Spring 2012

Globally optimal image segmentation incorporating region, shape prior and context information

Qi Song
University of Iowa

GLOBALLY OPTIMAL IMAGE SEGMENTATION INCORPORATING REGION, SHAPE PRIOR AND CONTEXT INFORMATION

by

Qi Song

An Abstract

Of a thesis submitted in partial fulfillment of the
requirements for the Doctor of Philosophy
degree in Electrical and Computer Engineering
in the Graduate College of
The University of Iowa

May 2012

Thesis Supervisor: Associate Professor Xiaodong Wu

ABSTRACT

Accurate image segmentation is a challenging problem in the presence of weak boundary evidence, large object deformation, and serious mutual influence between multiple objects. In this thesis, we propose novel approaches to multi-object segmentation, which incorporates region, shape and context prior information to help overcome the stated challenges.

The methods are based on a 3-D graph-theoretic framework. The main idea is to formulate the image segmentation problem as a discrete energy minimization problem. The prior region, shape and context information is incorporated by adding additional terms in our energy function , which are enforced using an arc-weighted graph representation. In particular, for optimal surface segmentation with region information, a ratio-form energy is employed, which contains both boundary term and regional term. To incorporate the shape and context prior information for multi-surface segmentation, additional shape-prior and context-prior terms are added, which penalize local shape change and local context change with respect to the prior shape model and the prior context model. We also propose a novel approach for the segmentation of terrain-like surfaces and regions with arbitrary topology. The context information is encoded by adding additional context term in the energy. Finally, a co-segmentation framework is proposed for tumor segmentation in PET-CT images, which makes use of the information from both modalities. The globally optimal solution for the segmentation of multiple objects can be obtained by computing a single

maximum flow in a low-order polynomial time.

The proposed methods were validated on a variety of applications, including aorta segmentation in MRI images, intraretinal layer segmentation of OCT images, bladder-prostate segmentation in CT images, image resizing, robust delineation of pulmonary tumors in MVCBCT images, and co-segmentation of tumors in PET-CT images. The results demonstrated the applicability of the proposed approaches.

Abstract Approved: _____
Thesis Supervisor

_____
Title and Department

_____
Date

GLOBALLY OPTIMAL IMAGE SEGMENTATION INCORPORATING REGION, SHAPE PRIOR AND CONTEXT INFORMATION

by

Qi Song

A thesis submitted in partial fulfillment of the
requirements for the Doctor of Philosophy
degree in Electrical and Computer Engineering
in the Graduate College of
The University of Iowa

May 2012

Thesis Supervisor: Associate Professor Xiaodong Wu

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

PH.D. THESIS

This is to certify that the Ph.D. thesis of

Qi Song

has been approved by the Examining Committee for the thesis requirement for the Doctor of Philosophy degree in Electrical and Computer Engineering at the May 2012 graduation.

Thesis Committee: _____
Xiaodong Wu, Thesis Supervisor

_____
Milan Sonka

_____
Mona K. Garvin

_____
Punam K. Saha

_____
R. Alfredo C. Siochi

ACKNOWLEDGEMENTS

Many thanks to all people who contributed to my Ph.D study.

First, I would like to thank my advisor, Dr. Xiaodong Wu, for the guidance and support he has given me throughout the five years of my Ph.D work. His excellent mentorship, invaluable insights and wide knowledge have a remarkable influence on my future career. I am very fortunate for the opportunity to work under his instruction. Dr. Milan Sonka also deserves a special mention as someone who has been a significant help for his valuable suggestions and consistent support throughout my Ph.D study.

Special thanks also go to Dr. Mona Garvin, who provided valuable instructions for my thesis work and also shared the OCT data for my study. I am most gracious to Dr. R. Alfredo C. Siochi, who exported and reconstructed MVCBCT data, which is very time consuming and I greatly appreciate it. I would like to thank Dr. Punam Saha, who served on my defense committee and provided valuable suggestions.

I would like to thank Dr. Honghai Zhang, Dr. Gary Christenesn and Dr. Andreas Wahle in the Department of Electrical and Computer Engineering for their help during my study. I would also like to thank Dr. Ryan Flynn, Dr. Yusung Kim, Dr. John Buatti and Dr. John Bayouth in the Department of Radiation Oncology for their support to provide us the experimental data and also the instruction on radiotherapy.

Sincere thanks go to all my lab members, Xin Dou, Yunlong Liu, Junjie Bai,

and other students in the Deparment of ECE and the Department of BME, including Bhavna Antony, Kunlin Cao, Zhiyun Gao, Weichen Gao, Zhihong Hu, Yinxiao Liu, Hu Qiao, Sanhui Sun, Ruikai Wang, Ziyue Xu, Xiayu Xu, Ping Yu and Cheng Zhang. Special thanks to Mingqing Chen, who helped us to prepare the data.

This work was supported, in part, by NSF grants CCF-0830402 and CCF-0844765, and NIH grants R01-EB004640 and K25-CA123112.

Last but not least, I would like to thank all my family members' love and support during the study.

ABSTRACT

Accurate image segmentation is a challenging problem in the presence of weak boundary evidence, large object deformation, and serious mutual influence between multiple objects. In this thesis, we propose novel approaches to multi-object segmentation, which incorporates region, shape and context prior information to help overcome the stated challenges.

The methods are based on a 3-D graph-theoretic framework. The main idea is to formulate the image segmentation problem as a discrete energy minimization problem. The prior region, shape and context information is incorporated by adding additional terms in our energy function , which are enforced using an arc-weighted graph representation. In particular, for optimal surface segmentation with region information, a ratio-form energy is employed, which contains both boundary term and regional term. To incorporate the shape and context prior information for multi-surface segmentation, additional shape-prior and context-prior terms are added, which penalize local shape change and local context change with respect to the prior shape model and the prior context model. We also propose a novel approach for the segmentation of terrain-like surfaces and regions with arbitrary topology. The context information is encoded by adding additional context term in the energy. Finally, a co-segmentation framework is proposed for tumor segmentation in PET-CT images, which makes use of the information from both modalities. The globally optimal solution for the segmentation of multiple objects can be obtained by computing a single

maximum flow in a low-order polynomial time.

The proposed methods were validated on a variety of applications, including aorta segmentation in MRI images, intraretinal layer segmentation of OCT images, bladder-prostate segmentation in CT images, image resizing, robust delineation of pulmonary tumors in MVCBCT images, and co-segmentation of tumors in PET-CT images. The results demonstrated the applicability of the proposed approaches.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

Figure

CHAPTER 1
INTRODUCTION

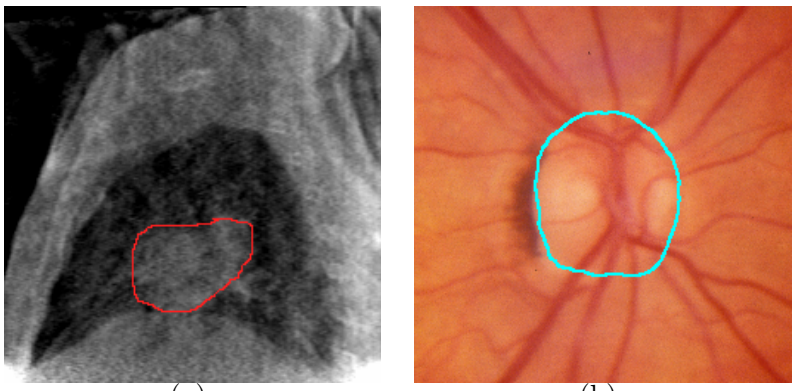
1.1 Introduction and Specific Aims

Image segmentation is widely accepted as one of the most important steps for high-level image understanding. The main purpose is to divide the image into regions with certain common characteristics, which have a strong correlation with objects or areas of the real world [51]. The possible application of image segmentation includes machine vision, face recognition, image compression, *etc.* Recently, medical image segmentation has become a very hot topic. Due to the widespread use of diagnostic imaging devices, there is an urgent need to develop proper computer-aided techniques to handle the huge amount of medical image data generated by the acquisition devices. In particular, automatic segmentation of anatomical structures and regions allows to analyze the medical image data in a quantitative manner. Thus it plays a vital role in numerous biomedical applications including computer aided diagnosis, treatment planning and computer-integrated surgery [17]. Though intensive research has been done, accurate segmentation of organs and other structures of interest is still a challenging problem. The main problems lie in the following aspects[14, 53]:

First, target objects often lack strong boundaries in the presence of high noise or object occlusions, *e.g.,* the lung tumor in reconstructed MV-CBCT image (Fig. 1.1(a)) and the optic disc in the stereo color image (Fig. 1.1(b)), *etc.*

Second, human organs often present a large variation in both shape and size.

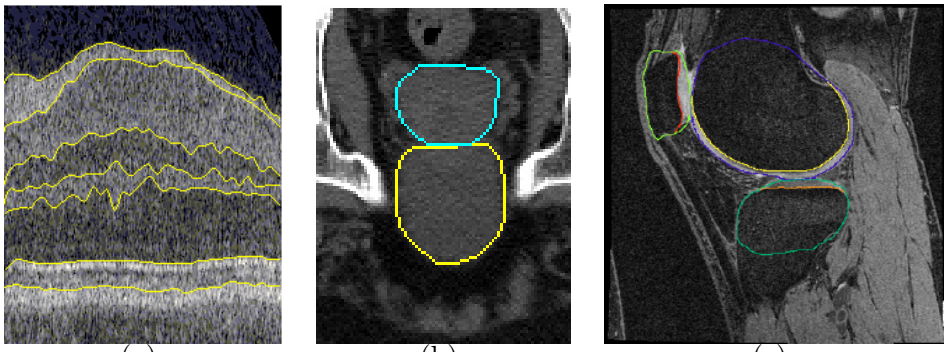
Third, many target surfaces have serious mutual influence between each other. Fig. 1.2 shows some examples of mutually interacting surfaces.



(a) (b)

Figure 1.1: Surfaces with weak boundary information: (a) Tumor surface in reconstructed MV-CBCT image.(b) Optic disc in the stereo color image.

To solve these problems, incorporation of prior knowledge is a good choice. In this thesis, we propose novel approaches for globally optimal image segmentation, which are mainly based on the graph search framework developed by Wu *et al.* [82] and Li *et al.* [37], and the graph cut framework developed by Boykov *et al.* [85, 84]. The main idea is to formulate image segmentation problems as discrete energy optimization problems, which can be solved by a graph-based method. Region, shape and context information is incorporated into the energy function, which can be enforced through proper graph construction utilizing the weights of both graph nodes and arcs. Specifically, a ratio form of energy function is employed in the graph search frame-



(a) (b) (c)

Figure 1.2: Multiple surfaces with serious mutual influence: (a) Intraretinal layers of OCT images. (b) Bladder (yellow) and prostate (blue) surfaces. (c) Bone and cartilage surfaces.

work, which allows regional information such as volume, homogeneity and texture to be included with boundary information without introducing additional parameters. To encode the shape and context information, pair-wise energy terms are added into the graph search energy. In particular, a shape-prior term is used to penalize local shape changes and a context-prior term is used to penalize local surface-distance changes from a model of the expected shape and surfaces distances, respectively. We also report a novel solution to a segmentation problem, in which target objects of arbitrary topology mutually interact with terrain-like surfaces, which widely exists in the medical imaging field. The context information between objects are incorporated by adding a context term into the energy function, which is enforced by introducing weighted inter-graph arcs to the graph. Finally, a co-segmentation approach is proposed based on the graph cut framework, which allows simultaneous segmenta-

tion of objects in multi-modality imaging. The context information is incorporated by adding weighted-arcs between corresponding nodes in the graph. To solve the energy minimization problem, maximum-flow algorithm is employed. A globally optimal solution can be achieved by solving a single maximum flow problem in low-order polynomial time, which yields an accurate and efficient segmentation of volumetric medical images.

For validation, the proposed approaches have been applied for several medical imaging applications. Our ratio-form energy with regional information was employed for the segmentation of aorta in MR images. The framework of optimal multiple surfaces segmentation with shape and context information was validated on intraretinal layer segmentation of OCT images and simultaneous segmentation of the bladder and the prostate in CT images. To show the diverse applications of the proposed framework, our approach was also validated in image resizing problem with shape control. The surface-region segmentation framework was evaluated in robust delineation of lung tumors in Mega-Voltage Cone Beam CT (MVCBCT) images. Finally, a co-segmentation of tumors in PET-CT images was performed following our graph-based co-segmentation framework.

In summary, the specific aims of the proposed research are as follows:

- Aim 1: Develop and validate a method for optimal surface segmentation incorporating both boundary and regional information using a ratio-form energy.
- Aim 2: Develop and validate a method for optimal multiple surfaces segmentation with both shape and context priors.

- Aim 3: Develop and validate a method for optimal surface-region segmentation with context constraints.
- Aim 4: Develop and validate a method for optimal co-segmentation in multi-modality imaging with context constraints.

1.2 Thesis Overview

The thesis is organized as follows.

- Chapter 1 gives a brief overview of the proposed methods and the specific aims. The related work of image segmentation with prior information is also discussed.
- Chapter 2 provides a review of the original graph search framework, which is the basis of the proposed work.
- Chapter 3 describes the developed method for the incorporation of regional information into the graph search framework based on a ratio-form energy. The validation on aortic MR image segmentation is also presented in this chapter.
- Chapter 4 shows how to incorporate both shape and context prior information for optimal multiple surface segmentation. The chapter also describes several applications of the proposed method, including intra-retinal segmentation of OCT images, bladder-prostate segmentation and image resizing.
- Chapter 5 provides a solution for a segmentation problem with target objects of arbitrary topology mutually interacting with terrain-like surfaces. The evaluation in lung tumor segmentation in Mega-Voltage Cone Beam CT (MVCBCT) is also described.

- Chapter 6 shows our co-segmentation framework with application on lung tumor segmentation in PET-CT images.
- Chapter 7 provides a general conclusion and possible future directions.

1.3 Related Work

Image segmentation is defined as the partitioning of an image into non-overlapping, constituent regions that are homogeneous with respect to certain characteristic such as intensity or texture. The regions separated by the segmentation correspond to distinct anatomical structures or regions of interest in the image [64, 63]. Segmentation of volumetric image data plays a key role in image processing, in particular as an important preprocessing step for quantitative analysis and object pattern recognition. In the following sections, we mainly focus on the segmentation method that can be formulated as an energy optimization problems. The majority of the methods can be divided in two large groups – the optimization in a continuous space and the optimization on a discrete set of variables. In both groups, the key problem lies in two aspects – how can the information (e.g., shape, context and region) can be encoded into the energy function and how the corresponding energy function can be optimized.

1.3.1 Segmentation via Optimization in Continuous Space

Methods based on energy minimization in continuous space date several decades back. In the framework of active contour models [48, 78], the boundary of the target object was explicitly represented through certain parametrization. These types of

segmentations had difficulties to handle multiple object segmentations or topological changes during the optimization process. To solve the problem, level-set based methods were proposed [69, 29]. The key idea is to represent the evolving contour using a signed function, where its zero level set corresponds to the motion of the contour. The level set framework has several highlighted features. It is implicit, intrinsic. The representation allows topology changes without introducing any specific parameterization. Furthermore, the level set framework allows the incorporation of a wide spectrum of prior information. Here we mainly focus on two aspects:

Level-set method with shape prior information: In [50], Rousson proposed a novel approach for introducing shape priors into level set representation. A global-to-local shape prior models are employed in the level set representations space according to probabilistic principles. A novel energetic term is added, which accounts for shape priors in level set representation. In [73], a level set based variational approach was proposed, which incorporates shape priors into Chan-Vese's model [71]. In their framework, a labelling level set function is added, which indicates the regions on which the prior shape should be compared. Both of these approaches focus on shape priors which are constant in time. In [12], Cremers introduced a novel method, which incorporates a nonlinear *dynamical* shape priors for level set based image segmentation.

Level-set method with context constraints: Level-set framework allows the incorporation of context constraints between multiple target surfaces. In [3], Tsai *et al.* presented a coupled multi-shape model based on level-set representation.

Multiple signed distance functions are employed as the implicit representations of the multiple shape classes. In [49], a bayesian formulation for coupled surface evolutions was introduced. A non-overlapping constraint between two target surfaces is incorporated in a Bayesian inference framework, which drives to the coupled surface evolution, where no overlap is possible. Recently, Ayed *et al.* [27] proposed a new algorithm for level-set image segmentation with a statistical overlap constraint. The segmentation biases the overlap between the distributions of photometric data within the segmentation regions to a statistical description learned *a priori*. The main drawback of the level-set formulation as well as of parametric deformable models is that the corresponding cost functions are usually non-convex. The optimization process based on gradient descent method may consequently easily be trapped in a local minimal solution. Furthermore, it is rarely known how far the obtained solutions are from the globally optimal ones.

Recently, continuous image segmentation has been formulated as a convex function minimization with guaranteed global optimality [72, 76, 31, 75]. While a number of attempts have been made, incorporating prior information into the energy function while maintaining convexity remains challenging. Cremers *et al.* [13] proposed a novel implicit representation of the shape that can be encoded in functionals, which are convex with respect to the shape deformations. The method can only be used for single object segmentation. Pock *et al.*'s method [75] allowed segmentation of multiple objects using a convex relaxation approach. However, no shape prior information was incorporated. In [66], a convex energy function was employed, which

incorporated the shape prior information into a multi-region probabilistic segmentation based on an isometric log-ratio transformation. No context prior information between multi-regions has been included in the energy function.

1.3.2 Segmentation via Optimization in Discrete Space

In recent years, segmentation through energy minimization in the discrete space has attracted considerable attention in computer vision [56, 84, 37]. Most approaches formulate the problem as a graph-based minimization problem. Nodes in the graph correspond to pixels or control points in the original image. Image intensity information as well as the prior knowledge is encoded by adding arcs with proper weights in the graph. The resulting graph is partitioned into areas in a way that certain specific criteria of the segmentation is optimized. Some typical algorithms of this category include normalized cuts proposed by Jianbo Shi *et al.*[33], minimum spanning tree-based algorithms introduced by Felzenszwalb *et al.*[56] and random walker developed by Leo Grady *et al.*[39].

One of the most influential work in related area is Boykov's graph cut algorithm based on minimum s-t cuts [86, 84], which is topologically flexible and shares some elegance with the level set methods. Recently, Wu and Li *et al.* reported a graph-based framework called "graph search" for globally optimal segmentation of multiple surfaces [82, 37]. Their methods allow simultaneous detection of mutually interacting surfaces by solving a single maximum-flow problem, which have been successfully applied in a variety of applications in medical imaging area [37, 74, 20, 62, 46, 91, 61,

83]. In this section we will give a brief introduction about Boykov's graph cut method. A detailed description of graph search method will be given on Chapter 2, which is the basis of the proposed approaches in this thesis. Both methods have a globally optimal guarantee and allow the incorporation of a wide spectrum of constraints.

1.3.2.1 Graph cut method

The general idea of using binary graph cut algorithms for object segmentation was first proposed by Boykov *et al.* in 2001 [85] and was extended in a number of interesting directions [87, 2, 10, 58, 88, 89, 41]. The main idea is to formulate the segmentation problem as a discrete optimization problem. An energy function is minimized over a finite set of integer-valued variables, which are associated with graph nodes corresponding to image pixels [84].

Let $\mathcal{I}$ denote the given image. l denotes the binary variables assigned for each voxel $\mathcal{I}_v$ in $\mathcal{I}$, indexed as l_v over voxels $\mathcal{I}_v \in \mathcal{I}$. Here $l_v = 1$ means that $\mathcal{I}_v$ belongs to the target region R and $l_v = 0$ means that $\mathcal{I}_v$ belongs to the background. The graph cut energy E is expressed as

$$E(R) = \sum_{v \in I} D_v(l_v) + \sum_{(v_i, v_j) \in \mathcal{N}_v} V_{i,j}(l_i, l_j) \ , \tag{1.1}$$

where D_v is the data term measuring how well the label l_v fits the voxel $\mathcal{I}_v$ given the image data, $\mathcal{N}_v$ defines the neighboring relationship between voxels, and the boundary energy term $V_{i,j}(l_i, l_j)$ is the penalty of assigning the neighboring voxels $(\mathcal{I}_i, \mathcal{I}_j)$ to labels l_i and l_j, respectively.

To solve the optimization problem, a graph $G(N, A)$ is defined as a set of

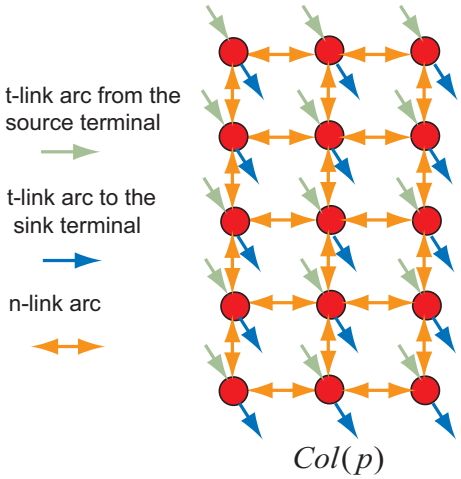


Figure 1.3: Graph construction for graph cut method.

nodes N and a set of arcs A connecting "neighboring" nodes. Every voxel $\mathcal{I}_v \in \mathcal{I}$ has a corresponding node $n \in N$. Two additional nodes, the source (object) s and the sink (background) t, are added. Each node n has one t-link to each of the source and sink, which enforces the data-term energy. Each pair of neighboring nodes is connected by an n-link, which encodes the boundary energy term. The minimum-cost s-t cut divides the graph G into two parts: all nodes belonging to the target object are included in the source set and all background nodes are in the sink set. The result minimizes the graph cut energy $E(R)$. Fig. 1.3 shows one example of the graph construction.

The major advantages of the graph cut method are as follows. First, compared with many other methods which are only applicable in 2-D, the graph cut method provides a globally optimal guarantee for N-dimensional images. Second, graph cut method is capable to handle the segmentation of arbitrary region, which means that

it does not have a specific preferred shape [6]. Furthermore, the graph cut framework allows the incorporation of both region information and boundary information, as well as many other prior information or constraints.

Graph cut with shape priors: Shape prior information can be embedded into the graph cut framework. Freedman *et al.* [15] devised an interactive shape prior segmentation method based on graph cut algorithms. The graph arc-weights were employed, which contained information about a level-set function of a shape template. Malcolm *et al.* [32] incorporated the prior shape information from kernel PCA into an iterative graph cut framework. Vu *et al.* [53] presented a multiple object segmentation framework. The shape energy based on a shape distance function is incorporated via the weights of the edges connected with the terminals.

Graph cut with context information: Graph cut method also allows the incorporation of context information between multiple objects. Boykov *et al.* [6] developed a multi-region framework for multi-object segmentation with mutual interaction. An interaction term is added into the energy function, which incorporated the geometric interactive relationship between different regions. A globally optimal solution can be obtained by solving a single maximum flow problem.

CHAPTER 2
OPTIMAL MULTI-SURFACE SEGMENTATION FRAMEWORK

Our method is mainly based on the graph search framework proposed in [37]. In this chapter, we give a detailed review of the graph search framework. To present the algorithm in a comprehensible manner, we consider the task of detecting multiple terrain-like surfaces representing boundaries of 3-D objects in a volumetric image. Note that the simpler principles used for this illustration are directly applicable to arbitrarily-irregularly meshed surfaces with a pre-segmentation step. A typical example is shown in Section 4.7.

2.1 Problem Formulation

Consider a volumetric image $\mathcal{I}(X, Y, Z)$ of size $X \times Y \times Z$. For each (x, y) pair, the voxel subset $\{\mathcal{I}(x, y, z)|0 \leq z < Z\}$ forms a column parallel to the z-axis, denoted by $p(x, y)$. Each column has a set of neighborhoods for a certain neighbor setting, e.g., four-neighbor relationship [37]. The target here is to find λ terrain-like surfaces which intersect each column $p(x, y)$ at exactly one vertex (See Fig. 2.1). Thus, the terrain-like surface S_i can be defined as a function $S_i(x, y)$, mapping (x, y) pairs to their z-values. A surface adjacency relationship is defined between pairs of surfaces following a given adjacency setting, e.g., two-neighbor relationship in z direction. A boundary-based cost $c_i(x, y, z)$ is assigned to each voxel $\mathcal{I}(x, y, z)$, which is inversely related to the likelihood that the desired surface S_i indeed contains this voxel.

A hard shape constraint is enforced for each surface S_i. Specifically, for any

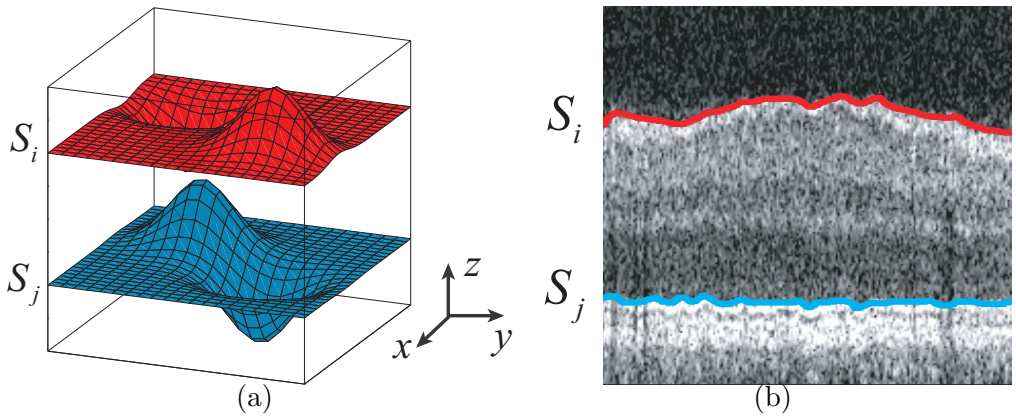


Figure 2.1: Typical example of terrain-like surfaces. (a) Terrain-like surfaces S_i and S_j intersect each (x, y) -column exactly one time. (b) Corresponding surfaces in a 2-D slice.

pair of neighboring columns p and q, the hard shape constraint is defined as follows: $\underline{L}^i_{pq} \leq S_i(p) - S_i(q) \leq \overline{L}^i_{pq}$, where $\underline{L}^i_{pq}$ and $\overline{L}^i_{pq}$ are the hard shape constraint parameters between p and q for S_i(Fig. 2.2(a)).

Surface context constraint is also applied to the column $p(x, y)$ for each pair of the adjacent surfaces S_i and S_j. For each $p(x, y)$, we have $\underline{H}^{ij}_p \leq S_i(p) - S_j(p) \leq \overline{H}^{ij}_P$, where $\underline{H}^{ij}_p$ and $\overline{H}^{ij}_P$ are two specified surface context parameters for S_i and S_j (Fig. 2.2(b)).

The optimal multi-surface segmentation problem is as follows. Given a 3-D image $\mathcal{I}$ and an integer $\lambda > 0$, find a set of λ surfaces $\mathcal{S} = \{S_1, S_2, \ldots, S_\lambda\}$, such that (a) each surface satisfies the hard shape constraint; (b) each pair of adjacent surfaces satisfies the hard surface context constraint; and (c) the following energy function

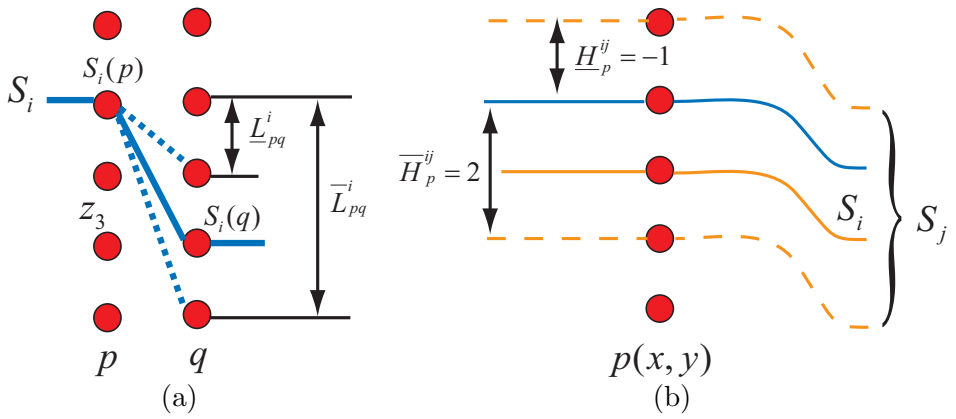


Figure 2.2: Typical examples of the hard shape constraints and the surface context constraints. (a) Hard shape constraints between columns p and q on surface S_i. (b) Surface context constraints. The blue solid line represents S_i. Dashed lines show the possible range of S_j. One example of S_j is represented by the yellow solid line.

$E(\mathcal{S})$ is minimized

$$E(\mathcal{S}) = \sum_{i=1}^{\lambda} \sum_{\mathcal{I}(x,y,z) \in S_i} c_i(x, y, z). \tag{2.1}$$

2.2 Graph Construction

In graph search framework, graph transformation scheme formulates the energy minimization problem as a single computation of maximum flow in the graph [82, 37]. A directed graph G containing λ node-disjoint subgraphs $\{G_i = (N_i, A_i) : i = 1, 2, ..., \lambda\}$ is defined, in which every node $n_i(x, y, z) \in N_i$ represents exactly one voxel $\mathcal{I}(x, y, z)$.

To enforce a variety of geometric constraints, following arcs are constructed.

Intra-column arcs: To ensure the monotonicity of the target surfaces (i.e., the

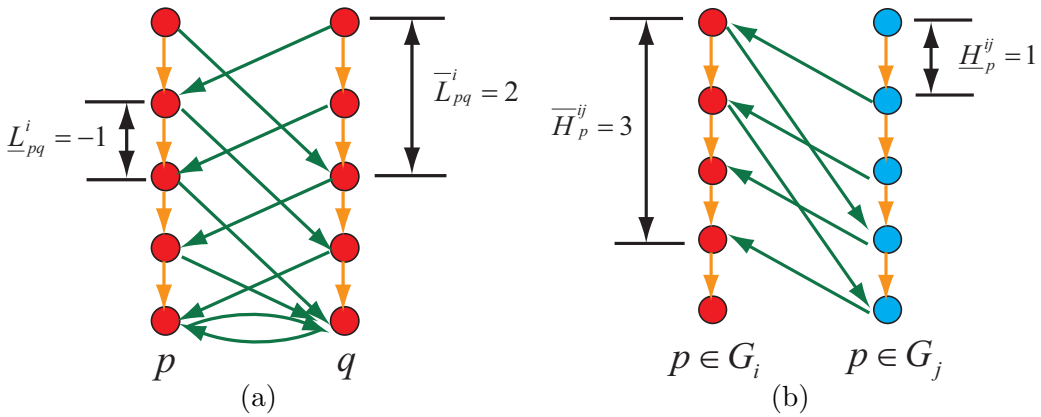


Figure 2.3: Graph construction. (a) Intra-column arcs (orange) and inter-column arcs (green). (b) Inter-surface arcs (green).

target surface intersects with each column exactly once), intra-column arcs are added. Along every column $p(x, y)$, each node $n_i(x, y, z)$ has a directed edge with $+\infty$ weight to the node immediately below it (i.e., $n_i(x, y, z-1)$). Fig. 2.3(a) shows a typical example.

Inter-column arcs: To add the hard shape constraints, the inter-column arcs are added as follows. Let $p(x_1, y_1)$ and $q(x_2, y_2)$ be two neighboring columns. The hard shape constraint has the following form: $\underline{L}^{i}_{pq} \leq S_i(p) - S_i(q) \leq \overline{L}^{i}_{pq}$. A directed arc with $+\infty$ weight is put from each node $n_i(x_1, y_1, z)$ to the node $n_i(x_2, y_2, z - \overline{L}^{i}_{pq})$. Meanwhile, we have a directed arc with $+\infty$ weight from the node $n_i(x_2, y_2, z)$ to $n_i(x_1, y_1, z + \underline{L}^{i}_{pq})$ (Fig. 2.3(a)).

