mathbb{R}^2$. A curve $\Gamma(p)$, parameterized by a parameter p, is a mapping:

$$\Gamma(p) : p \in [a, b] \rightarrow X$$

where X is a topological space within Ω, and $a, b \in \mathbb{R}$. The curve is *regular* if $\Gamma'(p)$ exists $\forall p \in [a, b]$ and $\Gamma'(p) \neq 0 \ \forall p \in [a, b]$. The curve is *closed* if $\Gamma(a) = \Gamma(b)$ and $\Gamma'(a) = \Gamma'(b)$. In this thesis, regular, closed curves will be referred to simply as "curves."

Curve evolution can then be thought of as constructing a family of related curves $\Gamma(p, t)$ that are indexed by an artificial time parameter t. Each curve at $t+1$ is described in terms of how it differs from the curve at t as follows

$$\Gamma(p, t+1) = \Gamma(p, t) + \frac{\partial \Gamma(p, t)}{\partial t}. \tag{2.1}$$

Here,

$$\frac{\partial \Gamma(p, t)}{\partial t} = \mathcal{V}(\Gamma(p, t), t) \tag{2.2}$$

and $\mathcal{V} : \mathbb{R}^2 \times \mathbb{R}^+$ is a vector field that describes the motion of each point along the curve. The partial differential equation (PDE) in Equation (2.2) dictates the motion of the curve, and is sometimes referred to as a "curve flow."

The vector field, $\mathcal{V}$, can be decomposed into a linear combination of the unit tangent $\mathbf{T}(p, t)$ and the inward unit normal $\mathbf{N}(p, t)$, which form a basis for $\mathbb{R}^2$. Hence one can re-write Equation (2.2) as

$$\frac{\partial \Gamma}{\partial t} = \langle \mathcal{V}, \mathbf{T} \rangle \mathbf{T} + \langle \mathcal{V}, \mathbf{N} \rangle \mathbf{N}$$

where $\langle, \rangle$ represents the inner product. However, any movement in the $\mathbf{T}$ direction is equivalent to a re-parameterization of the curve and does not result in any change in the curve's position. Hence, the coefficient on the tangent term can be reduced to zero to yield

$$\frac{\partial \Gamma}{\partial t} = \langle \mathcal{V}, \mathbf{N} \rangle \mathbf{N}, \tag{2.3}$$

because only the normal component of $\mathcal{V}$ contributes to the motion of the contour at each time step.

An active contour segmentation energy $E(\Gamma)$ is said to be "geometric" if it depends only on geometric properties of the contour and not the choice of parameterization [93]. Specifically, the gradient of $E(\Gamma)$ with respect to the curve must be a vector field that is always normal to Γ. Hence, the geometric gradient of $E(\Gamma)$ can be written

$$\nabla_\Gamma E = \langle \nabla_\Gamma E, \mathbf{N} \rangle \mathbf{N}.$$

The energy is then minimized by gradient descent by taking

$$\frac{\partial \Gamma}{\partial t} = -\nabla_\Gamma E = \langle -\nabla_\Gamma E, \mathbf{N} \rangle \mathbf{N}, \tag{2.4}$$

which satisfies the conditions for a geometric flow given that $\mathcal{V} = -\nabla_\Gamma E$. Geometric energies are desirable because the behavior of Γ should be independent of an arbitrary choice of parameterization.

2.2 Level Set Methods

To implement the variational active contours described in the previous section, one must choose a representation for the curve. Initially, much of the work in active contours used explicit, parametric representations of the curves based on control points connected with splines [22, 43]. More recently Osher and Sethian introduced an implicit, parameter-free representation based on level sets [72, 97] that has since become the standard choice due to four key advantages over the previous parameterized representations.

1. Inherent support for topological changes

2. Natural representation of surfaces in $\mathbb{R}^N$

3. Simple estimation of geometric properties

4. Straightforward numerical implementation

When using level sets as an implicit curve representation, Γ is embedded as the zero level set of a higher-order function $\phi : \Omega \to \mathbb{R}$. The signed distance function (SDF) is a

typical choice for ϕ meaning

$$\phi(x) = \begin{cases} -d(x, \Gamma) & \text{for } x \text{ inside } \Gamma \\ 0 & \text{for } x \in \Gamma \\ d(x, \Gamma) & \text{for } x \text{ outside } \Gamma \end{cases} \tag{2.5}$$

where

$$d(x, \Gamma) = \inf_{y \in \Gamma} (\|x - y\|).$$

While this is an arbitrary choice, it provides a simple, smooth function in which to embed the contour and aids in computing geometric properties of the curve. Furthermore, because the curve is embedded as the zero level of ϕ, it can be easily recovered as

$$\Gamma = \{x \in \Omega \mid \phi(x) = 0\}. \tag{2.6}$$

In Figure 2.1 several visualizations can be seen of 2-D curves and the corresponding signed distance functions. It is immediately apparent how topological changes can occur in the embedded curve. Also note that unlike the parametric representation described in Section 2.1, the dimensionality of Ω can be increased arbitrarily to represent surfaces in $\mathbb{R}^3$ and beyond.

Next, consider the computation of the unit normal, $\mathbf{N}$. Let us return to the parametric notation for Γ briefly so that the link between the two representations can be made clear. Bearing in mind that $\phi(\Gamma(p)) = 0 \ \forall \ p$, one has:

$$\frac{\partial \phi(\Gamma(p))}{\partial p} = \left\langle \nabla \phi, \frac{d\Gamma}{dp} \right\rangle = \left\langle \nabla \phi, \left\| \frac{d\Gamma}{dp} \right\| \mathbf{T} \right\rangle = 0$$
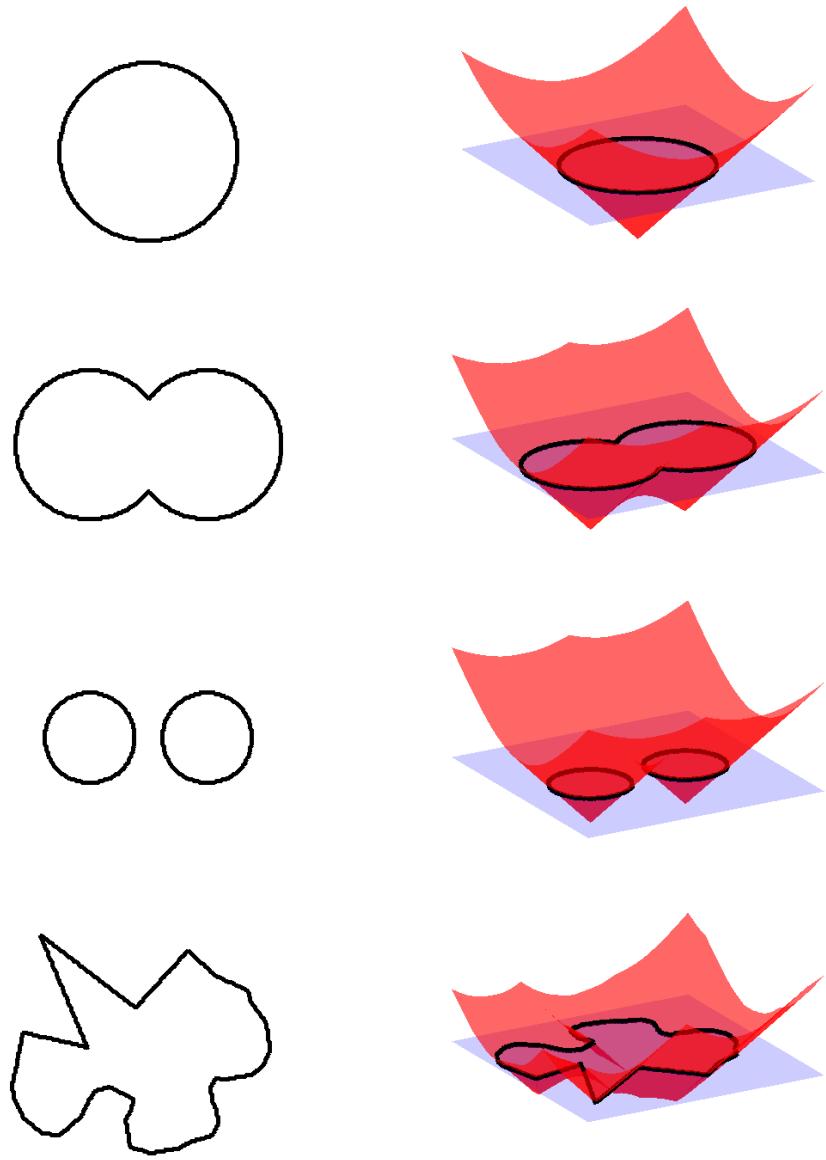
Hence, $\nabla \phi$ is orthogonal to the tangent plane and by normalizing, one obtains

$$\mathbf{N} = -\frac{\nabla \phi}{\|\nabla \phi\|},$$

where the negative sign is a result of the arbitrary choice of direction in the SDF from Equation (2.5).

Another useful geometric property is the curvature κ. This, too can be computed simply from the level set by finding the divergence of $\mathbf{N}$:

$$\kappa = \text{div}(\mathbf{N}) = \text{div}\left(\frac{\nabla \phi}{\|\nabla \phi\|}\right).$$

15

Figure 2.1: Various contours (left) and corresponding signed distance functions shown being intersected by the zero plane (right). Note the topological changes and complex shapes that can be represented using level sets.

Understanding the curvature of Γ is useful for regularization. By moving the curve proportionally to κ the arc-length (or surface area) of Γ can be reduced as described in [27]. A curvature-based regularization term is often used in addition to image-based terms in segmentation energies. This ensures that the resulting segmentation is smooth by penalizing areas of high curvature such as sharp corners.

To deform Γ while it is embedded in ϕ, one must compute the update function for ϕ that corresponds to Equation (2.3). However, one can use the level set representation to simplify this process. Instead of representing the energy as a function of Γ (i.e. $E(\Gamma)$), it is possible to re-write the energy directly in terms of ϕ as

$$E(\phi). \tag{2.7}$$

This approach was introduced in [18] and has become popular since derivations and numerical implementations become more straightforward with this representation. Specifically, if the segmentation energy is written as a function of ϕ, the update equation to minimize Equation (2.7) becomes

$$\frac{\partial \phi}{\partial t} = -\nabla_\phi E, \tag{2.8}$$

which is more easily computed (see Appendix A). This type of flow deforms Γ implicitly as ϕ is deformed explicitly. While this is not strictly equivalent to energies defined in terms of Γ, the steady-state value of ϕ is quite similar. Also note that because the implicit representation of Γ within ϕ is parameter-free, all energies defined in this manner are geometric. Within this thesis, all segmentation energies are be defined in terms of ϕ as in Equation (2.7) because this leads to significantly simpler derivations and more straightforward implementations.

One should note that the computed values of $\frac{\partial \phi}{\partial t}$ are only valid along Γ where $\phi = 0$. Thus, the values on the zero level set must be extended into the remaining level sets for the flow to continue to be valid. Originally, this was accomplished using fast marching methods [97]. However, propagating values from the zero level into the entire domain can add significant computational complexity, especially when the domain is large. If one makes the observation that the structure of ϕ is only important near the zero level set, it

follows that one should use a *narrow band* representation of ϕ [19, 72]. In this scenario, ϕ is only updated at points close to the zero level set. The idea of narrow bands has been extended and made extremely efficient using a sparse representations of ϕ and specialized data structures [103]. This efficient and numerically accurate method has been used as the default implementation for level sets within this thesis. Specific details of the algorithm and its implementation can be found in Appendix B.

2.3 Segmentation Approaches

Active contour methods have become popular in recent years, and have found applications in a wide range of problems including image segmentation and visual tracking; see [12, 75, 76, 107] and the references therein. As previously stated, the basic idea is to allow a contour to deform in order to minimize a given energy functional and produce the desired segmentation; see [65, 73, 93, 97]. Two main categories exist for active contour segmentation energies: edge-based and region-based.

Edge-based active contour models utilize image gradients to identify object boundaries [17, 44]. In one such model known as geodesic (or conformal) active contours the segmentation energy is designed to find a shortest-length curve on a Riemannian space with a metric based on image edges. An example of this type of energy is

$$E(\Gamma) = \oint_0^L \Psi(\Gamma(s))ds, \tag{2.9}$$
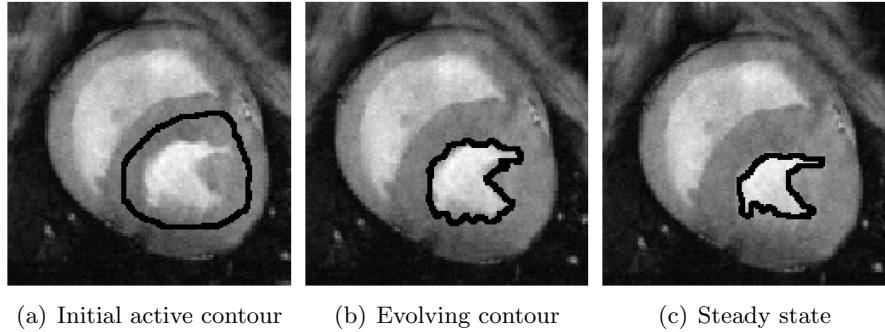
where $s \in [0, L]$ is a parameter chosen to correspond exactly with the length L of Γ and Ψ is a weighting function that is smallest where strong edges exist within the image. A common choice for Ψ is

$$\Psi = \frac{1}{1 + |\nabla(\mathcal{G} * I)|^2},$$

where I represents the image, $\mathcal{G}$ is a Gaussian smoothing kernel, and $*$ is the convolution operator. Minimizing Equation (2.9) with the curve flow

$$\frac{\partial \Gamma}{\partial t} = (\nabla\Psi - \kappa)\mathbf{N}$$

will drive the contour to rest on the strong edges near the contour's initial position. An

18

(a) Initial active contour   (b) Evolving contour   (c) Steady state

Figure 2.2: An MRI image of the heart being segmented with a geodesic edge-based active contour. The initial contour, an intermediate step and the final contour are shown in (a), (b), and (c) respectively.

example of a contour segmentation driven by a geodesic edge-based active contour can be seen in Figure 2.2.

This type of segmentation is based on the assumption that object boundaries would be located at areas that show rapid intensity changes. While this assumption is often true, rapid intensity changes also occur where object boundaries do not exist. In edge-based models no consideration is given to the appearance of regions inside or outside the contour; only information located close to the evolving contour is examined.

This type of highly localized image information is adequate in some situations but has been found to be sensitive to image noise and highly dependent on initial curve placement. One benefit of this type of flow is that no global constraints are placed on the image. Thus, the foreground and background can be heterogeneous and a correct segmentation can still be achieved when strong edges are present and Γ is initialized near the true object boundary.

Recently, work in active contours has been focused on region-based flows inspired by the region-competition work of Zhu and Yuille [108]. These approaches model image regions statistically and find an energy optimum where the model best fits the image. Energies of this type minimize a function f inside of a region R enclosed by Γ and are of the form
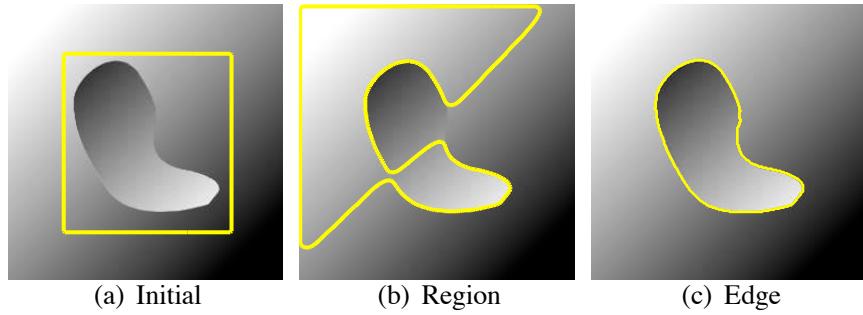
$$E(\Gamma) = \int_R f(x)dx. \tag{2.10}$$

This energy can be minimized by the curve flow

$$\frac{\partial \Gamma}{\partial t} = -f\mathbf{N}.$$

19

Some well-known and widely used region-based active contour models are based on the assumption that image regions have constant intensity [18, 90, 104, 105]. More advanced techniques attempt to model regions by known distributions, intensity histograms, texture maps, or structure tensors [23, 45, 63, 91].

Because region-based approaches examine all information within an image they have many advantages over edge-based methods. These advantages include robustness against initial curve placement and insensitivity to image noise. However, techniques that attempt to model regions using global statistics are usually not ideal for segmenting heterogeneous objects. In cases where the object to be segmented cannot be easily distinguished in terms of global statistics, region-based active contours may lead to erroneous segmentations. Consider the synthetic image in Figure 2.3. Here a situation arises where the foreground and

(a) Initial   (b) Region   (c) Edge

Figure 2.3: A synthetic image of a blob with heterogeneous intensity on a similar background. (a) Initial contour. (b) Unsuccessful result of region-based segmentation. (c) Successful result of edge-based segmentation.

background are heterogeneous and share nearly the same statistical model. The construction of this image causes it to be segmented improperly by a standard region-based algorithm [18] but segmented correctly by an edge-based algorithm [17]. Heterogeneous objects frequently occur in natural and medical imagery. To accurately segment these objects, a new class of active contour energies should be considered which utilizes local information and also incorporates the benefits of region-based techniques. The next chapter presents a framework designed to accomplish this goal.

CHAPTER III

A LOCALIZED ACTIVE CONTOUR MODEL

This chapter presents a natural framework that allows any region-based active contour segmentation method to be re-formulated so that it relies on information from local - rather than global - image regions. Using this framework *localized active contours* are introduced. Active contours of this type are capable of segmenting objects with heterogeneous feature profiles that would be difficult to capture correctly using standard global methods. The presented localization technique is versatile enough to be used with any global region-based active contour energy and instill in it the benefits of localization.

In the following section the need for the localized active contour model is motivated, and Section 3.2 reviews relevant work from the literature. Section 3.3 gives a detailed description of the proposed framework for localizing region-based segmentation energies, and Section 3.4 demonstrates how to use the framework to localize three well known segmentation models. An extension that allows simultaneous segmentation of multiple regions is discussed in Section 3.5. Key implementation details are given in Section 3.6 after numerous experiments compare local flows with corresponding global flows, analyze key parameters, discuss limitations of the technique, and show examples of accurate segmentations on challenging images. Section 3.7 offers concluding remarks about this new class of localized active contours.

3.1 Motivation for Locally Optimal Solutions

It is typically challenging to represent the foreground and background of an image using global statistical models. This is true, to some extent, in all but the simplest images. In order to successfully segment these challenging images using existing methods, one must create intricate statistical models that describe the object's appearance. However, this becomes more difficult as the complexity of the image increases. Consider the images in Figure 3.1. The AIRPLANE image is simple to segment because the object (the airplane)

21

(a) AIRPLANE   (b) HUG

Figure 3.1: (a) The AIRPLANE image is simple to segment because the foreground and background are quite different. (b) The HUG image is much more challenging due to the complexity of the image.

can be easily distinguished from the background by the brightness of the image pixels. However, in the HUG image the animals and the background have similar intensities that are not consistent or easily distinguishable from one another. As a result, the HUG image is difficult to segment using existing methods.

The key is to recognize that in complex images such as the HUG image it is only difficult to find object boundaries when the entire foreground and background are examined together. If this criteria is relaxed, and one examines each point along an ideal curve separating the animals from the background, it is apparent that the foreground and background have easily distinguishable properties. Thus, the localized active contour framework presented in this chapter aims to explore the ramifications of the following observation: **Along the boundary, the object and background will have different appearances** *locally.*

As discussed in Chapter 2, active contours are variational segmentation methods that rely on energy minimization to drive a contour toward an appropriate segmentation. Typically region-based active contour energies have only a few minima and the resulting segmentations can be though of as "almost-global" solutions. Conversely, edge-based active contours have a tremendous number of local minima, which makes them sensitive to initialization. In this work, a scale parameter is introduced that can control the number of

local minima by changing the degree of spatial localization. This allows one to create segmentation energies that have a moderate number of desirable local minima rather than too many in the case of edge-based active contours or too few in the case of region-based active contours.

3.2   Prior Work and Contributions

Before the localized active contour model is presented, this section reviews several relevant papers from the literature that address similar problems. Paragios and Deriche [77] presented a method in which edge-based energies and region-based energies were explicitly summed to create a joint energy, which was then minimized. In [99] Sum and Cheung take a similar approach and minimize the sum of a global region-based energy and a energy based on local image contrast. The idea of incorporating localized statistics into a variational framework begins with the work of Brox and Cremers [13] who show that segmenting with local means is a first order approximation of the popular piecewise smooth simplification [68] of the Mumford-Shah functional [69]. This focus on the piecewise smooth model is also presented in several related works described below.

Li *et al.* [54] analyze the localized energy of Brox and Cremers and compare it to the piecewise smooth model in much more detail. However, they offer no explicit analysis of the appropriate scale on which to localize. Piovano *et al.* [84] focus on fast implementations employing convolutions that can be used to compute localized statistics quickly, and hence yield results similar to piecewise-smooth segmentation in a much more efficient manner. The effect of varying scales is noted but not discussed in detail. The work of An *et al.* [2] also mentions the efficiency of localized approaches versus full piecewise smooth estimation. That work introduces a way in which localizations at two different scales can be combined to allow sensitivity to both coarse and fine image features. Finally, in [49], the author proposed a flow based on computing geodesic curves in the space of localized means rather than an approximating a piecewise-smooth model.

Contributions: The framework in this chapter makes three major contributions:

1. A novel framework that can be used to localize any region-based energy.

2. A natural way for localized contours to interact and create n-ary segmentations.

3. An in-depth study of the effect of the localization radius on segmentation results.

The localization framework presented allows any region-based energy to be localized in a fully variational way. The significant improvement of localization within this framework is that objects that have complex appearances can be successfully segmented with localized energies when corresponding global energies fail. The framework is also used to derive three localized energies. The first, presented in Section 3.4.1, is similar to those in the works mentioned above. Two additional region-based segmentation energies and their localized counterparts are formulated in Sections 3.4.2 and 3.4.3. To best of this author's knowledge localization of energies other than the Chan and Vese energy have not been shown outside of this work. These are provided as examples to demonstrate how any energy can be localized in a similar manner. The primary claim is that localization in the presented framework can improve the segmentations provided by a global energy in certain circumstances. The author does not suggest that one of the proposed localized energies is superior to the others, just that in many cases localizing a global energy in the manner suggested will improve performance.

Because binary segmentation is sometimes an insufficient result for higher-level vision problems, a novel method is included that allows n localized active contours to naturally compete to segment different objects that may or may not share borders. This new method extends the work of Brox and Weickert [14], so that it can be successfully utilized with localized active contours.

Lastly, this study discusses the significance of the scale parameter, which is common to all localized statistical models. This parameter controls the degree of localization and is fundamentally important to the use of localized energies. Experiments are provided that explain its effect and give guidelines to assist in choosing this parameter correctly.

Additional experiments are also presented to analyze the strengths and limitations of the technique.

3.3   Localization Framework

This section contains a description of the localized region-based active contour segmentation framework. Within this framework, segmentations are based on information from local interior and exterior regions along the contour rather than global regions. This allows one to remove the assumption that the foreground and background regions are distinguishable based on their global statistics. Instead, the only assumption is that interior and exterior regions will be different locally.

The analysis of local statistics leads to the construction of a family of local regions about each point on the curve. In order to optimize localized energies, each point moves to optimize the component of the energy added from its local region. Each point's component of the overall energy is computed by splitting the local neighborhoods into local-interior and local-exterior using the evolving curve and a characteristic function as described below.

As in Chapter 2, I will denote a given image defined on the domain Ω, and Γ will denote a closed contour represented as the zero level set of a signed distance function ϕ (i.e. $\Gamma = \{x | \phi(x) = 0\}$) [72, 73]. The interior of Γ is specified by the following approximation of the smoothed Heaviside function

$$\mathcal{H}\phi(x) = \begin{cases} 1 & \phi(x) < -\epsilon \\ 0 & \phi(x) > \epsilon \\ \frac{1}{2}\{1 + \frac{\phi}{\epsilon} + \frac{1}{\pi}\sin(\frac{\pi\phi(x)}{\epsilon})\} & \text{otherwise} \end{cases} \quad (3.1)$$

Similarly, the exterior of Γ is defined as $(1 - \mathcal{H}\phi(x))$.
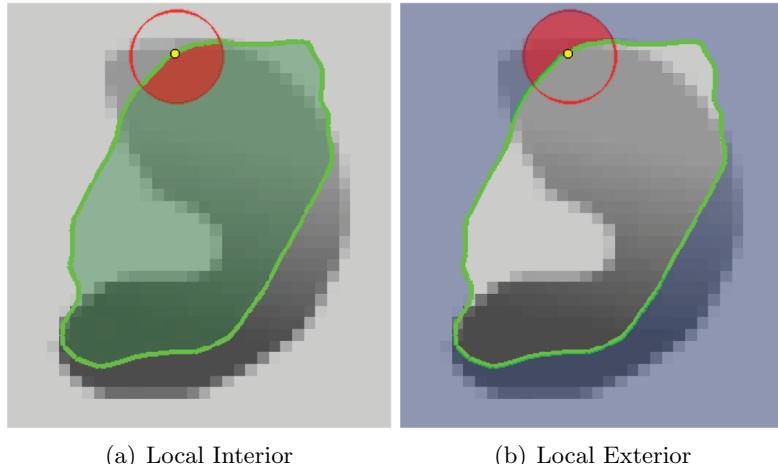
To specify the area immediately adjacent to the curve, the derivative of $\mathcal{H}\phi(x)$ is used. Hence, the curve is specified by a smoothed version of the Dirac delta:

$$\delta\phi(x) = \begin{cases} 1 & \phi(x) = 0 \\ 0 & |\phi(x)| < \epsilon \\ \frac{1}{2\epsilon}\{1 + \cos\left(\frac{\pi\phi(x)}{\epsilon}\right)\} & \text{otherwise} \end{cases} \quad (3.2)$$

Now a second spatial variable y is introduced. In the remainder of this paper x and y are used as independent spatial variables each representing a single point in Ω. Using this notation, the characteristic function

$$\mathcal{B}(x, y) = \begin{cases} 1 & \|x - y\| < r \\ 0 & \text{otherwise} \end{cases} \tag{3.3}$$

is introduced in terms of a radius parameter r. This characteristic function is used to identify local regions. This function will be 1 when the point y is within a ball of radius r centered at x and 0 otherwise. The interaction of $\mathcal{B}(x, y)$ with the interior and exterior of Γ is illustrated in Figure 3.2.

(a) Local Interior   (b) Local Exterior

Figure 3.2: A ball is considered at each point along the contour. This ball is split by the contour into local interior and local exterior regions. In both images, the point x is represented by the small dot. The $\mathcal{B}(x, y)$ neighborhood is represented by the larger red circle. In (a) the local interior is the shaded part of the circle, and in (b) the shaded part of the circle indicates the local exterior.

Using $\mathcal{B}(x, y)$, an energy functional is defined in terms of a generic internal energy functional, F. The resulting energy is given as follows:

$$E(\phi) = \int_{\Omega_x} \delta\phi(x) \int_{\Omega_y} \mathcal{B}(x, y) \cdot F(I, \phi, x, y) dy \, dx. \tag{3.4}$$

The functional, F is a generic internal energy measure used to represent local adherence to a given model at each point along the contour. In Section 3.4, several possible candidates for F are examined to show how any region-based energy can be modified and included as F this framework.

