SURFACE-BASED IMAGE SEGMENTATION USING APPLICATION-SPECIFIC PRIORS

by

Gopalkrishna Veni

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Computing

School of Computing

The University of Utah

December 2016

The University of Utah Graduate School

STATEMENT OF DISSERTATION APPROVAL

The dissertation of **Gopalkrishna Veni**

has been approved by the following supervisory committee members:

Ross T. Whitaker	, Chair	**7/7/2016** Date Approved
Suyash P. Awate	, Member	**7/19/2016** Date Approved
Thomas P. Fletcher	, Member	**7/7/2016** Date Approved
Jeffrey M. Phillips	, Member	**7/7/2016** Date Approved
Robert S. MacLeod	, Member	**7/7/2016** Date Approved

and by **Ross T. Whitaker** , Chair/Dean of

the Department/College/School of **Computing**

and by David B. Kieda, Dean of The Graduate School.

ABSTRACT

Image segmentation entails the partitioning of an image domain, usually two or three dimensions, so that each partition or segment has some meaning that is relevant to the application at hand. Accurate image segmentation is a crucial challenge in many disciplines, including medicine, computer vision, and geology. In some applications, heterogeneous pixel intensities; noisy, ill-defined, or diffusive boundaries; and irregular shapes with high variability can make it challenging to meet accuracy requirements. Various segmentation approaches tackle such challenges by casting the segmentation problem as an energy-minimization problem, and solving it using efficient optimization algorithms. These approaches are broadly classified as either region-based or edge (surface)-based depending on the features on which they operate.

The focus of this dissertation is on the development of a surface-based energy model, the design of efficient formulations of optimization frameworks to incorporate such energy, and the solution of the energy-minimization problem using graph cuts. This dissertation utilizes a set of four papers whose motivation is the efficient extraction of the left atrium wall from the late gadolinium enhancement magnetic resonance imaging (LGE-MRI) image volume. This dissertation utilizes these energy formulations for other applications, including contact lens segmentation in the optical coherence tomography (OCT) data and the extraction of geologic features in seismic data.

Chapters 2 through 5 (papers 1 through 4) explore building a surface-based image segmentation model by progressively adding components to improve its accuracy and robustness. The first paper defines a parametric search space and its discrete formulation in the form of a multilayer three-dimensional mesh model within which the segmentation takes place. It includes a generative intensity model, and we optimize using a graph formulation of the *surface net* problem. The second paper proposes a Bayesian framework with a Markov random field (MRF) prior that gives rise to another class of surface nets, which provides better segmentation with smooth boundaries. The third paper presents a maximum a posteriori (MAP)-based surface estimation framework that relies on a generative image model by incorporating global shape priors, in addition to the MRF, within the Bayesian formulation. Thus, the resulting surface not only depends on the learned model of shapes,

but also accommodates the test data irregularities through smooth deviations from these priors. Further, the paper proposes a new shape parameter estimation scheme, in closed form, for segmentation as a part of the optimization process. Finally, the fourth paper (under review at the time of this document) presents an extensive analysis of the MAP framework and presents improved mesh generation and generative intensity models. It also performs a thorough analysis of the segmentation results that demonstrates the effectiveness of the proposed method qualitatively, quantitatively, and clinically.

Chapter 6, consisting of unpublished work, demonstrates the application of an MRF-based Bayesian framework to segment coupled surfaces of contact lenses in optical coherence tomography images. This chapter also shows an application related to the extraction of geological structures in seismic volumes. Due to the large sizes of seismic volume datasets, we also present fast, approximate surface-based energy minimization strategies that achieve better speed-ups and memory consumption.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

First, I thank my advisor, Ross Whitaker, who guided me during my PhD program. His invaluable guidance and support have shaped my PhD research over the years. I would also like to thank my committee members, Suyash Awate, Tom Fletcher, Jeff Phillips, and Rob McLeod, for their valuable feedback, helpful suggestions, and productive discussions.

Many thanks to Shireen Elhabian for her insightful suggestions in addressing some technical challenges related to this research. Special thanks to Suyash Awate, Josh Cates, Kristen Zygmunt, Ayla Khan, Shridharan Chandramouli, and Zhisong Fu, with whom I collaborated on various projects. Special thanks to Christine Pickett and Avantika Vardhan for proofreading my papers and this thesis and for their helpful suggestions.

I thank my fellow labmates for all the appealing discussions and the fun we had during this PhD span. I also thank the Scientific Computing and Imaging Institute for creating a great research environment and providing wonderful amenities.

Thanks to my parents and my brother for their love and belief in me and for supporting my decisions. I thank my wife, Poonam Ekhelikar, for her continuous support, love, and patience on this very long PhD journey.

CHAPTER 1

INTRODUCTION

Image segmentation deals with the partition of an image into multiple segments or regions. The extent of this partition depends on the context of the application at hand. Accurate segmentation is a ubiquitous demand in a variety of applications, including medicine, ophthalmology, and seismology. However, this accuracy is affected by significant challenges depending upon the application. These challenges include highly variable pixel intensities, noisy/ill-defined diffusive boundaries, and irregular shapes with high variability. Further, the objects of interest are often surrounded by undesired structures of similar appearance. In a broad sense, this dissertation focuses on developing a segmentation framework that aims to segment desired objects accurately and in an efficient manner.

Some standard and powerful image segmentation approaches include deformable models such as active contours [1] and geodesic active contours [2]; statistical models such as active shape models [3] and active appearance models [4]; path-based graph methods such as intelligent scissors [5], live-wire [6], and random walker [7]; spectral analysis techniques such as normalized cuts [8] and its variants; combinatorial graph cuts-based methods such as grab-cuts [9] and lazy-snapping [10], etc. All these approaches can be broadly classified into two groups based on the underlying image features on which they operate: *region-based* and *edge-based*. Region-based methods, also referred to as pixel-labeling methods, primarily focus on regional features by associating each image pixel with one of the coherent regions. Such features typically correspond to the intensity distributions of the different regions being segmented. On the other hand, edge-based methods (also referred to as surface-based methods) explicitly act upon the internal and external characteristics of the object boundaries. The internal characteristics try to keep the boundaries smooth, whereas the external characteristics drive the solution toward the desired objects of interest. In this dissertation, we deal with the second class of segmentation methods, because for the applications under consideration, the features of the desired objects are encoded in the object boundaries (across edges), especially when neighboring structures attain similar region-based features. Applications in this dissertation are the left atrial wall, the contact lens, and the horizons.

1.1 Motivation

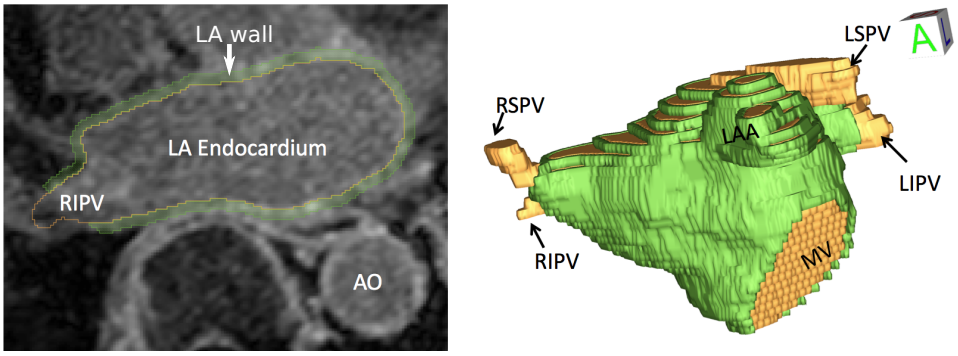
This research is motivated by a class of segmentation problems whose goal is to precisely extract either one or multiple surfaces. The specific applications dealt with in this dissertation include the automatic segmentation of the left atrium wall from late gadolinium enhancement magnetic resonance imaging (LGE-MRI) data in patients suffering from atrial fibrillation, segmentation of the contact lens in optical coherent tomography (OCT) image data, and extraction of horizons in geological data. In the following subsections, we will provide more details about each application, along with the corresponding challenges.

1.1.1 Left Atrial Wall from LGE-MRI Data

Atrial fibrillation (AF) is the most common type of cardiac arrhythmia, associated with an increased risk of mortality due to heart failure, dementia, and stroke [11–13]. It is characterized by irregular, rapid electrical impulses that are generated inside the atrial chambers of the heart, producing fibrosis (tissue remodeling) within the atrial wall. Catheter-based radio frequency (RF) ablation is considered to be one of the prominent procedures to treat AF, which entails scarring of the fibrotic tissue [14]. A successful RF ablation aids in isolation of the fibrosis-affected regions, thus helping cure atrial fibrillation.

The LGE-MRI imaging technique has proven effective in highlighting fibrosis, as well as in quantifying and tracking the influence of the RF ablation in the formation of scar tissue within the left atrial (LA) wall [15]. Thus, the assessment of the abnormal atrial tissue involves the following: LGE-MRI imaging, segmentation of the LA wall, and analysis of the fibrosis/scar regions within the wall. Current methods manually segment the LA wall, followed by the study of fibrosis/scar regions.

The manual segmentation of the LA wall requires the delineation of its inner and outer surfaces, namely the *endocardium* and *epicardium*, respectively. However, this method is time-consuming and requires sufficient anatomical expertise and extensive manpower. Further, it suffers from reproducibility among different expert image analysts. Automatic segmentation, on the other hand, tries to overcome these drawbacks by producing faster and more consistent results. As shown in Fig. 1.1, automatic LA wall extraction is challenging due to its thin structure, with noisy and occluded boundaries; complex geometry accompanied by localized intensity variations; surrounding tissues; and low signal-to-noise (SNR) ratio.



(a) (b)

Figure 1.1. Segmentation of the left atrial (LA) wall. (a) LA segmentation slice of LGE-MRI volume, (b) LA segmentation in three dimensions. RIPV - Right inferior pulmonary vein, AO - aorta, RSPV - right superior pulmonary vein, LSPV - left superior pulmonary vein, LAA - left atrium appendage, LIPV - left inferior pulmonary vein, and MV - mitral valve. Image courtesy: Joshua Cates with the Comprehensive Arrhythmia Research and Management (CARMA) Center, University of Utah.

1.1.2 OCT Data of Contact Lens

Optical coherence tomography (OCT) is a noninvasive imaging modality that uses light waves to capture micron-scale cross-sectional images of an object [16]. Due to its ability to produce submicrometer resolution, OCT imaging is extensively used in the quantitative assessment of cardiovascular- [17], dermatological- [18], ophthalmic-, and optometric-related applications, for example, in measuring the thickness of retinal layers [19], cornea [20], and contact lens [21], etc.

The second area of application involves the automatic segmentation of a contact lens in OCT images, which reflects the accurate delineation of inner and outer boundaries of various lens types characterized by their curvatures and thicknesses. Accurate segmentation of lens boundaries helps in performing quality control of the shapes of these lenses. The problem with OCT lens images is that they are corrupted from speckle noise with low SNR [22]. Further, the lens images especially suffer from correlated and uncorrelated noise and imaging artifacts, making the segmentation problem difficult to handle. Fig. 1.2 shows examples of OCT lens images that exhibit these challenges.

1.1.3 Geological Feature Extraction in Seismic Data

The third application is related to the extraction of horizons from the seismic data. Seismic horizons are unique layers of the earth's subsurface that lie between a sequence of sedimentary rocks of distinct mineral densities and porous characteristics [23]. They represent a transition in rock properties and play a key role in geoscientific interpretation. In

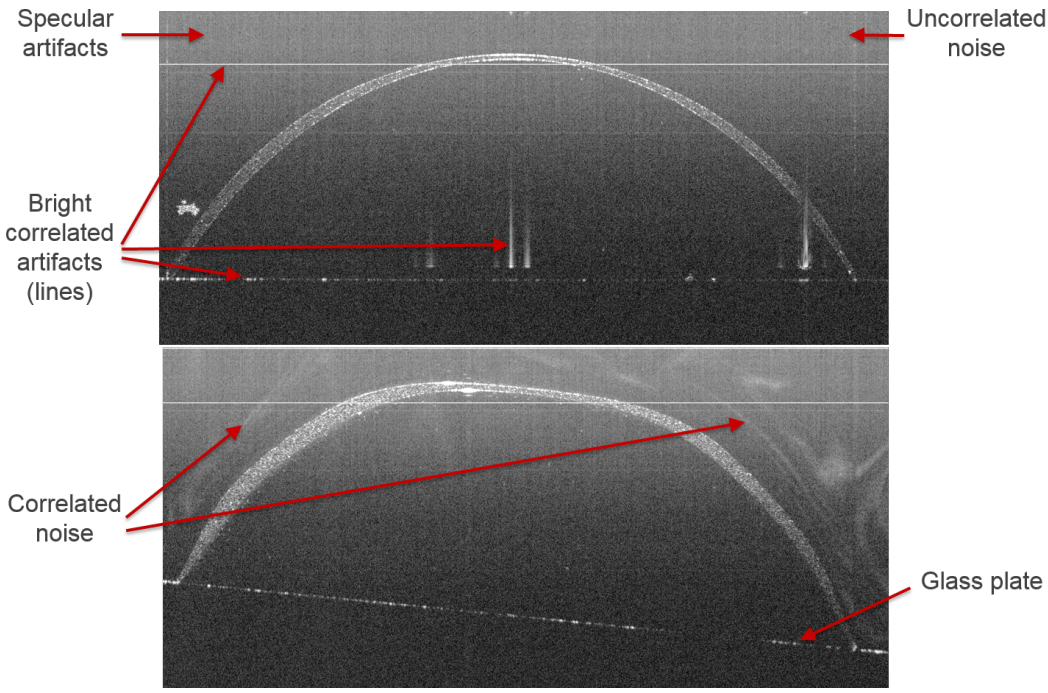


Figure 1.2. OCT lens images showing the challenges in segmentation.

this application domain, the segmentation task involves effective and efficient extraction of horizon surfaces in large 3D seismic volumes. The geometric interpretation of the extracted horizons leads to the identification of faults, channels, and gas leakages within the earth's subsurface [23]. Faults are the subvertical structures with shifted horizons on either side, thus affecting the continuity of horizons. Fig. 1.3 shows a slice of the seismic volume comprising a series of horizons and a couple of faults running across these horizons. The difficulty with segmenting the desired horizons is that they suffer from nonuniform intensity patterns and are surrounded by many unwanted horizons that possess similar intensity structures.

1.2 Segmentation as Energy Minimization

Real-world segmentation challenges often require prior knowledge of the object being segmented. This prior information could be either local, concerning the attributes of the desired surface, or global, which focuses on the general shape of the surface and its variability. For example, most real-world objects possess smooth boundaries. Such an inherent property could be incorporated into a segmentation algorithm by enforcing regularization constraints to ensure smoothness. In some applications, including medical images such as the LA wall, the object boundary gets diffused because the surrounding

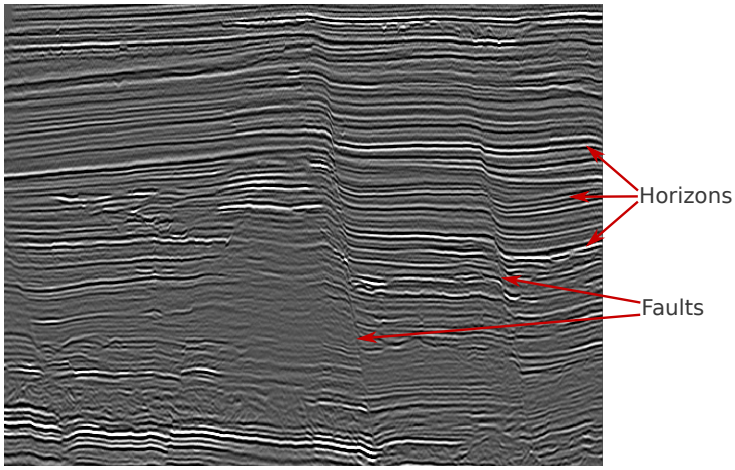


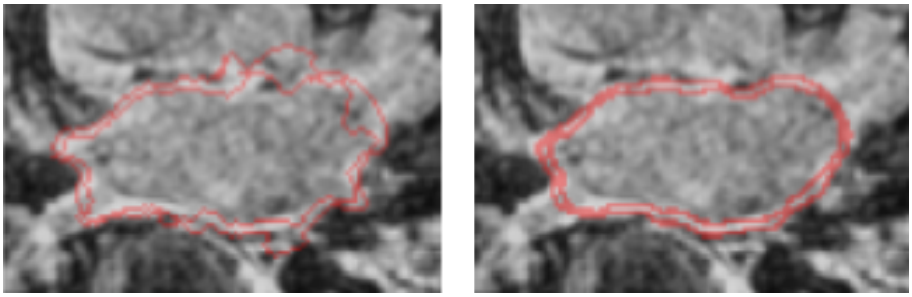
Figure 1.3. Two-dimensional (2D) slice of a $551 \times 426 \times 426$-sized seismic volume showing a series of horizons.

tissues obscure portions of the desired object, thus confusing any algorithm in restoring the actual boundary around these regions. Such weak-boundary segmentation problems are often solvable if one has some prior information of the required shapes. This notion of adding a prior triggers a strong motivation to use a Bayesian framework for segmentation.

One elegant way to solve the image segmentation problem within the Bayesian framework is by casting it as a maximum a posteriori (MAP) estimation. This estimation requires the minimization of the posterior energy such that it corresponds to the desired segmentation. These energies typically combine regional/surface properties of desired objects along with some prior information. The energy-based segmentation methods can be classified into two groups depending upon the type of energy function they use and the optimization approach to minimize it.

The first category defines an energy functional in the space of continuous domain $\mathbb{R}^{\infty}$. These functionals are minimized using variational approaches and gradient descent schemes, which are numerically solved approximately on discrete grid by using finite differences or finite elements. The methods that follow this approach are generally referred to as deformable-surface models. Examples of such models include snakes [24], geodesic active contours [2], and methods relying on level sets [25]. Typically, these methods are unstable as they are prone to get trapped in local minima - a problem that we explicitly address in this research. Fig. 1.4a explains these drawbacks by showing the segmentation results of the left atrium wall boundaries using level sets. Further, they are not robust in the case of low-quality segmentation problems and are sensitive to initializations.

The second category formulates the solution of energy function in the finite-dimensional



(a) (b)

Figure 1.4. Motivation behind using global over local optimization strategies. Surface boundaries of the LA wall obtained by using (a) local optimum (level sets) vs. (b) global optimal methods (graph cuts).

(discrete) space $\mathbb{Z}^D$. Such energies are minimized based on combinatorial optimization algorithms. Some of these algorithms include brute-force search, dynamic programming, graph-based, and graph cuts-based approaches. These approaches produce globally optimal solutions by searching the entire solution spaces and are, therefore, insensitive with respect to initializations. However, the search space can be exponential ($O(k^N)$, where k represents a categorical variable and N denotes the number of variables in the domain. For instance, a brute-force search takes an exponential time to solve an energy function by seeking every possible solution in the search space. Thus, we need smarter algorithms that provide better efficiency than exponential time algorithms.

Graph cuts are one class of algorithms that can solve specific, discrete energy functions in polynomial time (see Subsection 1.3). Considering the applications at hand, where objects to be segmented suffer from ill-defined and missing boundaries due to imaging aspects and adjacent structures with similar boundary features, these kinds of segmentation problems can be formulated as a particular class of energy minimization problems. Next, these energies are solved by transforming them into graph cuts problems that can be solved in polynomial time. Fig. 1.4b shows an advantage of using graph cuts in accurately segmenting the left atrium wall.

1.2.1 Graph Cuts

Graph cuts are classic combinatorial optimization algorithms defined over a graph, G, comprising vertices, $\mathcal{V}$, and edges, $\mathcal{E}$, that connect them. The edges carry nonnegative costs/weights that are determined based on problem specifications. A *cut* is defined over a set of edges that separates the entire vertex set into two disjoint subsets. Fig. 1.5 illustrates a graph example and a cut passing through a set of edges. An optimal cut is determined by

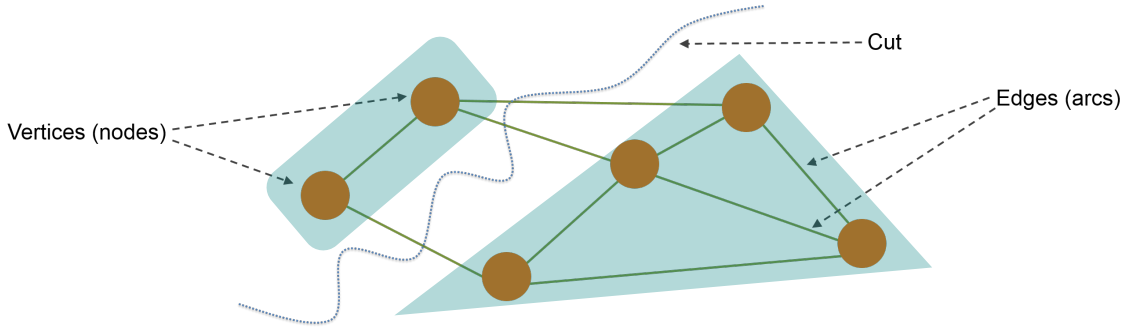


Figure 1.5. Schematic showing a graph along with a cut that separates graph vertices into two disjoint subsets (blue polygons).

finding a set of edges that produces an aggregate minimum cost. Therefore, graph cuts are also called minimum cut (*min-cut*) problems. The relationship between an energy function and a graph is maintained by encoding the likelihood energy on the vertices of the graph and the prior on the graph edges. This energy is then minimized by finding the min-cut in the derived graph.

In the network flow theory, finding a min-cut in the graph, G, is equivalent to computing its maximum flow (*max-flow*) [26] in the transformed graph, G_{st} by means of two distinguished vertices. These vertices are commonly referred to as the source, s, and the sink, t. Given a flow network, a flow, f, is defined as a real-valued function that satisfies the capacity and flow conservation constaints [27]. It quantifies the amount of flow transfer from the source to the sink. The value of a flow is the aggregate flow either from the source to vertices or from the vertices to the sink:

$$|f| = \sum_{v \in V} f(s, v) = \sum_{v \in V} f(v, t). \tag{1.1}$$

The goal is to maximize $|f|$ through a set of nonsaturated edges, $\mathcal{E}$, until no more flow can be sent. Max-flow algorithms are used in numerous applications related to the bipartite matching, algorithm design, industrial, and commercial problems, etc.

A variety of combinatorial optimization algorithms are available to solve min-cut, or its dual, max-flow algorithms. The next subsection covers the basic mechanism behind the standard optimization algorithms that are relevant to this research. Specifically, this familiarity helps one understand how to efficiently optimize graph cuts on large-sized image volumes, as we will discuss in Chapter 6 of this dissertation. Initially, we will summarize two general classes of optimization algorithms: **augmented path** and **push-relabel**. This summary is followed by a variant of the augmented path algorithm, namely the Boykov-

Kolmogorov (**BK**) algorithm, which is especially designed to solve the image analysis-related problems.

1.2.2 Maximum Flow Algorithms

Generally, maximum flow algorithms are classified into two groups depending upon the way they perform. *Augmented path*-based methods [26] repetitively search for augmented paths through a set of unsaturated edges in order to push the flow from *source* to *sink*. This process continues until all augmented paths are exhausted. *Push-relabel*-based methods [28] take a local approach to push the flow by working on one node at a time. These methods maintain a *height* and *excess* flow for each node in order to perform push and relabel operations. The process starts by pushing an excess flow from the source to its neighbors. Next, it continues to adjust the heights of nodes by relabeling them while pushing flow between nodes until all the flow reaches the sink. Since the push relabel algorithm involves local operations, it is suitable for massive-sized data applications, such as horizon extraction, where all the data can be divided into multiple blocks and are run in parallel.

In [29], Boykov and Kolmogorov proposed a modified version of the augmented-path algorithm (*BK* algorithm) in order to improve its empirical performance in vision-related problems. The algorithm essentially maintains two nonoverlapping search trees starting from the source and the sink. When the two trees meet each other through a set of nonsaturated edges, an augmented path is formed and the flow is sent from source to sink. In the case of computer vision applications, the BK algorithm achieves the best empirical performance compared to any other max-flow algorithms [29]. Therefore, it has been used in all our serial max-flow algorithms.

1.2.3 Energy Minimization and Graph Cuts

Greig et al. [30] pioneered in applying graph cuts to the energy-minimization-based vision problems. They minimized a specific class of energy functions by formulating them as a maximum a posteriori, MAP, estimate of a Markov random field (MRF). For a given image, $\mathcal{I}$, this energy class is specific to a set of binary variables, $x_i \in \{0, 1\}$. The corresponding energy is represented as

$$E(x) = \sum_{i \in \mathcal{I}} E_d(x_i) + \sum_{\{i,j\} \in \mathcal{N}} E_c(x_i, x_j) \tag{1.2}$$

where $\mathcal{N}$ represents a set of pairwise neighboring pixels. E_d represents a data-dependent likelihood function that quantifies the desired features of x with respect to the given image domain, and E_c represents a prior, which models a pairwise interaction term in the form

of MRF. Typically, E_c defines the coherency among neighborhood features. To globally minimize the binary-valued energy defined in (1.2), Greig et al. used a graph cuts-based optimization technique. The basic strategy involves the construction of a graph such that the min-cut of the graph corresponds to the minimization of the energy. Later, Boykov and Jolly [31] extended this idea to general-purpose image segmentation problems, in which an image segmentation is represented as a pixel-labeling problem over which the energy is defined. The likelihood function evaluates the regional properties of a given image, and the prior quantifies the discontinuity in labeling among neighbors. Fig. 1.6 illustrates a simple 2D example that shows the equivalence of image segmentation and pixel-labeling-based graph cuts.

Global minimization of an arbitrary energy function is NP-hard [32]. Under certain conditions, energy minimization can be reduced to a graph cut problem, for which the global minimum can be computed in polynomial time. Kolmogorov and Zabih [32] characterize a certain class of energy functions and establish the fact that regularity is the necessary and sufficient condition for these energy functions to be graph representable. For example, consider a binary-labeled energy function, $E(x)$, as expressed in (1.2), belonging to the $\mathcal{F}^2$ class, where each $x_i \in \{0, 1\}$. We call this function regular only if every term of (1.2) satisfies the following inequality constraint:

$$E_c(0,0) + E_c(1,1) \leq E_c(0,1) + E_c(1,0) \tag{1.3}$$

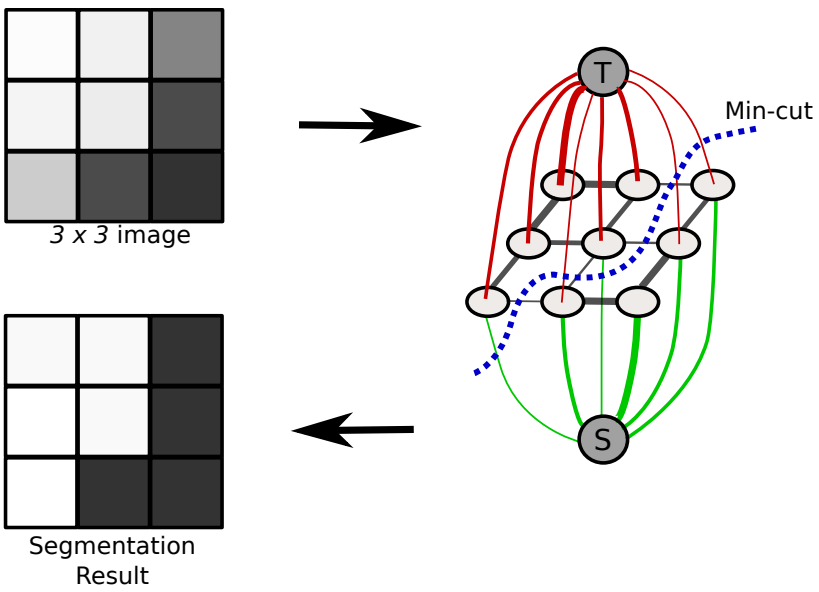


Figure 1.6. Schematic showing a transformation of the 2D image into an s-t graph, G_{st}. The min-cut of this graph corresponds to the segmentation of that image. The thickness of the edges reflects their costs.

that is, the average energy cost imposed by assigning different feature variables must be at least the average energy cost imposed by assigning the same feature variables. According to the continuous optimization theory, a set of functions that satisfies (1.3) is referred to as a *convex function*. Analogously, in the case of discrete optimization, these functions are often called *submodular functions* [33], [34].

In the case of pixel-labeling problems, apart from the submodularity property, efficient global energy minimization is possible via graph cuts only on condition that the corresponding labels are binary valued [35]. In order to handle the multilabel case, one needs to resort to the approximation algorithms by repeatedly minimizing an energy function with respect to two labels at a time. Some efficient approximation methods include α-expansion, $\alpha-\beta$-swap [35], and range moves [36]. Although these methods are fast and empirically effective in many image segmentation-based applications, they do not achieve the global minimum, but do reach local minima (in a strong sense) [35]. Thus, they might not be effective for a set of problems involving segmentation of multiple labeling regions, such as the left atrial wall extraction and the segmentation of the layers of a contact lens. Ishikawa [37] pointed out that the global optimum for E is possible even in the general (multilabel) case by enforcing convexity with respect to a linearly ordered label set in the prior term, E_c. Nevertheless, these labels need to be ordered and are restricted to convex interaction penalties over label-pairs.

Another drawback related to the graph cuts-based pixel-labeling-type energy functions is that their likelihood costs correspond to the regional properties instead of boundary properties, which is essential in the case of weak-boundary segmentation problems. In addition, their prior term characterizes coherency among neighboring pixels instead of boundary/surface attributes, thus disturbing its regularity. Further, it is very difficult to incorporate shape information in their graph structure, which is necessary in the case of the left atrial wall segmentation.

To overcome the deficiencies of the pixel-labeling-based graph cut approaches, we formulate the energy in (1.2) with respect to the surface, $\mathcal{S}$, instead of the label. Here, a surface $\mathcal{S}$ is defined as a subset of vertices, $\mathcal{S} \subset \mathcal{V}$, in a specially designed graph. Thus, the corresponding energy becomes

$$E(\mathcal{S}) = \sum_{i \in \mathcal{I}} E_d(s_i) + \sum_{\{i,j\} \in \mathcal{C}} E_c(s_i, s_j) \tag{1.4}$$

where $\mathcal{C}$ represents the clique, which defines the neighborhood structure while constructing a graph. The likelihood term, $E_d(s_i)$, determines features related to the desired sur-

face/surfaces, and the prior term $E_c(s_i, s_j)$ defines its inherent characteristics. The primary advantage of considering these types of energy functions is that they are capable of simultaneously extracting multiple interacting surfaces in a globally optimal manner.

Surface-based image segmentation takes place using a search space (parameterization) whose structure approximately conforms with the resulting surface estimate. Since the search space required to formulate the surface estimation problem is in the form of a graph structure, a sampling strategy is used, where the image data are sampled at locations that are associated with the geometric model, as shown in Fig. 1.7(*left*). Based on the combinatorial optimization theory, this class of energy minimization can be cast as an optimal *net surface* problem.

In order to define a net surface, a special geometric graph construction is required whose vertices are arranged logically as a set of parallel columns with each column carrying the same number of vertices. A graph that follows this kind of vertex arrangement is defined as a *multicolumn* graph. A net surface (*net*) is a subgraph in a multicolumn graph that intersects with each column at exactly one vertex. Fig. 1.7(*right*) illustrates an example of a net surface in a multicolumn graph. To find an optimal net surface in a general multicolumn graph is a NP-hard problem [38]. Under certain conditions, Wu and Chen introduced two net surface problems [38], namely V-weight and VCE-weight net surface problems, and proved their global optimality in low-order polynomial time. These conditions correspond to the properly ordered configuration of edges between adjacent columns of the multicolumn graph [38]. V-weight net surface problems basically deal with vertices that carry real-valued weights, whereas VCE-weight net surface problems deal with vertices as well as constrained edges that carry real-valued weights to find an optimal net. Both algorithms use graph cuts to obtain polynomial time solutions ($T(n, m)$ with $O(n)$ vertices and $O(m)$ edges) in a properly ordered multicolumn graph. The ramifications of transforming these net surface problems as graphs cuts problems are explained in [38].

1.3 Proposed Surface-Based Image Segmentation

This section discusses the pitfalls of standard optimal surface estimation models and outlines a series of solutions to address each issue in achieving a better segmentation.

In order to build a multicolumn graph, which maintains a topological structure similar to that of the desired surface, one needs to account for the complexity of the surface. For regular surfaces, some standard approaches involving cylindrical coordinate transform [39] and electrical lines of forces [40] can be used. However, in the case of complex and irregular anatomical structures, for example the left atrium, these approaches are inapplicable. The

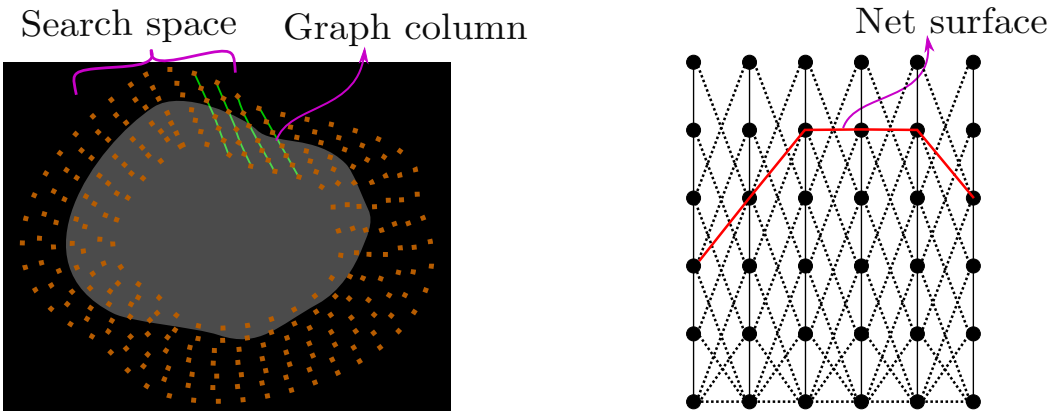


Figure 1.7. (*Left*) Schematic showing a search space whose topology approximately complies with the surface being estimated. (*Right*) Multicolumn graph and a net surface.

graph constructed by using these approaches produces straight columns that tangle with each other, thus resulting in a self-intersecting surface extraction. To address the column-tangling problem, Chapter 2 proposes a new sophisticated multilayer mesh-generation strategy based on an iterative relaxation of dynamic particles [41]. Chapter 2 also discusses the choice of the likelihood term, $E_d(\mathcal{S})$, in (1.4) in the form of a generative intensity model. Such a model is derived from a set of intensity profiles around the object boundaries learned from the training data.

Li and Sonka used the V-weight net surface approach to simultaneously extract multiple interacting surfaces in a number of image segmentation-based applications [42–45]. Although their approach produces efficient globally optimal solutions, it is not cast in the form of an energy minimization problem. As a result, it cannot be formulated as a MAP estimation problem. In other words, it limits the incorporation of prior information inherent to the surface being segmented. Chapter 3 proposes a Bayesian formulation with a Markov random field (MRF) prior. Such a framework brings about a second type of surface net problem, namely *VCE*-weight net, which encodes submodular (convex) soft penalties on a properly ordered edge configuration between adjacent columns. The energy minimization involving the MRF prior that is convex over properly ordered edges of a multicolumn graph can be transformed into a min-cut problem, which can be solved in a low-order polynomial time [37], [38]. This prior is designed not only to maintain surface regularity, but also to define the interrelations in the case of multiple surfaces. The immediate benefit of this approach is its ability to extract smoother as well as more accurate surfaces.

As previously mentioned in Section 1.1, the MRF-based Bayesian surface segmentation framework has been applied in all three areas of applications. Specifically, Chapter 3

analyzes the capability of the proposed framework on the left atrial data. Chapter 6 discusses the implementation details and results corresponding to the segmentation of various types of lenses characterized by their curvature and thickness, followed by the interactive extraction of horizons based on the user-specified landmarks.

Enforcing MRF-based soft geometric constraints in the Bayesian segmentation framework essentially favors the local properties of the surface by regularizing them. Nonetheless, it will be advantageous to also incorporate the prior information about the shapes of the desired object. Such global characteristics pertaining to the shapes will further improve the segmentation accuracy and robustness, especially in the case of weak-boundary segmentation problems, such as LA. In other words, inclusion of global shape priors within the Bayesian formulation addresses ill-posed segmentation problems by biasing solutions toward an object class of interest. Chapters 4 and 5 present a shape-based image segmentation algorithm based on a generative image model by incorporating both local and global shape priors. The proposed method uses MAP formulation that relies on graph cuts and parameterized shape prior estimation, thus providing a global updates-based optimization strategy. Further, the resulting solution not only relies on the shape prior but also accommodates irregularities in the test data through controlled deviations from the prior, thus maintaining regularity on output surfaces. These deviations are in turn derived from the submodular MRF-based local priors.

Along with the detailed formulation of shape-based surface segmentation modeling, Chapter 5 presents a fast nested mesh-building strategy and a modified generative intensity model. The new nested mesh-building strategy is based on the gradient vector flow [46] over an implicit shape representation in the form of signed distance transform. With respect to the generative intensity model that operates on intensity profiles, collected from the training data, there naturally exists some covariance among these profiles. The modified framework accounts for this covariance structure to build a better intensity model.

While working on massive image volumes, as is the case with horizon extraction, the corresponding graph structure, generated by using the VCE-weight net surface algorithm, is equivalently large. This form of graph structure reflects the computation time and memory usage of the graph cut algorithm, restricting its applicability to the real-time practice. It is thus necessary to increase the speed of the algorithm either by parallelizing graph cuts or by adopting approximation schemes that produce desirable outputs. Chapter 6 discusses a set of attempts that are made to improve the performance of the graph cuts-based energy minimization schemes. Our first attempt tries to parallelize graph cuts by dividing the

multicolumn graph into multiple subgraphs and processing them in parallel on multicore processors. We use a push-relabel-based max-flow strategy, which takes a local approach by operating on individual nodes one at a time. Further, we adopt a level-synchronized global relabeling (LSGR) strategy to minimize false sharing of vertices between multiple cores. The second attempt uses fast-approximate efficient graph-cut algorithms, namely α-expansion, $\alpha - \beta$-swap [35], that show compelling performance in terms of running time and memory usage on large-image volumes.

