2008

Automated 3-D segmentation and analysis of retinal optical coherence tomography images

Mona Kathryn Garvin
University of Iowa

Recommended Citation

Garvin, Mona Kathryn. "Automated 3-D segmentation and analysis of retinal optical coherence tomography images." PhD (Doctor of Philosophy) thesis, University of Iowa, 2008.
https://doi.org/10.17077/etd.crsmiciw

Follow this and additional works at: https://ir.uiowa.edu/etd
Part of the Biomedical Engineering and Bioengineering Commons

AUTOMATED 3-D SEGMENTATION AND ANALYSIS OF RETINAL OPTICAL COHERENCE TOMOGRAPHY IMAGES

by

Mona Kathryn Garvin

An Abstract

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

May 2008

Thesis Supervisor: Professor Milan Sonka

ABSTRACT

Optical coherence tomography (OCT) is becoming an increasingly important modality for the noninvasive assessment of a variety of ocular diseases such as glaucoma, diabetic macular edema, and age-related macular degeneration. Even though individual layers of the retina are visible on OCT images, current commercial quantitative assessment is limited to measuring the thickness of only one layer. Because each intraretinal layer may be affected differently by disease, an intraretinal layer segmentation approach is needed to enable quantification of individual layer properties, such as thickness or texture. Furthermore, with the latest generation of OCT scanner systems producing true volumetric image data, processing these images using 3-D methods is important for maximal extraction of image information.

In this thesis, an optimal 3-D graph search approach for the intraretinal layer segmentation of OCT images is presented. It is built upon the optimal 3-D multiple surface graph-theoretic approach presented by Li *et al.* (K. Li, X. Wu, D. Z. Chen, and M. Sonka, “Optimal surface segmentation in volumetric images – a graph-theoretic approach,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 1, pp. 119–134, 2006). In this method, multiple surfaces can be found simultaneously by transforming the 3-D segmentation problem into finding a minimum-cost closed set in a corresponding vertex-weighted geometric graph. However, the original formulation of this approach did not incorporate varying feasibility constraints or true regional information, two extensions that would aid in the intraretinal layer segmentation of OCT images.

Thus, the major contributions of this thesis include: 1) extending the optimal 3-D graph-theoretic segmentation approach to allow for the incorporation of varying feasibility constraints and regional information, 2) developing a method for learning varying constraints and cost functions from examples for use in the approach, 3) developing and validating a method for the 3-D segmentation of intraretinal layers in both time-domain and spectral-domain OCT images (achieving error rates comparable to that of multiple human experts), and 4) analyzing layer thickness properties in normal subjects and in patients with anterior ischemic optic neuropathy (AION).

Abstract Approved: ______________________________

Thesis Supervisor

____

Title and Department

____

Date

AUTOMATED 3-D SEGMENTATION AND ANALYSIS OF RETINAL OPTICAL COHERENCE TOMOGRAPHY IMAGES

by

Mona Kathryn Garvin

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

May 2008

Thesis Supervisor: Professor Milan Sonka

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

Mona Kathryn Garvin

has been approved by the Examining Committee for the thesis requirement for the Doctor of Philosophy degree in Biomedical Engineering at the May 2008 graduation.

Thesis Committee: ______________________________
Milan Sonka, Thesis Supervisor

____
Michael D. Abràmoff

____
Edwin L. Dove

____
Joseph M. Reinhardt

____
Todd E. Scheetz

ACKNOWLEDGEMENTS

First, I would like to thank my adviser, Dr. Milan Sonka, for the guidance and support he has given me throughout this work. Dr. Michael Abràmoff also deserves special mention as someone who has been a significant help. I would like to thank all of the ophthalmologists I have worked with for both their advice and informative discussions. For example, I have interacted with Dr. Randy Kardon throughout the entire project and I have learned a great deal from him. Other collaborators from the Department of Ophthalmology and Visual Sciences include Dr. Stephen Russell, Dr. Young Kwon, Dr. Chris Johnson, and Dr. Vinay Shah. Dr. Xiaodong Wu is another person I would like to thank for his role in graph-based discussions.

A special thanks also goes to Dr. Michael Abràmoff, Dr. Stephen Russell, and Dr. Randy Kardon for their willingness to manually trace time-domain OCT images for validation purposes. I am most gracious to Dr. Stephen Russell and Dr. Michael Abràmoff for also tracing spectral-domain OCT images. This is something that took a significant amount of time and I appreciate it greatly. Dr. Trudy Burns deserves a special thanks for helping with the statistical analysis of the spectral OCT results. Towards the end of my Ph.D. work, I also enjoyed working with Dr. Meindert Niemeijer. In addition, I would like to thank Dr. Edwin Dove, Dr. Joseph Reinhardt, and Dr. Todd Scheetz for serving on my committee.

I would like to think a number of organizations and individuals for providing permission for me to include previously published material. For example, Springer Science and Business Media has kindly allowed me to include figures and tables from my conference papers. IEEE has allowed me to include figures and tables from my *IEEE Transactions on Medical Imaging* paper. In addition, MIT Press has allowed me to include the pseudocode of algorithms from an algorithms textbook. Mark Erickson and Helga Kolb have both given me permission to reuse an image for the background chapter of my thesis.

This work was supported, in part, by NIH-NIBIB grant R01-EB004640-01, NIH-NEI grant R01-EY017066, a grant from the Veterans Administration (Merit Review and Rehabilitation Division), and an unrestricted grant from Research to Prevent Blindness, New York, NY. In addition, the images used for the spectral-domain OCT segmentation and analysis were kindly provided by Carl Zeiss Meditec, Inc.

Finally, last but not least, I would like to extend a special thanks to my husband, Justin W. Garvin, for his love and support throughout this process.

TABLE OF CONTENTS

CHAPTER

LIST OF TABLES

Table

LIST OF FIGURES

Figure

LIST OF ALGORITHMS

Algorithm

CHAPTER 1
INTRODUCTION

Since its first introduction in 1991 [1], optical coherence tomography (OCT) has become an increasingly important modality. One of its greatest impacts has been within the ophthalmic community where the high-resolution cross-sectional images resulting from OCT scanners are being used for the diagnosis and management of a variety of ocular diseases such as glaucoma, diabetic macular edema, and age-related macular degeneration. Currently used time-domain commercial scanners (e.g., Stratus OCT-3, Carl Zeiss Meditec, Inc., Dublin, CA) provide the ability to scan up to six cross-sectional slices of the retina, each with an axial resolution of 8–10 μm. The rapid progress of OCT technology, such as the development of Fourier-domain acquisition schemes [2], has enabled the latest generation of commercial scanners to acquire many more slices in a similar time-frame, thus allowing the acquisition of true volumetric image data.

With increasing image data becoming available to ophthalmologists, the need and potential impact of image analysis methods is high. Although intraretinal layers are visible on such images, current quantitative analysis typically only focuses on measuring the thickness of one layer of the retina, thus leaving potentially clinically important information unexplored. For example, in cases where the layers may be affected differently by disease processes (e.g., one layer may thin due to neuron loss while another may thicken due to edema), it would make the most sense to be able to quantify the properties of the layers individually.

Furthermore, with volumetric data available, taking advantage of contextual information and performing the intraretinal layer segmentation in 3-D (as opposed to repeatedly performing a 2-D segmentation approach) will be important. However, as is well known in the medical imaging community for applications in other imaging modalities such as computed tomography (CT) or magnetic resonance imaging (MRI), the ability to efficiently and accurately segment images in 3-D or 4-D remains a challenging problem. For example, in order to make many 3-D segmentation approaches practical, optimality of the resulting segmentation is often not guaranteed because of the computational demands in searching a large solution space [3–5]. Nevertheless, Li *et al.* [6] recently presented a low-polynomial time graph-based approach for the *optimal* multi-surface segmentation of 3-D or higher dimensional data. This was an extension of the approach for the optimal detection of single surfaces presented by Wu and Chen [7] to the multiple surface case.

Such a 3-D graph search method [6, 7] is well-suited for the intraretinal layer segmentation for a variety of reasons. Perhaps the most important include: 1) the ability to find an optimal set of surfaces with respect to a cost function in 3-D and 2) the ability to find multiple surfaces simultaneously. However, as originally presented, the optimal graph search approach did not allow for the incorporation of true regional image information or varying constraints. Such extensions would be important for flexibility in the cost function design and incorporating prior shape knowledge, respectively.

Thus, after extending the graph search approach to allow for varying constraints and the incorporation of regional image information, this thesis is concerned with developing and validating a 3-D method for the intraretinal layer segmentation of OCT images. In addition, the developed approach is utilized in applications of clinical interest to help show how an intraretinal analysis may provide additional clinical information over that of the traditional method of analyzing only one layer of the retina.

1.1 Specific aims

In particular, the specific aims of the work described in this thesis were as follows:

- **Aim 1:** Develop a method for the incorporation of regional information and varying constraints into the optimal 3-D graph search.
- **Aim 2:** Develop a method for learning cost functions and varying constraints from examples for use in the 3-D graph search.
- **Aim 3:** Develop and validate a method for the 3-D segmentation of intraretinal layers from OCT images.
- **Aim 4:** Use the intraretinal layer segmentation approach to (a) help determine which macular layers show significant changes in thickness in unilateral anterior ischemic optic neuropathy (AION) subjects and (b) perform pilot analysis of localized thickness and thickness variability of individual layers in normal subjects.

Note that in addition to the contributions to the ophthalmic community provided by all the aims, the medical image analysis community will also be particularly interested in Aims 1 and 2 because of the ability to apply the graph search extensions to other applications as well.

1.2 Thesis overview

This thesis is divided into 8 chapters. An overview of each of the remaining chapters follows:

- Chapter 2 provides additional background information, including a discussion of age-related eye diseases, optical coherence tomography, the layers of the retina, and prior approaches for the segmentation of OCT images.
- Chapter 3 provides a review of optimal graph search concepts and describes two extensions to the optimal 3-D graph search: the incorporation of varying constraints and the incorporation of regional information.
- Chapter 4 provides a summary of the methods and validation completed for segmenting time-domain OCT images.

- Chapter 5 provides a summary of the methods and validation completed for segmenting spectral-domain OCT images.
- Based on the segmentation approaches described in Chapters 4 and 5, Chapter 6 provides a regional analysis of localized thickness measurements in unilateral AION subjects and a pilot study of the localized thickness and thickness variability in normal subjects.
- Chapter 7 provides a general discussion and future directions.
- Chapter 8 provides some concluding remarks.

In addition, the appendix provides some additional background on relevant graph-theoretic concepts.

CHAPTER 2
BACKGROUND AND SIGNIFICANCE

2.1 Age-related eye diseases

Based on demographics from the 2000 US Census, 1 in 28 Americans older than 40 years are currently affected by blindness or low vision, with the prevalence expected to increase markedly in the next 20 years as the population ages [8]. The four leading causes of blindness and visual impairment affecting older Americans are age-related macular degeneration (AMD), cataract, diabetic retinopathy, and glaucoma [9]. While cataract affects the front of the eye (clouding of the lens), the remaining three major eye-related diseases affect the retina and/or optic nerve head at the back of the eye. An important difference between cataract and the other three eye diseases is that cataract is usually noticed by the patient early enough for adequate treatment, while the early forms of AMD, glaucoma, and diabetic retinopathy are usually not noticed by the patient, causing substantial damage unless diagnosed early.

Fig. 2.1 provides a simulated view of a scene as might be observed by a patient with each of these three diseases. As illustrated Fig. 2.1(b), age-related macular degeneration (AMD) often affects the area of central vision. There are actually two forms of AMD: dry AMD and wet AMD. Dry AMD is the chronic, usually slowly progressive form, while wet (exudative, neovascular) AMD is the less common, but more acute and severe form of the disease. Wet AMD, if diagnosed too late and/or left untreated, will rapidly progress to major visual loss. This process, known as choroidal neovascularization, involves the growth of abnormal, leaking blood vessels under the macula (part of the retina responsible for central vision – see Fig. 2.2), ending with atrophy and scarring of the macula.