While computing E, only contributions from the points within the radius r of the contour are considered. By ignoring image complexity that may arise outside of this radius, the ability to capture a much broader range of objects is achieved. In Equation (3.4) this is accomplished by multiplication with the Dirac function $\delta\phi(x)$ in the outer integral over x. Note that this term also ensures that the curve will not change topology by spontaneously developing new contours, although it still allows for contours to split and merge

For every point x with non-zero $\delta\phi(x)$, the local region inside of $\mathcal{B}(x, y)$ is selected so that F only operates on local image information. Thus, the total contribution of the first term of the energy is the sum of F values for every $\mathcal{B}(x, y)$ neighborhood along the zero level set.

Finally, in order to keep the curve smooth, a regularization term is added as is commonly done in active contour segmentation energies. The arclength of the curve is penalized and weighted by a parameter λ. The final energy is given as follows:

$$
\begin{aligned}
E(\phi) &= \int_{\Omega_x} \delta\phi(x) \int_{\Omega_y} \mathcal{B}(x, y) \cdot F\big(I, \phi, x, y\big) dy \; dx \\
&\quad + \lambda \int_{\Omega_x} \delta\phi(x) \|\nabla\phi(x)\| dx.
\end{aligned} \tag{3.5}
$$

By taking the first variation of this energy with respect to ϕ the following evolution equation is obtained[1]:

$$
\begin{aligned}
\frac{\partial \phi}{\partial t}(x) &= \delta\phi(x) \int_{\Omega_y} \mathcal{B}(x, y) \cdot \nabla_{\phi(y)} F\big(I, \phi, x, y\big) dy \\
&\quad + \lambda \delta\phi(x) \, \mathrm{div}\left(\frac{\nabla\phi(x)}{|\nabla\phi(x)|}\right) \|\nabla\phi(x)\|.
\end{aligned} \tag{3.6}
$$

Notice that the only restriction on the internal energy, F is that its first variation with respect to ϕ can be computed. This ensures that any region-based segmentation energy can be put into this framework.

3.4 Various Internal Energies

Having formulated the framework in terms of a generic internal energy measure F, three specific energies are now introduced: the *uniform modeling energy*, the *mean separation*

---

[1]See Appendix A for details on the derivation.

energy, and the *histogram separation energy*. These three energies are presented as examples of how any energy can be localized and improved using this framework. However, no claim is made that one energy out performs the others in all cases. This section describes each global energy, gives an intuitive description of its behavior, and shows how it can be incorporated into the generic framework described above.

Two well known techniques [18, 105] make use of the mean intensities of the entire interior and exterior regions to perform segmentations. These means are denoted as μ_{in} and μ_{out} respectively:

$$\mu_{\text{in}} = \frac{\int_{\Omega_y} \mathcal{H}\phi(y) \cdot I(y) dy}{\int_{\Omega_y} \cdot \mathcal{H}\phi(y) dy} \tag{3.7}$$

$$\mu_{\text{out}} = \frac{\int_{\Omega_y} (1 - \mathcal{H}\phi(y)) \cdot I(y) dy}{\int_{\Omega_y} \cdot (1 - \mathcal{H}\phi(y)) dy}. \tag{3.8}$$

In Sections 3.4.1 and 3.4.2, internal energy functions are discussed that rely on local mean intensities to separate regions. Hence, the localized equivalents of μ_{in} and μ_{out} are now defined in terms of the $\mathcal{B}(x, y)$ function. The localized versions of the means $\mu_{\text{in}}(x)$ and $\mu_{\text{out}}(x)$,

$$\mu_{\text{in}}(x) = \frac{\int_{\Omega_y} \mathcal{B}(x, y) \cdot \mathcal{H}\phi(y) \cdot I(y) dy}{\int_{\Omega_y} \mathcal{B}(x, y) \cdot \mathcal{H}\phi(y) dy} \tag{3.9}$$

$$\mu_{\text{out}}(x) = \frac{\int_{\Omega_y} \mathcal{B}(x, y) \cdot (1 - \mathcal{H}\phi(y)) \cdot I(y) dy}{\int_{\Omega_y} \mathcal{B}(x, y) \cdot (1 - \mathcal{H}\phi(y)) dy} \tag{3.10}$$

represent the intensity mean in local interior and exterior regions around a point x. These localized statistics are needed to compute how much F will contribute to the overall energy at each point along the contour.

3.4.1   Uniform Modeling (UM) Energy

A well-known example of an energy that uses a constant intensity model energy is the Chan-Vese energy [18], referred to here as the *uniform modeling energy*:

$$E_{\text{UM}} = \int_{\Omega_y} \mathcal{H}\phi(y)(I(y) - \mu_{\text{in}})^2 + (1 - \mathcal{H}\phi(y))(I(y) - \mu_{\text{out}})^2 dy. \tag{3.11}$$

This energy models the foreground and background as constant intensities represented by their means, μ_{in} and μ_{out}. The corresponding internal energy, F is formed by replacing

global means μ_{in} and μ_{out} by their local equivalents from Equation (3.9) and (3.10) as follows:

$$F_{\text{UM}} = \mathcal{H}\phi(y)(I(y) - \mu_{\text{in}}(x))^2 + (1 - \mathcal{H}\phi(y))(I(y) - \mu_{\text{out}}(x))^2. \tag{3.12}$$

This F can be substituted directly into Equation (3.5) to form a complete localized energy. In order to obtain the evolution equation for ϕ, the derivative of F is taken with respect to $\phi(y)$. The derivative can be written immediately as

$$\nabla_{\phi(y)} F_{\text{UM}} = \delta\phi(y) \left((I(y) - \mu_{\text{in}}(x))^2 - (I(y) - \mu_{\text{out}}(x))^2 \right). \tag{3.13}$$

By inserting this into Equation (3.6), one obtains the curvature flow for the localized version of the uniform modeling energy:

$$\frac{\partial \phi}{\partial t}(x) = \delta\phi(x) \int_{\Omega_y} \mathcal{B}(x, y)\delta\phi(y) \left((I(y) - \mu_{\text{in}}(x))^2 - (I(y) - \mu_{\text{out}}(x))^2 \right) dy$$
$$+ \lambda\delta\phi(x) \operatorname{div}\left(\frac{\nabla\phi(x)}{|\nabla\phi(x)|} \right) \|\nabla\phi(x)\|. \tag{3.14}$$

The uniform modeling flow finds its minimum energy when the interior and exterior are best approximated by means μ_{in} and μ_{out}. In the localized version, the minimum is obtained when each point on the curve has moved such that the local interior and exterior about every point along the curve is best approximated by local means $\mu_{\text{in}}(x)$ and $\mu_{\text{out}}(x)$.

3.4.2 Mean Separation (MS) Energy

Another important region-based energy that uses mean intensities is the one proposed by Yezzi *et al.* [105], referred to here as *mean separation energy*:

$$E_{\text{MS}} = \int_{\Omega_y} -(\mu_{\text{in}} - \mu_{\text{out}})^2. \tag{3.15}$$

This energy relies on the assumption that foreground and background regions should have maximally separate mean intensities. Optimizing the energy causes the curve to move so that interior and exterior means have the largest possible difference. There is no restriction on how well the regions are modeled by μ_{in} and μ_{out}. A corresponding F is formed from the global energy by substituting local means for global ones as shown here:

$$F_{\text{MS}} = -(\mu_{\text{in}}(x) - \mu_{\text{out}}(x))^2. \tag{3.16}$$

By substituting the derivative of F_{MS} into Equation (3.6) one obtains the following local region-based flow:

$$
\begin{aligned}
\frac{\partial \phi}{\partial t}(x) \;=\;& \delta\phi(x) \int_{\Omega_y} \mathcal{B}(x,y)\delta\phi(y) \cdot (\mu_{\mathrm{in}}(x) - \mu_{\mathrm{out}}(x)) \left(\frac{I(y) - \mu_{\mathrm{in}}(x)}{A_{\mathrm{in}}(x)} + \frac{I(y) - \mu_{\mathrm{out}}(x)}{A_{\mathrm{out}}(x)} \right) dy \\
& + \lambda\delta\phi(x) \operatorname{div}\left(\frac{\nabla\phi(x)}{|\nabla\phi(x)|} \right) \|\nabla\phi(x)\|.
\end{aligned}
\tag{3.17}
$$

where $A_{\mathrm{in}}(x)$ and $A_{\mathrm{out}}(x)$ are the areas of the local interior and local exterior regions respectively given by

$$
A_{\mathrm{in}}(x) \;=\; \int_{\Omega_y} \mathcal{B}(x,y) \cdot \mathcal{H}\phi(y)dy
\tag{3.18}
$$

$$
A_{\mathrm{out}}(x) \;=\; \int_{\Omega_y} \mathcal{B}(x,y) \cdot (1 - \mathcal{H}\phi(y))dy.
\tag{3.19}
$$

The optimum of this energy is obtained when $\mu_{\mathrm{in}}(x)$ and $\mu_{\mathrm{out}}(x)$ are the most different at every point x along the contour. In some cases, mean separation is more desirable than fitting a constant mean model. In the former case, the local interior and exterior means are encouraged to be different rather than constant. This allows contours driven by the mean separation energy to settle on image edges without being distracted by non-uniform local regions.

3.4.3 Histogram Separation (HS) Energy

Next, a more complex energy is considered that looks beyond means and compares the full histograms of the foreground and background. Its incorporation into the framework is shown to be as simple as the two previous energies. Consider $P_{\mathrm{in}}(z)$ and $P_{\mathrm{out}}(z)$ to be two smoothed intensity histograms computed from the global interior and exterior regions of an image I using z intensity bins such that

$$
P(z) = \int_{\Omega_x} \mathcal{H}(I(x) - z_{\mathrm{min}}) - \mathcal{H}(I(x) - z_{\mathrm{max}})dx.
\tag{3.20}
$$

The Bhattacharyya coefficient $\mathcal{B}$ is a measure used to compare probability density functions that results in a scalar corresponding to the similarity of the two histograms [11]. Recently Michailovich $et\ al.$ [63] proposed an image segmentation energy,

$$
E_{\mathrm{HS}} = \mathcal{B} = \int_z \sqrt{P_{\mathrm{in}}(z)P_{\mathrm{out}}(z)}dz
\tag{3.21}
$$

based on minimizing this measure. This will be called the *histogram separation energy*. It works by separating intensity histograms of the regions inside and outside of the curve, and thus allows interior and exterior regions to be complex as long as their intensity profiles are different.

In the localized case, $P_{\text{in},x}(z)$ and $P_{\text{out},x}(z)$ will represent the intensity histograms in the local image regions $\mathcal{B}(x,y) \cdot \mathcal{H}\phi(y)$ and $\mathcal{B}(x,y) \cdot (1 - \mathcal{H}\phi(y))$, respectively. As before the internal energy measure F_{HS} is formed by substituting the local equivalents for $P_{\text{in}}(z)$ and $P_{\text{out}}(z)$ yielding the following expression:

$$F_{\text{HS}} = \int_z \sqrt{P_{\text{in},x}(z)P_{\text{out},x}(z)}dz. \tag{3.22}$$

By substituting the first variation of F_{HS} into Equation (3.6) the evolution equation for the localized version of this flow,

$$\frac{\partial \phi}{\partial t}(x) \quad = \delta\phi(x)\frac{1}{2}\int_{\Omega_y} \mathcal{B}(x,y)\delta\phi(y) \cdot$$
$$\left[F_{\text{HS}}\left(\frac{1}{A_{\text{out}}} - \frac{1}{A_{\text{in}}}\right) + \int_z K(z - I(y))\left(\frac{1}{A_{\text{in}}}\sqrt{\frac{P_{\text{out},x}(z)}{P_{\text{in},x}(z)}} - \frac{1}{A_{\text{out}}}\sqrt{\frac{P_{\text{in},x}(z)}{P_{\text{out},x}(z)}}\right) dz \right] dy$$
$$+ \lambda\delta\phi(x) \operatorname{div}\left(\frac{\nabla\phi(x)}{|\nabla\phi(x)|}\right) \|\nabla\phi(x)\|, \tag{3.23}$$

is obtained with K representing a Gaussian kernel.

By using the Bhattacharyya measure to quantify the separation of intensity histograms, the global version of this flow is capable of segmenting objects that have non-uniform intensities. However, the intensity profile of the entire object and the entire background must still be separable. In the localized version, this global constraint is removed but the segmentation remains capable of effectively separating locally non-uniform regions. An example of when this property is useful is discussed in Section 3.6.1 and shown in Figure 3.7.

3.5 Coupling Localized Evolutions

In cases where multiple foreground objects exist, simple separation into foreground and background is not sufficient. This section shows how to extend the proposed localized region based framework to allow simultaneous segmentation of multiple objects. Inspiration

is drawn from the work of Brox and Weickert [14] who proposed a simple but effective algorithm for multiple region segmentation based on the idea of competing regions.

In a standard single-level-set evolution scheme, the energy update equation can be though of as having two competing components: *advance* and *retreat*. The advance component a is always positive and tries to move the curve outward along its normal. Alternatively, the retreat component r is always negative and tries to move the curve inward along its normal. The relative magnitudes of a and r govern curve evolution. Hence, the update equation for ϕ can be expressed as

$$\frac{\partial \phi}{\partial t}(x) = \delta\phi(x)(a + r). \tag{3.24}$$

To give an example, consider the update equation for the local uniform modeling energy in Equation (4.7). This equation may be re-written in terms of the forces a and r with

$$a = \int_{\Omega_y} \mathcal{B}(x,y)\delta\phi(y) \cdot (I(y) - \mu_{\text{in}}(x))^2 dy + \frac{\lambda}{2}\, \delta\phi(x)\, \text{div}\left(\frac{\nabla\phi(x)}{|\nabla\phi(x)|}\right) \|\nabla\phi(x)\|$$

$$r = -\int_{\Omega_y} \mathcal{B}(x,y)\delta\phi(y) \cdot (I(y) - \mu_{\text{out}}(x))^2 dy + \frac{\lambda}{2}\, \delta\phi(x)\, \text{div}\left(\frac{\nabla\phi(x)}{|\nabla\phi(x)|}\right) \|\nabla\phi(x)\|.$$

Note how the length penalty used for curve regularization is included in both the a and r terms. Inclusion of this term ensures that all forces acting on the curve are represented solely by a and r.

The inherent competition of a and r in this formulation of curve evolution allows multiple signed distance functions to interact. Consider n signed distance functions, $\{\phi_i\}_{i=1}^n$ representing n evolving curves. In [14] the goal is to evolve every ϕ_i such that all points in the domain are eventually in the interior of exactly one curve. To accomplish this, each ϕ_i moves according to

$$\frac{\partial \phi_i}{\partial t}(x) = \delta\phi_i(x)\left(r_i - \max_{\substack{\delta\phi_i(x) > 0 \\ j \neq i}} (r_j, r_i - 1) \right). \tag{3.25}$$

This scheme compares the retreat portion of the typical evolutions and causes all of the curves interacting at a point to move together according to the strongest retreat force. When only one curve is present, it advances outward toward the uninhabited region. The simplicity of this method is appealing, and it is capable of producing complete segmentations

of an image. This behavior is well suited for global energies but a modification of this scheme is needed when used with the proposed techniques due to their local nature.
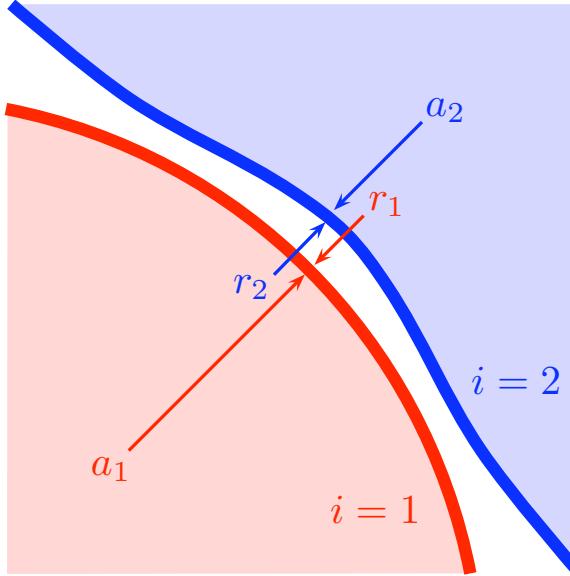
The proposed localized techniques are capable of segmenting complex, non-uniform objects. Thus, if they are used directly in the framework of [14] to produce complete segmentations the proposed localized active contours could incorrectly capture multiple objects' boundaries within the same contour. Specifically, the tendency of curves to advance into uninhabited areas could cause a local-looking contour to move far from its intended object and fail to capture it correctly. Thus, the method is modified to work more naturally with localized contours by allowing regions of the domain to remain uncovered while still preventing overlaps.

The goal is to allow multiple contours to compete with each other at an interface, but also allow them to compete with themselves when no other contours are nearby. This allows contours to stop, either as they would in a single-contour framework, or by competing with adjacent contours. To do this the notion of competition between advance and retreat forces is retained and combined differently to produce a new set of i update equations:

$$\frac{\partial \phi_i}{\partial t}(x) = \delta \phi_i(x) \left(\max_{\substack{\delta \phi_i(x) > 0 \\ j \neq i}} (a_i, -r_j) + \min_{\substack{\delta \phi_i(x) > 0 \\ j \neq i}} (r_i, -a_j) \right). \tag{3.26}$$

In this formulation, the advance force of the current contour is compared to the corresponding retreat forces of all adjacent contours. Similarly, the retreat force is compared to adjacent advance forces. By choosing the strongest candidates in each case all contours at an interface will move together in order to find the best joint solution, and lone contours will continue to evolve as before. Figure 3.3 shows two interacting contours and illustrates the advance and retreat forces acting on each contour.

The ELEPHANTS image in Figure 3.4 shows the results of an n-ary segmentation produced by allowing two contours to deform according to Equation (3.26). The segmentation is initialized with one contour on the large elephant and one contour on the small elephant. Each contour uses the local mean separation energy to find the optimal boundary. However as the contours move they compete with themselves and with one another to ensure that

Figure 3.3: Advance, a_i and retreat, r_i forces are shown as they affect two interacting contours.
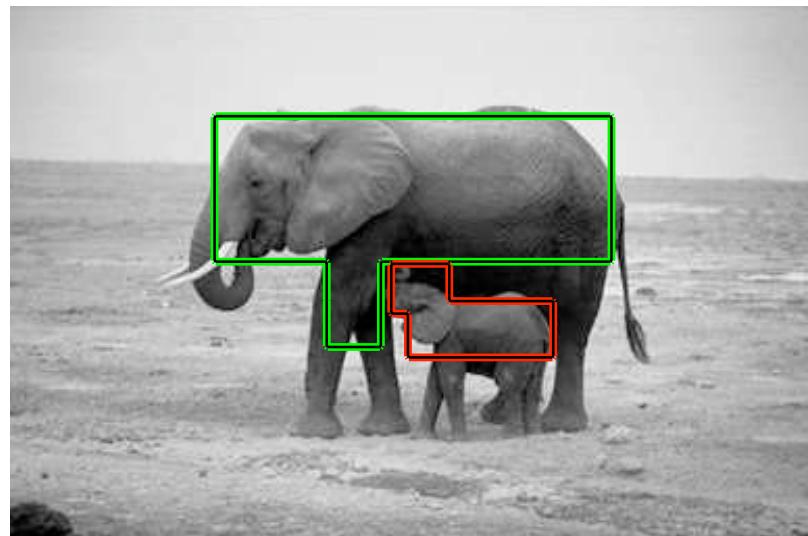
they stop at the appropriate boundary and never overlap. This ensures that both elephants are captured correctly and uniquely. In the final trinary segmentation, the curves come to rest alone on certain parts of the elephants' boundaries and together on shared boundaries, making for a reasonable segmentation result.

3.6 *Experiments and Analysis*
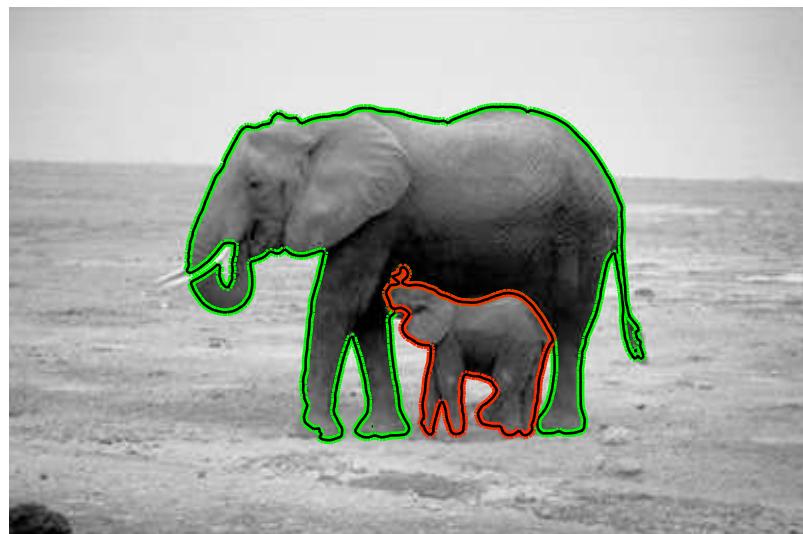
In order to demonstrate the strengths and limitations of localized active contours several experiments are presented. First, the three presented localized energies are compared with their global counterparts to show the improvements offered by localization. Next, the effects of local radius selection and contour initialization are studied. Finally, the speed and convergence properties of the proposed method are examined.

3.6.1 Comparison With Global Energies

In Section 3.4 three global energies were presented and it was shown how they could be localized using the framework described in this chapter. Here the improvements offered by localization are demonstrated. As with all segmentation techniques, the three global techniques behave somewhat differently from one another. This is due to differences in the

(a) Initialization

(b) Final Segmentation
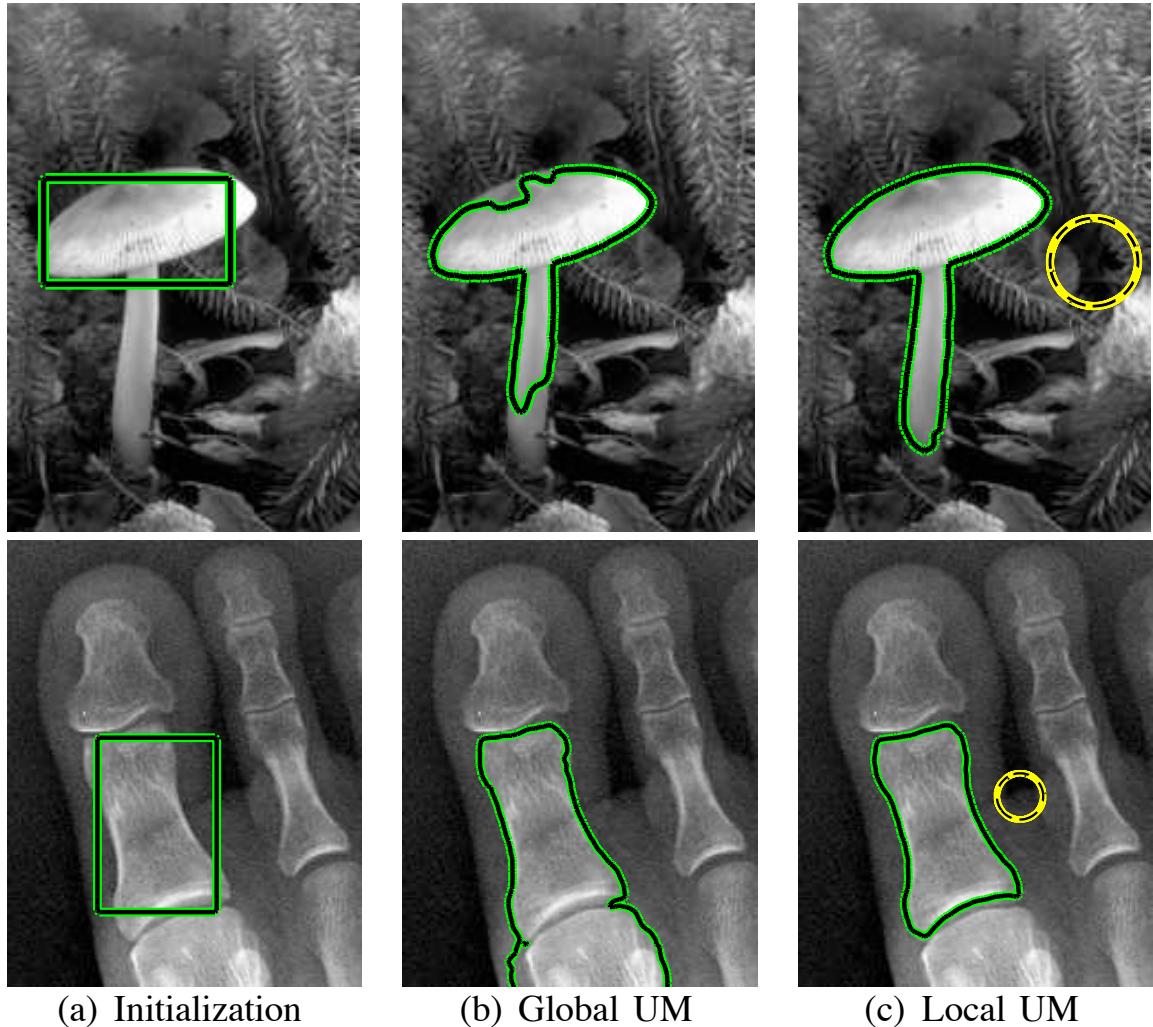
Figure 3.4: The ELEPHANTS image demonstrates how two contours can interact with each other to find a correct segmentation.

underlying assumptions about the image that are inherent in each energy. Likewise there are differences in the behavior of the corresponding localized energies. The purpose of the experiments below is to demonstrate that localization can improve the performance of a given global energy, not to specifically compare the original global energies themselves. In each case, the global energies find segmentations that are consistent with their underlying assumptions about image content but are ultimately incorrect. Only the localized methods are capable of obtaining a correct segmentation in these cases.
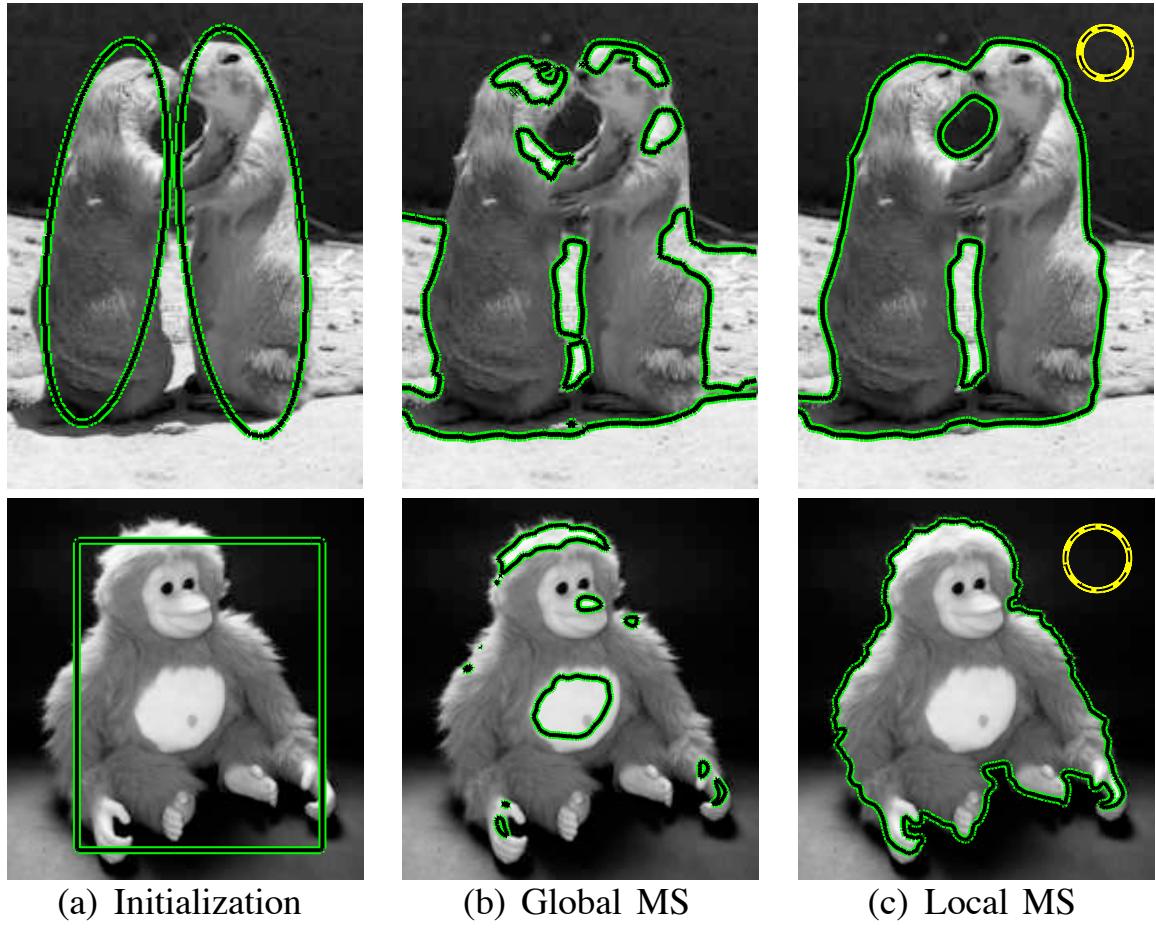
Uniform Modeling Energy: Initially, the uniform modeling energy from Section 3.4.1 is considered. In Figure 3.5, the benefits of localization are seen. The localized active contour is capable of extending further to find true object boundaries in the MUSHROOM image and is capable of stopping earlier on true object boundaries in the X-RAY image. These examples show how global techniques can experience significant problems even when the images seem simple. The slight intensity gradients in these images prevent global region based methods from correctly capturing the objects.