1.4 Overview and Summary of Contributions

Chapters 2 through 5 in this dissertation are reprints of published articles, accepted and under review. Although there exist some mathematical notation inconsistencies between papers, we tried to maintain the content as it stands in the original version. To summarize,

- Chapter 2 develops a novel, sophisticated mesh-generation strategy to construct a properly ordered multicolumn graph that avoids surface tangling and a learned model of intensity patterns to induce costs on graph vertices.

- Chapter 3 formulates the segmentation problem as a MRF-based Bayesian surface extraction algorithm by inducing soft penalties to maintain regularities on surface smoothing and surface coupling.

- Chapter 4 designs a shape-based surface estimation system that relies on a generative image model by incorporating both local and global shape priors.

- Chapter 5 proposes a gradient-vector flow-based nested mesh-building strategy in order to construct a multicolumn graph and a covariance-carrying generative intensity model in the shape-based surface-segmentation framework.

- Chapter 6 presents the application of a MRF-based Bayesian surface extraction algorithm to lens segmentation in OCT data and horizon extraction in geological data. It also discusses other graph cuts-based optimization strategies to improve the performance of energy-specific minimization schemes for large-sized structured data.

1.5 References

[1] A. Blake and M. Isard, *Active Contours: The Application of Techniques from Graphics, Vision,Control Theory and Statistics to Visual Tracking of Shapes in Motion*, 1st ed. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 1998.

[2] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," *Int. J. Comp. Vision*, vol. 22, no. 1, pp. 61–79, 1997.

[3] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models-their training and application," *Comput. Vision and Image Understanding*, vol. 61, no. 1, pp. 38–59, 1995.

[4] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 23, no. 6, pp. 681–685, 2001.

[5] E. N. Mortensen and W. A. Barrett, "Interactive segmentation with intelligent scissors," *Graphical Models and Image Process.*, vol. 60, no. 5, pp. 349–384, 1998.

[6] A. X. Falcão, J. K. Udupa, S. Samarasekera, S. Sharma, B. E. Hirsch, and R. d. A. Lotufo, "User-steered image segmentation paradigms: Live wire and live lane," *Graphical Models and Image Process.*, vol. 60, no. 4, pp. 233–260, 1998.

[7] L. Grady, "Random walks for image segmentation," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 28, no. 11, pp. 1768–1783, 2006.

[8] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 22, no. 8, pp. 888–905, 2000.

[9] C. Rother, V. Kolmogorov, and A. Blake, "Grabcut: Interactive foreground extraction using iterated graph cuts," *ACM Trans. Graph. (ToG)*, vol. 23, no. 3, pp. 309–314, 2004.

[10] Y. Li, J. Sun, C.-K. Tang, and H.-Y. Shum, "Lazy snapping," *ACM Trans. Graph. (ToG)*, vol. 23, no. 3, pp. 303–308, 2004.

[11] E. J. Benjamin, D. Levy, S. M. Vaziri, R. B. D'Agostino, A. J. Belanger, and P. A. Wolf, "Independent risk factors for atrial fibrillation in a population-based cohort: The Framingham heart study," *JAMA*, vol. 271, no. 11, pp. 840–844, 1994.

[12] A. S. Go, E. M. Hylek, K. A. Phillips, Y. Chang, L. E. Henault, J. V. Selby, and D. E. Singer, "Prevalence of diagnosed atrial fibrillation in adults: National implications for rhythm management and stroke prevention: The anticoagulation and risk factors in atrial fibrillation (atria) study," *JAMA*, vol. 285, no. 18, pp. 2370–2375, 2001.

[13] K. Higuchi, M. Akkaya, N. Akoum, and N. F. Marrouche, "Cardiac mri assessment of atrial fibrosis in atrial fibrillation: Implications for diagnosis and therapy," *Heart*, vol. 100, no. 7, pp. 590–596, 2014.

[14] P. Jaïs, R. Weerasooriya, D. C. Shah, M. Hocini, L. Macle, K.-J. Choi, C. Scavee,

M. Haïssaguerre, and J. Clémenty, "Ablation therapy for atrial fibrillation (af)," *Cardiovascular Res.*, vol. 54, no. 2, pp. 337–346, 2002.

[15] M. Daccarett, T. J. Badger, N. Akoum, N. S. Burgon, C. Mahnkopf, G. Vergara, E. Kholmovski, C. J. McGann, D. Parker, J. Brachmann *et al.*, "Association of left atrial fibrosis detected by delayed-enhancement magnetic resonance imaging and the risk of stroke in patients with atrial fibrillation," *J. Amer. College Cardiology*, vol. 57, no. 7, pp. 831–838, 2011.

[16] J. G. Fujimoto, C. Pitris, S. A. Boppart, and M. E. Brezinski, "Optical coherence tomography: An emerging technology for biomedical imaging and optical biopsy," *Neoplasia*, vol. 2, no. 1, pp. 9–25, 2000.

[17] H. G. Bezerra, M. A. Costa, G. Guagliumi, A. M. Rollins, and D. I. Simon, "Intracoronary optical coherence tomography: A comprehensive review: Clinical and research applications," *JACC: Cardiovascular Interventions*, vol. 2, no. 11, pp. 1035–1046, 2009.

[18] S. Chua, "High-definition optical coherence tomography for the study of evolution of a disease," *Dermatology Bulletin*, vol. 26, no. 1, pp. 2–3, 2015.

[19] E. Z. Blumenthal, J. M. Williams, R. N. Weinreb, C. A. Girkin, C. C. Berry, and L. M. Zangwill, "Reproducibility of nerve fiber layer thickness measurements by use of optical coherence tomography," *Ophthalmology*, vol. 107, no. 12, pp. 2278–2282, 2000.

[20] S. Sin and T. L. Simpson, "The repeatability of corneal and corneal epithelial thickness measurements using optical coherence tomography," *Optometry & Vision Sci.*, vol. 83, no. 6, pp. 360–365, 2006.

[21] B. J. Kaluzny, W. Fojt, A. Szkulmowska, T. Bajraszewski, M. Wojtkowski, and A. Kowalczyk, "Spectral optical coherence tomography in video-rate and 3d imaging of contact lens wear," *Optometry & Vision Sci.*, vol. 84, no. 12, pp. E1104–E1109, 2007.

[22] B. R. Davidson and J. K. Barton, "Application of optical coherence tomography to automated contact lens metrology," *J. Biomed. Optics*, vol. 15, no. 1, pp. 016 009–016 009, 2010.

[23] D. Patel, S. Bruckner, I. Viola, and E. M. Groller, "Seismic volume visualization for horizon extraction," in *Vis. Symp. (PacificVis), 2010 IEEE Pacific.* IEEE, 2010, pp. 73–80.

[24] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *Int. J Comput. Vision*, vol. 1, no. 4, pp. 321–331, 1988.

[25] R. T. Whitaker, "Volumetric deformable models: Active blobs," in *Visualization Biomedical Comp.* Inter. Society for Optics and Photonics, 1994, pp. 122–134.

[26] L. R. Ford and D. R. Fulkerson, "Maximal flow through a network," *Can. J. Math.*, vol. 8, no. 3, pp. 399–404, 1956.

[27] T. H. Cormen, *Introduction to Algorithms.* Cambridge, MA: MIT Press, 2009.

[28] A. V. Goldberg and R. E. Tarjan, "A new approach to the maximum-flow problem," *J. ACM (JACM)*, vol. 35, no. 4, pp. 921–940, 1988.

[29] Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 26, no. 9, pp. 1124–1137, 2004.

[30] D. Greig, B. Porteous, and A. H. Seheult, "Exact maximum a posteriori estimation for binary images," *J. Roy. Statist. Soc. Series B (Methodological)*, vol. 51, no. 2, pp. 271–279, 1989.

[31] Y. Y. Boykov and M.-P. Jolly, "Interactive graph cuts for optimal boundary & region segmentation of objects in nd images," in *Comput. Vision, 2001. ICCV 2001. Proc. Eighth IEEE Int. Conf. on*, vol. 1. IEEE, 2001, pp. 105–112.

[32] V. Kolmogorov and R. Zabin, "What energy functions can be minimized via graph cuts?" *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 26, no. 2, pp. 147–159, 2004.

[33] L. Lovász, "Submodular functions and convexity," in *Mathematical Programming The State of the Art*. Springer, 1983, pp. 235–257.

[34] A. Delong, "Advances in graph-cut optimization: Multi-surface models, label costs, and hierarchical costs," Ph.D. dissertation, University of Western Ontario, LDN, ON, Oct 2011.

[35] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Trans. Patt. Anal. Mach. Intell*, vol. 23, no. 11, pp. 1222–1239, 2001.

[36] O. Veksler, "Graph cut based optimization for mrfs with truncated convex priors," in *Comput. Vision and Patt. Recognition, 2007. IEEE Conf. on*. IEEE, 2007, pp. 1–8.

[37] H. Ishikawa, "Exact optimization for markov random fields with convex priors," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 25, no. 10, pp. 1333–1336, 2003.

[38] X. Wu and D. Z. Chen, "Optimal net surface problems with applications," in *Automata, Languages and Programming: 29th International Colloquium, ICALP 2002 Málaga, Spain, July 8–13, 2002 Proc.* Berlin, Heidelberg: Springer Berlin Heidelberg, 2002, pp. 1029–1042.

[39] M. Sonka, V. Hlavac, and R. Boyle, *Image Processing, Analysis, and Machine Vision*. Boston, MA: Cengage Learning, 2014.

[40] Y. Yin, Q. Song, and M. Sonka, "Electric field theory motivated graph construction for optimal medical image segmentation," in *Graph-Based Representations in Patt. Recognition*. Berlin, Heidelberg: Springer, 2009, pp. 334–342.

[41] M. Meyer, R. M. Kirby, and R. Whitaker, "Topology, accuracy, and quality of isosurface meshes using dynamic particles," *IEEE Trans. Vis. Comput. Graphics*, vol. 13, no. 6, pp. 1704–1711, 2007.

[42] K. Li, X. Wu, D. Z. Chen, and M. Sonka, "Optimal surface segmentation in volumetric images-a graph-theoretic approach," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 28, no. 1, pp. 119–134, 2006.

[43] M. K. Garvin, M. D. Abràmoff, X. Wu, S. R. Russell, T. L. Burns, and M. Sonka, "Automated 3d intraretinal layer segmentation of macular spectral-domain optical coherence tomography images," *IEEE Trans. Med. Imag.*, vol. 28, no. 9, pp. 1436–1447, 2009.

[44] X. Dou, X. Wu, A. Wahle, and M. Sonka, "Globally optimal surface segmentation using regional properties of segmented objects," in *Comput. Vision and Patt. Recognition, 2008, IEEE Conf. on.* IEEE, 2008, pp. 1–8.

[45] I. Oguz and M. Sonka, "Logismos-b: Layered optimal graph image segmentation of multiple objects and surfaces for the brain," *IEEE Trans. Med. Imag.*, vol. 33, no. 6, pp. 1220–1235, 2014.

[46] C. Xu and J. L. Prince, "Snakes, shapes, and gradient vector flow," *IEEE Trans. Image Process.*, vol. 7, no. 3, pp. 359–369, 1998.

CHAPTER 2

PROPER ORDERED MESHING OF COMPLEX SHAPES AND OPTIMAL GRAPH CUTS APPLIED TO ATRIAL-WALL SEGMENTATION FROM DE-MRI

This chapter is reprinted, with permission, from

Gopalkrishna Veni, Zhisong Fu, Suyash P. Awate, and Ross T. Whitaker, **Proper ordered meshing of complex shapes and optimal graph cuts applied to atrial-wall segmentation from DE-MRI.** In Proceedings of the 2013 IEEE 10th International Symposium on Biomedical Imaging (ISBI), April 2013.

2013 IEEE 10th International Symposium on Biomedical Imaging:
From Nano to Macro
San Francisco, CA, USA, April 7-11, 2013

PROPER ORDERED MESHING OF COMPLEX SHAPES AND OPTIMAL GRAPH CUTS APPLIED TO ATRIAL-WALL SEGMENTATION FROM DE-MRI

Gopalkrishna Veni, Zhisong Fu, Suyash P. Awate, Ross T. Whitaker

Scientific Computing and Imaging (SCI) Institute, University of Utah

ABSTRACT

Segmentation of the left atrium wall from delayed enhancement MRI is challenging because of inconsistent contrast combined with noise and high variation in atrial shape and size. This paper presents a method for left-atrium wall segmentation by using a novel sophisticated mesh-generation strategy and graph cuts on a proper ordered graph. The mesh is part of a template/model that has an associated set of learned intensity features. When this mesh is overlaid onto a test image, it produces a set of costs on the graph vertices which eventually leads to an optimal segmentation. The novelty also lies in the construction of proper ordered graphs on complex shapes and for choosing among distinct classes of base shapes/meshes for automatic segmentation. We evaluate the proposed segmentation framework quantitatively on simulated and clinical cardiac MRI.

Index Terms— Atrial Fibrillation, Mesh Generation, Geometric Graph, Minimum s-t cut, Optimal surfaces.

1. INTRODUCTION

In the context of imaging, delayed enhancement MRI (DE-MRI) produces contrast in myocardium and in regions of fibrosis and scarring, which are associated with risk factors and treatment of atrial fibrillation (AF). DE-MRI is therefore useful for evaluating the potential effectiveness of radio ablation therapy and for studying recovery. Automatic segmentation of the heart wall in this context is quite important; in a single clinic, hundreds of man hours are spent per month in manual segmentation.

Automatic segmentation of the heart wall in DE-MRI is quite challenging, because of relatively low and inconsistent contrast, high level of unwanted texture and noise, and high variability of atrial shape. Figure 1(a) shows typical DE-MRI images of the left atrium (LA). Several conventional segmentation methods have been ineffective.

The authors would like to acknowledge the Comprehensive Arrhythmia Research and Management (CARMA) Center, and the Center for Integrative Biomedical Computing (CIBC) by NIH Grant P41 GM103545-14, and National Alliance for Medical Image Computing (NAMIC) through NIH Grant U54 EB005149, for providing Utah fibrosis data, and CIBAVision.

Several papers address the problem of segmentation of blood pool in images from MRI angiography (MRI-A) protocols [1, 2]. These methods take advantage of the relatively homogeneous brightness of the blood pool in MRI-A, which is well suited for deformable models or registration-based approaches. However, high-quality properly-aligned blood-pool images are often *not* readily available from DE-MRI protocols. Further, the atrial wall is relatively thin in DE-MRI, confounding algorithms like template registration that often rely on coarse anatomical features. Deformable-surface methods that rely on gradient descent optimizations, including level sets, are unable to deal with the large variations in boundary contrast. Statistical models, such as active shape models, which rely on a low-dimensional subspace of learned models, have been proven to be too inflexible in dealing with the small and large-scale shape variability, and they also suffer from being trapped in local minima during optimization. While recent developments to address this problem (such as [3]) are promising, they rely on deformable models and/or image registration approaches. In our experience, they tend to get caught in local minima, and are *not* particularly reliable — a problem that we explicitly address in this paper.

The difficulty of optimizing shape or surface models in the presence of weak signal, high variability, and high noise, suggests that this problem would benefit from an optimization strategy that seeks global optima. Wu and Chen [4] described a strategy that represents a segmentation problem as a minimum *s-t* cut on a proper ordered graph, which is solved (globally) by a polynomial-time algorithm. Later, it was extended by Li *et al.* [5] to simultaneously segment multiple coupled surfaces, by incorporating offset constraints via the graph construction. The approach has demonstrated some success in several challenging segmentation problems [6, 7].

The standard proper ordered graph technique is not applicable on complex and irregular anatomical structures, particularly LA. The graph constructed from these structures results in "tangling" between columns. This does not comply with the underlying assumption of topological smoothness which breaks the graph-cut model. Thus, these proper ordered graph-cut methods require a careful construction of the underlying graph. We propose a new method for the construction of a proper ordered graph that avoids tangling. The construction is carried out by a nested set of triangular meshes

978-1-4673-6455-3/13/$31.00 ©2013 IEEE

through a set of prisms, which form columns of a proper order graph. The feature detectors on each node of the graph are also learned from the input data. Because of the variability in shape, we cluster the training examples into a small collection of shape templates. The algorithm automatically selects the best template for a particular test image based on the correlation. The evaluation has been carried out on a set of synthetic examples and LA DE-MRI images with hand segmentations as the ground truth.

2. METHODS

A graph is a pair of sets $G = (V, E)$, which are vertices, $\{v_i\}$, and edges, $\{e_{i,j} = (v_i, v_j)\}$, respectively. For a proper ordered graph, the vertices are arranged logically as a collection of parallel columns that have the same number of vertices. The position of each vertex within the column is denoted by a superscript, e.g. v_i^l. The collection of vertices at the same position across all columns is called a *layer*. We let N be the number of columns and L be the number of vertices in each column (number of layers).

The construction of the derived directed graph is based on the method proposed by [8]. Here, the weight of each vertex in the innermost layer, the *base* layer, is given by $w_i^0 = c_i^0$. Every vertex in this layer is connected by a directed edge to every other vertex with a cost $+\infty$ in its adjacent columns. This makes the base layer strongly connected. For each vertex in layer $l \in [1, L-1]$, a weight of $w_i^l = c_i^l - c_i^{l-1}$ is assigned to each vertex. Again, a directed edge $e_{i,i}^{l,l-1}$ with a cost $+\infty$ is connected from that vertex to the one below it.

A pair of directed edges $e_{i,j}^{l,l-\Delta s}$ and $e_{j,i}^{l,l-\Delta s}$ with costs $+\infty$ go from a vertex v_i^l to a vertex $v_j^{l-\Delta s}$ and from v_j^l to a vertex $v_i^{l-\Delta s}$ making them an ordered pair. The Δs parameter controls the deviation in cuts between one column and its neighbors. To transform this graph into the *s-t* graph, G_{st}, two special nodes, called the *source* and the *sink* are added. The edges connecting each vertex to either the source or sink depend upon the sign of its weight. In case the weight on the vertex is negative, an edge with capacity equal to the absolute values of the weights of the corresponding vertex is directed from a source to that vertex; otherwise an edge is directed from that vertex to the sink. For simultaneous segmentation of multiple interacting surfaces, disjoint subgraphs are constructed as above and are connected with a series of directed edges defined by Δl and Δu parameters. These edges enforce the lower and upper inter surface constraints (described in [8]).

These edge capacities combined with the underlying topology of the graph determine the *minimum s-t cut* of the graph. The optimal surface is obtained by finding a minimum closed set Z^* in G^{st} [4]. Thus, a k^{th} surface in each sub graph is recovered by the intersection of the uppermost vertex of each layer in its respective sub graph and the minimum closed set Z^*. The computation of Z^* is done using *minimum s-t cut* algorithm, which produces global optima in polynomial time.

2.1. Building a Valid Proper Ordered Mesh

The previous section describes the graph topology based on a triangle structure within a layer, relying on vertex cost associated with image properties. To do so, we associate with each vertex v_i^l, a 3D position $\vec{x}_i^l$, which corresponds to a position in the image (volume). Here we describe how to assign 3D positions to mesh vertices and to triangulate each layer so that the layers form a nested set of watertight meshes in 3D. We call this set of vertices, their 3D positions, and the prismatic topology of the nested meshes, a proper ordered (PO) mesh.

For constructing the PO-mesh we use an extension of the dynamic-particle-system method proposed by Meyer *et al.* [9] for meshing implicit surfaces. We build this mesh using a template shape (described in the next section), which approximates very roughly the LA that we intend to segment. This template shape is represented as the zero level set of a signed distance transform in a volume. We first describe, very briefly how to build a mesh for the zero level set of this template.

Point or particles are distributed on an implicit surface by interactively minimizing a potential function. The potential function is defined pairwise between points and decreases monotonically with distance, $U_{i,j}^{l,l} = \Phi(|\vec{x}_i^l - \vec{x}_j^l|)$, and thus particles repel each other. We denote the sum of this collection of repulsive potentials within each layer as $\mathcal{R}$. These particle systems have been shown to form consistent, nearly regular packings on complex surface [9]. Once points have been distributed on an implicit surface (with sufficient density), a Delaunay tetrahedralization scheme can be used to build a water-tight triangle mesh of the surface [10].

To build a nested set of surface meshes, we require a collection of offset surfaces, both inside and out, that not only inherent the topology of the base surface, but also represent valid, watertight 3D triangle meshes. This is essential, because the cuts, which mix vertices from different layers, must also form watertight triangle meshes. Thus, it results to bend the columns in order to avoid tangling of columns/triangles as the layers extend outward from the mean shape. For this, we introduce a collection of particle systems, one for each layer in the graph/mesh, and we couple these particles by an attractive force (Hooks law) between layers. Thus, there is an additional set of potentials of the form $U_{i,i}^{l,l+1} = |\vec{x}_i^l - \vec{x}_i^{l+1}|^2$, and we denote the sum of the attractive forces of neighboring particles between layers as $\mathcal{A}$.

To optimize a collection of particle systems for L layers, we perform gradient descent, using asynchronous updates, as in [9], on the total potential $\mathcal{R} + \beta\mathcal{A}$. Figure 1(c) shows a nester 3-layered mesh for one of the LA templates.

The parameter β controls the relationship between attraction across layers and repulsion within layers and is tuned to

prevent tangling. For this paper, we have used $\beta = 10$. The optimization requires an initial collection of particles. So, we place a particle at each point where the adjacent voxels have values on either side of the level set. This gives an average density of approximately one particle per unit surface area (in voxel units). The physical distance between layers must be inversely proportional to the particle density within layers, and this is a compromise between the tangling that results from large offsets and the extra computation associated with many thin layers. This corresponds to around 14,000 particles for heart images and 2000 particles for simulated images.

2.2. Learning Template Meshes and Feature Detectors

This section describes the construction of template shapes and the mechanism for computing costs on nodes from input images. Due to the high variability of shapes of LA, we rely on a *training set* of 32 DE-MRI images with LA segmentations. These training images provide (i) a way of constructing a collection of PO graphs, so that new images can be segmented as cuts through one of these graphs and (ii) examples of intensity profiles for the features that define epi- and endo-cardial surfaces, which lead to costs at each node in the PO graph.

We begin by clustering the examples (roughly) based on shape. For this, we compute distance transforms, represented as volumes, of each endocardial surface. Training images are aligned via translation so that they all have the same center of mass for the blood pool (region bounded by the endocardium). We then compute clusters using k-means with a metric of mean-squared-distance between these volumes. We choose the number of clusters, based on the cluster residual curve, to be 5, and we removed one of these clusters from the test, because it contained only two (high distorted) examples. Surface meshes associated with the distance-transform means of these 4 clusters are shown in Figure 1(b).

The cost assigned to each vertex is designed to reflect the degree to which that vertex is a good candidate for the desired boundary or surface, which will be found via a graph cut. Here we use the training data to derive an intensity profile along a line segment, or *stick*, perpendicular to the surface associated with each vertex. We sample the stick at a spacing of one voxel. The intensity along each stick on each vertex of each template is computed by a weighted average of intensities of sticks for each feature point in each training image. Thus, for a particular vertex in a particular cluster, the intensities on a stick would be a Gaussian weighted average, with standard deviation of 2 voxels, of several nearby sticks from different images (that share the same blood-pool center). The costs are computed via a normalized cross correlation. Figure 1(d) shows a diagram of the stick configuration and several stick intensity profiles for parts of a particular template.

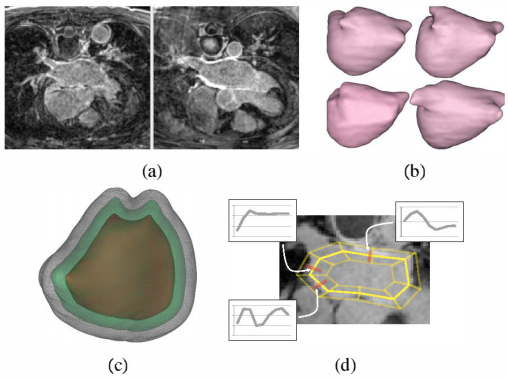


Fig. 1. (a) MRI examples showing low contrast and uneven background. (b) Examples of average shapes, derived from k-means clustering on distance transforms of training images, around which the PO-meshes are constructed. (c) An example of several layers of PO-meshes for the LA. (d) A mock up of a simplified PO-mesh in 2D with examples of feature detectors learned from the training data (actually, PO-meshes for the LA have over 400,000 vertices)

2.3. Segmentation by Graph Cuts

To segment a particular test image, we rely on user input to position the template, by specifying the center of the atrium. The algorithm is robust to this position, as long as the layers of the template do not lie outside or inside the desired surface. We sample the input image along all of the sticks at all nodes and compute the correlation with the template. This results in costs, weights, edge capacities, and then an optimal cut. The geometric parameters used for constructing our graph include $\Delta s = 4$, $\Delta l = 3$ and $\Delta u = 10$ which reflect the complexity of surfaces and the inter-surface separation between them. These values have been employed on all datasets including synthetic and LA images that proves algorithm's robustness. We employ all of the learned templates to the input image, choosing the segmentation that produces the best average correlation with the local intensity models for the optimal cut. Once the segmented mesh is retrieved from the cut, it is scan converted to reproduce the segmented volume(s). The processing time to find an optimal cut is few seconds.

3. EXPERIMENTS AND RESULTS

The experiments were evaluated on 100 simulated images of size $64 \times 64 \times 96$ voxels each and 32 DE-MRI images of the LA of size $400 \times 400 \times 107$ voxels. The simulated images include two oblong non-crossing surfaces with the inner surface translated randomly (Gaussian distribution) in 3D to mimic variations in heart-wall thickness; each image was corrupted with Rician noise ($\sigma = 20$ for the underlying Gaussian model) and a smoothly-varying bias field.

In all of our experiments, 30 mesh layers were generated, spaced at 0.5 voxels each, which gives each template a cap-

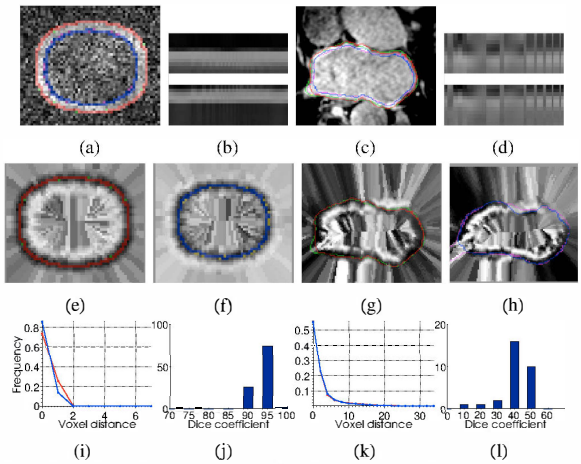


Fig. 2. (a),(c) Results of our algorithm on simulated and LA images (red and green curves represent outer boundaries extracted by our algorithm and ground truth, blue and purple curves represent inner boundaries extracted by our algorithm and ground truth) (b),(d) intensity profiles for outer/epicardium and inner/endocardium surfaces. (e),(g) our algorithm results and ground truth boundaries overlaid on outer and epicardial cost function images. (f),(h) our algorithm results and ground truth boundaries overlaid on inner and endocardial cost function images. (i),(k) distance histogram plot (in mm) between the surfaces extracted by our algorithm and ground truth over all images for both surfaces. (j),(l) histogram of dice coefficients for the middle region/myocardium to represent the number of images against the percentage overlap.

ture range of approximately 15 voxels. In case of simulated data, 50 training data sets and 50 test sets were considered for analysis. We evaluate the segmentation accuracy for LA based on leave-one out strategy for a test dataset, against templates from the training data.

Figures 2(a,c,e,f,g,h) illustrate our segmented boundaries for epicardial (outer) and endocardial (inner) heart-wall surfaces. These boundaries are overlaid on the original data (a,c) and their corresponding cost functions (e-h). Figure 2(b,d) shows the intensity profiles for outer and inner models.

We quantify the segmentation accuracy using distance metric. The distance metric is based on the aggregate of pairwise distances between corresponding points on the ground truth and our segmentation. For each point on our segmented surface, we measure the distance to the nearest point on the ground truth; and vice versa. The histograms of these measured distances (Figure 2(i,k)), indicate the percentage of voxels, on either surface, which were a specific distance away from the other surface. For a perfect delineation of the boundary, all these distances would be zero. The curves indicate the power of our algorithm in extracting boundaries very close to the real surfaces even in such challenging conditions.

To evaluate the overlap quantitatively for myocardium (heart wall), we used Dice measures between the ground truth and our segmented regions. Figure 2(j,l) shows the histograms of Dice measures. For synthetic data, the Dice values indicate excellent matches. However, in the case of the myocardium, the Dice values are little lower due to its varying thinness (2-6mm) and undefined ground truth. The ground truth is a single hand segmentation from an expert. Therefore, much of the observed error is near the veins which is subjected to the indefinite cutoff between atrium and vessel. Otherwise, the results are always close. While more overlap with the wall is desirable, these results also reflect the difficulty of quantifying efficacy using overlap. For instance, the ground truth results for the wall do not always form a complete boundary around the blood pool (even ignoring the vessels), and experiments show that when using overlap, human raters disagree significantly, as much as 50% from the average, demonstrating very high overlaps in pairwise comparisons between raters. Thus, these results appear to be clinically usable for many cases, but will still require human experts to do quality control, making slight corrections with manual tools. Furthermore, we expect results will be improved if we can increase the set of training images and form more templates in order to better match a given input image.

4. REFERENCES

[1] John, M., Rahn, N.: Automatic left atrium segmentation by cutting the blood pool at narrowings. MICCAI **8** (2005) 798-805.

[2] Karim, R., Mohiaddin, R., Rueckert, D.: Left atrium segmentation for atrial fibrillation ablation. SPIE. Volume 6918. (2008)

[3] Gao, Y., Gholami, B., MacLeod, R., Blauer, J., Haddad, W., Tannenbaum, A.: Segmentation of the endocardial wall of the left atrium using local region-based active contours and statistical shape learning. SPIE. Volume 7623. (2010)

[4] Wu, X., Chen, D.Z.: Optimal net surface problems with applications. In: Proc. Int. Colloquium on Automata, Languages and Programming. (2002) 1029–1042

[5] Li, K., Wu, X., Chen, D., Sonka, M.: Optimal surface segmentation in volumetric images-a graph-theoretic approach. IEEE Trans. PAMI, **28**(1) (2006) 119 –134

[6] Song, Q., Wu, X., Liu, Y., Smith, M., Buatti, J., Sonka, M.: Optimal graph search segmentation using arc-weighted graph for simultaneous surface detection of bladder and prostate. MICCAI **12** (2009) 827–835

[7] Li, X., Chen, X., Yao, J., Zhang, X., Tian, J.: Renal cortex segmentation using optimal surface search with novel graph construction. MICCAI **14** (2011) 387–394

[8] Sonka, M., Hlavac, V., Boyle, R.: Image processing, analysis and computer vision (2008)

[9] Meyer, M., Kirby, R., Whitaker, R.: Topology, accuracy, and quality of isosurface meshes using dynamic particles. IEEE Trans. Vis. Comp. Graph. **12**(5) (2007) 1704–11

[10] Amenta, N., Bern, M., Eppstein, D.: The crust and the beta-skeleton: Combinatorial curve reconstruction. Graphic Models Image Proc. **60**(2) (1998) 125–135

CHAPTER 3

BAYESIAN SEGMENTATION OF ATRIUM WALL USING GLOBALLY-OPTIMAL GRAPH CUTS ON 3D MESHES

This chapter is reprinted with permission from

Bayesian Segmentation of Atrium Wall Using Globally-Optimal Graph Cuts on 3D Meshes. *Gopalkrishna Veni, Zhisong Fu, Suyash P. Awate, and Ross T. Whitaker.* In Proceedings of the Information Processing in Medical Imaging (IPMI). Springer Berlin Heidelberg, pages 656-667, 2013.

Bayesian Segmentation of Atrium Wall using Globally-Optimal Graph Cuts on 3D Meshes

Gopalkrishna Veni, Zhisong Fu, Suyash P. Awate, Ross T. Whitaker*

Scientific Computing and Imaging (SCI) Institute, University of Utah

Abstract. Efficient segmentation of the left atrium (LA) wall from delayed enhancement MRI is challenging due to inconsistent contrast, combined with noise, and high variation in atrial shape and size. We present a surface-detection method that is capable of extracting the atrial wall by computing an optimal a-posteriori estimate. This estimation is done on a set of nested meshes, constructed from an ensemble of segmented training images, and graph cuts on an associated multi-column, proper-ordered graph. The graph/mesh is a part of a template/model that has an associated set of learned intensity features. When this mesh is overlaid onto a test image, it produces a set of costs which lead to an optimal segmentation. The 3D mesh has an associated weighted, directed multi-column graph with edges that encode smoothness and inter-surface penalties. Unlike previous graph-cut methods that impose hard constraints on the surface properties, the proposed method follows from a Bayesian formulation resulting in soft penalties on spatial variation of the cuts through the mesh. The novelty of this method also lies in the construction of proper-ordered graphs on complex shapes for choosing among distinct classes of base shapes for automatic LA segmentation. We evaluate the proposed segmentation framework on simulated and clinical cardiac MRI.

Keywords: Atrial Fibrillation, Bayesian segmentation, Minimum s-t cut, Mesh Generation, Geometric Graph.

1 Introduction

Segmentation of the heart's left atrium (LA) is a highly relevant problem in the clinical domain. In the context of medical imaging, delayed enhancement MRI (DE-MRI) has been shown to produce contrast in myocardium (heart wall) and in regions subjected to fibrosis and scarring [1]. So, these regions are associated with risk factors and treatment of atrial fibrillation (AF). Imaging with DE-MRI is therefore useful for the evaluation of potential effectiveness of radio-ablation therapy and for studying recovery. This AF recovery includes analysis of scarring as well as atrial shape and structural remodeling (SRM) after treatment. Automatic segmentation of the heart wall in this context is quite important; in

* The authors would like to acknowledge the Comprehensive Arrhythmia Research and Management (CARMA) Center, and the Center for Integrative Biomedical Computing (CIBC) by NIH Grant P41 GM103545-14, and National Alliance for Medical Image Computing (NAMIC) through NIH Grant U54 EB005149, for providing Utah fibrosis data, and CIBAVision.

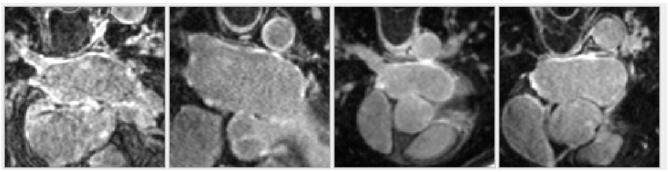


Fig. 1. Slices of left atrium DE-MRI images showing the challenges in segmentation.

a single clinic, hundreds of man hours are spent per month in manual segmentation. In DE-MRI images, this becomes a challenging task, because of relatively low and inconsistent contrast, high level of unwanted texture and noise, and high variability of atrial shape. Moreover, this problem gets aggravated by inaccuracies in cardiac gating and the SRM in chronic AF. Thus, this is a difficult image analysis problem, which also represents an ubiquitous challenge in a 3D medical segmentation: segmenting in the presence of relatively poor signal, high noise, and large variations in shape.

Several papers address the problem of segmenting the blood pool in MRI angiography (MRI-A) images [2, 3]. These methods make use of the relatively homogeneous brightness of the blood pool in MRI-A, which is well suited for deformable models or registration-based approaches. However, high-quality properly-aligned blood-pool images are often *not* readily available from DE-MRI protocols. Furthermore, due to thinness of the atrium wall, algorithms based on template registration fail as they often rely on coarse anatomical features. Figure 1 shows examples of DE-MRI images of the LA that depict its varying, low-contrast boundaries, high level of correlated noise, and high shape variability.

A variety of conventional segmentation methods have proven to be ineffective. One strategy to address these challenges is to introduce a prior on the segmentation problem, either in the form of probability on specific kinds of shapes or more generally on shape properties, such as smoothness. These priors are combined with image matching terms or simply feature detection to find some ideal compromise between the prior and the data. Level-set methods [4] rely on gradient-descent optimizations, which are sensitive to initializations and local minima. We have found such local optimizations to be particularly ill suited to this problem. Statistical models, such as active shape models [5] have been proven to be effective, but are also limited in their ability to deal with the small and large-scale shape variability. Generally, coarse-to-fine optimization strategies can help avoid local minima, but have proven inadequate for this segmentation problem, mostly because the features of interest (thin, brighter regions and small dark gaps between the atrium and nearby tissues) do not hold up under blurring. While recent developments addressing this problem [6] are promising, they rely on deformable models and/or image registration approaches that tend to also get caught in local minima.

The difficulty of segmentation in this context suggests that this problem would benefit from a global optimization strategy. Recently, Wu and Chen [7]

described a scheme by which the problem of finding an optimal function value on a discrete grid (a *surface net* problem) is represented as a minimum$s$-t cut on a proper-ordered graph. Optimal solutions to the s-t cut are given by relatively efficient, polynomial-time algorithms. Li *et al.* [8] applied a version of this surface-net formulation to simultaneously segment multiple coupled surfaces in noisy images by including image-based costs and geometric constraints of the underlying graph. That approach has demonstrated some success in several challenging image segmentation problems [8–10]. This surface-net relies on the construction of a properly-ordered graph, which also defines the topology of the resulting segmentation. The construction of such graphs is challenging for complex and irregular anatomical structures, such as LA. Using naive offsets from a base mesh results in "tangling" between columns, and resulting cuts are not guaranteed to be valid surfaces or regions. Thus, these proper ordered graph-cut methods require a careful construction of the underlying graph.

The contributions of this paper are as follows. We show that a Bayesian formulation with a Markov random field prior can give rise to a certain type of surface-net problem, namely, a *VCE-Net*, which is solvable by the algorithm of Wu and Chen [7]. This formulation gives rise to soft penalties on surface smoothing and surface coupling, which, as we will show, is superior to the hard contraints described by Li *et al.*. The Bayesian formulation also gives rise to a set of *learned* feature detectors, so that the method does not rely on user-defined methods for characterizing edges or regions. We also propose a new method for the graph construction on irregular surfaces that avoids tangling. To address the variability in shapes, we process training examples into clusters to form multiple shape templates, that *compete* in our optimization scheme for the best segmentation. We evaluate the method on a set of synthetic examples and LA DE-MRI images with hand segmentations as the ground truth.