Inter-surface arcs: To enforce the surface context constraints $\underline{H}^{ij}_{p} \leq S_i(x, y) - S_j(x, y) \leq \overline{H}^{ij}_{p}$ between any two adjacent surfaces S_i and S_j, the inter-surface arcs

are added. For each $p(x, y)$, a directed arc with $+\infty$ weight is added from each node $n_i(x, y, z)$ to $n_j(x, y, z - \overline{H}_p^{ij})$ to make sure that $S_i(x, y) - S_j(x, y) \leq \overline{H}_p^{ij}$. On the other hand, each node $n_j(x, y, z)$ also has a directed arc with $+\infty$ to $n_i(x, y, z + \underline{H}_p^{ij})$ such that $\underline{H}_p^{ij} \leq S_i(x, y) - S_j(x, y)$ (Fig. 2.3(b)).

As described in [82, 37], the weight of node $n_i(x, y, z)$ for surface S_i is set as following

$$w_i(x, y, z) = \begin{cases} c_i(x, y, z) & \text{if } z = 0 \\ c_i(x, y, z) - c_i(x, y, z - 1) & \text{if } 0 < z \leq Z - 1 \end{cases} \tag{2.2}$$

.

Here $c_i(x, y, z)$ denotes the boundary-based cost for node $n_i(x, y, z)$ on surface S_i.

2.3 Optimization

With the constructed graph G, an optimal cut $\mathcal{C}^* = (A^*, \bar{A}^*)$ $(A^* \cup \bar{A}^* = N)$ can be found in G, minimizing the total weight of nodes in A^*. The optimal cut corresponds to the optimal set of λ surfaces in $\mathcal{I}$, which minimizes our energy $E(\mathcal{S})$. As described in [82, 37], this optimal cut can be found by solving a maximum flow problem in a low-order polynomial time in a transformed graph, as follows. A source node s and a sink node t are added to G. Each node n is connected to either the source s by the arc with weight $-w_n$ if $w_n < 0$ or the sink t by the arc with weight w_n if $w_n > 0$. Using this construction, a minimum s-t cut of the resulting graph G_{st} can be computed using a maximum flow algorithm, which corresponds to the optimal set of λ surfaces in $\mathcal{I}$, which minimizes our energy $E(\mathcal{S})$. The optimal λ surfaces can be

recovered from these optimal cut as follows. For each target surface i ($i = 1, \ldots, \lambda$), if node $n_i(x, y, z^*)$ has the largest z-coordinate for all nodes belonging to the source set in column $p(x, y)$, then corresponding image voxel $\mathcal{I}(x, y, z^*)$ is on the optimal surface S_i. In this way, the optimal cut in the graph uniquely defines optimal λ surfaces in $\mathcal{I}$.

CHAPTER 3
OPTIMAL SURFACE SEGMENTATION INCORPORATING BOTH BOUNDARY AND REGIONAL INFORMATION USING RATIO-FORM ENERGY

3.1 Introduction

In the original graph-searching framework described in Chapter 2, the cost of a feasible surface (or a set of feasible surfaces) is defined as the total sum of cost values associated with voxels on the surface(s) (i.e., the cost of a voxel with respect to a particular surface reflects the unlikeliness that the voxel would be part of the surface). While such boundary-based costs are commonly used for segmentation, in many cases the object boundaries in medical image data may lack strong edges, e.g.,several objects overlap together or multiple adjacent objects with similar intensity profiles are present in an image. In addition, the segmentation results may be too sensitive to image noise if only using image edge information. Garvin *et al.* [47, 46] first incorporated the regional information into the graph-searching framework by using an energy functional with a linear combination of an boundary-cost term and a region-cost term, as in the standard active contour energies. However, it is not an easy task to select appropriate parameters (i.e., co-efficients) for such a linear energy combination and the segmentation results may highly depend on the parameters selected.

To address those problems, we propose in this chapter a new form of energy function for globally optimal segmentation of a *single* surface, which incorporates

both boundary and regional information, for the optimal surface detection. The energy takes a ratio form of the "on-surface" cost and the "in-region" cost. The on-surface cost is a summation of cost values associated with voxels on the surface (i.e., the cost of a voxel with respect to a particular surface reflects the unlikeliness that the voxel would be part of the surface); The in-region cost is the total sum of cost values associated with voxels in the region bounded by the surface. We thus introduce an optimal surface segmentation model allowing regional information such as volume, homogeneity and texture to be included with boundary information such as intensity gradients. Other types of boundary and regional information can also be incorporated for certain applications. Unlike the linear combination in the standard active contour energies, our ratio-form energy function allows the incorporation of both image boundary and regional information without introducing additional parameter. More important, this energy admits a globally optimal solution in low-order polynomial time. By judiciously characterizing the intrinsic structure of the problem, We are able to globally optimize the energy functional by computing a parametric maximum flow [22]. The ratio-form energy functions were used in the literature before. However, previous approaches either cannot guarantee global optimal solution (e.g. Grady *et al.* [40]), or is very hard, if not impossible, to incorporate shape information (e.g. Kolmogorov *et al.* [79]), or only works for lower dimension segmentation, e.g., Jermyn *et al.* [28] and Stahl *et al.* [34]. Althogh Shi and Malik's normalized cut method [33] works for higher dimensional space, it is computationally intractable.

To validate the method, we applied the new algorithm for the segmentation of 15 3-D aortic MR images. The experiments showed that our method can produce highly accurate and consistent segmentation results.

3.2 Energy Function

Given a volumetric image $\mathcal{I}(X, Y, Z)$ of size $X \times Y \times Z$, we try to find a terrain-like surface S intersecting each column $p(x, y)$ exactly one time. A hard shape constraint is enforced for surface S between pairs of neighboring columns p and q, as described in Section 2.2. An boundary-based *on-surface* cost $c(x, y, z)$ is assigned to each voxel $\mathcal{I}(x, y, z)$, which is inversely related to the likelihood that the desired surface S_i contains the voxel $\mathcal{I}(x, y, z)$. The total on-surface cost of the feasible surface S, denoted $h(S) = \sum_{\mathcal{I}(x,y,z)\in S} c(x, y, z)$, is the total on-surface cost of all voxels on the surfaces.

To incorporate both boundary and regional information into the optimal surface segmentation, we introduce a ratio-form energy. Let $R(S)$ denotes all the voxels bounded by a feasible surface S, consisting of all voxels either below or above the surface with respect to z-axis. Without loss of generality, we define the subset of voxels below as the region bounded by S, that is, $R(S) = \{\mathcal{I}(x, y, z)|z \leq S(x, y)\}$. Each voxel is assigned a so-called *in-region* cost $r(x, y, z)$ measuring, for example, the homogeneity and/or texture property of the image, which reflects the likeliness of the voxel in the desired region. Let $g(S)$ denote the in-region cost of the region bounded by surfaces S with $g(S) = \sum_{\mathcal{I}(x,y,z)\in R(S)} r(x, y, z)$. The energy functional is defined

as

$$E(S) = \frac{h(S)}{g(S)}. \tag{3.1}$$

Our goal is to find an optimal surface S such that (a) surface S satisfies the hard shape constraint; and (b) the energy (3.1) is minimized.

3.3 Energy Optimization

To optimize the ratio-form energy(3.1), we adopt the parametric search technique. The essential procedure in parametric search is, for a given parameter θ, to compute a desired surface S^* while minimizing $h(S) - \theta g(S)$, that is, to solve the following closely related problem:

$$\mu(S^*, \theta) = \min\{\mu(S, \theta) = h(S) - \theta g(S)\}. \tag{3.2}$$

It is well-known that S^* solves Problem (3.1) if and only if (S^*, θ^*) solves Problem (3.2) for $\theta = \theta^* = \theta(S^*)$ giving the value $\mu(S^*, \theta^*) = 0$ [22]. The following algorithm generates a sequence of solutions until ths condition is met:

ALGORITHM **RegionGraphSearch**.

Step 1. Set $k = 0$. Select an initial surface S^0. Compute $\theta_0 = \frac{h(S^0)}{g(S^0)}$.

Step 2. Solve Problem $\mu(S^{k+1}, \theta_k) = \min_S \mu(S, \theta_k) = h(S) - \theta_k g(S)$.

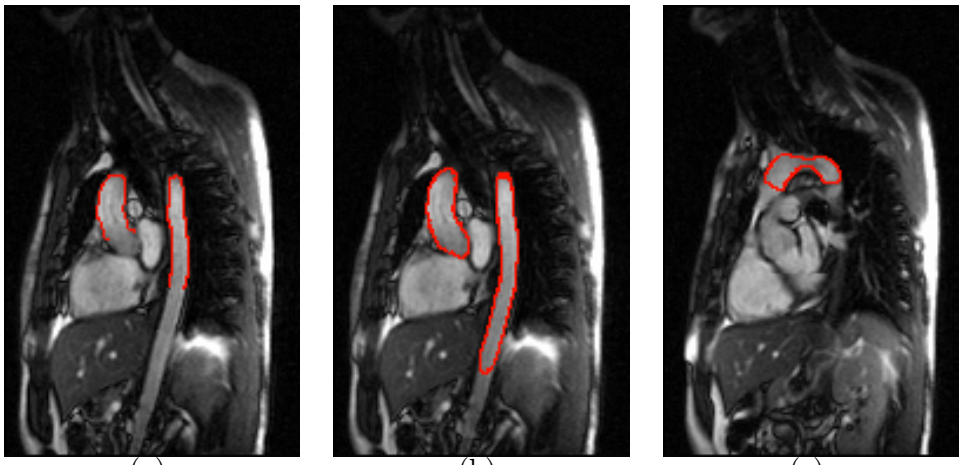
Step 3. If $\mu(S^{k+1}, \theta_k) = 0$, stop and set $S^* = S^k$. Otherwise, let $\theta_{k+1} = \frac{h(S^{k+1})}{g(S^{k+1})}$, replace k by $k + 1$ and go to *Step* 2.

The applicability of the parametric search technique in ALGORITHM **RegionGraphSearch** depends on (1) solving Problem (3.2) in Step 2; and (2) the termination of the algorithm.

To solve Problem (3.2) for a given paramter θ_k, the graph search method in Chapter 2 is applied based on the following observation. For each voxel $\mathcal{I}(x, y, z)$, we re-define its on-surface cost as

$$c'(x, y, z) = c(x, y, z) - \theta_k \sum_{l=0}^{z} r(x, y, l).$$

Then, the cost $h'(S)$ of a surface S under the new voxel costs $c'(\cdot)$ equals to $h(S) - \theta_k g(S)$. Thus, an optimal surface S^{k+1} with respect to the parameter θ_k in Step 2 can be obtained by computing a maximum flow in a transformed graph $G_{st}(\theta_k)$. Furthermore, a very good property of ALGORITHM **RegionGraphSearch** is that the sequence of parameters $\{\theta_k | k = 0, 1, 2, \ldots\}$ generated is decreasing [22]. We thus can view the sequence graphs $\{G_{st}(\theta_k) | k = 0, 1, 2, \ldots\}$ as a monotone parametric network, in which the cost of any arc from source s is a non-decreasing function of a real-valued parameter θ (i.e., $c(x, y, z-1) - c(x, y, z) + \theta r(x, y, z)$) and those of all other arcs are constant with respect to θ. Thus, we can apply Gallo *et al.*'s parametric maximum flow algorithm [22] to compute the sequence of maximum flows in $\{G_{st}(\theta_k) | k = 0, 1, 2, \ldots\}$, yielding a sequence of feasible surfaces $\{S^k | k = 0, 1, 2, \ldots\}$ with the bounded regions $R(S^0) \supseteq R(S^1) \supseteq R(S^2) \supseteq \ldots$. Observe that when $S^{k+1} = S^k$ (i.e., $R(S^{k+1}) = R(S^k)$), ALGORITHM **RegionGraphSearch** terminates. Hence, the algorithm terminates in $\Theta(n)$ steps with $n = X \times Y \times Z$. Thus, using Gallo *et al.*'s algorithm [22], the globally optimal surface S^* can be obtained in the time complexity of computing a *single* maximum flow.



(a) (b) (c)

Figure 3.1: Three slices of the manual segmentation results for aortic segmentation

3.4 Application for Aortic Images Segmentation

Aortic aneurysms and dissections cause 1.3% of all deaths among men aged $65-85$ years in developed countries [52]. Accurate segmentation of the aorta in cardiovascular magnetic resonance (MR) image data is important for early detection of congenital aortic disease. However, manual segmentation of the aorta in 3-D images requires expert knowledge and is tedious and time-consuming (Fig. 3.1). Automated segmentation is becoming a key tool. De Bruijne [45] applied an active shape model for tubular structure segmentation, which requires a training set and a manual initialization model. Bodur [54] introduced a semi-automatic aorta analysis based on isoperimetric partitioning, which was conducted on a series of 2-D cross-sections perpendicular to the centerline of the aorta. Zhao and Zhang [21, 20] adopted a similar graph-based framework for aortic segmentation as we use. However, their energy

functional does not incorporate regional information.

We apply our new globally optimal surface detection algorithm for the aorta segmentation, which consists of the following steps (Fig.3.2): Aortic surface presegmentation, centerline extraction, image resampling and accurate segmentation using our new algorithm.

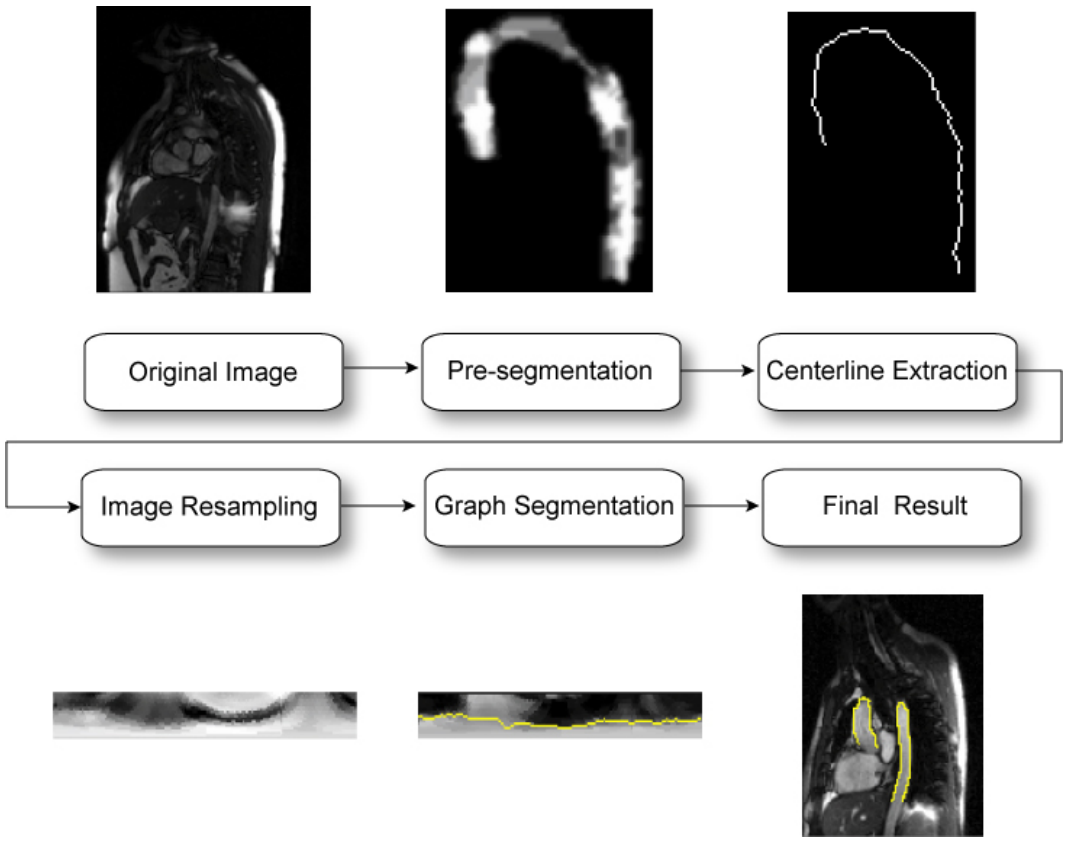


Figure 3.2: Main steps for aortic segmentation.

3.4.1 Pre-processing

Our new algorithm works on 3-D terrain-like surface. Thus we need to first find the centerline of the aorta, then resample the volume image along the plane perpendicular to the centerline and unfold the cylindrical surface into a terrain-like surface.

In order to compute the centerline of the aorta, a 3-D fast marching method [65] is employed for pre-segmentation. Then a skeletonization algorithm [38] is applied to the pre-segmented surface to extract the centerline of the aorta. In order to build the 3-D terrain-like surface, the cross section is obtained by resampling the image perpendicular to the centerline using cubic B-spline interpolation, forming a new volumetric image $\mathcal{I}(X, Y, Z)$.

3.4.2 Graph-based Segmentation

In our aortic segmentation, the gradient-based cost function is employed for on-surface costs. The negative magnitude of the gradient of the image $\mathcal{I}(X, Y, Z)$ is computed at each voxel:

$$c(x, y, z) = -\sqrt{(\frac{\partial \mathcal{I}(x, y, z)}{\partial x})^2 + (\frac{\partial \mathcal{I}(x, y, z)}{\partial y})^2 + (\frac{\partial \mathcal{I}(x, y, z)}{\partial z})^2} \tag{3.3}$$

The in-region cost $r(x, y, z)$ of each voxel $\mathcal{I}(x, y, z)$ is assigned a unit value. Thus, the denominator $g(S)$ in the energy functional actually measures the volume of the region bounded by the surface S. The hard shape constraints are set as 1 in each direction.

3.5 Experimental Methods

3.5.1 Data

15 MR aortic images of 3-D candy-cane view (as shown in Fig. 3.3) at the R-wave peak were acquired from 15 subjects (8 normal, 7 patients). The MR imaging was performed either on a GE Signa or Siemens Avanto 1.5 T scanners, in both cases using a 2-D cine true steady-state free precession gradient-recalled echo imaging. The resolution of image ranges from $1.5 \times 1.5 \times 1.5\ mm^3$ to $2.0 \times 2.0 \times 2.0\ mm^3$. For validation, expert-defined manual segmentation result is obtained as the gold standard.

3.5.2 Validation

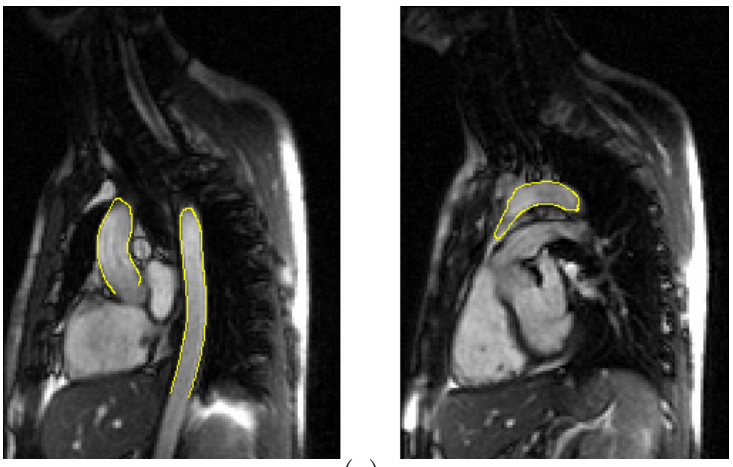
To validate our algorithm, the experiments were carried out on the testing set of 15 MR images. The unsigned surface positioning errors were calculated as the distances between the computed surfaces and the surfaces of the gold standard.The results were reported as mean $\pm$ standard deviation.

3.6 Results

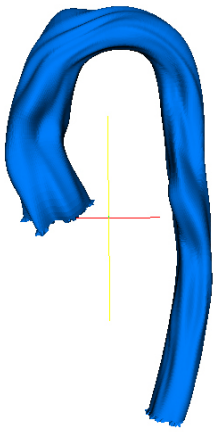
The segmentation results were compared with the expert-defined independent standard. The mean unsigned surface positioning error was computed for quantitative analysis.

All 15 images were successfully segmented. The output results agreed to the manual segmentation quite well . The overall mean unsigned surface positioning error was 0.76 ± 0.88 voxels (1.29 ± 1.50 mm). Fig. 3.3.(a) shows two slices of

segmentation result in transverse view. Fig. 3.3.(b) is the 3-D representation of the aortic segmentation.



(a)

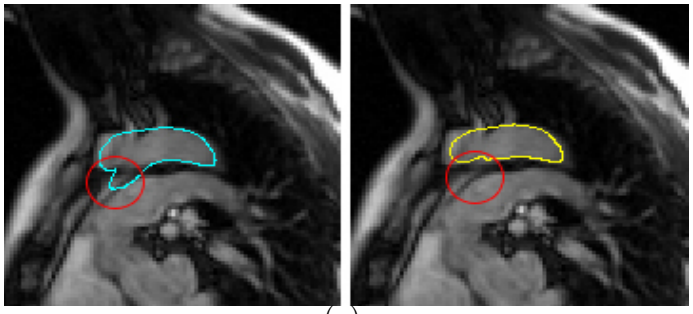


(b)

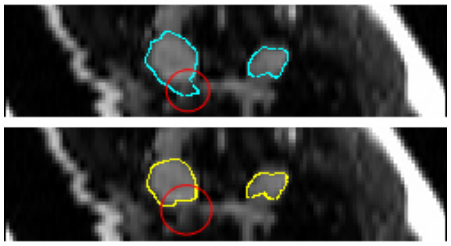
Figure 3.3: The aortic segmentation results in candy cane view. (a) Two slices of segmentation result. (b) 3-D representation of aortic segmentation.

To test the effectiveness of the region-cost term, we segmented the same MR

images with and without using the region-cost term. Fig. 3.4 shows example segmentation results. From both transverse and coronal view, the output boundaries without using the region-cost term were negatively influenced by the adjacent tissues, while those with incorporating the region information successfully avoided such interference.



(a)



(b)

Figure 3.4: The segmentation result with region information (yellow outline) and without region information (blue outline). (a) Transverse View. (b) Coronal View.

3.7 Discussion and Conclusion

In this chapter, we propose a novel algorithm employing a new form of energy function, which incorporates both boundary and regional information, for the optimal surface detection. There are at least three notable unique features of our method: First, our algorithm allows easy incorporation of shape information (e.g. terrain-like or tubular shape) by proper graph modeling in our graph search framework. Recent developments of the unfolding techniques allow the segmentation of more complex shape of object, such as liver, knee bone and cartilage, heart, pulmonary airway trees, and vascular trees. Second, our energy functional is of ratio-form. Thus the incorporation of both image boundary and regional information is parameter free, unlike the linear combination in the standard active contour energies. Third, and the most important, our method guarantees a globally optimal solution in low-order polynomial time and solves a parametric maximum flow problem in the time complexity of computing a single maximum flow.

The utility and performance of our algorithm was demonstrated in 3-D aortic MR image segmentation. Experiment results showed that our method successfully avoided the interference of adjacent tissues.

CHAPTER 4
OPTIMAL MULTIPLE SURFACES SEGMENTATION WITH SHAPE AND CONTEXT PRIORS

4.1 Introduction

In the original graph search framework described in Chapter 2, only node weights are employed in the graph to represent the desired segmentation properties, which limits the ability to incorporate a broader variety of a prior knowledge. To make full use of prior information, we propose an arc-weighted graph representation, which utilizes the weights of both graph nodes and arcs to incorporate a wider spectrum of constraints. Two additional pair-wise terms are added into the energy function, which encode the shape and the context prior information using a set of convex functions. For optimization, the new pair-wise terms are enforced by adding specific weighted arcs in the graph. A globally optimal solution is then computed by solving a single maximum flow problem in the graph, which corresponds to optimal surfaces.

In our framework, the sought surfaces are allowed to pass through the same voxel (Fig. 4.1(a)). However, in some applications, e.g., the image resizing problem, multiple height-field surfaces (seams) in the image need to be removed or inserted to change the image size. Those surfaces may cross each other, but cannot pass through the same voxel (the same voxel cannot be removed twice), as shown in Fig. 4.1(b). In Section 4.4, we prove that when applying the same cost function and shape prior function to all surfaces, the problem of finding multiple crossing surfaces without passing through the same voxel can be transformed into a problem of finding

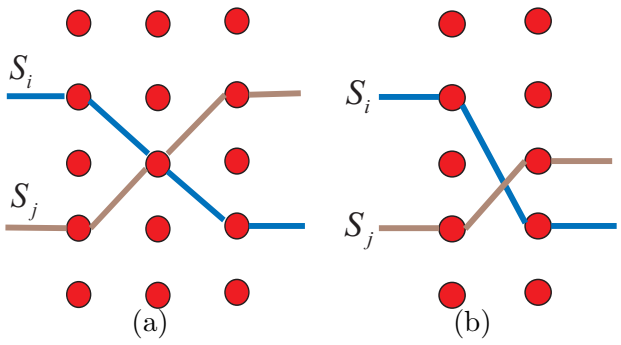


Figure 4.1: Crossing surfaces. (a) Two surfaces cross at a specific voxel. (b) No two crossing surfaces pass through the same voxel.

an optimal solution of multiple non-crossing surfaces, which can be solved in our framework by adding proper surface interacting constraints.

To assess the power of our approach, the proposed framework was applied for three challenging applications, including intraretinal layer segmentation of OCT images (Section 4.6), simultaneous segmentation of the bladder the prostate (Section 4.7), and image resizing problem (Section 4.8). The experiments showed the power of our algorithm.

4.2 Incorporation of Shape and Context Priors

To present our method in a comprehensible manner, in this section we consider the task of detecting multiple *terrain-like* surfaces incorporating shape and context prior knowledge. Note that this simple principle used for this illustration is directly applicable to arbitrarily-irregularly meshed surfaces (see Section 4.7). Consider a volumetric image $\mathcal{I}(X, Y, Z)$ of size $X \times Y \times Z$. For each (x, y) pair, the voxel subset

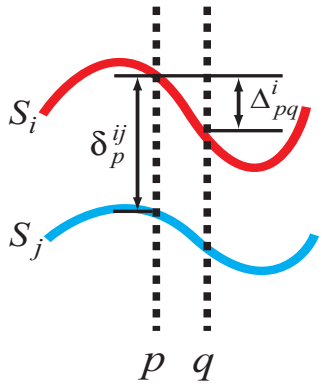


Figure 4.2: The shape representation Δ_{pq}^i between neighboring columns p and q on surface S_i; and the context representation δ_p^{ij} between surfaces S_i and S_j on column p.

$\{\mathcal{I}(x,y,z)|0 \leq z < Z\}$ forms a column parallel to the z-axis, denoted by $p(x,y)$. Each column has a set of neighboring columns for a certain neighbor setting $\mathcal{N}_s$, e.g., four-neighbor relationship. The target is to find λ terrain-like surfaces, each of which intersects each column $p(x,y)$ at exactly one voxel. Thus, the terrain-like surface S_i can be defined as a function $S_i(x,y)$, mapping (x,y) pairs to their z-values. An on-surface cost $c_i(x,y,z)$ is assigned to each voxel $\mathcal{I}(x,y,z)$ for surface S_i, which is inversely related to the likelihood that the desired surface S_i contains the voxel.

4.2.1 Shape prior constraints

In this thesis, the shape of surface S_i are defined as the surface height changes between pairs of neighboring columns. Specifically, for any pair of neighboring columns p and q, the shape of surface S_i between (p,q) relies on $\Delta_{pq}^i = S_i(p) - S_i(q)$ (see Fig.

4.2). Suppose $\overline{\Delta}^i_{pq}$ represents the learned shape model. The shape deformation between current shape Δ^i_{pq} and the prior shape model $\overline{\Delta}^i_{pq}$ can be expressed as $\Delta^i_{pq} - \overline{\Delta}^i_{pq}$. Two kinds of shape constraints are enforced: the hard shape constraint and the shape prior penalty. The hard shape constraint defines the possible range of the shape deformation with the form: $|\Delta^i_{pq} - \overline{\Delta}^i_{pq}| \leq L^i_{pq}$, where L^i_{pq} is the shape constraint parameter between columns p and q for surface S_i. The shape-prior penalty is enforced using a convex function $f_s(\Delta^i_{pq} - \overline{\Delta}^i_{pq})$, which penalizes the shape deformation inside the range of the hard shape constraint.

4.2.2 Context prior constraints

For a set of target surfaces, the context prior penalty is enforced to penalize the surface distance change between two adjacent surfaces. Suppose S_i and S_j are two adjacent surfaces denoted as $(S_i, S_j) \in \mathcal{N}_c$, where $\mathcal{N}_c$ is a given surface adjacency setting, e.g., two-neighbor relationship in z direction. The context information between (S_i, S_j) on column p is defined as $\delta^{ij}_p = S_i(p) - S_j(p)$ (Fig. 4.2). Let $\overline{\delta}^{ij}_p$ denote the learned prior context model. The change between current context and the prior context model can be represented by $\delta^{ij}_p - \overline{\delta}^{ij}_p$. Similarly like the shape prior constraints, two kinds of context prior constraints are employed. The possible distance between two surfaces (S_i, S_j) are defined by hard context constraints as $|\delta^{ij}_p - \overline{\delta}^{ij}_p| \leq H^{ij}_p$, where H^{ij}_p is the context constraint parameter for column p between surfaces (S_i, S_j). The context-prior penalty is given by a convex function $f_c(\delta^{ij}_p - \overline{\delta}^{ij}_p)$ to penalize the context change between current context and the prior model.

Now the overall energy of the set $\mathcal{S}$ of λ surfaces S_i's takes the form:

$$E(\mathcal{S}) = \sum_{i=1}^{\lambda} \sum_{\mathcal{I}(x,y,z)\in S_i} c_i(x,y,z) + \sum_{i=1}^{\lambda} \sum_{(p,q)\in\mathcal{N}_s} f_s(\Delta_{pq}^i - \overline{\Delta}_{pq}^i) + \sum_{p} \sum_{(i,j)\in\mathcal{N}_c} f_c(\delta_p^{ij} - \overline{\delta}_p^{ij}) \quad (4.1)$$

The first term is the boundary energy term, which is equal to the summation of on-surface cost for all voxles on surfaces, as used in the original graph search framework. The boundary energy term drives the surface set towards the best fit to the current image data. The second and the third term are the shape-prior penalty term and the context-prior penalty term proposed in this work, which measure how well the surface set fulfills the prior shape model and the context model, respectively. Using this formulation, we strive to find the optimal terrain-like surfaces such that (a) each surface satisfies the hard shape constraint; (b) each pair of surfaces satisfies the hard context constraints; and (c) the energy in Eq. (6.6) is minimized.

4.3 Arc-Weighted Graph Construction

The basic idea for solving the energy minimization problem is to reduce it into a maximum flow problem. A directed graph G containing λ node-disjoint subgraphs $\{G_i = (N_i, A_i) : i = 1, 2, ..., \lambda\}$ is defined, in which every node $n_i(x,y,z) \in N_i$ represents exactly one voxel $\mathcal{I}(x,y,z)$.

To enforce a variety of hard geometric constraints, the following arcs with $+\infty$ weight are constructed.

Intra-column arcs: To ensure the monotonicity of the target surfaces (i.e., the target surface intersects with each column exactly one time), the intra-column arcs are added as described in [37]. Along every column $p(x,y)$, each node $n_i(x,y,z)(z > 0)$

has a directed arc with $+\infty$ weight to the node immediately below it (i.e., $n_i(x, y, z-1)$).