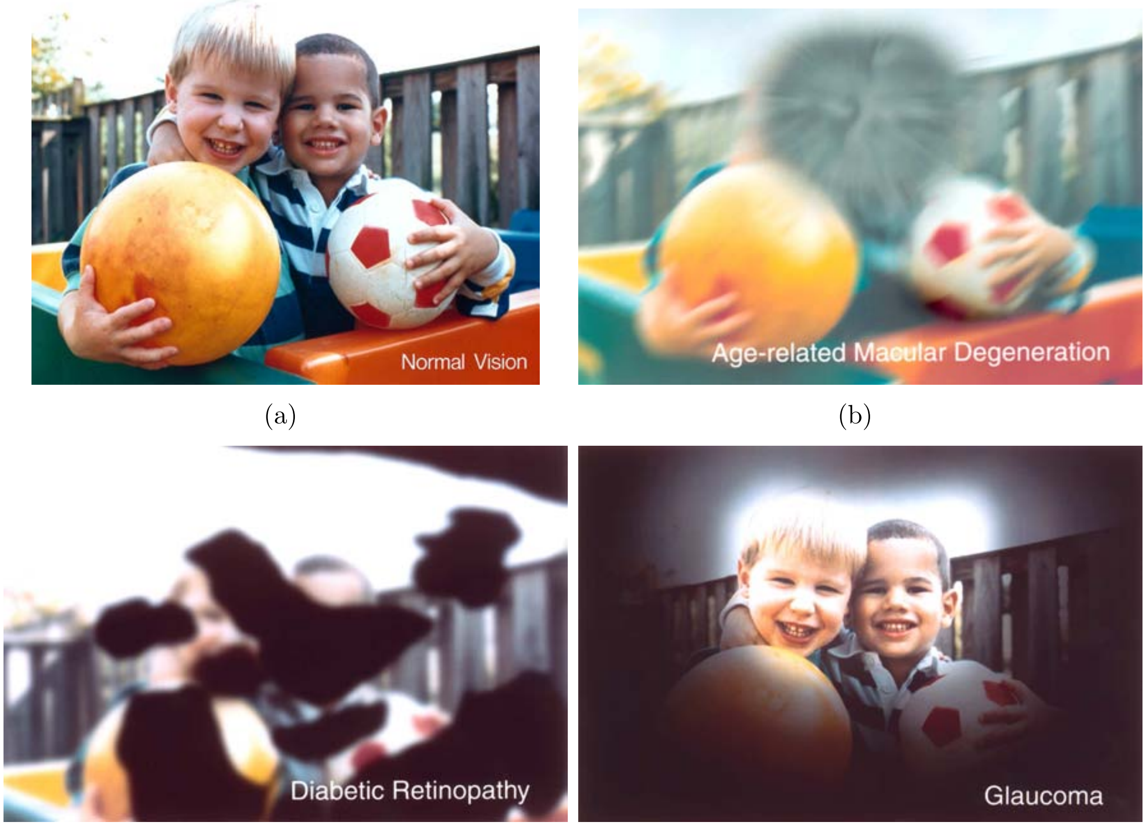
Diabetic retinopathy (Fig. 2.1(c)) is a complication of diabetes, which if undiagnosed or left untreated, can lead to hemorrhage, scarring and atrophy in the retina through vascular damage and ischemia. In proliferative diabetic retinopathy, abnormal blood vessels lead to hemorrhages and destruction of retinal tissue. Diabetic macular edema (DME) occurs when the abnormal blood vessels leak, leading to swelling of the macular tissue and damage to photoreceptors.

Glaucoma (Fig. 2.1(d)) is an optic neuropathy, involving gradual damage to the retinal ganglion cells and their axons that lead into the optic nerve and further into the visual areas of the brain. Because of the redundancy of the human visual system, which is more pronounced in central vision, this initially leads to peripheral vision loss.

2.1.1 Current clinical management

The three eye diseases mentioned in the previous section are usually managed with a combination of early diagnosis, medical and surgical treatment, and close follow-up. For AMD, new effective treatments have recently become available in the form of anti-VEGF (VEGF = vascular endothelial growth factor) treatments. A combination of fundus photography, indirect ophthalmoscopy, stereo biomicroscopy of the macula, fluorescein angiography, and optical coherence tomography are used to make the diagnosis and manage the patient with anti-VEGF drugs as well as laser surgery. For diabetic retinopathy, fundus photography, indirect ophthalmoscopy, stereo biomicroscopy of the macula, fluorescein angiography, and optical coherence tomography are used to diagnosis and manage the patient. While proliferative diabetic retinopathy is usually managed by laser and vitreoretinal surgery, DME is usually managed by laser surgery, anti-inflammatory drugs, and anti-VEGF drugs.

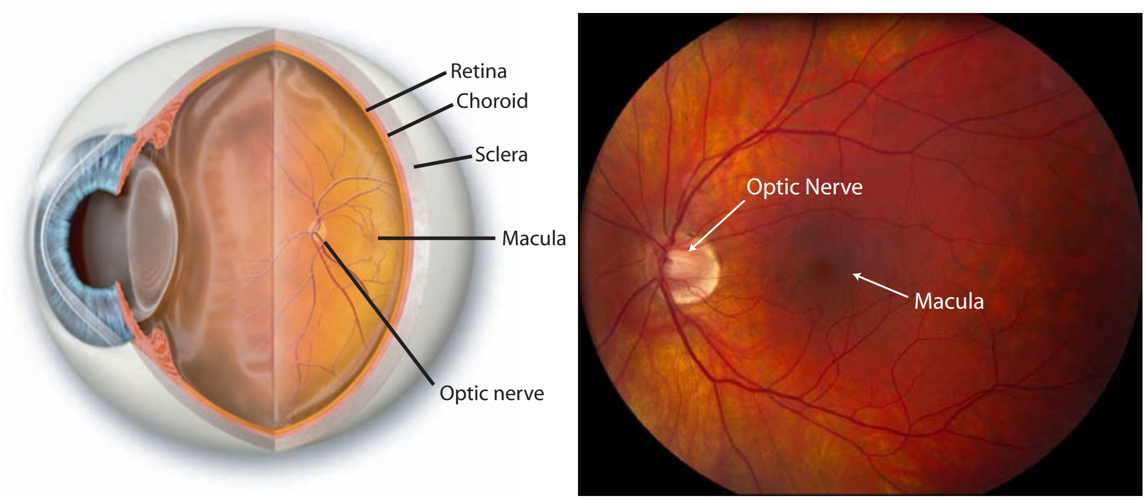
In glaucoma, intraocular pressure (IOP) measurements (tonometry), examining the optic nerve head at the back of the eye using indirect ophthalmoscopy and optic disc stereo biomicroscopy, examining the front of the eye using anterior segment slit lamp biomicroscopy, checking whether the angle between the iris and cornea is "open" or "closed" using gonioscopy, and measuring visual fields using static or kinetic



(a) (b)

(c) (d)

Figure 2.1: Simulated view of a scene as might be perceived with patients with age-related eye diseases. (a) Normal view. (b) View from the perspective of someone with wet age-related macular degeneration. (c) View from the perspective of someone with diabetic retinopathy. (d) View from the perspective of someone with glaucoma. (Images courtesy of National Eye Institute, National Institutes of Health.)



(a) (b)

Figure 2.2: The eye. (a) 3-D illustration of the eye (obtained courtesy of Mark Erickson from JirehDesign.com). (b) Fundus photograph with labeled macula and optic nerve.

perimetry are used, as well as retinal never fiber layer imaging techniques including scanning laser polarimetry (e.g., GDx), confocal scanning laser ophthalmoscopy (e.g., Heidelberg Retinal Tomography or HRT), and optical coherence tomography (OCT).

Most of these conventional techniques are based on en-face examination of the retina. However, all three diseases affect the thickness of structures as well, such as macular thickening in diabetic macular edema and AMD, and retinal nerve fiber and ganglion cell layer thinning in glaucoma. Though stereo biomicroscopy can result in a subjective assessment of thickness differences of retinal structures, only with the advent of scanning laser polarimetry, confocal SLO and especially OCT, objective measurements of the thickness of retinal structures are now clinically achievable [10]. More importantly, evidence is now available that these objective measures are clinically relevant.

2.2 Optical coherence tomography

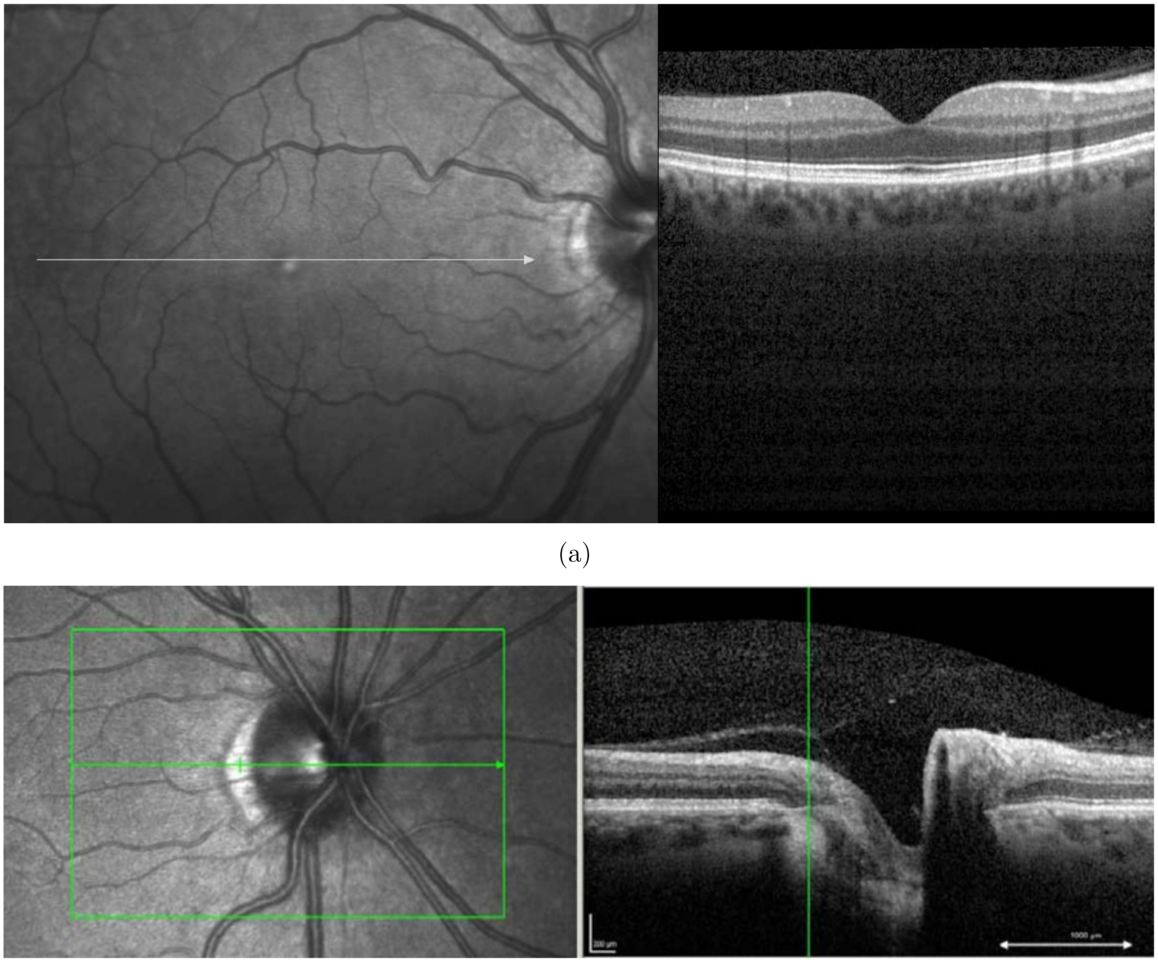
With its ability to obtain high-resolution cross-sectional views of the retina or optic nerve head (Fig. 2.3), optical coherence tomography is becoming an increasingly important modality for assessing ocular diseases such as AMD, DME, and glaucoma. Until recently, all commercial systems were time-domain systems (Fig. 2.4(a)). Now spectral-domain systems (Fig. 2.4(b)) have also become available, thus allowing the acquisition of substantially more data. A brief overview of each of these systems is provided in the next few sections.