2 Methods

2.1 A Bayesian Formulation of Graph-Cut Segmentation

We treat the problem of segmentation as a maximum a-posteriori estimation. The proposed work differs from many previous Bayesian methods in two important aspects. First, we formulate the segmentation as estimation problem on a graph structure, rather than the image directly. Secondly, we obtain a global optimum to this problem by means of a graph-cut algorithm. The data for this formulation is the image data sampled at locations that are associated with the model. The prior is expressed as a Markov random field (MRF) on the location of the cut in the graph which is related to the formulation introduced by Ishikawa [11]. The graph, which forms a 3D mesh, must approximately adhere the shape to be segmented. It introduces a topological structure on the problem over which the Markov property is introduced.

We begin with a description of the graph structure and associated notation. The graph G is a proper-ordered graph with a set of columns, a neighborhood structure on those columns, and a consistent topological structure as one moves

up and down the columns. We define the *base graph* $G^0 = (V^0, E^0)$, as a set of vertices $\{v_i^0 \in V^0\}$, and edges, $\{e_{i,j}^0 = (v_i^0, v_j^0) \in E^0\}$. For a proper-ordered graph, the vertices are arranged logically as a collection of (conceptually) parallel columns that have the same number of vertices. The entire graph G consists of an ordered set of copies of the base graph, and each vertex can be referenced by its column i and the position within that column l, e.g. v_i^l. The collection $G^l = (V^l, E^l)$ of vertices and edges at the same position l across all columns is called a *layer*. For ease of notation, an edge or vertex without a superscript, v_i or e_{ij}, is considered with respect to the base layer, which defines the topology of all columns. We let N be the number of columns and L be the number of vertices in each column (number of layers). The neighboring columns of the ith column are denoted as the set $\mathcal{N}_i$.

Above is the topological structure of the graph; here we describe its geometry. Each node in the graph has an associated position in the 3D volume/image, which we denote as $x_i^l = x(v_i^l) \in \mathbb{R}^3$. Associated with each x_i^l, there is a *set* of image coordinates, which form an image patch for that vertex, which we call $\mathcal{P}_i^l$. Associated with each patch is a probabilistic model of the intensity patterns one would find in the image at those locations, which is like the formulation of [5].

We now model a set of image measurements associated with a segmentation on the graph. We introduce a probabilistic model with respect to a single segmentation and extend that to coupled surfaces subsequently. We define the surface segmentation as a subset of nodes in the graph $\mathcal{S} \subset V$. Because we restrict the optimal cut to have only one vertex per column, we can parametrize the cut with respect to the base mesh, thus $\mathcal{S}$ can be represented as the function $S : V^0 \mapsto [0, \dots L-1]$. Furthermore, $S(i)$, combined with the topology introduced by the base mesh and the 3D coordinates of the vertices describes a surface in 3D. Thus, we are describing a *surface estimation* problem.

For any given vertex in the graph, v_i^l, we can sample the image I as prescribed by the patch $\mathcal{P}_i^l$. We call the set of image patches for all vertices in the graph as I^V and the set of patches associated with segmentation to be I^S. For a particular segmentation, there is an associated patch I_i^S for each column i.

Now we introduce the probabilistic model, the posterior probability of a segmentation conditioned on image data as follows. Using Bayes rule and considering only terms in the optimization we have:

$$P(\mathcal{S}|I^V) \propto P(I^V|\mathcal{S})P(\mathcal{S}) \tag{1}$$

Next we introduce specific models. For the image intensity model we assume independence of image patches and use an isotropic Gaussian, with a mean for each column that is learned from a set of training examples. That is,

$$P(I^V|\mathcal{S}) \propto \Pi_{i=1}^N P(I_i^S) = \Pi_{i=1}^N \exp\left(-\frac{1}{2\sigma^2} \parallel I_i^S - \mu_i \parallel^2\right) \tag{2}$$

where μ_i is an average patch template learned for surface with physical locations of column i in training examples, and σ is a standard-deviation parameter associated with this data.

For the surface prior, we use a MRF on the function $S(i)$. Let $\mathcal{C} \subset V^0 \times V^0$ be the set of cliques in the base graph, defined by the neighborhood structure, and $C(S(j), S(k))$ is the pairwise clique potential. We use a Gibbs potential on these cliques for the MRF prior, which gives:

$$P(\mathcal{S}) = \exp\Big(- \sum_{(v_j, v_k) \in \mathcal{C}} C(S(j), S(k)) \Big), \tag{3}$$

where the clique potential $C(\cdot)$ typically takes the form $f(|S(j) - S(k)|)$. Here f is monotonic and convex (for optimization to be feasible). In this paper, we use $f(d) = \alpha d^{1+\gamma}; \gamma > 0$.

We minimize the negative log posterior to get the optimal segmentation as:

$$\mathrm{argmin}_{\mathcal{S}} \Big[\lambda \sum_{\mathrm{i=1}}^{\mathrm{N}} \parallel \mathrm{I_i^S} - \mu_{\mathrm{i}} \parallel^2 + \sum_{(\mathrm{v_j}, \mathrm{v_k}) \in \mathcal{C}} \mathrm{C(S(j), S(k))} \Big], \text{ where } \lambda = 1/(2\sigma^2). \tag{4}$$

Segmentation of LA wall requires extraction of epicardial and endocardial surfaces. So, we extend the model to two surfaces/segmentations, $\mathcal{S}^1, \mathcal{S}^2$:

$$P(\mathcal{S}^1, \mathcal{S}^2 | I^V) \propto P(I^V | \mathcal{S}^1, \mathcal{S}^2) P(\mathcal{S}^1, \mathcal{S}^2) \tag{5}$$

We use the same independence assumption with different mean patches for the different surfaces. As we use the MRF for intra-surface smoothness, we propose an inter-surface probability to model interactions between surfaces.

$$\begin{aligned} P(\mathcal{S}^1, \mathcal{S}^2) = \exp\Big(- \sum_{(v_j, v_k) \in \mathcal{C}} C(S_j^1, S_k^1) \Big) \exp\Big(- \sum_{(v_j, v_k) \in \mathcal{C}} C(S_j^2, S_k^2) \Big) \\ \exp\Big(- \sum_{j=1}^{N} g(S_j^1 - S_j^2 - \Delta_j) \Big), \end{aligned} \tag{6}$$

where Δ_j is the ideal inter-surface distance, which may vary with column and learned from training examples, and $g(S_j^1 - S_j^2 - \Delta_j)$ must meet the same conditions of $f()$ in the clique penalty, but must also enforce $S_j^1 < S_j^2$. For this work we use

$$g(d) = \begin{cases} \alpha' d^{1+\gamma'} & d > -\Delta_j \\ \infty & d \leq -\Delta_j \end{cases} \tag{7}$$

The optimization problem for coupled surfaces is therefore:

$$\begin{aligned} &\mathrm{argmin}_{\mathcal{S}^1, \mathcal{S}^2} \Big[\lambda \sum_{i=1}^{N} \Big(\parallel I_i^{S^1} - \mu_i^1 \parallel^2 + \parallel I_i^{S^2} - \mu_i^2 \parallel^2 \Big) \\ &+ \sum_{(v_j, v_k) \in \mathcal{C}} \big(C(S^1(j), S^1(k)) + C(S^2(j), S^2(k)) \big) + \sum_{j=1}^{N} g(S_j^1 - S_j^2 - \Delta_j) \Big] \end{aligned} \tag{8}$$

2.2 Graph Cut Formulation

From the objective functions in the previous section, we now construct a revised graph and define an optimal graph cut that is equivalent to the above optimization. The construction of the derived graph follows, generally, the method proposed by [7] for converting this optimization into an *s-t* cut. Wu *et al.* [7] detail general strategies for solving *surface-net* problems of the type described by Eq. 8. They describe both the Vnet problem, which imposes hard constraints on inter-column behavior and the VCEnet problem, which allows for soft penalties. Previous work including [9, 10] shows the use of the Vnet solution for image segmentation. The Bayesian formulation in the previous section leads to a VCEnet problem, which we also extend to coupled surfaces.

We now briefly review the conversion to the graph-cut problem. The weights on vertices and edges on the extended graph are denoted by $w(v)$ and $c(e)$, respectively. Every vertex in the base layer is connected by a directed edge with a cost $+\infty$ to every other base vertex in its adjacent (neighboring) columns. This makes the base layer strongly connected. For each vertex in layer $l \in [1, L-1]$, a weight of $w_i^l = c_i^l - c_i^{l-1}$ is assigned. A directed edge $e_{i,j}^{l,l-1}$ with a cost $+\infty$ is let from that vertex to the one below it.

The MRF property is incorporated as follows. For every pair of adjacent columns in G, a sequence of directed edges, $e_i^{l,l-d}, d = \{l, ..., 0\}$ go from a vertex v_i^l in i-column to vertices v_j^{l-d} for all $j \in \mathcal{N}_i$, as shown in Figure 2(a). For notational convenience we first define an intermediate function to edge weights

$$q(e_{i,j}^{l,l-d}) = f(d), \quad d = 0, \ldots, l, \tag{9}$$

where $f(d)$ is the penalty, which derives from the clique potential, on the difference in the "height" of adjacent cuts. The weights on these edges are defined through a finite-difference scheme for second derivatives (along columns) of q:

$$w(e_{ij}^{l,0}) = q(e_{ij}^{l,1}) - q(e_{ij}^{l,0}) \tag{10}$$

$$w(e_{ij}^{l,m}) = q(e_{ij}^{l,m+1}) + q(e_{ij}^{l,m+1}) - 2q(e_{ij}^{l,m}), \quad m = 1, \ldots l-1. \tag{11}$$

For the penalty on *inter-surface* distance, we extend the method of [8] to the VCEnet construction We construct two identical disjoint subgraphs, using the procedure above, one for each surface. In addition, a set of directed arcs are added between a pair of subgraphs such that the consistency is maintained between a pair of mutually interacting surfaces. To achieve this interaction, we include a set of arcs between corresponding columns of two subgraphs which are penalized by soft constraints. The formulation resembles the one above; however, all edges are between corresponding columns in the two subgraphs. For ease of notation, all references to vertices associated with the second/inner surface will have a hat (i.e., $\hat{\cdot}$). So, v_i^l and $\hat{v}_i^l$ are corresponding vertices on the two subgraphs. We denote edges between the two surface graphs with a $\tilde{\cdot}$.

Part of our design for this segmentation problem is that one surface should always lie inside the other surface (or "below", if we imagine all columns standing

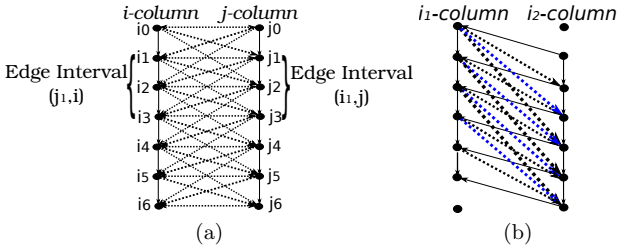


Fig. 2. (a) Inter-column arcs. (b) Inter-surface arcs. Blue arrows from column i_1 to i_2 represent arcs subjected to ideal inter-surface distance, Δ_i

vertically). To achieve this, we include a directed edge between graphs, $\tilde{e}_{i,i}^{l,l+1} = (\hat{v}_i^l, v_i^{l+1})$ with weight $w(e_{i,i}^{l,l+1}) = +\infty$. Similarly, we construct a set of weighted edges that capture the second derivative of the inter-surface penalty when the inner/outer constraint is met as shown in Figure 2b.

$$\tilde{e}_{i,i}^{l,l+\Delta_i+d} = (v_i^l, \hat{v}_i^{l+\Delta_i+d}), \quad d = -\Delta_i + 1, \ldots \tag{12}$$

$$\text{and } w(\tilde{e}_{i,i}^{l,l+\Delta_i+d}) = g(d+1) - 2g(d) + g(d-1). \tag{13}$$

Subsequently, we obtain optimal segmentation of coupled surfaces by finding a minimum s-excess set in the derived graph, as described in Wu *et al.* [7]. This minimum s-excess set is computed by applying a minimum s-t cut in the transformed graph, G_{st}.

2.3 Building a Valid 3D Mesh

In the previous section, we described topology of the underlying graph based on a triangle structure per each layer. Here we describe the assignment of 3D positions to mesh vertices and triangulation of each layer so that these layers form a nested set of watertight meshes in 3D. This complete collection including a set of vertices, their 3D positions, and the prismatic topology of the nested meshes form a proper-ordered (PO) mesh.

For constructing the PO-mesh, we use an extension of the dynamic-particle-system method proposed by Meyer *et al.* [12]. This method computes thin-layers of triangular prisms that conform to shapes. A mesh is built using a template shape (described in the next section), which approximates the LA that we intend to segment. This template shape is represented as the zero level-set of a signed distance transform in the volume. So the following paragraph describes how to generate layers of high-quality meshes on top of this template.

The meshing strategy uses a cluster of points called particles. These particles are distributed on an implicit surface by interactively minimizing a potential function. The potential function based on pairwise distances defines a repulsive interaction between particles as, $U_{i,j}^{l,l} = \Phi(|\boldsymbol{x}_i^l - \boldsymbol{x}_j^l|)$. We denote the sum of this

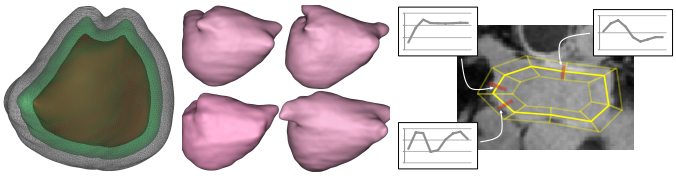
collection of repulsive potentials within each layer as $\mathcal{R}$. These particle systems have been shown to form consistent, nearly regular packings on complex surface [12]. Once points have been distributed on an implicit surface (with sufficient density), a Delaunay tetrahedralization scheme can be used to build a water-tight triangle mesh of the surface [13].

To build a nested set of surface meshes, we require a collection of offset surfaces, both inside and out, that not only inherit the topology of the base surface, but also represent valid, watertight 3D triangle meshes. This is crucial, because the cuts, which pass through vertices from different layers, must also form watertight triangle meshes. Thus, it results to bend the columns in order to avoid tangling of columns/triangles as the layers extend outward from the mean shape. For this, we introduce a collection of particle systems, one for each layer in the graph/mesh, and we couple these particles by an attractive force (Hooks law) between layers. Thus, there is an additional set of potentials of the form $U_{i,i}^{l,l+1} = |\boldsymbol{x}_i^l - \boldsymbol{x}_i^{l+1}|^2$, and we denote the sum of the attractive forces of neighboring particles between layers as $\mathcal{A}$.

To optimize an ensemble of particle systems for L layers, we perform gradient descent, using asynchronous updates, as in [12], on the total potential $\mathcal{R} + \beta\mathcal{A}$. Figure 3a illustrates a nested 3-layered mesh for one of the LA templates. The parameter β controls the relationship between attraction across layers and repulsion within layers and is tuned to prevent tangling. For this paper, we have used $\beta = 10$. The optimization requires an initial collection of particles. So, we place a particle at each point where the adjacent voxels have values on either side of the level set. This gives an average density of approximately one particle per unit surface area (in voxel units). The physical distance between layers must be inversely proportional to the particle density within layers. This is a compromise between the tangling that results from large offsets and the extra computation associated with many thin layers. Since a good mesh constraints the topology and the set of possible segmentations, we try different meshes based on the assumption that all good segmentations can be represented as spatially varying offsets of a mean. This corresponds to around 14,000 particles per each mesh layer for heart images and 2000 particles for simulated images. We have used a total of 30 layers, spaced at 0.5 pixels each, which gives each template a capture range of approximately 15 pixels.

2.4 Learning Template Meshes and Feature Detectors

Here we describe the construction of template shapes and the mechanism for computing costs on nodes from input images. The shapes of LA in the context of AF are highly variable. To address this, we rely on a *training set* of presegmented images. For this paper, the training set consisted of 32 segmented DE-MRI images of the LA. The work in this paper represents a prototype, and we anticipate a production-scale system that relies on hundreds of training images. These training images enable two things. First, training images give us a way of constructing a collection of PO-graphs, so that new images can be segmented as cuts through one of these graphs. Second, training images give us examples of



(a) (b) (c)

Fig. 3. (a) An example of several layers of PO-meshes for the LA.(b) Examples of average shapes, derived from k-means clustering on distance transforms of training images, around which the PO-meshes are constructed. (c) A mock up of a simplified PO-mesh in 2D with examples of feature detectors learned from the training data—actually P0-meshes for the LA have over 400,000 vertices.

patch profiles for the features that define epi- and endocardial surfaces, which leads to the costs at each node in the PO-graph.

We begin by clustering the examples based on their shapes. For this, we compute distance transforms of each endocardial surface. Training images are aligned via translation to ensure common center of mass for the blood pool (region bounded by the endocardium). This demands careful manual initialization of a template which will be handled in our future work by inducing other transformations. We then compute clusters using k-means using mean-squared distance metric between volumes. Based on the cluster residual curve, 5 clusters are chosen. However, one of the clusters has been removed from the test, because it contained only two (high distorted) examples. Surface meshes associated with the distance-transform means of these four clusters are shown in Figure 3b.

The cost associated with each vertex reflects the degree to which that vertex is a good candidate for a boundary, which will be found via a graph cut. At each vertex, the training data is used to derive a patch profile along a line segment, or *stick* perpendicular to the surface. We sample the stick at a spacing of one voxel. In our case, a patch size of 11 is considered along the normal direction of the surface. The intensity along each stick on each vertex of each template is computed by a weighted average of intensities of sticks for each feature point in each training image. Thus, for a particular vertex in a particular cluster, the intensities along a stick would correspond to an average of several hundreds of neighboring sticks from different images (that share the same blood-pool center). Thus the average stick at a vertex would be an isotropic Gaussian weighted average of all the nearby sticks (within the cluster) with standard deviation of 2 pixels. Figure 3c shows a diagram of the stick configuration and several stick intensity profiles for parts of a particular template.

3 Experiments and Results

For validation, we apply Bayesian framework based graph cut method on 100 simulated images of size $64 \times 64 \times 96$ voxels, and 30 DE-MRI images of the left

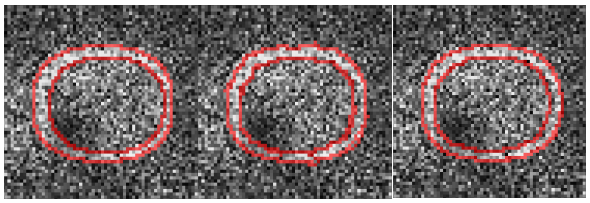


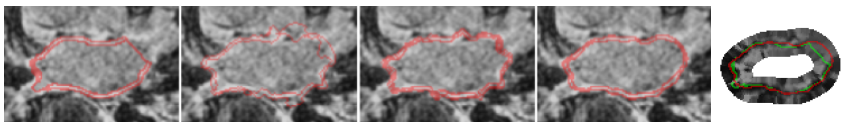
Fig. 4. Segmentation boundaries for outer and inner surfaces on synthetic data corresponding to ground truth, graph cuts with hard constraints and the proposed algorithm

atrium of size $400\times400\times107$ voxels. In all of our experiments, 30 mesh layers were generated, spaced at 0.5 voxels each, which gives each template a capture range of approximately 15 voxels. The scaling and exponential parameters, α and γ, for the convex function f of the graph's smoothness penalties are fixed as 300 and 2, respectively. The corresponding values for the function g of the graph's inter-surface penalties are set to 2 and 2, respectively. The values of these parameters reflect the complexity of surfaces and the inter-surface coupling between them.

To segment a given test image, we depend on the user input to position the template. The algorithm is robust to this position, as long as the nested mesh, corresponding to the template, does not lie outside or inside the desired surface (e.g. ±5 voxels). We sample the input image along all of the sticks at all nodes. Then, we compute a posterior probability on each test stick with the corresponding template stick. This results in the assignment of costs, weights, edge capacities, and then an optimal cut. Likewise, we employ all of the learned templates to the input image, choosing the segmentation that produces the best average probability with the local intensity models for the optimal cut. A pair of optimal mesh surfaces are then recovered from the computed minimum s-t cut. Based on the extracted topological mesh structure, defined by the cut, it is scan converted to reproduce segmented volume(s).

In case of the simulated data, 30 training datasets and 100 test sets were considered for analysis. All these images include two oblong non-crossing surfaces with the inner surface translated randomly (Gaussian distribution) in 3D to mimic variations in heart-wall thickness; each image was corrupted with Rician noise ($\sigma = 30$ for the underlying Gaussian model) and a smoothly-varying bias field. Figure 4 illustrates the effectiveness of the proposed method in extracting smoother boundaries for outer and inner surfaces as compared to hard penalties.

We evaluated the segmentation accuracy for LA based on leave-one out strategy for a test dataset, against templates from the training data. We compared the segmented boundaries of epicardial and endocardial surfaces using our method to that of hard constraints. Since the geometric constraints and soft penalties in the proposed graph cut formulation are analogous to the energy based formulation in deformable models [4], we compared our results with level set based methods. Figure 5a presents segmentation boundaries for epicardial and endocardial surfaces obtained by the proposed algorithm along with others. The cost



(a) (b)

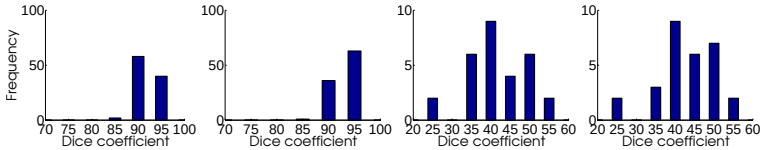
Fig. 5. (a) Surface boundaries of epi and endocardial surfaces corresponding to ground truth, level sets, graph cuts with hard constraints and proposed method. (b) Segmentation result of epicardial surface using the proposed algorithm (red) and ground truth boundary (green) overlaid on corresponding cost function image

function image, derived from a-posterior probability, creates a platform on which graph cuts work. Figure 5b illustrates our segmentation result on cost function image corresponding to the epicardial surface. The rationale behind presenting this result is to show how the algorithm is able to extract smoother and accurate boundaries in some areas of the image where even the costs, which are derived from the sophisticated feature detector, could not be defined properly.

The qualitative comparison between the proposed method with others clearly indicates that our method surmounts other techniques in not only extracting correct surfaces, but also in maintaining smoothness along the surfaces and consistency in between them. The irregularity in the surfaces that we notice due to the hard constraints were greatly eliminated.

To evaluate the segmentation accuracy quantitatively, we used distance metric. The distance metric is based on the aggregate of pairwise distances between corresponding points on the ground truth and our segmentation. For each point on our segmented surface, we measure the distance to the nearest point on the ground truth; and vice-versa. For a perfect delineation of the boundary, all these distances would be zero. In the case of simulated examples, this distance metric which was computed over all the images came out to be 0.1879 voxels for the outer surface and 0.2639 voxels for the inner surface. For LA data, we obtained this metric value of 2.5068 voxels for epicardium and 2.6321 voxels for the endocardium. This indicates that the segmentations acquired by the proposed method lie very close to the ground truth.

For quantitative comparison, we studied Dice measures on heart wall using soft against hard constraints. The Dice metric provides the percent overlap between the ground truth and segmented regions. Figure 6 shows the histogram of Dice measures. In both simulated as well as LA cases, the metric values by inducing soft penalties on geometric constraints overpowered hard penalties. For synthetic data, the Dice values indicate excellent matches. However, in the case of myocardium, the dice values are little lower due to its varying thinness (2-6 mm) and undefined ground truth. The ground truth is a single hand segmentation from an expert. Therefore, much of the observed error is near the veins, which are subject to inter-rater variability, as the cutoff between atrium and vessel is not well defined. Also the ground truths for the wall do not form a complete boundary around the blood pool (even ignoring the vessels). Furthermore, we



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

Fig. 6. Histogram of dice coefficients for the (a) middle region, graph cuts with hard constraints, (b) middle region, soft constraints, (c) heart wall, graph cuts with hard constraints, (d) heart wall, soft constraints

expect the improvement in results by increasing the number of training images so that more templates are formed in order to better match a given input image.

References

1. McGann, C.J., Kholmovski, E.G., Oakes, R.S., Blauer, J.J., Daccarett, M., Segerson, N., Airey, K.J., Akoum, N., Fish, E., Badger, T.J., DiBella, E.V., Parker, D., MacLeod, R.S., Marrouche, N.F.: New magnetic resonance imaging-based method for defining the extent of left atrial wall injury after the ablation of atrial fibrillation. J American College of Cardiology **52**(15) (2008) 1263–1271
2. John, M., Rahn, N.: Automatic left atrium segmentation by cutting the blood pool at narrowings. Med Image Comput Comput Assist Interv **8** (2005) 798–805
3. Karim, R., Mohiaddin, R., Rueckert, D.: Left atrium segmentation for atrial fibrillation ablation. In: SPIE Conference Series. Volume 6918. (2008)
4. Malladi, R., Sethian, J., Vemuri, B.: Shape modeling with front propagation: a level set approach. PAMI, IEEE Transactions on **17**(2) (feb 1995) 158 –175
5. Cootes, T., Taylor, C., Cooper, D., Graham, J.: Active shape models-their training and application. Computer Vision and Image Understanding **61**(1) (1995) 38 – 59
6. Gao, Y., Gholami, B., MacLeod, R., Blauer, J., Haddad, W., Tannenbaum, A.: Segmentation of the endocardial wall of the left atrium using local region-based active contours and statistical shape learning. In: SPIE. Volume 7623. (2010)
7. Wu, X., Chen, D.Z.: Optimal net surface problems with applications. In: Proc. Int. Colloquium on Automata, Languages and Programming. (2002) 1029–1042
8. Li, K., Wu, X., Chen, D., Sonka, M.: Optimal surface segmentation in volumetric images-a graph-theoretic approach. PAMI **28**(1) (2006) 119 –134
9. Dou, X., Wu, X., Wahle, A., Sonka, M.: Globally optimal surface segmentation using regional properties of segmented objects. In: CVPR. IEEE. (2008) 1–8
10. Song, Q., Wu, X., Liu, Y., Smith, M., Buatti, J., Sonka, M.: Optimal graph search segmentation using arc-weighted graph for simultaneous surface detection of bladder and prostate. MICCAI **12** (2009) 827–835
11. Ishikawa, H.: Exact optimization for markov random fields with convex priors. PAMI **25**(10) (oct. 2003) 1333 – 1336
12. Meyer, M., Kirby, R., Whitaker, R.: Topology, accuracy, and quality of isosurface meshes using dynamic particles. IEEE TVCG. **12**(5) (2007) 1704–11
13. Amenta, N., Bern, M., Eppstein, D.: The crust and the beta-skeleton: Combinatorial curve reconstruction. Graphic Models Image Proc. **60**(2) (1998) 125–135

CHAPTER 4

A BAYESIAN FORMULATION OF GRAPH-CUT SURFACE ESTIMATION WITH GLOBAL SHAPE PRIORS

This chapter is reprinted, with permission, from

Gopalkrishna Veni, Shireen Y. Elhabian, and Ross T. Whitaker, **A Bayesian Formulation of Graph-Cut Surface Estimation with Global Shape Priors.** In Proceedings of the 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI), April 2015.

A BAYESIAN FORMULATION OF GRAPH-CUT SURFACE ESTIMATION WITH GLOBAL SHAPE PRIORS

Gopalkrishna Veni[1], *Shireen Y. Elhabian*[1,2], *Ross T. Whitaker*[1]

[1] Scientific Computing and Imaging (SCI) Institute, University of Utah, USA
[2] Faculty of Computers and Information, Cairo University, Egypt

ABSTRACT

In this paper, we propose a formulation of graph-cut segmentation that relies on a generative image model by incorporating both local and global shape priors. With surface estimation, rather than pixel classification, we cast the segmentation problem as a maximum *a posteriori* estimation from the image intensities via a *cut* through a multi-layer three-dimensional mesh model that preserves the topology of the shape class of interest. Methods that rely on local optimization techniques and/or local shape penalties, e.g., smoothness, have been proven to be ineffective to address challenging segmentation problems, such as noisy/ill-defined boundaries and irregular shapes. On the other hand, our method relies on graph cuts as well as a new formulation to estimate shape parameters in a closed form that provides a global updates-based optimization strategy. We demonstrate our formulation on synthetic datasets as well as the left atrial wall segmentation from late-gadolinium enhancement MRI, which is useful in atrial fibrillation to identify fibrosis, but presents local contrast and noise within the wall.

Index Terms— Bayesian Segmentation, Atrial Fibrillation, Minimum *s-t* Cut, Mesh Generation, Geometric Graph.

1. INTRODUCTION

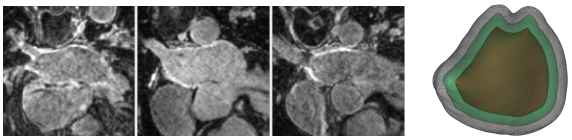
Accurate segmentation of medical images is a ubiquitous demand in clinical studies for diagnosis and analysis. Nonetheless, specific classes of organs/tissues present a challenging combination of features including heterogeneous pixel intensities, noisy diffuse boundary features and irregular shapes, which often arise during the soft tissue imaging for a variety of applications in neuroscience/neurology, cardiology and oncology, to name a few. One such challenge is the segmentation of the left atrium (LA) wall from late-gadolinium enhancement cardiac MRI (LGE-MRI) in patients suffering from atrial fibrillation (AF). LA wall segmentation is crucial to characterize fibrosis (tissue remodeling) in patients prior to radiofrequency ablation therapy (pre-ablation) and for follow-up studies of LA wall scarring in patients after radiofrequency ablation (post-ablation) [1]. Examples of such images are shown in Figure 1a.

Weak boundary segmentation problems are often solvable only if one has some *a priori* notion of shapes in question; motivating a Bayesian framework for segmentation being cast as a surface estimation problem. In the extreme case, one might restrict the solution to some low-dimensional set of shapes defined through training data, as is done with *active shape models* (ASM) [2], active appearance models (AAM) [3] and many variations derived from these approaches. Nonetheless, segmentations must follow correct boundaries rather closely, and yet coarsely similar to a reasonable summary of the training dataset. Furthermore, many surface-based approaches rely on some type of incremental fitting algorithm, such as gradient descent, which is prone to get trapped in local minima, especially if one allows for a very expressive prior to accurately capture details. Based on learning the topology from the shape prior paradigm, our work is built on the ASM approach by projecting onto a low-dimensional, statistical description from a set of training shapes/images. However, the proposed method also accommodates shapes that are not represented in this low-dimensional description. Thus, the method offers optimization schemes that do not rely on incremental updates to shapes or parameters, as is the case of a gradient descent scheme.

Graph cut methods, on the other hand, have recently been extended to incorporate shape priors. For example, [4, 5] rely on the variations of *pixel labeling* formulation for the graph-cut segmentation of Boykov and Jolly [6]. The idea is to identify pixels based on the specified intensity distributions, and then project those pixel classifications onto shape priors, so that the priors affect pixel classification probabilities. The optimization schemes are quite effective, but still approximate. In [7], Gorelick *et al.* use an approximate model to minimize segmentation energies but their method is designed for single surface, which does not fit our application. To address multi-surface segmentation, [8] and [9] rely on intensity or statistical differences across boundaries. Although surface-separation constraints are enforced, they do not use soft penalties with preferred separation distance, and thus their approaches suffer from extracting irregular surfaces. Moreover, no parametrized shape priors are embedded in their graph-structure. Such priors are necessary for weak boundary segmentation problems. Further, [9] uses an approximation graph-cuts scheme that converges to a locally optimal solution while reconstructing the entire graph during each iteration.

The proposed framework builds on the work related to the optimal net surfaces [10] and their applications to medical image segmentation [11, 12]. The segmentation problem is specified as an optimal net surface through a proper ordered (PO) graph, where that net is defined by a minimum *s-excess* set and reduced to a minimum *s-t* cut, for which a global solution is efficient. Veni *et al.* [13] used this surface-net approach to solve for a similar problem. However, they do not incorporate any shape information which, we believe, is essential while dealing with low quality images accompanied with ill-defined boundaries. Song et al. [12] introduced shape priors into such a surface estimation scheme, but the priors were strictly local, implemented as modifications to smoothness parameters in the optimal cut formulation. Here we propose a surface-based segmentation algorithm that relies on a fully generative model that incorporates a global shape prior, controlled smooth deviations from that prior and a combination of a linear system and graph-cut approach

The authors would like to acknowledge Josh Cates, the CARMA Center, CIBC by NIH Grant P41 GM103545-14 and NAMIC through NIH Grant U54 EB005149, for providing Utah fibrosis data.



(a) (b)

Fig. 1. (a) Slices of left atrium DE-MRI images showing the challenges in segmentation. (b) An example of several layers of a PO-mesh for the LA

to maximize *a posteriori* optimization. Thus, the resulting optimal segmentation not only is based on the prior but also accommodates irregularities in the test data in order to best match image intensities. Moreover, we estimate the shape parameters, in closed form, for segmentation as a part of the optimization process.

2. METHODOLOGY

2.1. Bayesian Formulation

We begin by defining a segmentation $\mathcal{S}$ as a surface whose vertices are defined on a 3D grid and are connected by edges that form a triangular mesh. Associated with each vertex is an intensity profile (e.g., a 3D patch or a 1D array of intensities) oriented along the approximate normal to the surface, as in [13]. Thus, a particular surface induces a probability on a set of image intensities near each point on the surface. The segmentation takes place on a set of columns, which are connected via a triangular mesh structure as described in [13]. Each column has a set of 3D vertices along its length. Thus, the columns span the 3D space in which the segmentation takes place, as seen in Figure 1b. A particular vertex at the jth position along ith column has a 3D position x_{ij}. The graph has M columns and N vertices along each column. In principle, this column set is a discrete approximation of an underlying continuous parametrization of a subset of $\Re^3$ that relies on surface integrals within which the surface estimation takes place. This *parametrization* also defines the topology of the resulting segmentation. For simplicity, we present only the discrete formulation. A segmentation $\mathcal{S}$ is the set of nodes $\{s(i) \in [1, N]\}\ \forall i \in [1, M]$ representing the position of the surface.

We use a Bayesian formulation to express the posterior probability of a segmentation $\mathcal{S}$ and underlying shape parameters β as a function of the full image I.

$$P(\mathcal{S}, \beta|I) \propto P(I|\mathcal{S}, \beta)P(\mathcal{S}, \beta) = P(I|\mathcal{S}, \beta)P(\mathcal{S}|\beta)P(\beta). \quad (1)$$

$P(I|\mathcal{S}, \beta)$ is a generative model of intensities in the vicinity of a segmentation. We use a Gaussian distribution centered around a mean intensity profile, μ_i, at each point on the associated surface, where μ_i for column i is estimated from the training data. We consider the set of intensities[1] p_{ij} at a vertex x_{ij} to be statistically independent from one vertex to another, which results in a product of probabilities and a sum for the log of the conditional probability.

$$-\log P(I|\mathcal{S}, \beta) = \frac{1}{2\sigma_I^2} \sum_i ||p_{i,s(i)} - \mu_i||^2 \quad (2)$$

$P(\mathcal{S}|\beta)$ describes the probability of a surface estimate as a function of the shape parameters. We represent the surface as a base shape $\mathcal{S}_\beta$ and an offset $\mathcal{S}_o$ where $\mathcal{S} = \mathcal{S}_\beta + \mathcal{S}_o$. For a given set of shape parameters β, the surface probability depends only on the offset, which needs to be small (the shape should be similar to the base shape) and smooth (nearby columns should have similar offsets). Therefore the log-probability can be written as:

$$-\log P(\mathcal{S}|\beta) = \frac{1}{2\sigma_o^2} \sum_i s_o^2(i) + \frac{1}{2\sigma_\mathcal{C}^2} \sum_{(i,j)\in\mathcal{C}} (s_o(i) - s_o(j))^2 . \quad (3)$$

where σ_o and $\sigma_\mathcal{C}$ are standard deviations of surface offsets learned from the training set and the neighborhood structure, defined by $\mathcal{C}$, which is the clique in the base graph defined over the base shape.

For the shape parameters, we use PCA on a set of distance transforms (DT) of the training data [14], and choose a low-dimensional subspace that maintains 97% of data variability to represent the prior. The Gaussian model for the shape space gives:

$$-\log P(\beta) = \frac{1}{2}\beta^T \Lambda^{-1} \beta, \quad (4)$$

where β is a vector of shape parameters and Λ is the diagonal matrix of eigenvalues from the PCA. The eigenvectors (eigen DT) d_{e_k} from PCA, along with the shape parameters β, are used to construct the corresponding DT d_β.

$$d_\beta(x) = d_\mu(x) + \sum_k \beta_k d_{e_k}(x), \quad (5)$$

where $x \in \Re^3$ and d_μ denotes the mean DT. We use d_β to express the conditional surface probabilities in (3). The DT for a particular β plays two roles. For a given β, we can build the surface $\mathcal{S}_\beta$ by computing the intersection of each column i with the zero crossing of d_β, to give a vertex and 3D positions for $s_\beta(i)$. The DT also gives a convenient way to compute the penalty on the magnitude of the offset from $\mathcal{S}_\beta$, which becomes $\sum_i d_\beta^2(s_i)$, where $s_i = x_{s(i)}$, the sum of square distances of the cuts along each column i from the β-shape.

Thus, we have a full, shape-based generative model of an image. A point β in the low-dimensional shape space, derived from PCA on training data, chosen from a multivariate normal distribution, gives rise to a DT, whose zero level set is the β-shape $\mathcal{S}_\beta$. An offset from this β-shape is chosen from a distribution defined by a Markov random field (MRF) model that favors smaller yet smoother offsets. Notice, this offset model extends the learned probability distribution from the low-dimensional PCA subspace into the entire shape space in a way that favors smooth shapes that are very near, but do not exactly resemble, the training set. Finally, each point on that surface gives rise to an array (patch or vector) of image intensities that are drawn from (identically, independently) a Gaussian distribution, with a mean and covariance that are learned from the training data along each column.

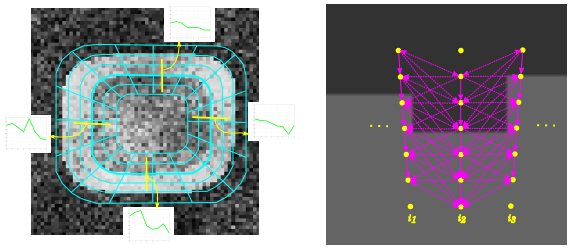
It is worth noting that the LA wall segmentation problem consists of a pair of nested surfaces (epicardium and endocardium). Thus, the above formulation is applied such that the distance transforms are modeled jointly. Similar to [13], we introduce a penalty on the inter-surface distances (expressed in the same probabilistic framework, learned from the data and tied to the β-shape-pair) such that the surfaces are estimated optimally in a joint formulation.