Mean Separation Energy: In Figure 3.6, the global mean separation energy from Section 3.4.2 and its corresponding localization are compared. Notice that the curves driven by the global energy find only the brightest parts of the image while curves driven by the localized mean separation energy come to rest on object boundaries. Both the HUG image and the MONKEY image show objects and backgrounds that are multi-modal and have varying intensities that change either smoothly or rapidly in different parts of the image. In the HUG image in Figure 3.6, the proposed method is initialized with two ellipses corresponding to a single level set. The contour changes topology as the two ellipses merge to capture both animals. An initial placement like this is necessary to segment the spaces between the animals. Because the level set is only updated in areas specified by $\delta\phi(x)$, it is not possible for new contours to spontaneously emerge.

Histogram Separation Energy: Finally, Figure 3.7 compares the global histogram separation energy from Section 3.4.3 to its localization. Again, a clear improvement is shown.

(a) Initialization (b) Global UM (c) Local UM
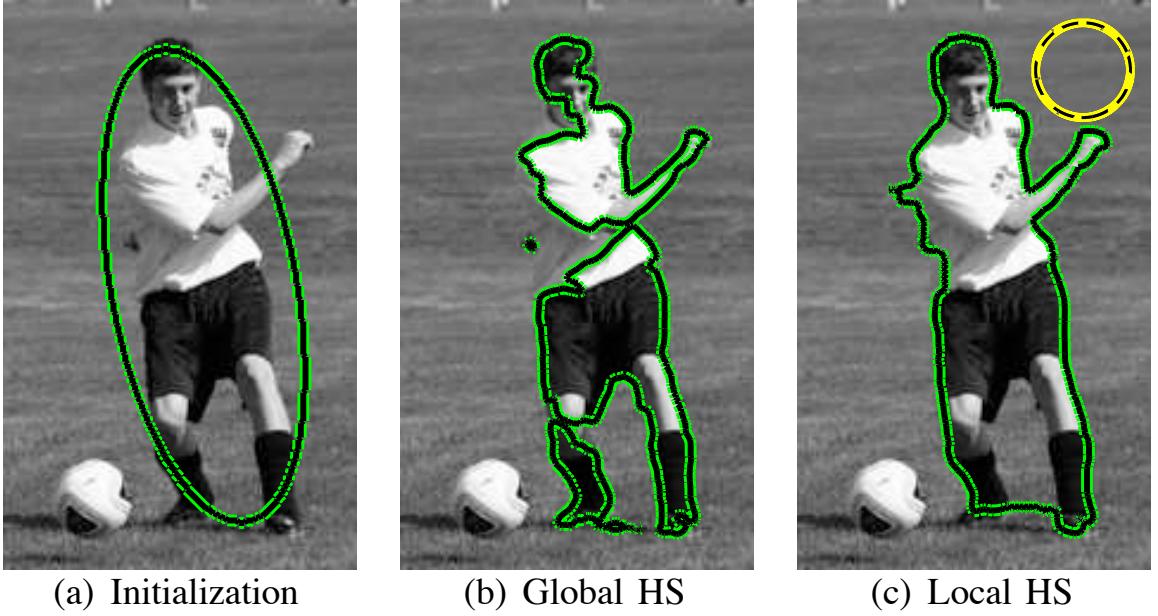
Figure 3.5: Segmentations of the MUSHROOM and X-RAY image using global and local uniform modeling energy. The dashed yellow circle in (c) represents the localization scale. Considerable improvement due to localization can be seen.

(a) Initialization   (b) Global MS   (c) Local MS

Figure 3.6: Segmentations of the HUG and MONKEY image using global and local mean separation energy. The dashed yellow circle in (c) represents the localization scale. Considerable improvement due to localization can be seen.

While the localized contour does not capture the area between the player's legs, the segmentation found over the rest of the player is much more accurate than in the global case. The PLAYER image shows sharp changes in intensities within the foreground. Strong edges such as those between the player's shirt and pants, and between the player's socks and legs, make the localized histogram separation energy a good choice for this image. Here, the foreground is sometimes locally bimodal meaning that energies based on local means would have trouble segmenting the image. It can be seen that the histogram separation energy performs well, though.

(a) Initialization   (b) Global HS   (c) Local HS
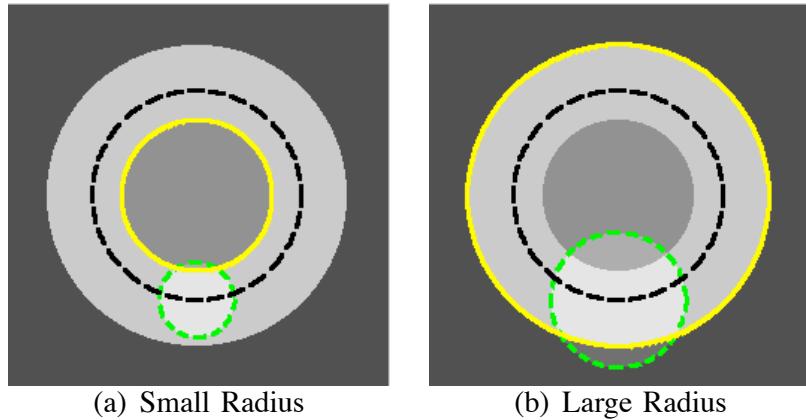
Figure 3.7: Segmentations of the PLAYER image using global and local histogram separation energy. The dashed yellow circle in (c) represents the localization scale. Considerable improvement due to localization can be seen.

For consistency in these experiments, the influence of contour smoothness is weighted with $\lambda = 0.1 \max(|\frac{\partial \phi}{\partial t}|)|_{t=0}$. The specific value for λ is determined separately in each experiment at the first iteration. This choice, which is dependent on the specific energy being used, was made so that the relative influence of smoothness was consistent across the different energies tested. The size of the local radius is shown by the dashed yellow circle drawn on the results for the localized methods. The radius was set empirically for each experiment; information on choosing this parameter can be found in Section 3.6.2. For

results in Figure 3.7, 256 bins were used when computing local or global image histograms. All segmentations were allowed to run until convergence.

3.6.2   Selecting the Localization Radius

The radius of the ball selected by $\mathcal{B}(x, y)$ is an important parameter to be considered when using localized energies. Its size determines the degree of localization that the resulting segmentation will have. As such it should be chosen based on the scale of the object(s) of interest and the presence and proximity of the surrounding clutter. For example, when attempting to capture objects that are small with nearby clutter, a small localization radius should be used. Larger radii are useful when attempting to segment large objects with less nearby clutter.

(a) Small Radius   (b) Large Radius
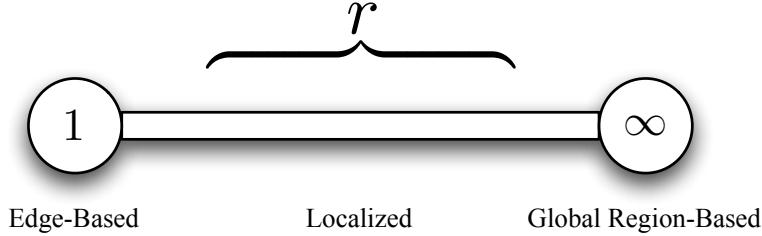
Figure 3.8: In both images, the initialization is the same (black dashed line). However, the ball described by $\mathcal{B}(x, y)$ (green dashed line) is optimized for small structures in (a) and larger structures in (b). Final segmentations with the mean separation energy are shown by a solid yellow line. With a larger radius, a more global final solution is found than with a smaller radius.

The synthetic example in Figure 3.8 illustrates the effect of different localization radii. In this example, there is no reason to prefer one circle over the other as the "correct" segmentation. With the same initialization it is possible to obtain two different results by varying the radius r of the localizing ball. With a small radius, one can capture the local result (the smaller, dark disk), and with a large radius one captures the global result (both disks together). This is a useful property for images where multiple correct segmentations may exist. Indeed, depending on the nature of the objects to be segmented the proposed

method can be tuned to capture fine scale or coarse scale results.

The parameter r also enforces the smoothness of statistics inside and outside the contour. The local statistics along the curve are forced to change smoothly by the fact that $\mathcal{B}(x, y)$ neighborhoods overlap. Thus, the larger the radius, the smoother the change in local statistics must be along the curve.
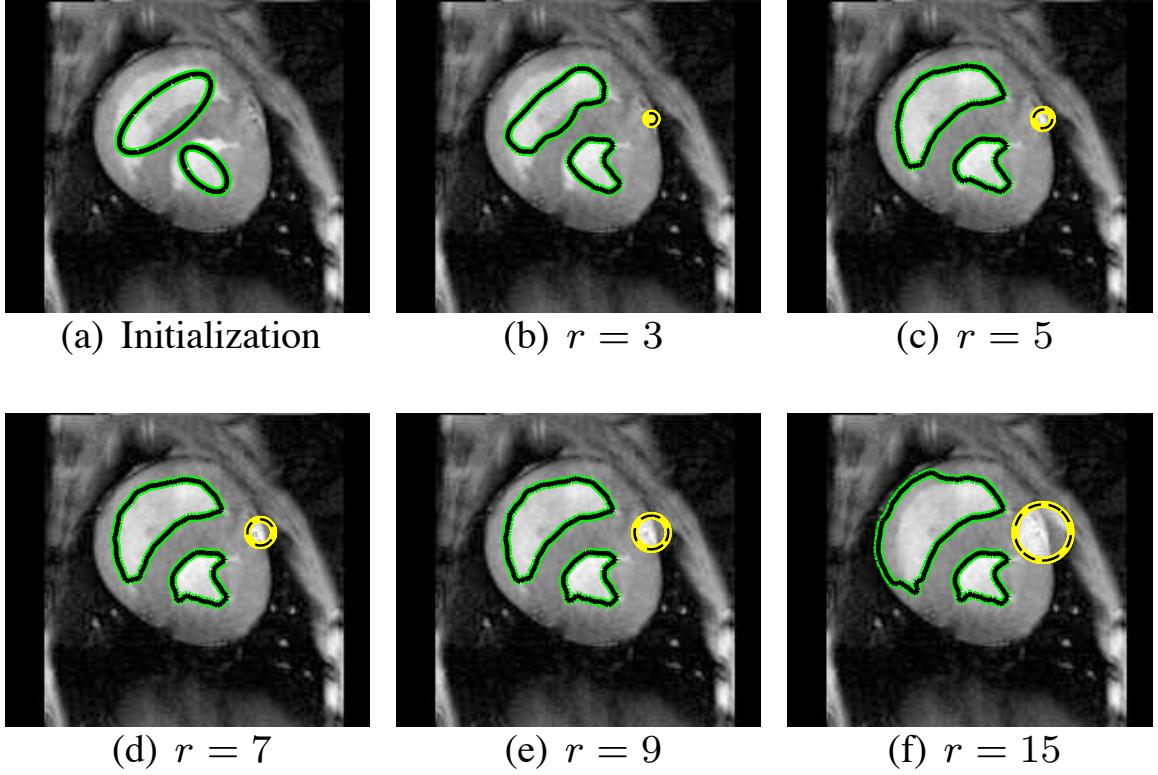
Next, consider the behavior of localized contours when the radius of $\mathcal{B}(x, y)$ is either 1 or infinity. One notices that the ideas of local and global flows are blended together by this technique. If the proposed energy is evaluated with radius of 1, then F becomes, in essence, an edge detector based on the statistics of the pixels immediately adjacent to the center of the ball. On the other hand, if one lets the radius grow until $\mathcal{B}(x, y)$ includes the entire image, then the local region statistics are equal to the global region statistics and are shared by all points in the image. In this case, the behavior will be the same as the original region-based flow. Hence, by tuning the parameter r, one can choose the degree to which local and global behavior are blended. An illustration of this can be seen in Figure 3.9

Figure 3.9: Illustration of how edge-based methods correspond to a $r = 1$ while global region-based methods correspond to $r = \inf$.

An additional experiment, designed to show the effect radius size on a real image is shown in Figure 3.10. The HEART image is segmented five times with different local radii but the same initialization. Here the desired result is to capture the bright ventricle. In the example, the smallest radius size results in an incorrect segmentation that is too local. Each of the three intermediate radius sizes all result in accurate segmentations, and the largest radius size results in an incorrect segmentation that is too global.
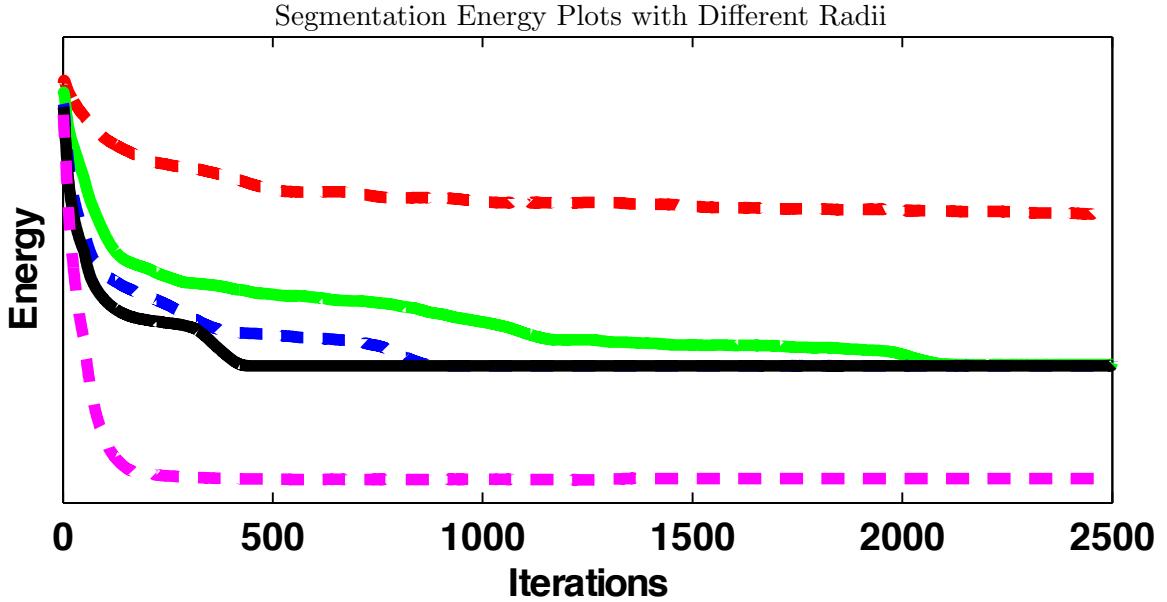
The convergence properties of these five segmentations is also of interest. Figure 3.11 reveals that the speed of convergence is often a function of radius size. The smallest radius

(a) Initialization	(b) $r = 3$	(c) $r = 5$
(d) $r = 7$	(e) $r = 9$	(f) $r = 15$

Figure 3.10: (a) Shows the initialization on the HEART image. (b-f) show the resulting segmentations with mean separation energy using localizing radii of different sizes ($\mathcal{B}(x, y)$ neighborhood shown as the dashed yellow circle in each image).

takes the longest to converge, eventually arriving at an incorrect local minimum. The three intermediate radii converge to the same final energy but at different speeds. The convergence speed is slower when decisions are more localized because the movement of the contour is based on less information. Finally with the largest radius, the segmentation converges quickly to an incorrect energy value that is too global for the task at hand. Thus, this experiment demonstrates the tradeoff between speed of convergence and local radius size and shows that radius sizes that are too big or too small may lead to incorrect segmentations.

Figure 3.12 illustrates two key distances that should be taken into account when choosing the size of the localization radius. Again, the bright ventricle is the intended object to be segmented. Important distances to consider are those between the initialization and the desired boundary (small blue arrows) and those between the desired boundary and nearby clutter (large red arrows). The localization radius should be chosen so that the
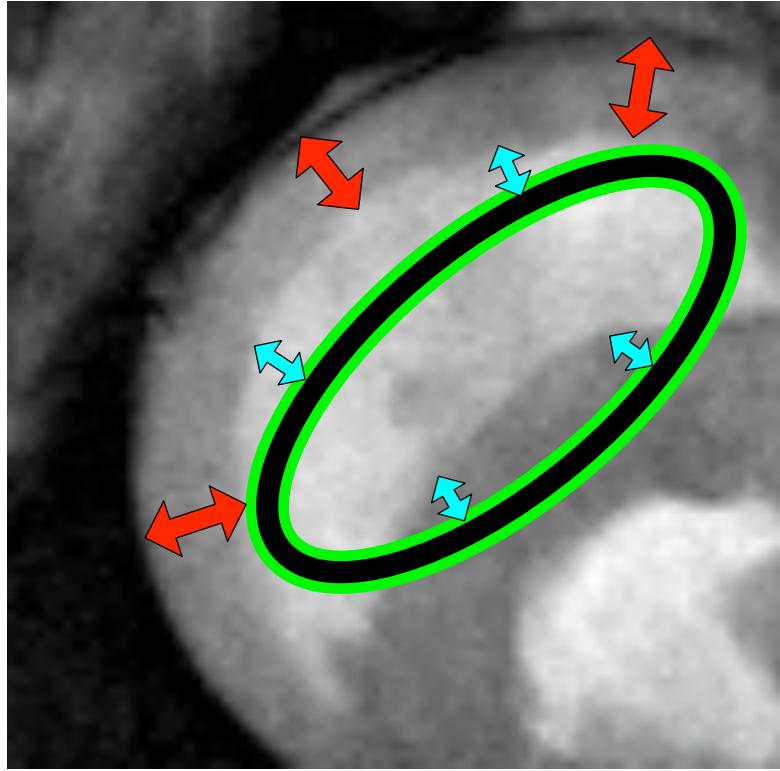
Figure 3.11: Energies for the segmentations in Figure 3.10. Energies shown (from top down) for radius sizes 3, 5, 7, 9, and 15 pixels. The three energies associated with intermediate radii converge to the same correct energy. The other two obtain incorrect results.

local neighborhood is large enough to include the desired boundary, but small enough that nearby clutter does not distract the contour once the desired segmentation is achieved.

3.6.3   Analyzing Sensitivity to Initialization

One limitation of the proposed method is that it has a greater sensitivity to initialization than typical, global region-based methods. This is an inherent trade-off of the proposed localization. The more image data that is analyzed, the more robust the technique is against poor initialization. In other words, global methods will typically be more robust to initialization than local ones. However, analyzing large amounts of image data can lead to erroneous solutions as seen in the previous experiments. Figure 3.13 shows several initializations and the resulting segmentations on a synthetic image of a bimodal box.
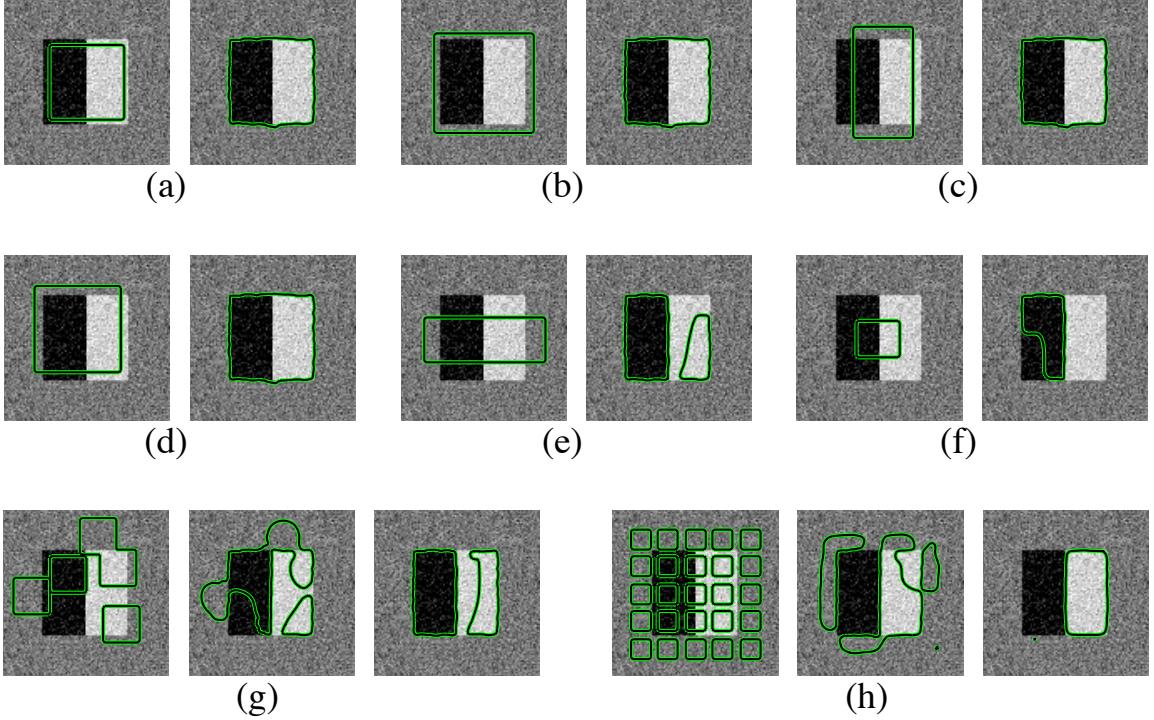
Experiments (a-d) in Figure 3.13 show a correct segmentation obtained from various initializations. Experiments (e-h) show incorrect segmentations resulting from other initializations. These incorrect results are due to two limitations of the technique. First, localized contours may not use enough information. This is shown in (f), where the contour, initialized with a small square, does not grow to fully cover the black half of the box.

Figure 3.12: Important distances to consider when choosing the localizing radius: Small, blue arrows are the distance to the ideal result. Large, red arrows are the distance from the desired result to nearby clutter. An ideal radius size will be between the small, blue arrows and the large, red arrows.

The other major limitation is the "flipping" that may occur. In (h), the contour is initialized with a grid of small squares. Initially the contour finds all the borders of the box, but it finds the *inside* of the white half with the *inside* of the contour and the *inside* of the black half with the *outside* of the contour. Eventually, the part of the contour around the outside of the black half collapses due to a lack of support. Notice that in (g) that a similar effect occurs, but the contour partially recovers.

However, the increased sensitivity of the proposed method does not mean that extreme care must be taken when initializing the active contours. The experiment in Figure 3.14 shows a slice of two MRI brain scans in which the corpus callosum has been correctly segmented with several different initializations. In this experiment, localized analysis allows the segmentation to accurately separate this structure from the rest of the brain despite its thin structure and the presence of nearby structures of similar intensity.
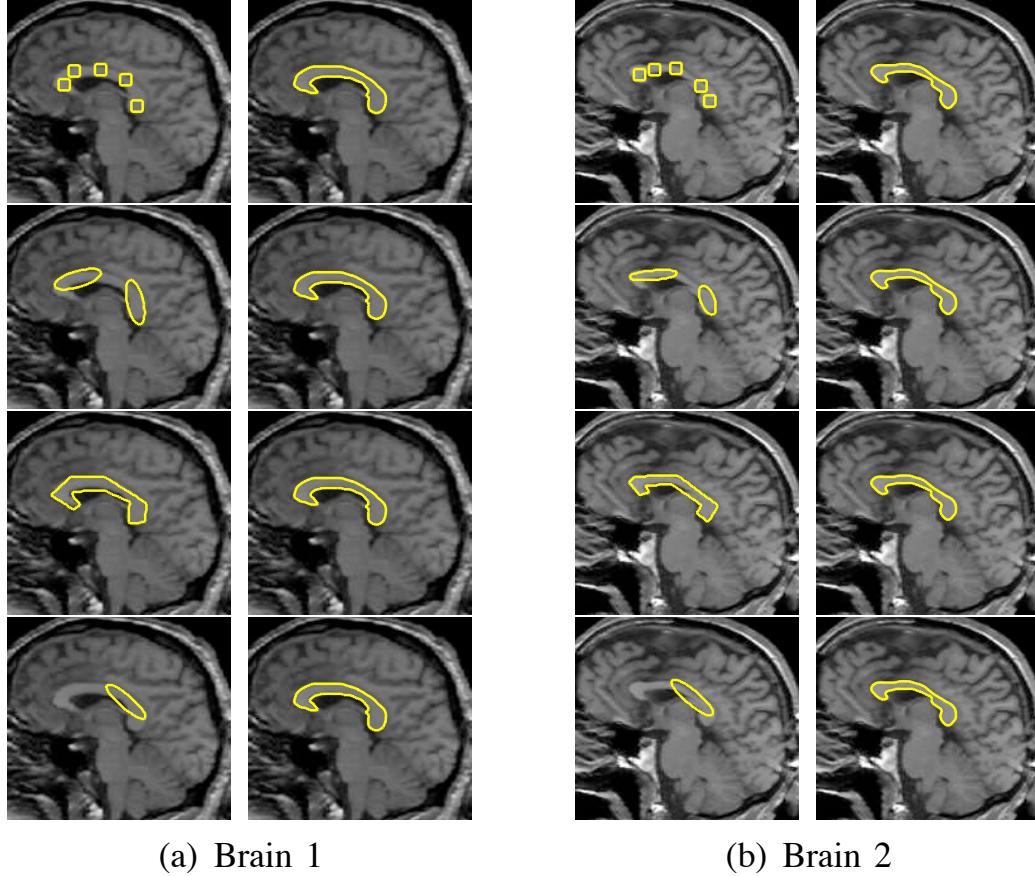
Figure 3.13: Each group shows the initialization and final segmentation with the local mean separation energy. Although the method is robust to some variation in initialization, it is not impervious and should begin relatively close to the desired boundary. (a-d) Correct segmentation, (c-h) incorrect segmentation. Parameter $r = 10$ for each experiment

3.6.4 Computational Efficiency and Convergence

In the proposed method, local region statistics must be computed for each point along the evolving curve. This increases the computational complexity of the algorithm and the processing time beyond that of standard global methods. Nevertheless, this increase is not substantial and can be reduced by using efficient update schemes for the localized statistics. The following describes the efficient computation of local statistics in two parts: initializations and updates.