2.2.1 Time-domain OCT

The retinal images of most current commercial OCT systems such as the Stratus OCT-3 (Fig. 2.4(a)) of Carl Zeiss Meditec are acquired in the time-domain. The acquisition speed of such systems limits the number of slices that can be reasonably acquired in a scanning sequence to about six. As illustrated in Fig. 2.5, the two regions of the retina commonly scanned are the macula and the peripapillary region (region near the optic disc). One common scanning protocol for acquiring scans in the macular region involves the acquisition of six linear radial scans in a “spoke pattern” centered at the fovea (e.g., the Fast Macular protocol on the Stratus OCT-3). When acquiring scans surrounding the optic disc, it is common to use a number of circular scans. An example set of six radial images from a macular OCT series can be found in Fig. 2.6.

2.2.2 Spectral-domain OCT

The latest generation (available in late 2007) of commercial scanners are spectral OCT scanners [2], thus allowing true volumetric data to be acquired. An example spectral-domain OCT machine (Cirrus, from Carl Zeiss Meditec, Inc., Dublin, CA) is shown in Fig. 2.4(b). For example, while a typical time-domain OCT image from a current commercial scanner might have dimensions of $6 \times 128 \times 1024$ pixels, a typical


(a)

(b)

Figure 2.3: Cross-sectional views of the retina and optic nerve head using OCT. (a) Slice of macula using Spectralis (Heidelberg Engineering). (b) Slice of optic nerve head using Spectralis (Heidelberg Engineering).

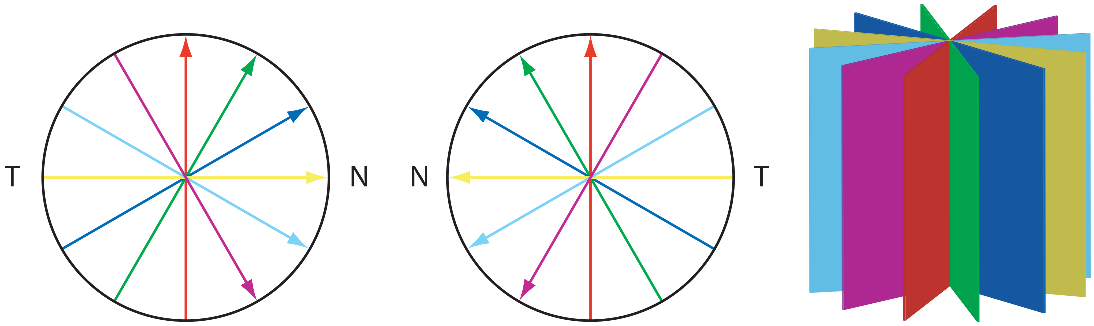
(a) (b)

Figure 2.4: Example OCT scanner systems. (a) Stratus OCT-3 (time-domain system by Carl Zeiss Meditec, Inc., Dublin, CA). (b) Cirrus (spectral-domain system by Carl Zeiss Meditec, Inc., Dublin, CA).

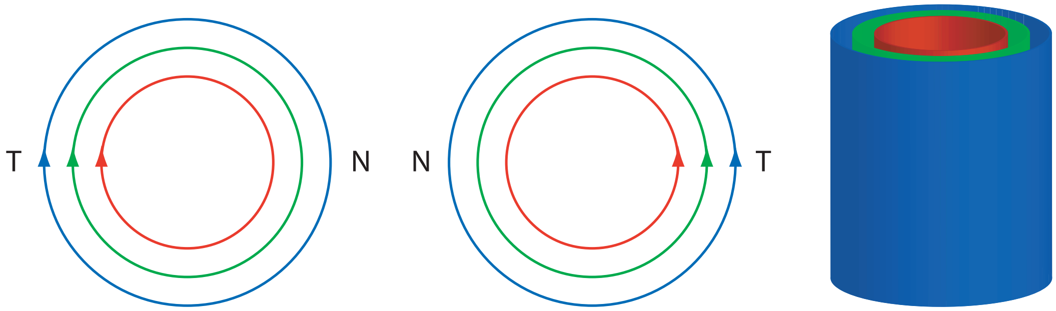
spectral OCT image might have dimensions of 200 × 200 × 1024 pixels (Fig. 2.7). Other enhancements to OCT systems, such as the use of adaptive optics [11] and the use of better light sources for producing "ultrahigh-resolution" images [12], have also been reported in the literature.

2.3 The layers of the retina

Covering the inside of most of the eye (Fig. 2.8), the retina is a multilayered structure responsible for transforming light energy into neural signals for further use by the brain. In very general terms, the processing of light starts with the light-sensitive photoreceptor cells (rods and cones), which are actually located in the outer portion of the retina (away from the incoming light). These cells convert the light signal into action potentials that are transmitted by the bipolar neurons in the central layers of the retina to the ganglion cells of the inner retina. It is the axons of the ganglion cells that eventually exit the eye to form the optic nerve. Other cells in the retina, such as horizontal cells, amacrine cells and interplexiform neurons, also help in the processing of the neural signal at a local level. Neuroglial cells (such as



(a) Macular OD (b) Macular OS (c)



(d) Circular OD (e) Circular OS (f)

Figure 2.5: Schematic view of the macular (a–c) and circular (d–f) scanning protocols on the Stratus OCT-3. (a) Scans in macular series on the right eye. (N = nasal, T = temporal.) (b) Scans in macular series on the left eye. (e) Visualization of acquired macular scans for one eye in 3-D. Each color represents a different 2-D scan. (d) Scans in peripapillary circular series on the right eye. (e) Scans in peripapillary circular series on the left eye. (f) Visualization of acquired circular scans for one eye in 3-D.

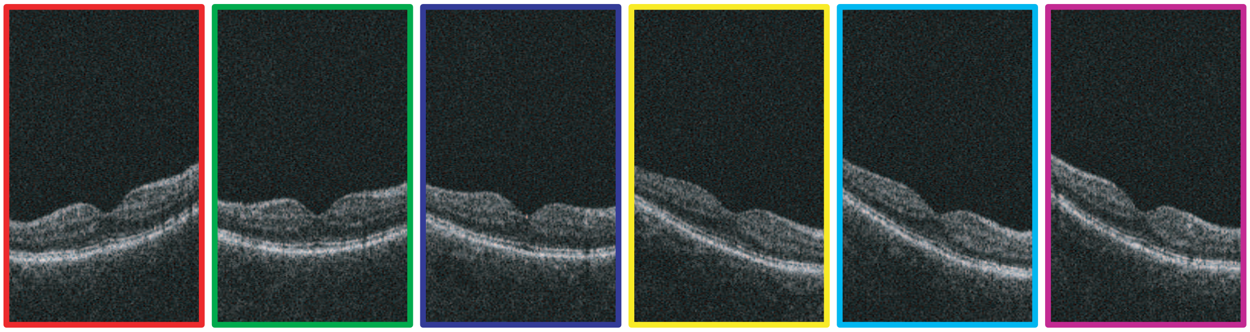


Figure 2.6: Example six raw scans in a macular scan series. Note that the colored borders correspond to those found in Fig. 2.5(a)–(c).

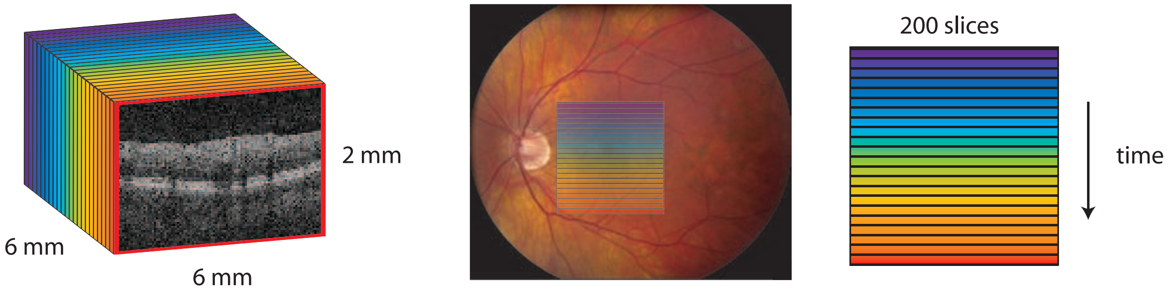
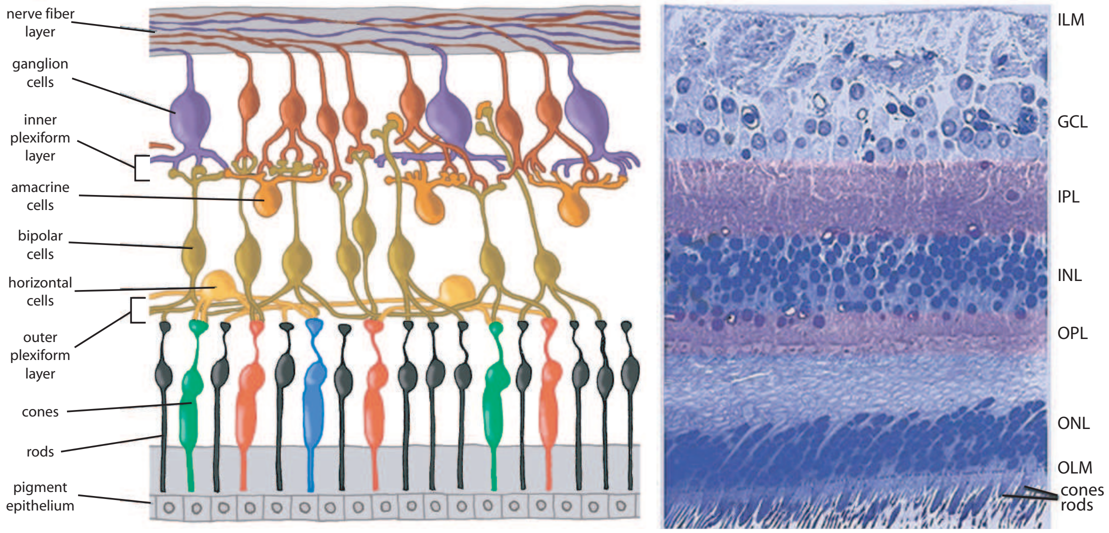


Figure 2.7: Example dimensions of a 3-D spectral OCT image. A typical spectral OCT images covers a volume of $6 \times 6 \times 2$ mm^3 ($200 \times 200 \times 1024$ cubic voxels) using one of the protocols on the Cirrus machine by Carl Zeiss Meditec.