2.2. Optimization

We present a global optimization solution by following an alternating scheme in which we iterate between two fast *global* optimization phases on two sets of parameters.

In the first phase, we optimize for $\mathcal{S}$ given β where we use graph cuts to solve the VCE-weight net surface problem using [10]. We use the MRF formulation [13], which is related to the derivation introduced in [15]. The basic strategy is to express the above penalties

[1]This set of intensities could be organized in any way around a node on the mesh, but for this work we use a 1D array, a kind of *stick* configuration, which is shown in Figure 2a.



(a) (b)

Fig. 2. (a) 2D illustration of a simplified PO-mesh and *sticks* (yellow) at different mesh points, (b) Arc construction between adjacent columns with respect to $\mathcal{S}_\beta$

as a set of weights on vertices and edges of a PO-graph and reduce the surface estimation problem to a *s-excess* cut on the graph. The *s-excess* cut produces an optimal low-order polynomial time solution– and in practice runs very fast, even on large graphs. The details of the graph construction follow from [10] with some modifications, and therefore we give a high-level view of the algorithm here.

Weights/costs on the graph vertices arise from two sources: (1) a matching term in the image probability, by sampling image patches around each vertex and comparing them against the learned mean, and (2) sampling the squared DT of the β-shape at each vertex position. These terms are weighted and added to generate vertex costs. Edges are inserted along graph-columns and also between the nodes of adjacent columns. The inter-column edges model the pairwise clique potential $C(s_i, s_j)$, associated with the offsets, rather than the absolute positions of the surface, as in previous work. For the optimization to be feasible, we use $C(d) = \alpha d^2$, where $d = (s_i - s_j)$. The β-shape $\mathcal{S}_\beta$ gives a position on each column and $\mathcal{S}_o$ is computed with respect to $\mathcal{S}_\beta$. Thus, the inter-column arcs are adjusted, relative to the $\mathcal{S}_\beta$. Figure 2b summarizes the graph construction of the above configuration of intra- and inter-column arcs. The resulting graph on which we estimate the cut becomes a numerical construct and does not define or heavily influence the shapes of the objects to be segmented. If a mesh is sufficiently fine grained, it only serves to restrict the topology of the surface. The $\mathcal{S}_\beta$ defines an arbitrary surface through the mesh and the offsets allow smooth perturbations on that shape. For a given β, the surface estimate $\mathcal{S}$, computed using the graph-cut, is the global optimum of the posterior in (1).

In the second phase we optimize for β in a closed form given a segmentation $\mathcal{S}$. Here we take the derivative of the log posterior with respect to β and equate it to zero.

$$\frac{1}{\sigma_o^2}\sum_i d_\beta(s_i)\frac{\partial d_\beta}{\partial \beta}(s_i) + \Lambda^{-1}\bar{\beta} + \frac{1}{\sigma_c^2}\sum_{(i,j)}\left(d_\beta(s_i) - d_\beta(s_j)\right)\left(\frac{\partial d_\beta}{\partial \beta}(s_i) - \frac{\partial d_\beta}{\partial \beta}(s_j)\right) = 0. \quad (6)$$

From (5), d_β is a weighted sum of images. The images $\partial d_\beta / \partial \beta$ are the principle components (eigen DT) from the PCA decomposition of the distance transforms in the training set. By isolating β in (6), we obtain a closed-form algebraic solution on vectors/matrices in the space of shape parameters.

$$\bar{\beta} = -(\sigma_c^2 P^a + \sigma_o^2\sigma_c^2\Lambda^{-1} + 2\sigma_o^2(P^a - P^b))^{-1} \times (\sigma_c^2 q^a + 2\sigma_o^2(q^a - q^b)) \quad (7)$$

$$q_k^a = {d_\mu}^T d_{e_k}, \text{and} \quad q_k^b = {d_\mu}^T \mathcal{C}_\mathcal{N} d_{e_k} \quad (8)$$

$$P_{kl}^a = {d_{e_k}}^T d_{e_l}, \text{and} \quad P_{kl}^b = d_{e_k}\mathcal{C}_\mathcal{N} d_{e_l}. \quad (9)$$

Fig. 3. Left: the original image. Middle: Boykov's result overlaid on the ground-truth segmentation. Right: Our result.

where $\mathcal{C}_\mathcal{N}$ represents the $M \times M$ indicator matrix, with its elements equal to 1 if $j \in \mathcal{N}_i$ (neighborhood of i), else 0. This is the *optimal* β for a given surface segmentation, which finds a compromise between fitting the current segmentation and choosing a likely β based on the Gaussian distribution.

When we consider the algorithm in full, this method differs significantly from others that, in general, either work strictly within the shape space [2] or project intermediate solutions onto that subspace. Here, the optimization strategy offers two advantages. First, it allows intermediate results to deviate from the low-dimensional shape space and to pull the βs along. Second, in each iteration we perform a global optimization of $\mathcal{S}$ and β, thus relying less on local image properties near the current solution. Thus, the algorithm converges with relatively few iterations.

3. EXPERIMENTS AND RESULTS

In order to make a reasonable comparison of our approach with the state-of-art graph-cut algorithms including Boykov's, we considered 2D superquadric examples, corrupted with Rician noise and a smoothly varying bias field. The corresponding energy function to be minimized is given by [6]:

$$E = E_{Data} + E_{Smooth} = \sum_{p\in\mathcal{P}} D_p(x_p) + \sum_{(p,q)\in\mathcal{C}} V_{p,q}(x_p, x_q). \quad (10)$$

The data costs were computed with respect to the mean DT of the training set and the smoothness costs were computed as the normalized cross correlation between the intensity profile at each point and the corresponding model profile that was again derived from the training set. Figure 3 shows the effect of weights on data and smoothness terms, and the corresponding result using our approach. Such incorrect object delineation in Boykov's case is due to the absence of the shape prior, lack of well-defined smoothness structure that is derived from the MRF property and incompetence of allowing only one *cut* per column.

To evaluate our framework, we applied it to 100 simulated images (64×64×96 voxels) and 58 clinical LGE-MRI (29 pre- and 29 post-ablation, 400×400×107) images that were obtained retrospectively from an AF patient image database at the University of Utah's Comprehensive Arrhythmia Research and Management (CARMA) Center. Based on the capture range of two surfaces, 50 mesh layers were used for simulated data and 30 layers were generated for LA examples. We used the dynamic particle system to generate the initial mesh and respective layers (see [13] for details). The mesh does not get rebuilt during segmentation as it has enough layers and resolution to allow shape flexibility. It is the cut through the mesh that is modified to embed shape parameters.

For a given test example, the algorithm relies on the user input to locate its approximate center based on which the nested-mesh is positioned. The algorithm is robust enough as long as the mesh does not lie either outside or inside the desired surface. The input image is sampled along all sticks at all nodes. Next, the posterior probability is computed using (1) which reflects the assignment of

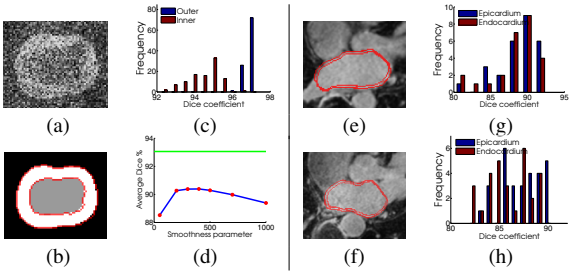


Fig. 4. (a) 2D slice of a perturbed, noisy synthetic example, (b) Segmentation boundaries overlaid on ground-truth (c,g,h) dice histograms of two surfaces on synthetic, pre- and post-ablation datasets (d) Average dice comparison of our method (green) vs VCEnet without shape priors for different α (e) LA wall 1 (f) LA wall 2

costs, weights, edge capacities and then an optimal cut. A pair of optimal mesh surfaces is then recovered from the minimum *s-t* cut. Using this cut, the shape parameters are optimized. This procedure is repeated until (1) converges[2]. Based on the extracted topological mesh structure defined by the final cut, it is scan-converted to reproduce segmented volume(s). The scaling parameter α for the potential function C is fixed to 300 for the synthetic case and 100 for the LA dataset. These values reflect the complexity of surfaces.

For the simulated data, we considered 30 training examples and 100 test cases, which include two oblong mutually interacting surfaces. To mimic the variations of the LA-wall thickness, we incorporated shape changes by perturbing both of the surfaces in each image by means of a randomly (Gaussian distributed with $\mu = 11$, $\sigma^2 = 36$) corrupted DT to generate wiggles on the surfaces. Next, each image was corrupted with Rician noise (with $\sigma = 30$) and a smoothly varying bias field (see, e.g., Figure 4a). Figure 4b shows our segmentation result on the corresponding ground-truth to prove the effectiveness of our approach in retrieving both the boundaries correctly, even in such a challenging environment.

For the quantitative evaluation, we used the Dice metric to analyze the overlapping percentage between our segmentation result and the ground-truth. Figure 4c shows the histogram plot of Dice measures on both surfaces. To prove the effectiveness of embedding global shape priors into the VCE-weight net framework, we compared the average Dice of the middle layer over all images to the one that does not include the prior. As shown in Figure 4d, the average Dice percentage with the shape prior is 93.06%. By contrast, a maximum of only 90.39% is achieved by varying the α value of VCE-weight net without the prior [13]. This analysis illustrates the advantage of including the shape prior in extracting desired boundaries.

Figures 4e and 4f show the segmented boundaries of the endocardial and epicardial surfaces of a LA in an example slice from an LGE-MRI image. Together, these two surfaces define the extent of the LA wall and were compared with manual segmentations from experts at the CARMA Center. Average Dice coefficients in this analysis include 0.89 and 0.87 for the epi- and endocardial boundaries in the pre-ablation images, respectively, and 0.87 and 0.86 for the epi- and endocaridal boundaries in the post-ablation images. The histogram plots of Dice measures on endo- and epicardial surfaces in the pre- and post ablated studies are shown in Figure 4g and 4h respectively. We also computed Hausdorff distances to the ground-truth, which provides more insight into the accuracy of both the *shape* and the *local alignment* of our automatic segmentations. For the pre-ablation surfaces, we found an average Hausdorff distance of $1.89 \pm 0.58mm$ and $1.91 \pm 0.55mm$ for epi and endocardium. For the post-ablation surfaces, average distances were $2.36 \pm 0.56mm$ and $2.45 \pm 0.56mm$ for epi- and endocardium, respectively. These errors are consistent with manual segmentation variability reported by CARMA expert segmenters in LGE-MRI images of similar quality.

[2]Our algorithm converges in $\sim$7 iterations on average.

4. REFERENCES

[1] Marrouche, N.F., Wilber, D., Hindricks, G., Jais, P., Akoum, N., Marchlinski, F., Kholmovski, E., Burgon, N., Hu, N., Mont, L., et al.: Association of atrial tissue fibrosis identified by delayed enhancement mri and atrial fibrillation catheter ablation: The decaaf study. JAMA **311**(5) (2014) 498–506

[2] Cootes, T., Taylor, C., Cooper, D., Graham, J.: Active shape models-their training and application. Computer Vision and Image Understanding **61**(1) (1995) 38 – 59

[3] Cootes, T.F., Edwards, G.J., Taylor, C.J.: Active appearance models. PAMI **23**(6) (Jun 2001) 681 – 685

[4] Cremers, D.: Dynamical statistical shape priors for level set-based tracking. PAMI **28**(8) (aug. 2006) 1262 –1273

[5] Malcolm, J., Rathi, Y., Tannenbaum, A.: Graph cut segmentation with nonlinear shape priors. In: ICIP. Volume 4. (16 2007-oct. 19 2007) IV –365 –IV –368

[6] Boykov, Y., Jolly, M.P.: Interactive graph cuts for optimal boundary amp; region segmentation of objects in n-d images. In: ICCV. Volume 1. (2001) 105 –112 vol.1

[7] Gorelick, L., Schmidt, F.R., Boykov, Y.: Fast trust region for segmentation. In: CVPR, 2013 IEEE Conference on, IEEE (2013) 1714–1721

[8] Delong, A., Boykov, Y.: Globally optimal segmentation of multi-region objects. In: Computer Vision, 2009 IEEE 12th International Conference on, IEEE (2009) 285–292

[9] Schmidt, F.R., Boykov, Y.: Hausdorff distance constraint for multi-surface segmentation. In: Computer Vision–ECCV 2012. Springer (2012) 598–611

[10] Wu, X., Chen, D.Z.: Optimal net surface problems with applications. In: Proc. Int. Colloquium on Automata, Languages and Programming. (2002) 1029–1042

[11] Li, K., Wu, X., Chen, D., Sonka, M.: Globally optimal segmentation of interacting surfaces with geometric constraints. In: CVPR. Volume 1. (2004) I–394 – I–399

[12] Song, Q., Bai, J., Garvin, M., Sonka, M., Buatti, J., Wu, X.: Optimal multiple surface segmentation with shape and context priors. TMI IEEE **32**(2) (Feb 2013) 376–386

[13] Veni, G., Fu, Z., Awate, S.P., Whitaker, R.T.: Bayesian segmentation of atrium wall using globally-optimal graph cuts on 3d meshes. In: IPMI, Springer (2013) 656–667

[14] Leventon, M.E., Grimson, W.E.L., Faugeras, O.: Statistical shape influence in geodesic active contours. In: CVPR. Volume 1., IEEE (2000) 316–323

[15] Ishikawa, H.: Exact optimization for markov random fields with convex priors. PAMI **25**(10) (oct. 2003) 1333 – 1336

CHAPTER 5

SHAPECUT: BAYESIAN SURFACE ESTIMATION USING SHAPE-DRIVEN GRAPH

5.1 Abstract

A variety of medical image segmentation problems present significant technical challenges, including heterogeneous pixel intensities, noisy/ill-defined boundaries, and irregular shapes with high variability. The strategy of estimating optimal segmentations within a statistical framework that combines image data with priors on anatomical structures promises to address some of these technical challenges. However, methods that rely on local optimization techniques and/or local shape penalties (e.g., smoothness) have been proven to be inadequate for many difficult segmentation problems. These challenging segmentation problems can benefit from the inclusion of global shape priors within a maximum-a-posteriori estimation framework, which biases solutions toward an object class of interest. In this paper, we propose a maximum-a-posteriori formulation that relies on a generative image model by incorporating both local and global shape priors. The proposed method relies on graph cuts as well as a new shape parameters estimation that provides a global updates-based optimization strategy. We demonstrate our approach on synthetic datasets as well as on the left atrial wall segmentation from late-gadolinium enhancement magnetic resonance imaging (MRI), which has been shown to be effective for identifying myocardial fibrosis in the diagnosis of atrial fibrillation. Experimental results prove the effectiveness of the proposed approach in terms of the average surface distance between extracted surfaces and the corresponding ground-truth, as well as the clinical efficacy of the method in the identification of fibrosis and scars in the atrial wall. [1]

[1]This chapter is a submitted copy of **Shapecut: Bayesian Surface Estimation Using Shape-driven Graph.** *Gopalkrishna Veni, Shireen Y. Elhabian, and Ross T. Whitaker.* to J. Med. Image Anal., 2016.

5.2 Introduction

Medical image segmentation typically deals with partitioning an image into multiple regions representing anatomical objects of interest. A three-dimensional segmentation can be viewed as either *pixel classification* (defining regions) or a *surface estimation* (boundaries of regions), and in this paper, we deal with the latter approach. Heterogeneous pixel intensities, noisy diffuse boundaries, and irregular shapes are some factors that make it difficult to develop accurate, robust segmentation algorithms for certain classes of images and anatomy. The proposed method is motivated by a particular problem of this type, which is the segmentation of the left atrium (LA) wall, also referred to as the *myocardium*, from the late-gadolinium enhancement cardiac MRI (LGE-MRI) in patients suffering from atrial fibrillation (AFib).

AFib is the most common cardiac arrhythmia and is characterized by irregular heartbeat due to rapid and asynchronous electrical impulses generated in the upper (atrial) chambers of the heart. AFib is a significant risk factor for stroke as well as other forms of heart disease [1], [2]. Radio frequency (RF) ablation is one of the conventional medical procedures used to treat AFib. The conventional wisdom is that the scar tissue resulting from the ablation prevents abnormal myocardial tissue from improperly propagating the circulating electrical waves that give rise to arrhythmia. However, RF ablation is an expensive, invasive procedure that has generally demonstrated effectiveness of less than about 50%, thereby limiting its overall utility.

Recent clinical data have shown that fibrosis in the atrial wall is linked to AFib, and that the degree of fibrosis is a strong predictor of the efficacy of RF ablation—patients with less fibrotic tissue are more likely to respond to ablation. LGE-MRI imaging highlights abnormal tissue, including fibrosis, within the heart wall [3], and is used for both quantifying the degree of fibrosis prior to RF ablation, as well as quantifying and tracking the effects of the RF ablation in the formation of scar tissue. The quantification of abnormal atrial tissue for any particular patient entails LGE-MRI imaging, the delineation or segmentation of the atrial wall, and the identification of the (brighter) fibrotic or scar tissue within that region [4], [5]. Due to the large number of patients presenting with AFib, the Department of Cardiology at the University of Utah processes several hundred patients a year, which includes the analyses of atrial heart tissue through MRI imaging, creating a huge workload for the technicians who painstakingly identify the atrial wall using a set of drawing tools within a 3D image editing software application.

The manual segmentation of the left atrium wall entails the delineation of its inner and

outer surfaces, namely the *endocardium* and *epicardium*. These manual segmentations can be time consuming (approximately one hour per image) and require significant manpower with sufficient domain-specific expertise. The expertise is necessary in order to apply anatomical knowledge in cases in which the image contrast is particularly weak or where nearby anatomy confounds the ability to locate a clear boundary for the atrium. Furthermore, these segmentations are prone to inconsistencies or discrepancies among experts. Automatic segmentation, on the other hand, promises to be an economical solution to such problems, with the benefit of reducing human variability or bias. Fig. 5.1 shows samples of such images where one can appreciate the challenges of extracting a thin wall structure (2-5 mm [6]) with weak boundaries and complex geometry. Further, these structures suffer from contrast variation, low signal-to-noise ratio, and motion artifacts [7].

To address the challenges of LA segmentation from LGE-MRI data, we propose a surface estimation scheme with the following aspects. First, it should handle various degrees of noise and estimate surfaces with ill-defined boundaries. Second, it needs to simultaneously estimate multiple surfaces in a globally optimal manner. Next, it should preserve the topology of the desired surface by taking into account the shape population of interest. Finally, the reconstructed surfaces are required to be regularized to compensate for misleading intensity information.

Current segmentation approaches handle such problems, but they are lacking in one way or another. For example, weak boundary segmentation problems are often solvable only by inducing an a priori notion of the shape characteristics in question. This triggers a strong motivation for a Bayesian framework for segmentation, which includes statistical models such as the *active shape models* (ASM) [8], and many variations derived from that approach. However, these models restrict the solution to lie *only* in a low-dimensional shape space defined through the training data. In the case of the LA, clinicians require

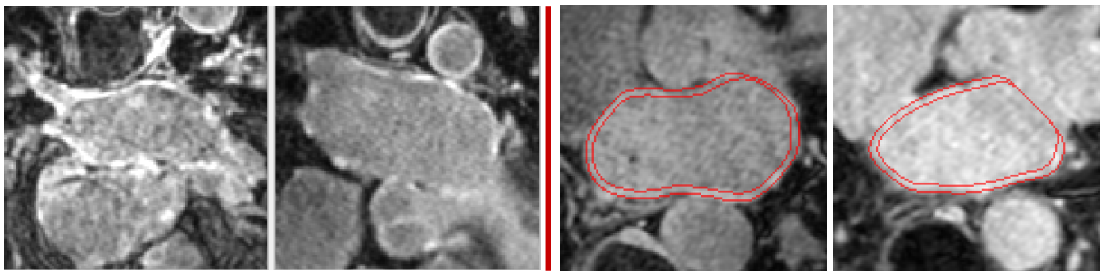


Figure 5.1. LA wall challenges: Slices of LGE-MRI images showing the challenges in segmentation. The left two 2D slices are shown without manual delineations of the LA wall, and the right two slices show manual delineations.

the output surfaces to follow correct boundaries rather closely and yet coarsely similar to a significant summary of the training dataset. Lower-level deformable shape methods, such as level sets [9], [10], which typically rely on gradient descent optimizations, have been shown to be unable to handle such large variations in the boundary contrast. Moreover, statistical and deformable methods rely on some sort of incremental fitting algorithm that is prone to get trapped in local minima during optimization, which has been demonstrated in Veni et al. [11]. The difficulty in optimizing shape models in the presence of such a challenging environment motivates us to consider an optimization strategy that seeks a globally optimal solution. In particular, graph-cuts provide a flexible global optimization tool with significant computational efficiency. Boykov and Kolmogorov [12], [13] developed efficient algorithms by using graph-cuts to segment regions optimally in the n-D image data. However, these methods are region-based and are not necessarily meant to extract features specific to the desired surface as verified by Veni et al. [14]. Moreover, they do not preserve the topology of the desired surface and use approximations while dealing with multiregion segmentation [15]. To extract multiple interlinked surfaces, Sonka et al. [16] used an optimal surface net approach. However, the technique embeds hard geometric constraints on surfaces smoothness and their interrelations that lead to irregular output surfaces.

In this paper, we address these issues by developing a surface-based segmentation algorithm with several important features that address unmet technical challenges in a range of applications. The primary features of our formulation include: 1) the reliance on a fully generative model that incorporates a global shape prior, 2) controlled smooth deviations from that prior that lead to soft penalties and eventually result in regular output surfaces, 3) a learned model of intensity patterns, 4) a combination of shape-based linear system and graph-cut approach to maximize *a posteriori* optimization (and hence we refer to this framework as *ShapeCut*), and 5) simultaneous extraction of multiple surfaces. Thus, the resulting optimal segmentation not only relies on the prior information but also accommodates irregularities in the input data in order to best match the observed image intensities.

The remainder of the paper is organized as follows: In Section 5.3, we provide the necessary background related to the presented work. In Section 5.4, we provide the theory behind ShapeCut, which includes shape-based generative modeling and the two fast global optimization strategies. Section 5.5 presents experimental results followed by their analysis in Section 5.6. Finally, we conclude with a summary of our approach in Section 5.6.4 along

with future directions.

5.3 Related Work

The relevant work falls into several different categories based on the approach and the effectiveness/efficiency for tasks relevant to the goals of this paper. We begin with the image segmentation methods based on shape priors that are learned from a set of training examples. This category includes active shape models (ASM) [8] and active appearance models (AAM) [17], which rely on training data with correspondences/landmarks and assume a Gaussian density in the low-dimensional subspace of landmark-based shape representations. Variations include nonparametric shape priors in the subspace defined by principle component analysis (PCA) of signed distance transforms [18]. Extending these methods that learn a low-dimensional shape prior, the proposed work relies on a low-dimensional subspace derived from the training shapes, but also accommodates perturbations from this relatively confined space. Furthermore, the proposed method relies on minimization schemes that optimize segmentations and shape parameters in a global fashion, rather than relying on incremental, gradient descent-based schemes.

The second class of segmentation methods falls under the category of *pixel-labeling*-based energy minimization schemes. Boykov and Kolmogorov [12], [13] used this scheme to segment objects in N-dimensional images. Energies typically include a regional term that quantifies the incompatibility of each pixel with a given region (e.g., foreground, background), and a smoothness term that defines the extent to which a region/labeling is piecewise smooth. To optimize such energy, graph-cuts have been shown to be efficient and globally optimal. However, these methods are best when they rely either on an initial set of seed points or on region-based penalties [19]. Further, they do not readily incorporate shape-based information. The resulting segmentation tends to be biased towards regions similar to the seed points and may not preserve the topology of the underlying surface. Thus, pixel-labeling methods cannot be applicable for the LA wall segmentation due to the localized intensity patterns surrounding the left atrial region and its shape complexity.

Malcolm et al. [20] extended the pixel-labeling graph-cuts framework by incorporating nonlinear shape priors. They used kernel PCA from a set of training shapes followed by iterative graph-cuts to segment the desired object, thus loosing the global nature of the original graph-cuts optimization. Vu and Manjunath [21] used a discrete version of the signed distance function to represent the shape prior within the level-set framework. Then, they merged this shape prior with the regional term of the minimizing energy and optimized

it using graph-cuts. Recently, Gorelick et al. [22] proposed a trust region approach to optimize the pixel-labeling-based segmentation energy that includes the nonlinear region terms, surface area penalties, and shape priors based on image moments. Within the trust region, they solved the nonlinear approximation of the underlying energy via graph-cuts. On the other hand, we view the segmentation problem as a surface estimation problem, rather than a pixel-labeling problem, that accommodates shape complexity and localized intensity patterns surrounding the surface being extracted. Thus, we introduce the shape priors in a more systematic, Bayesian framework with learned surface features (rather than region penalties), which are particularly important for the class of problems studied in this paper. In addition, the proposed framework simultaneously extracts multiple interlinked surfaces in a globally optimum manner. Although Delong et al. [23] and Schmidt et al. [24] proposed a pixel-labeling-based segmentation framework that extracts multiple regions by using a single graph-cut, it is very difficult to encode parameterized shape priors in their graph structure. Further, their framework lacks enforcing of smoothness constraints over the individual surfaces, which results in producing irregular boundaries. This has been shown in Section 5.5.4.

In order to deal with surface based features, deformable-surface methods like level sets [10], [9] can be used along with shape priors to extract the desired surfaces. For example, Rousson and Paragios [25], Chan and Zhu [26], and Cremers et al. [27] used shape priors in their level sets-based segmentation framework. However, like statistical shape models, deformable surface methods use gradient descent-based optimization approaches that are sensitive to initializations and are prone to get trapped in local minima.

The work most relevant to the proposed method is based on the optimal net surface problems on a special geometric graph structure introduced by Wu and Chen [28] and its applications to medical image segmentation [29], [30]. Wu and Chen designed a graph in the form of a multicolumn properly ordered (PO) graph, and they proved the equivalence between the image segmentation problem and *min-cut* on the $s-t$ graph via optimal net surface problems, thus providing a globally optimal and efficient solution. These surface-net problems induce costs on surfaces rather than on regions and do not rely on approximate (suboptimal) solutions when applied on multiple surfaces [31], motivating its practical use in surface-based segmentation approaches. Sonka et al. [31], [19], [32], [33], [34] used one of the surface-net approaches in the form of a *V-weight* net surface problem to simultaneously extract multiple interrelated surfaces in multiple clinical applications. Song et al. [30] extended this idea by introducing shape priors into such surface estimation schemes, but

the priors were strictly local, implemented as modifications to smoothness parameters in the optimal cut formulation. Veni et al. [11] used another form of the surface-net approach, namely, a *VCE-weight* net surface problem by enforcing soft constraints on surface smoothness and interrelations between coupled surfaces. However, they do not incorporate parameterized shape priors in their Bayesian formulation, which is essential for the LA wall segmentation.

The proposed work introduces a parameterized shape space and estimates global shape parameters in a closed form for segmentation as a part of the optimization process. Furthermore, the integration of shape space keeps it separate from the mesh/graph. Thus, the graph becomes, essentially, a discrete approximation to an underlying continuous parameterization of a subset of the 3D volume. The work in this paper builds on the previous work of the authors [11], [14]. The paper also discusses further improvements in the generative intensity model by embedding the covariance structure that exists between intensity profiles. We also talk about the gradient vector flow (GVF)-based nested mesh building strategy, which, in turn, defines the multicolumn graph. Further, the paper provides an extensive formulation of the shape prior model on coupled surfaces. With respect to the evaluation of results, we performed a thorough analysis of every aspect of the underlying model. Apart from the qualitative and quantitative analysis of segmentation results, we carried out their clinical evaluation, which could elucidate the effectiveness of the proposed work even from a clinical viewpoint.

5.4 Methodology

We treat the image segmentation as an energy minimization problem over the surface $\mathcal{S}$ to be estimated. Within a Bayesian framework, this energy minimization can be cast as a maximum-a-posteriori (MAP) estimation, where the log-posterior of the desired surface is maximized for a given image. However, the proposed work differs from many other Bayesian methods in that the underlying search space is an approximation of a continuous *parameterization* of the set of possible surfaces, as shown in Figure 5.2a. We discretize the underlying continuous parameterization in the form of a special form of geometric graph structure, in order to simultaneously estimate multiple, interacting surfaces in a globally optimal manner. Figure 5.2b depicts the strategy representing the continuous parameterization in the form of a discrete graph whose design involves a set of nested *layers* and *columns*. Here, each layer maintains a topological structure similar to that of the desired surface, and each column necessarily ensures the estimated surface to pass through

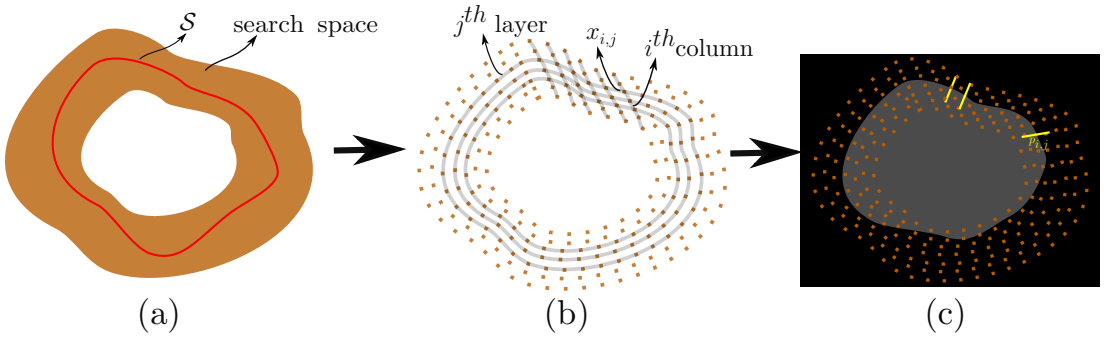


Figure 5.2. Parametric search space: Schematic showing a (a) continuous parameterization of the surface estimate $\mathcal{S}$, (b) discrete approximation of the underlying continuous parameterization within which the surface estimation takes place, and (c) overlay of discrete grid on a given image and intensity profiles, $p_{i,j}$ (*yellow*) at different grid points $x_{i,j}$.

it. For the 2D surface estimation, these layers represent 2D discrete contours, and in the case of 3D, these layers become meshes.

We begin by defining the graph structure along with the associated notations. The graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is represented by a set of columns and a neighborhood structure among adjacent columns, and hence is referred to as a *multicolumn* graph. This multicolumn graph is formed by a set of vertices $\mathcal{V}$ and a set of edges $\mathcal{E}$ that are connected in a properly ordered fashion [28]. The graph has I columns and J vertices along each column, forming J layers. Each column has a set of 3D vertices along its length. Thus, the columns span the 3D space in which the segmentation takes place. A particular vertex at the j^{th} position along i^{th} column is associated with a 3D location $x_{i,j}$ in a given volume. Thus, a surface $\mathcal{S}$ is the set of nodes $\mathcal{S} = \{s(i) \in [1, J]\},\ \forall i \in [1, I]$ representing the position of the surface along each column. Associated with each vertex is an intensity profile that may consist of either a 3D patch or a 1D array of intensities $p_{i,j}$ (e.g., oriented along the approximate normal to the surface), as shown in Figure 5.2c. Thus, a particular surface (set of vertices) induces a probability on a set of image intensities around each point on the surface.

5.4.1 Bayesian Formulation

Here we introduce the Bayesian model to express the posterior probability of a surface $\mathcal{S}$ as a function of the given image I. Using Bayes rule, we have

$$P(\mathcal{S}|I) \propto P(I|\mathcal{S})P(\mathcal{S}). \tag{5.1}$$

The image likelihood, $P(I|\mathcal{S})$, is a generative model of intensities in the vicinity of a particular surface estimate. [11] defined the image intensity model by means of a set of training images with an isotropic Gaussian distribution around a learned mean. Here we

model the intensity profiles for a surface at a column location with a column-dependent mean, μ_i, and covariance Σ_i. Following [11], the generative models for intensities are computed by analyzing the intensity profiles that lie within a predefined radius r of points on the column. We model the intensity profiles at different points on the surface as independent, and therefore, the log of the conditional probability is

$$-\log P(I|\mathcal{S}) \propto \sum_{i=1}^{I} \Big(\log|\Sigma_i| \; + \; (p_{i,s(i)} - \mu_i)^T {\Sigma_i}^{-1} (p_{i,s(i)} - \mu_i) \Big). \tag{5.2}$$

The surface prior, $P(\mathcal{S})$, is a generative model for surfaces. For this work, we use a linear model and an associated set of latent variables, β, to represent global shape information. This linear model is learned over training shapes, and captures low-dimensional properties of shape representation, as in [8]. To account for the actual complexity of the left atrium, the generative shape model allows smooth deviations or offsets from the low-dimensional shape distribution, formulated as a Markov random field (MRF) prior on their configuration. As a result, the proposed surface prior, $P(\mathcal{S})$, captures global and local characteristics of the desired surface. Ideally, because β is a latent variable, we formulate the complete prior in terms of a marginalization over β.

$$P(\mathcal{S}) = \int_{\beta} P(\mathcal{S}, \beta) d\beta = \int_{\beta} P(\mathcal{S}|\beta) P(\beta) d\beta. \tag{5.3}$$

Computing $P(\mathcal{S})$ as an integral over all instances of β is generally intractable. However, in the case of Gaussian distributions, we can assume that the integral in (5.3) is sharply peaked at a most probable latent setting β_{MP} where classical dimensionality reduction techniques provide a closed-form solution. Hence, we approximate the integral of (5.3) using a point estimate of β, and thus we have

$$P(\mathcal{S}) = P(\mathcal{S}|\beta_{MP}) P(\beta_{MP}), \tag{5.4}$$

where the point estimate β_{MP} can be interpreted as the mean of latent posterior, $p(\beta|\mathcal{S})$, as in [35]. This approach becomes the conventional *maximize-maximize* formulation of parameter estimation in the context of latent variable models. Hereafter, we will refer to the latent point estimate as β for notational convenience.

Figures 5.3a and 5.3b show a simplified geometric interpretation of the model. The linear subspace, which is learned from training data, is relatively low-dimensional, but allows the segmentation algorithm to operate effectively when parts of the anatomy are not well delineated in the image data, typically because of low-contrast and high-correlated and/or uncorrelated noise. A shape is generated from a position β on this low-dimensional

shape representation (shown as a 2D space in Figure 5.3), and we refer to β as *shape parameters*.

To represent and learn this shape space, we use the signed distance transform (DT) of the training data. Thus, for each training segmentation, namely d, whose zero level set corresponds to the surface boundary, we use principal component analysis (PCA) on the set of training DTs and choose a low-dimensional linear subspace that maintains 95% of the shape variability to model shape parameters [36]. From the training data, we also estimate a Gaussian distribution over the shape parameters, and therefore

$$P(\beta) \propto \exp(-\frac{1}{2}\beta^T \Lambda_K^{-1} \beta), \tag{5.5}$$

where Λ_K is the diagonal matrix of K largest eigenvalues derived from the PCA. The eigenvectors (eigen DT) d_{e_k} from PCA, along with a specific set of shape parameters β, give rise to a particular DT, which we denote d_β

$$d_\beta(x) = d_\mu(x) + \sum_{k=1}^{K} \beta_k d_{e_k}(x), \tag{5.6}$$

where $x \in \Re^3$ is the particular image location and d_μ denotes the mean DT. The zero level set of the base distance transform, $\mathcal{S}_\beta = \{x | d_\beta(x) = 0\}$ plays an important role in expressing the conditional surface probabilities, in terms of both a base shape and an offset.

Unlike previous shape-prior-based image segmentation methods, which restrict the surface estimate to remain in the low-dimensional subspace, the proposed method also allows offsets from this subspace while introducing a covariance structure on these offsets. To allow the offset, we model the conditional surface probability, $P(\mathcal{S}|\beta)$, in the form of MRFs. The MRF model enforces correlation among nearby points on the surface while prioritizing smooth deviations from the linear model learned from a set of training shapes. We let $\mathcal{S}_o$ be defined as an offset from the base shape, $\mathcal{S}_\beta$, which lies in the subspace. Thus, the surface estimate $\mathcal{S}$ is given by $\mathcal{S}_\beta + \mathcal{S}_o$ (as in Figure 5.3c). For a given set of shape parameters β, the surface probability depends only on the offset ($\mathcal{S} \equiv \mathcal{S}_o$), which we penalize so that it should be small (the shape should be similar to the base shape) and smooth (nearby columns should have similar offsets). Therefore, the negative log conditional probability of the surface for a given β can be defined as

$$-\log P(\mathcal{S}|\beta) = \frac{1}{2\sigma_o^2} \sum_{i=1}^{I} s_o^2(i) + \frac{1}{2\sigma_\mathcal{C}^2} \sum_{(i_1,i_2)\in\mathcal{C}} \left(s_o(i_1) - s_o(i_2)\right)^2, \tag{5.7}$$

where σ_o and $\sigma_\mathcal{C}$ are standard deviations of surface offsets learned from the training set and the neighborhood structure, defined by $\mathcal{C}$, which is the clique in the base graph defined over

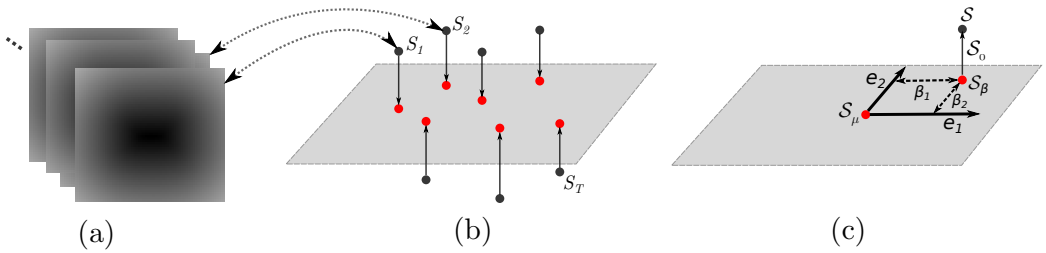


Figure 5.3. Shape representation within and off the subspace: Schematic showing (a) training shapes implicitly represented by distance transforms, (b) projection of shapes onto a linear subspace, and (c) estimated surface $\mathcal{S}$ derived from the deviation, $\mathcal{S}_o$, of the base shape $\mathcal{S}_\beta$.

the base shape. According to the MRF, the first term of (5.7) defines the unary potential of the negative *log*-likelihood of the surface, $\mathcal{S}$, given shape parameters, β, so that it encodes relatively small offsets from the linear subspace. The second term defines the pairwise clique potential, which encodes smooth deviation of $\mathcal{S}$ from the linear subspace. The combined effect of the unary and pairwise clique potentials allows the surface estimate to deviate from a low-dimensional linear subspace.