Inter-column arcs: The hard shape constraint is incorporated by adding inter-column arcs between neighboring columns in the graph. Specifically, let $p(x_1, y_1)$ and $q(x_2, y_2)$ be two neighboring columns. To enforce the hard shape constraint $|\Delta_{pq}^i - \overline{\Delta}_{pq}^i| \leq L_{pq}^i$, a directed arc with $+\infty$ weight is put from each node $n_i(x_1, y_1, z)$ of $p(x_1, y_1)$ to the node $n_i(x_2, y_2, min(Z-1, max(z - \overline{\Delta}_{pq}^i - L_{pq}^i, 0)))$ of $q(x_2, y_2)$. Meanwhile, we have a directed arc with $+\infty$ weight from the node $n_i(x_2, y_2, z)$ to $n_i(x_1, y_1, min(Z-1, max(z + \overline{\Delta}_{pq}^{\ i} - L_{pq}^i, 0)))$.

Inter-surface arcs: The hard context constraint can be enforced by adding inter-surface arcs between different sub-graphs. Suppose S_i and S_j are two neighboring surfaces. The hard context constraint $|\delta_p^{ij} - \overline{\delta}_p^{ij}| \leq H_p^{ij}$ on column $p(x, y)$ is incorporated by adding a directed arc with $+\infty$ weight from each node $n_i(x, y, z)$ to the node $n_j(x, y, min(Z-1, max(z - \overline{\delta}_p^{ij} - H_p^{ij}, 0)))$. On the other hand, each node $n_j(x, y, z)$ also has a directed arc with $+\infty$ weight to the node $n_i(x, y, min(Z-1, max(z + \overline{\delta}_p^{ij} - H_p^{ij})))$.

The remaining challenge is how to incorporate the shape-prior penalty term and the context-prior penalty term into the graph search framework. To solve the problem, additional weighted arcs are introduced in the graph. We start from the incorporation of shape prior penalties.

Weighted inter-column arcs: Let $p(x_1, y_1)$ and $q(x_2, y_2)$ be two neighboring columns. To "distribute" the convex shape prior penalty $f_s(\Delta_{pq}^i - \overline{\Delta}_{pq}^i)$ to arcs between neighboring columns (p, q) in G_i, we make use of the (discrete equivalent of) second

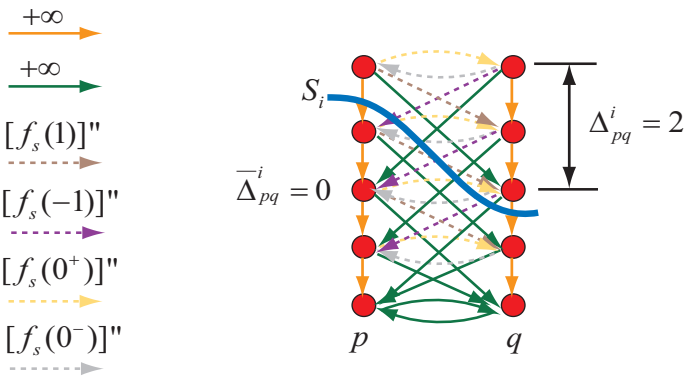


Figure 4.3: Arc-weighted graph construction for the incorporation of the shape prior penalty on surface S_i between neighboring columns p and q. The intra-column arcs are shown in orange with $+\infty$ weight. The hard shape constraint $|\Delta^i_{pq} - \overline{\Delta}^i_{pq}| \leq L^i_{pq}$ is enforced by green arcs. Here we suppose $\overline{\Delta}^i_{pq} = 0$ and $L^i_{pq} = 2$. The shape prior penalty is incorporated by arcs with dashed lines (brown, purple, yellow and gray). Here we assume $\overline{\Delta}^i_{pq} = 0$, $[f_s(0)]' = 0$ and $f_s(0) = 0$. Target surface S_i cuts arcs with weight $[f_s(1)]''$ (brown) and $[f_s(0^+)]'' = f_s(1) - f_s(0)$ (yellow). The total weight is equal to $f_s(2)$.

derivative of $f_s(\cdot)$ with the form $[f_s(h)]'' = [f_s(h+1) - f_s(h)] - [f_s(h) - f_s(h-1)]$. Since $f_s(h)$ is a convex function, $[f_s(h)]'' \geq 0$. Let $[f_s(h)]' = f_s(h+1) - f_s(h)$ denotes the first derivative of $f_s(\cdot)$. For each $h = \Delta^i_{pq} - \overline{\Delta}^i_{pq}$, where $-L^i_{pq} < h < L^i_{pq}$, if $[f_s(h)]' \geq 0$, an arc is added from $n_i(x_1, y_1, z)$ to $n_i(x_2, y_2, z - \overline{\Delta}^i_{pq} - h)$ carrying an arc-weight of $[f_s(h)]''$. If $[f_s(h)]' \leq 0$, an arc from $n_i(x_2, y_2, z)$ to $n_i(x_1, y_1, z + \overline{\Delta}^i_{pq} + h)$ has the weight of $[f_s(h)]''$. Fig. 4.3 shows one typical graph construction. Note that if $h = h_0$, where $[f_s(h_0)]' = 0$, we let $[f_s(h_0^+)]'' = f_s(h_0 + 1) - f_s(h_0)$ for arcs from

$n_i(x_1, y_1, z)$ to $n_i(x_2, y_2, z - \overline{\Delta}^i_{pq} - h_0)$ and $[f_s(h_0^-)]'' = f_s(h_0 - 1) - f_s(h_0)$ for arcs from $n_i(x_2, y_2, z)$ to $n_i(x_1, y_1, z + \overline{\Delta}^i_{pq} + h_0)$.

In the following, we show that using this construction, the total weight of the arcs that are cut by S_i between two neighboring columns p and q equals to the shape prior penalty $f_s(\Delta^i_{pq} - \overline{\Delta}^i_{pq})$. Note that the arc is *cut* by the surface if and only if the arc starts from the node below the surface and ends at the node above the surface. WLOG, we assume that $f_s(h_0) = 0$ (otherwise we can subtract $f_s(h_0)$ from the weight of each arc from p to q without essentially affecting the total arc-weights). According to our arc construction, if $[f_s(h)]' > 0$, where $[f_s(h)]' = f_s(h+1) - f_s(h)$, then S_i cuts the arcs with a total weight of $\sum_{j=1}^{h-h_0} j[f_s(h-j)]'' = f_s(h) - f_s(h_0) = f_s(h)$. If $[f_s(h)]' < 0$, then S_i cuts the arcs with a total weight of $\sum_{j=1}^{h_0-h} j[f_s(h+j)]'' = f_s(h) - f_s(h_0) = f_s(h)$. If $[f_s(h)]' = 0$, no arc is cut, the total weight is 0. Using this construction, the total weight of the arcs that are cut by S_i between neighboring columns p and q equals to the shape prior penalty $f_s(\Delta^i_{pq} - \overline{\Delta}^i_{pq})$, which penalizes the shape change of the surface S_i between the two columns.

Weighted inter-surface arcs: The context-prior penalty is enforced in a similar way by adding weighted inter-surface arcs between corresponding sub-graphs. Suppose S_i and S_j are two adjacent surfaces. The context-prior penalty is distributed between subgraph G_i and G_j on the same column $p(x, y)$. For each $d = \delta^{ij}_p - \overline{\delta}^{ij}_p$, where $-H^{ij}_p < d < H^{ij}_p$, if $[f_c(d)]' \geq 0$, an arc is added from $n_i(x, y, z)$ to $n_j(x, y, z - \overline{\delta}^{ij}_p - d)$ with weight $[f_c(d)]''$. If $[f_c(d)]' \leq 0$, an arc is assigned from $n_j(x, y, z)$ to $n_i(x, y, z + \overline{\delta}^{ij}_p + d)$ with weight $[f_c(d)]''$. The graph construction is shown

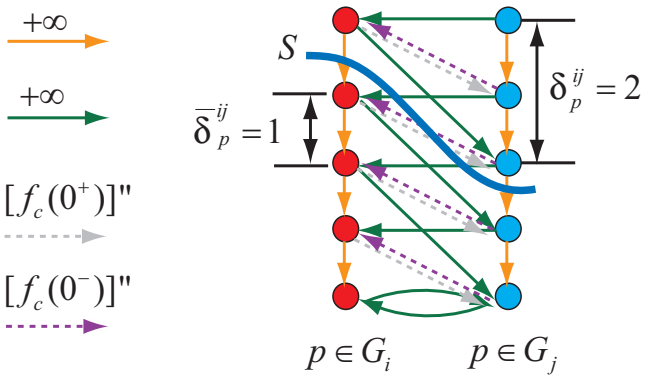


Figure 4.4: Arc-weighted graph construction for the incorporation of the context prior constraints between sub-graphs G_i (red) and G_j (blue) on column p. The hard context constraint $|\delta_p^{ij} - \overline{\delta}_p^{ij}| \leq H_p^{ij}$ is incorporated by green arcs. Here $\overline{\delta}_p^{ij} = 1$, $H_p^{ij} = 1$. The context-prior penalty is enforced by gray and purple arcs. We assume that $[f_c(0)]' = 0$ and $f_c(0) = 0$. Target surface S (connecting S_i and S_j) cuts arcs with weight $[f_c(0^+)]''$ (gray). The total weight is equal to $f_c(1)$.

in Fig. 4.4. Using this construction, the total weight of the arcs cut by the surface set S between two subgraphs G_i and G_j on column p equals to the context prior penalty $f_c(\delta_p^{ij} - \overline{\delta}_p^{ij})$.

To encode the on-surface cost, the weight of each node in the graph is designed using a similar way as described in [37]. Suppose a voxel $I(x', y', z')$ is on a surface S_i. Then all nodes $n_i(x', y', z)$ with $z \leq z'$ in column $p(x', y')$ are viewed as being inside the surface S_i. The node weight is assigned such that the total weight of all nodes n_i inside the surface S_i equals to the boundary energy term $\sum_{\mathcal{I}(x,y,z)\in S_i} c_i(x, y, z)$, where

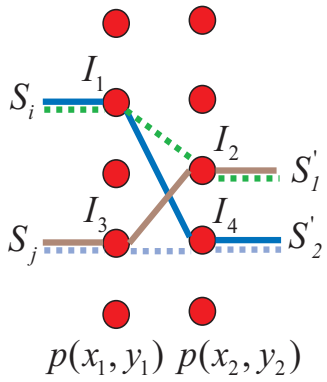


Figure 4.5: Surface transform from the crossing surfaces (blue and brown) into non-crossing surfaces (green and gray).

$i = 1, \ldots, \lambda$:

$$w_i(x, y, z) = \begin{cases} c_i(x, y, z) & \text{if } z = 0 \\ c_i(x, y, z) - c_i(x, y, z-1) & \text{if } 0 < z \leq Z - 1 \end{cases}. \tag{4.2}$$

4.4 Detecting Multiple Surfaces with No Common Voxels

In this section, we consider the multiple surface detection problem, in which no two surfaces share a common voxel. Note that in this surface detection problem, no surface context constraint is enforced. We prove that if using the same cost functions and shape prior constraints for all sought surfaces, an optimal solution to this problem exists, which consists of non-crossing surfaces with minimum distance of 1 between any two adjacent ones.

Consider a volumetric image $\mathcal{I}(X, Y, Z)$ of size $X \times Y \times Z$. For each (x, y) pair, the voxel subset $\{\mathcal{I}(x, y, z) | 0 \leq z < Z\}$ forms a column parallel to the z-axis, denoted by $p(x, y)$. Let $\mathcal{S} = \{S_1, S_2, ..., S_\lambda\}$ denote the optimal set of λ sought terrain-like

surfaces. $\mathcal{V}$ represents the union of all voxels on those surfaces. Let $S_1'(p)$ denote the maximum z-coordinate of the voxels in $\mathcal{V}$ on each column $p(x, y)$. This results in a new surface S_1' in $\mathcal{I}$. Remove all voxels on S_1' from $\mathcal{V}$. Then let $S_2'(p)$ denote the maximum z-coordinate of the voxels in $\mathcal{V} - S_1'$. This results in another new surface S_2'. By doing this iteratively, we can obtain a new set of λ surfaces, denoted as $\mathcal{S}' = \{S_1', S_2', ..., S_\lambda'\}$. In the following, we prove that $\mathcal{S}'$ is an optimal solution.

First, from the construction of the $\mathcal{S}'$, each surface S_i' intersects each column $p(x, y)$ at exactly one voxel. We next prove by induction that $\mathcal{S}'$ satisfies the hard shape constraint and the total shape-prior penalty on $\mathcal{S}'$ are no larger than that on $\mathcal{S}$.

Consider two neighboring voxels $\mathcal{I}_1 = \mathcal{I}(x_1, y_1, z_1)$ and $\mathcal{I}_2 = \mathcal{I}(x_2, y_2, z_2)$ on S_1'. The same hard shape constraint $|\Delta_{12} - \overline{\Delta}_{12}| \leq L_{12}$ is applied for all surfaces. If $\mathcal{I}_1$ and $\mathcal{I}_2$ both belong to the same surface S, they satisfy the hard shape constraint and the shape prior penalty on the arc $(\mathcal{I}_1, \mathcal{I}_2)$ does not change. Otherwise, suppose $\mathcal{I}_1 \in S_i$ and $\mathcal{I}_2 \in S_j$ $(i \neq j)$. Let $\mathcal{I}_3 = \mathcal{I}(x_1, y_1, z_3)$ denote the voxel on S_j, which intersects with $p(x_1, y_1)$, and $\mathcal{I}_4 = \mathcal{I}(x_2, y_2, z_4)$ denote the voxel on S_i, which intersects with $p(x_2, y_2)$ (Fig. 4.5). Since $S_1'(x, y)$ is the maximum z-coordinate of the voxels in $\mathcal{V}$ on each column $p(x, y)$, we have $z_1 > z_3 \geq 0$ and $z_2 > z_4 \geq 0$. Note that S_i and S_j are feasible surfaces which satisfy the hard shape constraint $|\Delta_{12} - \overline{\Delta}_{12}| \leq L_{12}$. Thus $-L_{12} + \overline{\Delta}_{12} \leq z_1 - z_4 \leq L_{12} + \overline{\Delta}_{12}$ and $-L_{12} + \overline{\Delta}_{12} \leq z_3 - z_2 \leq L_{12} + \overline{\Delta}_{12}$. Therefore, $-L_{12} + \overline{\Delta}_{12} \leq z_1 - z_2 \leq L_{12} + \overline{\Delta}_{12}$, which indicates that S_1' satisfies the hard shape constraint. In the meanwhile, we have $-L_{12} + \overline{\Delta}_{12} \leq z_3 - z_4 \leq L_{12} + \overline{\Delta}_{12}$. Thus,

$\mathcal{V} - S'_1$ forms a set $\mathcal{S}_{\lambda-1}$ of $\lambda - 1$ feasible surfaces.

We next look into the change of the total shape prior penalty. The same shape-prior penalty function $f_s(\Delta_{12} - \overline{\Delta}_{12})$ is applied for all surfaces. WLOG, we assume that $\overline{\Delta}_{12} = 0$ (otherwise we can set a new penalty function with the form $f'_s(\Delta_{12}) = f_s(\Delta_{12} - \overline{\Delta}_{12})$). From the computation of S'_1, we replace the old arcs $(\mathcal{I}_1, \mathcal{I}_4)$ and $(\mathcal{I}_3, \mathcal{I}_2)$ by two new arcs $(\mathcal{I}_1, \mathcal{I}_2)$ and $(\mathcal{I}_3, \mathcal{I}_4)$ to obtain S'_1 and $\mathcal{S}_{\lambda-1}$. The total shape prior penalty on the new arcs $(\mathcal{I}_1, \mathcal{I}_2)$ and $(\mathcal{I}_3, \mathcal{I}_4)$ is $f_s(z_1 - z_2) + f_s(z_3 - z_4)$ and that on the old arcs $(\mathcal{I}_1, \mathcal{I}_4)$ and $(\mathcal{I}_3, \mathcal{I}_2)$ is $f_s(z_1 - z_4) + f_s(z_3 - z_2)$. Since $z_1 > z_3$ and $z_2 > z_4$, we have $z_1 - z_2 < z_1 - z_4$, $z_1 - z_2 > z_3 - z_2$, $z_3 - z_4 < z_1 - z_4$ and $z_3 - z_4 > z_3 - z_2$. Due to the convexity of the function $f_s(\cdot)$, we have $f_s(z_3 - z_4) \leq \{(z_1 - z_3)/(z_1 + z_2 - z_3 - z_4)\} f_s(z_3 - z_2) + \{(z_2 - z_4)/(z_1 + z_2 - z_3 - z_4)\} f_s(z_1 - z_4)$ and $f_s(z_1 - z_2) \leq \{(z_2 - z_4)/(z_1 + z_2 - z_3 - z_4)\} f_s(z_3 - z_2) + \{(z_1 - z_3)/(z_1 + z_2 - z_3 - z_4)\} f_s(z_1 - z_4)$. Adding them together, we got

$$f_s(z_1 - z_2) + f_s(z_3 - z_4) \leq f_s(z_1 - z_4) + f_s(z_3 - z_2).$$

Hence, the total shape prior penalty of the resulting λ surfaces (i.e., S'_1 and $\lambda - 1$ surfaces in $\mathcal{S}_{\lambda-1}$) is no larger than that of $\mathcal{S}$.

Let $\mathcal{S}_{\lambda-l}$ denote the remaining $\lambda - l$ surfaces in the solution after we compute the first l surfaces $S'_1, S'_2, \ldots, S'_l$. Assume that all those l surfaces satisfy the hard shape constraint, and the total shape prior penalty of those l surfaces and the $\lambda - l$ surfaces in $\mathcal{S}_{\lambda-l}$ does not increase comparing to that of $\mathcal{S}$. We need to prove that after computing the $(l+1)$st surface S'_{l+1}, the non-increasing property of the total shape prior penalty is still hold and S'_{l+1} satisfy the hard shape constraint. Notice

that S'_{l+1} is computed from the vertex set $\mathcal{V} - \cup_{k=1}^{l} S'_k$, which forms the feasible $\lambda - l$ surfaces in $\mathcal{S}_{\lambda-l}$. We can use exactly the same argument as we do for S'_1, which shows that S'_{l+1} satisfies the hard shape constraint and the non-increasing property of the total shape prior penalty is hold. We thus prove that $\mathcal{S}'$ is a feasible solution and the total shape prior penalty of $\mathcal{S}'$ is no larger than that of $\mathcal{S}$.

Since $\mathcal{S}$ is an optimal solution, no two surfaces in $\mathcal{S}$ pass through the same voxel. It is obvious that no common voxels may present in any two surfaces in $\mathcal{S}'$. The union of the voxels on the surfaces in $\mathcal{S}$ is the same as the union of the voxels on the surfaces in $\mathcal{S}'$. Thus, $\sum_{i=1}^{\lambda} \sum_{\mathcal{I}_i(x,y,z)\in S_i} c_i(x,y,z) = \sum_{i=1}^{\lambda} \sum_{\mathcal{I}_i(x,y,z)\in S'_i} c_i(x,y,z)$.

Hence, the total energy $E(\mathcal{S}')$ of the λ surfaces in $\mathcal{S}'$ is no larger than that of $\mathcal{S}$, that is, $E(\mathcal{S}') \leq E(\mathcal{S})$. Since $\mathcal{S}$ is an optimal set of surfaces, $\mathcal{S}'$ is an optimal solution. We conclude that there exists an optimal solution consisting of λ non-crossing surfaces with the minimum distance between any two adjacent surfaces no less than 1.

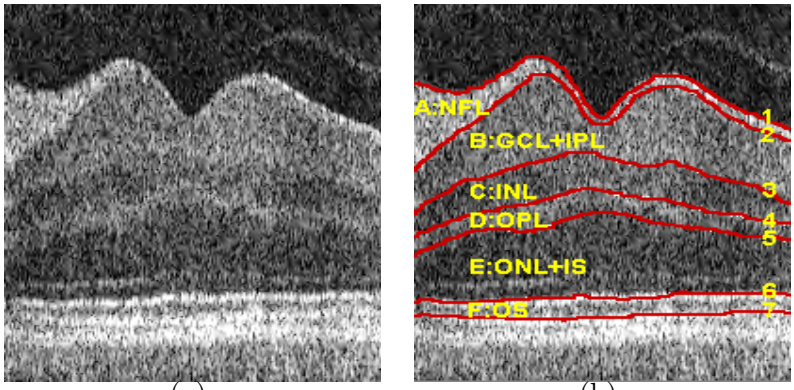
4.5 Optimization

With the constructed graph G, we can find an optimal cut $\mathcal{C}^* = (A^*, \bar{A}^*)$ ($A^* \cup \bar{A}^* = N$) in G, minimizing the total weight of nodes in A^* and the total arc weight of $\mathcal{C}^*$. As described in [82, 18], this optimal cut can be found by solving a maximum flow problem, as follows. Two additional nodes s and t are added into the graph G. Each node $n_i(x,y,z)$ is connected to either the source s by the arc with weight $-w_i(x,y,z)$ if $w_i(x,y,z) < 0$ or the sink t by the arc with weight $w_i(x,y,z)$ if $w_i(x,y,z) > 0$. The computed minimum cut separates the graph G into two parts:

the source set $A^* \cup s$ and the sink set $\bar{A}^* \cup t$. The optimal cut corresponds to the optimal set of λ surfaces in $\mathcal{I}$, which minimizes our energy $E(\mathcal{S})$. The optimal λ surfaces can be recovered as follows. For each target surface i ($i = 1, \ldots, \lambda$), if node $n_i(x, y, z^*)$ has the largest z-coordinate for all nodes belonging to the source set in column $p(x, y)$, then corresponding image voxel $\mathcal{I}(x, y, z^*)$ is on the optimal surface S_i. In this way, the optimal cut in the graph uniquely defines optimal λ surfaces in $\mathcal{I}$.

4.6 Application on Intraretinal Layer Segmentation

Now we demonstrate the applications of our framework on challenging clinical problems. In this section we focus on the automated segmentation of 3-D intraretinal layers in OCT images, which is in urgent need to facilitate quantification of individual retinal layer properties [46]. The seven intraretinal layers are shown in Fig. 4.6(a) and (b). Note that the original images have already been flattened using the method described in [46] so that the surfaces near the retinal pigment epithelium (RPE) became approximately a flat plane. Surfaces 1, 6, and 7 with relatively strong boundaries were simultaneously detected first using the original graph search approach without incorporating the shape and context prior information [46]. Our new approach with the shape and context penalties was then used to simultaneously segment the remaining surfaces 2, 3, 4, and 5, which lack clear boundaries and have substantial interactions in between. Fig. 4.7 shows our main workflow.



(a) (b)

Figure 4.6: Intraretinal Layer in 3-D OCT images. (a) One 2-D slice from the center of an OCT volumetric image. (b) Seven surfaces (labeled 1-7) and six corresponding intralayers on one typical slice. (NFL: nerve fiber layer; GCL+IPL: ganglion cell layer and inner plexiform layer; INL: inner nuclear layer; OPL: outer plexiform layer; ONL+IS: outer nuclear layer and photoreceptor inner segments and OS: photoreceptor outer segments.)

4.6.1 Cost Function Design

4.6.1.1 Boundary cost

For the boundary energy term, we use the gradient-based on-surface cost $c_i(x, y, z)$ for voxel $\mathcal{I}(x, y, z)$ with respect to surface S_i, as reported in [47, 46]. A sobel kernel is used to favor a dark-to-light transition for surface 4 and a light-to-dark transition for surfaces 2, 3, 5. Fig. 4.8 shows typical slices of boundary cost images.

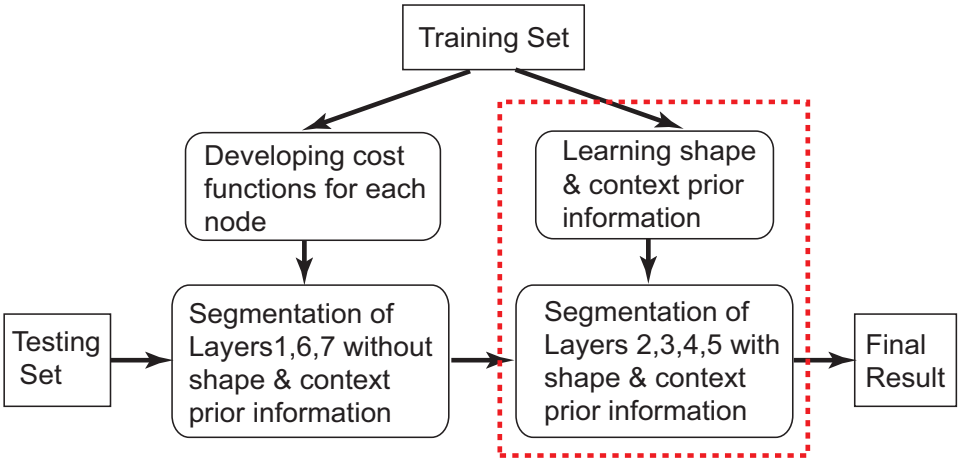
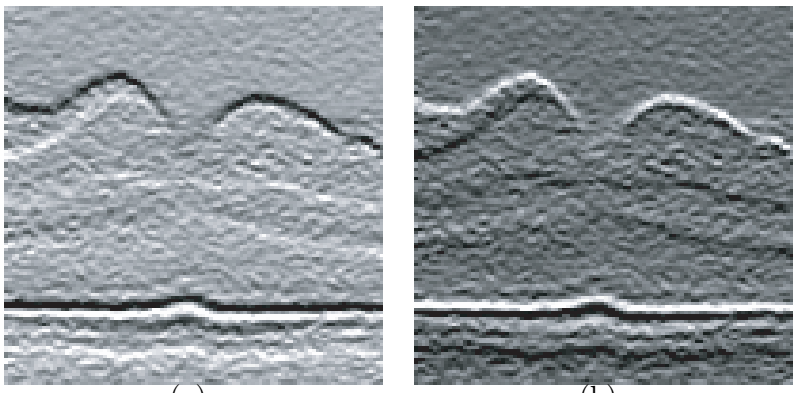


Figure 4.7: Workflow for OCT intraretinal layer segmentation. Here we mainly focus on the second step (indicated by red dashed box) for the segmentation of surfaces 2, 3, 4, 5 using the proposed approach with shape and context priors.

4.6.1.2 Shape prior constraints

To incorporate the proper shape prior information, we learn the shape model from the manual tracing results of the training dataset. The distribution of the shape Δ_{pq}^{i} on the surface i between neighboring columns p and q roughly fits a gaussian model with the mean $\overline{\Delta}_{pq}^{i}$ and the standard deviation σ_{pq}^{i}. To allow for 99% of the shape from column p to column q, the hard shape constraint is set as $|\Delta_{pq}^{i} - \overline{\Delta}_{pq}^{i}| \leq 2.6 \cdot \sigma_{pq}^{i}$. The shape-prior penalty function $f_s(\Delta_{pq}^{i} - \overline{\Delta}_{pq}^{i})$ is designed to penalize the shape deviation between the current shape Δ_{pq}^{i} and the original shape model as follows:

$$f_s(\Delta_{pq}^{i} - \overline{\Delta}_{pq}^{i}) = -\ln e^{-\frac{(\Delta_{pq}^{i} - \overline{\Delta}_{pq}^{i})^2}{2(\sigma_{pq}^{i})^2}} = \frac{(\Delta_{pq}^{i} - \overline{\Delta}_{pq}^{i})^2}{2(\sigma_{pq}^{i})^2} \tag{4.3}$$



(a) (b)

Figure 4.8: Boundary cost image. Low intensity value indicates the small value of the boundary cost. (a) Cost image for surface 4. Gradient-based image with a Sobel kernel favoring dark-to-light transition. (b) Cost image for surface 2, 3, 5. Gradient-based image with a Sobel kernel favoring light-to-dark transition.

4.6.1.3 Context prior constraints

The context prior information is enforced using a similar manner as the shape prior constraints. Let $\overline{\delta}_p^{ij}$ and σ_p^{ij} denote the mean and the standard deviation of the distance between surfaces i and j on column p, respectively. The hard context constraint has the form $|\delta_p^{ij} - \overline{\delta}_p^{ij}| \leq 2.6 \cdot \sigma_p^{ij}$. The context-prior penalty function is set as

$$f_c(\delta_p^{ij} - \overline{\delta}_p^{ij}) = -\ln e^{-\frac{(\delta_p^{ij} - \overline{\delta}_p^{ij})^2}{2(\sigma_p^{ij})^2}} = \frac{(\delta_p^{ij} - \overline{\delta}_p^{ij})^2}{2(\sigma_p^{ij})^2} \tag{4.4}$$

4.6.2 Experimental Methods

4.6.2.1 Data

We used the same data as reported in [46]. Macula-centered 3-D OCT volumes ($200 \times 200 \times 1024$ voxels, with each of $6 \times 6 \times 2$ mm^3) were acquired from the right eyes of 27 normal human subjects. A 3-D OCT volume of the right eye of the first 13 subjects was obtained using one machine for the training dataset. The right eyes of the remaining subjects (14-27) were scanned twice using two different machines. The obtained 28 volumetric datasets were used for performance assessment. For each 3-D volumetric image in the test set, 10 random slices were traced independently by two ophthalmologists. The average of these tracings were used as the gold standard for validation.

4.6.2.2 Parameter setting

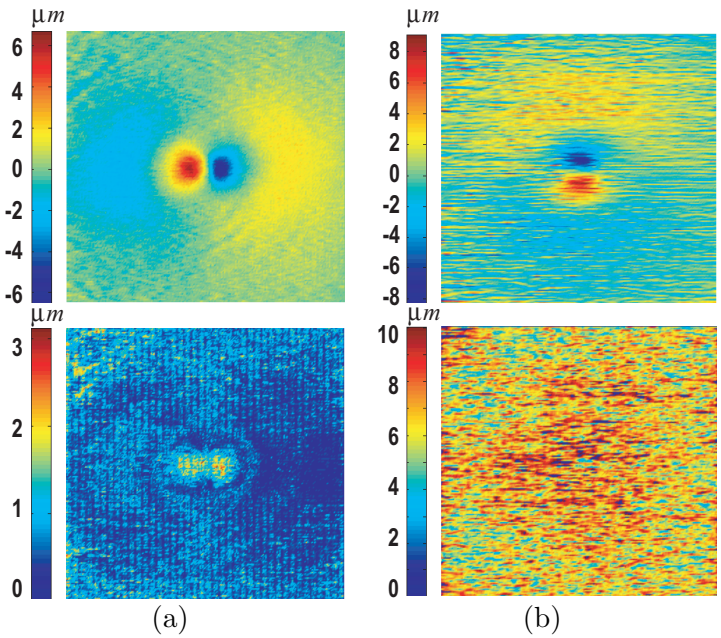
As described in Section 4.2, our energy function contains three terms: the boundary term, the shape-prior term and the context-prior term. The combination of three terms can be described by two parameters α and β as follows:

$$E(\mathcal{S}) = E_{boundary}(\mathcal{S}) + \alpha E_{shape}(\mathcal{S}) + \beta E_{context}(\mathcal{S}). \tag{4.5}$$

In our experiments, two parameters were set as $\alpha = 0.9$ and $\beta = 0.1$, which were identified according to the tests on the training dataset.