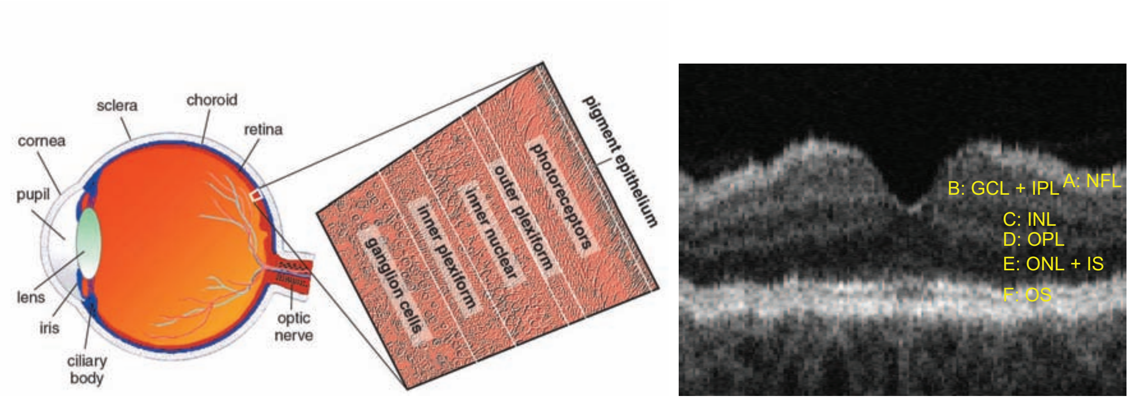
Müller cells) provide structure and support [13]. Many of the cells of the retina are illustrated in Fig. 2.8(a).

Based on its appearance from light microscopy (Fig. 2.8(b)), the retina is traditionally considered to be composed of the following ten major "layers" (starting with the outermost layer) [13]:

- Reginal pigment epithelium (RPE): single layer of pigmented hexagonal cells.
- Photoreceptor layer: the outer (containing the light-sensitive discs) and inner segments of rods and cones.
- External (or outer) limiting membrane (ELM or OLM): intercellular junctions between photoreceptor cells and between photoreceptor and Müller cells (not an actual membrane).
- Outer nuclear layer (ONL): rod and cone cell bodies.
- Outer plexiform layer (OPL): synapses between photoreceptor cells and cells from the inner nuclear layer.
- Inner nuclear layer (INL): cell bodies of bipolar cells, horizontal cells, amacrine cells, interplexiform neurons, Müller cells, and some displaced ganglion cells.
- Inner plexiform layer (IPL): synaptic connections between bipolar cell axons and ganglion cell dendrites.



(a) Sketch of retinal layers (b) Light micrograph



(c) 2-D illustration of eye (d) OCT scan

Figure 2.8: Different views of the retina. (a) Schematic illustration of cellular layers of retina (copied with permission from Helga Kolb [14]). (b) Light micrograph of a vertical scan through central human retina (copied with permission from Helga Kolb [15]) (c) Cross section of the eye with illustration of the retina (modified with permission from Helga Kolb [15]) (d) OCT view of macular retina.

- Ganglion cell layer (GCL): mostly ganglion cell bodies.
- Nerve fiber layer (NFL): ganglion cell axons.
- Internal limiting membrane (ILM): innermost membrane of retina separating the retina from the vitreous.

Intraretinal layers are also visible from optical coherence tomography images as shown in the example image of Fig. 2.8(d). Note that the precise anatomical correspondence of the layers visible in OCT images is not known. The anatomical labels are our current presumption based on comparisons with histology and images from higher-resolution OCT scanners published in the literature [16].

2.4 Clinical motivation for intraretinal layer segmentation

One general clinical motivation for the availability of an intraretinal layer segmentation approach is to enable the creation of better quantitative tools for helping ophthalmologists diagnose and monitor ocular diseases. Because of the large amounts of information provided within the images, the potential impact of such a 3-D segmentation approach for many different applications is high. Some example applications include:

1. Earlier detection of ganglion cell loss in cases of concurrent optic nerve swelling: In cases of optic disc edema (as is often associated with acute stages of optic neuropathies), the thickness of the retinal nerve fiber layer on peripapillary scans (i.e., circular scans surrounding the optic disc) may not reflect the true neural loss. It is possible that information in the macular scans, such as the thickness of the ganglion cell layer from an intraretinal segmentation, may be able to detect such loss at an earlier stage.
2. Better understanding of what quantitative features on OCT images predict visual disfunction in diseases such as AMD, DME, and glaucoma: As a long term goal, it would be desirable to be able to predict future disfunction from a

patient's current history of OCT images and other data. This would help guide ophthalmologists on treatment options and understanding the progression of the diseases. However, it would also be useful just to understand what features from OCT images are currently associated with visual disfunction. To do this, an intraretinal layer segmentation would be necessary in order to extract individual layer properties such as thickness or texture.

More applications and additional clinical background information will be provided in later chapters (e.g., see Chapter 7).

2.5 Prior approaches for the segmentation of OCT images

Even though multiple layers of the retina are identifiable on OCT images, commercially available systems currently only segment and provide thickness measurements for one layer of the retina (i.e., the total retina on macular scans and the retinal nerve fiber layer on peripapillary scans). Although we do not know the proprietary details of the segmentation approach used, it is most likely a 2-D approach. Similarly, to the best of our knowledge, the reported approaches by others [17–24] for the segmentation of OCT images have been two-dimensional in nature (i.e., if multiple 2-D slices are available in a particular scanning sequence they are segmented independently). Thus, applying these approaches to the segmentation of 3-D images (e.g., by repeatedly applying the 2-D approach to each slice) would not take advantage of any 3-D contextual information. In fact, many of the reported approaches do not even take full advantage of 2-D contextual information as they first rely on finding intensity peaks from each preprocessed A-scan (column) separately [17–20, 22, 23]. While variations to each of the prior approaches exist for the segmentation of retinal boundaries, a typical approach proceeds as follows:

- Preprocess the image (e.g., with a median filter as in [17–20] or anisotropic diffusion filter as in [22]).

- Perform a 1-D peak detection algorithm on each A-scan (column) of the processed image to find points on each border of interest.
- (For only a few approaches) Process the points further to correct for possible discontinuities in the 1-D border detection approaches (e.g., use Markov modeling to connect smaller segments to the largest continuous segment followed by spline-fit as in [17, 18]).

Other prior approaches include the use of 2-D dynamic programming by Baroni *et al.* [24] and manually-initialized deformable models for the segmentation of fluid-filled regions by Cabrera Fernández [21].

The prior segmentation approaches by others have attempted to find different numbers of boundaries of the retina. In particular, Koozekanani *et al.* [17, 18] found 2, Baroni *et al.* [24] found 3, Shahidi *et al.* found 4 [23], Ishikawa *et al.* [19, 20] found 5, and Cabrera Fernández found 7 [22]. Because most authors reported limited quantitative validation, it is difficult to assess the robustness of the approaches. Nevertheless, based on the given results and personal communication with some of the authors, it is clear that better segmentation approaches are needed for greater accuracy.

2.6 Optimal 3-D segmentation

While it is often simpler to repeatedly perform a 2-D approach for the 3-D segmentation of surfaces, the lack of 3-D contextual information can produce segmentation errors [6]. However, many traditionally used 3-D surface segmentation approaches such as level-sets [25–29] and active-shape/active-appearance models (AAM/ASM) [30–33] involve converging to a local minimum, and thus do not guarantee finding a globally-optimal solution. Similarly, some other 3-D surface segmentation approaches must rely on heuristic solutions because of the computational demands of finding an optimal solution [3, 34]. Other approaches that do not explicitly involve an optimization problem, such as region-growing, are prone to their own set of problems such as leaking.

The recent work by Wu and Chen [7] and later extended to the multiple surface case by Li *et al.* [6] reflects a notable exception that allows for the optimal detection of surfaces in low-polynomial time. In these approaches, the surface (or multisurface) segmentation task is transformed into that of finding a minimum-cost closed set in a constructed vertex-weighted geometric graph. The edges of the graph are defined so that each closed set in the graph corresponds to a feasible surface (or set of feasible surfaces). Furthermore, the vertex costs are assigned so that the cost of each closed set directly corresponds to the cost of the set of surfaces. The closed set with the minimum cost (corresponding to the optimal set of surfaces) is then determined by finding a minimum s-t cut in a closely-related graph.

Note that even though the surfaces are ultimately found by finding a minimum-cost s-t cut in a constructed graph, these approaches [6,7] are fundamentally different than the "graph cut" methods of Boykov *et al.* (e.g., [35]), which represent another option for optimally performing segmentation tasks in 3-D. However, perhaps most importantly for the purposes of intraretinal layer segmentation, one advantage of using the optimal graph search approaches in [6,7] over that of the graph cut methods [35], is the ability to find multiple surfaces simultaneously. Other potential advantages include: 1) the ability to constrain the topology of the found surfaces, 2) the fact that the segmentation problem is directly formulated as finding surfaces instead of assigning "labels" to voxels, and 3) the clear separation of the surface feasibility constraints and the surface costs in the graph representation.

CHAPTER 3
MULTI-SURFACE IMAGE SEGMENTATION – EXTENSIONS OF OPTIMAL 3-D GRAPH-BASED SURFACE DETECTION

In this chapter, we review the original optimal 3-D graph search and present two extensions: the incorporation of regional information (presented at IPMI 2007 [36]) and the incorporation of varying constraints (presented at MICCAI 2007 [37]). Note that the appendix provides some additional background on relevant graph concepts.

3.1 Review of optimal 3-D graph search concepts

Recall that the optimal 3-D graph search approach is designed to solve what we will call the "multiple surface segmentation problem." An overview of the steps used in this approach are illustrated in Fig. 3.1. In very general terms, the multiple surface segmentation problem can be thought of as an optimization problem with the goal being to find the set of surfaces with the minimum cost such that the found surface set is feasible. Thus, there are two major components to the problem specification: 1) the specification of the constraints to require surface set feasibility and 2) the formalization of the cost of a set of surfaces. The first step in the graph search approach is to construct a graph such that the minimum-cost closed set of the graph corresponds to the set of surfaces with the minimum cost. (A closed set is a subset of the vertices of a graph such that no directed edges leave the set.) This is done by 1) ensuring that there is a one-to-one correspondence between each closed set in the constructed graph and each feasible surface set and 2) ensuring that the cost of each closed set in the graph corresponds (within a constant) to the cost of a set of (feasible) surfaces. Thus, it is the structure of the graph that reflects the feasibility constraints and vertex weights of the graph that reflect the cost functions. Finally, the actual minimum-cost closed set is found by finding minimum-cost s-t cut in a closely related graph [6, 7] (also see the appendix).

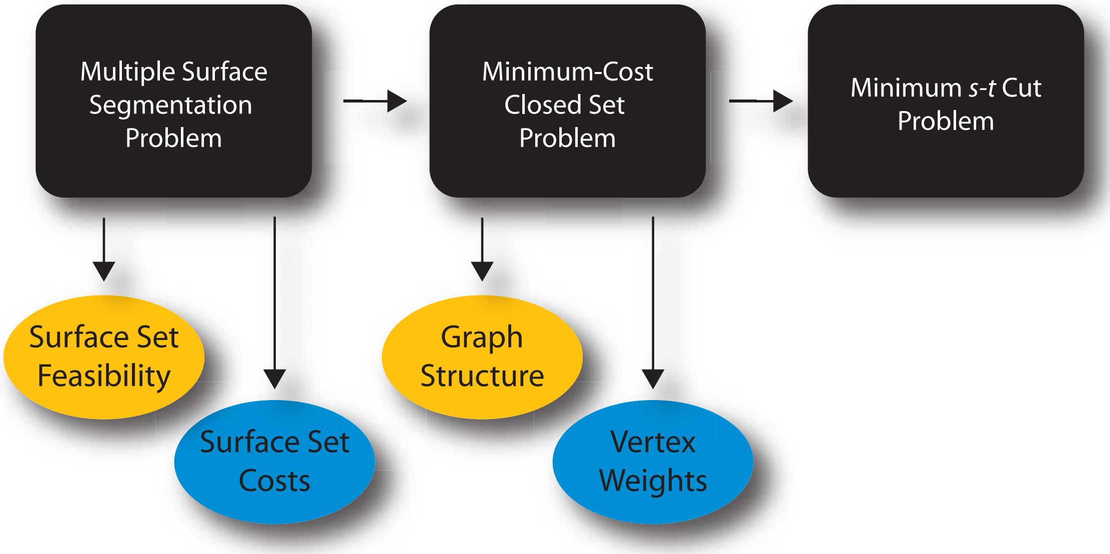


Figure 3.1: Review of optimal 3-D graph search approach. First, the multiple surface segmentation problem is transformed into the graph-theoretic problem of finding a minimum-cost closed set in a geometric graph. The graph is constructed so that the structure of the graph reflects the feasibility constraints and the vertex weights of the graph reflect the cost function. Finally, the minimum-cost closed set is found in the constructed graph by finding a minimum s-t cut in a closely related graph.