To express $P(\mathcal{S}|\beta)$ in (5.7) on the proposed grid, we use the following approximation. For a given β, we build the base shape $\mathcal{S}_\beta$ by computing the intersection of each column i with the zero crossing of d_β, to give a vertex and 3D positions for $s_\beta(i)$. The magnitude of the offset $\mathcal{S}_o$ from $\mathcal{S}_\beta$ is the distance along each column relative to the base shape, and the contribution to the log-prior is therefore $\sum_i s_o^2(i) = \sum_i d_\beta^2(s_i)$, where $s_i = x_{i,s(i)}$, the sum of square distances of the vertices along each column i from the base shape.

Combining (5.2), (5.5), and (5.7), we derive the negative log-posterior of a surface $\mathcal{S}$, given an input image I

$$\begin{aligned} -\log P(\mathcal{S}|I) = &\sum_{i=1}^{I} \left(\log|\Sigma_i| \; + \; (p_{i,s(i)} - \mu_i)^T {\Sigma_i}^{-1} (p_{i,s(i)} - \mu_i) \right) \\ &+ \frac{1}{2\sigma_o^2} \sum_{i=1}^{I} s_o^2(i) + \frac{1}{2\sigma_{\mathcal{C}}^2} \sum_{(i_1,i_2)\in\mathcal{C}} (s_o(i_1) - s_o(i_2))^2 \\ &+ \frac{1}{2} \beta^T \Lambda_K^{-1} \beta. \end{aligned} \tag{5.8}$$

The above derivation results in a full, shape-based generative model of an image. A point β in the low-dimensional PCA subspace, chosen from a multivariate normal distribution, gives rise to a distance transform, whose zero level set is the base shape, $\mathcal{S}_\beta$. An offset from $\mathcal{S}_\beta$ is chosen from an MRF distribution, which extends the learned probability distribution from the PCA subspace into the entire shape space in a way that favors smooth shapes

that are very near to but do not exactly resemble the training set. Finally, each point on that surface gives rise to an array (patch or vector) of image intensities that are drawn from a (identically, independently) Gaussian distribution, with a mean and covariance that are learned from the training data along each column.

5.4.2 Extension of Bayesian Formulation to Multiple Surfaces

The LA wall segmentation problem consists of a pair of nested surfaces—the epicardium and endocardium—which we denote $\mathcal{S}_1$ and $\mathcal{S}_2$, respectively. We extend the above formulation so that the distance transforms are modeled jointly. Similar to [11], we introduce a penalty on the intersurface distances (expressed in the same probabilistic framework, learned from the data, and tied to the β-shape-pair, that is β^1 and β^2) such that the surfaces are estimated optimally in a joint formulation. Therefore, the combination of (5.1) and (5.4) can be modified for two surfaces as follows

$$\begin{aligned} P(\mathcal{S}^1, \mathcal{S}^2|I) &\propto P(I|\mathcal{S}^1, \mathcal{S}^2)P(\mathcal{S}^1, \mathcal{S}^2) \\ &= P(I|\mathcal{S}^1, \mathcal{S}^2)P(\mathcal{S}^1|\mathcal{S}^2, \beta^1, \beta^2)P(\mathcal{S}^2|\beta^2, \beta^1)P(\beta^1, \beta^2). \end{aligned} \tag{5.9}$$

For this work, we assume that the intensity models for each column corresponding to epi- and endocardial surfaces are independent of each other with different mean intensity profiles μ_i^1 and μ_i^2 and covariance structures Σ_i^1 and Σ_i^2, respectively. Likewise, we apply the independence assumption over shape parameters β^1 and β^2. Following this, the probabilities for intensity and shape models are applied in a similar way for both surfaces as expressed in (5.2) and (5.5). Since β^2 is implicitly related to $\mathcal{S}^2$ while simplifying for $P(\mathcal{S}^1|\mathcal{S}^2, \beta^1, \beta^2)$ and β^1 remains fixed for $P(\mathcal{S}^2|\beta^2, \beta^1)$, (5.9) can be further simplified as

$$P(\mathcal{S}^1, \mathcal{S}^2|I) \propto \left[\prod_{l=1}^{2} P(I|\mathcal{S}^l)P(\beta^l)\right] P(\mathcal{S}^1|\mathcal{S}^2, \beta^1)P(\mathcal{S}^2|\beta^2). \tag{5.10}$$

To express the conditional surface probabilities, $P(\mathcal{S}^1|\mathcal{S}^2, \beta^1)$ and $P(\mathcal{S}^2|\beta^2)$, we account for the mutual relationship among the coupled surfaces apart from the individual surface estimates. By modeling $\mathcal{S}^1$ with a normal distribution on distance to $\mathcal{S}^2$, the negative *log* of the joint surface estimate is

$$
\begin{aligned}
&-\log P(\mathcal{S}^1|\mathcal{S}^2,\beta^1) - \log P(\mathcal{S}^2|\beta^2) \\
&\qquad = \sum_{l=1}^{2}\left[\frac{1}{2{\sigma^l}_o^2}\sum_{i=1}^{I}{s^l}_o^2(i) + \frac{1}{2{\sigma^l}_{\mathcal{C}}^2}\sum_{(i_1,i_2)\in\mathcal{C}}\Big({s^l}_o(i_1) - {s^l}_o(i_2)\Big)^2\right] \\
&\qquad + \frac{1}{2\sigma_{\mathcal{C}'}^2}\sum_{i=1}^{I}\Big(s^1(i) - \Big(s^2(i) + \Delta(i)\Big)\Big)^2, \qquad (5.11)
\end{aligned}
$$

where $\Delta(i)$ denotes the ideal intersurface distance for each column i and $\sigma_{\mathcal{C}'}$ represents the standard deviation of this distance, both of which are learned from the training data. For simplicity, let the intersurface relationship between the epi- and endocardial surfaces, $s^1(i) - \Big(s^2(i) + \Delta(i)\Big)$, be denoted by ρ. By following the fact that the epicardial surface is always above the endocardial surface, we let $\rho < -\Delta(i)$. This condition will make sure that $s^1(i) < s^2(i)$ for any given mesh column and eventually avoids either crossing or overlapping of coupled surfaces. Therefore, the negative *log*-posterior of coupled surfaces $\mathcal{S}^1$ and $\mathcal{S}^2$, and shape parameters, β, with an input image I, is given by

$$
\begin{aligned}
-\log P(\mathcal{S}^1,\mathcal{S}^2|I) = \sum_{l=1}^{2}\Bigg[&\sum_{i=1}^{I}\Big(\log|{\Sigma^l}_i| \;+\; (p_{i,s(i)} - {\mu^l}_i)^T {\Sigma^l}_i^{-1}(p_{i,s(i)} - {\mu^l}_i)\Big) \\
&+ \frac{1}{2{\sigma^l}_o^2}\sum_{i=1}^{I}{s^l}_o^2(i) + \frac{1}{2{\sigma^l}_{\mathcal{C}}^2}\sum_{(i_1,i_2)\in\mathcal{C}}\Big({s^l}_o(i_1) - {s^l}_o(i_2)\Big)^2 \\
&+ \frac{1}{2}\beta^{lT}{\Lambda^l}_K^{-1}\beta^l\Bigg] + \frac{1}{2\sigma_{\mathcal{C}'}^2}\sum_{i=1}^{I}\Big(s^1(i) - \Big(s^2(i) + \Delta(i)\Big)\Big)^2. \qquad (5.12)
\end{aligned}
$$

5.4.3 3D Mesh Construction

The proposed Bayesian formulation over $\mathcal{S}$ on a surface-conforming multicolumn graph structure requires a 3D nested mesh that defines a 3D subdomain that contains the surfaces to be estimated. We propose a mesh building strategy that relies on a set of training samples.

As seen in Figure 5.4, the shapes of the left atrium are highly variable, especially within and around the veins that come out of it. This variability is captured by the statistical model, described above, and the use of training data for the estimation of the computational domain.

In order to construct a 3D nested mesh structure, we initially need to build a base mesh that complies with the training set of segmented shapes. We begin by rigidly aligning the training set of segmented images, using FSL-FLIRT [37] to perform the pairwise registration step to a group atlas. In particular, with 7 degrees-of-freedom (translation, rotation, and scaling), we align each training shape to an iteratively evolving average (template) until

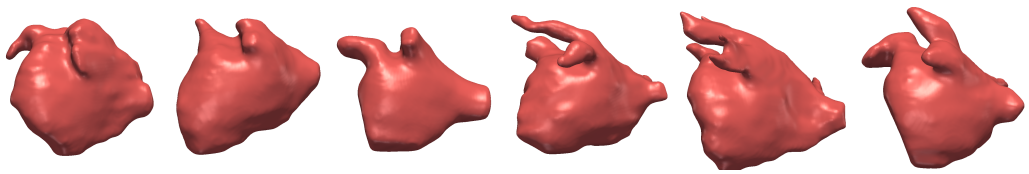


Figure 5.4. Variability of left atrium shapes.

convergence (i.e., a rigid with-scale unbiased atlas building strategy over all training shapes). To build our nested mesh model, we use the mean signed distance transform (DT), d_μ from the aligned training shapes. The zero-level set of d_μ gives rise to a new distance transform $\tilde{d}_\mu$ with respect to the average shape, using the Fast Marching Method (FMM). Using marching cubes and subsequent decimation and resampling to improve quality [38], we have a triangulated *base mesh* with relatively good mesh quality (aspect ratios and triangle sizes).

Once the base mesh is generated, we construct nested mesh layers that represent a collection of offset surfaces, both inside and out. For this we use a gradient vector flows (GVF)-based nested mesh building strategy [39] derived from the distance field, $\tilde{d}_\mu$. A small amount of smoothing in estimating the gradients and sufficiently small integration steps ensure that columns do not cross or tangle. Starting at each node of the base mesh, we traverse along the direction of the gradient vectors while creating new nodes successively. These nodes are eventually used to construct graph columns. Specifically, consider a node in a mesh that is located at $x_{i,j}^{old}$. To acquire a corresponding node $x_{i,j}^{new}$ for the next mesh layer, either inside or outside the base mesh, we apply

$$x_{i,j}^{new} = \begin{cases} x_{i,j}^{old} + \delta \bigtriangledown \tilde{d}_{\mu_{i,j}}^{old}, & \text{if } \textit{inside} \\ x_{i,j}^{old} - \delta \bigtriangledown \tilde{d}_{\mu_{i,j}}^{old}, & \text{otherwise} \end{cases} \tag{5.13}$$

where δ denotes the distance between two successive mesh layers and $\tilde{d}_{\mu_{i,j}}^{0}$ represents the zero-level set of $\tilde{d}_\mu$. A collection of all the nodes forms a new mesh layer that maintains the topology of the base mesh. This action continues so that the nested mesh layers cover the required computational domain. For our application, the number of mesh nodes per each layer corresponds to 12665 for the heart data and 6262 for the synthetic data. Based on the capture range of two surfaces, we generated 75 mesh layers, spaced at 0.5 voxels each, which occupy a capture range of 38 voxels for the final surface estimate.

5.4.4 Optimization

To seek globally optimal solution for $\mathcal{S}$ and β, we follow a conditional maximization scheme in which we iterate between two relatively fast, *global* optimization phases. In the

first phase, we optimize for $\mathcal{S}$ given β, and in the second phase, we optimize for β given $\mathcal{S}$.

For the first phase, we use graph-cuts to solve the *VCE-weight* net surface problem, using [28]. Here we use the MRF formulation [11], which follows the derivation introduced in [40]. The strategy is to express the above penalties as a set of weights on vertices and edges of a multicolumn properly ordered (PO) graph, and reduce the surface estimation problem to an *s-excess* cut on the graph. This produces an optimal low-order polynomial time solution–and in practice runs very fast, even for large graphs. The details of the multicolumn PO graph construction follow from [28], and we present here how our energy minimization-based probabilities get encoded on the vertices and edges of the PO graph.

Weights/costs on the graph vertices arise from two sources. The first source of costs is derived from the intensity model, $P(I|\mathcal{S})$, by sampling image patches around each vertex $p_{i,j}$, and computing the negative log-likelihood from the Gaussian model, (μ_i, Σ_i). For notational convenience, we denote the negative log-likelihood by W^I. The second part comes from the shape model by sampling the squared distance transform d_β of the base shape $\mathcal{S}_\beta$ at each vertex position (i, j) in the form of surface offset $\mathcal{S}_o$ and is denoted by $W^{\mathcal{S}_\beta}$. These terms are weighted by a dominance factor ω, and added to induce vertex costs $W_{i,j}$:

$$W_{i,j} = \omega W^I_{i,j} + (1-\omega) W^{\mathcal{S}_\beta}_{i,j}; (0 \leq \omega \leq 1). \tag{5.14}$$

Typically, for low-contrast images, ω is chosen to be less than 0.5 in order to emphasize shape-based costs over the image matching-based costs.

To build the PO graph, edges are inserted between adjacent vertices along each column carrying $+\infty$ costs. They also go in between vertices of adjacent columns to model the pairwise clique potential $C(s_o(i_1), s_o(i_2))$, associated with the offsets, rather than the absolute positions of the surface, as in [11]. For the optimization to be feasible, we use $C(d) = \alpha d^2$, where α represents the scaling factor on the clique potential. Typically, α regulates the penalty on the smoothness of the estimated surface, and $d^2 = (s_o(i_1) - s_o(i_2))^2$ comes from the surface smoothness part in (5.7). The configuration of the edges between adjacent columns is inherently managed by the β-shape, $\mathcal{S}_\beta$. To be specific, the intersection of $\mathcal{S}_\beta$ with each graph column gives a reference position, s_β. The s_o's between adjacent columns, defined by the clique potential C, are adjusted relative to s_β's positions. Figure 5.5 illustrates the arc configuration with respect to the intra- and intercolumn arcs. In this way, the nested mesh on which we perform graph-cuts becomes a numerical construct and does not define or heavily influence the shapes of the objects to be segmented. If a mesh is sufficiently fine grained, it serves only to restrict the topology of the surface. For a given β,

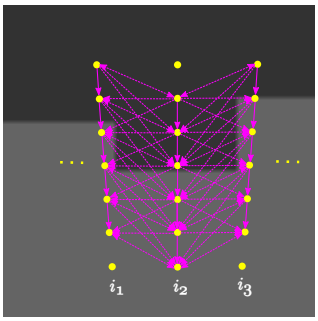


Figure 5.5. Shape complying properly ordered graph construction: Schematic showing arc configuration between adjacent columns being overlaid on a given image. The entire arc configuration assembly is deformed with respect to the base shape $\mathcal{S}_\beta$ during each iteration.

the surface estimate $\mathcal{S}$, computed using graph-cuts, is the global optimum of the posterior in (5.1).

In the second phase of the optimization, we compute β for a given surface $\mathcal{S}$ by minimizing the negative log-posterior. This computation is done by taking the derivative of (5.8) with respect to β and equating it to zero. Since the negative log-posterior is quadratic in terms of β, we obtain a closed-form solution.

$$\begin{aligned} &\frac{1}{\sigma_o^2}\sum_i d_\beta\left(s(i)\right)\frac{\partial d_\beta}{\partial \beta}\left(s(i)\right) + \Lambda_K^{-1}\beta \\ &+ \frac{1}{\sigma_c^2}\sum_{(i_1,i_2)}\left(d_\beta\left(s(i_1)\right) - d_\beta\left(s(i_2)\right)\right)\left(\frac{\partial d_\beta}{\partial \beta}\left(s(i_1)\right) - \frac{\partial d_\beta}{\partial \beta}\left(s(i_2)\right)\right) = 0. \end{aligned} \tag{5.15}$$

Here d_β represents the base distance transform and is expressed in (5.6). The images $\partial d_\beta/\partial\beta$ are the principal components. By rearranging terms in (5.15), the optimal shape parameters, β^*, are given by

$$\beta^* = -(\sigma_c^2\mathcal{P}^a + \sigma_o^2\sigma_c^2\Lambda_K^{-1} + 2\sigma_o^2(\mathcal{P}^a - \mathcal{P}^b))^{-1}(\sigma_c^2 q^a + 2\sigma_o^2(q^a - q^b)) \tag{5.16}$$

$$q_k^a = {d_\mu}^T(\mathcal{S})d_{e_k}(\mathcal{S}), \text{and } \; q_k^b = {d_\mu}^T(\mathcal{S})\mathcal{C}_\mathcal{N} d_{e_k}(\mathcal{S}), \; \forall k \in [1, K] \tag{5.17}$$

$$\mathcal{P}_{k_a k_b}^a = d_{e_{k_a}}^T(\mathcal{S})d_{e_{k_b}}(\mathcal{S}), \text{and } \; \mathcal{P}_{k_a k_b}^b = d_{e_{k_a}}(\mathcal{S})\mathcal{C}_\mathcal{N} d_{e_{k_b}}(\mathcal{S}), \; \forall k_a, k_b \in [1, K]. \tag{5.18}$$

In equations (5.17) and (5.18), $\mathcal{C}_\mathcal{N}$ represents $I \times I$ symmetric, indicator matrix, with the corresponding elements, $i \in I$, equal to 1 if $i' \in \mathcal{N}_i$ (neighborhood of i), else 0. Thus, (5.16) gives the *optimal* β for a given surface segmentation that finds a compromise between fitting the current segmentation and choosing a likely β based on the Gaussian distribution.

When we consider the algorithm in full, the proposed method differs significantly from others that either work strictly within the shape subspace [8] or project intermediate solutions onto that subspace. Here, the optimization strategy offers two advantages. First,

it allows intermediate results to deviate from the low-dimensional shape subspace and to pull the shape parameters, β, toward this segmented surface. Second, in each iteration, we perform a global optimization of $\mathcal{S}$ and β, thus relying less on local image properties near the current solution. Thus, the algorithm converges with relatively few iterations (e.g., $\sim$5 iterations on average for results in this paper).

While dealing with the multisurface problem, we follow the same iterative conditional maximization scheme where we optimize for coupled surfaces, $\mathcal{S}^1$ and $\mathcal{S}^2$, for a given β-shape-pair in the first phase, and in the second phase, we optimize shape parameters β^1 and β^2 independently for $\mathcal{S}^1$ and $\mathcal{S}^2$, respectively. At each iteration, for a given $\mathcal{S}^l$, the mutual interaction between surfaces does not influence the estimation of shape parameters, β^l, allowing each β^l to be estimated independently in the same way as in (5.16). However, to estimate coupled surfaces for a given β-shape-pair, we consider the mutual interaction estimate and optimize the VCE-weight net surface problem using graph-cuts.

5.5 Results

To assess the performance of the proposed method, we applied ShapeCut on the left atrial (LA) wall segmentation from late-gadolinium enhancement MRI. As a proof-of-concept, we start by demonstrating the algorithm's performance on 3D simulated parametric shape examples that have coupled-surface structure similar to the LA wall.

5.5.1 Datasets

5.5.1.1 Synthetic Dataset

We considered a class of 3D parametric shapes termed *supershapes* [41], [42] whose 2D parameterization in polar coordinates can be written as

$$r(\phi) = \left[\left| \frac{\cos(m\phi/4)}{a} \right|^{n_2} + \left| \frac{\sin(m\phi/4)}{b} \right|^{n_3} \right]^{-1/n_1}. \tag{5.19}$$

The 3D extension was obtained by a spherical product:

$$\begin{aligned} x &= r(\theta)\cos\theta \; r(\phi)\cos\phi \\ y &= r(\theta)\sin\theta \; r(\phi)\cos\phi \\ z &= r(\phi)\sin\phi \end{aligned} \tag{5.20}$$

where $\theta \in [-\pi/2, \pi/2]$, $\phi \in [-\pi, \pi]$ and we set $a = b = 1$. We generated shapes of four rotational symmetries, $m = 4$, to mimic pulmonary vein structures coming out of the LA. The associated shape parameters (n_1, n_2, n_3) were drawn from χ^2 distributions with

four degrees of freedom. We used the implicit representation of rational supershapes in [42], which yields smoother shapes compared to the standard ones in [41] that suffer from sharp corners. We imitated coupled-surfaces using two level sets of $r(\phi)$: the zero-level set associated to the generated supershape, and a positive level set to simulate the inner surface. Using this strategy, we produced 170 volumes of size $200 \times 200 \times 200$ where each volume contained three regions: the region inside the innermost surface (*innermost region*), the region in between the inner and outer surfaces (*middle region*), and the background. To mimic variations associated with the LA-wall thickness, we perturbed the inner and outer surfaces by introducing correlated random noise determined by convolution of random noise with a Gaussian filter of standard deviation 25 to produce smooth wiggles along both surfaces. To simulate image intensities that reflect appearance differences between distinct regions, we considered a Gaussian-distributed appearance model within each region of variance 9 and means 40, 70, and 10 for the innermost, middle, and background regions, respectively. To replicate some of the MRI imaging challenges, each image was corrupted with Rician noise with variance 30 and a smoothly varying bias field.

5.5.1.2 LGE-MRI Cardiac Dataset

For the real-data experiments, we apply ShapeCut on the LA wall segmentation where 72 LGE-MRI volumes (37 pre- and 35 postablation, each of size $300 \times 400 \times 107$ and isotropic resolution of 0.625 mm) were obtained retrospectively from an AFib patient image database at the University of Utah's Comprehensive Arrhythmia Research and Management (CARMA) Center.[2] Along with the LGE-MRI volumes, we were provided with expert-delineated binary segmentations for the epicardium and endocardium. Of the 72 samples dataset, 62 volumes were served with a single-handed segmentations from an expert, and the remaining 10 volumes were provided with binary segmentations from three experts. These segmentations were used to generate the properly ordered (PO) graph, extract intensity features, learn the shape model, and for the purpose of evaluation.

5.5.2 Shape Prior Construction: Training Sample Size

Learning shape priors poses convergence challenges for high-dimensional shape spaces and relatively low training sample size. Hence, it is critical to show that the underlying shape variability in a given population can be captured by a finite, probably small, number of training shape exemplars. Such training samples account for linear subspace approximation,

[2] http://healthsciences.utah.edu/carma/carma_team/index.html

for which even fewer eigenmodes, relative to the training sample size, would capture the population shape variability. To assess the algorithm's performance against the number of training samples, we used a repeated random-subsampling-based resampling strategy [43] where repeated *trials* (subsets) were drawn to compute performance statistics. Figure 5.6 shows average scree plots from repeated random training subsets of different sample sizes for the synthetic and LA datasets. It can be observed that 95% of the shape variability can be captured by a smaller number of eigenmodes relative to the training sample size. Note that for high-dimensional shape spaces, we do not require an exact characterization of the covariance structure in the off-subspace. For a segmentation task, we allow for deviation from the shape subspace in a regularized manner in the form of MRF-based clique potentials.

Another important aspect of building shape priors is *subspace consistency* where the eigenmodes are reliably estimated with the given finite training set. To this end, we adopted two shape-based PCA measures, *projection error* and *subspace error*, that were computed from data subsets. These measures reflect the representation capability of the learned shape prior in terms of deriving shapes (i.e., $\mathcal{S}_\beta$) that resemble the training set while determining the deviation from the prior's subspace (i.e., $\mathcal{S}_o$), which is needed to accurately represent a testing shape. Smaller errors, relative to the off-subspace noise variance, eventually help in obtaining a reliable segmentation. We define these measures as follows.

5.5.2.1 Projection Error

This measure quantifies the distance between the original set of centered samples $\tilde{X}$ and the reconstructed samples derived from a learned shape prior. Mathematically, it is given

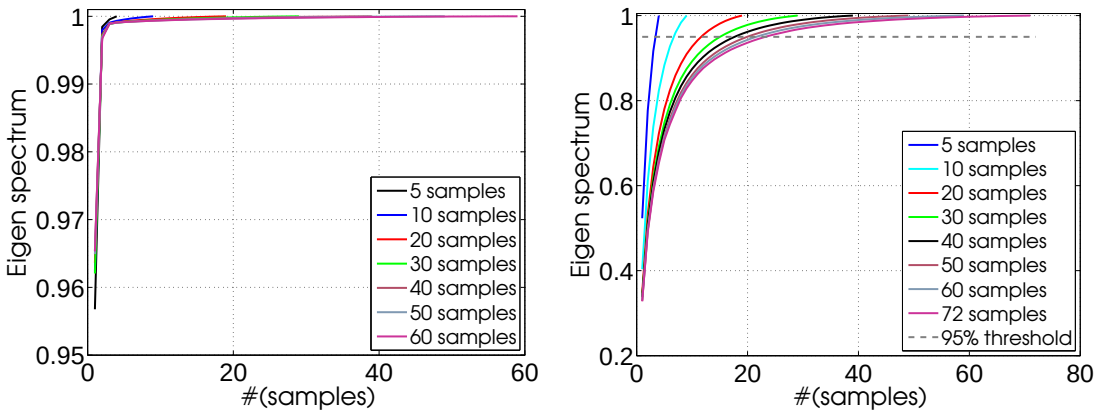


Figure 5.6. Eigen spectrum: Average scree plots for different training set sizes for (a) the synthetic dataset and (b) the left atrium.

by

$$||\tilde{X} - U_k^j {U_k^j}^T \tilde{X}||_F^2. \tag{5.21}$$

where, U_k^j represents a set of k significant eigenvectors derived from the PCA on subsampled data for the j^{th} trial and F represents the Frobinious norm. From the segmentation standpoint, this error provides the surface offset $\mathcal{S}_o$ from a within-subspace shape reconstruction $\mathcal{S}_\beta$ that is needed to represent a test shape. Since the surface probability demands small and smooth offsets, it is desirable to have a small projection error at the cost of the number of training samples.

5.5.2.2 Subspace Error

This measure computes the distance between the reconstructed samples derived from the eigen decomposition of the subsampled data against the reconstructed samples derived from the entire data. Mathematically, it is given by

$$||W_k {W_k}^T \tilde{X} - U_k^j {U_k^j}^T \tilde{X}||_F^2. \tag{5.22}$$

where W_k represents a set of k significant eigenvectors derived from the PCA with respect to the entire dataset. From the segmentation standpoint, it provides the distance between base shapes $\mathcal{S}_\beta$ that are defined over the subspaces spanned by subsets of the dataset and the entire dataset itself. This measure encodes the consistency of prior learning using smaller training sizes as compared to the full dataset.

For both these measures, we again used a repeated random-subsampling-based resampling strategy to draw multiple random samples for different sample sizes. For synthetic as well as the left atrium datasets, we chose training samples sizes ranging from 10 through 60 in the interval of 10 and drew 150 random subsets for each training size. Finally, for each subset, sample statistics were calculated from these trials. Figure 5.7 shows the mean and standard deviation (in noise variance units considering all training samples) of these measures by considering different numbers of subsamples. As expected, the errors decrease with the number of samples. In particular, the projection error levels off at or beyond the off-subspace noise variance for both datasets, which suggests that it served to learn shape priors with relatively small-sized training datasets (e.g., as would be available in clinical applications). Similarly, the convergence of subspace error to less than noise variance (especially for the LA data) is indicative of the shape prior robustness and consistency in learning eigenmodes that capture shape variability with a finite subset of the training shapes.

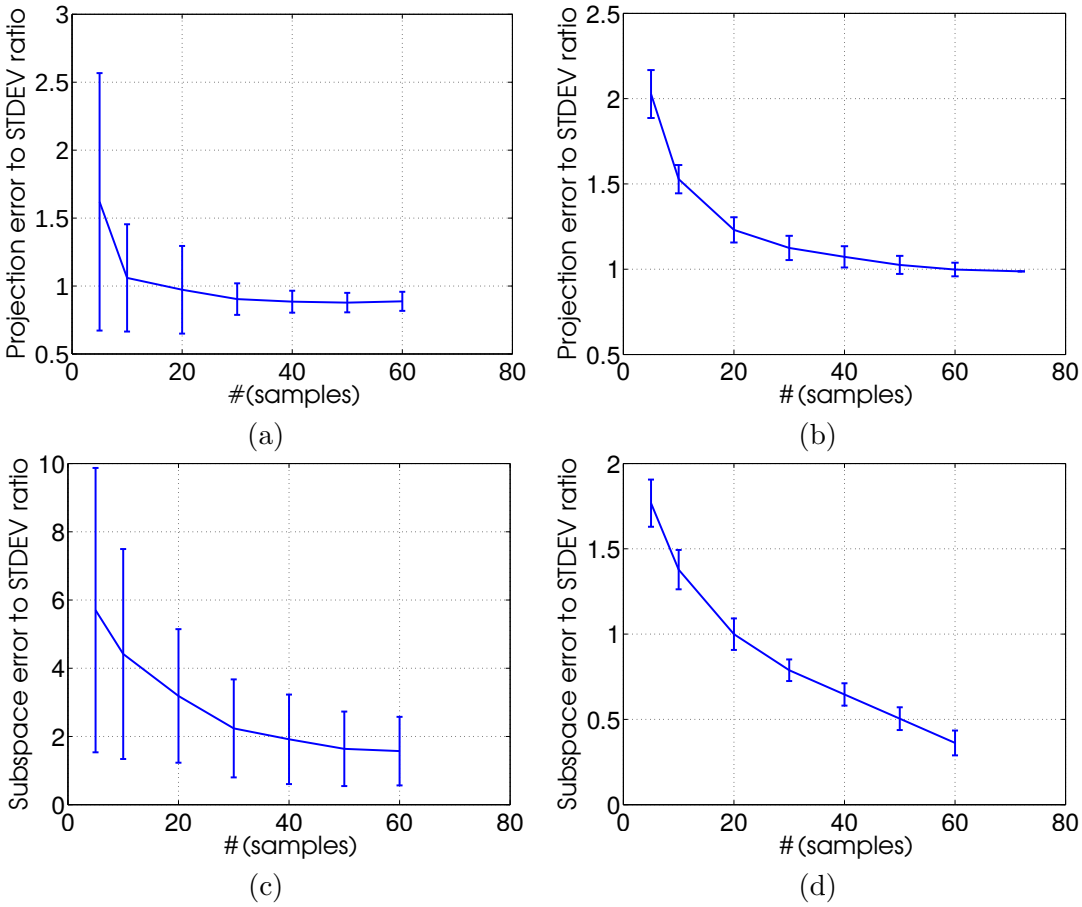


Figure 5.7. Shape-based PCA measures: Projection error (*top*) for (a) the synthetic dataset and (b) the left atrium, subspace error (*bottom*) for (c) the synthetic dataset and (d) the left atrium.

5.5.3 Algorithm Flow

For a given test example, ShapeCut relies on user input to locate the model's approximate center, based on which every point of the 3D nested mesh is positioned. This approach is viable because, according to acquisition protocols deployed by CARMA, subjects are generally aligned while being imaged in a scanner. Although the transformation parameters associated with the alignment can be estimated as a part of the optimization process, we opt to focus on the shape parameter estimation aspect, considering the problem at hand.

The algorithm is robust to this position, as long as the nested mesh model does not fall either outside or inside the desired surface. The input image is then sampled along the patch positions at all mesh nodes. Next, the negative log-posterior is computed based on (5.12), which reflects the assignment of costs, weights, edge capacities, and an optimal cut. A pair of optimal mesh surfaces is then recovered from the minimum s-t cut. Using this *cut*, the shape parameters are optimized. This procedure is repeated until convergence. We

use Boykov's maximum flow algorithm [44] to perform graph-cuts.

5.5.4 Evaluation

Based on the training sample size selection results from the synthetic dataset, we chose 20 samples for training and the remaining 150 samples for testing purposes. Figure 5.8 shows some example slices of supershapes and the corresponding noise- and bias-corrupted simulated images with the segmentation boundaries overlaid. These results demonstrate the effectiveness of ShapeCut in retrieving the inner as well as outer surfaces correctly even in the presence of correlated noise that ambiguates region boundaries. However, the proposed model does not exactly delineate the inner boundary on the top portion of the rightmost image because of the high noise level in that specific region for the actual boundary to be distinguishable, as can be noticed in its simulated image. In these occluded boundary regions, the PCA-based shape model dominates the energy minimization process in driving the output towards the representative shape $\mathcal{S}_\beta$ that is derived from the learned prior.

We compared ShapeCut against one of the state-of-the-art pixel-labeling-based energy minimization approaches [23, 24, 45]. These methods also use graph-cuts to efficiently optimize the corresponding energies. However, the difference is that their energies rely on region labelings rather than on surfaces. Specifically, we chose [23] for comparison as it deals with multiregion segmentation by using a single graph-cuts. To mimic the multiregion energy terms with the proposed method, the data term comprises the intensity-based likelihood model and the shape term, which is derived from the mean distance transform, d_μ. In order to define the smoothness term, we used a pairwise intensity difference-based Gaussian model [12]. Figure 5.9 compares the segmentation results of the middle region on two test examples. Based on these results, one can observe that the ShapeCut surpasses the multiregion approach in producing accurate surfaces with regularized boundaries. This superiority is achieved because, in the areas that correspond to the ill-defined boundaries, ShapeCut takes advantage of using parameterized shape priors in its graph structure to guide the segmentation process. Another reason is related to the MRF-based smoothness constraints that are enforced on individual surfaces, rather than regions, maintain their regularity.

With regard to the LA wall segmentation, Figure 5.10 shows the extracted boundaries of epicardial and endocardial surfaces on some example slices from LGE-MRI image volumes. Due to the lack of ground-truth information, these surfaces are compared against the manually segmented images. Although the 2D contours of our automatic segmentation are shown on the respective slices of three images, the actual 3D surfaces produced by the

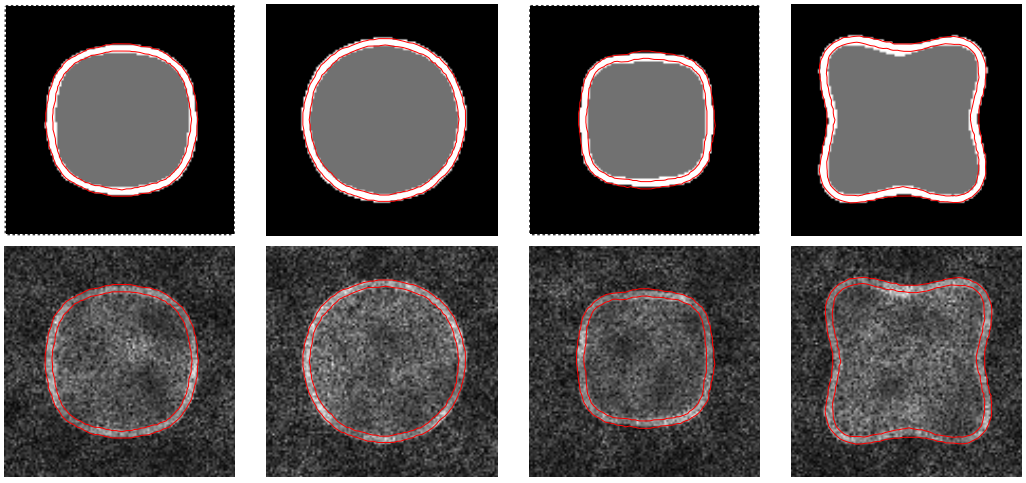


Figure 5.8. Supershape qualitative results: 2D contours derived from the proposed method overlaid on supershape ground-truths (*top*) and the corresponding noisy test examples (*bottom*).

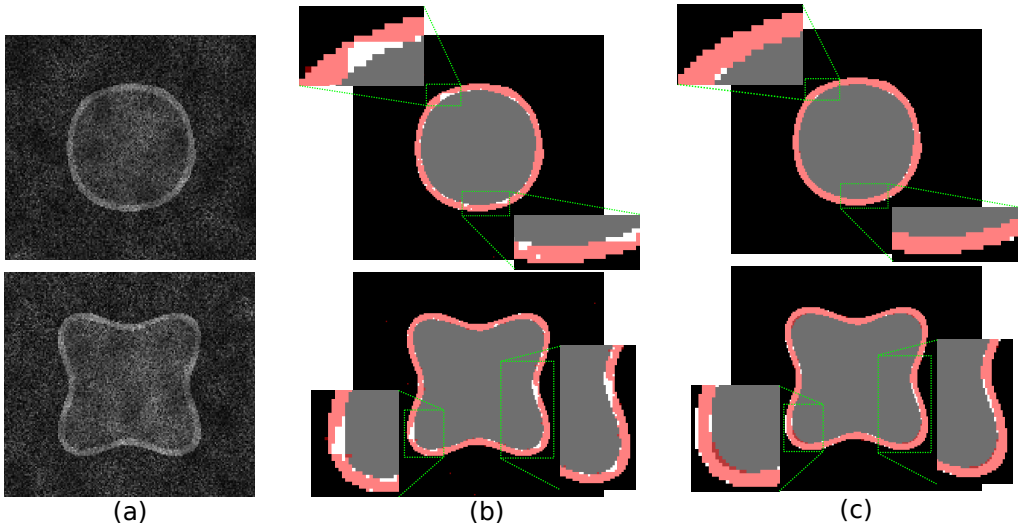


Figure 5.9. Supershape qualitative comparison: (a) 2D slices of the 3D supershape test examples. Segmentation results of the middle region on the corresponding ground truth by using (b) multiregion graph-cuts [23] and (c) ShapeCut.

proposed method are accurate enough in the core left atrial regions except near the veins, which, in any case, do not belong to the LA. Moreover, the delineation of these veins is subject to interrater variability, as the cutoff between atrium and veins is not well defined.