4.6.2.3 Shape & context prior model

The shape prior model as well as the context prior model were learned from the training set of 13 OCT volumes. Fig. 4.9 shows the mean and the standard deviation



(a) (b)

Figure 4.9: Visualization of the shape priors for surface 2 learned from the training set. The mean and the standard deviation are illustrated in the first row and the second row, respectively. (a) X-direction. (b) Y-direction.

of the shape priors for surface 2 in two directions (e.g., $\Delta^i_{pq} = S_i(x+1, y) - S_i(x, y)$ for the x-direction and $\Delta^i_{pq} = S_i(x, y+1) - S_i(x, y)$ for the y-direction). Fig. 4.10 illustrates the surface context priors between surface pairs $(2, 3)$, $(3, 4)$ and $(4, 5)$.

4.6.2.4 Validation

The proposed algorithm was carried out on the testing set of 28 volumetric images for validation. The unsigned surface positioning errors were calculated as the distance between the computed surface and the surface of the gold standard (the average surface of two manual tracings by ophthalmologists) in each column

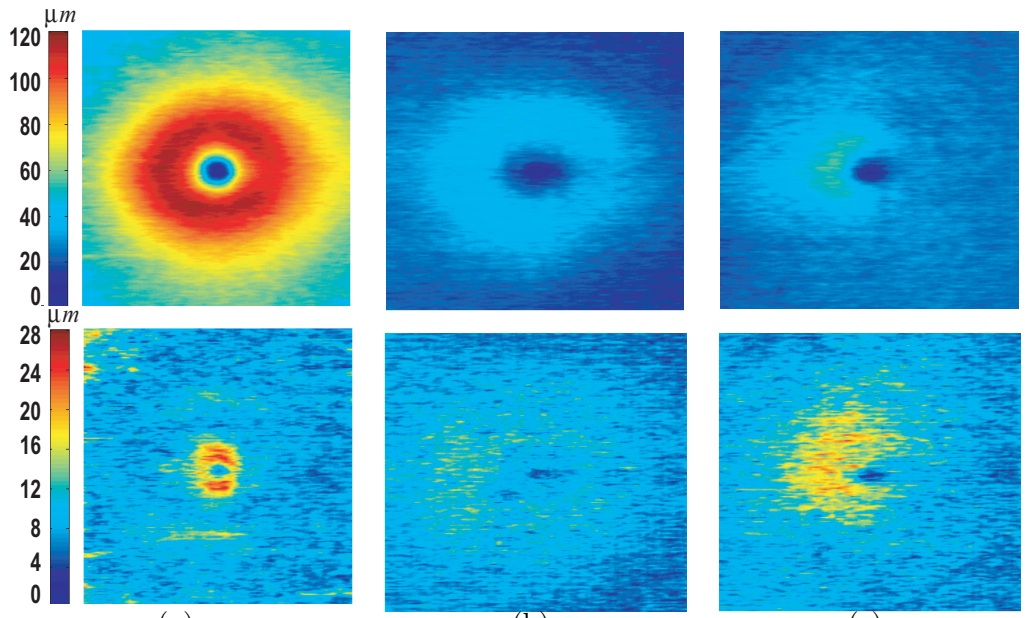


Figure 4.10: Learned surface context priors in the form of the mean (the first row) and the standard deviation (the second row) between (a) surfaces $(2, 3)$; (b) surfaces $(3, 4)$; and (c) surfaces $(4, 5)$.

for quantitative measurement. The results were reported in mean $\pm$ standard deviation fashion. The unsigned surface distance between two manual contours were also computed as the measurement of the inter-observer variability between two ophthalmologists. As comparative validations, the performance of the proposed method were compared with the inter-observer variability and also the results reported in [46], which used the conventional graph search method with only hard constraints. Statistical significance of the observed differences was determined using the Student t-test for which p value of 0.05 is considered significant.

4.6.3 Results

4.6.3.1 Quantitative validation

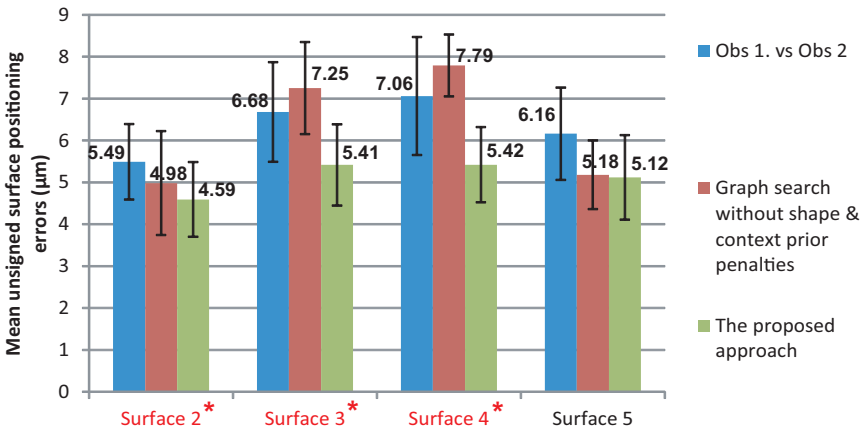


Figure 4.11: Quantitative and comparative performance evaluation based on the unsigned surface positioning errors in 28 volumetric OCT images. Both the results of the original graph search method with only hard constraints (blue) [46] and the results of the proposed method (red) are shown.

The computed unsigned surface positioning error for surfaces 2-5 are summarized in Table. 4.1. For all four surfaces, the resulting errors are significantly smaller than the corresponding inter-observer variabilities ($p < 0.001$).

Fig. 4.11 shows the performance comparison of the proposed approach and the original graph search method used in [46]. Our method produced a significant lower surface positioning errors for surface 2 ($p = 0.01$), surface 3 ($p < 0.001$) and surface

Surface	Algo. vs. Avg. Obs	Obs. 1 vs Obs. 2
2	4.59 ± 0.89	5.49 ± 0.90
3	5.41 ± 0.97	6.68 ± 1.19
4	5.42 ± 0.90	7.06 ± 1.41
5	5.12 ± 1.01	6.16 ± 1.10
Overall	5.14 ± 0.99	6.35 ± 0.92

Mean±SD in μm.

Table 4.1: Summary of the mean unsigned surface positioning errors.

4 ($p < 0.001$). The errors for surface 5 are not significantly different between the proposed algorithm and the original graph search method with only hard constraints.

4.6.3.2 Qualitative results

Qualitatively, the proposed algorithm produced a good segmentation. Fig. 4.12 shows typical segmentation results for surfaces 2, 3, 4, 5 from one dataset with the median overall mean unsigned surface distance error. Fig. 4.13 shows the illustrative results of the proposed algorithm (Fig. 4.13(d)) in comparison with the traditional graph search method using only hard constraints (Fig. 4.13(c)) on one 2-D slice from the 3-D volume. It demonstrates that the proposed method provides a better shape and context control in the presence of the high noise, which leads to an more accurate and smoother segmentation compared with the original graph search method.

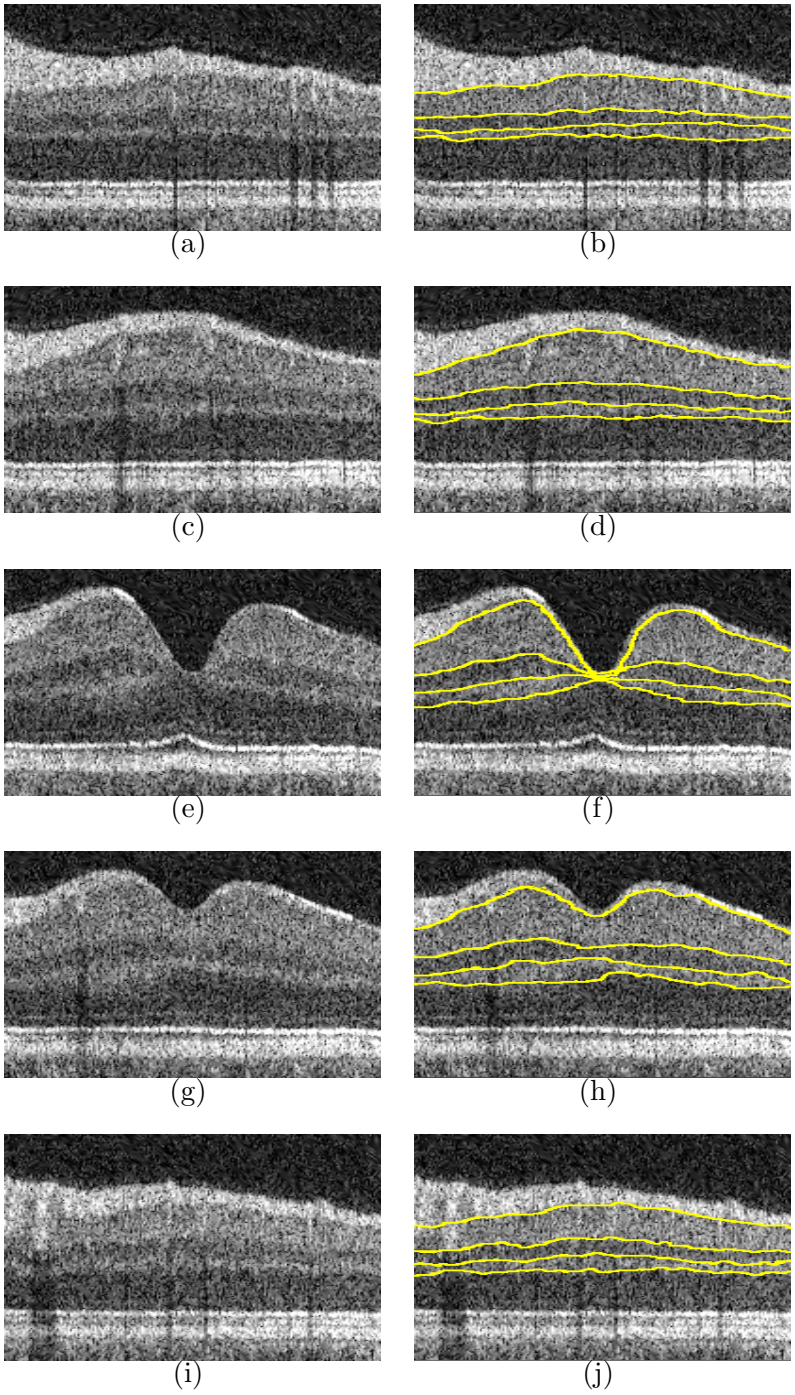


Figure 4.12: Intraretinal layer segmentation results for surfaces 2, 3, 4, 5 shown on five slices selected from one image set (median case according to overall mean unsigned surface distance error). Original slices are shown on the left column and our segmentation results are shown on the right column. (a),(b) Slice 20. (c),(d) Slice 51. (e),(f) Slice 96. (g),(h) Slice 117. (i),(j) Slice 172.

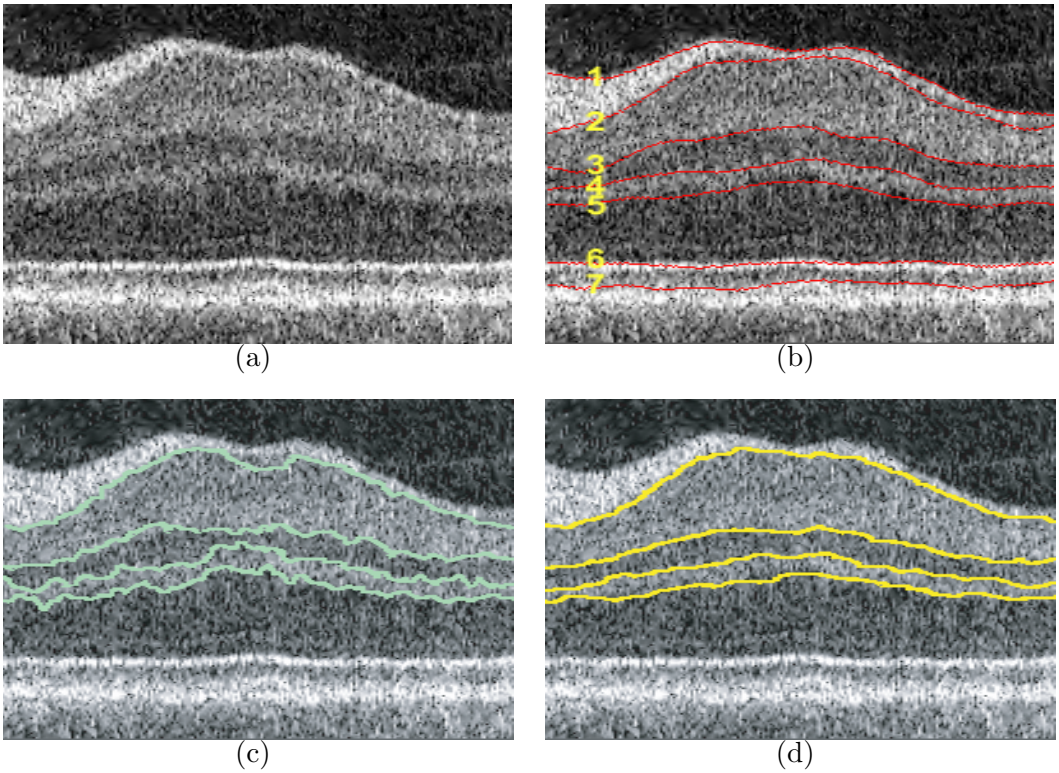


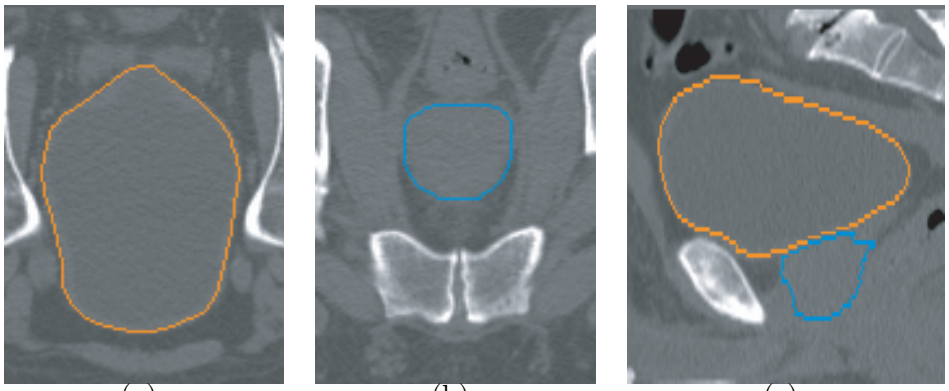
Figure 4.13: Comparative results of intraretinal layer segmentation in 3-D OCT images. (a) One 2-D slice of retinal OCT. (b) Seven manually labeled surfaces (1-7). (c) One example 2-D slice of the graph searching result for surfaces 2, 3, 4, 5 using only hard constraints. (d) The proposed algorithm with shape & context prior penalties.

4.6.3.3 Execution time

Our algorithm was implemented in C++ on a Linux workstation (3GHz, 32GB memory). Current non-optimized implementation required about one and a half hours. The possible way to reduce the processing time includes a multi-scale graph-search approach, as described in [36]. The basic idea is to detect the intraretinal layer

in a relative small area of the highest resolution image, which is defined according to the surfaces segmented in the lower resolution images. Another choice is using a parallel implementation for the maximum-flow algorithm [5].

4.7 Application on Bladder & Prostate Segmentation



(a) (b) (c)

Figure 4.14: Bladder and prostate in 3-D CT images. (a) One 2-D slice of the bladder in transverse view. (b) Prostate in transverse view. (c) Bladder (orange) and prostate (blue) in sagittal view.

In Section 4.6, we have presented how to apply the proposed framework for the segmentation of multiple terrain-like surfaces. In this section, we show how to use our approach to deal with the segmentation of arbitrarily-irregular surfaces. The basic idea is presented as follows. A shape model is first built for each target object based on the prior information from the training datasets, which reflects the approximate

topological structure information of the target surfaces. The graph is then constructed based on this shape model. Multiple constraints (i.e., shape priors constraints, context prior constraints) are incorporated into the graph using the method proposed in Section 4.2. Here we validate our approach for simultaneous segmentation of bladder and prostate in CT images.

In the United States, prostate cancer is one of the most common cancers in men, accounting for about 28% of all newly diagnosed cases [7]. Precise target delineation is critical for a successful 3-D radiotherapy treatment planning for prostate cancer treatment. Automatic segmentation techniques are urgently needed due to large amounts of 3-D image data that require increased time for manually contouring.

The segmentation of pelvic structure is of particularly difficulty. It involves soft tissues that present a large variability in shape and size. Those soft tissues also have similar intensity and have seriously mutual influence in position and shape. Many attempts have been tried in this area, such as registration approach (e.g. [77]), implicit and explicit models (e.g. [14, 19, 44]), and bayesian formulation (e.g. [49]). None of these methods can produce a globally optimal solution with respect to a task-specific objective function. Freedman *et al.* [15] developed an interactive approach based on graph cut method [84]. Their method incorporating soft shape prior allows for a global optimum. However, it focuses on only single object (bladder) segmentation and is at least non-trivial to incorporate mutually interacting surface constraints to simultaneously segment multiple object surfaces.

Here we apply the proposed approach for globally optimal segmentation of

bladder and prostate. Both shape prior information and hard surface context constraints are incorporated using arc-weighted graph representation. Besides, a regional term is also added into our energy function, which encodes the learned appearance information.

Our approach mainly consists of two stages: (a) Initial model building of the objects of interest. An approximate model of target surfaces can be obtained, which gives useful information about the topological structures of the target objects. (b) Accurate delineation using graph optimization based on the mesh of the initial model. The use of the model allows our method to readily incorporate shape and context priors. The workflow is shown in Fig. 4.15.

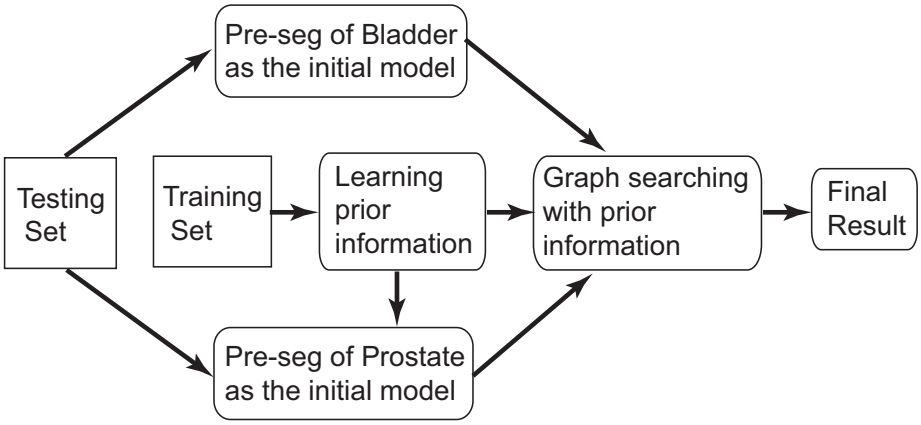


Figure 4.15: Workflow for simultaneous segmentation of the bladder and the prostate

4.7.1 Initial Model Building

A pre-segmentation step is performed to construct an initial model for the target object, which contains the basic topological information. A 3-D geodesic active contour method [78] is conducted for pre-segmentation of the bladder. Three user-defined points are required as an initial input. The prostate shows a much better coherency in shape than bladder. Hence we computed the mean shape of the prostate from the training set of eight 3-D manual segmentations. Then an approximate bounding box of interest for prostate is interactively defined and the obtained mean shape is roughly fitted into the never-before seen CT images using a rigid transformation as the initial model of the prostate. Fig. 4.16 shows one typical example of the pre-segmentation result in three views. Note that the model only serves to provide a basic topological structure information, thus we do not require accurate segmentation at this stage. Overlapping between the model of the bladder and the prostate is also allowed, which can be resolved in the graph optimization step. From pre-segmentation results, two triangulated meshes $M_1(V_1, E_1)$ and $M_2(V_2, E_2)$ are constructed respectively using isosurfacing algorithm (e.g., marching cubes), where V_i $(i \in 1, 2)$ denotes the vertex set of M_i and E_i denotes the edge set of M_i.

4.7.2 Graph Construction

The weighted graph $G_i(N_i, A_i)$ is built from the triangulated mesh M_i as follows. For each vertex $v \in V_i$, a column of K nodes $n_i(v, k)$ is created in G_i, denoted by $p(v)$ (Fig. 4.17(b)). The positions of nodes reflect the positions of corresponding

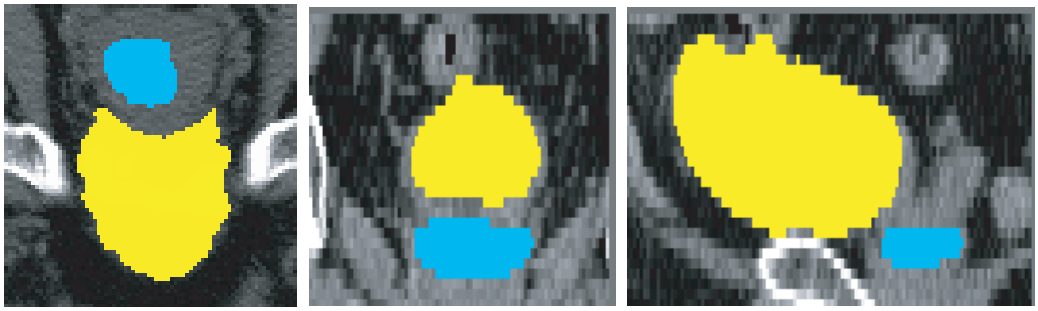


Figure 4.16: Pre-segmentation for bladder (yellow) and prostate (blue) in the transverse (left), coronal (middle) and sagittal (right) views.

voxels in the image domain. The length of the column is set according to the required search range. The number of nodes K on each column is determined by the required resolution. The direction of the column is set as the triangle normal. The nodes on the same column are connected by the intra-column arc from $n_i(v, k)$ to $n_i(v, k-1)$ with infinity weight, as described in Section 4.3. Each column also has a set of neighbors, i.e., if $(v, u) \in E_i$, then $p(v)$ and $p(u)$ are neighboring columns. The feasible surface S_i in the graph G_i is defined as the surface containing exact one node in each column. To avoid the overlapping of two target surfaces, a "partially interacting area" is defined according to the distance between two meshes, which indicates that the two target surfaces may mutually interact each other at that area. To model the context relation, for each column $p(v_1) \in G_1$ in the partially interacting area, these exists a corresponding column $p(v_2) \in G_2$ with the same position in $\mathcal{I}$, and the target surfaces S_1 and S_2 both cut those columns, as shown in Fig. 4.17(b). For implementation, a one-to-one correspondence between two surface meshes needs to be computed on

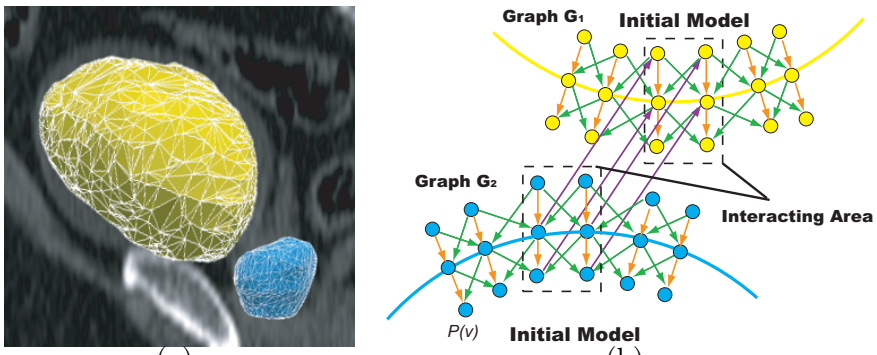


Figure 4.17: Graph construction from initial models. (a) Triangulated meshes for the bladder (yellow) and the prostate (blue) based on initial models. (b) Corresponding graph construction. An example 2-D slice is presented. $p(v)$ represents the column with respect to the vertex v on the mesh. Dots represent nodes $n_i \in G_i$. Two sub-graphs G_1 and G_2 are constructed for the segmentation of the bladder and the prostate, respectively. Note that in the interacting region (dashed box), for each column $p(v_1) \in G_1$, there exists a corresponding column $p(v_2) \in G_2$ with the same position. The inter-surface arcs (purple) between corresponding columns enforce the surface context constraints in the interacting region.

partially interacting area. We project the pre-segmented prostate surface mesh on the interacting region to the mesh of the pre-segmented bladder boundary surface. Then we use the projected mesh patch to replace the original bladder surface mesh on the interacting region. Thus, a one-to-one mesh correspondence on interacting region is established since two new meshes on that area have exactly the same topological structure. In this way, two graphs share the same nodes' positions in the partially

interacting area. The non-overlapping constraint is enforced in the area as the context prior information by adding arcs between corresponding columns using the approach proposed in Section 4.3.

The optimal set S of two surfaces corresponding to the bladder and the prostate can then be found by minimizing the following energy through the constructed graph:

$$E(S) = \sum_{i=1}^{2} E_{boundary}(S_i) + \sum_{i=0}^{2} E_{region}(R_i) + \sum_{i=1}^{2} E_{shape}(S_i) \tag{4.6}$$

The boundary energy term serves as an external force, which drives the mesh towards the best fit to the image data. The shape energy term functions as an internal force, which keeps the shape of the original model and restricts the flexibility of the mesh. To incorporate the learned regional information, an additional region energy term is added in our energy function. Specifically, two surfaces for the bladder and the prostate naturally divide the volume into 3 regions denoted by R_0, R_1 and R_2, which corresponds to the region enclosed by the bladder surface S_1, one between S_1 and the prostate surface S_2 at the partially interacting area, and the region enclosed by S_2, respectively. Our region energy term $E_{region}(R_i)$ reflects the region property of all voxels inside R_i. Note that in this application, a hard non-overlapping constraint is employed as a hard context constraint, which enforces that the segmented bladder and the prostate will not overlap with each other. No soft context prior penalties are introduced in our energy.

4.7.3 Cost Function Design

Cost function design plays an important role in accurate surface delineation. In our method, four kinds of cost function are involved.

4.7.3.1 Boundary cost design

For each node $n_i(v,k) \in G_i$, we have an on-surface cost $c_i(v,k)$ with respect to surface S_i, which is inversely related to the possibility that the corresponding voxel belongs to the target surface S_i. The boundary energy term $E_{boundary}(S_i)$ is defined as the summation of the on-surface costs associated with all nodes on surface S_i, i.e., $E_{boundary}(S_i) = \sum_{n_i \in S_i} c_i(v,k)$.

In soft-tissue segmentation, there is often no clear boundary information between adjacent tissues. The intensity gradient at the interface of adjacent tissues is also low. To overcome the difficulty, we incorporate the object class-uncertainty information into our cost function. Given *a prior* knowledge of intensity probability distributions of each object, the class-uncertainty can be computed for each intensity value, which yields the uncertainty level of the classification. Suppose the image is divided into m objects, denoted by o_i, $i \in [0, 1, ..., m-1]$. For any given node $n_i(v,k)$ with corresponding image intensity $g(v,k)$, the posteriori probability $p(o_i|g)$ is obtained from the training set using Bayes rules. The uncertainty measure for the classification that $n_i(v,k)$ with an associated intensity of g falls into any object is the entropy of all posteriori probability values, with the form: $h(g) = \sum_i -p(o_i|g)log(p(o_i|g))$. As demonstrated in [60, 59], intensities with high class uncertainty measures tend

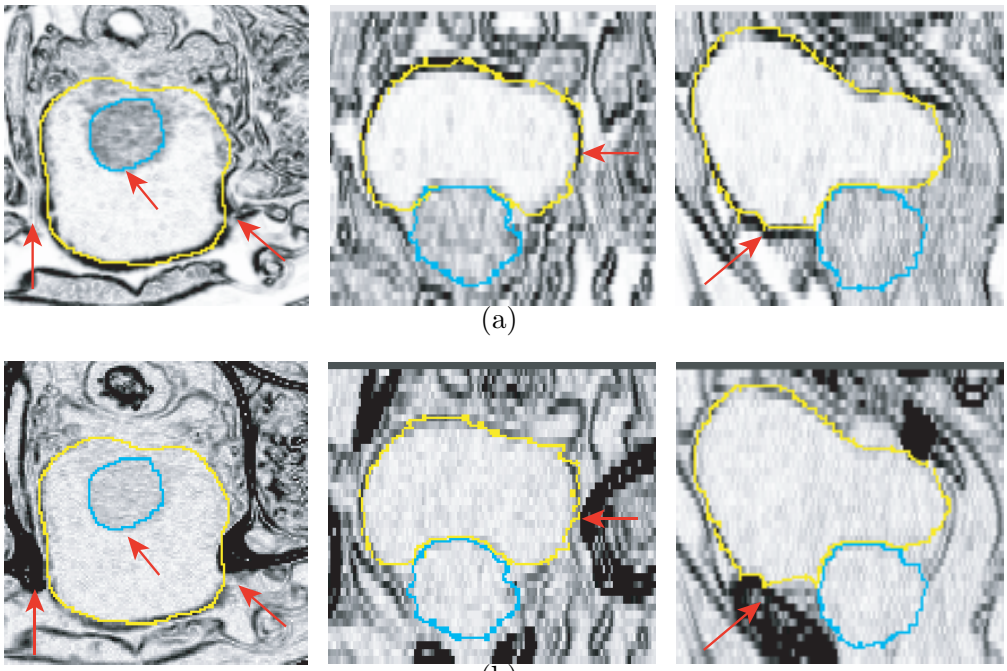


Figure 4.18: Boundary cost in transverse (left), coronal (middle) and sagittal (right) views. Low intensity value indicates the small value of the boundary cost. The manual contours of bladder and prostate are shown in yellow and blue, respectively. (a) Typical image slices of boundary cost combining both gradient information and class uncertainty information. (b) Typical image slices of boundary cost using only gradient information. The red arrows indicate the improved cost by incorporating class uncertainty information.

to appear close to object boundaries rather than inside homogeneous regions. Combined with the gradient information, the uncertainty measure is expected to provide an enhanced on-surface cost at the locations absent a clear boundary. Let $\nabla(v, k)$

denote the gradient magnitude at node $n_i(v,k)$, our boundary cost function has the following form:

$$c_i(v,k) = -(\nabla(v,k) + \alpha h(g(v,k))) \tag{4.7}$$

where α is a constant parameter. Here we set $\alpha = 0.5$ according to the experiments on the training set. Fig. 4.18(a) shows one example slice of the boundary cost combining both gradient information and class uncertainty information. Fig. 4.18(b) is the result only using gradient information.