3.1.1 Limitations

However, in its original formulation, the 3-D optimal graph search method employed surface feasibility constraints that were constant in each direction. For example, the smoothness constraints for a particular surface $f(x, y)$ were represented by two parameters, Δ_x and Δ_y, reflecting the allowed change in surface height when moving from one neighboring surface point to the next in the x-direction and y-direction, respectively. Similarly, the surface interaction constraints (reflecting the allowed minimum and maximum distances between surface pairs) were constant. More flexibility in constraining surfaces to particular shapes would be obtained if varying constraints were allowed. Such a change would especially be important for surfaces in which the needed constraints are expected to change based on location (e.g., the foveal region in OCT images).

Furthermore, in its original formulation [6, 7], the cost of a set of surfaces was defined as a summation of cost values associated with voxels on the surfaces (i.e., the cost of a voxel with respect to a particular surface reflected the unlikeliness that the voxel would be part of the surface). While such "on-surface" costs can incorporate both image edge and regional information (e.g., see [38, 39] and variation 1 of Chapter 4), the incorporation of regional information is often limited to a region immediately surrounding the voxel for which the cost is defined (especially in cases of multiple surface detection). In some applications, better cost functions could likely be defined if "true" regional information could be incorporated. This involves extending the definition of the cost of a set of surfaces to also include the summation of in-region cost values in addition to the on-surface cost values. The in-region cost value for a voxel associated with a particular region would reflect the unlikeliness of that voxel belonging to the region. Thus, this chapter focuses on the development of an "extended version" of the optimal graph search approach that incorporates varying constraints and regional cost function terms.

3.2 Incorporating varying constraints

3.2.1 Surface set feasibility with varying constraints

Consider a volumetric image $I(x, y, z)$ of size $X \times Y \times Z$. We focus on the case in which each surface of interest can be defined with a function $f(x, y)$ mapping (x, y) pairs to z-values. Associated with each (x, y) pair is a column of voxels in which only one of the voxels — the voxel at $(x, y, f(x, y))$ — intersects the surface. Each column also has a set of neighbors. We use a “4-neighbor” relationship in which the set of neighbors for the column associated with (x, y) are the columns associated with $(x + 1, y)$, $(x - 1, y)$, $(x, y + 1)$, and $(x, y - 1)$. Other neighborhood relationships are also possible. One common example is to add a “circularity” neighbor relationship for images that are unwrapped from a cylindrical coordinate system. An example of this would be if the column associated with $(0, y_0)$ is considered to be a neighbor to the column associated with $(X - 1, y_0)$.

In prior work, a single surface is considered feasible if the difference in z-values of neighboring surface points is less than or equal to a constant parameter (Δ_x in x-direction, Δ_y in y-direction). For example, for neighboring columns $\{(x_1, y_1), (x_2, y_2)\}$ in the x-direction, this requires that

$$-\Delta_x \leq f(x_1, y_1) - f(x_2, y_2) \leq \Delta_x. \tag{3.1}$$

A similar constraint exists for neighbors in the y-direction.

Instead, we propose to allow the smoothness constraints to vary as a function of the column neighborhood pair. For a given neighborhood pair $\{(x_1, y_1), (x_2, y_2)\}$, the constraint becomes:

$$-\Delta^u_{\{(x_1,y_1),(x_2,y_2)\}} \leq f(x_1, y_1) - f(x_2, y_2) \leq \Delta^l_{\{(x_1,y_1),(x_2,y_2)\}}, \tag{3.2}$$

where $\Delta^u_{\{(x_1,y_1),(x_2,y_2)\}}$ reflects the maximum allowed increase in z-value when moving on a surface from column (x_1, y_1) to column (x_2, y_2) and $\Delta^l_{\{(x_1,y_1),(x_2,y_2)\}}$ reflects the maximum allowed decrease in z-value.

For a set of surfaces, additional constraints are added to model the desired relationships between the surfaces. For example, it may be known that one surface is always above another surface and that the distance between the surfaces is at least δ^l voxels, but no more than δ^u voxels (note the notational difference in δ used for surface interaction constraints and Δ used for smoothness constraints). Again, in prior work, these constraints were constant. Here, we allow these constraints to be a function of (x, y) so that in principle, a different interaction constraint can be used for each column.

3.2.2 Graph representation of surface set feasibility

The structure of the constructed graph reflects the feasibility constraints. Recall that this means that there must be a one-to-one correspondence between each feasible surface set and a closed set in the constructed graph. The graph is constructed in a similar manner as reported in [6] with the exception that the edges of the graph must take into account the varying constraints. First, one graph is created for each surface to be found. Then, intracolumn and intercolumn edges are added to enforce the surface smoothness constraints (Fig. 3.2(a-b)). Finally, the individual graphs are connected with intersurface edges to enforce the surface interaction constraints (Fig. 3.2(c-d)). Fig. 3.3 provides a 2-D toy example providing an intuitive illustration of how such a graph structure can enforce the feasibility constraints. Note that the "upper envelope" of the closed set corresponds to the surface.

More specifically, we will consider the added edges for one vertex (x_1, y_1, z_1) associated with a voxel towards the center of the image (i.e., a vertex not involved in boundary conditions). It will be associated with two intracolumn directed edges: one directed towards the vertex below it in the column and one from the vertex above it. Two intercolumn edges will also exist for each neighboring column (x_2, y_2):

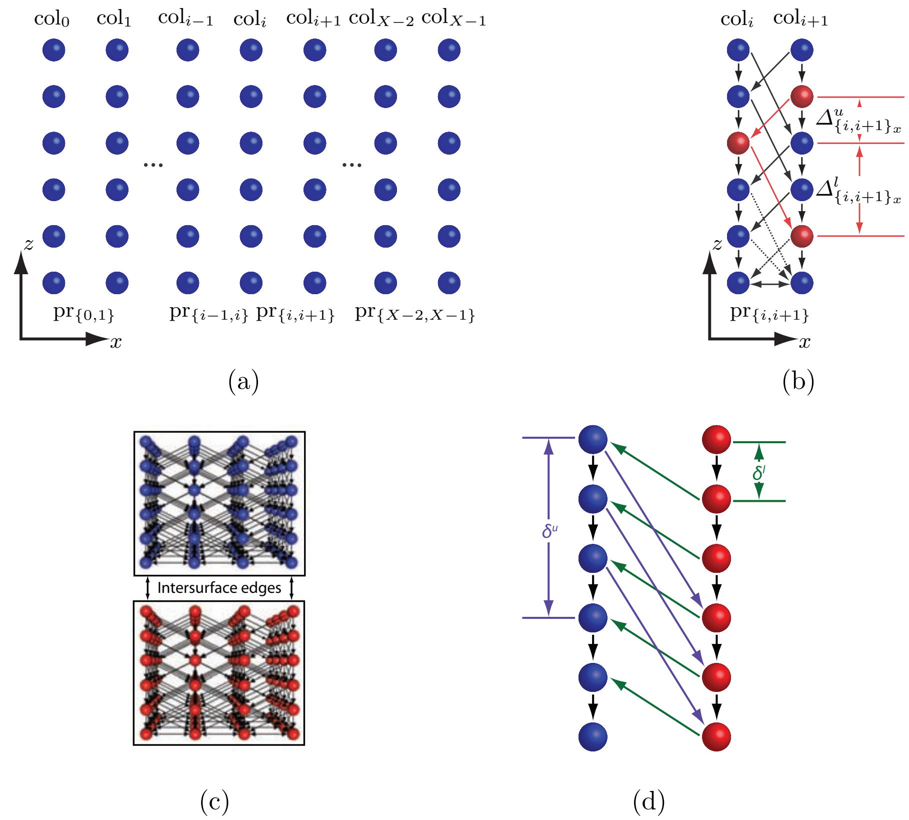
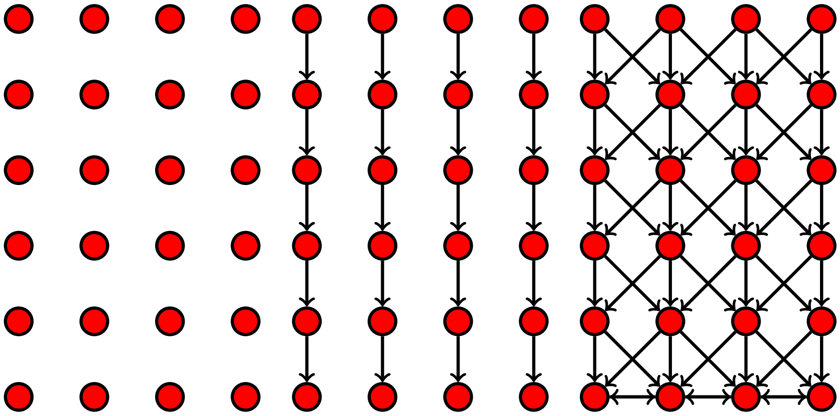
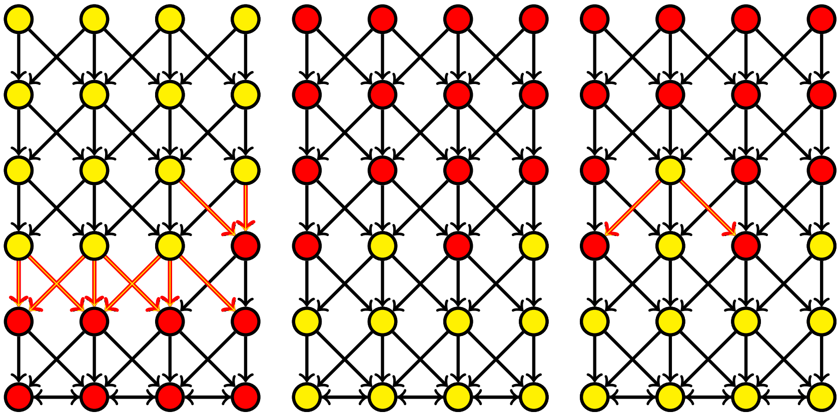


Figure 3.2: Graph representation of feasibility constraints. (a–b) Surface smoothness constraints (shown only in x-direction using simplified notation). (b–c) Intersurface edges are added between the surface subgraphs to enforce the surface interaction constraints.