To validate this hypothesis, we calculated statistics on pairwise surface distances between manually and automatically extracted surfaces across all test samples. These statistics include means and standard deviations of the pairwise surface distances that could be visualized as color maps on the base mesh. According to the left panels of Figures 5.11

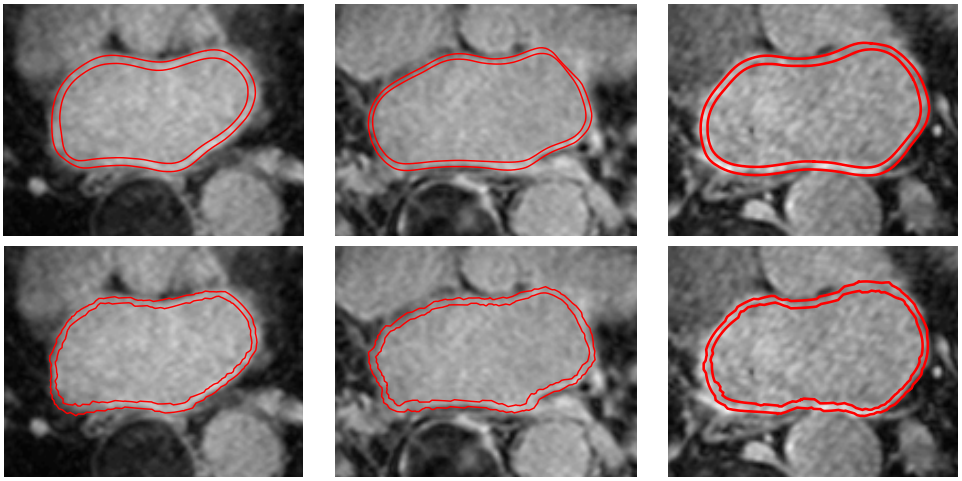


Figure 5.10. Left atrium qualitative results: 2D contours derived from manual segmentations (*top*) and segmentations obtained using the proposed method (*bottom*) over LGE-MRI slices.

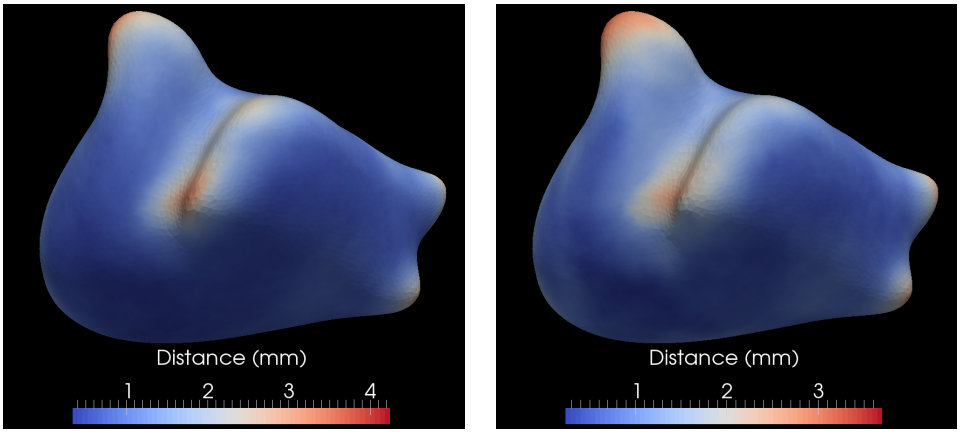
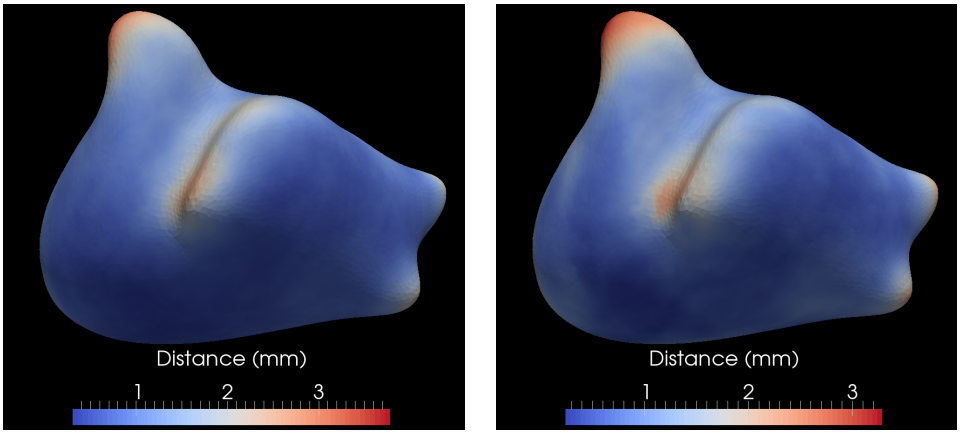


Figure 5.11. Epicardium quantitative results: Mean (*left*) and standard deviation (*right*) of epicardial surface distances across all the datasets shown as color maps on the base mesh surface.

and 5.12 that depict mean surface distances, most surface delineation incompatibilities are concentrated around veins. In addition, the right panels of Figures 5.11 and 5.12 show the deviation of pairwise surface distances across different datasets from the corresponding means, which is indirectly related to the interrater variability and testifies to how differently an expert can delimit LA boundaries, especially around vein regions.

Apart from the qualitative assessment of results, we performed quantitative analysis to evaluate the accuracy and precision of our segmentation results as compared to the ground-truths for simulated shapes and manual segmentations of the left atrium wall. Choosing



(a) (b)

Figure 5.12. Endocardium quantitative results: Mean (*left*) and standard deviation (*right*) of endocardial surface distances across all the datasets shown as color maps on the base mesh surface.

the correct evaluation measure is a crucial but nontrivial task and is needed to reflect the nature of the object to be segmented. Specifically, when the object size is small (as in the case of the LA wall) as compared to the image volume, overlap-based measures are generally not recommended due to their insensitive nature in providing the same metric value irrespective of the distance between two nonoverlapping segments being evaluated, eventually affecting the precision [46]. On the other hand, distance-based measures provide better insight into the precision and accuracy of both the *shape* and the *local alignment* of segmented regions. Apart from its small size, the LA wall is accompanied by the adjoining vein structures that exhibit significant inconsistency among manual segmenters and are not a part of the wall tissue, as mentioned earlier. Thus, they can be considered as *outliers*. Since Hausdorff distance (HD) is sensitive to outliers, it is not considered to be the right choice for the LA wall quantitative assessment. As such, we chose average surface distance (ASD) for the quantitative evaluation, which computes the average of all pairwise surface distances between two surfaces. Even in the case of ASD, the percentage of veins structures is large enough to influence the analysis. Hence, we tried to carefully eliminate these regions by acquiring help from the experts at the CARMA Center and computed ASD on the rest of the surface regions. Since the segmentation of the LA wall involves the segmentation of two surfaces, the epicardium and endocardium, we calculate the ASD of the LA wall as

$$ASD_{wall} = max\{ASD_{epi}, ASD_{endo}\}.$$

Unlike the case of the LA, the synthetic dataset does not possess outlier sensitivity issues. Hence, apart from the ASD, we have also calculated the Hausdorff distance (HD) between

the extracted surfaces using ShapeCut and the corresponding ground-truths. Similar to the LA dataset, the ASD and HD for the middle surface are computed using

$$\begin{aligned} ASD_m &= max\{ASD_o, ASD_i\} \\ HD_m &= max\{HD_o, HD_i\} \end{aligned}$$

where the subscripts m, o, and i on distance measures are used to represent the middle, outer, and inner regions, respectively.

For the synthetic dataset, we obtained an average ASD of (0.25 ± 0.04) pixels over all 150 test samples for the middle region with the corresponding histogram shown in Figure 5.13a. Further, an average HD over the test samples is (1.95 ± 0.38), with the corresponding histogram being shown in Figure 5.13b. For the LA wall, we obtained an average ASD of (0.66 ± 0.14) mm over all 72 test samples, with the corresponding histogram shown in Figure 5.13c.

5.5.5 Sensitivity Analysis

To assess the robustness of ShapeCut with respect to model parameters while analyzing its effect on the segmentation output, we performed sensitivity analysis on these parameters for the LA data. These model parameters include the patch length p, smoothness penalty α, and dominance factor ω. To carry out the quantitative analysis, we used an average pairwise surface distance (ASD) of the wall for evaluation. Initially, we analyzed patch profiles by choosing different lengths. Next, we fixed the patch length and varied the scaling factor of the smoothness penalty to analyze its effect on the segmentation accuracy. Finally, we

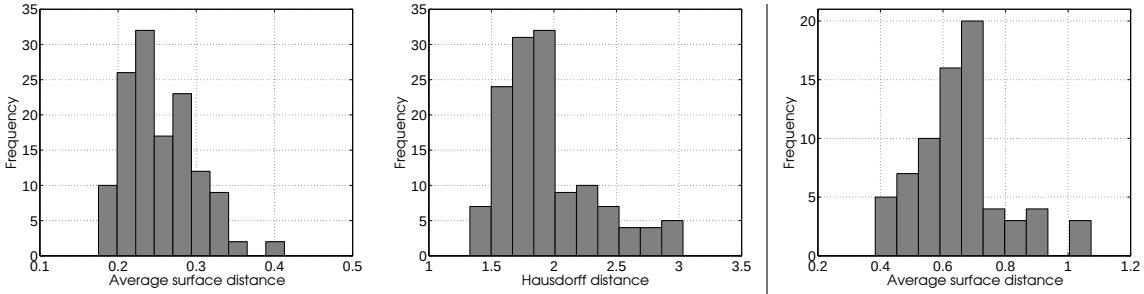


Figure 5.13. LA wall quantitative results: Histogram of (a) average surface distances (ASD) and (b) Hausdorff distance (HD) for the middle region of supershape test examples across the entire test dataset, mean ASD = (0.25 ± 0.14) pixels and mean HD = (1.95 ± 0.38) pixels. (c) Histogram of ASD for the LA wall across the entire dataset with mean ASD = (0.66 ± 0.14) mm.

assessed the effect of the dominance factor by fixing the patch length and smoothness factor to their best settings.

To assess the sensitivity of patch length p, we chose different patch lengths ranging from 5 through 21 in the interval of 4. Figure 5.14a shows the plot of means and standard deviations of ASDs over the entire dataset for different patch lengths. One can observe that the ASD measure does not vary much after a certain patch length ($p > 13$ mm), which suggests that as long as the patch length is sufficient to capture the localized intensity pattern around the left atrium wall, the algorithm produces reliable results.

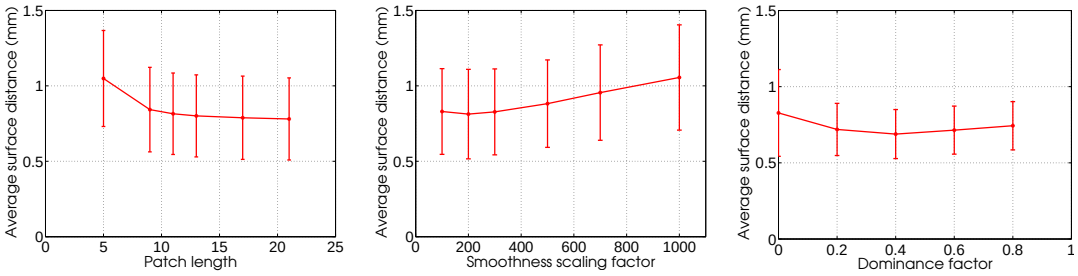
For the smoothness factor, we chose α values ranging from 100 to 1000. Values below 100 were shown to result in nonsmooth shapes [11]. Figure 5.14b shows the corresponding error plot on ASDs where the optimal α value that achieves the best segmentation result is 200. Although the error plot shows a small difference between $\alpha = 100$ versus $\alpha = 200$, we noticed that for a lower α, the segmented surfaces were irregular. As such, we did not intend to experiment with lower α values. Conversely, by increasing the α value beyond 500, the reconstructed surfaces became over-smoothed, thus affecting the accurate delineation of some high curvature left atrial regions.

For the dominance factor assessment, we chose ω values ranging from 0 to 0.8 in the interval of 0.2, where $\omega = 0$ implies no inclusion of the shape prior in the proposed Bayesian segmentation framework. Figure 5.14c shows the corresponding ASD plot that clearly manifests the necessity of the shape prior in acquiring better segmentation results. In particular, better ASD values are acquired for any other ω value as compared to $\omega = 0$. It is also worth noting that the shape prior-based segmentation model gives better performance in terms of providing smaller ASDs compared to models that do not include shape priors (by comparing three plots of Figure 5.14), which suggests the importance of the shape prior in the proposed method.

5.5.6 Algorithm Parameters

Based on the findings of the sensitivity analysis, we chose key parameters for the proposed model.[3] The patch length, $p_{i,j}$, for the intensity model was chosen as 13 mm for the LA dataset and 9 voxels for supershapes. The scaling factor, α, on the smoothness penalty was fixed to 200 for the LA as well as supershapes datasets. The dominance factor, ω, that allows control of the dominance of intensity over the shape prior was chosen to be 0.25 for the LA (in accordance with Figure 5.14c) and 0.4 for the supershapes, respectively.

[3]Similar analysis was conducted on supershapes to choose key parameteric values.



(a) (b) (c)

Figure 5.14. ShapeCut parameters sensitivity analysis: Error plot showing the effect of (a) patch length, (b) smoothness penalty scaling factor α, and (c) dominance factor ω on the segmentation result for the left atrial wall surface.

In the LA case, the ω value was slightly smaller as compared to the supershapes, because as mentioned in Section 5.2, the LA wall is surrounded by similar intensity structures. A smaller ω results in favoring the shape model over the intensity model, thereby biasing the solution for the representative shape, derived from the linear model, rather than letting it be attracted to the incorrect localized boundaries.

The remaining parameters include the standard deviations of the MRF-based offset model, σ_o, σ_c, and σ_c'. Based on our experimentation, the values corresponding to $\sigma_o = 5$, $\sigma_c = 5$, and $\sigma_c' = 1$ give optimal performance for both datasets. However, the results indicate the parameters are not highly sensitive to the choice of parametric values.

5.5.7 Clinical Evaluation

From the clinical perspective, the question would arise whether the proposed automated segmentation could replace an expert one to reduce manpower needed to derive a subsequent analysis. We therefore assessed fibrotic (structural remodeling) regions within the LA wall in preablated and LA wall scarring in the postablated LGE-MRI image volumes. The evaluation was carried out by comparing the fibrotic/scarred regions within the manually and automatically segmented wall regions.

To assess fibrosis in the preablated LGE-MRI scans, we segmented fibrosis regions within each wall based on the image-specific intensity threshold that was provided to us by the CARMA Center in the form of fibrosis percentages. Fibrosis percentage is defined as the percentage of the LA wall that is determined to be "enhanced" in the LGE-MRI image. A threshold-based algorithm [4] has been used to decide the enhanced regions.

For a given image, we used the given fibrosis percentage to determine the intensity threshold retrospectively based on the corresponding LGE-MRI image and its manual wall segmentation. Next, this threshold was used to extract the fibrotic regions in manual as well

as automatic wall segmentations. Figures 5.15 and 5.16 show the anterior-posterior (A-P) and posterior-anterior (P-A) views of fibrosis patterns on three manually and automatically extracted wall surfaces. These patterns clearly demonstrate the potential of the proposed method to correctly identify the fibrosis regions. Further, we analyzed the relationship of fibrosis percentages between manual and automatic wall segmentations. This relationship quantifies the amount of overlap between manual and automatically extracted fibrotic regions. To this end, we plotted the manual versus automatic fibrosis percentage for each preablated scan. Figure 5.17 shows the corresponding scatter plot. The linear relation with a very small error (mean square error, $MSE = 3.07$ and R-square value of 0.83) demonstrates that the manual and automatic fibrotic regions highly overlap.

In order to classify the LA wall scarring in the postablated LGE-MRI scans, we applied the multifeature-based K-means clustering algorithm proposed by [47]. The algorithm uses features including 14 Haralick's texture metrics, image-specific normalized voxel intensity within the left atrium and a Sobel edge map to separate scar regions from the nonscar regions. Figures 5.18 and 5.19 show the anterior-posterior (A-P) and posterior-anterior (P-A) views of scar patterns on three manually and automatically reconstructed wall surfaces. As for fibrosis, we computed scar percentages from the LA wall scarred regions and used them as the evaluation metric to study their relationship within the manually segmented LA walls against the automatic counterparts. Figure 5.20 shows a scatter plot comparing manual to automatic scar percentages for each postablated scan. The nearly linear relation with a small error (mean square error, $MSE = 3.1$ and R-square value of 0.56) between manual and automated scar percentages implies that the proposed method is effective in reproducing scarring results similar to the manual results.

5.5.8 Interrater Variability Analysis

As mentioned in subsection 5.5.1.2, 10 of 72 image samples were provided with the binary segmentations from three experts. To study the variability in segmentations among multiple experts and automatic segmentation, we carried out the interrater variability analysis of the LA wall on these images by using a simultaneous truth and performance-level estimation (STAPLE) algorithm [48]. STAPLE provides a probabilistic estimation model that not only evaluates the performance of a given segmentation with respect to the estimated true segmentation, but also quantifies the interrater variability. STAPLE estimates the true segmentation based on the weighted combination of segmentations being provided, a prior model derived from the spatial distribution of the segmented structures and spatial homogeneity. It computes the sensitivity (true positive fraction) of each class of segmentations

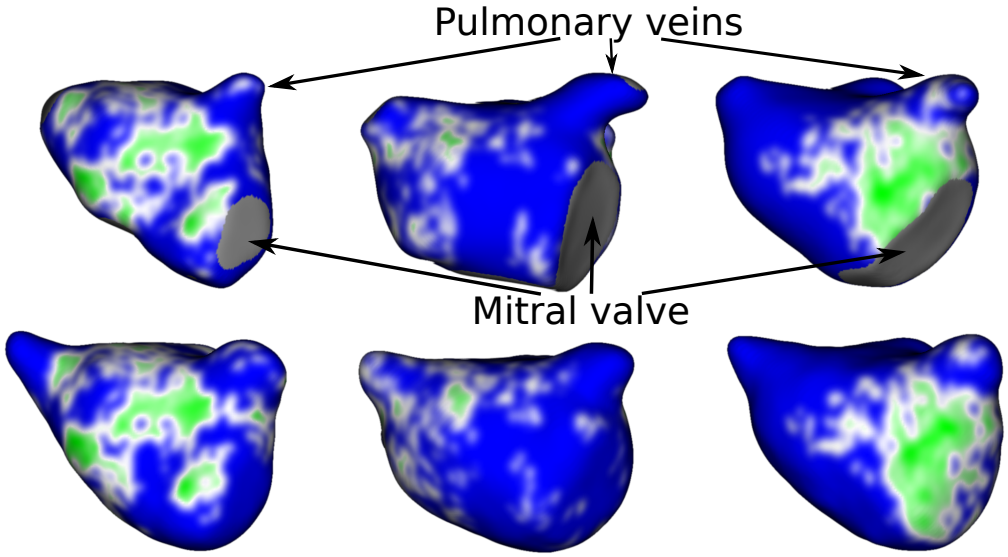


Figure 5.15. Fibrosis-based clinical evaluation (A-P view): Anterior-posterior view of LGE-MRIs depicting fibrosis patterns in manual (*top*) vs. the proposed method (*bottom*) wall regions. Fibrosis regions are displayed in green and healthy wall regions in blue.

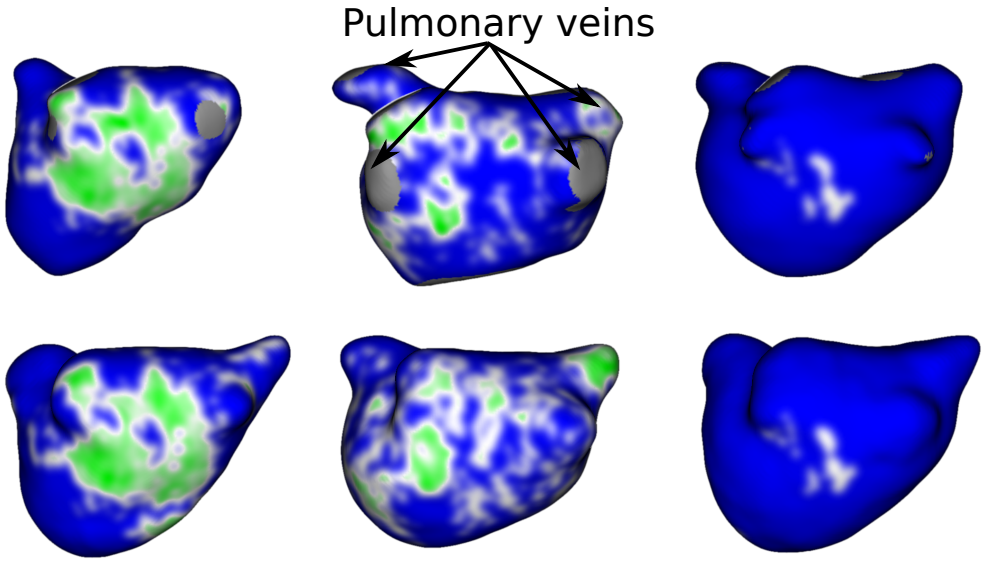


Figure 5.16. Fibrosis-based clinical evaluation (P-A view): Posterior-anterior view of LGE-MRIs depicting fibrosis patterns in manual (*top*) vs. the proposed method (*bottom*) wall regions. Fibrosis regions are displayed in green and healthy wall regions in blue.

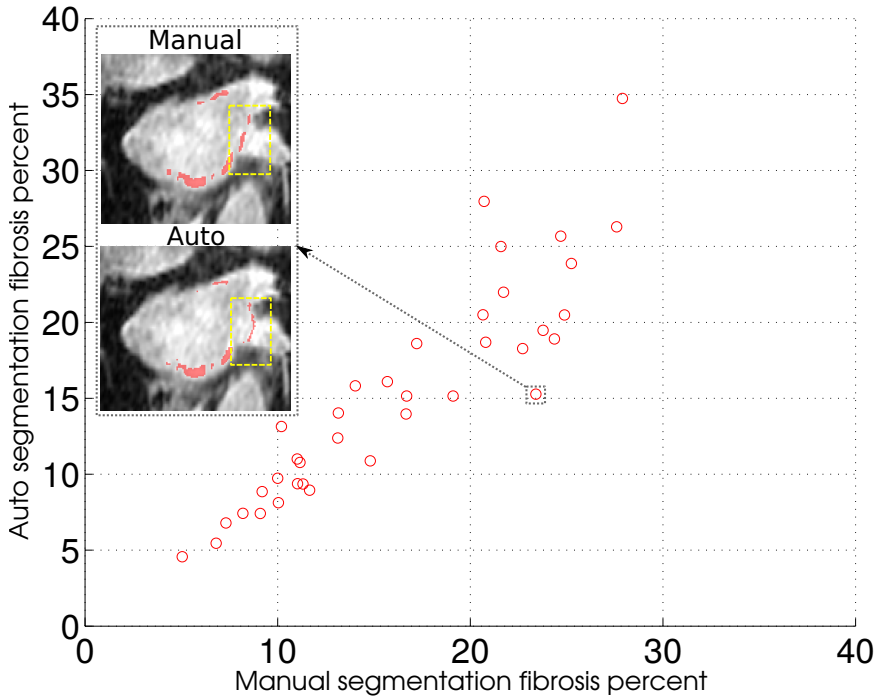


Figure 5.17. Manual vs. automatic fibrosis correlation: Scatter plot depicting the relation between manual vs. automatic fibrosis-wall percentages. An example slice shows fibrosis regions (*red spots*) and highlights the area (*yellow rectangle*) where manual and automatic segmentation disagrees near valves that come out of the left atrium.

as one of its evaluation measures.

Besides the fact that vein structures and mitral valve are not part of the LA, we noticed that they typically suffer from segmentation inconsistencies between raters. We therefore tried to remove vein structures with the help of experts from the CARMA Center. However, we did not attempt to exclude the mitral valve due to the difficulty involved in its extraction and its proximity to the surrounding fibrotic/scar regions.

The bar plot shown in the top panel of Figure 5.21 presents the sensitivity values for three manual segmentations and the automatic counterpart for each image, respectively. These uneven bar plots confirm that there exists inherent variability among raters. In order to determine the specific locations of the variability, we computed pairwise distance errors between automatically extracted surfaces and the surfaces generated from three manual segmentations. The bottom panel of Figure 5.21 shows distance errors on one of the data samples where the blue regions correspond to the low surface distance errors (ASD < 1 mm), and the red regions correspond to the high surface distance errors. As expected, the high surface distance errors are localized only around the mitral valve and veins, which are irrelevant to either fibrosis or scar analysis.

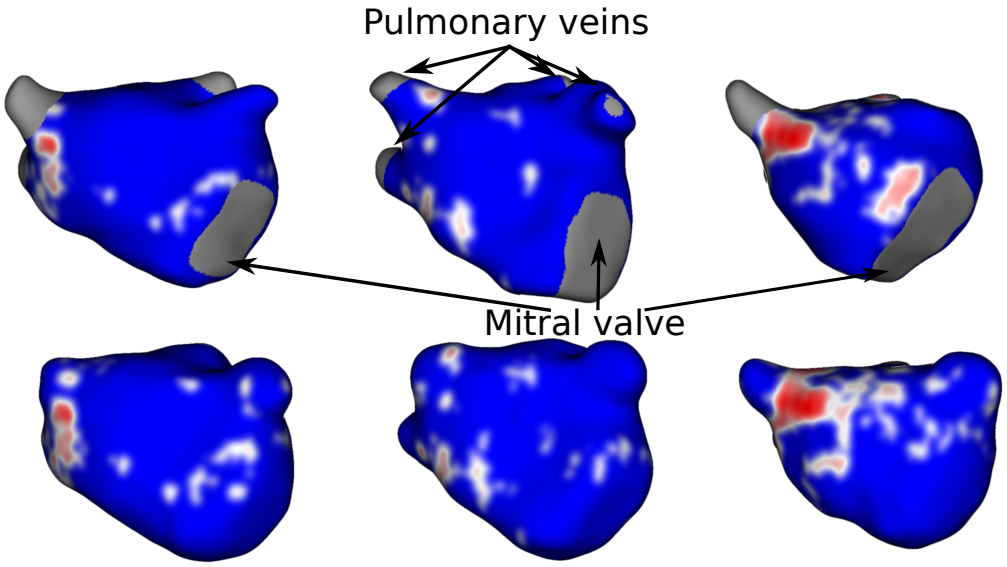


Figure 5.18. Scar-based clinical evaluation (A-P view): Anterior-posterior view of LGE-MRIs depicting scar patterns in manual (*top*) vs. the proposed method (*bottom*) wall regions. Scarred regions are displayed in red and healthy wall regions in blue.

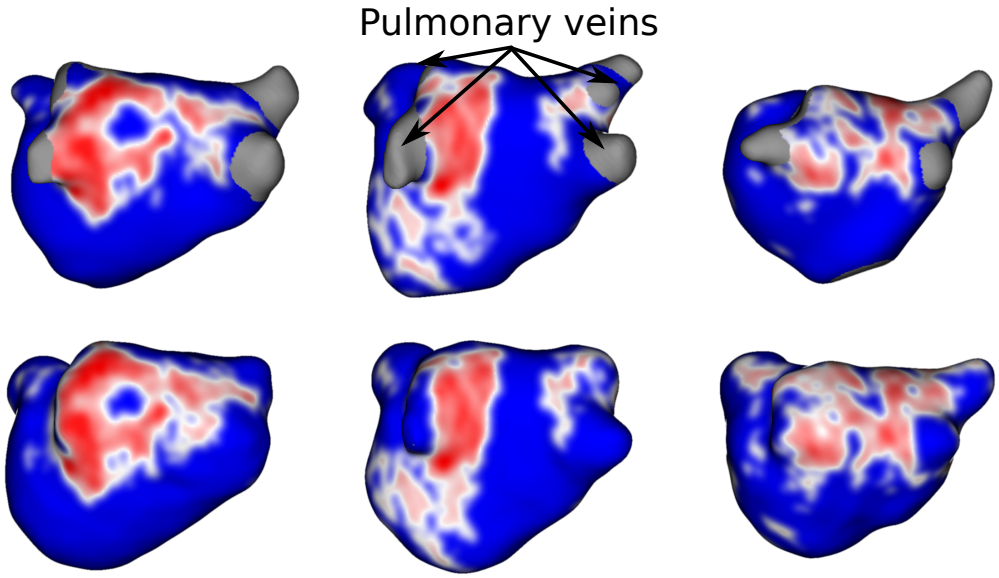


Figure 5.19. Scar-based clinical evaluation (P-A view): Posterior-anterior view of LGE-MRIs depicting scar patterns in manual (*top*) vs. the proposed method (*bottom*) wall regions. Scarred regions are displayed in red and healthy wall regions in blue.

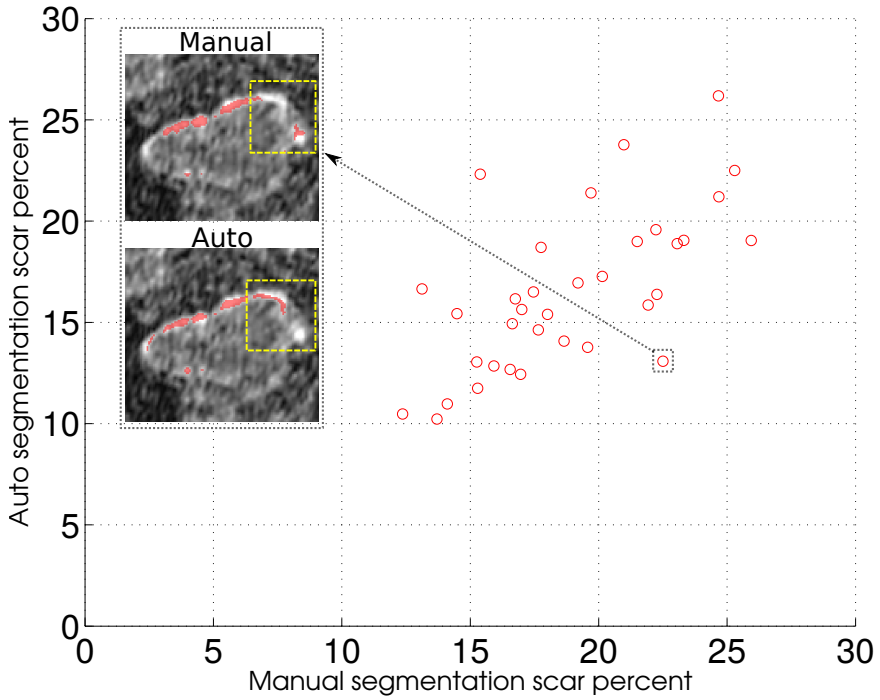


Figure 5.20. Manual vs. automatic fibrosis correlation: Scatter plot depicting the relation between manual vs. automatic scar percentages. An example slice shows scar regions (red spots) and highlights the area (yellow rectangle) where manual and automatic segmentation disagrees near valves that surround the left atrium.

5.6 Discussion and Conclusion

5.6.1 Clinical Impact and Performance

Based on the results presented in Section 5.5, ShapeCut is an effective method for the automatic extraction of multiple coupled surfaces such as the LA wall. Regarding the algorithm's run-time, which involves an iterative optimization of coupled surfaces and shape parameters, we obtained an average run-time of 2.3 minutes for the LA wall extraction. This run-time reflects the nested mesh layer model, which comprises 75 layers and 12665 mesh points per layer. Further, since we deal with two surfaces, these layers and the respective mesh points get doubled, for a total of $75 \times 12665 \times 2$ mesh points. Compared to the manual segmentations that typically take more than an hour, ShapeCut achieves significant time improvement apart from providing similar results, except around vein regions, not only qualitatively and quantitatively but also from a clinical perspective.

5.6.2 Limitations

One of the challenges inherent to multicolumn-based surface estimation methods lies in locating the approximate center of the 3D objects based on which the nested mesh structure

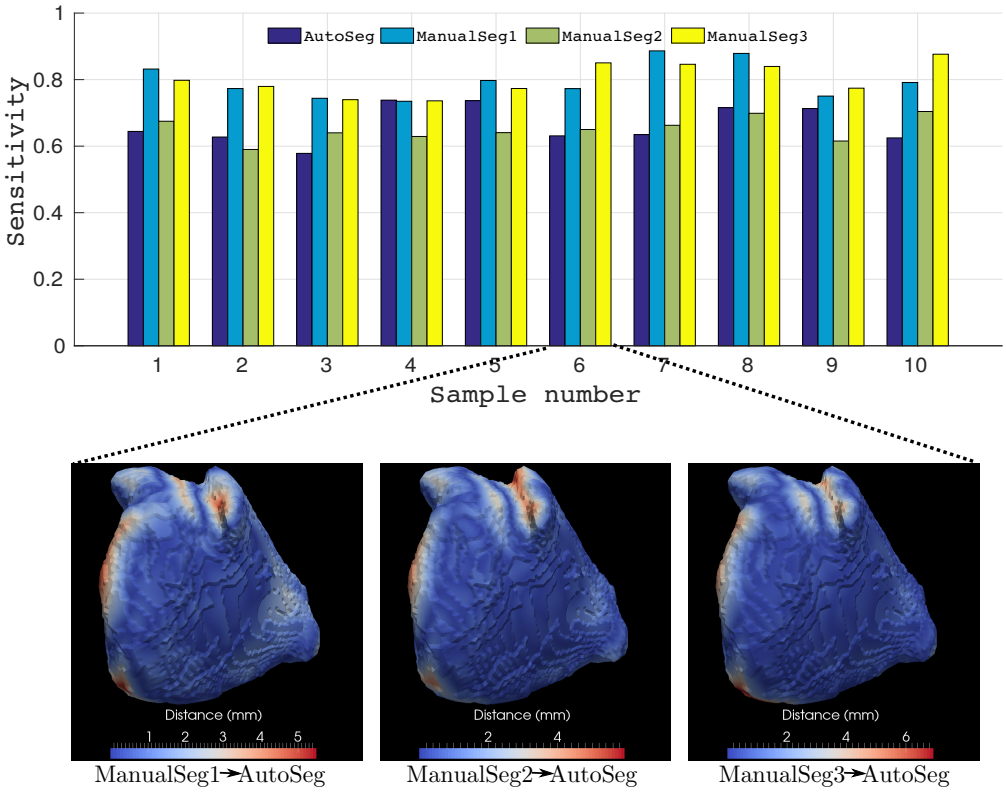


Figure 5.21. STAPLE analysis: (*Top*) Bar plot showing the sensitivity of the automatic segmented wall (*AutoSeg*) compared to the corresponding three manual walls (*ManualSeg*) on 10 samples. (The bar plot justifies the variability among interraters.) (*Bottom*) Pairwise surface distances between one automatically extracted endocardial surface and corresponding surfaces generated from three manual segmentations.

is positioned. Another challenge related to the nested mesh construction lies in knowing the range of the search domain. The search domain corresponds to the number of mesh layers needed to construct the nested mesh model. To achieve correct segmentation results at each point of the surface, this number must be large enough that all the surfaces being delimited fall well within its search space. An ideal way to circumvent this problem is to let the search space grow until its outermost mesh layer reaches the image boundaries and its innermost mesh layer diminishes to a single point. However, this process could increase the run-time of the algorithm. Another viable solution is to learn the surface bounds within the volume from a set of training shapes. This strategy not only helps determine the number of mesh layers, but also assists in positioning the approximate center of the desired objects to be segmented. The approximate center can be estimated by computing the centroid of an average shape that is derived from the training shapes.

5.6.3 Future Work

Based on the knowledge pertaining to the variability of LA shapes, we expect that this shape variation could be captured by using nonlinear shape priors. One way to achieve this nonlinearity is by modeling the shape priors as a mixture of Gaussians where the underlying model includes the linear subspace spanned by each mixture component. In order to estimate the desired surface in a given image, an expectation maximization (EM) algorithm can be used. The E-step estimates the distribution of the mixture component identity, and the M-step estimates the component-specific shape parameters and the surface using a conditional maximization scheme similar to the proposed method. This domain needs to be explored for future development of the underlying model.

Apart from the Gaussian-based generative intensity model, $P(I|\mathcal{S})$, which estimates the likelihood of the surface based on the statistics of the training intensity profiles, one can apply deep convolutional networks (ConvNets) to extract features [49] from these intensity profiles. This extraction can be followed by regression on the extracted features with respect to the test intensity profile in order to derive costs on the graph nodes. Further examination of feature extraction on the basis of ConvNets may result in a better likelihood model.

5.6.4 Conclusion

We present ShapeCut, a shape-based generative model for extracting multiple surfaces from a given image. ShapeCut is modeled by incorporating global shape information within the Bayesian framework, thus biasing the solution toward the desired shape. However, to accommodate a subclass of shapes that are not captured by means of the global shape priors, due to the smaller sample size, the proposed method introduces local shape priors in the form of MRFs. The optimization of the derived model is done in two phases that are performed in an iterative fashion: one for a multicolumn graph-based multisurface update and the other for closed form-based global shape refinement. The results demonstrated the effectiveness of our approach in the presence of weak boundaries, contrast variation, and a low signal-to-noise ratio.

5.7 References

[1] E. J. Benjamin, D. Levy, S. M. Vaziri, R. B. D'Agostino, A. J. Belanger, and P. A. Wolf, "Independent risk factors for atrial fibrillation in a population-based cohort: The Framingham Heart Study," *JAMA*, vol. 271, no. 11, pp. 840–844, 1994.

[2] A. S. Go, E. M. Hylek, K. A. Phillips, Y. Chang, L. E. Henault, J. V. Selby, and D. E. Singer, "Prevalence of diagnosed atrial fibrillation in adults: national implications for rhythm management and stroke prevention: The anticoagulation and risk factors in atrial fibrillation (ATRIA) study," *JAMA*, vol. 285, no. 18, pp. 2370–2375, 2001.

[3] K. Higuchi, M. Akkaya, N. Akoum, and N. F. Marrouche, "Cardiac mri assessment of atrial fibrosis in atrial fibrillation: Implications for diagnosis and therapy," *Heart*, vol. 100, no. 7, pp. 590–596, 2014.

[4] R. S. Oakes, T. J. Badger, E. G. Kholmovski, N. Akoum, N. S. Burgon, E. N. Fish, J. J. Blauer, S. N. Rao, E. V. DiBella, N. M. Segerson *et al.*, "Detection and quantification of left atrial structural remodeling with delayed-enhancement magnetic resonance imaging in patients with atrial fibrillation," *Circulation*, vol. 119, no. 13, pp. 1758–1767, 2009.

[5] N. F. Marrouche, D. Wilber, G. Hindricks, P. Jais, N. Akoum, F. Marchlinski, E. Kholmovski, N. Burgon, N. Hu, L. Mont *et al.*, "Association of atrial tissue fibrosis identified by delayed enhancement MRI and atrial fibrillation catheter ablation: The decaaf study," *JAMA*, vol. 311, no. 5, pp. 498–506, 2014.

[6] S. Y. Ho, J. A. Cabrera, and D. Sanchez-Quintana, "Left atrial anatomy revisited," *Circulation: Arrhythmia and Electrophysiology*, vol. 5, no. 1, pp. 220–228, 2012.