4.7.3.2 Region cost design

The graph search framework allows easy incorporation of region information by assigning proper weights for the graph nodes. An additional region energy term is added into the energy function. As described in Section 4.7.2, R_0, R_1, and R_2 denote the region enclosed by the bladder surface S_1, one between S_1 and the prostate surface S_2 at the partially interacting area, and the region enclosed by S_2, respectively. For each node $n_i(v,k)$, the in-region cost associated with R_i is assigned as $c_{R_i}(v,k)$ $(i = 0,1,2)$. The region energy term is then defined as the following form: $E_{region}(R_i) = \sum_{n_i(v,k)\in R_i} c_{R_i}(n_i(v,k))$. The in-region weight of node $n_i(v,k)$ in graph G_i , denoted by $w_{R_i}(v,k)$, is assigned such that the total weight of node in the source set of the graph G equals to the region energy term $\sum_{i=0}^{2} E_{region}(R_i)$ (with a constant difference)[47]:

$$w_{R_i}(v,k) = c_{R_{i-1}}(n_i(v,k)) - c_{R_i}(n_i(v,k)) \tag{4.8}$$

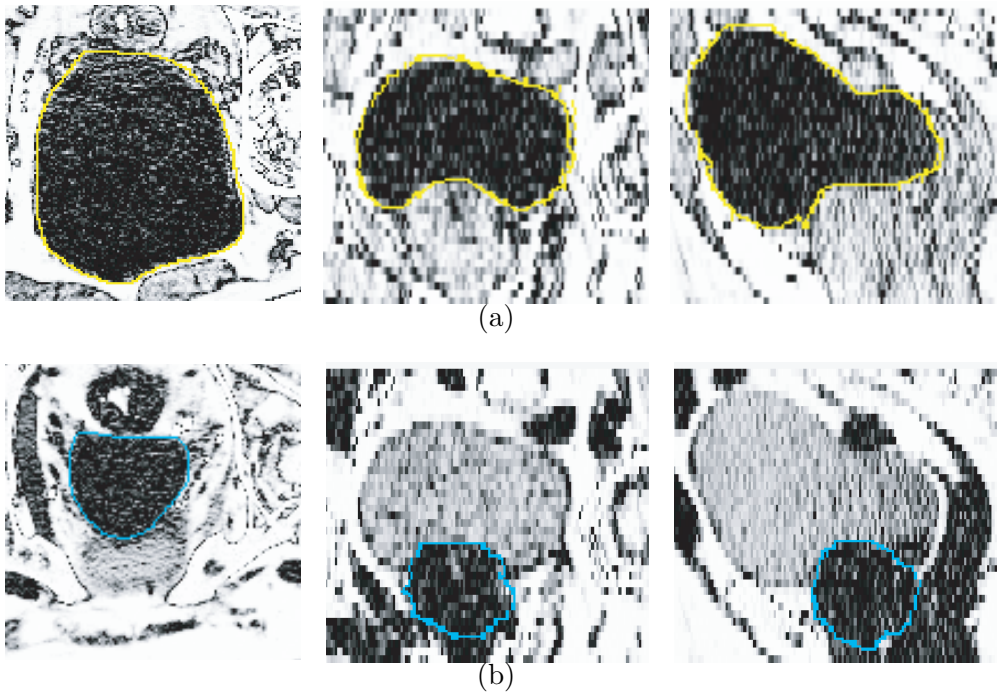


Figure 4.19: Region cost in transverse (left), coronal (middle) and sagittal (right) views. Low intensity value indicates the small value of the region cost. The manual contours of bladder and prostate are shown in yellow and blue, respectively. (a) Region cost of the bladder. (b) Region cost of the prostate.

For region cost design, the posterior probability learned from the training set is used with the form:

$$c_{R_i}(n_i(v,k)) = -p(n_i(v,k) \in R_i | g(v,k)), i \in \{0,1,2\} \tag{4.9}$$

Fig. 4.19 shows one example slice of region cost in three views.

4.7.3.3 Cost design for shape prior penalties

As described in Section 4.2.1, the shape prior penalties serve to keep the topology of the original shape model. Specifically, for any pair of neighboring columns p and q in $\mathcal{I}$, the shape prior penalties of surface S_i are set as $f^i_{pq}(S_i(p) - S_i(q))$, where f^i_{pq} is a convex function penalizing the shape changes of S_i on p and q.

For prostate and bladder segmentation, a second order shape prior penalty is employed with the form: $f(h) = \beta \cdot h^2$, where β is a constant parameter learned from the training set. In this project $\beta = 5$.

4.7.3.4 Context prior costraints

A hard surface distance constraint is incorporated to avoid the overlapping of two surfaces in the interacting area. Let $\delta^{12}_p = S_1(p) - S_2(p)$ denote the distance between the surface of the bladder (S_1) and the surface of the prostate (S_2) on column $p \in R$, where R denotes the possible interacting area obtained from the initial model. WLOG, the direction of intra-column arcs in the graph is set such that S_1 would be the upper surface and S_2 would be the lower surface in the interacting area. Then the hard context prior constraint is set as $\delta^{12}_p > 0$.

4.7.4 Experimental methods

4.7.4.1 Data

43 3-D CT images from different patients with the prostate cancer were employed for our experiments. The image size ranges from $80 \times 120 \times 30$ to $190 \times 180 \times 80$ voxles. The image spacing resolution ranges from $0.98 \times 0.98 \times 3.00\ mm^3$ to $1.60 \times$

$1.60 \times 3.00\ mm^3$. Out of 43 volumes, 8 were randomly selected as the training data and our segmentation was performed on the remaining 35 datasets. The contours of the bladder and the prostate for each dataset were manually traced by one clinical expert as the gold standard.

4.7.4.2 Parameter setting

For graph construction, the length of the column was set as $[-10, 10]$ voxels along the normal direction of the initial surfaces, which was determined by the observed inaccuracies of the initial surfaces on the training set. The number of nodes on each column was set as 21. In the on-surface cost function Eq. (4.7), the coefficients α for the class uncertainty term was set as 0.5. The coefficient β for the shape prior penalty function was set as 5. The parameters were tuned according to the tests on the training dataset.

4.7.4.3 Validation

The proposed algorithm was carried out on the testing set of 35 3-D CT images from different patients for validation. For surface distance error measurement, the unsigned surface positioning errors were calculated between the computed surface and the expert-defined contour. The results were reported as mean ± standard deviation in millimeters. For volumetric error measurement, the Dice similarity coefficient (DSC) was computed using $D = 2|V_m \bigcap V_c|/(|V_m| + |V_c|)$, where V_m denotes the manual volumetric result and V_c denotes the computed result.

$Surface$	DSC	$Mean$	$Maximum$
Bladder	0.93	0.72±0.15	4.80±1.71
Prostate	0.81	0.94±0.22	5.73±1.83

Mean±SD in mm.

Table 4.2: Overall quantitative results for the mean and the maximum unsigned surface positioning errors.

4.7.5 Results

4.7.5.1 Quantitative validation

The dice coefficients (DSC) and the unsigned surface positioning errors are shown in Table. 4.2. Our algorithm generally returned a high-quality segmentation (i.e., DSC 0.93 for the bladder, 0.81 for the prostate, unsigned surface positioning error 0.72 mm for the bladder and 0.94 mm for the prostate) considering the difficulty of the segmentation of the pelvic structure.

4.7.5.2 Qualitative results

The illustrative results in three views are displayed in Fig. 4.20(a)-(d). The 3-D representation is shown in Fig. 4.20(f). From all views, the proposed algorithm produced a very good delineation of both the bladder and the prostate in the 3-D space. The shape prior constraints succeeded to keep the original topological structure of the target organs. No overlapping of the bladder and the prostate is found due to the enforcement of the context constraints. Fig. 4.21 shows the comparison of

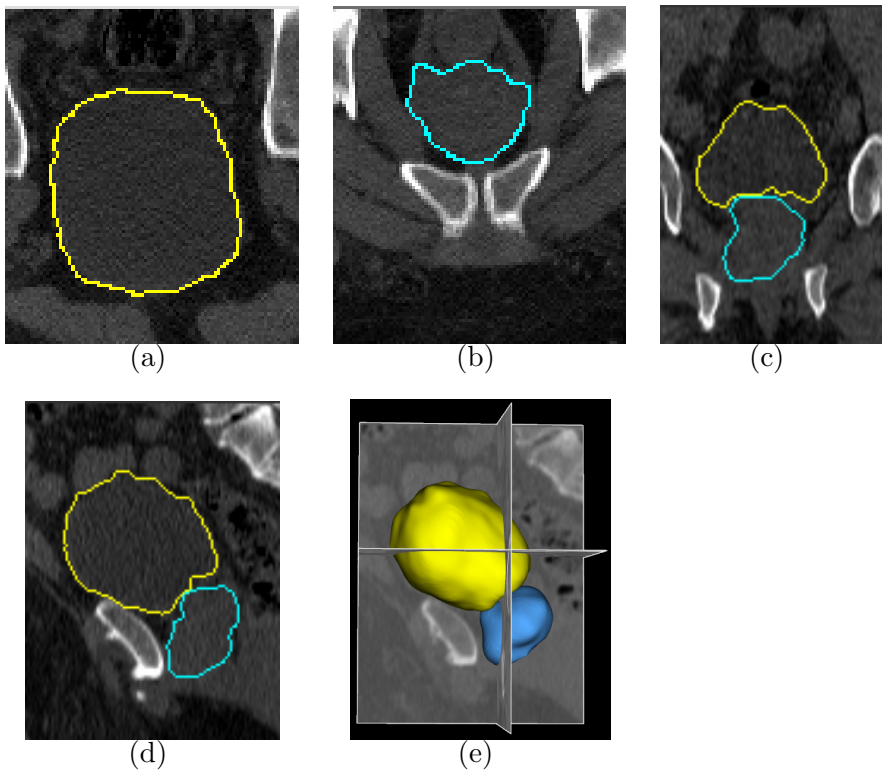


Figure 4.20: Typical slices of the simultaneous segmentation result of the bladder (yellow) and the prostate (blue) in 3-D CT images. (a),(b) Transverse view. (c) Coronal view. (d) Sagittal view. (f) 3-D representation of the segmentation result.

simultaneous bladder-prostate segmentation results with and without shape prior term. Our approach generally produced a smoother result and had a better local shape control.

4.7.5.3 Execution time

The execution time on a WinXP PC (2.13 × 2GHz, 2GB memory) was approximately 8 minutes per dataset, including both the pre-segmentation step and the

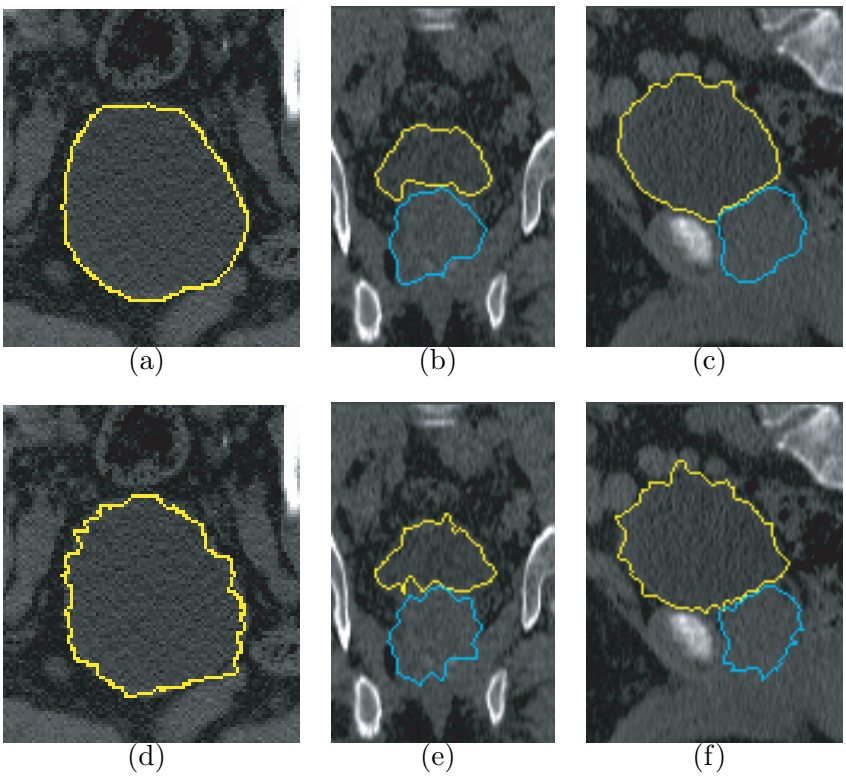


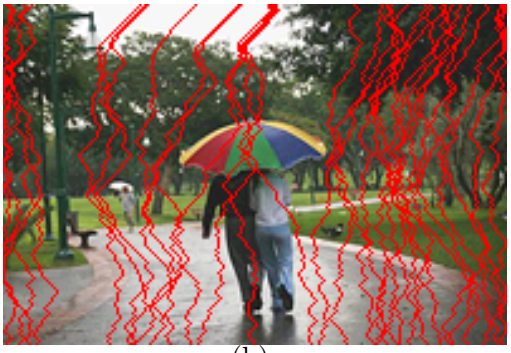
Figure 4.21: Comparison of simultaneous segmentation of the bladder (yellow) and the prostate (blue) in transverse (left), coronal (middle) and sagittal (right) views. (a)-(c) Results with shape prior penalty. (d)-(f) No shape prior penalty is used. All other parameters and cost functions are exactly the same.

graph search step.

4.8 Application for Image Resizing

In recent years, there has been increasing interest in image resizing due to its relevance to image display on devices with different resolutions [67, 16]. Our main idea for image resizing is similar to that used in [16]. The image resizing problem

(a)



(b)

(c)

Figure 4.22: Image resizing example using our graph search method. (a) Original image. (b) The detected surfaces using our approach. (c) Resizing result.

is formulated as an optimization problem, which seeks multiple surfaces whose total energy is minimized. By removing or inserting those surfaces, content-aware image

resizing is achieved. Fig. 4.22 shows one typical example.

The framework reported in [16] only used the node-weighted graph representation without incorporating shape prior penalties, which may cause artifacts due to the discontinuity between the neighboring columns. Here we use our algorithm in Chapter 4 to resolve that problem. The shape prior penalty is incorporated to penalize the shape deformation between the neighboring columns.

4.8.1 Methods

Our graph-based multi surfaces segmentation approach is employed to solve the image resizing problem. In this application, we consider the task of the image resizing in 2-D, which can be directly extended to the resizing for volumetric image or video in high dimensional spaces.

4.8.1.1 Problem formulation

Consider a 2-D image $\mathcal{I}(X, Y)$ of size $X \times Y$. For each x, the voxel subset $\{\mathcal{I}(x, y)|0 \leq y < Y\}$ forms a column parallel to the y-axis, denoted by $p(x)$. Each column has a set of neighborhoods for a certain neighbor setting, e.g., four-neighbor relationship. The target here is to find λ terrain-like surfaces, which intersect each column $p(x)$ at exactly one time. The image resizing is achieved by removing or inserting those surfaces such that the new image size in y direction will be $Y - \lambda$ for size reducing or $Y + \lambda$ for size increasing.

Similar like the original graph search framework, a boundary-based cost $c(x, y)$ is assigned to each pixel $\mathcal{I}(x, y)$, which is set according to the likelihood that the pixel

contains important information. To preserve the spatial coherence, a shape-prior penalty is assigned between neighboring columns, which serves as a soft smoothness control. The surface change is penalized, which ensures the surface continuity to keep the consistency of the target image.

4.8.1.2 Graph construction

Note that in image resizing problem, we require that the sought surfaces may cross each other, but share no common voxels. As proved in Section 4.4, if we apply the same cost functions and the same shape prior functions to all surfaces – which is reasonable in the image resizing problem, this problem can be transformed into the problem of finding an optimal solution for detecting multiple non-crossing surfaces. The minimum distance between any two adjacent surface is no less than 1, which can be enforced through hard context constraints. Thus our algorithm is applicable for solving this problem.

A directed graph G containing λ subgraphs $G_i(N_i, A_i)$ $(i = 1, 2, \ldots, \lambda)$ is constructed. As described in Section 4.3, in which every node $n_i(x, y)$ represents exactly one pixel $\mathcal{I}(x, y)$. Intra-column arcs, inter-column arcs and inter-surface arcs are added, which enforce the monotonicity, shape penalties and surface context constraints, respectively. The target λ surfaces can be found by solving a maximum flow problem in low-order polynomial time.

(a) (b)

Figure 4.23: Boundary cost image. (a) Original image. (b) The computed gradient-based cost image. Higher intensity value indicates that the corresponding pixel contains more important information.

4.8.1.3 Cost function design

For boundary cost, we simply use image gradient to capture the spatial features of the image. Let $\nabla(x, y)$ denote the gradient magnitude at pixel $\mathcal{I}(x, y)$, our boundary cost function is defined as $c(x, y) = \nabla(x, y)$. One typical cost image is shown in Fig. 4.23.

For shape prior penalty, a second order penalty function is employed with the form: $f(h) = \beta \cdot h^2$, which penalizes the change of the surface topology: $h = S_i(p) - S_i(q)$ for surface S_i between neighboring columns p and q. Here β is a constant parameter selected empirically. In this project $\beta = 10$.

A hard surface context constraint is incorporated to make sure that the minimum distance between two surfaces is 1. Let $\delta_p^{ij} = S_1(i) - S_2(j)$ denote the distance between two adjacent surfaces i and j. WLOG, the direction of intra-column arcs in

the graph is set such that S_i would be the upper surface and S_j would be the lower surface. Then the hard context prior constraint is set as $\delta_p^{ij} > 0$.

4.8.2 Experiments and Results

Fig. 4.24 shows our result on three 2-D images. Improvements indicated by arrows can be appreciated from the incorporation of the shape prior penalty.

4.9 Discussion

4.9.1 Method Properties, Novelty and Generality

In this chapter, we propose a novel approach for the segmentation of multiple surfaces with both shape and context prior information. Our method achieves several important goals. First, the approach allows the simultaneous segmentation of multiple objects. Second, the proposed energy function incorporates both shape prior information and context prior information through a set of convex functions. Third, our method has the global optimality guarantee. To the best of our knowledge, this is the first method that fulfills these three aims at the same time. Our approach can be directly extended to n-D. The possible extension includes 4-D object segmentation with object motion over time, co-segmentation of objects in multi-modality images, and so on. Our framework also allows easy incorporation of region information by assigning proper weights to the nodes, as described in [47].

(a) (b) (c)

Figure 4.24: Image resizing using graph search method. The first row is the original image. The second row is the graph search result without the shape prior penalty. The third row is the result with the convex shape prior penalties. Typical improvements are pointed out by arrows.

4.9.2 Limitation

The proposed framework can be easily applied for many image segmentation tasks. However, it also has several limitations. The major limitation lies in the fact

that it is non-trivial for our method to deal with the region of complicated topology without an approximate segmentation of the region's surface. For segmentation of objects with irregular topology, the quality of the pre-segmentation result may influence the final graph search segmentation. If the pre-segmented result is far away from the actual surface or the initial model does not reflect the true topological structure of the target object, the column of the constructed graph may not intersect with the desired surface, which will result in the failure of the segmentation. Possible improvements include following aspects. First, the pre-segmentation step for the objects with complicated topology can be improved. In [91], an AdaBoost localization followed by an iterative graph search approach is used for the pre-segmentation of the individual bone surfaces, which provides a sufficiently accurate results for the graph construction. Second, the construction of the graph would also be improved. In the current approach, the column is constructed following the normal direction of the pre-segmented surface. If the target object has a complicated topology, these columns may intersect with each other. To resolve this problem, Yin *et al.* [91] uses electric field line based method for column construction in the graph. In addition, the gradient vector flow (GVF) is employed for graph building to alleviate the problem [70].

Another limitation is the representation of the prior information. In our framework, a set of convex functions are employed to enforce the shape prior and the context prior penalties. If the distribution of the prior shape information or the prior context information can not be represented by a convex function set, we will have difficulty

to encode them into our framework. Also, our current framework only enforces the local prior information, e.g., local shape prior between neighboring columns. How to encode both local prior information and global constraint is a challenging task. Some attempts have been made towards the direction. In [70] and [26], the graph search framework is combined with the active shape model to provide a better global shape control.

4.10 Conclusion

We present a general framework for simultaneous segmentation of multiple surfaces. The prior shape and context information is incorporated into the energy function through a set of convex functions. An arc-weighted graph representation is employed. The optimal solution can be achieved by solving a maximum flow problem in a low order polynomial time. Compared with other graph-based methods, our approach provides more local and flexible control of the prior shape and context information. We also prove that our algorithm can handle the detection of multiple crossing surfaces with no shared voxels.

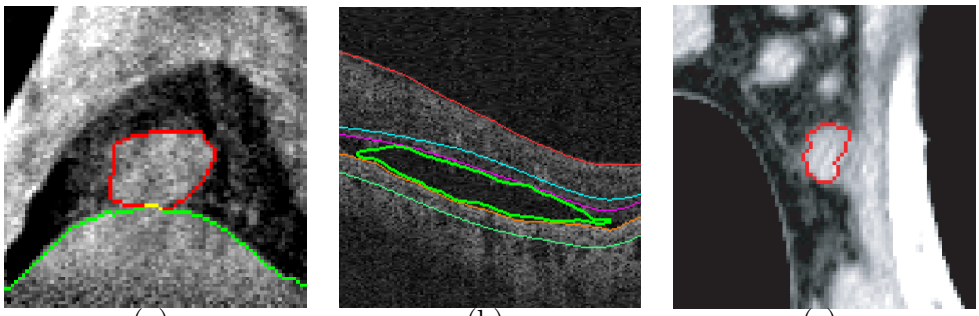
The proposed algorithm was validated on three proof-of-concept applications for intraretinal layer segmentation in OCT images, bladder-prostate segmentation in CT images and image resizing problem. The result demonstrated the applicability of the proposed algorithm.

CHAPTER 5
OPTIMAL SURFACE-REGION SEGMENTATION WITH CONTEXT CONSTRAINTS

5.1 Introduction

As described in previous chapters, accurate segmentation of medical images is a challenging job in the presence of multiple target objects with weak boundaries, similar intensity information and serious mutual interaction among each other. The incorporation of context information between different objects would be a great help for a successful segmentation. In this chapter, we consider a special structure consisting of mutually interacting terrain-like surfaces and regions of "arbitrary shape". Here "arbitrary shape" means that the target region itself does not have any specific preferred shape. Mutual interaction exists between the pairs of the terrain-like surfaces and the regions in a way that their topology and relative positions are usually known beforehand. Incorporating these interrelations into the segmentation plays a significant role for accurate segmentation when target surfaces and regions lack clear edge and have similar intensity distribution. Fig. 5.2(a),(b) and (c) show some typical medical imaging examples.

The main idea is to combine the advantages from both the graph cut method [84] and our graph search framework. As described in Section 1.3.2.1, the graph cut method formulates the segmentation problem as a discrete optimization problem. An energy function encoding both boundary information and region information is minimized over a finite set of integer-valued variables, which are associated with graph



(a) (b) (c)

Figure 5.1: Typical examples of interacting terrain-like surfaces and regions of arbitrary shape. (a) One example slice of the lung tumor (red) in megavoltage cone-beam CT. In this example the tumor is attached/adjacent to the diaphragm (green). (b) Fluid-filled region (green) in retinal optical coherence tomography (OCT). The fluid region is surrounded by intraretinal layers. (c) Lymph node (red) in X-ray CT data. The lymph node is connected to surrounding soft tissues.

nodes corresponding to image pixels. The graph cut method is topologically flexible, which means that it does not have any preferred shape. On the other hand, our graph search method tends to find a globally optimal segmentation of multiple *terrain-like surfaces*. As proved in Chapter 4, our framework allows easy incorporation of shape prior information and provides a better shape control. To make use of benefits from both methods, following strategy is employed. For regions of arbitrary topology, we construct a corresponding graph following the graph cut method making use of its topological flexibility. For terrain-like surfaces, the graph is built based on the graph search approach, which requires no initial seeds and has shape constraint. The key is to introduce an additional context term in the energy function, which can be implemented by adding inter-graph arcs between the pairs of different types of graphs, enforcing known context information between the target terrain-like surfaces and the regions of arbitrary topology. Then a globally optimal solution can be obtained by solving a *single* maximum flow problem, yielding the simultaneous segmentation of surfaces and regions.

The rest of the chapter is organized as follows. In next section, we propose our energy function. Section 5.3 gives a detailed description of the context term included in our energy. Section 5.4 shows the corresponding graph construction. In Section 5.5, we apply the proposed framework for simultaneous segmentation of the lung tumor and adjacent surfaces in Mega-Voltage Cone Beam CT (MVCBCT). Section 5.6 gives the conclusion.

5.2 Energy Function

To simplify the presentation and to ease the understanding of our method, let us first consider the task of detecting one terrain-like surface and one region of arbitrary topology with mutual interaction between each other. Note that the same principles used for this illustration are directly applicable to multiple pairs of surfaces and regions with interactions between those two kinds of targets.

Given a volumetric image $\mathcal{I}(X, Y, Z)$ of size $X \times Y \times Z$. For each (x, y) pair, the voxel subset $\{\mathcal{I}(x, y, z)|0 \leq z < Z\}$ forms a column parallel to the z-axis, denoted by $p(x, y)$. Each column has a set of neighborhoods for a certain neighbor setting $\mathcal{N}_c$. As described in Chapter 2, a *terrain-like surface* of particular interest, denoted as S_T, is the surface that intersects each column $p(x, y)$ at exactly one voxel, which can be defined as a function $S_T(x, y)$, mapping each (x, y) pair to its z-value. The target region of arbitrary shape, denoted as R_A, includes all the voxels inside the region. Fig. 5.2(a) shows one example 2-D slice from a 3-D image.

To explain the employed cost function, let us start with the energy terms used for detecting the terrain-like surface, which are similar in form to those described in Chapter 4. Suppose an boundary cost $c(x, y, z)$ is assigned to each voxel $\mathcal{I}(x, y, z)$, which is inversely related to the likelihood that the desired surface S_T indeed contains this voxel. For any pair of neighboring columns $(p, q) \in \mathcal{N}_s$, a convex function penalizing the surface shape change of S_T on p and q is expressed as $f_{pq}(S_T(p) - S_T(q))$.

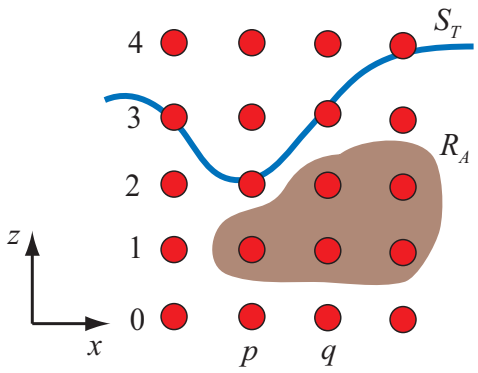


Figure 5.2: Example 2-D slice from a 3-D image. Terrain-like surface S_T is shown in blue and region of arbitrary shape R_A is shown in brown. $S_T(p) = 2$, $S_T(q) = 3$.

Then the energy term E_{gs} takes the form:

$$E_{gs}(S_T) = \sum_{\mathcal{I}(x,y,z)\in S_T} c(x,y,z) + \sum_{(p,q)\in\mathcal{N}_s} f_{pq}(S_T(p) - S_T(q)) \ . \tag{5.1}$$

For segmentation of the target region R_A, we employ the well-known binary graph cut energy [84]. Let l denote the binary variables assigned for each voxel, indexed as l_v over voxels $\mathcal{I}_v \in \mathcal{I}$. In our notation, $l_v = 1$ means that $\mathcal{I}_v$ belongs to the target region R_A and $l_v = 0$ means that $\mathcal{I}_v$ belongs to the background. The graph cut energy E_{gc} is expressed as:

$$E_{gc}(R_A) = \sum_{\mathcal{I}_v\in\mathcal{I}} D_v(l_v) + \sum_{(\mathcal{I}_i,\mathcal{I}_j)\in\mathcal{N}_v} V_{ij}(l_i, l_j) \ , \tag{5.2}$$

where D_v is the data term measuring how well the label l_v fits the voxel $\mathcal{I}_v$ given the image data, $\mathcal{N}_v$ defines the neighboring relationship between voxels, and the boundary energy term $V_{ij}(l_i, l_j)$ is the penalty of assigning the neighboring voxels $(\mathcal{I}_i, \mathcal{I}_j)$ to labels l_i and l_j, respectively.

As mentioned above, incorporation of known interrelations between terrain-like surface S_T and object R_A plays a key role for accurate segmentation. To enforce a priori object–surface context information, we add a context term $E_{context}(S_T, R_A)$ to our energy function and the energy function takes the form:

$$E(S) = E_{gs}(S_T) + E_{gc}(R_A) + E_{context}(S_T, R_A) \ . \tag{5.3}$$

Our objective is to find the optimal set $S = \{S_T, R_A\}$ such that the above energy is minimized.

5.3 Incorporation of Geometric Context Information

In this section, we specify the geometric context information incorporated in our framework between the target surface S_T and the region R_A, and show how to enforce as the context energy term $E_{context}(S_T, R_A)$. We start with the case that the region R_A tends to be lower than the terrain-like surface S_T with a given distance d. For any violating voxels in R_A, a penalty is given. More specifically, let $z(\mathcal{I}_v)$ denote the z coordinate for voxel $\mathcal{I}_v$. $S_T(p)$ is the z-value for the surface S_T on column p, representing the "height" of the surface on that column. Then for any voxel $\mathcal{I}_v \in p$, if $\mathcal{I}_v \in R_A$ and $S_T(p) - z(\mathcal{I}_v) < d$, a penalty w_v is given (Fig. 5.3(a)). Our context energy term takes the form:

$$E_{context}(S_T, R_A) = \sum_{p \in \mathcal{I}} \sum_{\substack{\mathcal{I}_v \in p \\ S_T(p) - z(\mathcal{I}_v) < d}} w_v l_v \ . \tag{5.4}$$

For the constraint that R_A is a priori expected to be positioned "higher" than the terrain-like surface S_T, a similar formulation is employed. For any voxel $\mathcal{I}_v \in p$,

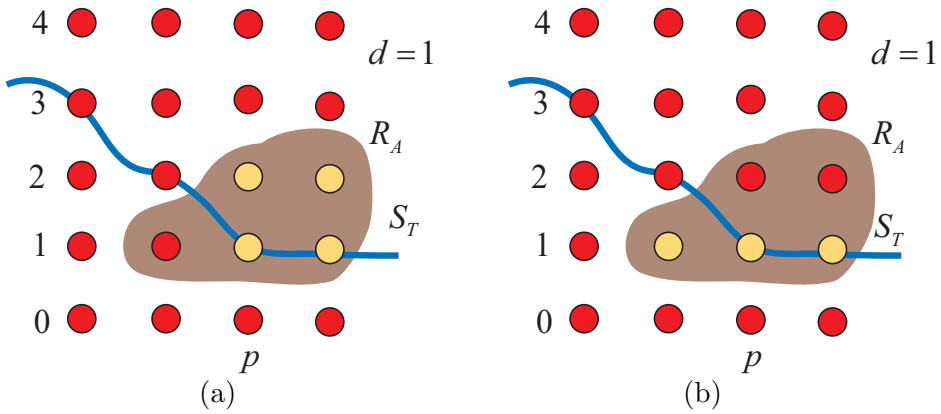


Figure 5.3: Example 2-D slice from a 3-D image. (a) Incorporation of context constraint that the region R_A tends to be positioned "lower" than the terrain-like surface S_T. For voxels $\mathcal{I}_v \in R_A$ and $S_T(p) - z(\mathcal{I}_v) < d$ (yellow voxels), a penalty w_v is given; d is set as 1. (b) Context constraint that R_A tends to be positioned "higher" than S_T. For voxels $\mathcal{I}_v \in R_A(l_v = 1)$ and $z(\mathcal{I}_v) - S_T(p) < d$ (yellow voxels), penalty w_v is given; d is set as 1.