The proposed local region-based method begins by initializing every pixel in the narrow band by computing local interior and exterior statistics. The nature of this operation varies depending on the statistical measures used. For instance, local means are simpler to compute and compare than local histograms. An additional cost occurs whenever the narrow band moves to include an uninitialized pixel. In this case, the local statistics of this new pixel must be initialized as well. The number of initialization operations performed

(a) Brain 1   (b) Brain 2

Figure 3.14: (a) and (b) show corpus callosums from MRI brain scans of two patients. Each row shows the initialization on the left and the final segmentation with the local mean separation energy on the right. Multiple initializations are shown to demonstrate robustness to initial curve placement. Parameter $r = 6$ for each experiment.

is therefore dependent on the length of the curve and how far it moves from its initial position to its final position. The initialization operation is only performed once for each pixel and therefore adds a constant complexity increase. However, depending on the size of the contour and the local radius, these computations can be significant.

The update step occurs whenever a pixel is crossed by the contour, thus moving it from the interior to the exterior or vice versa. In an ideal implementation local statistical models are kept in memory for every initialized pixel. When the interface crosses a pixel, the statistical models of all initialized pixels within the $\mathcal{B}(x, y)$ neighborhood are updated. For example, when local means are used, each pixel must maintain the number of pixels in the local regions both inside and outside of the curve as well as the sums of pixel intensities in
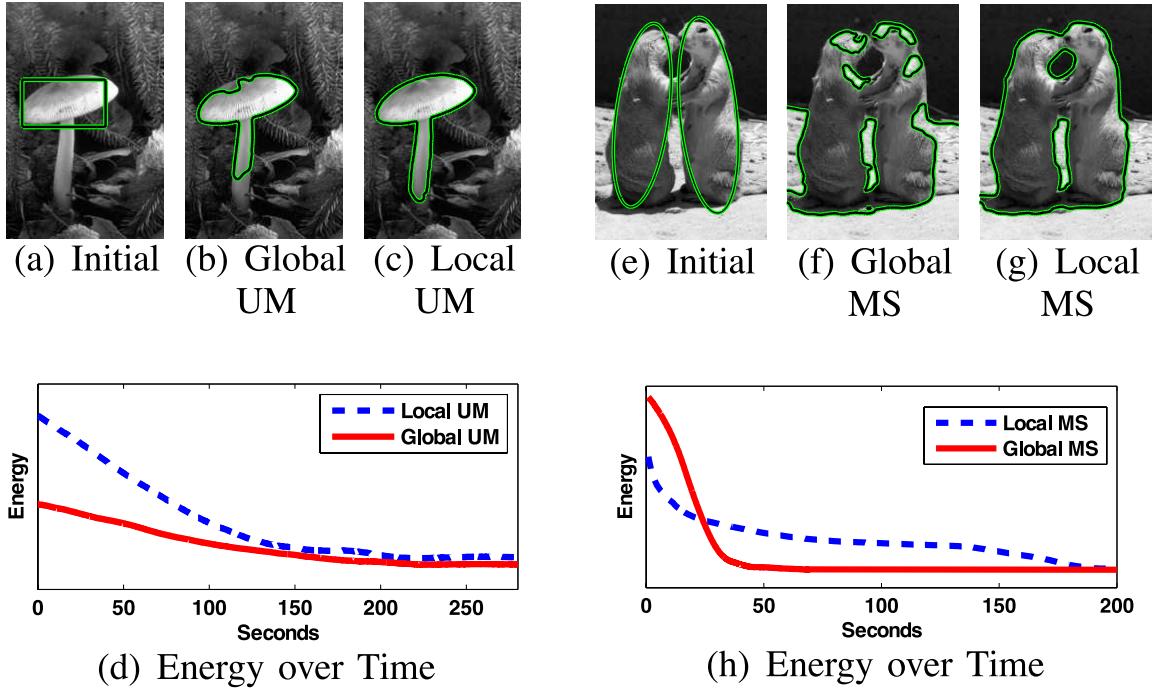
those regions. Updating this model consists of transferring values from the *inside* groups the *outside* groups or vice versa. For the histogram separation energy, a full histogram is kept of the local interior and exterior regions for each initialized pixel. While this requires significantly more memory to maintain than the means model, updates are just as simple: pixel intensities are subtracted from bins of the interior histogram and added to the same bin of the associated exterior histogram or vice versa.

Compared to global methods, local methods incur a linear increase in complexity because of the need to manage all of the local statistics. This increase is proportional to the area of $\mathcal{B}(x, y)$. Assume that at each iteration, m pixels are crossed by the evolving interface, and require an update of their statistics. A global region-based method would perform m statistical updates (one for each pixel), whereas the corresponding local region-based flow would perform $m \cdot n$ updates where n is the number of pixels that exist within the $\mathcal{B}(x, y)$ neighborhood. Experiments confirm this linear increase.

Convergence and Timing: Localized methods have a linear complexity increase over corresponding global methods as stated above. Through experimentation, computation times were found to be in the range of 3 to 5 times longer when the localizing radius is on the order of 10 to 20 pixels. In Figure 3.15, a comparison is given of local and global variants of the uniform modeling energy from Section 3.4.1 and the mean separation energy from Section 3.4.2. Note that in these figures the energies have been scaled to show them on the same graph. The actual energy values converged to are different for local and global methods.

Notice that in the MUSHROOM experiment, the local and global methods both converge in approximately 250 seconds. This is due to each method slowly expanding to capture the stem of the mushroom. Meanwhile, in the HUG image the global version converges to an incorrect result in approximately 60 seconds while the local method takes approximately 180 seconds to correctly segment the image.

As mentioned earlier, all active contour segmentations were implemented using the sparse-field representation [103]. Additional details can be found in Appendix B. In addition

47

(a) Initial   (b) Global   (c) Local   (e) Initial   (f) Global   (g) Local

UM   UM   MS   MS

(d) Energy over Time   (h) Energy over Time

Figure 3.15: (a,c) Initialization, (b,f) final result with global energies, (c,g) final result with local energies, (d,h) Convergence and timing properties of localized method (dashed line) and corresponding global method (solid line).

to the listed experiments, the proposed techniques were implemented in a fast approximate level set framework [57], which eliminates much of the overhead of curve evolution leaving statistical computations as the major computational cost. In this framework execution times on the order of a few seconds were achieved.

3.7   Remarks on Localized Segmentation

In this chapter, a novel framework for localizing region-based active contours was proposed. It was shown that localized active contours out-perform traditional region-based active contours in many cases because localization allows assumptions about the image to be relaxed resulting in more accurate segmentations. Three novel localized energies of this type were introduced, and the steps required to localize any global region-based energy were given. Furthermore, a method to allow these localized energies to interact in order to simultaneously segment multiple objects in a scene was also described.

Multiple experiments were shown that led to three interesting observations. First illustrative examples were shown of situations where global region-based energies failed while

48

localized versions produced reasonable segmentations. Next, experiments with various localization radii demonstrated that radii should be chosen based on the size of salient objects and the proximity of nearby clutter. It was also shown how convergence and timing are affected by radius size. Finally, analysis was performed to identify limitations of the technique including its increased sensitivity to initialization compared to global methods.

As this chapter has shown, localized segmentation energies are useful for segmenting a diverse group of natural and medical images, which may be difficult for existing global region-based segmentation techniques. The ability of localized segmentation methods to capture objects with varying intensities makes them ideal for use in many medical imaging applications. In this chapter, preliminary results were shown on imagery of the brain and heart. Both of these organs consist of large-scale and small-scale structures whose inner-workings are complex and critically important to understand. The ability of localized active contours to segment small-scale structures without being distracted by nearby clutter, and their ability to naturally accommodate structures whose intensities vary spatially can make them particularly well suited for analyzing bundles of neurons within the brain and blood vessels within the heart. Applications in neurology are discussed in Chapter 4 and applications in vessel analysis are presented in Chapter 5.

Furthermore, in the majority of realistic visual tracking scenarios the objects of interest are heterogeneous (e.g. vehicles, people, etc.). This makes frame to frame segmentation a challenging task with existing methods. This proposed segmentation approach in combination with existing contour trackers may allow these algorithms to keep track of an entire object rather than one region of homogeneous intensity. The methodology of localized contours could be improved by detecting and altering the size of the localization radius automatically. Work applying localization to tracking is discussed in Chapter 6.

CHAPTER IV

VOLUMETRIC QUANTIFICATION OF NEURAL PATHWAYS

In this chapter, the localization framework from Chapter 3 is used to develop a method to segment neural fiber bundles in diffusion tensor imaging (DTI)[1]. Different functional areas of the brain are connected by large bundles of neurons. Finding the shape and volume of these pathways is important in understanding brain function and identifying certain pathologies.

DTI data measures the orientation and anisotropy of water diffusion at each point in the brain. Because water is constrained to diffuse along neural fibers, the bundles can be identified in DTI data. However, their local orientation and corresponding DTI data values change drastically as these bundles curve through the brain. Hence a constant global diffusion model leads to inaccurate segmentations.

Accordingly, a method is proposed that uses localized statistics computed on orientation information to drive a variational active contour segmentation. This localized segmentation accurately models the non-homogeneous orientation information that exists along the bundles. Once initialized from a single fiber path, the proposed method can capture the surface shape of the entire bundle. Results of the technique are demonstrated by segmenting an important neural fiber bundle known as the cingulum bundle.

The chapter is briefly outlined as follows. First, in Section 4.1, a literature review is provided along with a background of tractography and fiber bundle segmentation algorithms. Second, Section 4.2 describes the algorithm for locally constraining region-based segmentations on orientation data. Next, Section 6.3 shows segmentation results on the cingulum bundle using the proposed localized active surface model. Finally, Section 4.4 offers concluding remarks and provides an explanation of how these ideas and results may be extended to improve performance in future work and other applications.

---

[1]See Section 1.1.2 for details on DTI.

4.1 Tractography Background and Contributions

Since the advent of diffusion imaging, a great deal of research has been devoted to finding and characterizing neural connections between brain structures. DTI data provides information about the directions that water diffuses at each point in a volume. Hence, it is capable of lending information about the orientation of neural axons. This is because axons, which are essentially long narrow tubes, restrict diffusion perpendicular to the axon and only allow diffusion along it. By analyzing DTI data, therefore, one can deduce where individual neurons, or groups of neurons, exist. Currently, image resolution is high enough that major white matter tracts, or bundles of densely packed axons, are several voxels in cross-sectional diameter [67]. The goal of tractography is to segment these fiber bundles from the DTI datasets.

High angular resolution diffusion imaging data typically consists of many different directional measurements at each point in the imaged volume. While some methods make use of the measurements directly, others first model the measurement using a mathematical object known as a tensor [9]. The diffusion tensor is one of the simplest diffusion models. It is estimated from multiple diffusion measurements to reduce computational and representational complexity. In the three-dimensional case the diffusion tensor is a 3×3 symmetric, positive definite matrix representing the orientation, magnitude, and anisotropy of the underlying diffusion data. For a detailed description of tensor modeling see [9].

Finding Single-Neuron Paths: Early tractography methods were based on streamlines, which employed local decision-making based on the principal eigenvector of diffusion tensors [8, 21, 66, 98]. In these techniques, tracts are propagated from a starting point until they reach some termination criterion. Due to the local decision-making process, these methods have been shown to perform poorly in noise and often stop prematurely. These techniques also fail to provide a measure of connectivity for the resulting tracts. Despite the shortcomings of this approach, streamlines have become the most popular method for fiber segmentation.

To infer volumetric information about fiber bundles identified from streamline tractography results, several methods have been proposed for clustering the detected fibers. The goal of clustering is to capture a group of streamlines and use their collective behavior to drive segmentation of the entire bundle. The end result of clustering algorithms accurately captures many neural fiber bundles [64, 71].

Recently, another line of work has emerged that seeks to avoid the use of the problematic streamlines. Advances in tractography have been made that allow extraction of optimal connectivity maps that connect predefined seed regions. These methods are more robust to noise and may be able to make use of all diffusion measurements rather than the simplified tensors. These approaches can be subdivided into stochastic and energy-minimization approaches.

Stochastic approaches produce probability maps of connectivity between a seed region and the rest of the brain. Parker *et al.* developed a probabilistic index for standard streamline techniques [78]. Perrin *et al.* presented probabilistic techniques for untangling fiber crossings [81]. While they do provide a measure of connectivity between brain regions, these stochastic approaches do not provide an explicit segmentation of the fiber bundle itself and often fail to provide the optimal connection between regions of the brain.

Energy-minimization techniques address some of these problems. Parker *et al.* proposed fast marching tractography, which minimizes an energy based on the position and direction of a propagating front [79]. O'Donnell *et al.* cast the tractography problem in a geometric framework by finding geodesics on diffusion tensor manifolds [74]. Similarly, Prados *et al.* and Lenglet *et al.* demonstrated a Riemannian based technique, for computing geodesics using a variant of Fast Marching adapted for directional flows [53, 86]. Jackowski *et al.* find geodesics using Fast Sweeping as given by Kao *et al.* [36, 41, 42]. Pichon *et al.* and Melonakos *et al.* use the more general Finsler metric to find optimal connections between brain regions [60, 61, 82, 83]. Finally, Fletcher *et al.* demonstrated solving Hamilton-Jacobi-Bellman systems on the graphics processing unit to find geodesics in near real-time speeds [28]. In each of these methods, an optimal path is found that represents an optimal single-fiber path between two regions under the given metric.

Segmenting Bundles of Adjacent Neurons: A number of surface evolution approaches, which are briefly reviewed here, have been described in the literature for fiber bundle segmentation. Rousson *et al.* [92] use a multi-variate Gaussian distribution of the tensor components in a geodesic active region model to drive a surface evolution towards the segmentation of fiber bundles. The method is applied to the segmentation of the corpus callosum, but is unable to fully capture its curved shape as discussed by the authors. In a follow-up paper [52] a similar segmentation framework is used that yields superior segmentation results, especially when segmenting curved fiber bundles. Rathi *et al.* [88] demonstrated modeling the entire tensor distribution to capture inhomogeneity, and Malcolm *et al.* [56] extended the technique to segmentation using graph cuts. Additionally, Awate *et al.* [6] segmented curved fiber bundles using Markov random fields.

Jonasson *et al.* proposed two different ways to address the segmentation of curved fiber bundles in a surface evolution setting. First, a local approach is used in [38]. The surface evolution speed is influenced by the similarity of a tensor in comparison to its interior neighbors. Secondly, a region-based approach was proposed that used a similarity measure based on a single representative tensor within the segmented region [39]. In the latter case, capturing highly curved fiber bundles will be problematic because similarity is based on a single representative tensor. The approach proposed in this paper is similar to the work in [38] insofar as it uses local tensor similarities to drive the segmentation. However, Jonasson *et al.* use only a few adjacent pixels to determine local statistics. The present approach allows for larger local regions of pixels both inside and outside the evolving surface to compete thus making the technique more robust to noise and initialization.

Contributions: The work presented in this chapter is, in part, an extension of [62] where the authors perform segmentation of the cingulum bundle using a greedy flood-fill algorithm to find the tensor bundle from an initial anchor tract. In this previous work, an anatomical prior based on distance to the anchor tract was used to prevent leaks into surrounding structures. By leveraging the power of localized active contours, the following improvements are made by the present work:

1. A fully variational active contour framework is used

2. Masking based on anatomical information is no longer needed

3. Tensor information is considered in a more principled way

4.2 *Localized Statistics in the Tensor Domain*
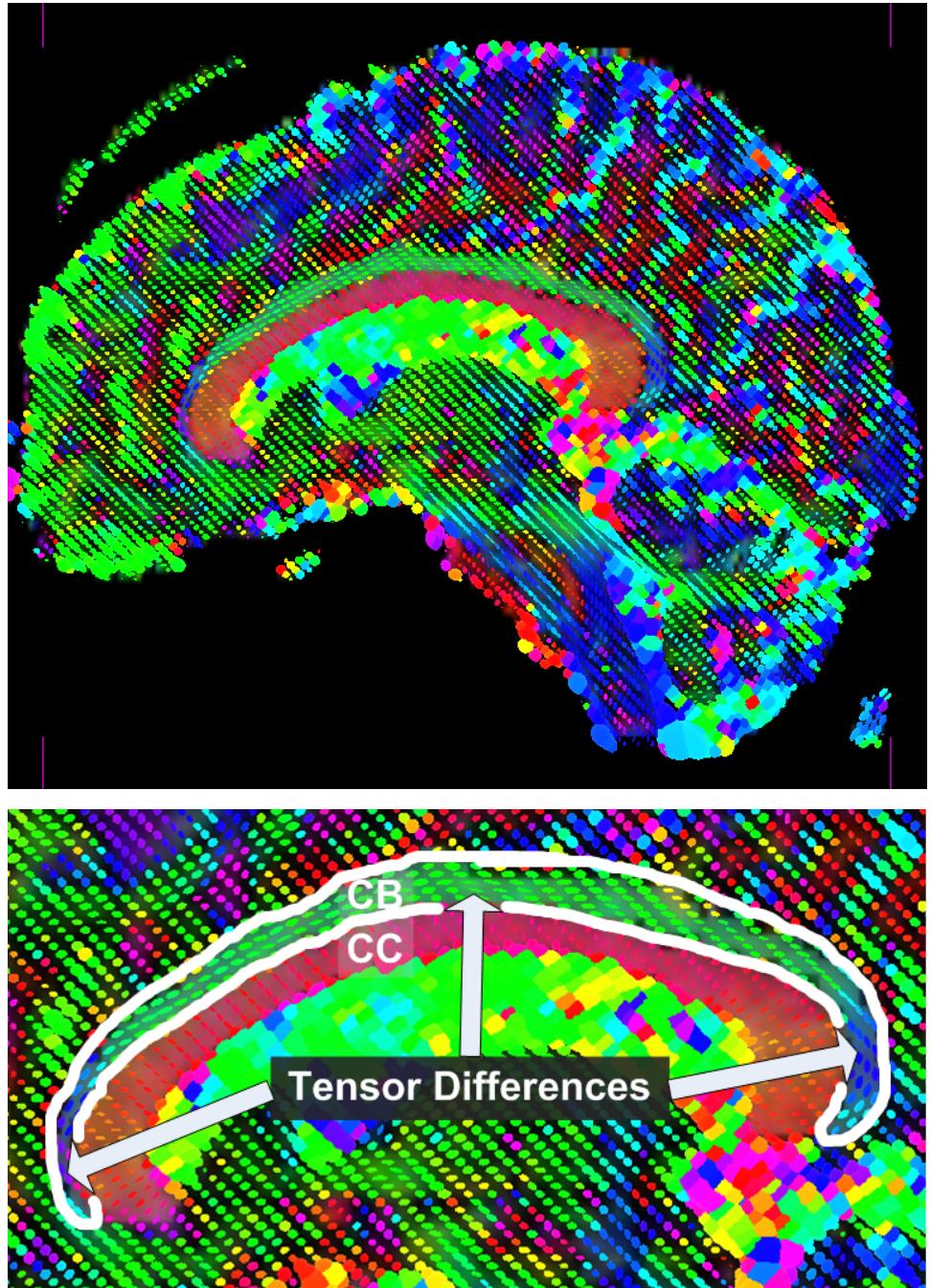
An implicit assumption of classical region-based approaches is that the interior region contains nearly homogeneous features. Note that this need not mean that the data is homogeneous, just that a statistical measure can be constructed that effectively separates the object of interest from the rest of the image. Under this assumption, segmentation algorithms proceed by evolving the closed curve or surface to minimize an energy defined over these global features.

If features within the object are heterogeneous across the object of interest or similar to the exterior it becomes difficult to define a global region-based approach that will accurately segment the image. Neural fiber bundles are a perfect example of this scenario. For instance, in the case of the cingulum bundle, which curves strongly, the tensors along the bundle vary in orientation along the entire length, as shown in Figure 4.1. In this sagittal view, one can see that it is difficult to define a feature on the space of tensors that uniquely separates the entire interior of the cingulum bundle from the exterior. This is due to two reasons:
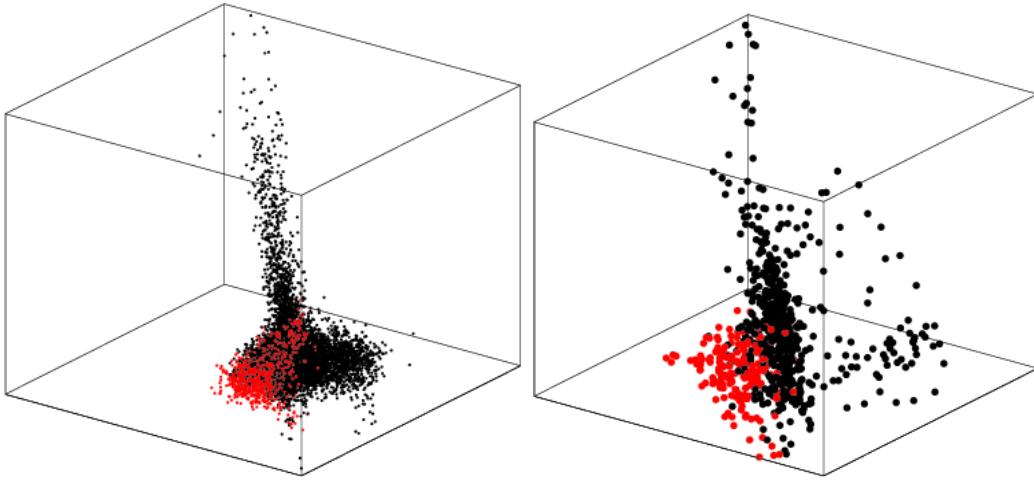
1. The tensors change drastically along the length of the bundle

2. Tensor statistics inside the bundle are also observed outside the bundle

However, one notices that the tensor orientation and anisotropy vary smoothly along the bundle and that the background is *locally* different from the interior at every point. Hence, by looking locally along the bundle one can define tensor features that are distinguishable from the exterior. In Figure 4.2 this is illustrated by examining tensors globally as well as within a particular local region. Notice how interior and exterior points are clustered more strongly when a local region is examined instead of examining the entire image domain.

The basis of the proposed technique is this key assumption: tensors in of the interior of the bundle will be statistically different from tensors in the exterior of the bundle if

Figure 4.1: Example of the need for local constraints on region-based segmentation algorithms that attempt to segment the cingulum bundle. Notice that tensor anisotropy and orientation vary across the length of the cingulum bundle. In these images, tensor color corresponds to orientation and roundness corresponds to anisotropy.

(a) Interiors and exteriors using global statistics (b) Interiors and exteriors using local statistics
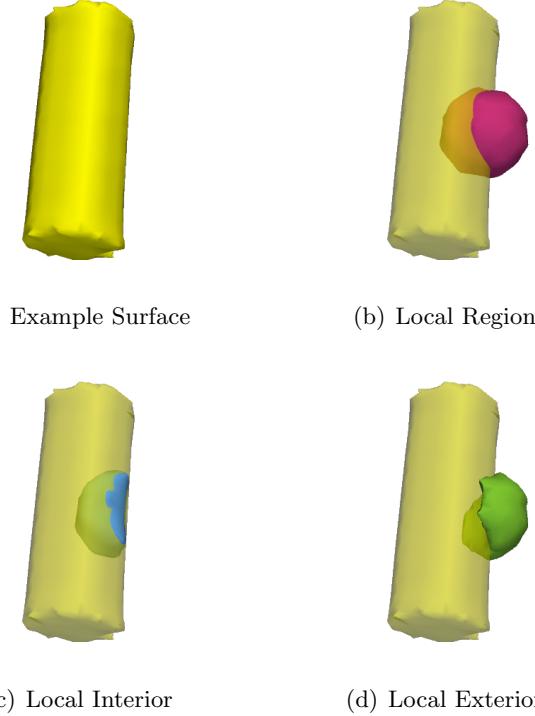
Figure 4.2: A visualization of the high-dimensional tensors in 3D by projection onto a unit vector. Tensors from the interior of the segmented fiber bundle are shown in *red*, and tensors from the exterior are shown in *black*. (a) Globally, interior and exterior tensors are mixed and not easily separable. (b) When a local region centered on a point on the surface of the fiber bundle is examined, interior and exterior tensors appear separable.

examined locally. In order to cope with the non-homogeneous nature of fiber bundles and the surrounding brain structures, standard global statistical modeling techniques are re-formulated in terms of localized statistics. First, the localized statistics are described and then it is shown how they can be incorporated into a single energy functional that is minimized with a fully variational level set energy minimization framework.

Let $\mathbf{T}$ denote a volume of tensors, normalized such that $\|\mathbf{T}(x)\| = 1$ for every x in the domain Ω_x. Also let Γ be a closed surface represented as the zero level set of a signed distance function ϕ as described in Chapter 2.

In order to describe localized statistics in the tensor domain, the characteristic ball function from Equation (3.3) is used with the Heaviside function and Dirac functions from Equation (3.1) and Equation (3.2). Rendered 3-D views of localized interior and exterior regions are shown as (c) and (d) in Figure 4.3. These representations are helpful in defining local statistics on the orientation of the tensors so that they can be easily incorporated into an active surface optimization model.

Unlike the segmentation examples in Chapter 3 that operated on intensity data, the

(a) Example Surface   (b) Local Region

(c) Local Interior   (d) Local Exterior

Figure 4.3: Visualization of the local regions over which tensor statistics are computed in order to segment fiber bundles with heterogeneous statistics.

present segmentation energy operates on tensors, which do not lie in a linear vector space, but instead lie on a conical manifold. In order to simplify tensor arithmetic, a log-Euclidean mapping is used to place these tensors in a linear vector space [4]. Then, standard linear arithmetic can be used to compute distances and calculate statistics such the mean. In this transformation one computes the matrix logarithm of the 3×3 positive definite symmetric tensor. The logarithm maps the tensors onto the log-Euclidean tangent plane where linear operations may be performed directly. The inverse mapping to return the conical manifold is a matrix exponential. Thus, in this space, the integral of a group of tensors, $\mathbf{T}(x)$ is computed simply as

$$\exp\left(\int_x \log(\mathbf{T}(x)) \ dx\right) \tag{4.1}$$

Therefore, the localized interior and exterior tensor means are defined as

$$\mu_{\mathrm{in}}(x) = \exp\left(\frac{\int_{\Omega_y} \mathcal{B}(x,y) \ \mathcal{H}\phi(y) \ \log\left(\mathbf{T}(y)\right) dy}{\int_{\Omega_y} \mathcal{B}(x,y) \ \mathcal{H}\phi(y) dy}\right), \tag{4.2}$$

57

and

$$\mu_{\text{out}}(x) = \exp\left(\frac{\int_{\Omega_y} \mathcal{B}(x,y)\ (1 - \mathcal{H}\phi(y))\ \log\left(\mathbf{T}(y)\right) dy}{\int_{\Omega_y} \mathcal{B}(x,y)\ (1 - \mathcal{H}\phi(y)) dy}\right), \tag{4.3}$$

respectively. One can also define a log-Euclidean distance,

$$d_{LE}[\mathbf{T}_1, \mathbf{T}_2] = \|\log(\mathbf{T}_1) - \log(\mathbf{T}_2)\|^2, \tag{4.4}$$

using a standard L2 norm after mapping the tensor data onto a linear space.