[7] D. C. Peters, J. V. Wylie, T. H. Hauser, K. V. Kissinger, R. M. Botnar, V. Essebag, M. E. Josephson, and W. J. Manning, "Detection of pulmonary vein and left atrial scar after catheter ablation with three-dimensional navigator-gated delayed enhancement MR imaging: Initial experience 1," *Radiology*, vol. 243, no. 3, pp. 690–695, 2007.

[8] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models-their training and application," *Comput. Vision and Image Understanding*, vol. 61, no. 1, pp. 38–59, 1995.

[9] R. Malladi, J. A. Sethian, and B. C. Vemuri, "Shape modeling with front propagation: A level set approach," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 17, no. 2, pp. 158–175, 1995.

[10] R. T. Whitaker, "Volumetric deformable models: Active blobs," in *Visualization Biomedical Comp.* Inter. Society for Optics and Photonics, 1994, pp. 122–134.

[11] G. Veni, Z. Fu, S. P. Awate, and R. T. Whitaker, "Bayesian segmentation of atrium wall using globally-optimal graph cuts on 3d meshes," in *Inf. Proc. in Med. Imag.: 23rd Int. Conf., IPMI 2013, Asilomar, CA, USA, June 28–July 3, 2013. Proc.* Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 656–667.

[12] Y. Y. Boykov and M.-P. Jolly, "Interactive graph cuts for optimal boundary & region

segmentation of objects in N-D images," in *Comp. Vision, 2001. ICCV 2001. Proc. Eighth IEEE Int. Conf. on*, vol. 1. IEEE, 2001, pp. 105–112.

[13] V. Kolmogorov and R. Zabih, "What energy functions can be minimized via graph cuts," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 26, pp. 65–81, 2004.

[14] G. Veni, S. Y. Elhabian, and R. T. Whitaker, "A Bayesian formulation of graph-cut surface estimation with global shape priors," in *Biomed. Imag., 2015 IEEE 12th Int. Symp. on*. IEEE, 2015, pp. 368–371.

[15] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Trans. Patt. Anal. Mach. Intell*, vol. 23, no. 11, pp. 1222–1239, 2001.

[16] M. Sonka, V. Hlavac, and R. Boyle, *Image Processing, Analysis, and Machine Vision*. Boston, MA: Cengage Learning, 2014.

[17] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 23, no. 6, pp. 681–685, 2001.

[18] J. Kim, M. Çetin, and A. S. Willsky, "Nonparametric shape priors for active contour-based image segmentation," *Signal Processing*, vol. 87, no. 12, pp. 3021–3044, 2007.

[19] K. Li, X. Wu, D. Z. Chen, and M. Sonka, "Optimal surface segmentation in volumetric images-a graph-theoretic approach," *IEEE Trans. Patt. Anal. and Mach. Intell.*, vol. 28, no. 1, pp. 119–134, 2006.

[20] J. Malcolm, Y. Rathi, and A. Tannenbaum, "Graph cut segmentation with nonlinear shape priors," in *IEEE Int. Conf. Image Process.*, vol. 4. IEEE, 2007, pp. IV–365.

[21] N. Vu and B. Manjunath, "Shape prior segmentation of multiple objects with graph cuts," in *Comput. Vision and Patt. Recognition, 2008. CVPR 2008. IEEE Conference on*. IEEE, 2008, pp. 1–8.

[22] L. Gorelick, F. R. Schmidt, and Y. Boykov, "Fast trust region for segmentation," in *Comput. Vision and Patt. Recognition (CVPR), 2013 IEEE Conf. on*. IEEE, 2013, pp. 1714–1721.

[23] A. Delong and Y. Boykov, "Globally optimal segmentation of multi-region objects," in *Comput. Vision, 2009 IEEE 12th Int. Conf. on*. IEEE, 2009, pp. 285–292.

[24] F. R. Schmidt and Y. Boykov, "Hausdorff distance constraint for multi-surface segmentation," in *Computer Vision – ECCV 2012: 12th European Conf. on Comput. Vision, Florence, Italy, October 7-13, 2012, Proc., Part I*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 598–611.

[25] M. Rousson and N. Paragios, "Shape priors for level set representations," in *Comput. Vision — ECCV 2002: 7th European Conf. on Comp. Vision Copenhagen, Denmark, May 28–31, 2002 Proc., Part II*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2002, pp. 78–92.

[26] T. Chan and W. Zhu, "Level set based shape prior segmentation," in *Comput. Vision and Patt. Recognition, 2005. CVPR 2005. IEEE Computer Society Conf. on*, vol. 2. IEEE, 2005, pp. 1164–1170.

[27] D. Cremers, S. J. Osher, and S. Soatto, "Kernel density estimation and intrinsic alignment for shape priors in level set segmentation," *Int. J. Comput. Vision*, vol. 69, no. 3, pp. 335–351, 2006.

[28] X. Wu and D. Z. Chen, "Optimal net surface problems with applications," in *Proc. Int. Colloquium on Automata, Languages and Programming*, 2002, pp. 1029–1042.

[29] K. Li, X. Wu, D. Z. Chen, and M. Sonka, "Globally optimal segmentation of interacting surfaces with geometric constraints," in *Comput. Vision and Patt. Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, vol. 1. IEEE, 2004, pp. I–394.

[30] Q. Song, J. Bai, M. Garvin, M. Sonka, J. Buatti, and X. Wu, "Optimal multiple surface segmentation with shape and context priors," *IEEE Trans. Med. Imag.*, vol. 32, no. 2, pp. 376 –386, 2013.

[31] K. Li, S. Millington, X. Wu, D. Z. Chen, and M. Sonka, "Simultaneous segmentation of multiple closed surfaces using optimal graph searching," in *Inf. Proc. Med. Imag.: 19th Int. Conf., IPMI 2005, Glenwood Springs, CO, USA, July 10-15, 2005. Proc.* Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 406–417.

[32] X. Dou, X. Wu, A. Wahle, and M. Sonka, "Globally optimal surface segmentation using regional properties of segmented objects," in *Comput. Vision and Patt. Recognition, 2008. CVPR 2008. IEEE Conference on.* IEEE, 2008, pp. 1–8.

[33] Q. Song, X. Wu, Y. Liu, M. Smith, J. Buatti, and M. Sonka, "Optimal graph search segmentation using arc-weighted graph for simultaneous surface detection of bladder and prostate," in *"Med. Image Comput. and Computer-Assisted Intervention – MICCAI 2009: 12th Int. Conf., London, UK, September 20-24, 2009, Proc., Part II.* Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 827–835.

[34] X. Li, X. Chen, J. Yao, X. Zhang, and J. Tian, "Renal cortex segmentation using optimal surface search with novel graph construction," in *Med. Image Comput. and Computer-Assisted Intervention – MICCAI 2011: 14th Int. Conf., Toronto, Canada, September 18-22, 2011, Proc., Part III.* Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 387–394.

[35] M. E. Tipping and C. M. Bishop, "Probabilistic principal component analysis," *J. Roy. Statist. Soc.: Series B (Statist. Methodology)*, vol. 61, no. 3, pp. 611–622, 1999.

[36] M. E. Leventon, W. E. L. Grimson, and O. Faugeras, "Statistical shape influence in geodesic active contours," in *CVPR*, vol. 1. IEEE, 2000, pp. 316–323.

[37] M. Jenkinson and S. Smith, "A global optimisation method for robust affine registration of brain images," *Med. Image Anal.*, vol. 5, no. 2, pp. 143–156, 2001.

[38] S. Valette, J.-M. Chassery, and R. Prost, "Generic remeshing of 3D triangular meshes with metric-dependent discrete voronoi diagrams," *IEEE Trans. Vis. and Comput. Graphics*, vol. 14, no. 2, pp. 369–381, 2008.

[39] C. Bauer, S. Sun, and R. Beichel, "Avoiding mesh folding in 3D optimal surface segmentation," in *Advances in Visual Comput.: 7th Int. Symp., ISVC 2011, Las Vegas, NV, USA, September 26-28, 2011. Proc., Part I.* Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 214–223.

[40] H. Ishikawa, "Exact optimization for Markov random fields with convex priors," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 25, no. 10, pp. 1333–1336, 2003.

[41] J. Gielis, "A generic geometric transformation that unifies a wide range of natural and abstract shapes," *Amer. J. Botony*, vol. 90, no. 3, pp. 333–338, 2003.

[42] Y. Fougerolle, A. Gribok, S. Foufou, F. Truchetet, and M. Abidi, "Rational supershapes for surface reconstruction," in *Eighth Int. Conf. on Quality Control by Artificial Vision.* International Society for Optics and Photonics, 2007, pp. 63 560M–63 560M.

[43] D. N. Politis and J. P. Romano, "Large sample confidence regions based on subsamples under minimal assumptions," *Ann. Stat.*, pp. 2031–2050, 1994.

[44] Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision," *PAMI*, vol. 26, pp. 359–374, 2001.

[45] R. Bensch and O. Ronneberger, "Cell segmentation and tracking in phase contrast images using graph cut with asymmetric boundary costs," in *Biomed. Imag., 2015 IEEE 12th Int. Symp. on.* IEEE, 2015, pp. 1220–1223.

[46] A. A. Taha and A. Hanbury, "Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool," *BMC Med. Imag.*, vol. 15, no. 1, p. 29, 2015.

[47] D. Perry, A. Morris, N. Burgon, C. McGann, R. MacLeod, and J. Cates, "Automatic classification of scar tissue in late gadolinium enhancement cardiac MRI for the assessment of left-atrial wall injury after radiofrequency ablation," in *SPIE Med. Imag.* International Society for Optics and Photonics, 2012, pp. 83 151D–83 151D.

[48] S. K. Warfield, K. H. Zou, and W. M. Wells, "Simultaneous truth and performance level estimation (STAPLE): An algorithm for the validation of image segmentation," *IEEE Tran. Med. Imag.*, vol. 23, no. 7, pp. 903–921, 2004.

[49] T. Pfister, K. Simonyan, J. Charles, and A. Zisserman, "Deep convolutional neural networks for efficient pose estimation in gesture videos," in *Comput. Vision – ACCV 2014: 12th Asian Conf.on Comput. Vision, Singapore, Singapore, November 1-5, 2014, Revised Selected Papers, Part I.* Springer International Publishing, 2014, pp. 538–552.

CHAPTER 6

SURFACE EXTRACTION APPLICATIONS AND ALGORITHMS

This chapter discusses unpublished work related to the application of the MRF-based Bayesian surface extraction algorithm to segment layers of a contact lens in OCT images and horizon segmentation in seismic volumes. Section 6.1 deals with the procedural details and evaluation of lens segmentation, and Section 6.2 covers the algorithm's functionality in the extraction of horizons. This section also explains other graph cuts-based optimization algorithms to efficiently minimize the surface-based energies, which are extremely useful in the case of large-image volumes.

6.1 Contact Lens Segmentation in OCT Imaging

The segmentation of a contact lens entails the delineation of its inner and outer boundaries. Fig. 6.1 shows an OCT image example with the lens boundaries delineated. This delineation is carried out on various types of contact lenses (Type I - Type VII) characterized by different curvatures and thicknesses, thereby determining the optical properties of these lenses. As discussed in Chapter 1, OCT images suffer low SNR and imaging artifacts, including scan lines, vertical stray lines, and the glass plate, making lens segmentation challenging. To accurately delineate the inner and outer boundaries of these lenses in this challenging environment, we adopted the Bayesian surface extraction strategy.

6.1.1 Algorithm Components

Before employing the Bayesian surface extraction algorithm, we perform preprocessing of a given image in order to remove irrelevant information that corresponds to the imaging artifacts. This step is followed by building a parametric structure that complies with the lens boundaries. Building such a structure necessarily involves two steps. In the first step, a 2D curve is constructed that roughly aligns with the curvature of any lens type. The curve provides a basic structure of the multicolumn graph. Hence, it is referred to as the base curve (similar to the *base mesh* in 3D). The second step involves generation of nested

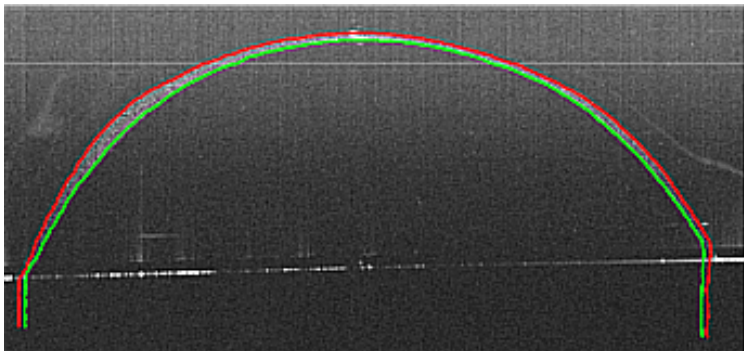


Figure 6.1. A sample OCT image showing delineation of the inner and outer boundaries of a contact lens.

curves that maintain a geometry similar to that of the base curve. These curves form a multicolumn graph within which the segmentation takes place. To formulate surface-based energy in the multicolumn graph, the nodes of the graph carry costs that correspond to the likelihood of features with respect to the lens boundaries, and the edges are used to impose constraints on the lens boundaries, such as smoothness and their interactions. Thus, the optimal lens boundaries are obtained by finding a *min-cut* in this special, geometric multicolumn graph. Details of the algorithm components are discussed in the following subsections.

6.1.2 Preprocessing

As shown in Fig. 6.1, the boundaries of the lens are surrounded by imaging artifacts, including unwanted scan lines, vertical stray lines, and the glass plate on which the lens rests. These unwanted structures affect the segmentation process and need to be removed by preprocessing the image. The preprocessing step also helps improve the lens model building process by retrieving only a collection of points that correspond to the lens. To carry out this process, we deal with the images derived from the variance of their intensities instead of dealing directly with the image intensities. Variance-based images are less susceptible to random noise and spatial inhomogeneity. Furthermore, a global threshold can be used to extract edge-related pixels in these kinds of images [1], [2].

Given an image, the variance-based image is generated by considering a predefined-sized window around each pixel, and computing the variance of pixel intensities within the window. The variance within each window is computed as:

$$Var(X) = \frac{\sum_{i=1}^{N}(x_i - \bar{x})}{N-1} \tag{6.1}$$

where $\bar{x}$ represents the mean of intensities and N represents the number of pixels. The size

of the window is empirically chosen as 21×21 for all images. Next, the following steps are used to remove the imaging artifacts from the variance-based image.

- Use a threshold on the variance-based image to acquire all the edge-related pixels, which gives a binary image. This threshold is empirically chosen as 15 for all examples.
- Apply the Hough transform [3] on this binary image to detect all straight lines. Fig. 6.2 shows various lines on the binarized variance image that are detected using the Hough transform.
- Extract lines that traverse from the first column of an image to its last column. These lines correspond to the scan lines and the glass plate, which are removed.
- Apply an annulus mask surrounding the lens structure. This mask removes pixels corresponding to the vertical stray lines that are present away from the lens, which eventually helps improve the lens-building model.

The preprocessing step provides a binary image with unwanted structures removed. Fig. 6.3 shows the preprocessing output (in blue) on two example slices. It can be seen that most of the points correspond to the lens structure, which is necessary for the next component of the segmentation process.

6.1.3 Building a Lens-Following Nested Curve

To build a lens-following topological model, we generate a set of nested layers that conform with the shape of the lens. Similar to Chapter 2 through Chapter 5, we rely on the training strategy, where a set of sample points is collected from the preprocessed binary images that correspond to the boundaries of the lens. Initially, a base layer is constructed over these sample points by using the least squares-based elliptical model system [4]. The primary reason for using this system is its ellipse-specific minimizing constraint while solving for a least squares-based optimization problem for a given set of sample points. This constraint specifically complies with the lens shape. Further, it is robust to noise. Once the base layer is created, a set of nested ellipses is generated by using an elliptical coordinate system [5]. By varying the values of elliptical coordinates, different-sized ellipses can be generated that are then used to construct multiple layers, both inside and outside the base layer. Fig. 6.4 shows the grid that is formed using this elliptical coordinate system. This grid provides the multicolumn graph structure.

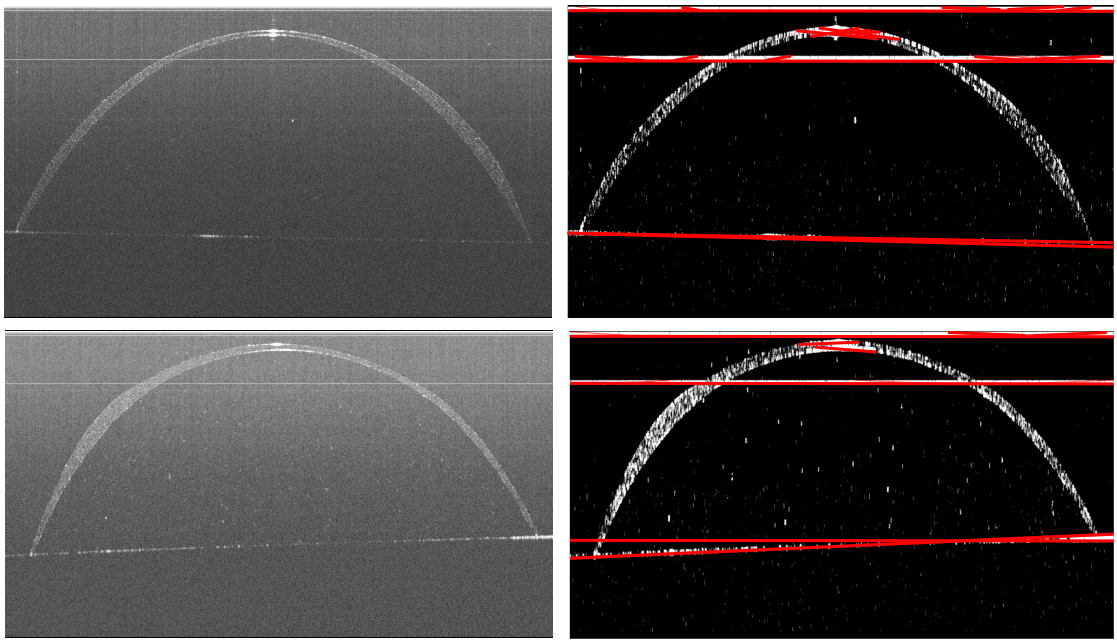


Figure 6.2. Hough lines (shown in red) on thresholded variance-based lens images.

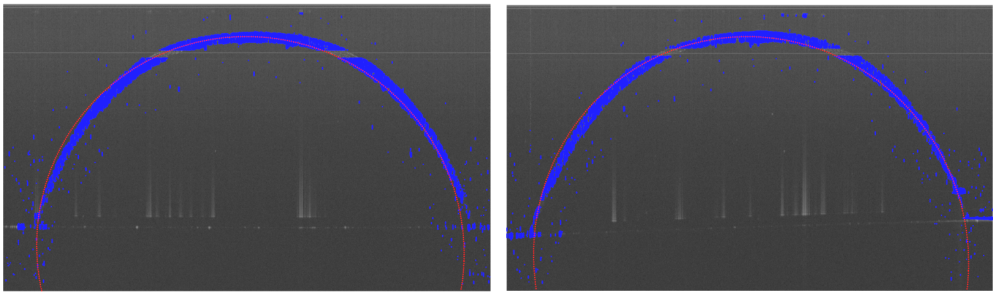


Figure 6.3. Outputs of the preprocessing stage (blue) overlaid on OCT examples. Notice that the preprocessing stage avoids pixels corresponding to the scan lines, glass plate, and vertical stray lines, which helps in building a better lens model whose base layer is shown in red.

6.1.4 Cost Assignments

Due to the variability in the shape of the lens, we again rely on the training strategy to learn intensity-related features and compute node costs. This strategy provides examples of patch profiles that define features along the inner and outer boundaries of the lens. A *patch profile* is a rectangular region of intensities, oriented in the normal direction to the surface/curve associated with each node. In the training phase, a set of preprocessed binary images is utilized to collect points belonging to the lens boundaries and model column-specific patch profiles. The intensity at each model patch location is computed by a distance-based weighted average of patch intensities for each boundary point in each

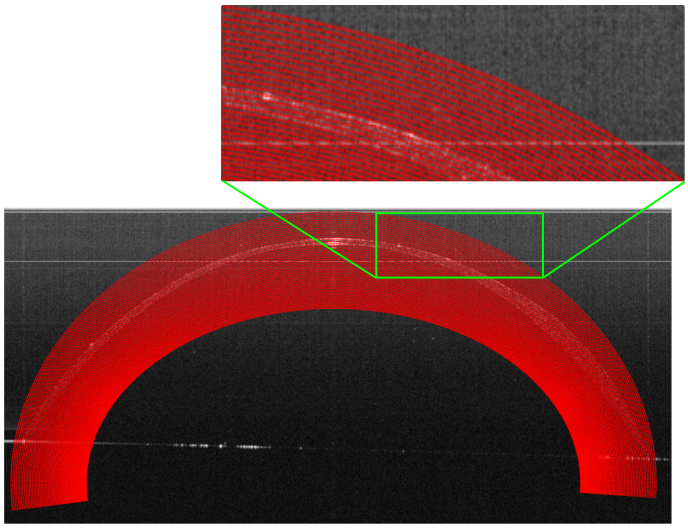


Figure 6.4. Elliptical nested-layer grid overlaid on a test example.

training image. Thus, for a column, a patch profile is a Gaussian-weighted average of several patch profiles from many images of different lens types. In the test phase, the patch profiles are acquired by sampling intensity patches around each node of the elliptical-shaped nested curve. Next, these patches are compared to the corresponding column-specific model patches to determine node costs.

The edges between adjacent columns of the multicolumn graph are configured in a properly ordered fashion. As described in Subsection 1.2.3, this configuration is crucial to formulate the surface-based energy minimization problem as a VCE-weight net surface problem. Further, we enforce smoothness constraints, while assigning costs on every PO edge-pair between adjacent columns. These constraints help achieve regular boundaries—a natural characteristic possessed by lens surfaces.

For a simultaneous segmentation of inner and outer lens boundaries, two disjoint PO graphs are constructed. These graphs are connected by an additional set of edges between corresponding columns that also follow the PO configuration with submodular constraints. These constraints maintain a regularized thickness between the inner and outer lens boundaries.

6.1.5 Lens Segmentation Using Graph Cuts

Once the geometric graph is constructed and the costs are assigned to its nodes and edges, it can be treated as a VCE-weight net surface problem. As described in [6], the optimal net surface corresponds to the optimal *s-excess* set in the derived graph, which is estimated by transforming the VCE-weight net surface problem into a graph cut problem.

Thus, the min-cut in the transformed graph eventually extracts two surfaces that correspond to the inner and outer boundaries of the lens.

6.1.6 Results and Discussion

The evaluation of the MRF-based Bayesian surface extraction approach has been carried out on 614 OCT examples, including six lens types. Table 6.1 summarizes the algorithm's success rate for a group of images pertaining to each lens type. Here, the term "*success*" is defined qualitatively based on a visual inspection such that both lens boundaries are correctly extracted from one end to the other at every point. Figs. 6.5 and 6.6 show the inner and outer boundaries of various lenses on the corresponding OCT images.

According to Table 6.1, the proposed algorithm produces successful results for almost all OCT examples. In a small number of cases, the algorithm extracts incorrect boundaries due to the vertical stray lines and very strong imaging artifacts, which do not get removed during the preprocessing step due to their proximity to the lens and which exhibit stronger intensity-based features compared to the lens, thus misleading the algorithm in choosing the correct boundaries. Such challenges can be handled by enhancing the likelihood model with better features, other than intensities, corresponding to the lens.

6.2 Horizon Segmentation in Seismic Volumes

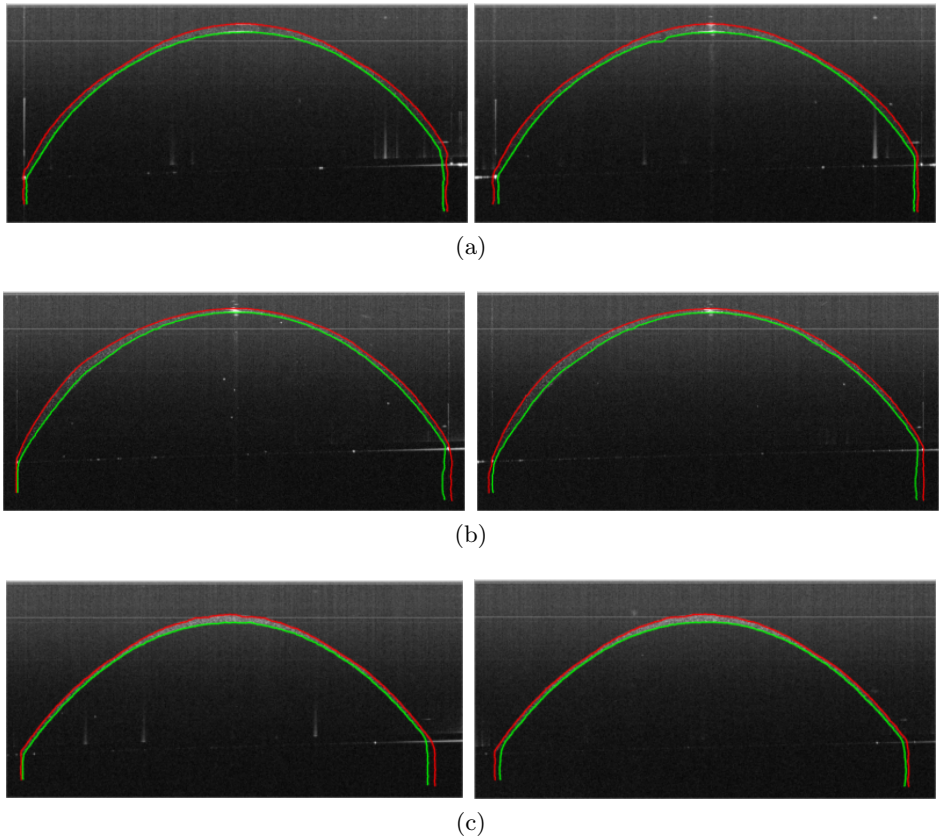
This section is related to the third area of application, which deals with the extraction of horizons in the seismic volumes. The geometric interpretation of the extracted horizons results in the detection of other important seismic structures such as faults, channels, and gas leakages [7]. Fig. 6.7 shows an example slice of the seismic volume comprising horizons and faults. As seen in the figure, the desired horizons suffer from heterogeneous intensities along their structures and are surrounded by other horizons that possess similar features. Based on our knowledge that the MRF-based Bayesian surface extraction algorithm is beneficial in addressing these challenges, it is used in the extraction of horizons.

To extract the desired horizon, a user places a set of landmarks, commonly referred to as *anchor points*, corresponding to the desired horizon. Using these landmarks, a surface-based energy is formulated by encoding its likelihood based on the features of the desired horizon and the MRF priors to enforce smoothness on the horizon. This energy formulation is executed on a properly ordered multicolumn graph structure. Finally, graph cuts are applied to minimize the horizon-specific Bayesian surface energy.

In order to build a multicolumn graph, we let each column of a given image correspond to the column of the graph and each image-row as its layer. In the case of 3D, these layers

Table 6.1. Summary of results for each type of lens

Lens type	Number of examples	Number of successful results
Type I :	108	108
Type II :	72	72
Type III :	92	92
Type IV :	89	89
Type V :	91	90
Type VI :	48	48
Type VII :	114	112



(a)

(b)

(c)

Figure 6.5. Segmentation results on (a) Type I, (b) Type II, and (c) Type III lens images.

are formed on the basis of rows and slices of the image. To determine the likelihood energy that quantifies costs on the nodes of the graph, intensity profile-based features are used, similar to Chapter 2. This process involves two phases. In the training phase, intensity profiles (*sticks*) that serve as the model profiles per column are computed. These model profiles are computed by initially sampling an array of intensities, both above and below the given landmarks. Next, the model profiles for the remaining columns are computed

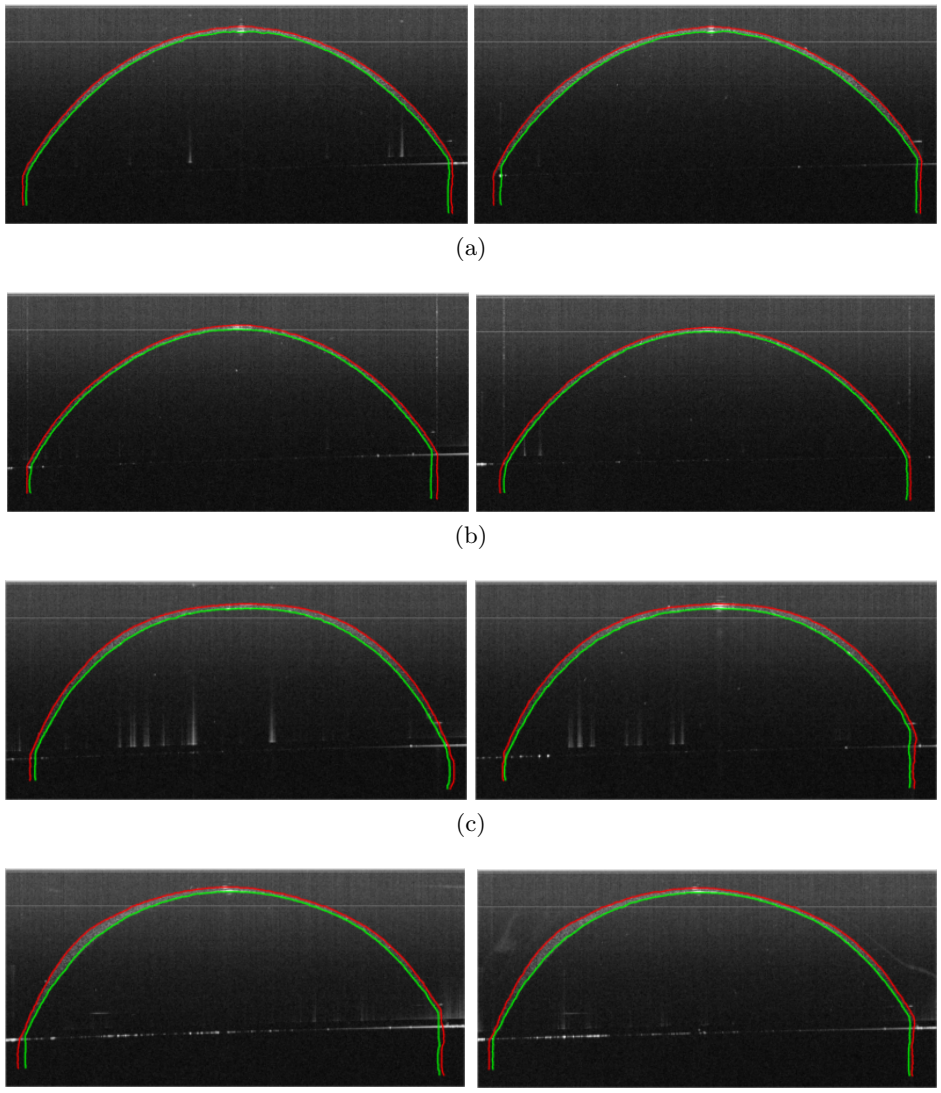


Figure 6.6. Segmentation results on (a) Type IV, (b) Type V, (c) Type VI, and (d) Type VII lens images.

based on the linear interpolation of the landmark-based model profiles. In the test phase, the intensity profiles are sampled around each point of the image and compared against the model profiles of the corresponding columns to determine node costs. The MRF-based smoothness prior is modeled by enforcing soft penalties among properly ordered edge-pairs between adjacent columns, which helps maintain the regularity on horizon surfaces.

To gain run-time and memory efficiency of the graph cut algorithm, we consider only a subset of rows, instead of all rows in a given image, thereby limiting the search space for finding the min-cut. The image formed by downsizing the search space is referred to as the *narrow band*, which is determined based on the user-specified landmarks and the length

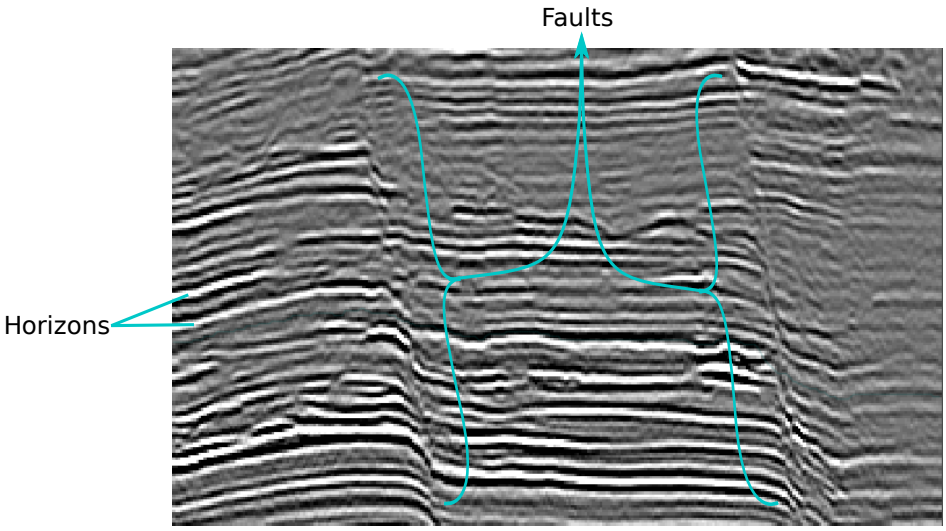


Figure 6.7. 2D slice of the seismic volume depicting horizons and faults.

of the stick. Specifically, given a stick length, the size of the narrow band is calculated by multiplying the stick length with an odd integer factor in order to accommodate equal numbers of rows above and below the stick length.

Fig. 6.8 shows the graphical user interface (GUI) that is specifically designed to extract 2D and 3D horizon surfaces based on the user-defined anchor points. The front-end module of the design has been developed with the help of Kristen Zygmunt and Ayla Khan, software developers at the University of the Utah.

Although multicolumn-based graph cuts yield globally optimal surfaces in a low-order polynomial time ($T(n, m)$ with $O(n)$ vertices and $O(m)$ edges), these algorithms are still impractical in this particular area of application due to the very large image volumes of size around $1000 \times 1000 \times 1000$. The graphs pertaining to these volumes typically carry billions of nodes and e edges per node (e = number of edges in the edge interval, which typically corresponds to the size of the column), and the serial Boykov Kolmogorov (BK)-algorithm (see Subsection 1.2.2) takes up to several hours to find the maximum flow. Further, the unstructured implementations often fail to take advantage of the special graph structure leading to a significant memory requirement in the order of a few tens of gigabytes. The following sections discuss various strategies to improve the performance of graph cuts in terms of run-time and memory consumption. Initially, we discuss a parallel version of graph cuts based on the push-relabel strategy. This work is done in collaboration with Shridharan Chandramouli, whose tasks involved implementation and analysis of the proposed approaches. Next, a label-based approximation approach is presented that uses graph cuts iteratively to minimize the surface-based energy model.

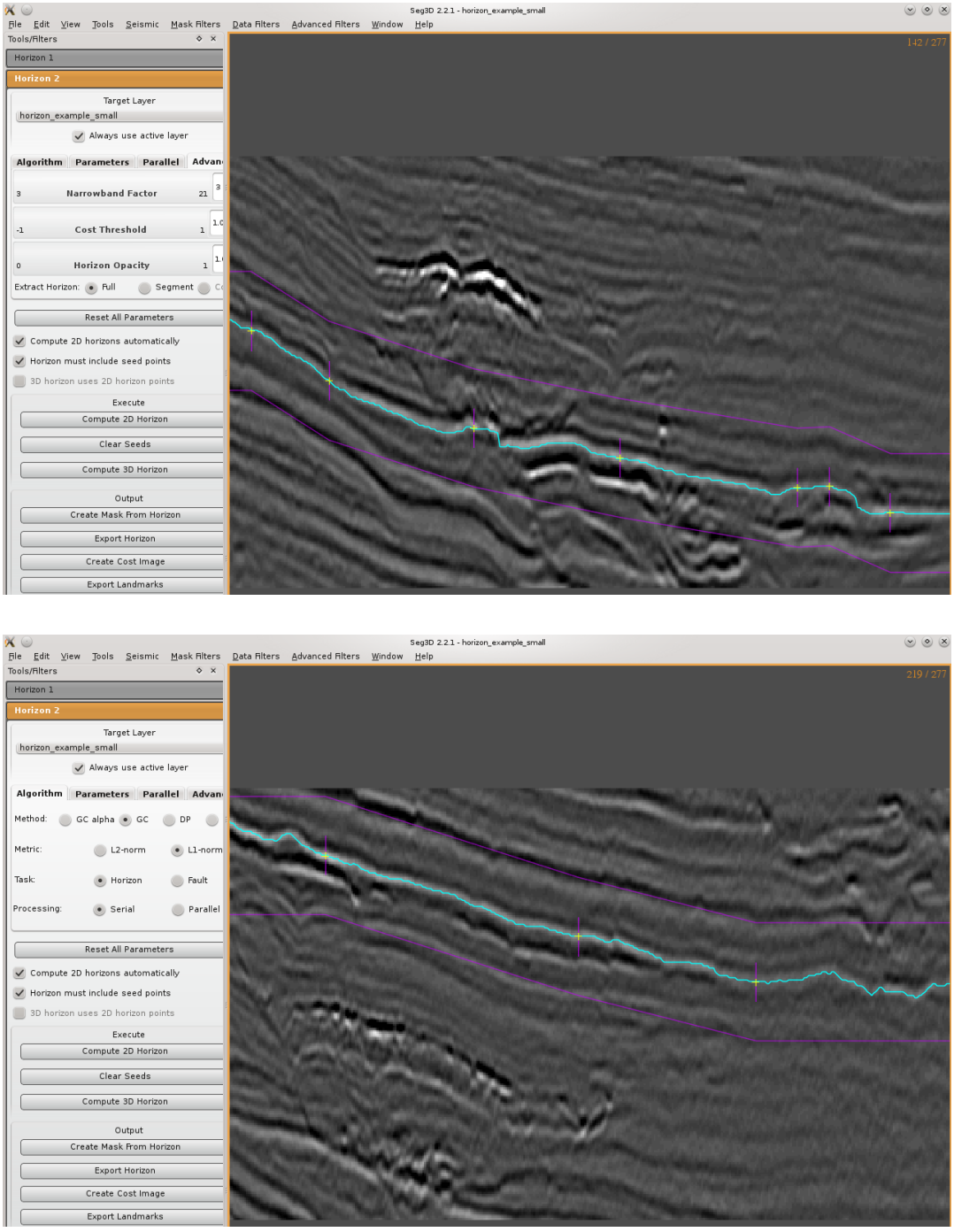


Figure 6.8. Horizon segmentation GUI design with two horizons extracted from different slices of a 3D geological volume. Yellow "+" symbols indicate anchor points and purple vertical lines represent anchor points-specific model profiles. The purple band represents the boundaries of the search space for the horizon extraction.