(a) Vertices (b) Intracolumn edges (c) Intercolumn edges



(d) Not a closed set (e) A closed set (f) Not a closed set

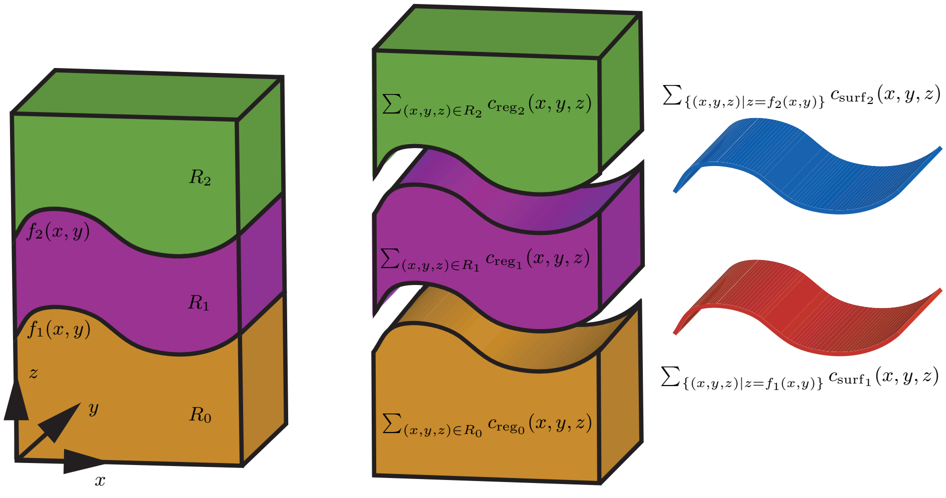
Figure 3.3: 2-D closed set toy example. (a–c) Creation of graph to enforce constant surface smoothness constraints ($\Delta_x = 1$). (d) Highlighted yellow vertices do not form a closed set because a number of edges (in red and yellow) leave the set, providing an intuitive illustration of why the intracolumn edges are important in enforcing surface feasibility. (e) Highlighted yellow vertices form a closed set because all of the directed edges stay within the set. (f) Highlighted yellow edges do not form a closed set because two edges (indicated in red and yellow) leave the set, providing an intuitive illustration of why the intercolumn edges are important in enforcing the smoothness constraints.

one directed to the vertex in the neighboring column that has a z-value that is $\Delta^l_{\{(x_1,y_1),(x_2,y_2)\}}$ smaller and one from the vertex in the neighboring column that has a z-value that is $\Delta^u_{\{(x_1,y_1),(x_2,y_2)\}}$ greater (Fig. 3.2(b)). Finally, for each corresponding column in the volume associated with a surface interaction constraint, two intersurface edges are associated with the vertex: one to the vertex in the corresponding column with a z-value that is $\delta^u(x_1, y_1)$ smaller and one from the vertex in the corresponding column with a z-value that is $\delta^l(x_1, y_1)$ smaller (assuming the surface for the given vertex is supposed to be "above" the interacting surface – see Fig. 3.2(c-d) with the blue surface graph being above the red surface graph; if not, the edge directions are reversed). Slightly different edges must be used in the "boundary cases" in which any of those vertices do not exist [6].

3.3 Incorporating regional information

3.3.1 Cost of a feasible surface set

Given a set of n non-intersecting surfaces $\{f_1(x,y),\ f_2(x,y),\ \ldots,\ f_n(x,y)\}$, the surfaces naturally divide the volume into $n+1$ regions (Fig. 3.4). Assuming the surfaces are labeled in "increasing" order, the regions can be labeled $R_0, \ldots, R_n$, where R_i reflects the region that lies between surface i and surface $i+1$ (with region boundary cases R_0 and R_n being defined as the region with lower z-values than surface 1 and the region with higher z-values than surface n, respectively). Each voxel can thus have $2n+1$ real-valued costs associated with it: n on-surface costs corresponding to the unlikeliness of belonging to each surface and $n+1$ in-region costs associated with the unlikeliness of belonging to each region. Let $c_{\mathrm{surf}_i}(x,y,z)$ represent the on-surface cost function associated with surface i and $c_{\mathrm{reg}_i}(x,y,z)$ represent the in-region cost function associated with region i. Then, the cost $C_{\{f_1(x,y),f_2(x,y),\ldots,f_n(x,y)\}}$



(a) (b)

Figure 3.4: Example schematic cost of two surfaces for the multiple surface segmentation problem. The two surfaces divide the volume into three regions.

associated with the set of surfaces can be defined as

$$C_{\{f_1(x,y),f_2(x,y),\ldots,f_n(x,y)\}} = \sum_{i=1}^{n} C_{f_i(x,y)} + \sum_{i=0}^{n} C_{R_i} \ , \tag{3.3}$$

where

$$C_{f_i(x,y)} = \sum_{\{(x,y,z)|z=f_i(x,y)\}} c_{\mathrm{surf}_i}(x,y,z) \ , \tag{3.4}$$

and

$$C_{R_i} = \sum_{(x,y,z)\in R_i} c_{\mathrm{reg}_i}(x,y,z) \ . \tag{3.5}$$

Note that $C_{f_i(x,y)}$ reflects the cost associated with voxels on surface i and C_{R_i} reflects the cost associated with voxels belonging to region i.

3.3.2 Graph representation of surface set costs

The cost of each vertex in the graph is set such that the cost of each closed set corresponds to the cost (within a constant) of the set of surfaces. (The cost of a closed set is the summation of the costs of all the vertices.) The weight $w_i(x,y,z)$

of each vertex ($i = 1, 2, \ldots, n$) can be defined as the summation of a term related to the on-surface costs ($w_{\text{on}-\text{surf}_i}(x, y, z)$) and a term related to the in-region costs ($w_{\text{in}-\text{reg}_i}(x, y, z)$):

$$w_i(x, y, z) = w_{\text{on}-\text{surf}_i}(x, y, z) + w_{\text{in}-\text{reg}_i}(x, y, z)\ . \tag{3.6}$$

3.3.2.1 On-surface costs

For on-surface costs, the cost of each vertex is assigned the on-surface cost of the corresponding voxel minus the on-surface cost of the voxel below it [6, 7]:

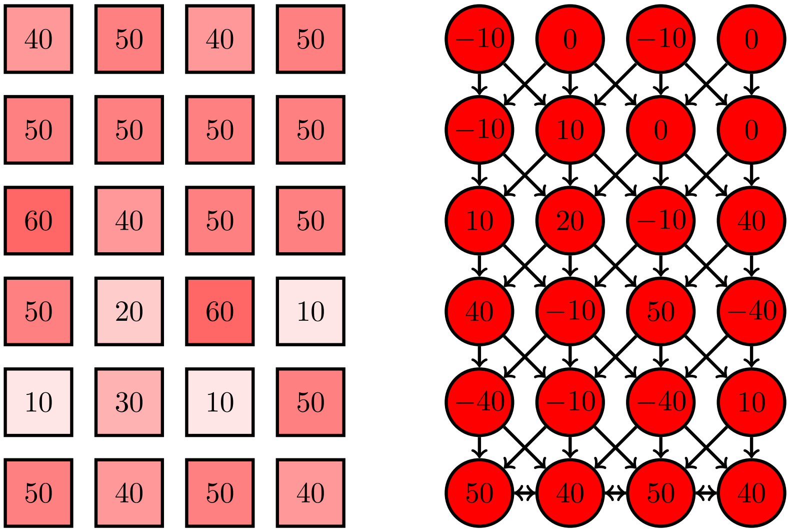
$$w_{\text{on}-\text{surf}_i}(x, y, z) = \begin{cases} c_{\text{surf}_i}(x, y, z) & \text{if} \quad z = 0 \\ c_{\text{surf}_i}(x, y, z) - c_{\text{surf}_i}(x, y, z - 1) & \text{otherwise} \end{cases}\ . \tag{3.7}$$

This is illustrated with a 2-D toy example in Fig. 3.5. In addition, to ensure that a nonempty closed set is ultimately determined, the cost of one of the vertices of the graph is modified so that the summation of all of the "base" vertices is negative, as illustrated in Fig. 3.6.

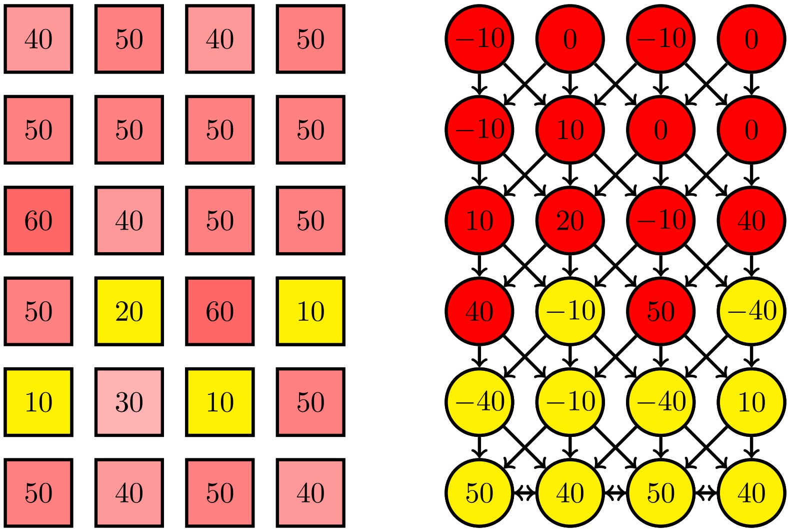
As a contrast, Fig. 3.7 illustrates the same toy example using a dynamic programming approach (finding a minimum-cost path in a graph as described in the appendix). As dynamic programming is a popular 2-D graph-based segmentation approach, this example is provided to make some of the similarities and differences more clear. However, it is important to remember that even though both approaches work for finding optimal boundaries in 2-D, the minimum-cost path approach does not scale well to multiple dimensions.

3.3.2.2 In-region costs

For in-region costs, the cost of each vertex is assigned the in-region cost of the region below the surface associated with the vertex minus the in-region cost of the

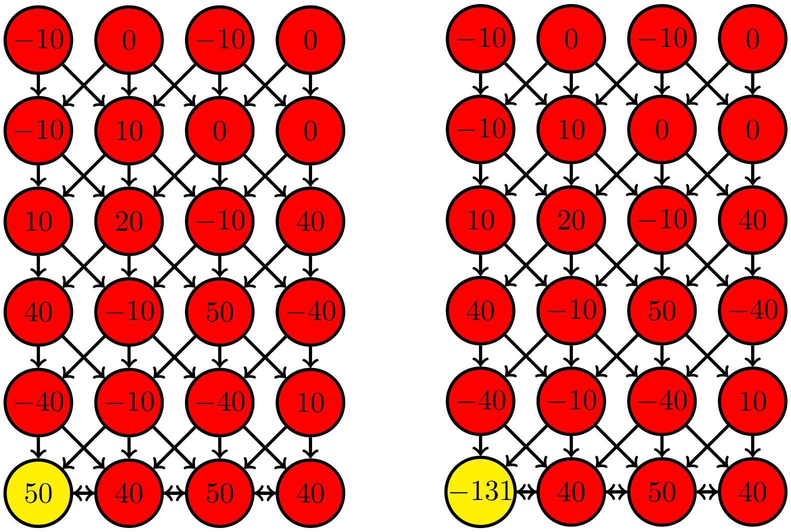


(a) Cost image (b) Graph



(c) Surface cost = 50 (d) Closed set cost = 50

Figure 3.5: Correspondence of image on-surface costs with closed set costs. (a) Example on-surface cost function. (b) Graph representation of on-surface costs. (c) Example surface highlighted in yellow (surface cost = 50). (d) Corresponding closed set highlighted in yellow (closed set cost = 50).



(a) Before (b) After

Figure 3.6: Additional cost transformation to ensure that a nonempty closed set will be found. The cost of one vertex in the base graph (bottom row of vertices in this figure) is modified so that the summation of all of the vertices in the base graph is negative.