With the localized statistics well defined, an energy functional is now proposed with which one can deform Γ, embedded in ϕ. This results in a segmentation based on modeling local regions by $\mu_{\text{in}}(x)$ and $\mu_{\text{out}}(x)$. This energy is defined in terms of the signed distance function, ϕ as before

$$\begin{aligned} E(\phi) &= \int_{\Omega_x} \delta\phi(x) \int_{\Omega_y} \mathcal{B}(x,y)\ F(\mathbf{T}, \phi, x, y) dy\ dx \\ &+ \lambda \int_{\Omega_x} \delta\phi(x) \|\nabla\phi(x)\| dx. \end{aligned} \tag{4.5}$$

where F is defined as

$$F = \mathcal{H}\phi(y)\ d_{LE}[\mathbf{T}(y), u_x] + (1 - \mathcal{H}\phi(y))\ d_{LE}[\mathbf{T}(y), v_x]. \tag{4.6}$$

The gradient descent curvature flow that will minimize E is then computed using the calculus of variations. This reveals the final evolution equation:

$$\begin{aligned} \frac{\partial\phi}{\partial t}(x) &= \delta\phi(x) \int_{\Omega_y} \mathcal{B}(x,y)\ \left(d_{LE}[\mathbf{T}(y), u_x] - d_{LE}[\mathbf{T}(y), v_x]\right)\ dy\vec{N} \\ &+ \lambda\delta\phi(x)\ \text{div}\left(\frac{\nabla\phi(x)}{|\nabla\phi(x)|}\right) \|\nabla\phi(x)\|. \end{aligned} \tag{4.7}$$

By evolving ϕ with this equation, the surface Γ is deformed so that it moves from its initialization to an optimal segmentation of the fiber bundle. The efficient level set implementation of [103] is employed to obtain fast and accurate results from the level set evolution in three dimensions.

The algorithm is initialized with a pre-computed anchor tract, representing the lowest cost path connecting two maximally spaced-out, pre-defined regions of interest on the fiber bundle to be segmented (the cingulum bundle). This anchor tract is computed using the

technique described in [60], but any similar tractography method could be substituted. The result is a single-pixel anchor tract that is then dilated using a 5 mm × 5 mm × 5 mm ball-shaped structuring element. This produces an initial surface with definite interior and exterior regions.

4.3 Segmentation Results on the Cingulum Bundle

This section presents experiments on the cingulum bundle (CB). First, a description is given of the medical importance of segmenting this particular structure and what makes that segmentation difficult. Subsequently, results of several experiments are shown that exemplify the favorable results obtained by this algorithm.

Cingulum Bundle Description and Motivation: Here, the problem of segmenting the cingulum bundle is motivated. The cingulum bundles in either hemisphere loop around through the middle of the brain (front to back) forming a pathway by which many different functional areas communicate. Specifically, the cingulum bundle is a 5-7 mm in diameter neural pathway that interconnects all parts of the limbic system. It originates within the white matter of the temporal pole, and runs posterior and superior into the parietal lobe, then turns, forming a "ring-like belt" around the corpus callosum. Finally, it continues the turn into the frontal lobe and terminates anterior and inferior to the genu of the corpus callosum in the orbital-frontal cortex [95].

The cingulum bundle consists of long association fibers that directly connect temporal and frontal lobes as well as shorter fibers radiating into various other parts of the brain. The cingulum bundle also includes most afferent and efferent cortical connections of cingulate cortex including those of prefrontal, parietal and temporal areas, and the thalamostriatae bundle.
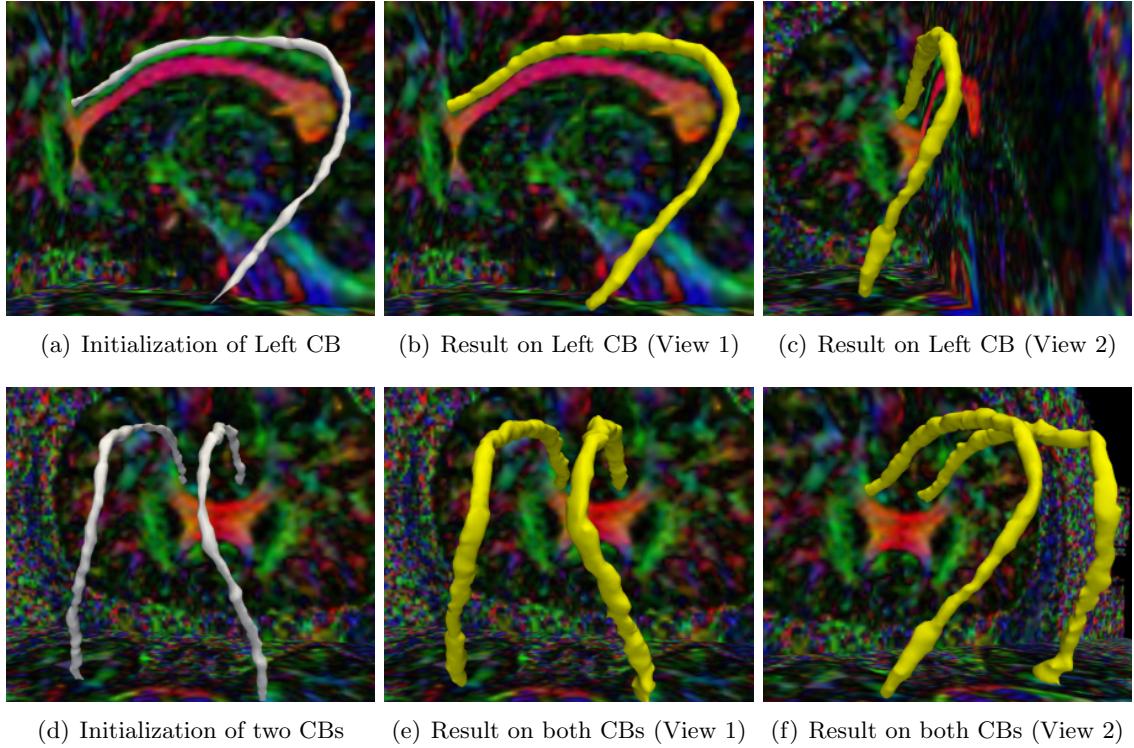
Lesion studies have documented a variety of neurobehavioral deficits resulting from lesions located on the cingulum bundle, including akinetic mutism, apathy, transient motor aphasia, emotional disturbances, attentional deficits, motor activation, and memory deficits. Because of its involvement in executive control and emotional processing, the

cingulum bundle has been investigated in several clinical populations, including depression and schizophrenia. Previous studies of schizophrenia, using diffusion tensor imagery, demonstrated that a decrease of fractional anisotropy in the anterior part of the cingulum bundle may give some diagnostic value [46, 102]. The studies were ultimately inconclusive, pointing to technical limitations that restricted the investigators from analyzing the entire fiber tract.

Experiments: Experiments were conducted on several DTI volumes of normal brains. Each dataset contains seed and target points drawn by experts, which were used to extract a single initial fiber tract within the bundle. From this initialization, the proposed technique is used to determine the surface segmentation of the entire cingulum bundle.

There is one key parameter used in this technique: the localization radius r, which describes the formation of the localized statistics that drive the segmentation. This parameter was chosen based on anatomical knowledge that the diameter size of a tensor bundle is at most 7 mm [95]. Hence, r was set at $r = 7$ mm to ensure that the entire bundle would be included in statistical computations. The other parameter used in this technique is the weighting coefficient λ in Equation (4.5). This parameter determines the intrinsic smoothness of the surface. Because the surface is necessarily high in curvature due to its fine structure, This parameter was set to $\lambda = 0.001 \max(|\frac{d\phi}{dt}|)|_{t=0}$ so that it would have little impact on the resulting segmentation. The values of λ and r were the only parameters used in the technique and were kept constant throughout all experiments.

Figure 4.4 shows the segmentation of the left cingulum bundle as well as the left and right cingulum bundles together. The initialization consists of a single fiber shown as a thin white volume. The final segmentations are shown as thicker yellow volumes. Two views of the final segmentations are provided. Using the proposed method, the bundle is segmented without leaking and captures the changing properties of the bundle as it bends around the corpus callosum. This is accomplished without any explicit shape prior to defend against leaks. Instead, leaks are avoided naturally because the localized statistical analysis is capable of resolving the boundaries of the cingulum bundle. Furthermore, the smoothness

(a) Initialization of Left CB   (b) Result on Left CB (View 1)   (c) Result on Left CB (View 2)

(d) Initialization of two CBs   (e) Result on both CBs (View 1)   (f) Result on both CBs (View 2)

Figure 4.4: Top Row: Segmentation of left cingulum bundle. (a) Shows initialization from single anchor tract. (b,c) Show multiple views of the final segmentation. Bottom Row: Segmentation of left and right cingulum bundles together. (d) Initialization from single tracts. (e,f) Multiple views of the final segmentation.

of the detected surface is ensured by the intrinsic properties of the evolving active surface.

4.4   Remarks on Localized Tensor Analysis

A novel segmentation method for diffusion tensor images was presented that works by utilizing the localized active contour model. The approach extends the original framework presented in Chapter 3 by incorporating tensor-specific mathematics. This fully variational surface evolution scheme is based on the local uniform modeling energy. The localization of the statistics allows it to capture directionally varying fiber bundles (specifically the cingulum bundle) simply and robustly. This is especially useful as many fiber bundles in the brain curve strongly (e.g., the cingulum bundle, arcuate fasciculus, and corpus callosum). Furthermore, the localized surface evolution models were shown to be able to discriminate regions and prevent leaks without the need for shape or distance priors that effectively mask off unwanted parts of the segmentation.

Several other extensions are possible, including the use of more sophisticated distance and similarity measures. The proposed localized method as well as global region-based segmentation methods will benefit from similarity metrics using complete diffusion information instead of the simplified tensor representation.

CHAPTER V

SOFT PLAQUE DETECTION IN CORONARY ARTERIES

As seen in previous chapters, localized active contour and active surface segmentation models are capable of segmenting many challenging images including natural images, medical images in 2-D, and 3-D orientation-based brain scans. Let us now consider the analysis of heart imagery. Specifically, this chapter considers how localized analysis can be used to accomplish segmentation and detection tasks related to assessing cardiovascular health of the coronary arteries.

First, the problem of segmenting vessels from computed tomography angiography (CTA) data[1] will be considered. Then using that segmentation as a starting point, a procedure to detect and measure non-calcified plaques (also known as *soft plaques*) will be demonstrated. These potentially dangerous lesions are known to be a major risk factor for cardiovascular disease, and it is likely that detecting them will improve physicians' ability to predict cardiac events. This is a particularly challenging problem in CTA imagery because plaques may appear similar to nearby blood and muscle tissue.
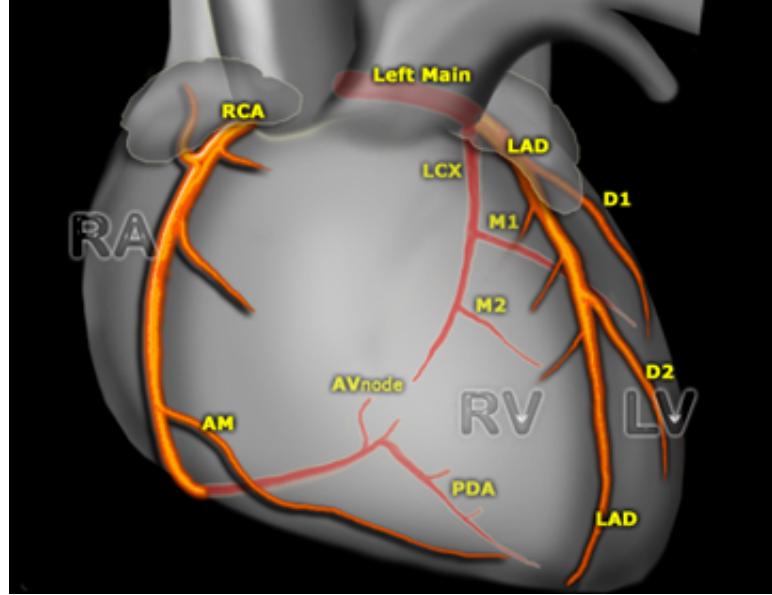
The proposed method for detecting soft plaques in CTA imagery identifies plaques that exist within the vessel wall by simultaneously segmenting the vessel from the inside-out and the outside-in using carefully chosen localized energies. This allows the complex appearances of plaques and vessels to be modeled with simple statistics. This method is shown to be an effective way to detect the minute variations that distinguish plaques from healthy tissue. Experiments demonstrating the effectiveness of the algorithm are performed on eight datasets, and results are compared with detections provided by an expert cardiologist.

The segmentation and detection algorithms are described in detail in Sections 5.2 through 5.5. In Section 5.6, experiments are shown on several vessel trees including the right coronary artery (RCA), left circumflex (LCX), and left anterior descending (LAD).

---

[1]See Section 1.1.2 for details on CTA imagery.

Illustrations of these arteries can be seen in Figure 5.1. The results are then compared with detections made by an expert cardiologist to validate the method. Finally, conclusions and directions for future work are given in Section 5.7.

Figure 5.1: This illustration shows the major arteries of the heart including: The right coronary artery (RCA), left anterior descending (LAD), left circumflex (LCX). Additionally, the right atrium (RA) is labeled as well as the right ventricle (RV) and left ventricle (LV).

5.1 Computational Cardiology Background and Contributions

Heart disease continues to be the leading cause of death in western countries across all ethnic groups. A number of recent studies have demonstrated that the presence of vascular plaques can be a significant indicator of risk for cardiac events [96]. The advent of multi-detector computed tomography (MDCT) and computed tomography angiography (CTA) allows practitioners and researchers to non-invasively image these plaques, which can be categorized as either *calcified*, *non-calcified*, or *mixed* based on their composition. Calcified plaques are easily discernible due to their high X-ray attenuation and corresponding bright appearance in computed tomography (CT) imagery. Consequently, several techniques have been presented to automatically detect calcified plaques with reasonable accuracy [16, 94]. Alternatively, non-calcified plaques, also known as *soft plaques*, have a CT attenuation similar to blood and myocardial tissue making them difficult to detect, even for trained
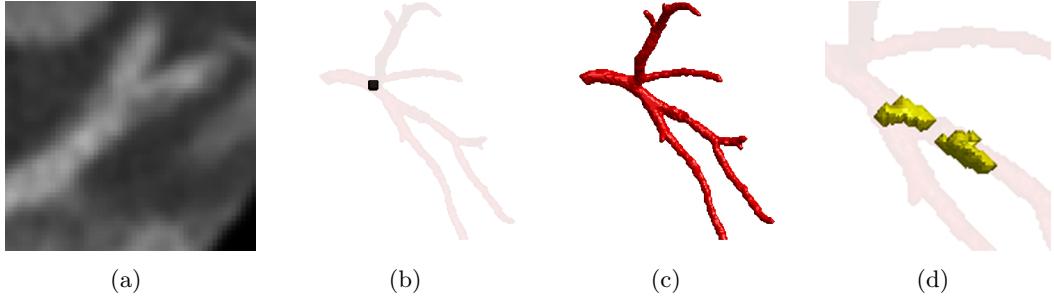
experts [1, 85]. Detection and segmentation of soft plaques is essential given that non-calcified plaques are much more likely than calcified plaques to rupture and cause a variety of acute coronary syndromes [101].

Currently, there is a paucity of published literature on automatic soft plaque detection in CTA imagery. The one recent paper explicitly addressing this problem by Renard and Yang [89] also uses localized information to segment the interior and exterior of the vessel wall. The method in [89] requires pre-and post processing of the volume, describes no way to detect plaques in branching vessels, and uses a simple adaptive thresholding scheme for segmentation. By contrast, the method proposed here recasts the problem in a fully-variational active contour framework that operates directly on the raw imagery, handles branching vessels naturally, and benefits from the geometric properties of active contour segmentations.

Indeed, this chapter presents an effective and straightforward technique for automatically detecting soft plaques in CTA imagery based on multiple segmentations of the vessel wall. The segmentations are performed using multiple active contours driven by spatially localized statistical models that allow the complex appearances of plaques and vessels to be described with simple statistics. First, the vessel tree is extracted using a single, user-provided initialization point. Next two surfaces are constructed that lie just inside and just outside the vessel wall. An active contour model driven by a localized energy designed for plaque detection is then employed to simultaneously segment the interior and exterior of the vessel wall. Finally, areas where these two segmentations do not match are identified as potential regions of non-calcified plaque. Figure 5.2 shows an example of the imagery used, the initialization, the segmented vessel, and the detected plaque.

5.2 Overview of Vessel Analysis Algorithms
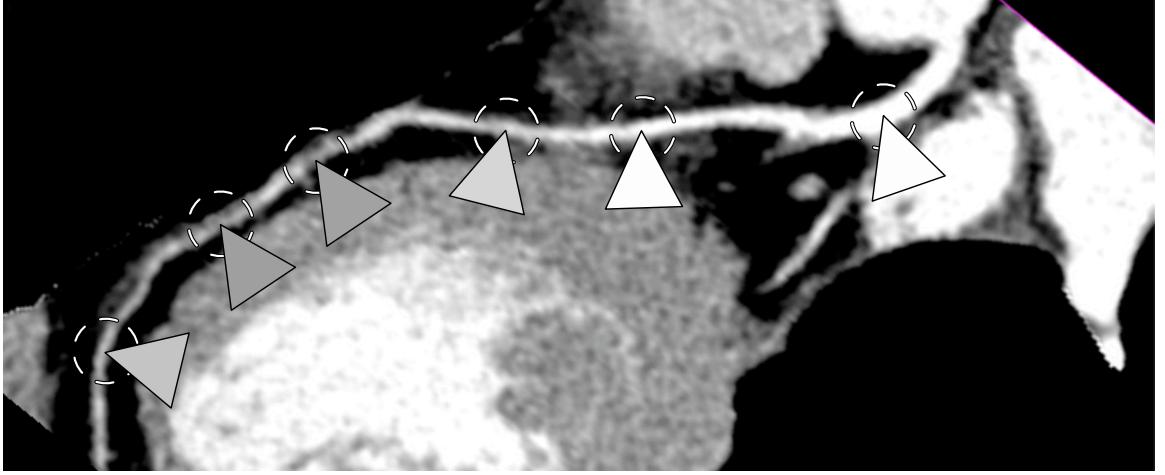
The algorithms presented throughout this chapter make extensive use of localized active contours as described in Chapter 3. This technique allows region-based segmentation energies to be spatially localized such that statistical models of the foreground and background adapt to image information as it changes over the domain of the image. This is particularly

(a)	(b)	(c)	(d)

Figure 5.2: Using CTA Imagery shown in (a) and a single-point initialization shown as the dark spot in (b), the vessel is segmented (c) and used to detect non-calcified plaques (d) in the vessel tree.

powerful for segmenting vessels, which often exhibit changing image intensities over their length as seen in Figure 5.3. Localized active contours are also used for the identification of non-calcified plaques, which typically have only slight intensity differences from the adjacent vessels and myocardium. To segment the vessels the uniform modeling energy from Section 3.4.1 is used while the mean separation energy in Section 3.4.2, is used for detection of plaques.

More specifically, for vessel segmentation, the *local uniform modeling* energy, presented in Section 3.4.1 is employed. With this choice of internal energy, a minimum is found when local interior and exterior regions are well characterized by $\mu_{\text{in}}(x)$ and $\mu_{\text{out}}(x)$, respectively.

Figure 5.3: A view of the LAD in a CTA volume re-oriented to be shown in a single slice. Shaded triangles show the mean intensity of the vessel within the dashed circles. It is clear that over the length of the vessel, the local mean intensity changes dramatically.

66

This energy is ideal for vessel segmentation because it allows the surface to expand into areas of similar local intensity as long as a larger difference exists between local interiors and exteriors. This allows rapid segmentation of vessels despite changing intensities along the length of the vessel.

For soft plaque detection, the *local mean separation* interior energy presented in Section 3.4.2 is utilized. This energy is minimized when the difference between $\mu_{\text{in}}(x)$ and $\mu_{\text{out}}(x)$ is maximized. Local means separation is well suited for detection because expansion into nearby regions that have slightly different intensities is discouraged, even if the local means are similar. This more stringent constraint is quite valuable when attempting to differentiate between vascular plaques and surrounding tissue.

5.3 Coronary Artery Segmentation Algorithm

The first step in detecting soft plaques is to segment a given vessel tree by minimizing the localized uniform modeling internal energy presented in Section 3.4.1:

$$F_{\text{UM}} = \mathcal{H}\phi(y)(I(y) - \mu_{\text{in}}(x))^2 + (1 - \mathcal{H}\phi(y))(I(y) - \mu_{\text{out}}(x))^2 dy. \tag{5.1}$$

This energy is particularly powerful for segmenting vessels. By looking locally, one is able to accommodate variations in intensities that occur over the length of vessels while also remaining sensitive to slight intensity differences that separate vessels from adjacent structures. Hence, the entire vessel can be captured from a simple initialization while at the same time preventing leaks into nearby heart chambers that may contain contrasted blood. This localized active contour approach also responds naturally to branches in the vessel, capturing them automatically without any special schemes to detect their presence.
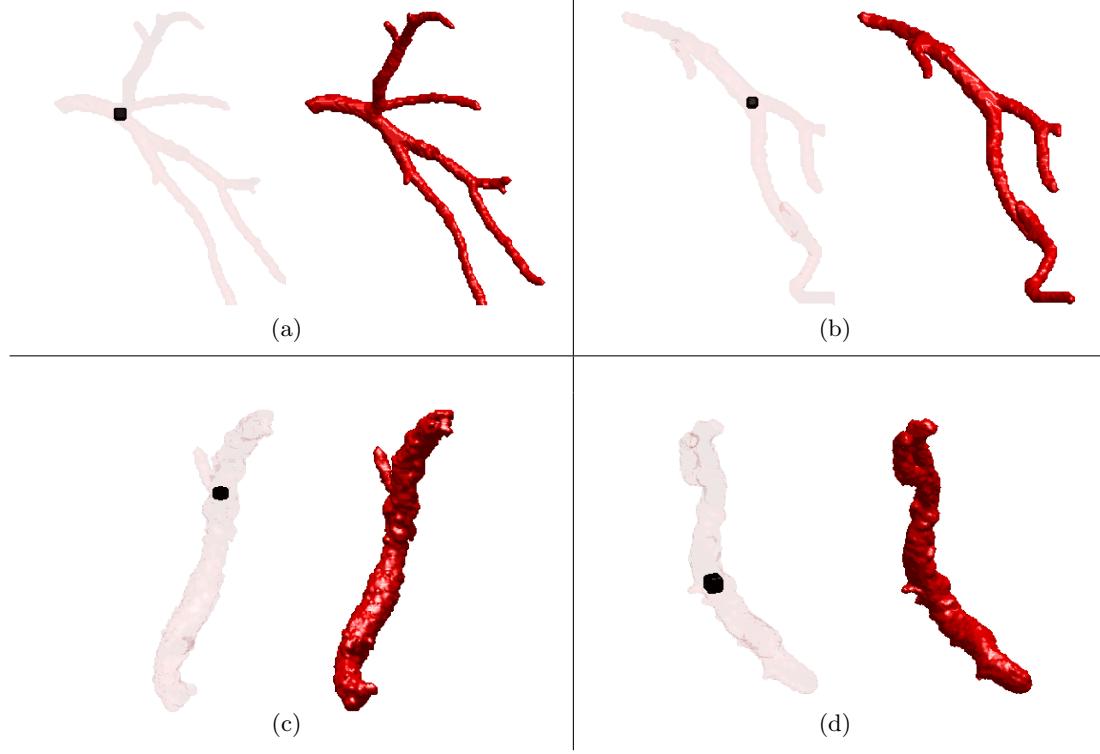
During vessel extraction and plaque detection a restricted domain $\tilde{\Omega}$ is used. The restricted domain only includes voxels[2] representing the heart and ignores dark voxels representing air present in the lungs. $\tilde{\Omega}$ is formed by excluding from Ω any points that fall below a threshold of -600 Houndsfield units (HU). This threshold is significantly below the ranges for blood, plaques, and myocardial tissues in all CTA data. Using a single-point

---

[2]Voxels are 3-D equivalents of pixels in 2-D images. Each voxel represents a data-point in a 3-D image.

initialization within the vessel, the entire vessel structure can be extracted by deforming the surface in order to minimize the localized uniform modeling energy.

For vessel segmentation, the smoothness parameter, $\lambda = 0.1\max(|\frac{d\phi}{dt}|)|_{t=0}$ and the localization radius, $r = 5$ mm. This choice for r is reasonable because it represents the maximum possible diameter for vessels in the vessel tree. Figure 5.4 shows the initializations, and resulting vessel segmentations on the LAD, LCX and RCA vessel trees.

(a) (b)

(c) (d)

Figure 5.4: 3-D renderings of initializations and segmentation results on the (a) LAD, (b) LCX, and (c, d) RCA. For each vessel, the initialization is shown on the left as a dark point and the resulting segmentation is shown on the right.

5.4   Construction of Initial Surfaces

Using the vessel segmentation as starting point, two surfaces are created that act as initializations for the detection step. These initializations are formed inside and outside of the initial vessel segmentation so that plaques, which form within the vessel wall, will be located between the two surfaces.

To create the interior surface a third localized energy is employed that relies only upon

the geometry of the surface and not image information. An internal energy defined as

$$F_{\text{SHRINK}} = \mathcal{H}\phi(y) \tag{5.2}$$

penalizes interior area locally. Thus as Equation (3.5) is minimized with F_{SHRINK}, the segmented vessel surface thins with each iteration, but the thickest parts of the vessel surface thin the fastest. Thus by deforming ϕ via Equation (3.6) several times, an interior surface is created that preserves the topology of the original surface and is always interior to it. This is similar to a medial axis (skeleton) except that relative thicknesses of vessel segments are preserved, meaning that although all segments become thinner, thick segments remain thicker than thin segments. Also, the surface resulting from this process will have a definite interior so that image statistics can be computed within it.