6.2.1 Parallelization of a Structured Graph

We begin by exploring various maximum flow strategies that have been used to parallelize graph cuts by taking advantage of today's advanced multicore machine architecture, graphics processor units (GPUs), and symmetric multiprocessors (SMPs).

6.2.1.1 Review of Parallel Max-Flow Algorithms

In order to parallelize graph cuts, push-relabel-based max-flow methods are typically preferred over the augmented path-based methods because of their local push and relabel operations on specific nodes, making the graph cuts capable of a parallel implementation. Goldberg and Tarjan [8] proposed the first parallel implementation for their push-relabel method by relying on the dynamic tree data structure [9]. The algorithm runs in $O(n^2 logn)$ time with n processors and $O(m)$ additional space on a moderate-sized general graph structure. The first practical implementation of the push-relabel algorithm was that of Anderson and Setubal [10] for shared-memory architectures. They introduced the concept of *waves* in order to perform global relabeling of vertices concurrently with the push and relabel operations. Here, global relabeling refers to the correct relabeling of vertices based on their exact distances from the sink in the residual graph [8]. A breadth-first search (BFS) algorithm is used to carry out the global relabeling task. Since the proposed approach uses this strategy to apply the push-relabel method on the structured graph setting, we provide further algorithmic details for this strategy.

In shared memory architectures, it is possible that multiple processors operate on the overlapping vertices, causing invalid labelings and thus limiting any advantage of incorporating global relabeling scheme while parallelizing graph cuts. To overcome this problem, Anderson and Setubal [10] use the notion of waves to periodically apply global relabeling after a specific number of discharge operations. Here, each node is affiliated with a wave number that gets updated based on the number of times the node has been globally relabeled. Further, each processor maintains a global shared queue and two local queues: an in-queue and an out-queue. Each time a processor needs to process vertices, it acquires vertices from the global shared queue and stores them in the local in-queue, where they are processed in the first-in, first-out (FIFO) fashion. After discharge operation, newly active vertices are placed in the local out-queue and are moved to the global shared queue only when the out-queue gets filled. Fig. 6.9 shows a schematic corresponding to this vertex processing operation. This strategy reduces the amount of false-sharing between shared memory systems, thereby improving the performance of the algorithm. All the push and relabel operations require locks to access any of these queues.

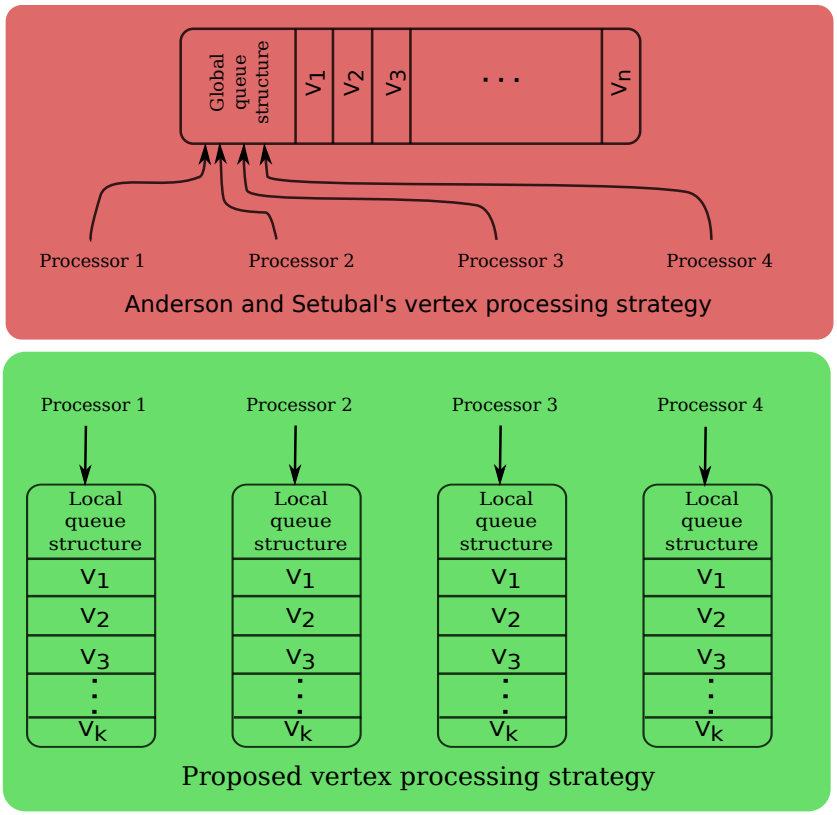


Figure 6.9. Anderson and Setubal's vertex processing strategy [10] and the proposed strategy. Each processor has a local processing queue instead of a shared global FIFO queue.

Vineet and Narayanan [11] implemented the first parallel version of the push-relabel algorithm for the GPUs. While the performance can be significantly improved by using GPUs over the shared memory-based graph cuts parameterization, these systems are bounded by the limited global memory sizes and are impractical in the case of larger graphs.

Generally, augmented path-based methods are not well suited for a parallel implementation of graph cuts due to their nonlocal operations while finding augmented paths to send the flow from source to sink. Liu and Sun [12] proposed a parallel version of the BK algorithm on a structured graph by splitting it into uniform segments and merging them adaptively after processing these segments in parallel. Since their graph structure is closely related to the graph being built for the image-based rectangular volumes, we implemented their parallel strategy on shared memory architecture.

In the following subsection, we will cover the algorithmic details of Liu and Sun's parallel version [12] of the BK-algorithm, Anderson and Setubal's parallel implementation of the push-relabel maximum flow algorithm, and the adaptations made to fit our application at hand.

6.2.1.2 Parallel BK Algorithm on Structured Graphs

We adopted Liu and Sun's adaptive hierarchical approach [12] by dividing the graph into small, disjoint, uniformly sized segments. In view of the multicolumn properly ordered graph structure that restricts only one cut per row, and the use of a narrow band that limits the number of rows to be considerably smaller than the number of columns or slices, the volume is divided across its columns and slices as shown in Fig. 6.10. Once the graph is partitioned, the algorithm simultaneously finds all the short-range augmented paths within the smaller segments. Next, these segments are adaptively merged while expanding the search span and augmenting the flow on longer paths. The process continues until all the paths are exhausted, thus reaching the maximum flow.

The advantage of this approach is memory locality, which improves cache-friendliness within smaller segments. However, we noticed only a threefold run-time improvement over the serial BK algorithm on a $551 \times 426 \times 426$-sized seismic volume. The rationale behind this meager improvement is that there exist more larger augmented paths than shorter paths, leaving a significant portion of the work being performed in the later stages of the max-flow process, thus limiting the run-time efficiency.

6.2.1.3 Parallel Push-Relabel Algorithm on Structured Graphs

To parallelize the push-relabel algorithm on a multicolumn properly ordered graph, we extended the idea of Anderson and Setubal [10] by modifying the concurrent global relabeling strategy with the level-synchronized global relabeling to improve the run-time

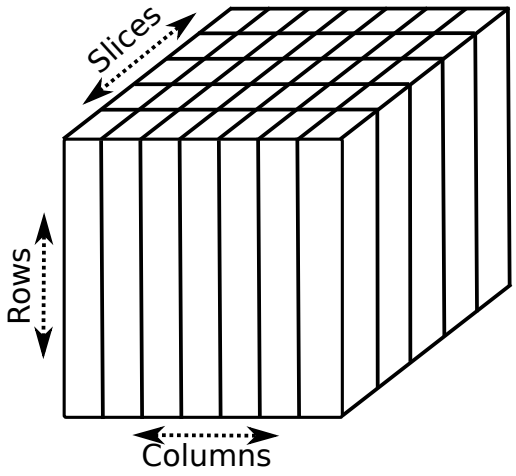


Figure 6.10. Schematic showing the partitioning of a structured graph into uniform segments across its columns and slices. These segments are adaptively merged to find longer augmented paths until the max-flow is reached.

performance of the algorithm. Further, we used an implicit addressing scheme to improve the algorithm's memory consumption.

Anderson and Setubal [10] use a queue or bucket structure, which involves transfer of nodes (vertices) between a local queue/bucket, managed by each processor, and a global queue, common to all processors. Since the nodes do not have any processor affinity structure, their algorithm requires frequent locking of vertices to enable push and relabel operations, thereby affecting the efficiency of the algorithm. We address this problem by taking advantage of the special graph structure inherent to the seismic volume. Based on the number of processors, we divide the graph into segments of equal size. Each segment is assigned to a separate processor, and, thus, the vertex affinity of a segment is set to that processor. At any given time, the nodes that are interior to the segment are processed by the same processor and, therefore, do not require a locking mechanism. However, there can be a transfer of excess flow between vertices of adjacent columns that belong to different segments. These vertices, which are referred to as *shared vertices*, are the only vertices guarded by a lock, as shown in Fig. 6.11. Using this strategy, two aspects become clear. First, vertex affinity replaces global queue structure with one local queue per processor. Second, any vertex addition to the queue is done either locally or from the columns corresponding to the immediately adjacent segments. This strategy provides the advantage of an improved vertex locking mechanism by limiting the communication between graph segments. Fig. 6.9 shows the proposed queue structure.

6.2.1.4 Level-Synchronized Global Relabeling

Similar to [10], we use the concept of waves to perform push and relabel operations. However, the concurrent global relabeling wave propagation is performed in a level-synchronized fashion. The algorithm works as follows. After a certain number of discharge operations, a new global relabeling wave starts (to effectively update the labeling of vertices to reflect

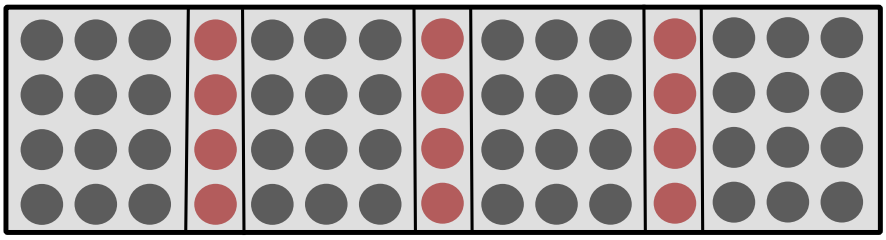


Figure 6.11. Vertex locking mechanism that illustrates shared vertices (red) between adjacent segments. Only these vertices need lock before processing. The interior vertices (gray) need lock only while interacting with the shared vertices.

their original distance to the sink), where every processor adds all the vertices, which are directly connected to the sink via nonsaturated edges, to the queue. This operation advances the frontier by one level. Once all the processors complete advancing the frontier, a software barrier is used to ensure all the operations are synchronized at the same depth before advancing to the next level. Since the frontier advancement is synchronized at every level, this scheme is referred to as the level-synchronized global relabeling (LSGR) scheme. Fig. 6.12 illustrates the frontier advancement mechanism using the LSGR scheme. A software barrier is required at each level of the BFS process; otherwise, there is the possibility that a shared vertex can be reached by the same wave at different levels from adjacent segments, resulting in the wrong label assignments to the node. The level-synchronized global relabeling strategy reduces the contention for queue and, thus, improves the cache coherency and false sharing of the push-relabel algorithm on a shared memory system.

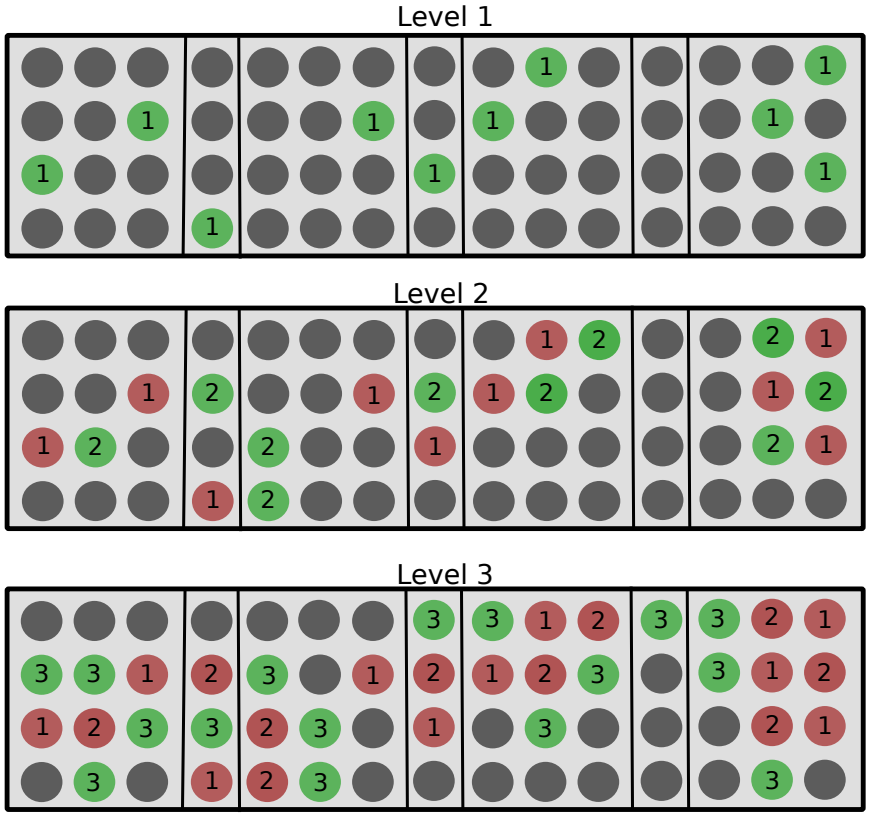


Figure 6.12. Frontier advancement during the level-synchronized global relabeling. For a particular wave, green represents vertices at the frontier and red represents relabeled vertices. The remaining vertices (gray) become a part of this wave as the frontier progresses.

6.2.1.5 Structured Graph and Implicit Addressing Scheme

In the case of general graph construction, every edge is associated with two pointers, one that points to the mate edge (reverse edge in the directed graph) and the other to the node that holds the tail of this mate edge. This data structure entails a considerable amount of memory usage while storing the information of all the edges in the graph. To reduce the memory consumption of the algorithm, we use the implicit addressing scheme. Based on the special configuration of the multicolumn properly ordered graph, any edge information can be exactly calculated based on its offset within a node's set of edges. Since we know the image dimensions beforehand, these implicit calculations can be cached in a 3D matrix to reduce the time required to compute them. The size of the cache is significantly smaller than having a mate stored as part of the edge by a factor equal to the number of graph-columns, which corresponds to a significant reduction in the amount of memory required to store the edge information.

6.2.1.6 Results and Scalability Analysis

To evaluate the efficiency of the proposed parallel push-relabel algorithm, we performed experiments on a $551 \times 426 \times 426$-sized seismic volume. All these experiments were run on shared memory architecture with 80 Intel(R) Xeon(R) CPU E-7-4870 2.4GHz processors and 750GB RAM space.

Fig. 6.13 shows the algorithm's run-time performance (in logarithmic scale) against the number of parallel threads spawned. According to the plot, the algorithm scales well on graphs with a large number of nodes. For moderate-sized graphs comprising 10-20 million nodes, we obtained a nearly linear decrease in the run-time by using 2 to 25 threads, before it leveled off, albeit with an increase in the number of threads. For large-sized graphs comprising 40-80 million nodes, a much better scaling performance scaling was obtained. For instance, we noticed that the run-time is almost halved by doubling the number of processors.

We also compared the run-time performance of the LSGR-based parallel push-relabel scheme against the serial BK algorithm on medium-sized graphs. Table 6.2 shows this comparison for two experiments with 11 and 21 million nodes. The current serial version of the BK algorithm failed to work on larger graphs due to the high memory requirements. Table 6.3 compares the memory usage of the algorithm by employing general and implicit addressing schemes. In the case of the general addressing scheme, each node consumes 128 bytes of memory, and each half-edge holds 32 bytes. Here, half-edges are defined as a set of edges that a node carries on one side of its adjacent columns in a multicolumn graph,

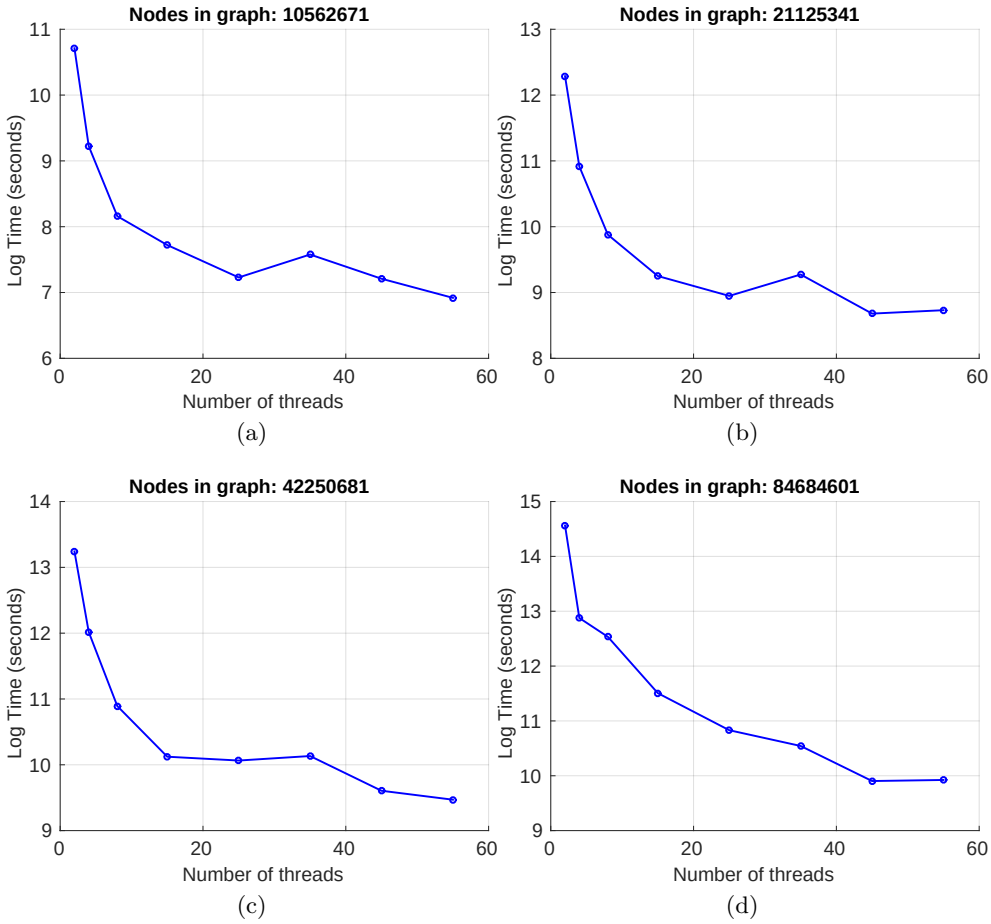


Figure 6.13. LSGR-based parallel push-relabel algorithm's run-time (in logarithmic scale) against the number of threads for graphs with varying numbers of nodes. (a) 10562671 nodes (image size: 551 × 45 × 426), (b) 21125341 nodes (image size: 551 × 90 × 426), (c) 42250681 nodes (image size: 551 × 180 × 426), and (d) 84684601 nodes (image size: 551 × 360 × 426).

Table 6.2. Comparison of serial run-time (BK) and parallel push-relabel algorithm.

Image	Number of nodes	Serial BK	Parallel push-relabel
Volume 1 :	10562671	3625 sec.	121 sec.
Volume 2 :	21125341	19228 sec.	410 sec.

Table 6.3. Comparison of memory usage by using general and implicit addressing schemes of the parallel push-relabel algorithm.

Image	Number of nodes	General addressing scheme	Implicit addressing scheme
Volume 1 :	10562671	32GB	6GB
Volume 2 :	21125341	59GB	13GB

as shown in Fig. 6.14. Although the memory consumed by each node by using the implicit addressing scheme is the same, each half-edge requires only 4 bytes of memory−a significant memory improvement over the general addressing strategy.

6.2.2 Fast Approximate Surface-Based Energy Minimization

Although the LSGR-based parallel push-relabel approach achieves better scalability, cache-friendliness, and memory efficiency by taking advantage of the special graph structure, it is still impractical for application to massive seismic volumes. This problem is due to the underlying VCE-weight net graph structure that comprises around 100 million nodes (taking narrow-band into consideration), with each node carrying up to the row-equivalent number of half-edges. Assume that the number of rows in a given volume is 100, which results in a total of about one trillion edges. Significant time and memory are required to compute max-flow on such a huge graph. Therefore, we need a segmentation model that

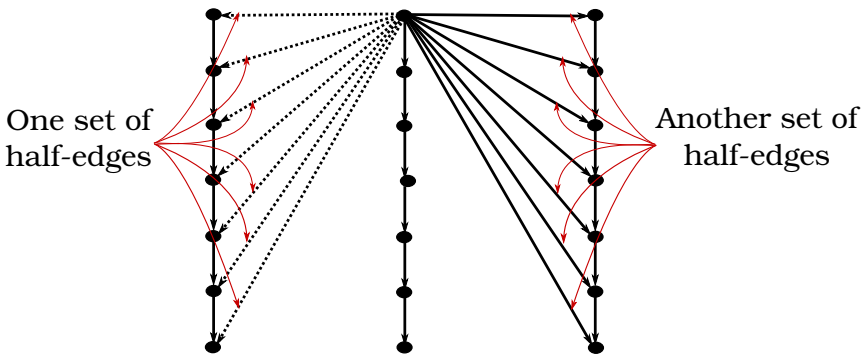


Figure 6.14. Schematic showing a set of half-edges on either side of a node in the multicolumn graph.

is faster and more efficient in memory than the MRF-based Bayesian surface extraction algorithm but does not degrade the segmentation accuracy.

To tackle this challenge, we resort to fast and efficient approximate energy minimization schemes that employ graph cuts iteratively [13]. As discussed in Chapter 1 (Subsection 1.2.3), these schemes work on certain types of energy that are derived from the Bayesian labeling of first-order MRFs [14]. These energies are specifically related to pixel labeling in a way that the minimization of these energies leads to the optimal assignment of a unique label to each pixel. Since the global minimization of these energies is NP-hard in the case of multiple labels, Boykov et al. [13] proposed two fast, highly efficient approximation algorithms. These approximate schemes produce locally optimal solutions when dealing with multiple labels. However, their empirical performance has been tested in producing acceptable solutions efficiently in various applications [13]. By relying on their work, the following work presents the approximate surface-estimation approaches to efficiently extract horizons from the large volumes. Nonetheless, the energy related to these approaches is primarily based on region. The following paragraph describes the underlying strategy behind these approximation schemes and the transformation of their energy formulation to accommodate surfaces.

The basic mechanism behind the approximate energy minimization algorithms is that given a set of labels, $l \in \mathcal{L}$, every pixel, $p \in \mathcal{P}$, in a given volume needs to be assigned a unique label. Now, the goal is to find a labeling such that the energy defined over the assignment of labels to the pixels is minimized. To achieve this, the algorithm relies on larger *moves* by simultaneously modifying the labels of a large collection of pixels. These moves are referred to as *expansion* moves and *swap* moves, and the corresponding algorithms are called $\alpha-$expansion and α-$\beta-$swap algorithms. The $\alpha-$expansion algorithm works by iteratively optimizing the labeling by changing one label at each iteration. Similarly, the α-$\beta-$swap algorithm works by iteratively optimizing the labeling with respect to one pair of labels at a time. During each iteration, both algorithms use graph cuts to find the global solution with respect to either individual labels ($\alpha-$expansion) or a pair of labels (α-$\beta-$swap). However, the basic difference between these two algorithms lies in their graph construction, which supports a specific type of interaction penalty and optimizes their energies. This penalty type defines the distance between a pair of labels in the label-set. To be specific, $\alpha-$expansion allows only a metric-type interaction penalty, whereas α-$\beta-$swap accepts metric as well as semimetric-type interaction penalties.

Based on the application requirements, the energy functions that are considered in this

work focus primarily on the surfaces. Thus, we extend the idea behind these approximation schemes in developing highly efficient, surface-based energy minimization algorithms by treating an intersection of each column and each slice within a given volume as a *site*. A set of all sites defines the search space. The rows along each site of the search space define the label-set. Now, the goal is to find a unique row (label) per site such that the total energy corresponding to the assignment of rows to the sites is minimum. This design can be interpreted as a multicolumn problem in which sites resemble columns and rows represent layers. Therefore, it is viable to adopt the energy terms corresponding to the multicolumn-based surface estimation technique in this approximation approach. This entails data costs, which reflect the assignment of a particular row to a site, and interaction costs/penalties based on the MRF-prior between a pair of rows in the adjacent sites. While expressing this MRF-prior in the form of a metric (for example, ℓ_1-norm), both approximation approaches can be used. On the other hand, if the MRF-prior is expressed in the form of a semimetric (ℓ_2-norm), only the α-$\beta-$swap surface estimation approach can be used. The advantage of these approaches over the multicolumn surface estimation approach lies in their energy minimization strategy, which deals with only one row or a row-pair at any given time. This strategy reduces the computation time as well as memory consumption by a factor corresponding to the number of rows in a volume.

6.2.2.1 Results and Discussion

To analyze the run-time and the memory consumption of the approximate surface estimation algorithms, we performed experiments using an $\alpha-$expansion-based approach on various seismic volumes of different sizes. Table 6.4 demonstrates the run-time and Table 6.5 shows a comparison of the memory consumption of this algorithm and the multicolumn-based globally optimal surface estimation technique. These results demonstrate the effectiveness of the approximation approaches in producing highly efficient solutions in terms

Table 6.4. Run-time comparison of $\alpha-$expansion and globally optimal surface estimation algorithms (— represents too much time).

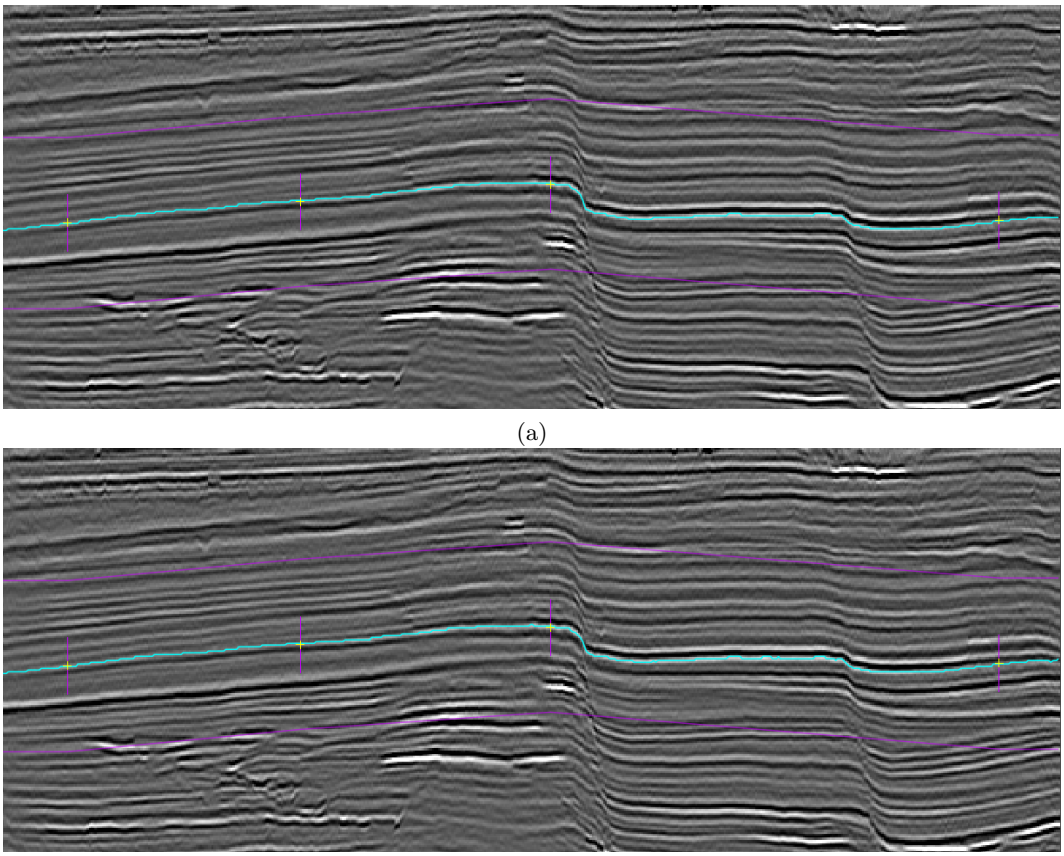
Image dimensions	$\alpha-$expansion	Global
$179 \times 45 \times 10$	0 sec	4 sec.
$187 \times 45 \times 385$	1 sec	167 sec.
$374 \times 39 \times 218$	6 sec	2273 sec.
$551 \times 90 \times 426$	15 sec	—
$1000 \times 75 \times 1024$	83 sec	—

Table 6.5. Comparison of memory usage between $\alpha-$expansion and globally optimal surface estimation algorithms (— did not fit in the memory).

Image dimensions	$\alpha-$expansion	Global
$179 \times 45 \times 10$	2MB	333MB
$187 \times 45 \times 385$	585MB	23.7GB
$374 \times 39 \times 218$	714MB	13.3GB
$551 \times 90 \times 426$	1GB	—
$1000 \times 75 \times 1024$	7GB	—

of time as well as memory. Specifically, the $\alpha-$expansion approach gains an exponential improvement in time and memory over the globally optimal surface estimation method.

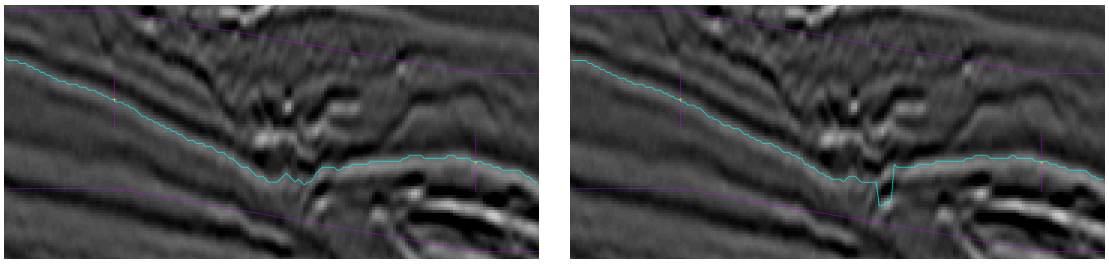
Besides the efficiency, we examined the quality of the extracted horizons in various challenging situations including the surrounding horizon structures that possess better data-related features than the desired horizon with respect to the user-defined landmarks. Figs. 6.15 and 6.16 show the slices of seismic examples with horizons extracted from a set of user-defined landmarks. As seen in Fig. 6.16b, $\alpha-$expansion algorithm produces undesired output around the region pertaining to the low-quality data-related features. In such kind of regions, the algorithm is attracted to the surrounding horizons. This undesired behavior of the algorithm is related to the initialization problem, because of which the solution has a tendency to get stuck in the local minimum.



(a)

(b)

Figure 6.15. Horizon extracted on $551 \times 90 \times 426$-sized volume using (a) globally optimal multicolumn and (b) $\alpha-$expansion surface estimation approaches.



(a)

(b)

Figure 6.16. Horizon extracted on $179 \times 45 \times 10$-sized volume using (a) globally optimal multicolumn and (b) $\alpha-$expansion surface estimation approaces.

6.3 References

[1] N. Otsu, "A threshold selection method from gray-level histograms," *Automatica*, vol. 11, no. 285-296, pp. 23–27, 1975.

[2] A. Z. Arifin and A. Asano, "Image segmentation by histogram thresholding using hierarchical cluster analysis," *Patt. Recognition Lett.*, vol. 27, no. 13, pp. 1515–1521, 2006.

[3] P. V. Hough, "Method and means for recognizing complex patterns," Tech. Rep., 1962.

[4] A. Fitzgibbon, M. Pilu, and R. B. Fisher, "Direct least square fitting of ellipses," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 21, no. 5, pp. 476–480, 1999.

[5] M. Hazewinkel, *Encyclopaedia of Mathematics: Volume 6: Subject IndexAuthor Index*. The Netherlands: Springer Science & Business Media, 2013.

[6] X. Wu and D. Z. Chen, "Optimal net surface problems with applications," in *Automata, Languages and Programming: 29th International Colloquium, ICALP 2002 Málaga, Spain, July 8–13, 2002 Proc.* Berlin, Heidelberg: Springer Berlin Heidelberg, 2002, pp. 1029–1042.

[7] D. Patel, S. Bruckner, I. Viola, and E. M. Groller, "Seismic volume visualization for horizon extraction," in *Vis. Symp. (PacificVis), 2010 IEEE Pacific*. IEEE, 2010, pp. 73–80.

[8] A. V. Goldberg and R. E. Tarjan, "A new approach to the maximum-flow problem," *J. ACM (JACM)*, vol. 35, no. 4, pp. 921–940, 1988.

[9] D. D. Sleator and R. E. Tarjan, "A data structure for dynamic trees," *J. Comput. Syst. and Sci.*, vol. 26, no. 3, pp. 362–391, 1983.

[10] R. Anderson and J. C. Setubal, "A parallel implementation of the push-relabel algorithm for the maximum flow problem," *J. Parallel and Distributed Comput.*, vol. 29, no. 1, pp. 17–26, 1995.

[11] V. Vineet and P. Narayanan, "Cuda cuts: Fast graph cuts on the gpu," in *Comp. Vision and Patt. Recognition Workshops, IEEE Computer Society Conf. on*. IEEE, 2008, pp. 1–8.

[12] J. Liu and J. Sun, "Parallel graph-cuts by adaptive bottom-up merging," in *Computer Vision and Patt. Recognition (CVPR), IEEE Conf. on*. IEEE, 2010, pp. 2181–2188.

[13] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Trans. Patt. Anal. Mach. Intell*, vol. 23, no. 11, pp. 1222–1239, 2001.

[14] S. Z. Li, *Markov Random Field Modeling in Computer Vision*. London, UK: Springer-Verlag, 1995.

CHAPTER 7

DISCUSSION

This chapter concludes the dissertation by summarizing the contributions of the research, followed by a discussion of its limitations and possibilities for future work.

7.1 Research Contributions

This dissertation develops a surface-based Bayesian image segmentation algorithm that relies on a generative image model by incorporating both global and local shape priors in the energy minimization framework. Unlike previous shape models that impose over-restrictive shape constraints by limiting the solution to some learned subspace of shapes, the proposed algorithm allows deviations from this subspace, in a rather controlled fashion by means of local shape priors. The major advantage in allowing subspace deviations lies in their ability to accommodate test data complexities that are not charaterized by the learned subspace. This ability is crucial when dealing with a limited number of training samples that may not be adequate in defining the underlying high shape variability of the desired object. Further, the framework employs the coordinate descent-based optimization strategy by globally maximizing the surface and parametric shape prior estimate in an iterative fashion. Due to these global updates, the energy to be minimized converges in a few iterations.

To seek a globally optimal surface, image segmentation takes place in a discrete parametric search space whose topology complies with the desired surface. A nested mesh model is required to define this search space. Various mesh-generation strategies have been explored to build good-quality nested meshes. This dissertation also develops a generative intensity model from a training set of intensity patterns with an anisotropic Gaussian distribution around a learned mean. Next, the dissertation extends this energy minimization framework to simultaneously extract multiple surfaces in a globally optimal manner.

We applied the proposed framework to segment the left atrium wall in LGE-MRI volumes. For the purpose of validation, this framework has been evaluated on not only the quality of the segmentation, but also its impact on clinical decision making. This validation helps produce a decision pipeline that has a much lower labor content.

Apart from the segmentation of the left atrium, the proposed framework has been applied to segment contact lenses in OCT images and to extract geological structures in seismic volumes. Due to the enormity of seismic volumes, the proposed framework takes a considerable amount of time and memory to seek globally optimal surfaces. Various strategies have been explored to achieve better speed-ups and memory consumption. Toward this goal, we tried to parallelize graph cuts by using various max-flow algorithms on shared memory multiprocessor systems. To further improve the run-time and memory usage, we adapted fast approximate energy minimization schemes [1] in a way that the energy to be minimized corresponds to the surfaces rather than to the regions.

7.2 Limitations and Future Work

This section discusses the limitations of the presented framework and proposes possible directions in which to extend this research.

7.2.1 Improvements in Global Shape Prior

As discussed in Chapters 2 through 5, the shapes of the left atrium that are acquired from LGE-MRI image volumes are highly variable. Modeling shape priors by using a unimodal Gaussian distribution may not be adequate to address this significant nonlinear shape variation.

To capture this shape variability, one can model shape priors by using a multimodal Gaussian distribution that considers the linear subspace spanned by each mixture component. In particular, global shape priors can be modeled as a mixture of Gaussians [2], whose parameters are learned in a high-dimensional shape space rather than preprojecting onto a low-dimensional subspace. The desired segmentation is estimated by treating the identity of a mixture component as a latent variable and marginalizing it within a generalized expectation maximization (EM) framework. In the E-step, mixture component probabilities are estimated for a given image and current surface estimate. In the M-step, a conditional maximization-based scheme can be used that alternates between two phases. In the first phase, component-specific shape parameters are estimated in closed form. In the second phase, the desired surface is estimated by marginalizing it with respect to different mixture components via a single graph cut.

7.2.2 Initialization of Fast Approximate Energy Minimization Schemes

Chapter 6, which addresses large volume segmentations, such as with seismic data, also discusses approximate energy minimization schemes that are faster and consume less memory than not only the global, surface-based energy minimization algorithm (Chapter 3) but also its parallel versions. However, these schemes have a tendency to get trapped in a local minimum if not initialized properly.

To overcome the local minima problem, a better initialization strategy is required. One way to initialize the approximate energy minimization algorithms is by providing the output of the globally optimal solution, which may be time and memory inefficient. This problem can be handled by extracting the globally optimal horizon from the downsampled volume and initializing the approximate approach based on the output acquired from the downsampling strategy. Other initialization strategies include outputs acquired from interpolation between landmarks and deformable models. Further, the accuracy of desired horizons can be improved by employing a better feature extraction strategy that can accommodate more features related to the geological structures.

7.3 References

[1] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Trans. Patt. Anal. Mach. Intell*, vol. 23, no. 11, pp. 1222–1239, 2001.

[2] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*. New York, NY: Springer-Verlag, 2006.