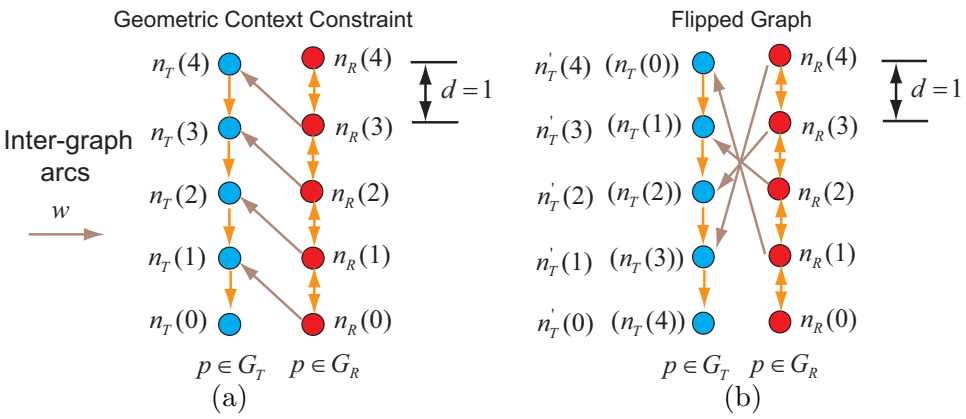


Figure 5.4: Graph construction with context constraints. (a) Incorporation of geometric context constraints by adding inter-graph arcs (brown arrows). Region R_A tends to be lower than the terrain-like surface S_T with a distance of $d = 1$; $n_T(i)$ and $n_R(i)$ correspond to the same voxel in original image. (b) Flipped graph for constraints that R_A tends to be higher than surface S_T with a distance of $d = 1$. $n'_T(i)$ and $n_R(Z - i - 1)$ correspond to the same voxel in the original image.

if $\mathcal{I}_v \in R_A(l_v = 1)$ and $z(\mathcal{I}_v) - S_T(p) < d$, a penalty w_v is given (Fig. 5.3(b)). Our context energy term then takes the form:

$$E_{context}(S_T, R_A) = \sum_{p \in \mathcal{I}} \sum_{\substack{\mathcal{I}_v \in p \\ z(\mathcal{I}_v) - S_T(p) < d}} w_v l_v \ . \tag{5.5}$$

5.4 Graph Construction

Our graph transformation scheme formulates the energy minimization problem as a single computation of maximum flow in the graph.

For the graph cut energy term $E_{gc}(S_T)$, a sub-graph $G_R(N_R, A_R)$ is built using

the method described in [84]. Every voxel $\mathcal{I}_v \in \mathcal{I}$ has a corresponding node $n_R \in N_R$. Two additional nodes, the source (object) s and the sink (background) t, are added. Each node n_R has one t-link to each of the source and sink, which enforces the data-term energy. Each pair of neighboring nodes is connected by an n-link, which encodes the boundary energy term. The minimum-cost s-t cut divides the graph G_R into two parts: all nodes belonging to the target object are included in the source set and all background nodes are in the sink set.

For the graph search energy term $E_{gs}(S_T)$, a sub-graph $G_T(N_T, A_T)$ is constructed, which follows the method described in Chapter 4. Every node $n_T(x, y, z) \in N_T$ corresponds to exactly one voxel $\mathcal{I}(x, y, z) \in \mathcal{I}$. The positions of nodes reflect the positions of corresponding voxels in the image domain. Two types of arcs are added to the graph: (1) The intra-column arcs with $+\infty$ weight serve to enforce the monotonicity of the target surface S_T; and (2) the inter-column arcs incorporate the shape-prior penalties f_{pq} between the neighboring columns p and q. Each node is assigned a weight w_n such that the total weight of a closed set in the graph G_T equals to the edge-cost term in E_{gs}. Then, as in [82, 62], each node n_T is connected to either the source s by the arc with weight $-w_n$ if $w_n < 0$ or the sink t by the arc with weight w_n if $w_n > 0$. Note that the source s and the sink t are the same nodes used in G_R for the implementation of the graph cut energy term. Using this construction, we merge the two sub-graphs G_R and G_T as a single s-t graph G. The original energy minimization can be achieved by solving a maximum flow problem in the graph. The target surface S_T can be defined by the minimum-cost s-t cut in the graph. All nodes

in G_T above surface S_T belong to the sink set and all nodes on or below S_T belong to the source set of the cut [82].

To incorporate geometric context constraints, additional inter-graph arcs are added between two sub-graphs. We begin with the case that region R_A tends to be lower than the terrain-like surface S_T with a given distance d. If $n_R(x_1, y_1, z_1)$ in the subgraph G_R belongs to the source set (labeled as "object") and $n_T(x_1, y_1, z_1 + d)$ in the subgraph G_T belongs to the sink set, which indicates that $S_T(x_1, y_1) - z_1 < d$, a penalty w contributes to the objective energy function. That can be enforced by adding a directed arc with a weight of w from each node $n_R(x, y, z)$ to $n_T(x, y, z + d)$, as shown in Fig. 5.4(a). Note that if we require that the distance between the target region R_A and the surface S_T be at least d, we can set $w = +\infty$.

To enforce the constraint that R_A tends to be higher than S_T with distance d, a "flip" operation is involved. A transformed graph $G'_T(N'_T, A'_T)$ is constructed, in which node $n'_T(x, y, z)$ corresponds to voxel $\mathcal{I}(x, y, Z - z - 1)$ in image $\mathcal{I}$. The context penalty is given by adding a directed arc with weight w from $n_R(x, y, z)$ to $n'_T(x, y, Z - z + d - 1)$, which is associated with voxel $\mathcal{I}(x, y, z - d)$ in original image. Fig. 5.4(b) shows the flipped graph.

Once the graph is constructed, a globally optimal solution can be found by solving a *single* maximum flow problem, which minimizes the total energy $E(\mathcal{S})$ in low-order polynomial time.

5.5 Application for Lung Tumor Segmentation

In this section, we apply our new approach incorporating surface-region context information for lung tumor segmentation from MVCBCT images. Lung cancer is the leading cause of cancer related deaths [9]. Radiotherapy, as one standard cancer therapy, is able to eradicate tumors when treated with relatively high doses. MVCBCT is a promising technique used in clinic for daily imaging of patients [55] for lung tumor radiotherapy. Successful segmentation of lung tumors from the respiratory correlated 3-D images reconstructed from the projection data of MVCBCT scans can provide important information of tumor motion and volume changes, which allows better delineation of lung tumors for radiation therapy [43, 42].

However, accurate segmentation of lung tumors from the reconstructed MVCBCT images is very challenging. First, the quality of the MVCBCT images is poor. Serious noise interference exists. Second, the lung tumor that moves a lot is frequently located next to the lung surface, e.g., the diaphragm. The adjacent tissues have a similar intensity profile, and no clear boundary exists in-between.

To overcome those difficulties, our new graph-based framework incorporating surface-region context information was employed for simultaneous segmentation of lung boundary and the tumor tissue.

5.5.1 Initialization

As the initialization step, one center point and two radii were required as the user input, from which two spheres were generated. The smaller one was required

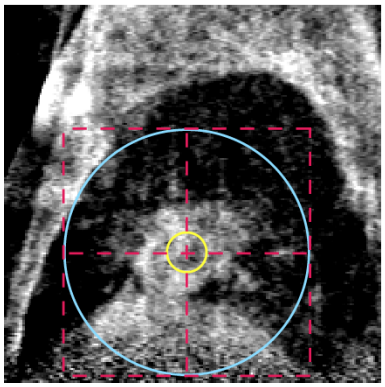


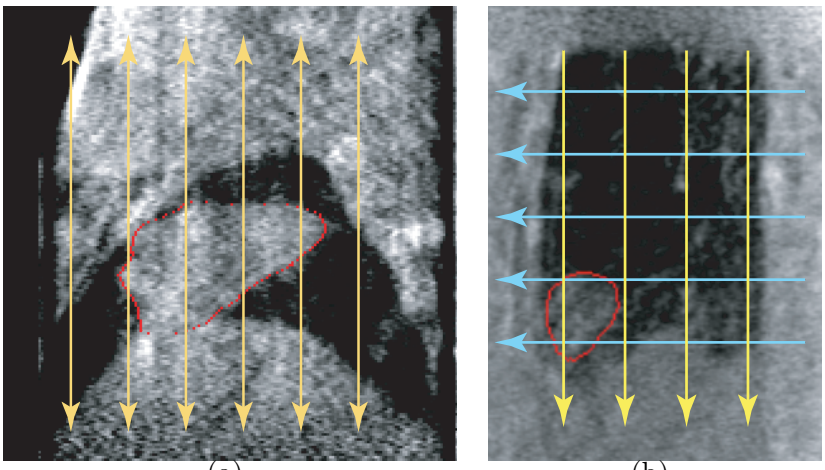
Figure 5.5: Typical example of the initialization step in sagittal view. Yellow sphere lies inside the tumor and blue sphere completely contains the tumor. The cross point indicates the given center point.

to be completely inside the tumor and the larger one was required to be completely outside the tumor. Fig. 5.5 shows a typical example. The segmentation was then conducted on the bounding box area of the larger sphere.

5.5.2 Graph Construction

For the segmentation of a single pair of surface and region, *s.t.*, one terrain-like surface and one region of arbitrary topology with mutual interaction between each other, the graph is constructed using the method described in Section 5.4, which contains two sub-graphs: G_T for lung boundary detection and G_R for tumor segmentation.

In certain cases, the target tumor may attach to lung surface from different directions, as shown in Fig. 5.6. In this situation, we segment the tumor and all



(a) (b)

Figure 5.6: Typical 2-D slices of 3-D MVCBCT image with tumors lying adjacent to the lung surface from different directions. (a) Tumor (red) lies adjacent to both the diaphragm and upper boundary of the lung. (b) Tumor (red) lies adjacent to the diaphragm and the chest wall.

surrounding surfaces simultaneously. Suppose we have i surrounding surfaces. Then $i+1$ graphs are constructed accordingly: i subgraphs $G_{T1}, \cdots, G_{Ti}$ for the detection of lung boundaries from different directions; G_R for tumor segmentation. The geometric interactions are enforced between each pairs of graphs: (G_{T1}, G_R), $\cdots$, (G_{Ti}, G_R).

5.5.3 Cost Function Design

5.5.3.1 Cost design for G_T:

For lung boundary detection, a gradient-based cost function was employed for edge-based costs. The negative magnitude of the gradient of the image $\mathcal{I}$ was computed at each voxel: $c(x, y, z) = -|\nabla \mathcal{I}|$. Note that the intensities inside the lung

are generally lower than surrounding tissues, a Sobel kernel was used to favor a bright-to-dark or dark-to-bright transition depending on the direction of the target surface. A second order shape prior penalty was employed with the form: $f(h) = \beta \cdot h^2$, where β was a constant parameter. The shape prior penalty penalized the change of the surface topology: $h = S_T(p) - S_T(q)$ between neighboring columns p and q.

5.5.3.2 Cost design for G_R:

For tumor detection, the data term D_v for voxel $\mathcal{I}_v$ was designed, as follows: For all voxels lying inside of the smaller circle, thus belonging to the object, $D_v(l_v = 1) = 0$, $D_v(l_v = 0) = +\infty$. Similarly, for all voxels outside the larger circle, $D_v(l_v = 1) = +\infty$, $D_v(l_v = 0) = 0$. For all other voxels, the intensity distribution for tumor followed Gaussian distribution. The mean intensity value $\bar{i}$ and standard deviation value σ were obtained from all the voxels inside of the smaller circle. For voxel $\mathcal{I}_v$ with intensity i_v, the data term was given as follows:

$$D_v(l_v = 1) = -\log Pr(i_v|l_v = 1) \propto \frac{(i_v - \bar{i})^2}{\sigma^2} , \tag{5.6}$$

$$D_v(l_v = 0) = -\log(1 - Pr(i_v|l_v = 1)) \propto -\log(1 - \exp(-\frac{(i_v - \bar{i})^2}{\sigma^2})) . \tag{5.7}$$

For the boundary penalty, a gradient-based cost function was used with the similar form as described in [84]:

$$w_{ij} = -\log(1 - \exp(\frac{-|\nabla\mathcal{I}|^2(i,j)}{2\sigma_g^2})) , \tag{5.8}$$

where w_{ij} corresponds to the weight of the n-link arc between neighboring voxels $\mathcal{I}_i$ and $\mathcal{I}_j$; $|\nabla\mathcal{I}|^2(i,j)$ denotes the squared gradient magnitude between $\mathcal{I}_i$ and $\mathcal{I}_j$.

5.5.3.3 Cost design for context constraint:

For the geometric context constraint, we required that the minimum distance between the target tumor and the lung boundary be at least one. A hard constraint is enforced with $d = 1$ and $w_v = +\infty$.

5.5.4 Experimental Methods

5.5.4.1 Data

The performance evaluation of the reported method was carried in 27 MVCBCT scans obtained from six patients with non-small cell lung cancer. For each scan, two volumetric images were reconstructed, one for full-exhalation phase and one for full-inhalation phase, resulting in 54 images [42]. The size of the reconstructed images was $256 \times 256 \times 256$ voxels with cubic voxel sizes of $1.07 \times 1.07 \times 1.07$ mm^3. Out of the 54 datasets, 2 were rejected for poor image quality by experts prior to any work reported here [42] and our experiments were conducted on the remaining 52 datasets. Surfaces of lung tumors were obtained by expert-defined manual tracing and served as the independent standard for assessing segmentation correctness of our approach.

5.5.4.2 Parameter setting

The employed values of the above-described segmentation method parameters were selected empirically. The same parameter values of $\beta = 5$, $d = 1$ were applied to all analyzed datasets. For σ_g used in the boundary cost of the graph-cut energy, we set $\sigma_g = 1$ for the tumor with a relative larger size (the approximate radius $R_T > 20$ mm) and $\sigma_g = 0.5$ for the tumor with a smaller size ($R_T < 20$ mm).

5.5.4.3 Validation

The segmentation performance was assessed using the Dice similarity coefficient (DSC), a broadly accepted measure of overlap between the segmented object and the independent standard. DSC $= 2|V_m \bigcap V_c|/(|V_m|+|V_c|)$, where V_m denotes volume of the independent standard and V_c denotes volume of the computer-determined object. All DSC values were computed in 3-D.

To determine the performance of our novel surface-and-region segmentation approach in comparison with a conventional approach of solely using a graph cut method to detect the tumor without simultaneously segmenting associated lung boundary surfaces, these two approaches were applied to all 52 MVCBCT images with identical spherical initialization (Fig. 5.5(a)) and the obtained DSC performance indices were compared for the two methods. Statistical significance of the observed differences was determined using Student t-test for which p value of 0.05 was considered significant.

5.5.5 Results

5.5.5.1 Quantitative validation

Our simultaneous surface–region segmentation method as well as Boykov's conventional graph-cut segmentation method were applied to all 52 MVCBCT images for which the independent standard was available. Our approach achieved tumor segmentation correctness characterized by DSC $= 0.82 \pm 0.07$ while the conventional approach yielded statistically significantly lower segmentation performance of

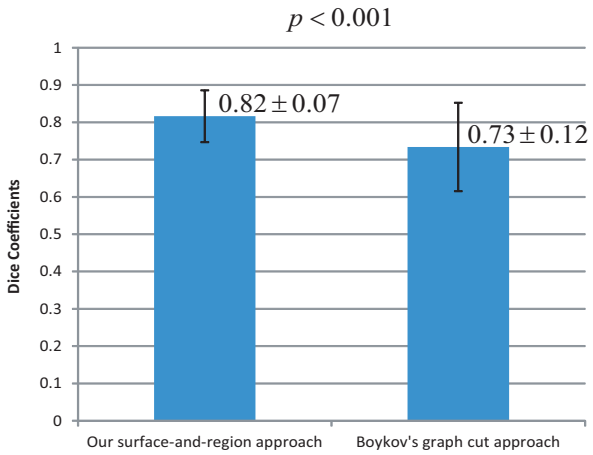


Figure 5.7: Quantitative and comparative performance evaluation based on computed DSC in 52 volumetric MVCBCT images. Overall tumor segmentation performance of Boykov's graph cut approach and our new surface-and-region method.

DSC $= 0.73 \pm 0.12$ ($p < 0.001$). Fig. 5.7 displays these overall results graphically as Mean±stdev.

Fig. 6.6 shows the pairwise performance comparisons for all 52 datasets, ordered according to the performance of our approach. Fig. 6.6 thus clearly demonstrates that the performance of the conventional method is very uneven in the analyzed data set. In contrary, our new approach not only shows an overall improvement of the segmentation, it also demonstrates that the segmentation performance is relatively consistent for all tumors in the entire set of the 52 analyzed images, thus showing a markedly higher robustness of the new approach resulting from the incorporation of image-based surface context.

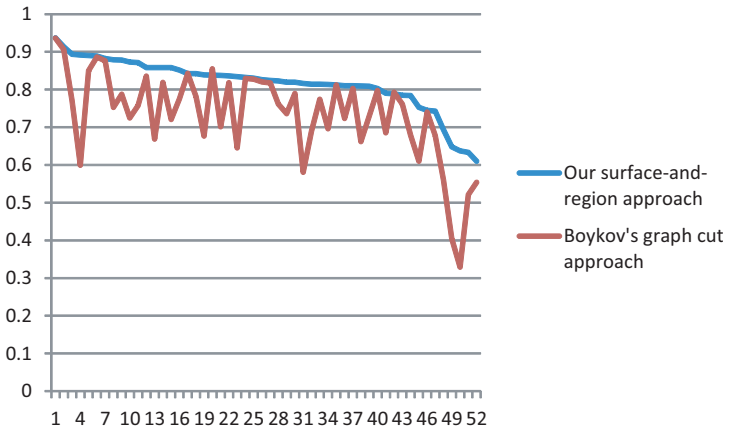
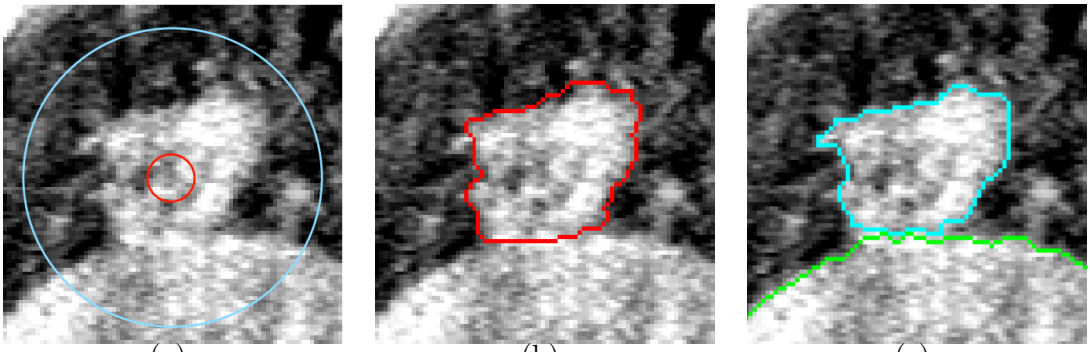


Figure 5.8: Dice coefficients for all 52 datasets ordered according to the performance of the conventional graph-cut approach (blue). Note the uneven character of the segmentation performance for the conventional method compared to the highly-robust segmentation performance of the new approach.

5.5.5.2 Qualitative results

Fig. 5.9 shows a typical outcome of our new tumor segmentation method and gives visual comparison with the independent standard. Fig. 5.11 shows tumor segmentation results obtained using the two compared methods in a difficult image, in which the tumor is closely adjacent to the lung surface from two directions. Fig. 5.11(b) shows a segmentation failure of the conventional approach while Figs. 5.11(c) show correctly segmented tumor using our new method.



(a) (b) (c)

Figure 5.9: Typical tumor segmentation example. (a) 2-D slice of 3-D MVCBCT image with outlines of spherical initialization. (b) Manual segmentation of the lung tumor – independent standard. (c) Simultaneous region-and-surface segmentation of the diaphragm (green) and lung tumor (blue) using our new approach showing excellent segmentation performance – DSC = 0.88.

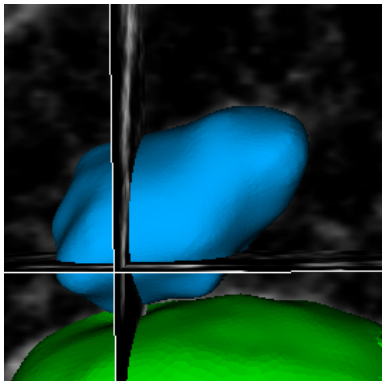


Figure 5.10: Typical tumor segmentation example: 3-D representation of the diaphragm (green) and the tumor (blue).

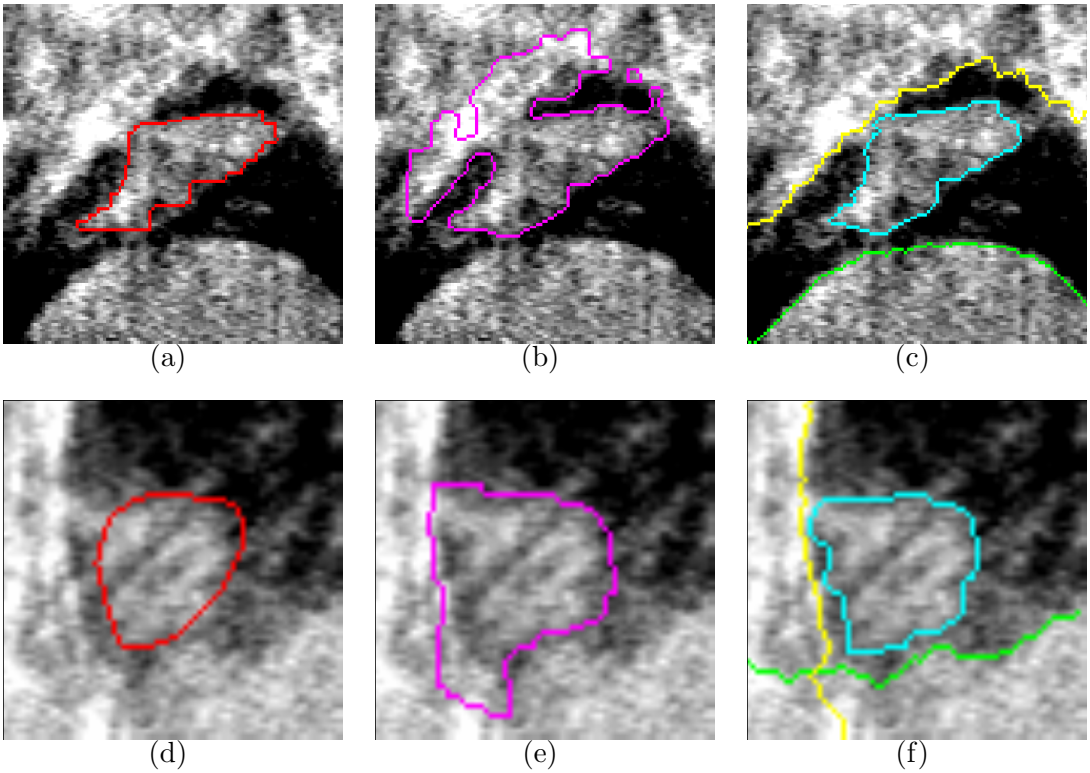


Figure 5.11: Performance comparison in two difficult images. (a),(d) Independent standard obtained by manual segmentation and shown in one 2-D slice of the 3-D volume. (b),(e) Tumor segmentation failure resulting from the conventional graph cut method – DSC = 0.69 for (b) and DSC = 0.70 for (e). (c),(f) Tumor segmentation obtained using our new method – segmented lung surfaces are shown in yellow and green, tumor surface is in blue – DSC = 0.81 for (c) and DSC = 0.84 for (f).

5.5.5.3 Execution time

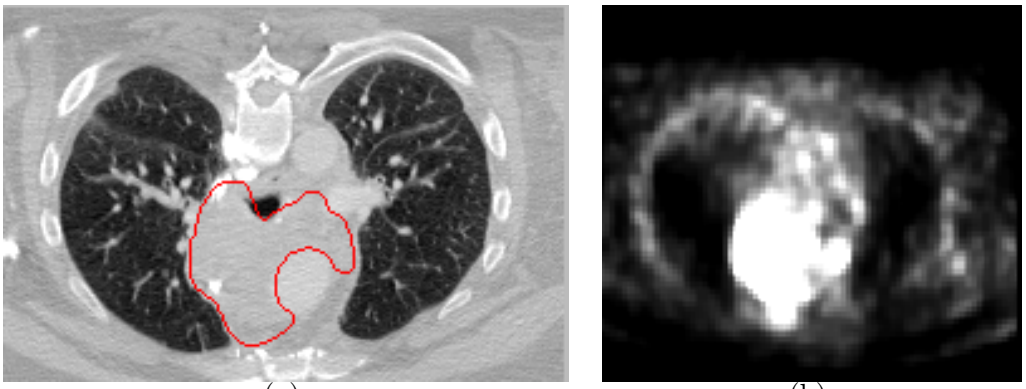
Our algorithm was implemented in C++ on a Linux workstation (3GHz, 32GB memory). Current non-optimized implementation required about 4 minutes.

5.6 Discussion and Conclusion

In this chapter, we report a novel solution to a segmentation problem, in which target objects of arbitrary topology mutually interact with terrain-like surfaces, which widely exists in the medical imaging field. A graph-based framework is employed and the known context information between terrain-like surfaces and regions of arbitrary topology is encoded by adding additional context terms in the energy function. A globally optimal solution can be achieved by solving a single maximum flow problem in low-order polynomial time, which leads to the simultaneous segmentation of surfaces and regions. For the tumor with arbitrary topology, we construct a corresponding sub-graph following the graph cut method making use of its topological flexibility. For lung surfaces, sub-graphs are built based on the graph search approach, which requires no initial seeds and provides good shape control. The context information between the tumor and the lung surface is encoded by adding weighted arcs between pair of sub-graphs. The simultaneous segmentation is achieved by solving a single maximum flow problem in the constructed graph. The evaluation showed that our method generated highly accurate tumor segmentations. Compared with the conventional graph-cut method, our new approach provided significantly better results ($p < 0.001$). The Dice coefficient obtained by the conventional graph-cut approach (0.73 ± 0.12) was improved to 0.82 ± 0.07 when employing our new method for pulmonary tumor segmentation.

CHAPTER 6
OPTIMAL CO-SEGMENTATION OF TUMOR IN PET-CT IMAGES WITH CONTEXT CONSTRAINTS

6.1 Introduction



(a) (b)

Figure 6.1: The PET-CT images. (a) One slice of CT image for the treatment planning of lung tumor (red line). (b) Corresponding PET image.

Accurate target delineation plays an important role in image-guided radiation therapy due to the high dose gradient inhere to the technology. Inaccurate delineation of target volumes by a small amount can result in either overdose of the surrounding tissues, dangerous underdose of the tumor, or both [8, 30]. For effective assessment of target volumes in cancer treatment, medical imaging modalities like Computed Tomography (CT) and Positron Emission Tomography (PET) are widely used. CT images usually have high resolution and give detailed anatomical information. How-

ever, no information about functionality can be obtained. Fig. 6.1(a) shows one typical example. Studies show that a large inter-observer variability has been observed for many tumor sites because the tumors usually have a very similar intensity distribution with surrounding soft tissues [1, 23].

To improve the visualization of tumor tissues, Positron Emission Tomograph (PET) imaging with the metabolic tracer ^{18}F-FDG (Fluro Deoxy Glucose) is introduced. In PET, target tumors have a high standardised uptake value (SUV) than the normal surrounding structures, which leads to accurate classification between normal tissues and malignant tumors. However, PET images usually have a lower resolution and a poorer boundary information, when comparing with CT images (see Fig. 6.1(b)). Thus accurate segmentation of tumors in PET alone is still problematic. For a better delineation, PET images have been combined with treatment planning CT for accurate target delineation in radiation therapy. In the treatment planning, PET images and CT images are registered together to locate the tumor and its relative location to the anatomy, which provides better localization of target volumes. In recent years, PET-CT has become a standard method in radiation therapy and has increasingly used for tumor delineation. Though the combination of PET-CT images helps to produce a better tumor contouring, manual delineation is still time consuming. Furthermore, observer variability in tumor delineation using FDG-PET is still high and often inconsistent with CT-based anatomically defined tumor contours. Accurate segmentation from co-registered PET-CT image is still a challenging problem.

In this chapter, we propose a novel method for co-segmentation of tumor in

PET and CT images, which makes use of the advantages from both modalities: the superior contrast of PET and the superior anatomical resolution of CT. Our method is based on the graph cut framework. As described in Section 1.3.2.1, the basic idea of the graph cut method is to formulate the segmentation problem as an energy minimization problem. The energy encodes both boundary and region information, which can be minimized in the discrete space using graph-based method. In this chapter, we propose a novel energy representation, which contains the information from both PET and CT images. In addition, a context term is also incorporated, which encodes the context constraint between PET and CT. The co-segmentation of PET and CT images can be achieved by solving a *single* maximum flow problem, which leads to a globally optimal solution in low-order polynomial time.

The rest of the chapter is organized as follows. In Section 6.2, we give a brief review of the related work in PET-CT image segmentation. In Section 6.3, we give a detailed description of the proposed method, including the formulation of the energy and the corresponding graph construction. Section 6.4 shows the application of the proposed framework for lung tumor segmentation. Section 6.5 draws the conclusion.

6.2 Related Work

The usage of both PET and CT images for accurate target delineation in radiotherapy has attracted considerable attention in recent years. Ashamalla *et al.* [23] employed integrated PET-CT for delineation of lung tumor. In [8], a comparative study was conducted using fused PET-CT for gross tumor volume delineation

in comparison with the approach of solely relying on CT. The result demonstrated the effectiveness of combining the information from both modalities. These methods mainly rely on manual contouring. Yu *et al.* [25] proposed a region-based method for the segmentation of head-and-neck tumor from PET-CT. Baardwijk *et al.* [1] showed an auto-contouring method in non-small-cell lung cancer based on automatic thresholding on PET images. The CT data set was only used for attenuation correction of PET images, which may not make full use of the information from CT. In [80], an automated delineation of tumor boundaries was achieved by a joint likelihood ratio test in CT image, which was computed based on an appearance model learned from an initial segmentation in PET. Gribben *et al.* [24] proposed a maximum *a posterior*-Markov random field (MAP-MRF) based approach for the segmentation of lung tumor in PET-CT images. Similarly, Xia *et al.* [90] also employed a MAP-MRF framework for simultaneous segmentation of joint PET-CT images. An iterated conditional modes (ICM) method is used to solve the energy minimization problem. The result lacks a globally optimal guarantee and might be trapped in a local minimum solution.

6.3 Method

The main idea of the proposed method is to make use of information from both PET and CT images. Two sub-graphs are constructed, one for PET and one for CT. The weight of arcs in the sub-graph encodes the boundary and region information from the corresponding image. The key point is the incorporation of context information

between PET and CT images, which is enforced by adding inter-graph arcs between correspondent nodes of two sub-graphs. Note that in the preprocessing step, we register the PET image to the CT image to make sure both images have the same size and there exists a one-to-one correspondence between voxels in these two images. Our co-segmentation is conducted between the input CT image and the registered PET image with the same size.

6.3.1 Energy Function

Consider a PET-CT image pair $(\mathcal{I}, \mathcal{I}')$, where $\mathcal{I}$ denotes an input CT image and $\mathcal{I}'$ denotes the registered PET image. For each voxel $v \in \mathcal{I}$, we have a correspondent voxel $v' \in \mathcal{I}'$. Let l denote the binary variables assigned for each voxel, indexed as l_v over voxels $v \in \mathcal{I}$ and $l_{v'}$ over voxels $v' \in \mathcal{I}'$. In our notation, $l = 1$ means that the voxel belongs to the target object. If $l = 0$, then the voxel is considered "background". For every voxel $v \in \mathcal{I}$ (resp., $v' \in \mathcal{I}'$), the assigned value of l defines the tumor region in the CT (resp., PET) image.