The exterior initialization is created by evolving the segmented vessel surface outward along its unit normal for several iterations using the update equation

$$\frac{d\phi}{dt}(x) = 1 + \lambda \int_{\Omega_x} \delta\phi(x)\|\nabla\phi(x)\|dx. \tag{5.3}$$
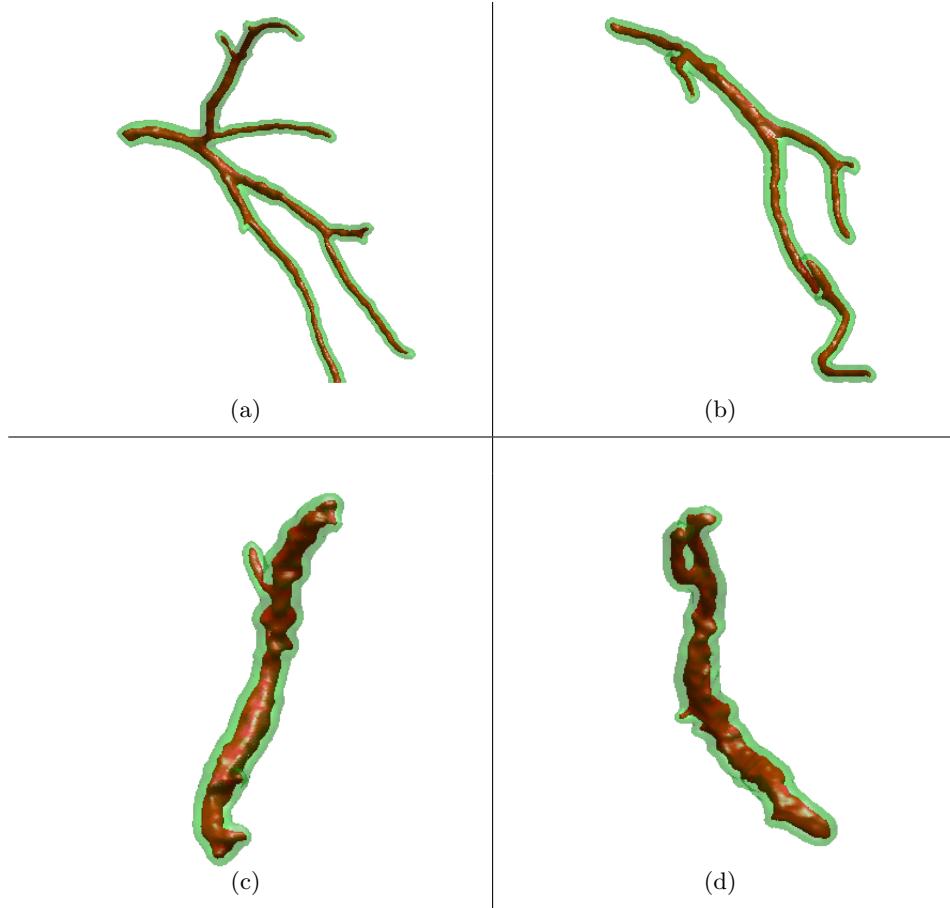
This is functionally equivalent to morphological dilation with a 5 mm×5 mm ball-shaped structuring element. The end result is a surface, which is always outside of the initial vessel segmentation. These initializations are particularly important due to the local-looking aspect of the active contours used for soft plaque detection.

5.5 *Plaque Detection Via Naive Segmentation*

The next step applies the localized mean separation internal energy from Section 3.4.2 to pull the two initial surfaces towards each other. Where no soft plaque exists, the two surfaces will meet on the vessel wall. However, if a soft plaque deposit exists between the two evolving surfaces they will each stop on the plaque boundary and remain separated from one another. By identifying areas where the two contours do not meet, soft plaques are detected.

The interior and exterior surfaces created in Section 5.4 are each deformed in order to minimize Equation (3.5) with the local mean separation internal energy,

$$F_{\text{MS}} = \int_{\Omega_y} -(\mu_{\text{in}}(x) - \mu_{\text{out}}(x))^2. \tag{5.4}$$

69

Figure 5.5: A rendered 3-D view of the interior (red) and exterior (green) initializations for four different volumes. All soft plaques should lie between these two surfaces.

The same parameters are used as in vessel segmentation. Minimizing this energy deforms the surface such that local means are pulled apart as much as possible.

Initially, the local interior regions of the inner surface will only include bright blood voxels. As the surface deforms, it will expand to capture more voxels containing blood but will not expand into slightly darker soft plaque voxels. Similarly, the exterior surface will not contract if doing so would cause soft plaque voxels, which are slightly brighter than myocardium voxels, to move into its local exterior regions. This behavior allows the two evolving surfaces to capture soft plaques between each other because neither will move into the plaque voxels when driven by the localized mean separation energy.

The two surfaces are evolved until convergence, and soft plaques are detected as areas where the distance between the two surfaces is larger than one voxel. A label map P, which

is 1 where plaques exist and 0 elsewhere, is defined as
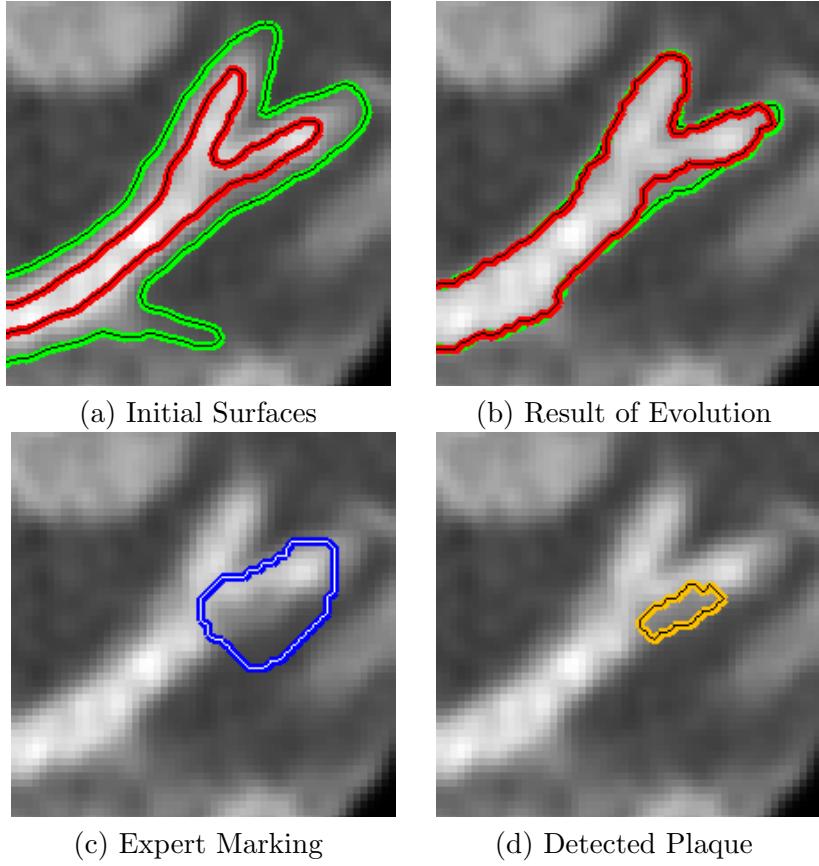
$$P = \big(1 - \mathcal{H}(\phi_{\text{interior}})\big) \cdot \mathcal{H}(\phi_{\text{exterior}} + 1). \tag{5.5}$$

5.6   Soft Plaque Detection Results

This section presents the results of experiments on MDCT scans of 8 patients undergoing CTA, for which expert markings of soft plaques were able to be obtained. All scans were taken with a Siemens Sensation 64-slice multi detector CT machine and examined by an expert cardiologist to mark soft plaque deposits. For each dataset plaques were detected using the algorithm presented in Section 5.2 based on a single-point initialization within the vessel tree of interest. The output of this detection algorithm was then compared to markings made by the expert investigator. Note that the expert markings denote large areas that are likely to contain plaques while the proposed detection algorithm finds the outline of the detected plaques.

First, examples of plaque detection are shown on 2-D slices where the behavior of the algorithm is easy to visualize. Figures 5.6 and 5.7 show examples of non-calcified plaques that cause negative and positive remodeling of the vessel wall, respectively. Each figure shows the initial interior and exterior surfaces, both surfaces after evolution with the local mean separation energy, the expert markings, and the detected plaque.
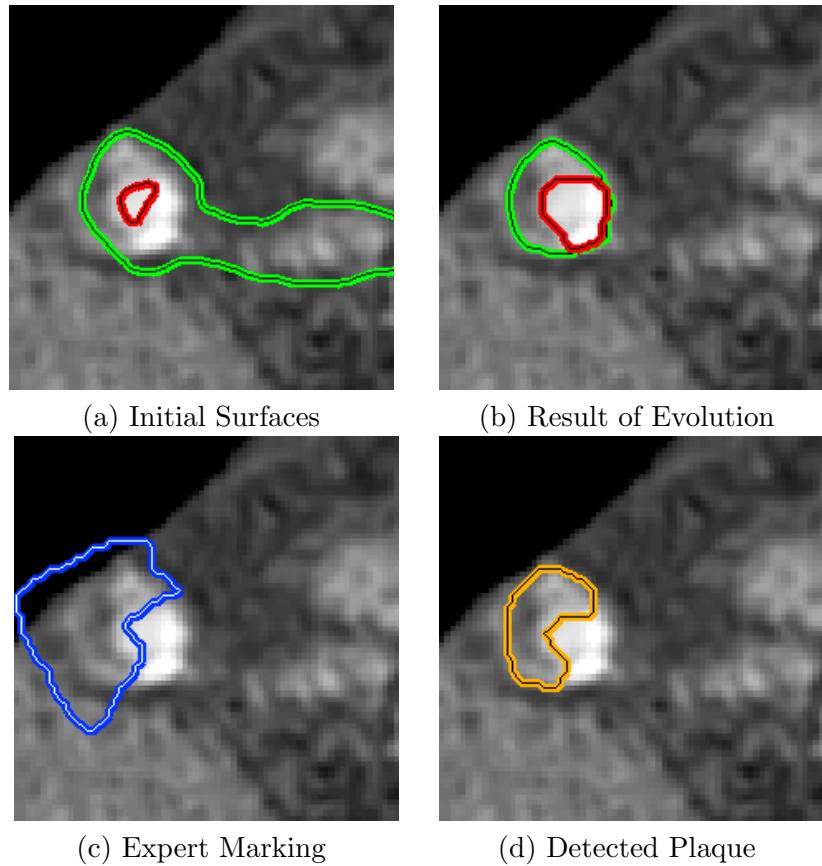
Figures 5.8 and 5.9 show rendered 3-D views of the results from all 4 datasets that contained soft plaques. For each vessel tree the segmented vessel is shown with truth markings and detected plaques. The datasets include plaques located in the LAD, LCX, and RCA vessel trees. In these experiments plaques were identified well based on a single-click input. In total 8 plaques were marked by the expert investigator within the 4 datasets. The proposed algorithm correctly identified 7 of 8 (87.5%) plaques with 1 false negative and 0 false positives. Results are summarized in Table 5.1. Additionally, the 4 datasets in which the LAD, LCX, and RCA were determined to be free of plaques were tested and, the algorithm correctly detected 0 plaques.

(a) Initial Surfaces   (b) Result of Evolution

(c) Expert Marking   (d) Detected Plaque
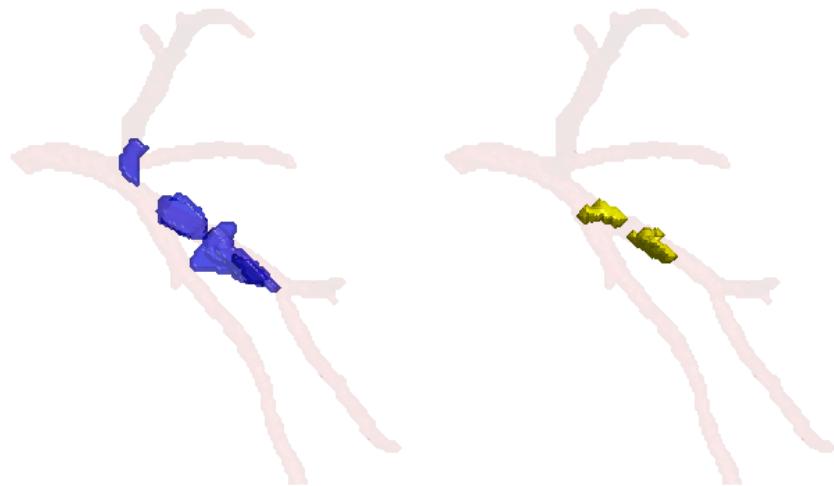
Figure 5.6: A 2-D view of detection results on the LAD (also shown in Figure 5.8-a). The vessel demonstrates *negative* remodeling in the presence of plaque. The images depict (a) the initial interior and exterior surfaces, (b) the result of evolution with the local mean separation energy, (c) the expert markings, and (d) the detected plaque.

Table 5.1: Results of soft plaque detection in Figures 5.8 and 5.9.

Plaque ID	Remodeling	Vessel Segment	Confirmed	Detected	Figure
#1	*negative*	LAD	×		Fig. 5.8-a
#2	*positive*	LAD	×	×	Fig. 5.8-a
#3	*positive*	LAD	×	×	Fig. 5.8-a
#4	*negative*	LCX	×	×	Fig. 5.8-b
#5	*positive*	LCX	×	×	Fig. 5.8-b
#6	*negative*	RCA	×	×	Fig. 5.9-a
#7	*negative*	RCA	×	×	Fig. 5.9-b
#8	*positive*	RCA	×	×	Fig. 5.9-b

(a) Initial Surfaces   (b) Result of Evolution

(c) Expert Marking   (d) Detected Plaque

Figure 5.7: A 2-D view of detection results on the RCA (also shown in Figure 5.9-b). The vessel demonstrates *positive* remodeling in the presence of plaque. The images depict (a) the initial interior and exterior surfaces, (b) the result of evolution with the local mean separation energy, (c) the expert markings, and (d) the detected plaque.
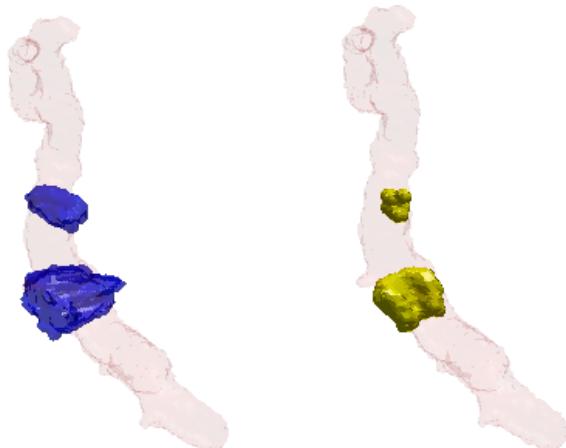
Figure 5.8: 3-D Renderings of detection results on the (a) LAD, (b) LCX. For each, the expert markings are shown in blue on the left, and the detection results are shown in yellow on the right.

Figure 5.9: 3-D Renderings of detection results on the RCA. Result (a) and (b) represent different datasets. For each, the expert markings are shown in blue on the left, and the detection results are shown in yellow on the right.

5.7 Remarks on Localized Coronary Analysis

This chapter presented a technique that is capable of segmenting vessel trees and detecting non-calcified plaques automatically from on a simple, one-click user input. Once again, the localized active contour framework was well-suited to handle these difficult medical image processing tasks. Simple statistical models based on local means were sufficient to accurately extract the vessel and find areas where soft plaques exist. Experiments produced results that match an expert's manual detections.

Additional future work should include a population study in which a larger number of datasets are analyzed. Furthermore, the results should be compared with intravascular ultrasound imagery to confirm the presence and composition of detected plaques. This method is among the first to address the problem of soft plaque detection. Hence, the author is optimistic that this work has the potential of being a profound step forward in automatically detecting non-calcified plaques, which have been clearly linked with the occurrence of heart attacks and stroke.

CHAPTER VI

TEMPORAL LOCALIZATION FOR VISUAL TRACKING

In the previous chapters, it was shown that localization can improve the performance of segmentation and detection in medical applications. However, the purview of computer vision encompasses more than these pursuits. This chapter, therefore, investigates visual tracking. As defined in Chapter 1 tracking is the process of continuously locating moving objects in time-varying images.

Tracking uses localized information by nature. The temporal coherency of video sequences implies that a tracked object will display little change between frames. However, this chapter takes this inherent localization a step further. In the same way that isolation of local *spatial* regions improves segmentation, isolation of local *temporal* regions can improve tracking results.

This chapter proposes a tracking system that is especially well-suited to tracking targets that change drastically in size or appearance, a situation that can defeat most trackers. To accomplish this, a fast two-phase template matching algorithm is employed to ensure accurate detection. Meanwhile, periodic template updates using temporal localization ensure accurate appearance models of the objects even as they change throughout the sequence. Furthermore, because of its design, the algorithm can continue to deliver real-time results even when targets grow to fill the entire field of view. The proposed method is demonstrated with good results on several sequences in which the tracked objects exhibit large changes in size, shape, and appearance.

This chapter is organized as follows: A background on relevant literature is given in Section 6.1. Next, Section 6.2 describes the proposed tracking algorithm in detail including the registration technique and temporally localized template update scheme. Section 6.3 shows several experiments demonstrating the capabilities of the algorithm and describes the influence of key parameters. Finally, concluding remarks and possible directions for further

77

development are offered in Section 6.4.

6.1   Visual Tracking Background and Contributions

This chapter addresses the problem of tracking objects that change drastically in size over time. When objects move closer or farther from the camera, significant changes in size, shape, and intensity profile occur, as seen in Figure 6.1. Methods that do not account for changing target appearance cannot accurately maintain track in these cases.

Figure 6.1: First, middle, and last frame from the LEAVES sequence. Notice the drastic change in size, shape, and appearance exhibited by the cluster of leaves as the camera approaches them.

Several techniques attempt to address this problem using the popular the mean-shift framework. Collins creates an additional feature space based on target scale characteristics and solves for scale and translation parameters simultaneously [20]. Peng *et al.* and Qian *et al.* adjust the window size and kernel bandwidth of the tracker based on estimations of target scale in successive frames [80, 87]. In these methods, localization is based on an unchanging intensity histogram most often taken from the first frame of the video sequence. When the target is far away, its estimated histogram may be quite different from a histogram taken when the target is closer. Because these previous methods do not account for this, they often lose track as the appearance changes.

Template tracking is another approach that addresses tracking targets of variable size [7, 55, 59]. The goal of these methods is to register a template image patch from a previous frame onto the current frame to determine movement. When registration parameters allow scaling, it is possible for the template tracker to follow the changing size of the target. However, and most importantly, this method is defeated if the object changes significantly in appearance as it changes in size.

78

Based on a review of the literature, there appear to be two sources of error that influence template models while the target is tracked. *Spatial drift* is change in the model such that the model and target are misaligned. *Feature drift* is the change of target appearance as it diverges from the appearance of the model over time.
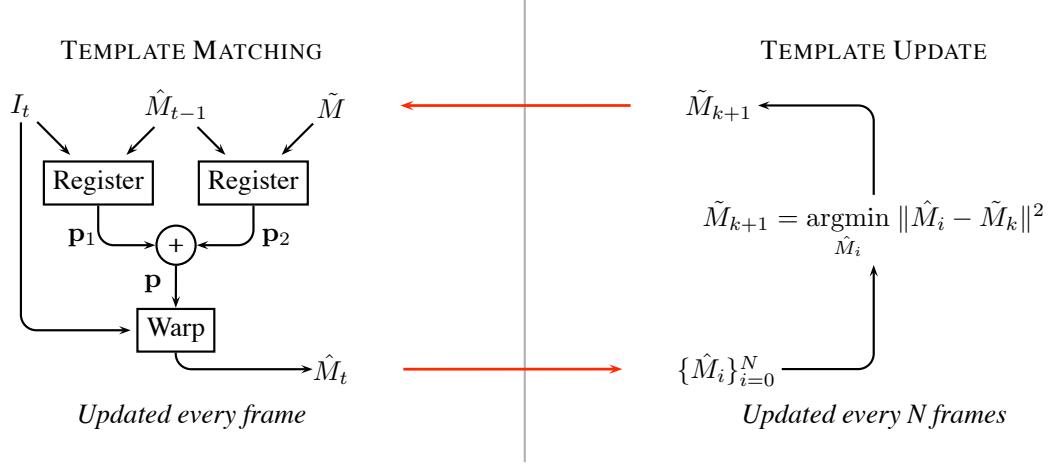
Proposed here is a method that deals with these two sources of error separately. Spatial drift is prevented by registering the previous target representation to the current frame and registering the previous frame to a periodically updated *key model*. This accounts for frame-to-frame movement and incorporates the influence of a stable model to minimize transient affects. Feature drift is accounted for by updating the key model based on local time-windows in a way that limits spatial drift and also allows smooth feature changes over time. This update allows the method to track features that are specific to the target's recent appearance and is robust against large changes in scale.

6.2   *Localized Template Matching Algorithm*

The proposed method is divided into two main components: template matching and localized template updating. Template matching, described below, prevents spatial drift in the target model by performing two registration procedures. First, translation parameters $\mathbf{p}_1$ are determined, which align the image I at the current frame to the model, $\hat{M}_t$ obtained from the previous frame. Here, models refer to image patches that represent a known appearance of the target in a previous frame. This first registration reduces frame-to-frame spatial drift.

Next, a second set of translation parameters $\mathbf{p}_2$ is found that aligns $\hat{M}$ to the key model $\tilde{M}$. The key model is the best match, in a sum of squared differences sense, of the *recent* $\hat{M}$ models and is periodically chosen in the template update step. This second registration reduces spatial drift. The two sets of translation parameters are summed to determine the final translation parameters $\mathbf{p}$ that specify the location of the target and the next model $\hat{M}_{t+1}$.

Template update, described in Section 6.2.2, is executed every N frames, and allows the key model $\tilde{M}$ to change over time in a manner that limits spatial drift yet allows a smooth

Figure 6.2: Flow chart representation of the proposed method

change in target appearance. Every N frames, the key model is replaced by the best model from the a set of the previous N models $\{\hat{M}_i\}_{i=1}^N$. The diagram in Figure 6.2 illustrates the entire tracking algorithm with template matching shown on the left and template update shown on the right.

6.2.1 Classic Template matching

The goal of template matching is to estimate transformation parameters $\mathbf{p} = \{p_1 \dots p_n\}$ that best align a reference template $T(x)$ to an image $I(x)$, where $x = (x, y)^T$ is a vector of template coordinates. Borrowing the notation of Baker and Matthews [7], $\mathbf{W}(x; \mathbf{p})$ is said to warp reference coordinates according to $\mathbf{p}$. This problem may then be formulated as the minimization of a sum of squared differences (SSD) expression,

$$\mathbf{p}^* = \underset{\mathbf{p}}{\operatorname{argmin}} \ \sum x \left[I(\mathbf{W}(x; \mathbf{p})) - T(x) \right]^2, \tag{6.1}$$

where the sum is computed over template coordinates x. First, this expression is reformed with the introduction of a step variable $\Delta\mathbf{p}$, and a truncated Taylor expansion is used to separate it from the warp:

$$= \ \sum x [I(\mathbf{W}(x; \mathbf{p} + \Delta\mathbf{p})) - T(x)]^2 \tag{6.2}$$

$$\approx \ \sum x [I(\mathbf{W}(x; \mathbf{p})) + \nabla I \cdot \frac{\partial \mathbf{W}}{\partial \mathbf{p}} \cdot \Delta\mathbf{p} - T(x)]^2. \tag{6.3}$$

Next the gradient with respect to $\Delta\mathbf{p}$ is taken and terms are rearranged to yield the step solution:

$$\Delta\mathbf{p} = \sum x H^{-1} \left[\nabla I \cdot \frac{\partial\mathbf{W}}{\partial\mathbf{p}}\right]^T [T(x) - I(\mathbf{W}(x; \mathbf{p}))] \tag{6.4}$$

where H is the $n \times n$ *Hessian* matrix:

$$H = \sum x \left[\nabla I \cdot \frac{\partial\mathbf{W}}{\partial\mathbf{p}}\right]^T \left[\nabla I \cdot \frac{\partial\mathbf{W}}{\partial\mathbf{p}}\right]. \tag{6.5}$$

If the Hessian matrix is examined, one notices that the terms that most effect its computation are those with large image gradients ∇I. These coordinate locations $X \subseteq I$ contain the most information affecting the transformation parameters $\mathbf{p}$. As suggested by Dellaert and Collins [24], using only this subset of pixels in the computations can increase speed by orders of magnitude with little loss in accuracy. Following this technique, at the start of each iteration, Equation (6.4), Equation (6.5), and necessary derivatives are computed only on the dominant subset X with the highest values of ∇I. Procedure 6.1 details the iterative estimation of $\mathbf{p}$ as described in [55, 59].

Procedure 6.1 Registration

    Determine dominant subset $X \subseteq I$
  repeat
    Compute $I(\mathbf{W}(x; \mathbf{p}))$
    Compute residual $[T(x) - I(\mathbf{W}(x; \mathbf{p}))]$
    Compute $\nabla I \cdot \frac{\partial\mathbf{W}}{\partial\mathbf{p}}$ and form the Hessian H as in Equation (6.5)
    Solve for $\Delta\mathbf{p}$ as in Equation (6.4)
    $\mathbf{p} \leftarrow \mathbf{p} + \Delta\mathbf{p}$
  until $\mathbf{p}$ has converged

This process is well suited for arbitrary transformation parameters, but experimentation suggests that allowing more complex transformations than translation introduces unnecessary degrees of freedom and can lead to inaccurate registrations. Hence, $\mathbf{p}$ is taken to be translation in the x and y direction, and

$$\mathbf{W}(x; \mathbf{p}) = x + \mathbf{p} \tag{6.6}$$

6.2.2 Temporally-Localized Update Scheme

The goal of template update is to allow the appearance of the reference image to change in order to reduce feature drift while at the same time preventing spatial drift from entering

the system. To achieve this, the key model $\tilde{M}$ is updated every N frames with the procedure outlined in Figure 6.3.
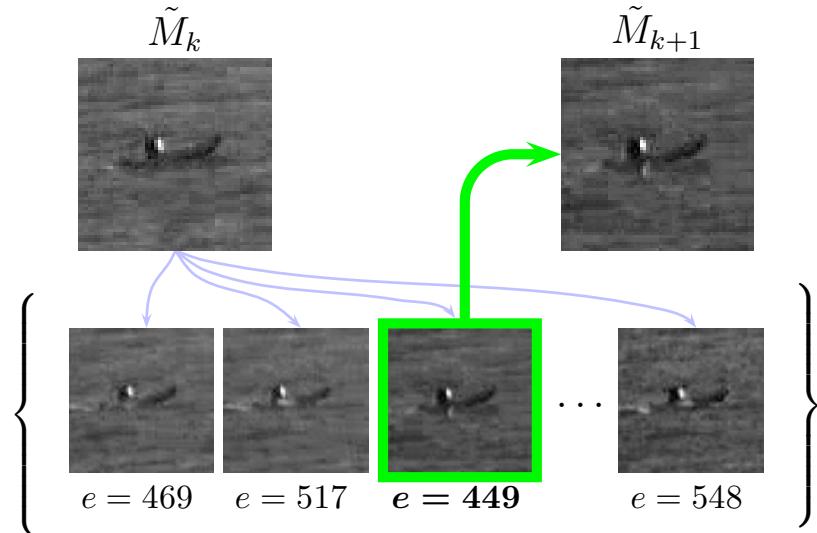
During the template matching step, the past N models are collected and denoted as $\{\hat{M}_i\}_{i=1}^N$. The new key model, $\tilde{M}_{k+1}$ is chosen as the best representative of this set as determined by a matching error, e

$$e(\hat{M}, \tilde{M}) = \|\hat{M} - \tilde{M}\|^2 \tag{6.7}$$

$$\tilde{M}_{k+1} = \underset{\hat{M}_i}{\operatorname{argmin}} \; e(\{\hat{M}_i\}_{i=1}^N, \tilde{M}_k) \tag{6.8}$$

The set of $\{\hat{M}_i\}_{i=1}^N$ represents optimally registered image patches from the previous N frames. However, many will exhibit image noise or slight misalignments that make them poor choices for a key model. The new key model is chosen as the candidate with the lowest matching energy so that the key model changes smoothly and without introducing spatial or feature drift.

Note that the size of the models never vary as they are updated. Although the location of the target is maintained as it changes in scale, the size of the template representation never changes. This is accomplished by tracking features available at the current scale of the target. Hence when the target is small, the key model may include the entire target,

Figure 6.3: The key model $\tilde{M}_k$ is updated periodically to prevent feature drift. The new key model $\tilde{M}_{k+1}$ is selected by choosing a model from the set $\{\hat{M}_i\}_{i=1}^N$ that has the lowest error e when compared to the previous key model $\tilde{M}_k$

but if the target grows such that it is larger than the model, the model holds finer details of the target's appearance that allow the algorithm to maintain its track point. This has the advantage of fixing the computational complexity of the algorithm even as the target's proportion of the frame increases. This, in turn, helps guarantee real-time speed even as the target becomes large.