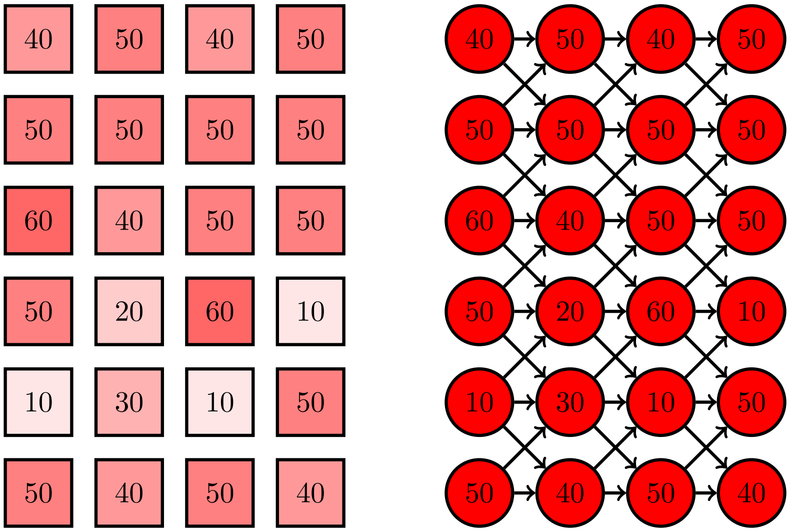
region above the surface associated with the vertex:

$$w_{\text{in-reg}_i}(x,y,z) = c_{\text{reg}_{i-1}}(x,y,z) - c_{\text{reg}_i}(x,y,z)\ . \tag{3.8}$$

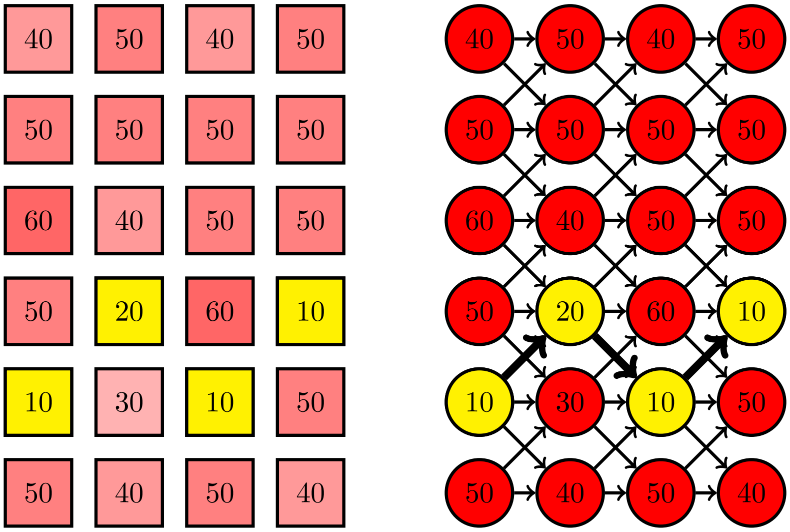
Because the use of in-region costs is new and perhaps less intuitive, Fig. 3.8 illustrates why such a transformation works and Figs. 3.9 and 3.10 provide a toy example in 2-D. The cost of the closed set $C(V_{CS_i})$ associated with surface i using the in-region costs becomes

$$C(V_{CS_i}) = \sum_{(x,y,z)\in R_0\bigcup\cdots\bigcup R_{i-1}} c_{\text{reg}_{i-1}}(x,y,z) - \sum_{(x,y,z)\in R_0\bigcup\cdots\bigcup R_{i-1}} c_{\text{reg}_i}(x,y,z)\ . \tag{3.9}$$

Recognizing that many of costs associated with each individual region cancel when added together and the fact that $\sum_{(x,y,z)\in R_0\bigcup\cdots\bigcup R_n} c_{\text{reg}_n}(x,y,z)$ is a constant K (e.g., 440 for the toy example illustrated in Figs. 3.9 and 3.10), the cost for the closed set



(a) Cost image (b) Graph



(c) Surface cost = 50 (d) Path cost = 50

Figure 3.7: Same toy example using a dynamic programming approach (finding a minimum-cost path). Note that while both a path-based and closed-set-based approach can be used to find an optimal surface in 2-D (as in this toy example), the path-based approach does not scale to multiple dimensions, but the closed-set approach does.

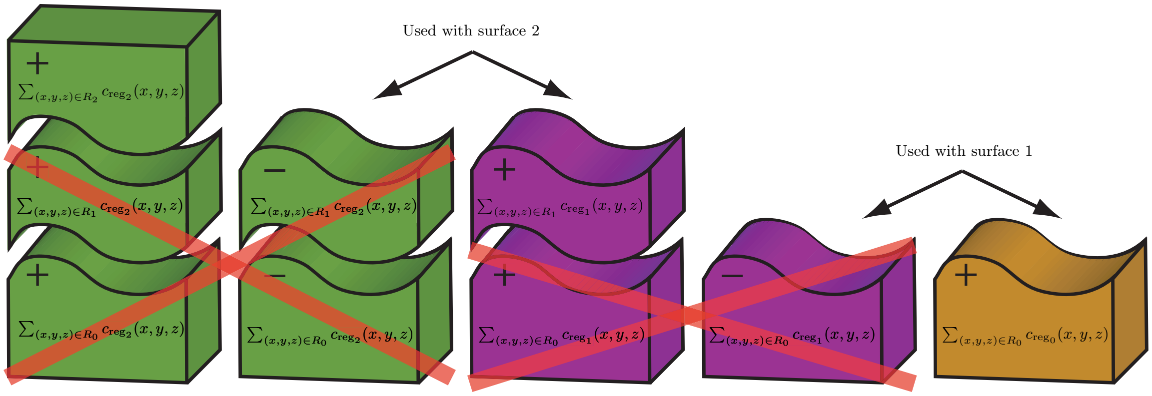


Figure 3.8: Schematic showing how the assignment of in-region costs to vertices produces the desired overall cost.

associated with the entire set of surfaces $C(V_{CS})$ reduces to

$$C(V_{CS}) = -K + \sum_{i=0}^{n} \sum_{(x,y,z)\in R_i} c_{\mathrm{reg}_i}(x,y,z) \ , \tag{3.10}$$

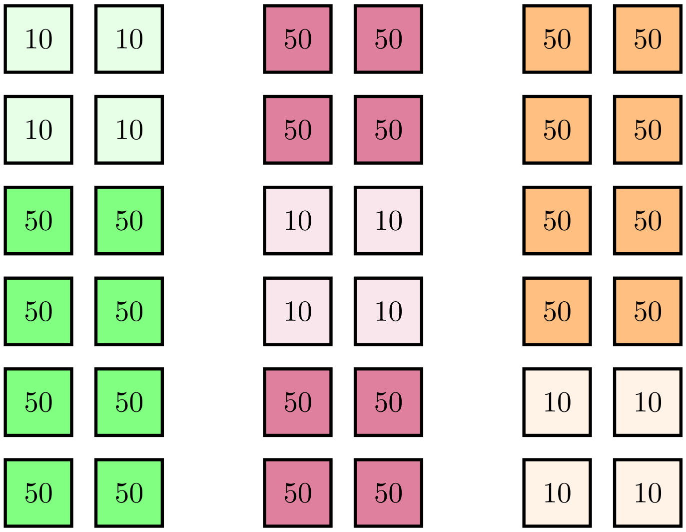
which, within a constant, is equivalent to the desired in-region component of the cost of the set of surfaces.

3.4 Incorporating prior knowledge into the graph search

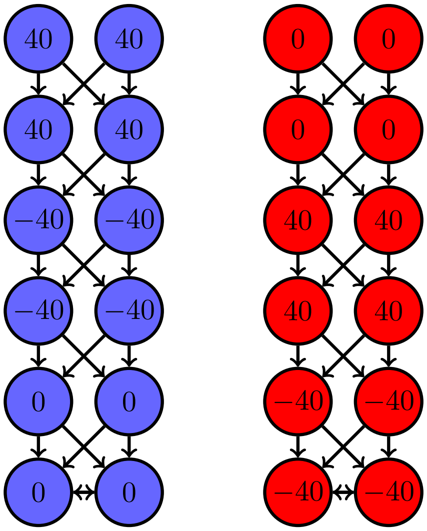
The clear separation between surface feasibility and surface set costs in the optimal 3-D graph search allows us to focus on each part independently. As an example, we conclude this chapter by presenting how surface feasibility constraints can be learned from examples (set of manually-traced boundaries) to incorporate prior shape information. Later chapters will address using prior knowledge in the cost-function design.

3.4.1 Learning varying constraints from examples

For purposes of learning the varying smoothness constraints for each surface, it may be easiest to think of each pair of neighboring columns $\{(x_1, y_1), (x_2, y_2)\}$ as

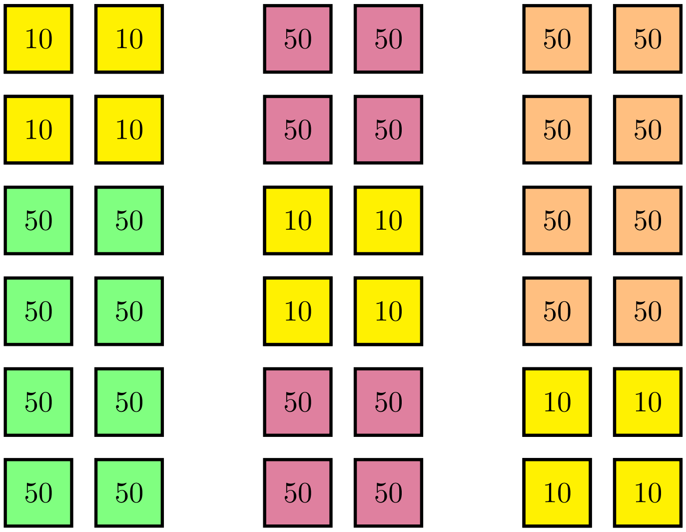


(a) Region 2 (b) Region 1 (c) Region 0



(d) Graph 2 (e) Graph 1

Figure 3.9: 2-D toy example illustrating graph representation of in-region costs. (a) Cost function for region 2. (b) Cost function for region 1. (c) Cost function for region 0. (d) Graph corresponding to surface 2. (d) Graph corresponding to surface 1. Note that the intersurface edges (that would exist between graph 1 and graph 2) are not shown.



10	10
10	10
50	50
50	50
50	50
50	50

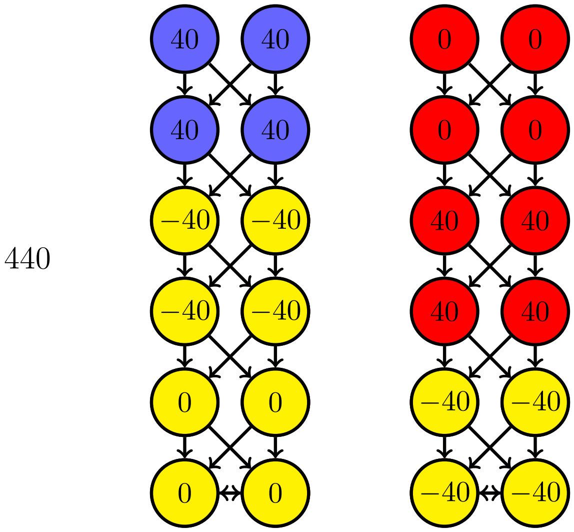
(a) Region 2

50	50
50	50
10	10
10	10
50	50
50	50

(b) Region 1

50	50
50	50
50	50
50	50
10	10
10	10

(c) Region 0



(d) Constant (e) Graph 2 (f) Graph 1

Figure 3.10: 2-D toy example illustrating how the cost of a closed set in the graph corresponds to the desired in-region cost terms. (a) Region 2 cost function with yellow highlighted vertices indicating vertices selected to be part of region 2 (for purposes of this example). (b) Region 1 cost function with yellow highlighted vertices indicating vertices selected to be part of region 1. (c) Region 0 cost function with yellow highlighted vertices indicating vertices selected to be part of region 0. (d) Summation of all of the vertices in the region 2 cost function (a constant). (e) Graph for surface 2 with vertices that are part of the closed set highlighted in yellow (summation of cost values = -160). (f) Graph for surface 1 with vertices that are part of the closed set highlighted in yellow (summation of cost values = -160). Note that $C(V_{CS}) = -320 = -440 + 120$.

having its own constraint that needs to be learned separately. The basic idea is use manual tracings in the training set to determine the mean and standard deviation of how the z-value changes when moving from column (x_1, y_1) to column (x_2, y_2). Let $\overline{d}$ reflect the mean deviation (i.e., the mean of $f(x_1, y_1) - f(x_2, y_2)$ from the reference standard) and σ reflect the standard deviation. Then, to allow for 99% of the expected changes in z-value when moving from column (x_1, y_1) to column (x_2, y_2) (assuming a normal distribution), the two parameters of the smoothness constraint, $\Delta^l_{\{(x_1,y_1),(x_2,y_2)\}}$ and $\Delta^u_{\{(x_1,y_1),(x_2,y_2)\}}$ can be set as follows:

$$\Delta^l_{\{(x_1,y_1),(x_2,y_2)\}} = \overline{d} + 2.6 * \sigma \ , \tag{3.11}$$

and

$$\Delta^u_{\{(x_1,y_1),(x_2,y_2)\}} = -(\overline{d} - 2.6 * \sigma) \ . \tag{3.12}$$

However, depending on the prior knowledge of the problem, this general case may not be necessary. For example, it may be sufficient to only learn the constraints in one direction (e.g., for column pairs in the x-direction) and use a constant constraint for columns in the other direction. Alternatively, the smoothness constraints can independently be learned in each direction. In essence, the concept is the same as in the general case, but fewer constraints are required to be learned, thus requiring less training data.