To express our energy function, we start with energy term E_{CT} for the segmentation of CT, which contains two familiar terms as the well-known binary graph cut energy [84]: a region term and a boundary term. For each voxel $v \in \mathcal{I}$, a region cost $D_v(l_v)$ is assigned, which measures how well the label l_v fits the voxel v given the image data. Let $\mathcal{N}_v$ defines the neighboring relationship between voxels. Then the boundary cost $V_{ij}(l_i, l_j)$ can be expressed as

$$V_{ij}(l_i, l_j) = \begin{cases} B_{ij} & \text{if } l_i \neq l_j \\ 0 & \text{if } l_i = l_j \end{cases}, \tag{6.1}$$

where B_{ij} is the penalty of assigning different labels l_i and l_j to neighboring voxels $(v_i, v_j) \in \mathcal{N}_v$. Our energy term for CT segmentation takes the form:

$$E_{CT}(l_v) = \sum_{v \in \mathcal{I}} D_v(l_v) + \sum_{(v_i, v_j) \in \mathcal{N}_v} V_{ij}(l_i, l_j) \ , \tag{6.2}$$

The energy term E_{PET} for the segmentation of PET image has the same form as energy used for CT. Let $D_{v'}(l'_v)$ denote the region cost for voxel $v' \in \mathcal{I}'$, $V_{i'j'}(l_{i'}, l_{j'})$ denote the boundary cost between neighboring voxels $(v'_i, v'_j) \in \mathcal{N}_{v'}$, where $\mathcal{N}_{v'}$ defines the neighboring relationship between voxels in $\mathcal{I}'$. Then the energy term E_{PET} can be expressed as

$$E_{PET}(l_{v'}) = \sum_{v' \in \mathcal{I}'} D_{v'}(l_{v'}) + \sum_{(v'_i, v'_j) \in \mathcal{N}_{v'}} V_{i'j'}(l_{i'}, l_{j'}) \ , \tag{6.3}$$

As described above, the key point for successful segmentation is to make use of information from both images. To incorporate context information between PET and CT images, a context term is added into the energy, which penalizes the segmentation difference between two image sets. Let (v, v') denote a pair of corresponding voxels from the CT image $\mathcal{I}$ and the PET image $\mathcal{I}'$. Then the context cost $W_{vv'}(l_v, l_{v'})$ is defined as

$$W_{vv'}(l_v, l_{v'}) = \begin{cases} C_{vv'} & \text{if } l_v \neq l_{v'} \\ 0 & \text{if } l_v = l_{v'} \end{cases} \ , \tag{6.4}$$

where $C_{vv'}$ is employed to penalize the disagreement between labels of corresponding voxels (v, v'). Our context energy term takes the form:

$$E_{context}(l_v, l_{v'}) = \sum_{(v, v')} W_{vv'}(l_v, l_{v'}). \tag{6.5}$$

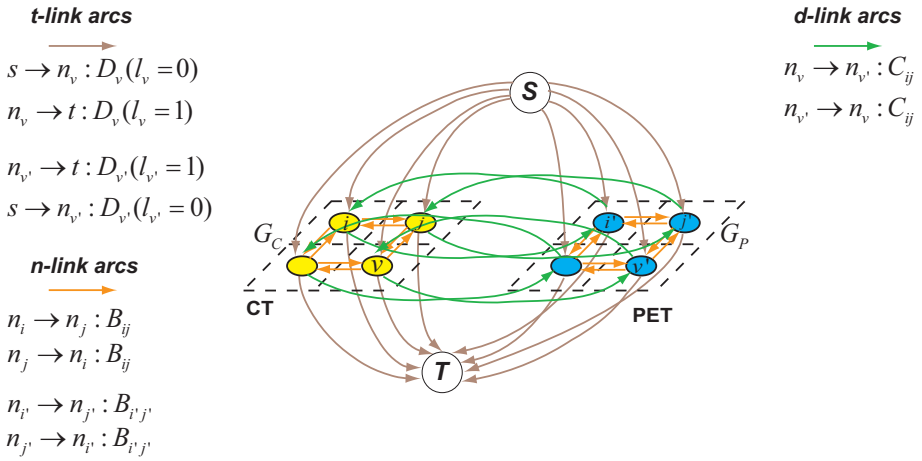


Figure 6.2: Graph construction of G with two sub-graphs G_C and G_P for the co-segmentation of PET-CT images. Three types of arcs are introduced. t-link arcs (brown) encode the region cost. n-link arcs (orange) encode the boundary cost. d-link arcs (green) enforce context information.

Now our energy has the overall form

$$E(l) = E_{CT}(l_v) + E_{PET}(l_{v'}) + E_{context}(l_v, l_{v'}), \tag{6.6}$$

Our target is to find an optimal set l such that the total energy is minimized, which corresponds to an optimal co-segmentation of PET-CT images.

6.3.2 Graph Construction

Our energy is minimized by solving a maximum flow problem in a corresponding graph, which allows a globally optimal solution in low-order polynomial time. In this section, we show how to encode our proposed energy terms through proper graph construction.

For energy terms E_{CT} and E_{PET}, a directed graph G is defined, which contains two node-disjoint sub-graphs $G_C(N, A)$ and $G_P(N', A')$, one for E_{CT} and one for E_{PET} (see Fig. 6.2). Note that two sub-graphs use exactly the same structure to incorporate the region and boundary term in the corresponding energy. For the construction of $G_C(N, A)$, every voxel $v \in \mathcal{I}$ has a corresponding node $n_v \in N$ in G_C. Two dummy nodes, a source s and a sink t, are added into the graph. To encode the region term $\sum_{v\in\mathcal{I}} D_v(l_v)$, we put a t-link arc from source s to each node n with the weight $D_v(l_v = 0)$ and a t-link arc from each node n to the sink t with the weight $D_v(l_v = 1)$. The boundary term $\sum_{(v_i,v_j)\in\mathcal{N}_v} V_{ij}(l_i, l_j)$ is enforced by adding n-links as follows. For each pair of neighboring voxels $(v_i, v_j) \in \mathcal{N}_v$, two n-link arcs are incorporated, one from n_i to n_j and one in the opposite direction from n_j to n_i. The weight of arcs is set as B_{ij}. The sub-graph $G_P(N', A')$ can be built using the same way to encode the region term $\sum_{v'\in\mathcal{I}'} D_{v'}(l_{v'})$ and the boundary term $\sum_{(v'_i,v'_j)\in\mathcal{N}_{v'}} V_{i'j'}(l_{i'}, l_{j'})$ for the energy term E_{PET}. To solve them in a single maximum flow, two sub-graphs share the same source s and sink t.

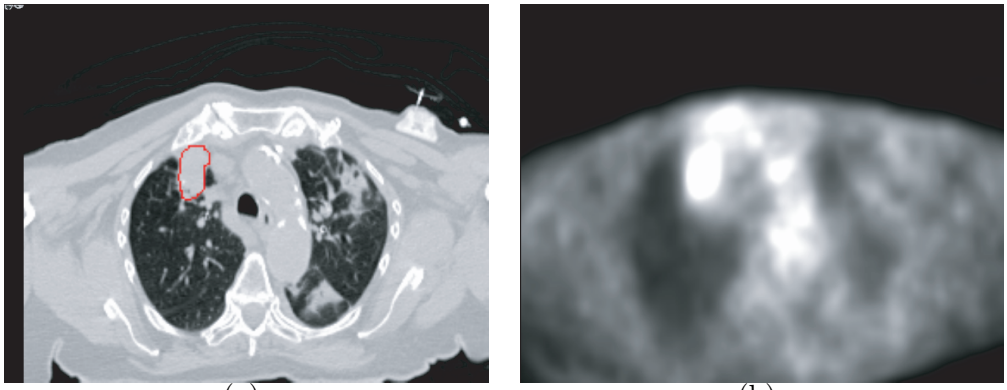
To enforce the context term $E_{context}(l_v, l_{v'}) = \sum_{(v,v')} W_{vv'}(l_v, l_{v'})$, additional inter-graph arcs are added between G_C and G_P, named as "*d-link*" arcs. For every pair of corresponding voxels (v, v'), where v belongs to CT image and v' is the corresponding voxel in PET image, two d-link arcs are added between corresponding nodes of two sub-graphs, one from n_v to $n_{v'}$ and one from $n_{v'}$ to n_v. The weight of arcs is set as $C_{vv'}$, which is the context cost penalizing the segmentation difference between two images (see the definition in Section 6.3.1).

We thus finish the construction of the graph. Fig. 6.2 shows one example. With the constructed graph G, we can find an optimal cut $\mathcal{C}^* = (A^*, \bar{A}^*)$, which separates the graph into two parts, the source set A^* and the sink set $\bar{A}^*$ ($A^* \cup \bar{A}^* = N \cup N' \cup \{s, t\}$) in G. This optimal cut can be found by solving a maximum flow problem in low-order polynomial time, minimizing our objective energy function Eq. (6.6). The target tumor area in CT image is defined by every voxel whose corresponding nodes in G_C belonging to the source set. Similarly, the segmented tumor area in PET image is given by every voxel whose associated nodes in G_P belonging to the source set.

6.4 Application on Lung Tumor Segmentation in PET-CT Images

In this section, we apply the proposed framework for a challenging task: lung tumor segmentation in PET-CT images. In non-small cell lung cancer (NSCLC), the tumor may invade into the adjacent area of the lung, including the chest wall, the mediastinal structure or the diaphragm, which largely increases the difficulty of the tumor detection with anatomical imaging. The incorporation of PET information is necessary for a successful tumor delineation. Here our algorithm is applied for co-segmentation of PET-CT images. Our method can be divided into two steps. In the pre-processing step, the PET image is registered with the CT image, which creates a one-to-one correspondence between the CT image and the registered PET image. Then, our graph-based co-segmentation is conducted based on PET and CT images.

6.4.1 Registration



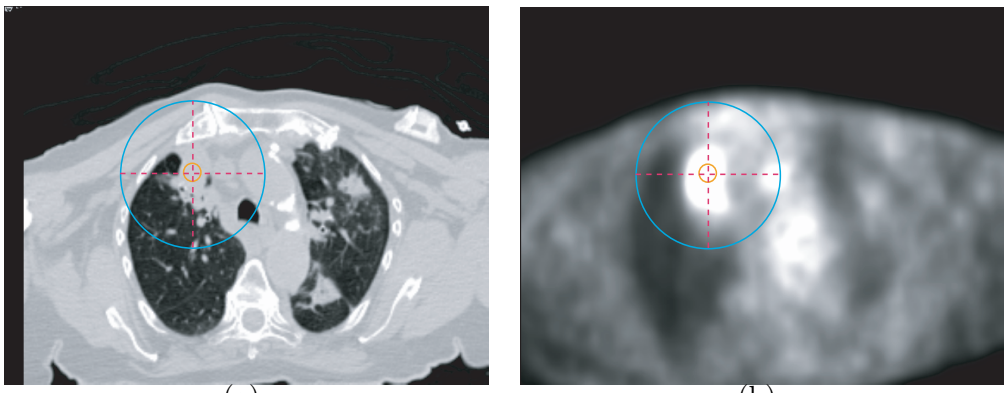
(a) (b)

Figure 6.3: Registration of PET and CT images. (a) One slice of CT image for the treatment planning of lung tumor (red). (b) Registered PET image using affine transform.

As a pre-processing, the PET image is registered with CT image using affine transformation based on the Elastix toolbox [68]. The registered PET image is interpolated to make sure that both images have the same size and there exists a one-to-one correspondence between voxels in these two images. Fig. 6.3 shows an typical PET-CT image pair after registration.

6.4.2 Initialization

Initial seed points are required for our graph-based co-segmentation. In our application, one center point and two radii are given by the user for each target



(a) (b)

Figure 6.4: Example slices of the initialization step in (a) CT and (b) PET images. The orange sphere completely lies inside the tumor and the blue sphere completely contains the tumor. The center point is given by the cross point of two red lines. Note that both images share the same initialization.

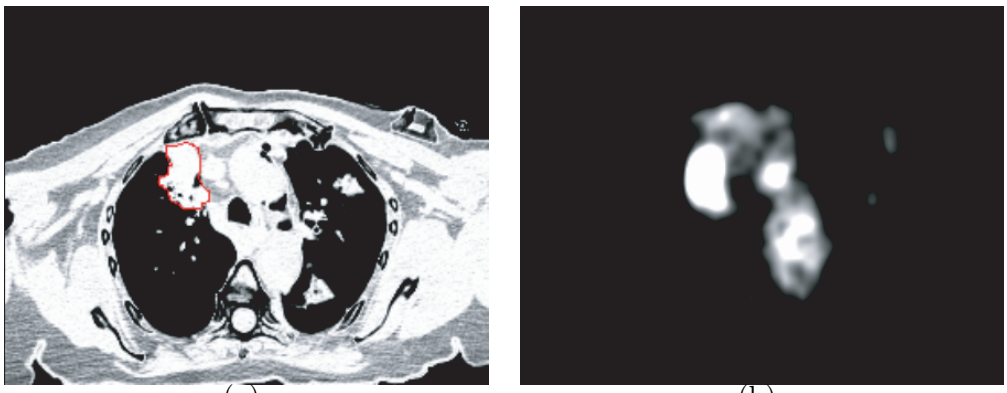
tumor, from which two spheres are generated in an approximate tumor area. We require that the smaller one must be completely inside the tumor and the larger one must be completely outside the tumor (see Fig. 6.4). In this way, all voxels inside the small circle are taken as the seed set of the object, denoted as $\mathcal{F}$ in CT image $\mathcal{I}$ and $\mathcal{F}'$ in PET image $\mathcal{I}'$; All voxels outside the large circle are considered to be the seed set of the background, denoted as $\mathcal{B}$ (resp., $\mathcal{B}'$) in $\mathcal{I}$ (resp., $\mathcal{I}'$). If there exists more than one tumors in the dataset, then for each tumor location, one center point as well as two radii are required as the initial input. Suppose we have M tumors in the image. Let $\mathcal{F}_i$ (resp., $\mathcal{F}_i'$) and $\mathcal{B}_i$ (resp., $\mathcal{B}_i'$) denote the seed set associated with the tumor i in CT (resp., PET), $i = 1, \ldots, M$. Then the overall seed set of the object

can be expressed as $\mathcal{F} = \mathcal{F}_1 \bigcup \mathcal{F}_2 \bigcup \ldots \bigcup \mathcal{F}_M$ (resp., $\mathcal{F}' = \mathcal{F}'_1 \bigcup \mathcal{F}'_2 \bigcup \ldots \bigcup \mathcal{F}'_M$). The overall seed set of the background has the form: $\mathcal{B} = \mathcal{B}_1 \bigcap \mathcal{B}_2 \bigcap \ldots \bigcap \mathcal{B}_M$ (resp., $\mathcal{B}' = \mathcal{B}'_1 \bigcap \mathcal{B}'_2 \bigcap \ldots \bigcap \mathcal{B}'_M$). Note that in this application, CT and PET images share the same seed spheres, which means that for each sphere in CT, we have a corresponding sphere in PET with the same center point and radius.

6.4.3 Graph-based Co-segmentation

The graph is constructed following the method described in Section 6.3.2. Two sub-graphs G_C and G_P are constructed for the segmentation of CT and PET images, respectively. The cost is designed as follows.

6.4.3.1 Cost function design for G_C



(a) (b)

Figure 6.5: One typical slice of cost image for the region term in (a) CT and (b) PET images. The target tumor area is labeled by the red contour. Higher intensity value indicates a larger probability that the voxel belongs to the tumor.

For every voxel inside the smaller circle ($v \in \mathcal{F}$), thus identified as belonging to the tumor by the user, a hard region cost is set as $D_v(l_v = 1) = 0$, $D_v(l_v = 0) = +\infty$, which makes sure that they will be labeled as the object. Otherwise an infinite cost will be introduced. Similarly, for every voxel outside the larger circle ($v \in \mathcal{B}$), which is identified as belonging to the background, $D_v(l_v = 1) = +\infty$, $D_v(l_v = 0) = 0$. Thus these voxels must be included in the background area.

For any voxel lying between two circles ($v \in \mathcal{I} - \mathcal{F} - \mathcal{B}$), the region cost is designed by how the intensity of voxel v fits into given intensity models. The image intensity of the tumor area in CT roughly follows a gaussian distribution. Thus the region cost can be assigned according to the intensity distribution learned from the seed set of the object. Let i_v denote the intensity value for voxel v. $\bar{i}$ is the mean intensity value computed from every voxel $v \in \mathcal{F}$. σ is the corresponding standard deviation. Then for any voxel $v \in \mathcal{I} - \mathcal{F} - \mathcal{B}$, the region term takes the form:

$$D_v(l_v = 1) = -\lambda_1 \log Pr(i_v | l_v = 1) \propto \frac{(i_v - \bar{i})^2}{\sigma^2} , \tag{6.7}$$

$$D_v(l_v = 0) = -\lambda_2 \log(1 - Pr(i_v | l_v = 1)) \propto -\log(1 - \exp(-\frac{(i_v - \bar{i})^2}{\sigma^2})) , \tag{6.8}$$

where λ_1 and λ_2 are given coefficients. Fig. 6.5(a) shows one typical slice of the cost image for the region term in CT.

For the cost design of the boundary term, a gradient-based cost is employed, which has a similar form as the well-known graph cut method in [84]. As described in Section 6.3.1, B_{ij} denotes the penalty of assigning different labels to neighboring

voxels $(v_i, v_j) \in \mathcal{N}_v$. Then B_{ij} can be expressed as

$$B_{ij} = -\lambda_3 \log(1 - \exp(\frac{-|\nabla \mathcal{I}|^2(i,j)}{2\sigma_g^2})) \ , \tag{6.9}$$

where $|\nabla \mathcal{I}|^2(i,j)$ denotes the squared gradient magnitude between v_i and v_j. λ_3 and σ_g are given parameters.

6.4.3.2 Cost function design for G_P

For every voxel $v' \in \mathcal{F}'$ and $v' \in \mathcal{B}'$, a hard region cost is set using the same form as we used for the CT image: $D_{v'}(l_{v'} = 1) = 0$, $D_{v'}(l_{v'} = 0) = +\infty$ for $v' \in \mathcal{F}'$ and $D_{v'}(l_{v'} = 1) = +\infty$, $D_{v'}(l_{v'} = 0) = 0$ for for $v' \in \mathcal{B}'$. For all other voxels $v' \in \mathcal{I}' - \mathcal{F}' - \mathcal{B}'$ between two circles, the region cost is computed based on the SUV value. As described in [35], the typical threshold value for tumor used in clinic application ranges from 15% to 50% of the maximum SUV in the image. Let $S(v')$ denote the SUV value for voxel v'. S_{max} denote the maximum SUV value in the image. Then for every voxel with a higher SUV value than 50% of S_{max}, they will have a high likelihood to belong to a tumor. Similarly, for every voxel with a lower SUV value than 15% of S_{max}, they most likely belong to the background. Based on this prior knowledge, our region cost function takes the form:

$$D_{v'}(l_{v'} = 1) = \begin{cases} 0 & \text{if } S(v') > S_u \\ \lambda_4 C_{max} \cdot (1 - (\frac{S(v') - S_l}{S_u - S_l})^2) & \text{if } S_l \leq S(v') \leq S_u \\ \lambda_4 C_{max} & \text{if } S(v') < S_l \end{cases} \tag{6.10}$$

$$D_{v'}(l_{v'} = 0) = \lambda_4 C_{max} - D_{v'}(l_{v'} = 1) \tag{6.11}$$

, where C_{max} is the maximum region cost allowed; $S_u = 50\% \cdot S_{max}$ and $S_l = 15\% \cdot S_{max}$ are the upper and lower threshold value for the possible tumor region. A quadratic function is employed to assign a high cost for a voxl with a low SUV value between S_l and S_u. λ_4 is the given coefficient. One typical slice of the cost image for the region term in PET is shown in Fig. 6.5(b).

The boundary term takes the same form as we have used in CT images. The penalty $B_{i'j'}$ of assigning different labels to neighboring voxels $(v'_i, v'_j) \in \mathcal{N}_{v'}$ can be expressed as

$$B_{i'j'} = -\lambda_5 \log(1 - \exp(\frac{-|\nabla \mathcal{I}'|^2(i', j')}{2\sigma'^2_g})) \,, \tag{6.12}$$

where $|\nabla \mathcal{I}'|^2(i', j')$ denotes the squared gradient magnitude between v'_i and v'_j. σ'_g is a given parameter.

6.4.3.3 Cost function design for the context term

The context term in the energy penalizes the segmentation difference between two images. As described in Section 6.3.1, $C_{vv'}$ encodes the penalty of assigning different labels to corresponding voxel pair (v, v') in CT and PET. Here we assign a constant value K to $C_{vv'}$, which encourages a consistent segmentation in PET and CT images. K is tuned empirically on one dataset and applied on all datasets.

6.4.4 Experimental Methods

6.4.4.1 Data

20 sets of 3-D FDG-PET-CT images were obtained from different patients with lung tumors. Each set of images contains one CT image and one corresponding PET

image acquired during the treatment planning of the radiotherapy. The reconstructed matrix size for each CT slice is 512×512, with a voxel size ranging from $0.98 \times 0.98 \times 2.00$ mm^3 to $1.37 \times 1.37 \times 2.00$ mm^3. For the PET images, the reconstructed matrix size ranges from 128×128 to 168×168. The voxel size ranges from $3.39 \times 3.39 \times 2.02$ mm^3 to $4.07 \times 4.07 \times 4.00$ mm^3. As described in Section 6.4.1, the PET images were co-registered to the CT images in the pre-processing step and the experiments were conducted on the CT and the registered PET images, which have exactly the same size as the corresponding CT images. The manual segmentation of tumors was conducted by one experienced radiation oncologist on the CT images by the guidance of the registered PET images, which was taken as the ground truth.

6.4.4.2 Parameter setting

In our experiments, following parameter setting was empirically employed for all analyzed datasets. For the segmentation of tumor in CT, boundary term provides an more important information compared with the region term. Thus we set coefficients as $\lambda_1 = \lambda_2 = 1$ for the regional term and $\lambda_3 = 5$, $\sigma_g = 0.5$ for the boundary term. For the segmentation in PET image, regional information based on the SUV thresholding plays a key role. Thus we set coefficients as $\lambda_4 = 1$ for the regional term and $\lambda_5 = 0.1$, $\sigma'_g = 0.5$ for the boundary term. The constant value K for the context term is set as 50.

6.4.4.3 Validation

The proposed algorithm was carried out on all 20 sets of PET-CT images for validation. The segmentation performance was evaluated between the manual contouring by experts and the computed results. For quantitative validation, the Dice similarity coefficient (DSC) was computed using $D = 2|V_m \bigcap V_c|/(|V_m| + |V_c|)$, where V_m denotes the manual volumetric result and V_c denotes the computed result. All DSC values were computed in 3-D.

To determine the performance of our novel co-segmentation approach in comparison with graph cut approaches of solely using CT or PET to detect the tumor, these three methods were applied to all 20 datasets with identical initialization and the same cost function for the boundary term and the regional term in corresponding image modalities. The computed DSC values were compared for all three methods. Statistical significance of the observed differences was determined using Student t-tests for which p values of 0.05 were considered significant.

6.4.5 Results

6.4.5.1 Quantitative validation

The computed dice coefficients are summarized in Fig. 6.6. The proposed approach combining information from both PET and CT achieved a segmentation performance of DSC = 0.80±0.07, which has a significant improvement in comparison with the graph cut method using only CT (DSC = 0.48 ± 0.26, $p < 0.001$) and the graph cut method using only PET (DSC = 0.63 ± 0.17, $p < 0.001$).

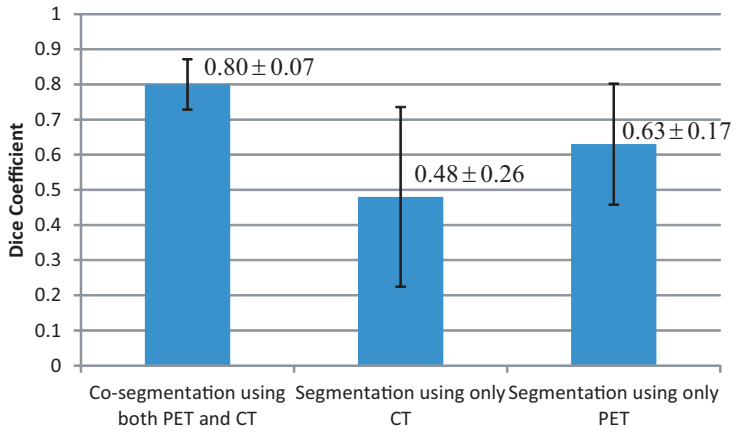


Figure 6.6: Quantitative and comparative performance evaluation based on computed DSC values in 20 pairs of volumetric PET-CT images. Our method showed a significant improvement compared with both the method using only CT ($p < 0.001$) and the method using only PET ($p < 0.001$).

6.4.5.2 Illustrative results

Illustrative results of our experiments are shown in Fig. 6.8 for three views. From all views, our co-segmentation works quite well in both CT and PET images. Fig. 6.8 shows the segmentation results obtained using the three compared methods. Our approach makes use of the advantages from both CT and PET images, exhibiting expected improvement in comparison with the graph-cut methods solely relying on CT or solely relying on PET.

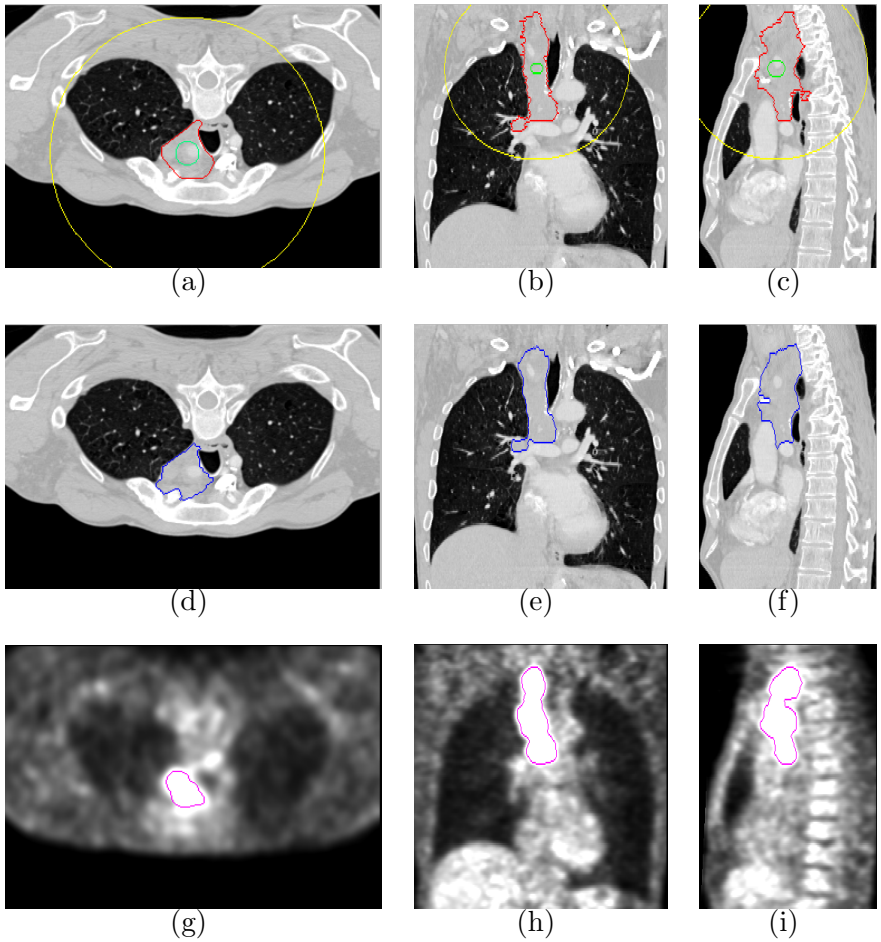


Figure 6.7: Typical tumor segmentation in transverse (left), coronal (middle) and sagittal (right) views. (a)-(c) 2-D slices of 3-D CT images with the ground truth (red) and outlines of spherical initialization (green and yellow). (d)-(f) Proposed co-segmentation results in CT images. (g)-(i) Co-segmentation results in PET images.

6.4.5.3 Execution time

Our algorithm was implemented in C++ on a Linux workstation (3GHz, 32GB memory). Current non-optimized implementation required about 4 minutes. The

execution time for the registration was about 60s. Our graph-based co-segmentation took about 3 minutes.

6.5 Conclusion

In this chapter, we propose a novel approach for the co-segmentation of the PET-CT images, which makes use of the strength from both modalities: the superior contrast of PET and the superior spatial resolution of CT. The target tumor in the PET and CT images is concurrently segmented with the help of the information acquired from the other modality. Our method was validated on a challenging task of the segmentation of lung tumors from 20 patients. The results clearly demonstrated the applicability and the improved performance of the proposed approach in comparison with the segmentation methods solely using PET or CT.

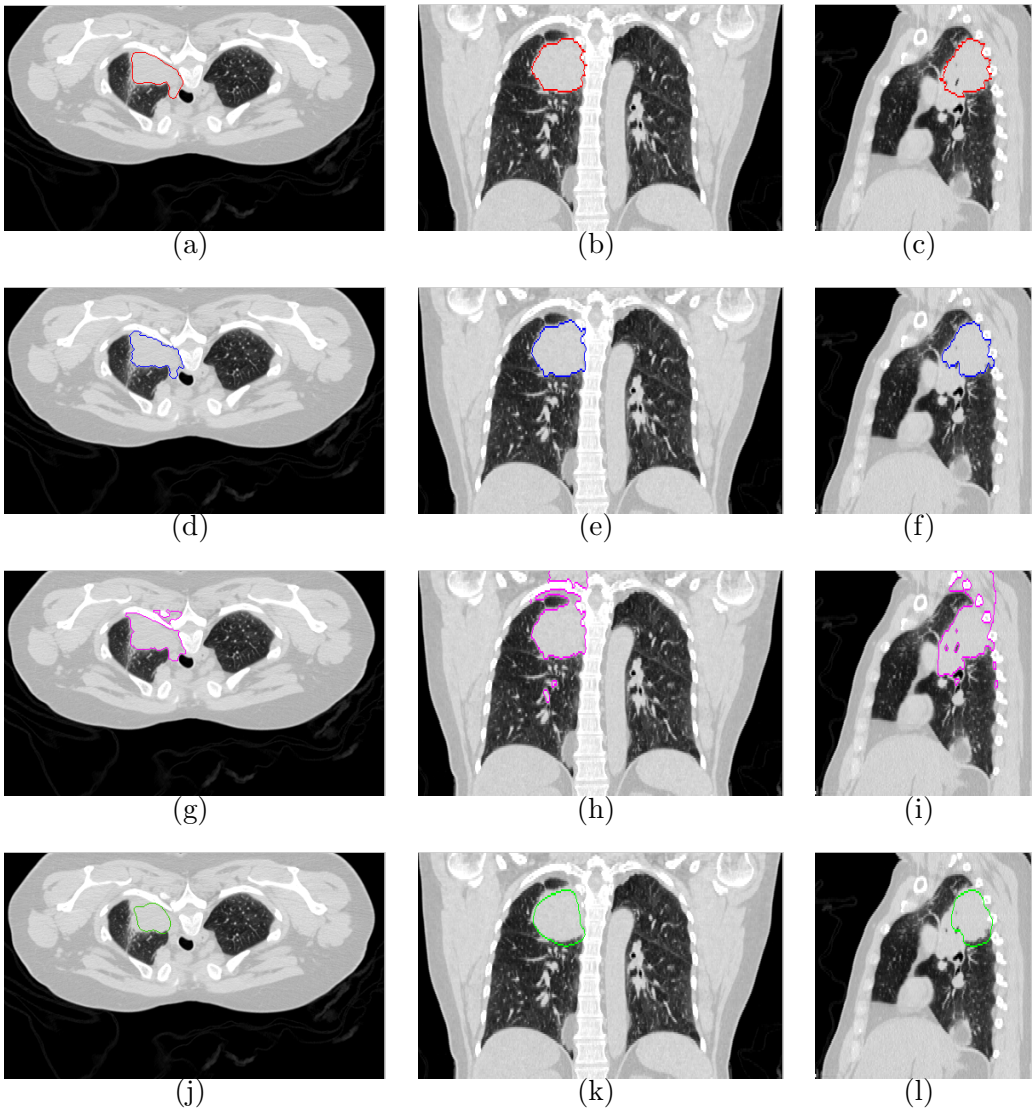


Figure 6.8: Comparative segmentation results in transverse (left), coronal (middle) and sagittal (right) views. (a)-(c) The manual segmentation results. (d)-(f) Our proposed co-segmentation results in CT images. (g)-(i) Segmentation results by the graph cut method solely using CT images. (j)-(l) Segmentation results by the graph cut method solely using PET images.