6.3   Experiments with the Proposed Tracker

The tracker is applied to several sequences chosen to demonstrate robustness against large changes in appearance of the target especially due to scale change. Additionally, a discussion is given regarding the effects of two key parameters that can affect system performance. Finally, computational complexity and achievable frame rates are discussed.
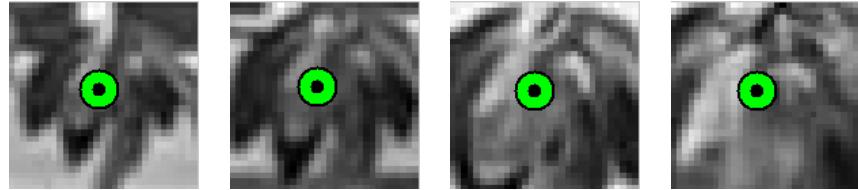
6.3.1   Results of the Tracking System

First, consider the LEAVES sequence, shown in Figure 6.4, where the proposed method was used to track a small group of leaves. In the first frame, when the track is initialized, the leaves appear small, and little detail can be seen. Over the course of the video, the camera approaches the leaves and their size and detail level increase dramatically. Three full frames from the video are shown with the tracking result on the first row of Figure 6.4. Rows two and three show models, $\hat{M}_t$ from selected frames. Notice how the information in the model adapts to represent features specific to the current scale of the target. For instance, originally the model holds leaves and background. Mid-way through the sequence, the model is mostly filled by the group of leaves being tracked. At the end, the model contains only the front-most leaf. In this way, the system maintains a stable track point throughout the sequence.

In Figure 6.5 tracking results for the VEHICLE sequence are shown. This tactical imagery shows a vehicle being tracked beginning from a far distance and closing quickly until the target becomes large in the field of view. In the initial frames, the target is only several pixels large, and can only be seen in the zoomed-in models shown in the second and third rows of the figure. The model adapts to the target's changing size, and the sequence is accurately tracked despite large changes in scale and appearance. Again, notice how the

(a) First, middle, and last tracked frame
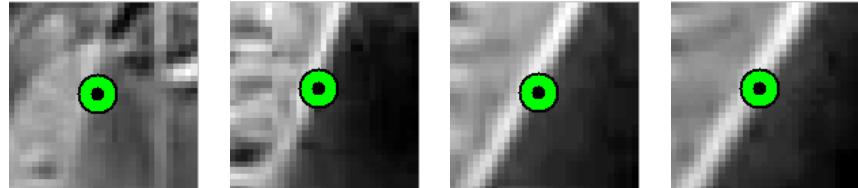
Frame 28   Frame 55   Frame 81   Frame 108

Frame 134   Frame 161   Frame 187   Frame 214
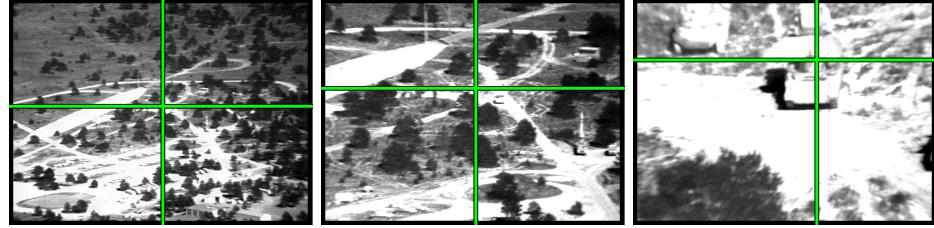
(b) Optimal models, $\tilde{M}$, at various frames
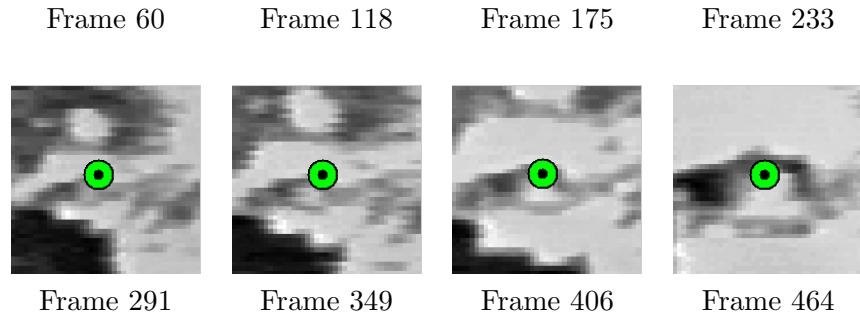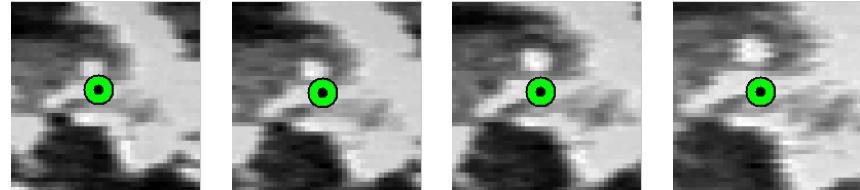
Figure 6.4: Tracking results on the LEAVES sequence demonstrating ability to track through large scale change. (a) Selected full-size frames. (b) Target models from selected frames throughout the sequence.

model changes dramatically over the course of the sequence. Feature drift is accounted for by the updating model, and spatial drift is prevented as well.

The BOAT sequence demonstrates the robustness of the technique against factors other than scale changes in Figure 6.6. The size change of the target is not as drastic as those in Figures 6.4 and 6.5, but this sequence exhibits poor resolution, sporadic illumination changes, significant image noise, and substantial target variability due to wave activity and the varying pose of the boat. Figure 6.6 shows three full size frames with tracking results and eight models from throughout the sequence.

(a) First, middle, and last tracked frame

| Frame 60 | Frame 118 | Frame 175 | Frame 233 |

| Frame 291 | Frame 349 | Frame 406 | Frame 464 |

(b) Optimal models, $\tilde{M}$, at various frames

Figure 6.5: Tracking results on the VEHICLE sequence demonstrating ability to track through large scale change. (a) Selected full-size frames. (b) Target models from selected frames throughout the sequence.

6.3.2 Understanding Key Parameters
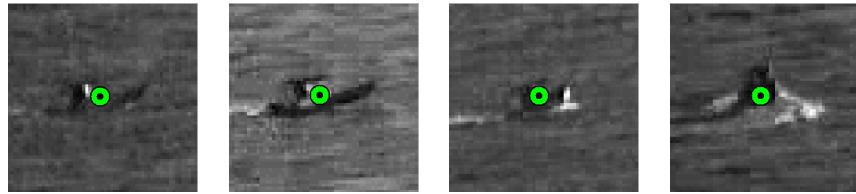
The proposed method has two parameters that control its behavior as described in Section 6.2:

1. The percentage of pixels that define the dominant subset $X \subseteq I$

2. The length, N of the history used for key model updates.
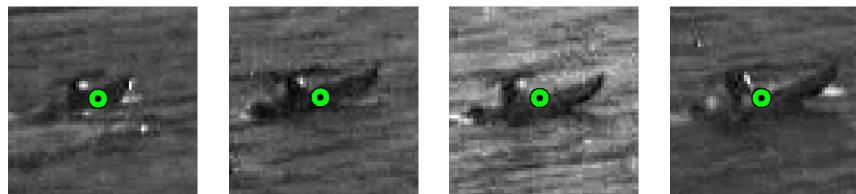
Regarding the percentage of pixels used in registration, it was found that using fewer pixels increases the speed of the algorithm significantly. It can also increase the accuracy of registration in some cases. In the BOAT sequence, the background water is noisy and low contrast. If 100% of the pixels are used in registration for this sequence, results would not

(a) First, middle, and last tracked frame

Frame 53   Frame 104   Frame 155   Frame 206

Frame 257   Frame 308   Frame 359   Frame 410

(b) Optimal models, $\tilde{M}$, at various frames

Figure 6.6: Tracking results on the BOAT sequence demonstrating ability to track through noise, illumination changes, and target variability. (a) Selected full-size frames. (b) Target models from selected frames throughout the sequence.

be as accurate due to many ambiguous background pixels being included in the registration. Instead, 25% of pixels were used with good results. In the LEAVES sequence the background is high-contrast. In this case using only a small percentage of the pixels can cause the template matching step to favor registration of the background rather than the target. Consequently, 100% of the pixels were used to ensure the object is tracked despite the high contrast background.

This parameter can be set in advance for most tracking scenarios based on prior knowledge about the application. For instance, in cluttered urban environments a large percentage is required as in the LEAVES example. Alternatively, in scenarios where the background is

simple and possibly noisy a low percentage is a more appropriate choice. The BOAT sequence is an example of this, as also would be tracking an aircraft against a background of sky.

The second parameter of interest is the length, N, of the model history used during the template update step. This parameter must be set to reflect how quickly the target is expected to change in appearance. For instance, in the LEAVES sequence, the target rapidly undergoes a large change in appearance and scale due to low frame-rate video. For this experiment a relatively short history of $N = 5$ is used. In the VEHICLE and BOAT sequences, the target changes more slowly, and a history of 20 frames is used.

6.3.3   Computational Efficiency

Finally, the efficiency inherent in this algorithm is examined. Other methods that are capable of tracking objects though scale changes may continuously change the size of their tracking window to include the entire object. Hence, the number of pixels analyzed in order to locate the target increases with target size. This increase can lead to slower frame rates when the target is close enough to take up a large portion of the field of view. Because the proposed algorithm uses a fixed window size and updates the features tracked within that window, the frame rate remains roughly constant despite changes in target size. Furthermore, the simplicity of the algorithm allows it to run at real time speeds. In a prototype Matlab implementation speeds between 17 and 26 Hz were achieved. Table 6.1 shows the frame rates achieved for each of the three experiments shown.

Table 6.1: Frame rates achieved during experiments

Sequence	Figure	Frame Rate
LEAVES	Figure 6.4	17.54 Hz
VEHICLE	Figure 6.5	25.46 Hz
BOAT	Figure 6.6	24.62 Hz

6.4   Remarks on Temporally Localized Tracking

The novel tracking algorithm presented applies the ideas of localization temporally rather than spatially to improve performance. The resulting system was capable of tracking targets

through significant changes in size and appearance. By using a two-phase template matching scheme and an intelligent, localized template update procedure, the system prevented spatial drift and feature drift in order to maintain accurate track point in challenging sequences. Furthermore, the efficiency of the algorithm makes it useful for real-time applications.

While this exploration into tracking was less in-depth than the medical segmentation examples seen in previous chapters, it does demonstrate that localization can indeed be useful in domains outside of segmentation and detection. Future work related to localized tracking methods may include using robust estimators rather than sum of squared differences to determine correlation between models and targets. Furthermore, one could investigate the inclusion of a localized active contour to segment tracked objects in addition to merely finding their location in each frame.

CHAPTER VII

CONCLUDING REMARKS

This thesis demonstrated that by solving computer vision problems with localized methods it becomes possible to relax assumptions about the image data and improve segmentation and tracking performance in most cases. The ability to accomplish image-processing tasks without the need for restrictive assumptions is important; solving complex problems with simple, robust algorithms paves the way to solve more complex problems that will arise in the future.

Within the thesis a novel framework was proposed that allowed active contour segmentation problems to be localized at various scales depending on the relevant image information. It was shown how this framework could take any region-based active contour segmentation energy and reformulate it so that a local scope could be used for statistical modeling instead of a global scope. This was followed with a study of how appropriate scales are chosen in different circumstances so that the reader may successfully apply the presented framework to general problems.

The major contributions of this framework are threefold. First, it provides a simple but effective way to improve global region-based energies by reducing the complexity of image assumptions through localized statistical analysis. Second, it introduced two novel segmentation energies developed using this framework. These two energies and their corresponding level set update equation were discussed in detail and experiments demonstrated how the localized versions out-perform their global counterparts. Finally, the framework is extended so that multiple localized contours can interact naturally to produce n-ary segmentations of an image.

With the framework well understood, two important applications were addressed from the medical imaging community. These applications, related to brain and heart imaging respectively, proved to be well suited for analysis with localized statistics. The success in

these applications showed the power of the presented localization method for solving real-life computer vision problems.

First, the segmentation of bundles of neuron fibers in DTI data was considered. DTI measures the orientation and anisotropy of water diffusion in various tissues in the body. Because the bundles curve significantly as they wind through the brain they are difficult to segment using existing methods. Fortunately, they are easily segmented when statistics are examined locally. In order to segment the neuron bundles the presented framework was utilized to create a localized energy capable of operating on diffusion tensors. Experiments confirmed that segmentations driven by a localized energy operating on tensor data could successfully segment these fiber bundles.

The second medical image processing application considered was that of automating analysis of CTA data to assess heart health. Specifically segmentation of the vessels and the automatic detection of potentially dangerous plaques were addressed. Coronary arteries exhibit changing appearance along their length and often exist near other structures with similar appearance. Localized segmentation proved to be successful in accurately segmenting these structures based on a simple initialization. While vessel segmentation is a useful task, it has been successfully achieved using other methods. Hence, a more important contribution is the detection of soft plaques.

It was shown that by performing a coupled segmentation of the interior and exterior of the vessel wall using a segmentation energy driven by local statistics it is possible to detect potentially dangerous soft plaques that exist within the coronary vessel tree. This challenging task is even difficult for human experts, and the presence of plaque is considered to be a major risk factor for cardiovascular disease. The detection algorithm was presented in detail and validated in experiments where automatic plaque detections were compared to those made by an expert cardiologist.

Finally, the ideas of localization were explored as they pertain to visual tracking. In this case, localization was applied temporally rather than spatially to allow the appearance model of a tracked object to be updated optimally over time. Experiments showed that the presented tracking technique is capable of continuously locating objects even as they

change drastically in size or appearance over the course of a video sequence.

The authors are optimistic that this work has contributed to the study of computer vision, and that it will help future researchers in their continuing efforts to improve the way computers interact with images. Possible directions for continuing the presented research would include automatically determining the radius of localization based on image data rather than prior knowledge, investigating additional localized segmentation energies, and conducting population studies to gauge the clinical usefulness of the presented medical image processing techniques.

APPENDIX A

DERIVATIONS OF CURVATURE FLOW

Several relevant derivations are given in this appendix. First, the derivation of a generic region-based active contour gradient flow is given, followed by the derivation of the general localized region-based active contour flow.

A.1 Derivation of a Generic Region-Based Flow

In this section, the gradient of a general region-based energy of the form

$$E(\phi) = \int_R F dx = \int_{\Omega_x} \mathcal{H}(\phi(x)) F dx \tag{A.1}$$

is computed. Here, F is an independent function and R is the region enclosed *inside* the curve Γ. Take ν to be a perturbation of ϕ (of same dimension as ϕ) and ξ to be a real number. One can then write

$$E(\phi + \xi\nu) = \int_{\Omega_x} \mathcal{H}(\phi(x) + \xi\nu(x)) F dx. \tag{A.2}$$

By taking the Gâteaux derivative of $E(\phi + \xi\nu)$ with respect to ξ, one obtains

$$\frac{dE(\phi + \xi\nu)}{d\xi}\bigg|_{\xi=0} = \int_{\Omega_x} \delta(\phi(x))\nu(x) F dx. \tag{A.3}$$

From this, one can invoke the Cauchy-Schwartz inequality to find the gradient of E with respect to ϕ,

$$\nabla_\phi E = \delta(\phi) F. \tag{A.4}$$

A.2 Derivation of the Localized Region-Based Flow

Using the derivation presented above, this section shows how to take the gradient of a generic localized active contour energy. Recall that a localized energy, E is defined in terms of a generic internal energy functional, F as shown in Equation (3.5):

$$E(\phi) = \int_{\Omega_x} \delta\phi(x) \int_{\Omega_y} \mathcal{B}(x,y) F(I, \phi, x, y) dy \, dx. \tag{A.5}$$

To compute the first variation of this term, a change of parameters is again performed to express $E(\phi)$ as $E(\phi + \xi\nu)$:

$$E(\phi + \xi\nu) = \int_{\Omega_x} \delta(\phi(x) + \xi\nu) \int_{\Omega_y} \mathcal{B}(x,y)F(I, \phi + \xi\nu, x, y)dy \ dx. \qquad (\text{A.6})$$

Next, the partial derivative of this energy is taken with respect to ξ evaluated at $\xi = 0$ to represent a tiny differential of movement. By the product rule,

$$\nabla_\xi\Big|_{\xi=0} E = \int_{\Omega_x} \delta(\phi(x)) \int_{\Omega_y} \nu\mathcal{B}(x,y)\nabla_{\phi(y)}F(I, \phi, x, y)dy \ dx +$$
$$\nu\int_{\Omega_x} \gamma\phi(x) \int_{\Omega_y} \mathcal{B}(x,y) \cdot F(I, \phi, x, y)dy \ dx \qquad (\text{A.7})$$

is obtained. Note that $\gamma\phi$ denotes the derivative of $\delta\phi$. On the zero level set $\gamma\phi$ evaluates to zero. As such, it does not affect the movement of the curve and thus, this term is ignored. Now the integral over y is moved outside the integral over x:

$$\nabla_\xi\Big|_{\xi=0} E = \int_{\Omega_y} \int_{\Omega_x} \nu\delta(\phi(x))\mathcal{B}(x,y)\nabla_{\phi(y)}F(I, \phi, x, y)dx \ dy. \qquad (\text{A.8})$$

At this point, the Cauchy-Schwartz inequality is used to show that the optimal direction to move ϕ is given by:

$$\frac{\partial\phi}{\partial t} = \int_{\Omega_y} \delta(\phi(x))\mathcal{B}(x,y)\nabla_{\phi(y)}F(I, \phi, x, y)dx \ dy. \qquad (\text{A.9})$$

Re-arranging the integrals once more gives the same equation in a form that is easier to understand. This yields the final level set update function

$$\frac{\partial\phi}{\partial t} = \delta\phi(x)\int_{\Omega_y} \mathcal{B}(x,y) \cdot \nabla_{\phi(y)}F(I, \phi, x, y)dy \ dx. \qquad (\text{A.10})$$

APPENDIX B

SPARSE FIELD LEVEL SET IMPLEMENTATION

The primary drawback of level set methods as described in Chapter 2 is that they are slow to compute. In their truest form, many computations are required to maintain a full signed distance function ϕ as the embedded contour changes. The key to reducing complexity is to notice that only the locations in ϕ where $\phi(x) \approx 0$ are important for accurate curve representation. This observation has led to proposals of various *narrow-band* algorithms that reduce the computational complexity by only performing calculations near the zero level set. The sparse field method (SFM) proposed by Whitaker in [103] remains one of the most efficient algorithms to maintain a minimal, but accurate representation of ϕ.

Whitaker's original paper outlines his algorithm and presents evidence that the method requires fewer computations and produces less error than other narrow-band techniques. However, in [103] the algorithm is described briefly and in abstract terms. The remainder of this appendix describes the SFM in detail and makes specific comments regarding implementation and natural extensions.

B.1 Sparse Field Algorithm

The sparse field method (SFM) uses lists of points that represent the zero level set as well as points adjacent to the zero level set. By using these lists and carefully moving points to and from the appropriate list an efficient representation of ϕ can be maintained.

These lists are implemented as *doubly-linked-lists*, which have the properties that elements can be added dynamically and removed from the middle of the list. This type of data-structure is available in most programming languages (for example, the `list` class in C++). Five lists are used in the SFM to represent five different levels:

$$L_0 \quad \rightarrow \quad [\text{-}0.5 \quad , \quad 0\,.5] \quad \rightarrow \quad \texttt{Lz}$$

$$L_{-1} \quad \rightarrow \quad [\text{-}1.5 \quad , \quad \text{-}0.5) \quad \rightarrow \quad \texttt{Ln1}$$

$$L_1 \quad \rightarrow \quad (\,0.5 \quad , \quad 1\,.5] \quad \rightarrow \quad \texttt{Lp1}$$

$$L_{-2} \quad \rightarrow \quad [\text{-}2.5 \quad , \quad \text{-}1.5) \quad \rightarrow \quad \texttt{Ln2}$$

$$L_2 \quad \rightarrow \quad (\,1.5 \quad , \quad 2\,.5] \quad \rightarrow \quad \texttt{Lp2}$$

Each list holds the x, y, and z location of pixels in the image. In addition to the lists, two arrays are used. The first is the ϕ array. This is the same size as the image domain and should be maintained at full floating point precision. The second array is a label map the same size as ϕ. This label map is used to record the status of each point and provides a way to determine which list a point belongs to with random access. The label map will only contain the values $\{-3, -2, -1, 0, 1, 2, 3\}$.

B.1.1 Initialization

Procedure B.2 takes in a binary image (`init`) and returns a fully initialized arrays for the label map (`label`), ϕ (`phi`), and the five lists representing levels -2 through 2. The binary image (`init`) should have values 0 and 1 where 1's represent foreground pixels, and 0's represent background pixels. Furthermore, the notation: `N(p)` is used to denote the neighborhood of an point p. This represents the 4-neighborhood in 2-D (up, down, left, right) and the 6-neighborhood in 3-D (up, down, left, right, forward, and backward).

B.1.2 Evolving the Contour

Once the data structures have been initialized with Procedure B.2, the level sets may be deformed in order to minimize some segmentation energy. There are many such energies, and the choice of appropriate energies is discussed throughout this thesis. For example purposes, consider the global uniform modeling energy [18]

$$E(\phi) = \int_{\Omega_x} \left(\mathcal{H}\phi(I - \mu_1)^2 + (1 - \mathcal{H}\phi)(I - \mu_2)^2 \right) dx, \tag{B.1}$$

which has the corresponding evolution equation

$$\texttt{F}(x) = \frac{\partial \phi(x)}{\partial t} = (I(x) - \mu_1)^2 - (I(x) - \mu_2)^2. \tag{B.2}$$

95

Procedure B.2 Initialization

```
    \\ Pre-condition labelmap and phi
 1: for each point p in the domain.
 2:    if(init(p) == 0) label(p) =  3, phi(p) =  3
 3:    if(init(p) == 1) label(p) = -3, phi(p) = -3

    \\ Find the zero-level set
 4: for each point p in the domain.
 5:    if(init(p) == 1 AND any point in N(p) == 0)
 6:        add p to Lz
 7:        label(p) = 0, phi(p) = 0

    \\ Find the +1 and -1 level set
 8: for each point p in Lz
 9:    for each point q in N(p)
10:        if(label(q)==-3) add q to Ln1, label(q)=-1, phi(q)=-1
11:        if(label(q)== 3) add q to Lp1, label(q)= 1, phi(q)= 1

    \\ Find the +2 and -2 level set
12: for each point p in Ln1
13:    for each point q in N(p)
14:        if(label(q)==-3) add q to Ln2, label(q)=-2, phi(q)=-2

15: for each point p in Lp1
16:    for each point q in N(p)
17:    if(label(q)== 3) add q to Lp2, label(q)= 2, phi(q)= 2
```

Note that F should only be computed along the zero level set and should be normalized such that $\|F\| < 0.5$ at each iteration.

Updating ϕ near the zero level set: Once F has been computed and normalized, Procedure B.3 describes how to update ϕ along the zero level set as well as the four other lists of adjacent points. Note the use of five additional lists that temporarily hold points that are changing status:

$$
\begin{aligned}
S_0 &= \texttt{Sz} &\rightarrow& \quad \text{Points moving to } L_0 \\
S_{-1} &= \texttt{Sn1} &\rightarrow& \quad \text{Points moving to } L_{-1} \\
S_1 &= \texttt{Sp1} &\rightarrow& \quad \text{Points moving to } L_1 \\
S_{-2} &= \texttt{Sn2} &\rightarrow& \quad \text{Points moving to } L_{-2} \\
S_2 &= \texttt{Sp2} &\rightarrow& \quad \text{Points moving to } L_2
\end{aligned}
$$

Procedure B.3 can be summarized as follows:

1. Scan through L_0 and add the corresponding value of `F` to the existing value of ϕ at each point.

2. Points where the new value of ϕ is outside of the range [-.5 .5] are removed from L_0 and added to S_{-1} or S_1 accordingly to denote that they will be changing status to L_{-1} or L_1 respectively.

3. L_{-1} and L_1 are scanned and ϕ values are updated so that they are exactly 1 unit from their nearest neighbor in L_0. If no L_0 neighbors exist, the point is moved to S_{-2} or S_2 respectively.

4. Points in L_{-1} and L_1 whose updated ϕ value fall outside of the specified range for L_{-1} or L_1 are moved to S_0, S_{-2}, or S_2 accordingly.

5. L_{-2} and L_2 are scanned and ϕ values are updated so that they are exactly 1 unit from their nearest neighbor in L_{-1} or L_1. If no L_{-1} or L_1 neighbors exist, the point is removed from all lists, and the value of ϕ and the label map is changed to -3 or +3 accordingly.

6. Points in L_{-2} and L_2 whose updated ϕ value is outside of the specified range are removed from L_{-2} or L_2 and either moved to S_{-1} or S_1 if their value is too low, or removed from all lists if their value is too high. Points that are removed from all lists should have their corresponding point in ϕ and the label map changed to -3 or +3 accordingly.

97

Procedure B.3 Update Level Set Lists

```
    \\ Update the zero level set
 1: for each point p in Lz
 2:     add F(p) to phi(p)
 3:     if(phi(p)> .5), remove p from Lz, add p to Sp1
 4:     if(phi(p)<-.5), remove p from Lz, add p to Sn1

    \\ Update -1 and +1 level sets
 5: for each point p in Ln1
 6:      if p has no neighbors q in N(p) with label(q) == 0
 7:         remove p from Ln1, add p to Sn2
 8:      else
 9:         M = MAX of points q within N(p) and with label(q) >= 0
10:         phi(p) = M-1;
11:         if(phi(p)>= -0.5) remove p from Ln1, add p to Sz
12:         if(phi(p)<  -1.5) remove p from Ln1, add p to Sn2

13: for each point p in Lp1
14:     if p has no neighbors q in N(p) with label(q) == 0
15:         remove p from Lp1, add p to Sp2
16:     else
17:         M = MIN of points q within N(p) and with label(q) <= 0
18:         phi(p) = M+1;
19:         if(phi(p)<=   0.5) remove p from Lp1, add p to Sz
20:         if(phi(p)>    1.5) remove p from Lp1, add p to Sp2

    \\ Update -2 and +2 level sets
21: for each point p in Ln2
22:     if p has no neighbors q in N(p) with label(q) == -1
23:         remove p from Ln2, label(p)=-3, phi(p)=-3
24:     else
25:         M = MAX of points q within N(p) and with label(q) >= -1
26:         phi(p) = M-1;
27:         if(phi(p)>= -1.5) remove p from Ln2, add p to Sn1
28:         if(phi(p)<  -2.5) remove p from Ln2, label(p)=-3, phi(p)=-3

29: for each point p in Lp2
30:     if p has no neighbors q in N(p) with label(q) ==  1
31:         remove p from Lp2, label(p)= 3, phi(p)= 3
32:     else
33:         M = MIN of points q within N(p) and with label(q) <=  1
34:         phi(p) = M+1;
35:         if(phi(p)<=  1.5) remove p from Lp2, add p to Sp1
36:         if(phi(p)>   2.5) remove p from Lp2, label(p)= 3, phi(p)= 3
```

Dealing with points that change status: Once all points in the lists L_0, L_{-1}, L_1, L_{-2}, and L_2 have been visited, it is necessary to process points that have changed status during the iteration. This is described in Procedure B.4, summarized as follows:

1. Scan S_0. Move each point onto L_0 and update the value of the label map to 0.

2. Scan S_{-1} and S_1. Move each point onto L_{-1} or L_1 accordingly and update the label map to -1 or 1. If points on L_{-1} or L_1 have neighbors with ϕ values equal to -3 or 3, add the neighbors to the appropriate S_{-2} or S_2 lists while setting their ϕ value to be 1 unit from the value of the current point.