CHAPTER 7
CONCLUSION AND FUTURE DIRECTIONS

7.1 Conclusion

Optimal image segmentation is a challenging problem in the presence of weak boundary, high noise and serious mutually influence. In this thesis, we present novel approaches incorporating region, shape prior and context information in 3-D graph-theoretic frameworks to overcome those difficulties. To summarize, we first recall our specific aims as follows:

- Aim 1: Develop and validate a method for optimal surface segmentation incorporating both boundary and regional information using a ratio-form energy.
- Aim 2: Develop and validate a method for optimal multiple surfaces segmentation with both shape and context priors.
- Aim 3: Develop and validate a method for optimal surface-region segmentation with context constraints.
- Aim 4: Develop and validate a method for optimal co-segmentation in multi-modality imaging with context constraints.

A novel method for surface segmentation incorporating both boundary and regional image information by a ratio-form energy (Aim 1) is described in Chapter 3. The globally optimal surface can be achieved by solving a parametric maximum flow problem in the time complexity of computing a *single* maximum flow. The developed method has been applied for aorta segmentation of 15 MR aortic images.

The results showed the effectiveness of our approach to alleviate the interference of adjacent objects.

In Chapter 4, we propose a novel arc-based graph representation to incorporate both shape priors and context priors for multiple surface segmentation (Aim 2). A shape-prior term and a context-prior term are added into the energy function, which penalize local shape change and surface distance change with respect to the learned shape and context models. A globally optimal solution can be achieved by solving a maximum flow problem. The proposed method was validated on intraretinal layer segmentation of OCT images, bladder-prostate segmentation and image-resizing problems. The results clearly demonstrated the applicability of the proposed methods.

In Chapter 5, we focus on the segmentation of a special structure consisting of mutually interacting terrain-like surfaces and regions of arbitrary shape (Aim 3). The prior context information between the pairs of the terrain-like surfaces and regions are incorporated into the segmentation, which helps to achieve an accurate result when target surfaces and regions lack clear boundary and have similar intensity distribution. A graph-based framework is constructed, which makes use of the advantages from both the graph cut method and the graph search method. Simultaneous segmentation of both surfaces and regions with context constraints is achieved by solving a single maximum flow problem. We applied the proposed framework for the segmentation of lung tumor in mega-voltage cone beam CT (MVCBCT) images, which showed the power of our algorithm.

Finally, in Chapter 6, we develop a co-segmentation framework for tumor segmentation in PET-CT images (Aim 4). Our method is mainly based on the graph cut framework. Two sub-graphs are constructed with arcs between each other, which encode the prior context information. Our formulation allows making use of the information from both modalities. The proposed method was validated for lung tumor segmentation in PET-CT images. The results showed significant improvement in comparing with solely using PET or CT.

7.2 Future Directions

The general frameworks presented in this thesis have a number of potential applications and extensions. Here, we discuss two interesting future directions in related areas.

7.2.1 Multi-Surface, Multi-Region Problem

In the thesis, we show how to encode context constraints between multiple surfaces (Chapter 4), surface-region (Chapter 5), and multiple regions (Chapter 6). A natural idea is to combine them into a single framework. Suppose we have m terrain-like surfaces $S_1, \ldots, S_m$ and n regions of arbitrary shape $R_1, \ldots, R_n$. Context constraints can be enforced between different pairs of surfaces (S_i, S_j), $i = 1, \ldots, m$, $j = 1, \ldots, m$, different pairs of regions (R_i, R_j), $i = 1, \ldots, n$, $j = 1, \ldots, n$ and different pairs of surface-regions (S_i, R_j), $i = 1, \ldots, m$, $j = 1, \ldots, n$. The target is to simultaneously segment all these surfaces and regions with context constraints.

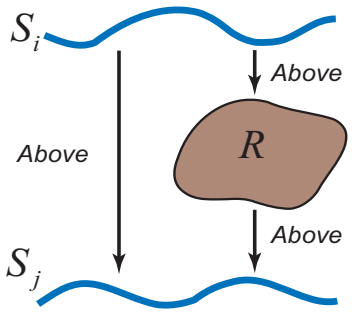


Figure 7.1: Typical context relationships that cannot be enforced using the proposed graph formulation.

To solve this problem, the basic idea is to employ a graph-theoretic framework, which has a similar form as what we proposed in the thesis. Multiple subgraphs are constructed, each of which corresponds to a target surface or region. The key part is how to encode the context information between pairs of objects. One possible solution is to introduce weighted arcs between pairs of sub-graphs. However, not all context constraints can be modeled in this way. Fig. 7.1 shows one example. Surface S_i is required to lie above the region R_i. Surface S_j is set to be below R. Meanwhile, there also exists context constraints between S_i and S_j. The above configuration cannot be trivially formulated using our current graph model. It introduces a *frustrated cycle* as described in [6], which cannot be optimized using graph-cut based method [11]. Specifically, Let $G_i(N_i, A_i)$ denote the sub-graph for surface S_i and $G_j(N_j, A_j)$ denote the sub-graph for surface S_j. If S_i is forced to lie above R_i, then G_i is constructed following the direction such that all nodes $n_i \in N_i$ below S_i belong to the source

set (see Section 5.4). Meanwhile, if S_j lies below R_i, a "flip" operation is involved and the graph G_j is built in an inverse direction such that all nodes above S_i belong to the source set (Section 5.4). The surface context constraint between S_i and S_j cannot be added between sub-graphs with opposite graph directions. In this situation, no globally optimal solution is guaranteed in polynomial time. However, some approximations can still be used, as described in [86, 11].

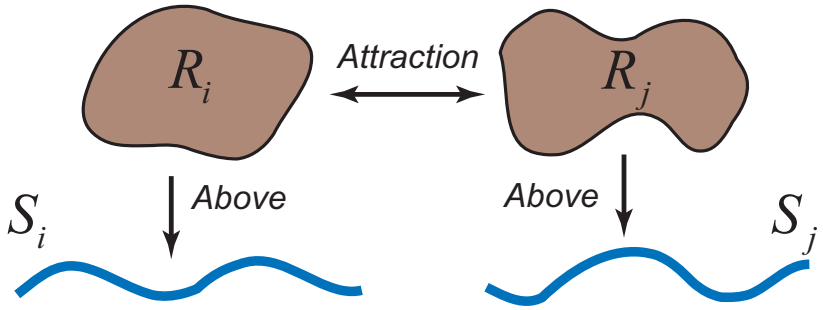
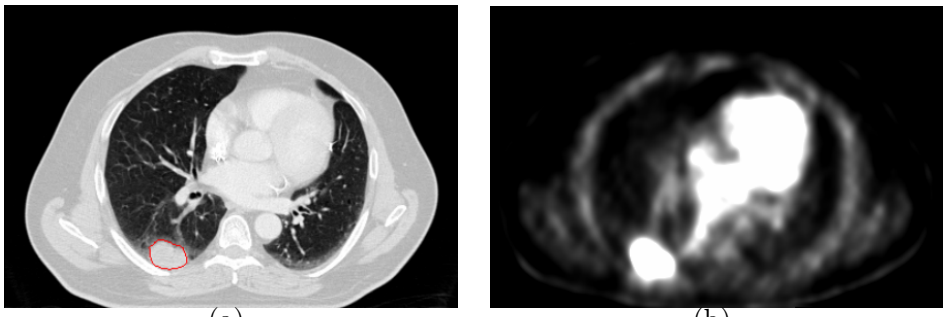


Figure 7.2: Typical context relationships that can be incorporated into the proposed graph framework.

Though not all context constraints can be modeled using our graph construction, there are still many useful context models, which allow a globally optimal solution in polynomial time. The example shown in Fig. 7.2 contains two regions R_i and R_j, and two terrain-like surfaces S_i and S_j. Four subgraphs G_{R_i}, G_{R_j}, G_{S_i} and G_{S_j} are constructed, respectively. Prior context information is enforced by adding weighted arcs between (G_{R_i},G_{S_i}), (G_{R_j},G_{S_j}) using exactly the same method as de-

scribed in Section 5.4. Here the attraction interaction between R_i and R_j means that the correspond nodes $n_{R_i} \in G_{R_i}$ and $n_{R_j} \in G_{R_j}$ tends to have the same labeling, which is incorporated by adding arcs between n_{R_i} and n_{R_j}, as described in Section 6.3.2.



(a) (b)

Figure 7.3: The tumor (red) attaches to the lung boundary in (a) CT image and (b) corresponding PET image.

The above configuration can be applied for the tumor segmentation in PET-CT images. As shown in Fig. 7.3, the tumor may attach to the lung boundary with similar intensity information, which increases the difficulty of the segmentation. Using the proposed multi-surface-multi-region framework, we can simultaneously segment tumor and adjacent lung boundary in both PET and CT images. The overlapping penalty is enforced between tumor and lung surface in PET and CT, separately. Furthermore, a context constraint is incorporated between PET and CT for tumor segmentation, which makes use of the information from both modalities. The whole

segmentation of four target objects (lung surfaces in PET and CT, tumors in PET and CT) can be achieved by solving a single maximum flow problem with a globally optimal guarantee.

7.2.2 Multi-Organ Localization

In many applications of medical image segmentation, an initial region of interest (ROI) is required, which indicates a rough position of the target objects. For example, our graph-based bladder-prostate segmentation needs a ROI to place the initial shape model, which is interactively defined by the user. Similarly, to detect lung tumors in MVCBCT and PET-CT images, an initial center point as well as two radii is needed, which gives seed areas for both objects and background. For a fully automatic segmentation framework, accurate object localization is in urgent need.

In [91], Yin *et al.* employed an Adaboost approach to define a ROI for each bone and its associated cartilage in the knee area, which was used as an initial localization for the graph-based cartilage segmentation. Unlike the approach used in the segmentation step which encoded the interaction constraints between different bones, his method conducted the localization for each object separately, which failed to make use of the context information between different target objects. In [4], Criminisi *et al.* presented a organ localization approach based on randomized decision forests. Long-range spatial context features were employed to incorporate the spatial correlations.

Here we propose a novel idea for multi-object localization. The basic idea is to

transform the objects localization problem into a model fitting problem. For each target object, a model is constructed, which incorporates both appearance information and geometric information. The prior context information between different pairs of target objects is encoded using a context model. The model fitting method is mainly based on Felzenswalb *et al.*'s work [57]. We formulate the model fitting problem into a discrete optimization problem, which can be solved using a dynamic programming based method. Specifically, let $L = l_1, \ldots, l_N$ denote the location for target objects $O_1, \ldots, O_N$, our energy takes the form:

$$E(L) = \sum_{i=1}^{N} m_i(l_i) + \sum_{(O_i,O_j)\in E} d_{ij}(l_i, l_j) \ , \tag{7.1}$$

where $m_i(l_i)$ is a fitting cost, which measures the degree of mismatch when object O_i is placed at location l_i. $d_{ij}(l_i, l_j)$ is a penalty function measuring the degree of deformation of the context model when object O_i is placed at location l_i and O_j is placed at location l_j. E defines the context relationship between pairs of objects, which is obtained from the context model. The optimal solution of this energy corresponds to the target locations for each objects.

In general, this energy minimization problem is an NP-hard problem. However, if the context relationship between pairs of objects is restricted to be a tree structure in the context model, the problem can be solved using dynamic programming. In [57], Felzenszwalb *et al.* proposed a new algorithm which runs in $O(M'N)$, where M' is the number of possible grid locations in a discrete space, which corresponds to the possible locations for the target objects. A globally optimal solution is then obtained using a dynamic programming method. A preliminary result of the proposed idea

was reported in [81], which proved the applicability of our approach.

REFERENCES

[1] A.Baardwijk, G.Bosmans, L.Boersma, J.Buijsen, S.Wanders, M.Hochstenbag, R.Suylen, A.Dekker, C.Dehing-Oberije, R.Houben, S.Bentzen, M.Kroonenburgh, P.Lambin, and D.Ruysscher. PET-CT-based auto-contouring in non-small-cell lung cancer correlates with pathology and reduces interobserver variability in the delineation of the primary tumor and involved nodal volumes. *International Journal of Radiation Oncology, Biology and Physics*, 68(3):771–778, 2007.

[2] A.Blake, C.Rother, M.Brown, P.Perez, and P.Torr. Interactive image segmentation using an adaptive gmmrf model. In *European Conference on Computer Vision*, 2004.

[3] A.Chan, W.Wells, C.Tempany, E.Grimson, and A.Willsky. Mutual information in coupled multi-shape model for medical image segmentation. 8(4):429–445, 2004.

[4] A.Criminisi, J.Shotton, and S.Bucciarelli. Decision forests with long-range spatial context for organ localization in CT volumes. In *MICCAI workshop on Probabilistic Models for Medical Image Analysis (MICCAI-PMMIA)*, 2009.

[5] A.Delong and Y.Boykov. A scalable graph-cut algorithm for N-D grids. In *Proc. IEEE International Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2008.

[6] A.Delong and Y.Boykov. Globally optimal segmentation of multi-region objects. In *Proc. IEEE International Conference on Computer Vision*, 2009.

[7] A.Jemal, R.Siegel, J.Xu, and E.Ward. Cancer statistics, 2010. *CA Cancer Journal for Clinicians*, 60(5):277–300, 2010.

[8] A.Riegel, A.Berson, S.Destian, T.NG, L.Tena, R.Mitnick, and P.Wong. Variability of gross tumor volume delineation in head-and-neck cancer using CT and PET/CT fusion. *International Journal of Radiation Oncology, Biology and Physics*, 65(3):726–732, 2006.

[9] A.Spira and D.Ettinger. Multidisciplinary management of lung cancer. *The New England Journal of Medicine*, pages 379–392, 2004.

[10] C.Rother, V.Kolmogorov, and A.Blake. Grabcut–interactive foreground extraction using iterated graph cuts. In *ACM Transactions on Graphics (SIGGRAPH)*, 2004.

[11] C.Rother, V.Kolmogorov, V.Lempitsky, and M.Szummer. Optimizing binary MRFs via extended roof duality. In *Proc. IEEE International Conference on Computer Vision and Pattern Recognition*, 2007.

[12] D.Cremers. Nonlinear dynamical shape priors for level set segmentation. In *Proc. IEEE International Conference on Computer Vision and Pattern Recognition*, 2007.

[13] D.Cremers, F.Schmidt, and F.Barthel. Shape priors in variational image segmentation: convexity, lipschitz continuity and globally optimal solutions. In *Proc. IEEE International Conference on Computer Vision and Pattern Recognition*, pages 1–6, 2008.

[14] D.Freedman, R.Radke, T.Zhang, Y.Jeong, D.Lovelock, and G.Chen. Model-based segmentation of medical imagery by matching distributions. *IEEE Trans. Medical Imaging*, 24(3), 2005.

[15] D.Freedman and T.Zhang. Interactive graph cut based segmentation with shape priors. In *Proc. IEEE International Conference on Computer Vision and Pattern Recognition*, volume 1, pages 755–762, 2005.

[16] D.Han, X.Wu, and M.Sonka. Optimal multiple surfaces searching for video/image resizing - a graph-threoretic approach. In *Proc. IEEE International Conference on Computer Vision*, 2009.

[17] D.Pham, C.Xu, and J.Prince. Current methods in medical image segmentation. 2:315–337, 2000.

[18] D.S.Hochbaum. An efficient algorithm for image segmentation, markov random fields and related problems. *Journal of the ACM*, 48:686–701, 2001.

[19] E.Dam, P.Fletcher, S.Pizer, G.Tracton, and J.Rosenman. Prostate shape modeling based on principal geodesic analysis bootstrapping. In *Proc. International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 1008–1016, 2004.

[20] F.Zhao, H.Zhang, A.Wahle, M.Thomas, A.Stolpen, T.Scholz, and M.Sonka. Congenital aortic disease: 4D magnetic resonance segmentation and quantitative analysis. *Medical Image Analysis*, 13(3):483–493, 2009.

[21] F.Zhao, H.Zhang, A.Wahle, T.D.Scholz, and M.Sonka. Automated 4D segmentation of aortic magnetic resonance images. In *British Machine Vision Conference (BMVA)*, volume 1, pages 247–256, 2006.

[22] G.Gallo, M.D.Grigoriadis, and R.E.Tarjan. A fast parametric maximum flow algorithm and applications. *SIAM J. Computing*, 18(1):30–55, 1989.

[23] H.Ashamalla, S.Rafla, K.Parikh, B.Mokhtar, G.Goswami, S.Kambam, H.Abdel-Dayem, A.Guirguis, P.Ross, and A.Evola. The contribution of integrated PET/CT to the evolving definition of treatment volumes in radiation treatment planning in lung cancer. *International Journal of Radiation Oncology, Biology and Physics*, 63(4):1016–1023, 2005.

[24] H.Gribben, P.Miller, G.Hanna, K.J.Carson, and A.R.Hounsell. MAP-MRF segmentation of lung tumors in PET/CT images. In *IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI)*, pages 290–293, 2009.

[25] H.Yu, C.Caldwell, K.Mah, I.Poon, J.Balogh, R.Mackenzie, N.Khaouam, and R.Tirona. Automated radiation targeting in head-and-neck cancer using region-based texture analysis of PET and CT images. *International Journal of Radiation Oncology, Biology, Physics*, 75(2):618–625, 2009.

[26] H.Zhang, A.Abiose, D.Campbell, M.Sonka, J.Martins, and A.Wahle. Left-ventricle segmentation in real-time 3D echocardiography using a hybrid active shape model and optimal graph search approach. In *Proc. SPIE Medical Imaging: Image Processing*, volume 7626, 2010.

[27] I.Ayed, S.Li, and I.Ross. Level set image segmentation with a statistical overlap constraint. In *Proceedings of the 21st Biennial International Conference on Information Processing in Medical Imaging*, 2009.

[28] I.H.Jermyn and H.Ishikawa. Globally optimal regions and boundaries as minimum ratio weight cycles. *IEEE Trans. Pattern Anal. Machine Intell*, 23(10):1075–1088, 2001.

[29] J.A.Sethian. *Level set methods and fast marching methods*. Cambridge University Press, 1999.

[30] J.Bradley, W.Thorstad, S.Mutic, T.Miller, F.Dehdashti, B.Siegel, W.Bosch, and R.Bertrand. Impact of FDG-PET on radiation therapy volume delineation in non-small-cell lung cancer. *International Journal of Radiation Oncology, Biology and Physics*, 59(1):78–86, 2004.

[31] J.Lellmann, F.Becker, and C.Schnorr. Convex optimization for multi-class image labeling with a novel family of total variation based regularizers. In *Proc. IEEE International Conference on Computer Vision*, pages 646–653, 2009.

[32] J.Malcolm, Y.Rathi, and A.Tannenbaum. Graph cut segmentation with nonlinear shape priors. In *Proc. IEEE International Conference on Image Processing*, 2007.

[33] J.Shi and J.Malik. Normalized cuts and image segmentation. *IEEE Trans. Pattern Anal. Machine Intell*, 22(8), 2000.

[34] J.S.Stahl and S.Wang. Edge grouping combining boundary and region information. *IEEE Trans. Image Process*, 16(10):2590–2606, 2007.

[35] K.J.Biehl, F.Kong, F.Dehdashti, J.Jin, S.Mutic, I.E. Naqa, B.A.Siegel, and J.D.Bradley. ^{18}F-FDG PET definition of gross tumor volume for radiotherapy of non-small cell lung cancer: is a single standardized uptake value threshold approach appropriate? *Journal of Nuclear Medicine*, 47(11):1808–1812, 2006.

[36] K.Lee, M.Niemeijer, M.Garvin, Y.Kwon, M.Sonka, and M.Abramoff. Segmentation of the optic disc in 3-D OCT scans of the optic nerve head. *IEEE Trans. Medical Imaging*, 29(1), 2010.

[37] K.Li, X.Wu, D.Z.Chen, and M.Sonka. Optimal surface segmentation in volumetric images - a graph-theoretic approach. *IEEE Trans. Pattern Anal. Machine Intell*, 28(1):119–134, 2006.

[38] K.Palagyi, E.Sorantin, E.Balogh, A.Kuba, C.Halmai, B.Erdhelyi, and K.Hausegger. A sequential 3D thinning algorithm and its medical applications. In *Proc. Biennial International Conference on Information Processing in Medical Imaging (IPMI)*, volume 2082 of LNCS, pages 409–415, 2001.

[39] L.Grady. Random walks for image segmentation. *IEEE Trans. Pattern Anal. Machine Intell*, 28(11), 2006.

[40] L.Grady and E.L.Schwartz. Isoperimetric graph partitioning for image segmentation. *IEEE Trans. Pattern Anal. Machine Intell*, 28(3):469–475, 2006.

[41] M.Bray, P.Kohli, and P.Torr. Posecut: Simultaneous segmentation and 3D pose estimation of humans using dynamic graph-cuts. In *European Conference on Computer Vision*, 2006.

[42] M.Chen and R.A.Siochi. A clinical feasibility study on respiratory sorted megavoltage cone beam CT. In *International Workshop on Pulmonary Image Analysis*, 2010.

[43] M.Chen and R.A.Siochi. Diaphram motion quantification in megavoltage conebeam CT projection images. *Medical Physics*, 37:2312–20, 2010.

[44] M.Costa, H.Delingette, S.Novellas, and N.Ayache. Automatic segmentation of bladder and prostate using coupled 3D deformable models. In *Proc. International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 252–260, 2007.

[45] M.D.Bruijne, B.V.Ginneken, M.A.Viergever, and W.J.Niessen. Interactive segmentation of abdominal aortic aneurysms in CTA images. *Medical Image Analysis*, 8(2):127–158, 2003.

[46] M.Garvin, M.D.Abramoff, X.Wu, S.R.Russell, T.L.Burns, and M.Sonka. Automated 3-D intraretinal layer segmentation of macular spectral-domain optical coherence tomography images. *IEEE Trans. Medical Imaging*, 28(9):1436–1447, 2009.

[47] M.Garvin, X.Wu, M.Abramoff, R.Kardon, and M.Sonka. Incorporation of regional information in optimal 3-D graph search with application for intraretinal layer segmentation of optical coherence tomograph images. In *Proc. Biennial International Conference on Information Processing in Medical Imaging (IPMI)*, 2007.

[48] M.Kass, A.Witkin, and D.Terzopoulos. Snakes: active contour models. *Int. Journal of Computer Vision*, 1(4):321–331, 1988.

[49] M.Rousson, A.Khamene, M.Diallo, J.Celi, and F.Sauer. Constrained surface evolutions for prostate and bladder segmentation in CT images. In *Computer Vision for Biomedical Image Applications (CVBIA)*, 2005.

[50] M.Rousson and N.Paragios. Shape priors for level set representations. In *European Conference on Computer Vision*, 2002.

[51] M.Sonka, V.Hlavac, and R.Boyle. *Image Processing, Analysis and Machine Vision*. Thomson Learning, 2007.

[52] N.Sakalihasan, R.Limet, and O.D.Defawe. Abdominal aortic aneurysm. *Lancet*, 365:1577–1589, 2005.

[53] N.Vu and B.Manjunath. Shape prior segmentation of multiple objects with graph cuts. In *Proc. IEEE International Conference on Computer Vision and Pattern Recognition*, 2008.

[54] O.Bodur, L.Grady, A.Stillman, R.Setser, G.Funka-Lea, and T.O'Donnell. Semi-automatic aortic aneurysm analysis. In *Medical Imaging 2007: Physiology, Function, and Structure from Medical Images. Proceedings of the SPIE*, volume 6511, page 65111G, 2007.

[55] O.Morin, A.Gillis, J.Chen, M.Aubin, M.Bucci, M.Roach, and J.Pouliot. Megavoltage cone-beam CT: system description and clinical applications. *Medical Dosimetry*, 31(1):51–61, 2006.

[56] P.Felzenszwalb and D.Huttenlocher. Efficient graph-based image segmentation. *Int. Journal of Computer Vision*, 59(2), 2004.

[57] P.Felzenszwalb and D.Huttenlocher. Pictorial structures for object object recognition. *International Journal of Computer Vision*, 61(1):55–79, 2005.

[58] P.Kohli and P.Torr. Efficiently solving dynamic markov random fields using graph cuts. In *Proc. IEEE International Conference on Computer Vision*, 2005.

[59] P.K.Saha, B.Das, and F.W.Wehrli. An object class-uncertainty induced adaptive force and its application to a new hybrid snake. *Pattern Recognition*, 40:2656–2671, 2007.

[60] P.K.Saha and J.K.Udupa. Optimum threshold selection using class uncertainty and region homogeneity. *IEEE Trans. Pattern Anal. Machine Intell*, 23:689–706, 2001.

[61] Q.Song, X.Wu, Y.Liu, M.Haeker, and M.Sonka. Simultaneous searching of globally optimal interacting surfaces with shape priors. In *Proc. IEEE International Conference on Computer Vision and Pattern Recognition*, 2010.

[62] Q.Song, X.Wu, Y.Liu, M.Smith, J.M.Buatti, and M.Sonka. Optimal graph search segmentation using arc-weighted graph for simultaneous surface detection of bladder and prostate. In *Proc. International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 827–835, 2009.

[63] R.Gonzalez and R.Woods. *Digital image processing*. Prentice Hall, 2002.

[64] R.Haralick and L.Shapiro. Image segmentation techniques. 29:100–132, 1985.

[65] R.Malladi and J.A.Sethian. A real-time algorithm for medical shape recovery. In *Proc. IEEE International Conference on Computer Vision*, pages 304–310. IEEE Computer Society Press, 1998.

[66] S.Andrews, C.McIntosh, and G.Hamarneh. Convex multi-region probabilistic segmentation with shape prior in the isometric log-ratio transformation space. In *Proc. IEEE International Conference on Computer Vision*, 2011.

[67] S.Avidan and A.Shamir. Seam carving for content-aware image resizing. *ACM Transactions on Graphics*, 26(3), 2007.

[68] S.Klein, M.Staring, K.Murphy, M.Vierggever, and J.Pluim. Elastix: a toolbox for intensity-based medical image registration. *IEEE Trans. Medical Imaging*, 29:196–205, 2010.

[69] S.Osher and J.Sethian. Fronts propagating with curvature dependent speed: algorithms based on the hamilton-jacobi formulation. 79:12–49, 1988.

[70] S.Sun, C.Bauer, and R.Beichel. Automated 3D segmentation of lungs with lung cancer in CT data using a novel reboust active shape model approach. *IEEE Trans. Medical Imaging (accepted)*.

[71] T.Chan and L.Vese. Active contours without edges. 79:12–49, 1988.

[72] T.Chan, S.Esedoglu, and M.Nikolova. Algorithms for finding global minimizers of image segmentation and denoising models. *SIAM Journal on Applied Mathematics*, 66(5):1632–1648, 2006.

[73] T.Chan and W.Zhu. Level set based shape prior segmentation. In *Proc. IEEE International Conference on Computer Vision and Pattern Recognition*, 2005.

[74] T.Heimann, S.Munzing, H.P.Meinzer, and I.Wolf. A shape-guided deformable model with evolutionary algorithm initialization for 3D soft tissue segmentation. In *Proc. Biennial International Conference on Information Processing in Medical Imaging (IPMI)*, volume LNCS-3565, pages 566–577, 2007.

[75] T.Pock, A.Chambolle, D.Cremers, and H.Bischof. A convex relaxation approach for computing minimal partitions. In *Proc. IEEE International Conference on Computer Vision and Pattern Recognition*, pages 810–817, 2009.

[76] T.Pock, T.Schoenemann, G.Graber, H.Bischof, and D.Cremers. A convex formulation of continuous multi-label problems. In *Proc. European Conference on Computer Vision*, pages 792–805, 2008.

[77] U.Malsch, C.Thieke, and R.Bendl. Fast elastic registration for adaptive radiotherapy. In *Proc. International Conference on Medical Image Computing and Computer-Assisted Intervention*, volume 2, pages 612–619, 2006.

[78] V.Caselles, R.Kimmel, and G.Sapiro. Geodesic active contours. *International journal on computer vision*, 22:61–97, 1997.

[79] V.Kolmogorov and Y.Boykov. Applications of parametric maxflow in computer vision. In *Proc. IEEE International Conference on Computer Vision*, pages 1–8. IEEE Computer Society Press, 2007.

[80] V.Potesil, X.Huang, and X.S.Zhou. Automated tumor delineation using joint PET/CT information. In *Proc. SPIE International Symposium on Medical Imaging: Computer-Aided Diagnosis*, volume 6514, 2007.

[81] X.Liu, Q.Song, P.Mendonca, X.Tao, and R.Bhotika. Organ labeling using anatomical model-driven global optimization. In *IEEE Internal Conference on Healthcare Informatics, Imaging and Systems Biology (HISB)*, pages 338–345, 2011.

[82] X.Wu and D.Z.Chen. Optimal net surface problems with applications. In *Proc. 29th International Colloquium on Automata, Languages and Programming*, pages 1029–1042, 2002.

[83] X.Xu, M.Niemeijer, Q.Song, M.Sonka, M.Garvin, and J.Reinhardt. Vessel boundary delineation on fundus images using graph-based approach. *IEEE Trans. Medical Imaging*.

[84] Y.Boykov and G.Funka-Lea. Graph cuts and efficient N-D image segmentation. *Int. Journal of Computer Vision*, 70(2):109–131, 2006.

[85] Y.Boykov and M.Jolly. Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images. In *Proc. IEEE International Conference on Computer Vision*, pages 105–112, 2001.

[86] Y.Boykov, O.Veksler, and R.Zabih. Fast approximate energy minimization via graph cuts. *IEEE Trans. Pattern Anal. Machine Intell*, 23(11), 2001.

[87] Y.Boykov and V.Kolmogorov. Computing geodesics and minimal surfaces via graph cuts. In *Proc. IEEE International Conference on Computer Vision*, volume 1, pages 26–33, 2003.

[88] Y.Li, J.Sun, C.Tang, and H.Shum. Lazy snapping. In *ACM Transactions on Graphics (SIGGRAPH)*, 2004.

[89] Y.Li, J.Sun, and H.Shum. Video object cut and paste. In *ACM Transactions on Graphics (SIGGRAPH)*, 2005.

[90] Y.Xia, S.Eberl, L.Wen, M.Fulham, and D.Feng. Dual-modality brain PET-CT image segmentation based on adaptive use of functional and anatomical information. *Journal of Computerized Medical Imaging and Graphics*, 2011.

[91] Y.Yin, X.Zhang, R.Williams, X.Wu, D.Anderson, and M.Sonka. LOGISMOS - Layered Optimal Graph Image Segmentation of Multiple Objects and Sur-

faces: Cartilage segmentation in the knee joints. *IEEE Trans. Medical Imaging*, 29(12):2023 – 2037, 2010.