3. Scan S_{-2} and S_2. Move each point onto L_{-2} or L_2 accordingly and update the label map to -2 or 2 depending on the sign.

By running Procedure B.3 and Procedure B.4 a full iteration is completed and F can be re-computed based on the new position of the contour. This process is repeated until convergence is reached.

Efficiently updating statistics: When computing the movement of the contour, it is often desirable to track when a point crosses the zero level set thus changing from an *interior* point to an *exterior* point or vice versa. This can be accomplished by checking the sign of ϕ before and after adding F to the value of ϕ for points on the zero level set. (Line 2 of Procedure B.3).

When a point changes sign from negative to positive it can be added to a list L_{in2out} or L_{out2in} if the sign changes from positive to negative. Then, these two lists can be processed after the iteration to update statistics more efficiently. For example, if interior and exterior means are used when computing F, it is much more efficient to alter the statistics based on the movement the few points that cross the interface than to re-compute the means inside and outside the contour at each iteration.

B.2 Strengths and Limitations

The SFM has several properties that make it a valuable tool for image processing. The fact that it uses lists to keep track of points near the zero level set means that the speed of

Procedure B.4 Deal with points that are changing status

```
    \\ Move points into zero level set.
 1: for each point p in Sz
 2:    label(p) = 0, add p to Lz, remove p from Sz

    \\ Move points into -1 and +1 level sets
    \\ and ensure -2, +2 neighbors
 3: for each point in Sn1
 4:    label(p) = -1, add p to Ln1, remove p from Sn1
 5:    for each point q in N(p)
 6:        if(phi(q)==-3) phi(q)=phi(p)-1, add q to Sn2

 7: for each point in Sp1
 8:    label(p) =  1, add p to Lp1, remove p from Sp1
 9:    for each point q in N(p)
10:        if(phi(q)== 3), phi(q)=phi(p)+1, add q to Sp2

    \\ Move points into -2 and +2 level sets
11: for each point p in Sn2
12:    label(p) = -2, add p to Ln2, remove p from Sn2

13: for each point p in Sp2
14:    label(p) =  2, add p to Lp2, remove p from Sp2
```

curve updates is dependent only on the length of the curve and not the size of the image domain. This is particularly valuable when working with large medical imaging volumes. Also, the accuracy of this method as shown in [103] means that by using the SFM, the behavior implemented energies can be well understood without concern that the specific implementation is affecting the outcome.

One limitation is that because forces are only computed along the zero level set, it is not possible for new curves to appear spontaneously. Instead, the initial contour may only grow, split, and merge to obtain its final shape. While this is the theoretically, "correct" behavior, many implementations make use of the ability for new contours to emerge spontaneously to fill holes and bridge gaps in image information.

REFERENCES

[1] ACHENBACH, S., MOSELEWSKI, F., ROPERS, D., FERENCIK, M., HOFFMANN, U., MacNEILL, B., POHLE, K., BAUM, U., ANDERS, K., JANG, I., DANIEL, W. G., and BRADY, T. J., "Detection of calcified and noncalcified coronary atherosclerotic plaque by contrast-enhanced, submillimeter multidetector spiral computed tomography a segment-based comparison with intravascular ultrasound," *Circulation*, vol. 109, no. 1, pp. 14–17, 2004. 65

[2] AN, J., ROUSSON, M., and XU, C., "γ-convergence approximation to piecewise smooth medical image segmentation," in *Proceedings of Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, vol. 4792, pp. 495–502, 2007. 23

[3] ANGENENT, S., PICHON, E., , and TANNENBAUM, A., "Mathematical methods in medical image processing," *Bulletin of the American Mathematical Society*, pp. 365–396, 2005. 2, 6, 12

[4] ARSIGNY, V., FILLARD, P., PENNEC, X., and AYACHE, N., "Log-euclidean metrics for fast and simple calculus on diffusion tensors," *Magnetic Resonance in Medicine*, vol. 56, pp. 411–421, 2006. 57

[5] ATLAS, S. W., *Magnetic Resonance Imaging of the Brain and Spine; Fourth Edition.* Lippincott WWilliams and Wilkins, 2008. 7

[6] AWATE, S. P., ZHANG, H., and GEE, J. C., "A fuzzy, nonparametric segmentation framework for dti and mri analysis: With applications to dti-tract extraction," *IEEE Transactions on Medical Imaging*, vol. 26, no. 11, 2007. 53

[7] BAKER, S. and MATTHEWS, I., "Lucas-kanade 20 years on: A unifying framework," *International Journal of Computer Vision*, vol. 53, no. 3, pp. 221–255, 2004. 78, 80

[8] BASSER, P., PAJEVIC, S., PIERPAOLI, C., DUDA, J., and ALDROUBI, A., "In vivo fiber tractography using DT-MRI data," *Magnetic Resonance in Medicine*, vol. 44, no. 4, pp. 625–632, 2000. 51

[9] BASSER, P. and PIERPAOLI, C., "Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI," *Journal of Magnetic Resonance B*, vol. 111, no. 3, pp. 209–219, 1996. 51

[10] BETSER, A., VELA, P., PRYOR, G., and TANNENBAUM, A., "Flying in formation using a pursuit guidance algorithm," in *Proceedings of American Control Conference (ACC)*, 2005. 8

[11] BHATTACHARYYA, A., "On a measure of divergence between two statistical populations dened by their probability distributions," *Bulletin of the Calcutta Mathematics Society*, vol. 35, pp. 99–110, 1943. 30

[12] BLAKE, A. and ISARD, M., *Active Contours.* Cambridge: Springer, 1998. 1, 12, 18

[13] BROX, T. and CREMERS, D., "On the statistical interpretation of the piecewise smooth mumford-shah functional," in *Workshop on Scale Space and Variational Methods in Computer Vision*, vol. 4485, pp. 203–213, 2007. 23

[14] BROX, T. and WEICKERT, J., "Level set segmentation with multiple regions," *IEEE Transactions on Image Processing*, vol. 15, pp. 3213–3218, October 2006. 24, 32, 33

[15] BRUCE, V., GREEN, P., and GEORGESON, M., *Visual perception: Physiology, psychology, and ecology.* Psychology Press, 2003. 1

[16] BRUNNER, G., KURKURE, U., CHITTAJALLU, D., YALAMANCHILI, R., and KAKADIARIS, I., "Toward unsupervised classification of calcified arterial lesions," in *Proceedings of Medical Imaging and Computing and Computer Assisted Intervention (MICCAI)*, pp. 144–152, 2008. 64

[17] CASELLES, V., KIMMEL, R., and SAPIRO, G., "Geodesic active contours," *International Journal of Computer Vision*, vol. 22, pp. 61–79, Feb. 1997. 18, 20

[18] CHAN, T. and VESE, L., "Active contours without edges," *IEEE Transactions on Image Processing*, vol. 10, pp. 266–277, Feb. 2001. 17, 20, 28, 95

[19] CHOPP., D. L., "Computing minimal surfaces via level set curvature flow.," *Journal of Computational Physics*, vol. 106, no. 1, pp. 77–91, 1993. 18

[20] COLLINS, R., "Mean-shift blob tracking through scale space," in *CVPR*, vol. 2, pp. 234–240, 2003. 78

[21] CONTURO, T., LORI, N., CULL, T., AKBUDAK, E., SNYDER, A., SHIMONY, J., MCKINSTRY, R., BURTON, H., and RAICHLE, M., "Tracking neuronal fiber pathways in the living human brain," *Proceedings of the National Academy of Science*, vol. 96, no. 18, pp. 10422–10427, 1999. 51

[22] CREMERS, D., KOHLBERGER, T., and SCHNOERR, C., "Diffusion snakes: introducing statistical shape knowledge into the Mumford-Shah functional," *International Journal of Computer Vision*, vol. 50, pp. 295–313, Feb. 2002. 14

[23] CREMERS, D., ROUSSON, M., and DERICHE, R., "A review of statistical approaches to level set segmentation: Integrating color, texture, motion, and shape," *International Journal of Computer Vision*, vol. 72, no. 2, pp. 195–215, 2007. 20

[24] DELLAERT, F. and COLLINS, R., "Fast image-based tracking by selective pixel integration," in *Workshop on Frame-Rate Vision (ICCV)*, 1999. 81

[25] DESOUZA, G. and KAK, A., "Vision for mobile robot navigation: A survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 2, pp. 237–267, 2002. 2

[26] DOUCET, A., DEFREITAS, N., and GORDON, N., *Sequential Monte Carlo Methods in Practice.* Springer, 2001. 8

[27] EVANS, L. C. and SPRUCK, J., "Motion of level sets by mean curvature.," *Int. Journal of Differential Geometry*, vol. 33, no. 3, pp. 635–681, 1991. 17

[28] FLETCHER, T., TAO, R., JEONG, W., and WHITAKER, R., "A volumetric approach to quantifying region-to-region white matter connectivity in diffusion tensor MRI," *IPMI*, 2007. 52

[29] GOLLAND, P., GRIMSON, W., SHENTON, M., and KIKINIS, R., "Detection and analysis of statistical differences in anatomical shape," *Medical Image Analysis*, vol. 9, no. 1, pp. 69–86, 2005. 4

[30] GONZALEZ, R. and WOODS, R., *Digital Image Processing*. Prentice Hall, 2001. 1

[31] HA, J., ALVINO, C., PRYOR, G., NIETHAMMER, M., JOHNSON, E., and TANNEN-BAUM, A., "Active contours and optical flow for automatic tracking of flying vehicles," in *Proceedings of American Control Conference (ACC)*, 2004. 8

[32] HABER, E. and MODERSITZKI, J., "A multilevel method for image registration," *SIAM Journal on Scientific Computing*, vol. 27, no. 5, pp. 1594–1607, 2006. 4

[33] HAKER, S., SAPIRO, G., TANNENBAUM, A., and WASHBURN, D., "Missile tracking using knowledge-based adaptive thresholding: Tracking of high speed projectiles," in *Proceedings of International Conference on Image Processing*, pp. 174–180, 2001. 8

[34] HORN, B., *Robot Vision*. MIT Press, 1986. 1

[35] HU, W., TAN, T., WANG, L., and MAYBANK, S., "A survey on visual surveillance of object motion and behaviors," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 34, no. 3, pp. 334–352, 2004. 2, 8

[36] JACKOWSKI, M., KAO, C., QIU, M., CONSTABLE, R., and STAIB, L., "White matter tractography by anisotropic wavefront evolution and diffusion tensors imaging," *Medical Image Analysis*, vol. 9, pp. 427–440, 2005. 52

[37] JOHNSON, E., PROCTOR, A., HA, J., and TANNENBAUM, A., "Visual search automation for unmanned aerial vehicles," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 41, pp. 219–232, Jan 2005. 8

[38] JONASSON, L., BRESSON, X., HAGMANN, P., CUISENAIRE, O., MEULI, R., and THIRAN, J., "White matter fiber tract segmentation in DT-MRI using geometric flows," *Medical Image Analysis*, vol. 9, no. 3, pp. 223–236, 2005. 53

[39] JONASSON, L., HAGMANN, P., POLLO, C., BRESSON, X., WILSON, C., MEULI, R., and THIRAN, J., "A level set method for segmentation of the thalamus and its nuclei in DT-MRI.," *Signal Processing*, vol. 87, no. 2, pp. 309–321, 2007. 53

[40] KALMAN, R. E., "A new approach to linear filtering and prediction problems," *Tranactions of ASME–Journal of Basic Engineering*, 1960. 8

[41] KAO, C., OSHER, S., and QIAN, J., "Lax–Friedrichs sweeping scheme for static Hamilton–Jacobi equations," *Journal of Computational Physics*, vol. 196, no. 1, pp. 367–391, 2004. 52

[42] KAO, C., OSHER, S., and TSAI, Y., "Fast sweeping methods for static Hamilton-Jacobi equations," *SIAM journal on numerical analysis*, vol. 42, no. 6, pp. 2612–2632, 2005. 52

[43] KASS, M., WITKIN, A., , and TERZOPOULOS, D., "Snakes: active contour models," *International Journal of Computer Vision*, vol. 1, pp. 321–331, 1987. 14

[44] KICHENASSAMY, S., KUMAR, A., OLVER, P., TANNENBAUM, A., and JR., A. Y., "Conformal curvature flows: From phase transitions to active vision," *Archive for Rational Mechanics and Analysis*, vol. 134, pp. 275–301, Sept. 1996. 18

[45] KIM, J., FISHER, J., YEZZI, A., CETIN, M., and WILLSKY, A., "A nonparametric statistical method for image segmentation using information theory and curve evolution," *IEEE Transactions on Image Processing*, vol. 14, pp. 1486–1502, Oct. 2005. 20

[46] KUBICKI, M., WESTIN, C., NESTOR, P., WIBLE, C., FRUMIN, M., MAIER, S., KIKINIS, R., JOLESZ, F., McCARLEY, R., and SHENTON, M., "Cingulate fasciculus integrity disruption in schizophrenia: a magnetic resonance diffusion tensor imaging study," *Biological Psychiatry*, vol. 54, no. 11, pp. 1171–1180, 2003. 60

[47] LANKTON, S. and TANNENBAUM, A., "Localizing region-based active contours," *IEEE Transactions on Image Processing*, vol. 17, pp. 2029–2039, Nov. 2008. 11

[48] LANKTON, S., MELONAKOS, J., MALCOLM, J., DAMBREVILLE, S., and TANNENBAUM, A., "Localized statistics for dw-mri fiber bundle segmentation," in *Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA)*, Jun 2008. 11

[49] LANKTON, S., NAIN, D., YEZZI, A., and TANNENBAUM, A., "Hybrid geodesic region-based curve evolutions for image segmentation," in *Proceedings of SPIE: Medical Imaging*, vol. 6510, p. 65104U, Mar. 2007. 23

[50] LANKTON, S., NAKHMANI, J. M. A., and TANNENBAUM, A., "Tracking through changes in scale," in *International Conference on Image Processing ICIP*, pp. 241–244, 2008. 11

[51] LANKTON, S., STILLMAN, A., RAGGI, P., and TANNENBAUM, A., "Automatic soft plaque detection and vessel segmentation," in *Workshop on Probabilitstic Models in Medical Image Analysis (PMMIA)*, Sept 2009. 11

[52] LENGLET, C., ROUSSON, M., and DERICHE, R., "DTI segmentation by statistical surface evolution," *IEEE Transactions on Medical Imaging*, vol. 25, no. 6, pp. 685–700, 2006. 53

[53] LENGLET, C., ROUSSON, M., DERICHE, R., FAUGERAS, O., LEHERICY, S., and UGURBIL, K., "A Riemannian Approach to Diffusion Tensor Images Segmentation," in *Proceedings of International Conference on Information Processing in Medical Imaging (IPMI)*, pp. 591–602, 2005. 52

[54] LI, C., KAO, C.-Y., GORE, J., and DING, Z., "Minimization of region-scalable fitting energy for image segmentation," *Image Processing, IEEE Transactions on*, vol. 17, pp. 1940–1949, Oct. 2008. 23

[55] LUCAS, B. and KANADE, T., "An iterative image registration technique with and application to stereo vision," in *Int. Joint Conf. on Art. Intel.*, pp. 674–679, 1981. 78, 81

[56] MALCOLM, J., RATHI, Y., and TANNENBAUM, A., "A graph cut approach to image segmentation in tensor space," in *Workshop on Component Analysis Methods*, 2007. 53

[57] MALCOLM, J., RATHI, Y., YEZZI, A., and TANNENBAUM, A., "Fast approximate surface evolution in arbitrary dimension," in *Proceedings of SPIE Medical Imaging*, 2008. 48

[58] MARTIN, D., TAL, C. F. D., and MALIK, J., "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *Proceedings of International Conference on Computer Vision (ICCV)*, vol. 2, pp. 416–423, 2001. 2, 4

[59] MATTHEWS, I., ISHIKAWA, T., and BAKER, S., "The template update problem," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, pp. 810–815, June 2004. 78, 81

[60] MELONAKOS, J., MOHAN, V., NIETHAMMER, M., SMITH, K., KUBICKI, M., and TANNENBAUM, A., "Finsler tractography for white matter connectivity analysis of the cingulum bundle," in *Proceedings of Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2007. 52, 59

[61] MELONAKOS, J., PICHON, E., ANGENENT, S., and TANNENBAUM, A., "Finsler active contours," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007. in press. 52

[62] MELONAKOS, J., NIETHAMMER, M., MOHAN, V., KUBICKI, M., MILLER, J., and TANNENBAUM, A., "Locally-constrained region-based methods for dw-mri segmentation," in *Workshop on Mathematical Methods in Biomedical Image Analysis (MM-BIA)*, 2007. 53

[63] MICHAILOVICH, O., RATHI, Y., and TANNENBAUM, A., "Image segmentation using active contours driven by the bhattacharyya gradient flow," *IEEE Transactions on Image Processing*, vol. 15, pp. 2787–2801, November 2007. 20, 30

[64] MOBERTS, B., VILANOVA, A., and VAN WIJK, J., "Evaluation of Fiber Clustering Methods for Diffusion Tensor Imaging," *IEEE Transactions on Visualization*, pp. 9–9, 2005. 52

[65] MOREL, J.-M. and SOLIMINI, S., *Variational Methods for Image Segmentation*. Boston, MA: Brikhauser, 1994. 12, 18

[66] MORI, S., CRAIN, B., CHACKO, V., and VAN ZIJL, P., "Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging.," *Annals of Neurology*, vol. 45, no. 2, pp. 265–9, 1999. 51

[67] MORI, S. and VAN ZIJL, P., "Fiber tracking: principles and strategies- a technical review," *NMR in Biomedicine*, vol. 15, no. 7-8, pp. 468–480, 2002. 7, 51

[68] MUMFORD, D. and SHAH, J., "Optimal approximations by piecewise smooth functions and associated variational problems," *Communications in Pure and Applied Mathematics*, vol. 42, pp. 577–685, 1989. 23

[69] MUMFORD, D., "A bayesian rationale for energy functionals," in *Geometry Driven Diffusion in Computer Vision* (ROMENY, B., ed.), Kluwer Academic, Dordrecht, pp. 141–153, 1994. 23

[70] NAIN, D., *Scale-based decomposable shape representations for medical image segmentation and shape analysis.* PhD thesis, Georgia Institute of Technology, 2006. 4

[71] O'DONNELL, L., *Cerebral white matter analysis using diffusion imaging.* PhD thesis, Massachusetts Institute of Technology, 2006. 52

[72] OSHER, S. and FEDKIW, R., *Level Set Methods and Dynamic Implicit Surfaces.* New York, NY: Cambridge University Press, 2003. 14, 18, 25

[73] OSHER, S. and TSAI, R., "Level set methods and their applications in image science," *Communications in Mathematics and Science*, vol. 1, no. 4, pp. 1–20, 2003. 18, 25

[74] ODONNELL, L., HAKER, S., and WESTIN, C., "New Approaches to Estimation of White Matter Connectivity in Diffusion Tensor MRI: Elliptic PDEs and Geodesics in a Tensor-Warped Space," in *Proceedings of Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, vol. 2488, pp. 459–466, 2002. 52

[75] PARAGIOS, N., CHEN, Y., and FAUGERAS, O., *Handbook of Mathematial Models in Computer Vision.* New York, NY: Springer, 2005. 1, 18

[76] PARAGIOS, N. and DERICHE, R., "Geodesic active contours and level sets for the detection and tracking of moving objects," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, pp. 226–280, Mar. 2000. 12, 18

[77] PARAGIOS, N. and DERICHE, R., "Geodesic active regions: A new framework to deal with frame partition problems in computer vision," *International Journal of Compututer Vision*, vol. 46, pp. 223–247, Feb. 2002. 23

[78] PARKER, G., HAROON, H., and WHEELER-KINGSHOTT, C., "A framework for a streamline-based probabilistic index of connectivity(PICo) using a structural interpretation of MRI diffusion measurements," *Journal of Magnetic Resonance Imaging*, vol. 18, no. 2, pp. 242–254, 2003. 52

[79] PARKER, G., WHEELER-KINGSHOTT, C., and BARKER, G., "Estimating distributed anatomical connectivity using fast marchingmethods and diffusion tensor imaging," *IEEE Transactions on Medical Imaging*, vol. 21, no. 5, pp. 505–512, 2002. 52

[80] PENG, N.-S., YANG, J., and CHEN, J.-X., "Kernel-bandwidth adaptation for tracking object changing in size," in *ICIAR*, vol. 3212, pp. 581–588, 2004. 78

[81] PERRIN, M., POUPON, C., COINTEPAS, Y., RIEUL, B., GOLESTANI, N., PALLIER, C., RIVIERE, D., CONSTANTINESCO, A., and LE BIHAN, D., "Fiber tracking in Q-ball fields using regularized particle trajectories," in *Proceedings of International Conference on Information Processing in Medical Imaging (IPMI)*, 2005. 52

[82] PICHON, E., *Novel methods for multidimensional image segmentation.* PhD thesis, Georgia Institute of Technology, 2005. 52

[83] Pichon, E., Westin, C., and Tannenbaum, A., "A Hamilton-Jacobi-Bellman approach to high angular resolution diffusion tractography," in *Proceedings of Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, vol. 3749, pp. 180–187, 2005. 52

[84] Piovano, J., Rousson, M., and Papadopoulo, T., "Efficient segmentation of piecewise smooth images," in *Workshop on Scale Space and Variational Methods in Computer Vision*, vol. 4485, pp. 709–720, 2007. 23

[85] Pohle, K., Achenbach, S., MacNeill, B., Ropers, D., Ferencik, M., Moselewski, F., Hoffmann, U., Brady, T., kyung Jang, I., and Daniel, W. G., "Characterization of non-calcified coronary atherosclerotic plaque by multi-detector row CT: Comparison to IVUS," *Atherosclerosis*, vol. 190, pp. 174–180, 2007. 65

[86] Prados, E., Lenglet, C., Pons, J., Wotawa, N., Deriche, R., Faugeras, O., and Soatto, S., "Control Theory and Fast Marching Techniques for Brain Connectivity Mapping," in *Proceedings of Computer Vision and Pattern Recognition (CVPR)*, pp. 1076–1083, 2006. 52

[87] Qian, H., Mao, Y., Geng, J., and Wang, Z., "Object tracking with self-updating tracking window," in *PAISI*, vol. 4430, pp. 82–93, 2007. 78

[88] Rathi, Y., Michailovich, O., and Tannenbaum, A., "Segmenting images on the tensor manifold," in *Proceedings of Computer Vision and Pattern Recognition (CVPR)*, June 2007. 53

[89] Renard, F. and Yang, Y., "Image segmentation for detection of soft plaques in multidetector CT images," in *Proceedings of Southwest Symposium on Image Analysis and Interpretation*, pp. 121–124, 2008. 65

[90] Rousson, M. and Deriche, R., "A variational framework for active and adaptive segmentation of vector valued images," in *Workshop on Motion and Video Computing*, p. 56, 2002. 20

[91] Rousson, M., Lenglet, C., and Deriche, R., "Level set and region based propagation for diffusion tensor MRI segmentation," in *Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA)*, pp. 123–134, 2004. 20

[92] Rousson, M., Lenglet, C., and Deriche, R., "Level Set and Region Based Surface Propagation for Diffusion Tensor MRI Segmentation," in *Workshop on Computer Vision and Mathematical Methods in Medical and Biomedical Image Analysis (CVMMBIA)*, vol. 35, p. 123, 2004. 53

[93] Sapiro, G., *Geometric Partial Differential Equations and Image Analysis*. New York, NY: Cambridge University Press, 2003. 14, 18

[94] Saur, S., Alkadhi, H., Desbiolles, L., Szekely, G., and Cattin, P., "Automatic detection of calcified coronary plaques in computed tomography data sets," in *Proceedings of Medical Imaging and Computing and Computer Assisted Intervention (MICCAI)*, pp. 170–177, 2008. 64

[95] Schmahmann, J. and Pandya, D., *Fiber Pathways of the Brain*. Oxford University Press, 2006. 59, 60

[96] SCHROEDER, S., KOPP, A., and BURGSTAHLER, C., "Noninvasive plaque imaging using multislice detector spiral computed tomography," *Seminars in Thrombosis and Hemostasis*, vol. 33, no. 2, pp. 203–209, 2007. 64

[97] SETHIAN, J., *Level Set Methods and Fast Marching Methods, Second Edition.* New York, NY: Springer, 1999. 14, 17, 18

[98] STIELTJES, B., KAUFMANN, W., VAN ZIJL, P., FREDERICKSEN, K., PEARLSON, G., SOLAIYAPPAN, M., and MORI, S., "Diffusion Tensor Imaging and Axonal Tracking in the Human Brainstem," *NeuroImage*, vol. 14, no. 3, pp. 723–735, 2001. 51

[99] SUM, K. and CHEUNG, P., "Vessel extraction under non-uniform illumination: A level set approach," *IEEE Transactions on Biomedical Engineering*, vol. 55, pp. 358–360, Jan. 2008. 23

[100] VELA, P., NIETHAMMER, M., PRYOR, G., TANNENBAUM, A., BUTTS, R., and WASHBURN, D., "Knowledge-based segmentation for tracking through deep turbulence," *IEEE Transactions on Control Systems Technology*, vol. 16, pp. 469–474, May 2008. 8

[101] VIRMANI, R., BURKE, A., FARB, A., and KOLODGIE, F., "Pathology of the vulnerable plaque," *Journal of the American College of Cardiology*, vol. 47, no. 8-C, pp. 13–18, 2006. 65

[102] WANG, F., SUN, Z., CUI, L., DU, X., WANG, X., ZHANG, H., CONG, Z., HONG, N., and ZHANG, D., "Anterior Cingulum Abnormalities in Male Patients With Schizophrenia Determined Through Diffusion Tensor Imaging," 2004. 60

[103] WHITAKER, R., "A level-set approach to 3D reconstruction from range data," *International Journal of Computer Vision*, vol. 29, no. 3, 1998. 18, 47, 58, 94, 100

[104] YEZZI, A., TSAI, A., and WILLSKY, A., "A statistical approach to snakes for bimodal and trimodal imagery," in *Proceedings of International Conference on Computer Vision (ICCV)*, vol. 2, pp. 898–903, 1999. 20

[105] YEZZI, A., TSAI, A., and WILLSKY, A., "A fully global approach to image segmentation via coupled curve evolution equations," *Journal of Visual Communication and Image Representation*, vol. 13, pp. 195–216, Mar. 2002. 20, 28, 29

[106] YILMAZ, A., JAVED, O., and SHAH, M., "Object tracking: A survey," *ACM Computing Survey*, vol. 38, no. 4, pp. 1–45, 2006. 2, 8

[107] ZHANG, T. and FREEDMAN, D., "Tracking objects using density matching and shape priors," in *Proceedings of International Conference on Computer Vision (ICCV)*, pp. 1950–1954, 2004. 18

[108] ZHU, S. C. and YUILLE, A., "Region competition: Unifying snakes, region growing, and bayes/mdl for multiband image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, pp. 884–900, Sept. 1996. 19

[109] ZITOVA, B. and FLUSSER, J., "Image registration methods: a survey," *Image and vision computing*, vol. 21, no. 11, pp. 977–1000, 2003. 4