The surface interaction constraints can be learned in a similar fashion. For each surface pairing, the minimum allowed distance between the surfaces at each column is set as the mean thickness minus 2.6 times the standard deviation (but truncated not so as not to go below 0). The maximum allowed distance between the surfaces at each column is set as the mean thickness plus 2.6 times the standard deviation.

CHAPTER 4
TIME-DOMAIN OCT SEGMENTATION

In this chapter, a summary the completed work related to segmenting time-domain OCT images in 3-D is presented [36–40].

4.1 Initial work with peripapillary scans

Our initial work in this area, published in MICCAI 2006 [40], involved the 3-D segmentation of the upper and lower surfaces of the retina on peripapillary OCT scans (Fig. 4.1) using the optimal 3-D graph search (with only on-surface costs and constant constraints). After preprocessing the images using a speckle-reducing anisotropic diffusion method [41], the upper surface (internal limiting membrane) was found first, followed by the lower surface (retinal pigment epithelium). A different cost function was used for each surface, but both involved combining an edge and localized regional term, as discussed in more detail below:

- Edge term: A 3x3 Sobel kernel was convolved with each preprocessed image, resulting in an estimate the strength of the edge in the z-direction for each voxel (giving positive values for light-to-dark transitions and negative values for dark-to-light transitions). Because we expected a dark-to-light transition to occur at both surfaces, we maintained the signed edge values in the cost function, thus favoring dark-to-light transitions and discouraging light-to-dark transitions.
- Localized regional term: The regional cost terms were added to help aid in the detection of the correct surface even when edge information was lacking. For the internal limiting membrane, we used a normalized cumulative image as our regional cost term. This was based on the observation that very few bright voxels exist above (have a smaller z-value than) the internal limiting membrane in each column. Incorporating a cumulative image into the cost also had the effect of discouraging the selection of voxels in the pigment epithelial surface, a necessity since we were finding the internal limiting membrane first.

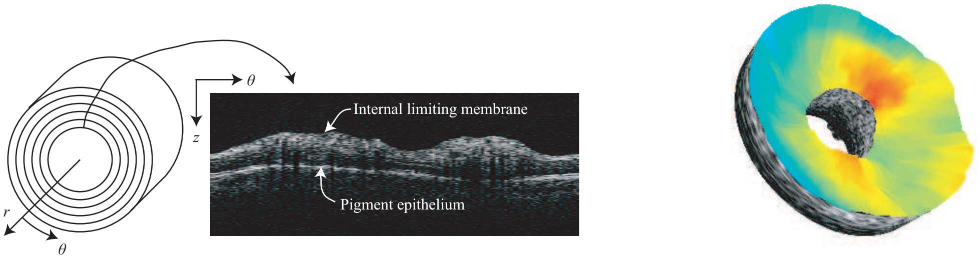
For the pigment epithelial surface, we incorporated a localized region-based term into the cost function that attempted to minimize the regional variances above and below the surface. As an example, suppose there are only two relatively-homogeneous regions in a 3-D image separated by a surface S with average intensities a_1 (from voxel locations (x, y, z) inside S) and a_2 (from voxel locations (x, y, z) outside S). Then a regional cost function term as presented by Chan and Vese [42] can be defined as

$$\begin{aligned} F(S, a_1, a_2) = &\int_{inside(S)} |I(x, y, z) - a_1)|^2 \, dx \, dy \, dz \\ &+ \int_{outside(S)} |I(x, y, z) - a_2|^2 \, dx \, dy \, dz. \end{aligned} \tag{4.1}$$

In our work, we maintained the idea of minimizing a variance-like term on either side of the surface, but we limited the computation to a region of limited size. Furthermore, because our cost function was voxel-based and not surface-based (the cost of a surface in this work was computed from a summation of voxel-based costs and thus could not include terms that depended on the overall surface properties), we added the variance-like term of the voxels above and below each voxel for which we were computing the cost:

$$\widehat{v}(r, \theta, z) = \sum_{k=z-\gamma}^{z-1} (I(r, \theta, k) - \bar{I}_{-})^2 + \sum_{k=z+1}^{z+\gamma} (I(r, \theta, k) - \bar{I}_{+})^2, \tag{4.2}$$

where γ was a constant defining how far above and below the voxel we computed the variance-like term, $\bar{I}_{-}$ was the average intensity of voxels with smaller z-values than the voxel at position (r, θ, z), and $\bar{I}_{+}$ was the average intensity of voxels with larger z-values than the voxel at position (r, θ, z). However, since Eq. (4.2) would also favor voxels that were not of interest (e.g., background



(a) (b)

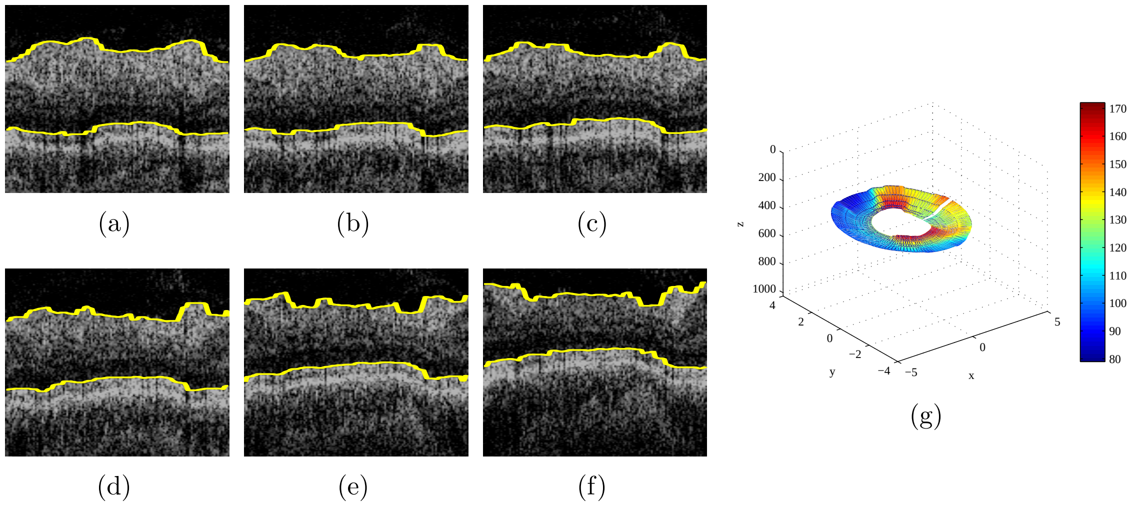
Figure 4.1: Circular scans used in MICCAI 2006 work [40]. (a) Each 3-D OCT image was composed of six 2-D circular scans surrounding the optic disc, as schematically shown on the left. One example circular scan is shown on the right. The internal limiting membrane is the "upper" surface of the retinal layer (smaller z-values) and the pigment epithelium is the "lower" surface of the retinal layer (larger z-values). (b) Example 3-D rendering of segmented image, with colors indicating total retinal thickness.

voxels), the actual regional cost term $v(r, \theta, z)$ would use the cost defined by Eq. (4.2) only if $\bar{I}_+$ were greater than $\bar{I}_-$:

$$v(r,\theta,z) = \begin{cases} \widehat{v}(r,\theta,z) & \text{if} \quad \bar{I}_+ + \varepsilon > \bar{I}_- \\ c & \text{otherwise,} \end{cases} \tag{4.3}$$

where c was a relatively large constant, thereby giving a large cost to potential surface voxels that do not separate a high intensity region "below" (larger z-values) from a low intensity region "above" (smaller z-values) the voxel.

The method was tested on 18 3-D data sets obtained using a Stratus OCT-3 scanner: 9 from patients with normal optic discs and 9 from patients with papilledema (optic nerve swelling). The dimensions of each data set were $6 \times 128 \times 1024$ ($R = 6$, $T = 128$, $Z = 1024$). The axial scan length was 2 mm, resulting in a voxel size of approximately 2 μm in the z-direction. The size of the voxels in the r and θ direction varied due to the radial nature of the scans. Expert analysis of all the data sets indicated that our method found the correct surfaces. Fig. 4.2 displays an example



(a) (b) (c) (d) (e) (f) (g)

Figure 4.2: Typical results for one 3-D data set [40]. The results are shown in increasing circular scan order such that the results for the inner circular scan are shown in (a), while the results for the outer circular scan are shown in (f). Note that for easier visualization of the borders, the top and bottom of the images have been cropped. (h) 3-D view of the internal limiting membrane with colors reflecting retinal thickness (in pixels).

of a typical segmentation result.

Furthermore, we achieved a decrease in analysis failures from the vitreo-retinal interface segmentation available using the Stratus OCT-3 method (4/108 minor failures using our method versus 19/108 minor and 2/108 complete failures using the Stratus OCT-3 method – see Table 4.1). A circular scan was considered to have a “minor failure” if the visible deviation of the result from the correct surface was small enough not to require correction in a clinical setting. A scan was considered to have a “complete failure” if the result had a large deviation that would require correction in order for the segmentation to be clinically useful. The complete failures of the Stratus system are shown in Fig. 4.3, with our results also shown for comparison. Admittedly, the lack of a reference standard made our validation only preliminary at this stage, but at minimum, it served as a good “proof-of-concept” study to demonstrate that the optimal 3-D graph search can work well for the segmentation of OCT images.

Table 4.1: Count of circular scan failures in our method and the Stratus OCT-3 method for the internal limiting membrane

Scan group	Method	Minor failures	Complete failures	Total failures
Normal scans	3-D Graph Search	0 (0.0%)	0 (0.0%)	0 (0.0%)
Normal scans	Stratus OCT-3	4 (7.4%)	0 (0.0%)	4 (7.4%)
Papilledema scans	3-D Graph Search	4 (7.4%)	0 (0.0%)	4 (7.4%)
Papilledema scans	Stratus OCT-3	15 (27.8%)	2 (3.7%)	17 (31.5%)
All scans	3-D Graph Search	4 (3.7%)	0 (0.0%)	4 (3.7%)
All scans	Stratus OCT-3	19 (17.6%)	2 (1.8%)	21 (19.4%)

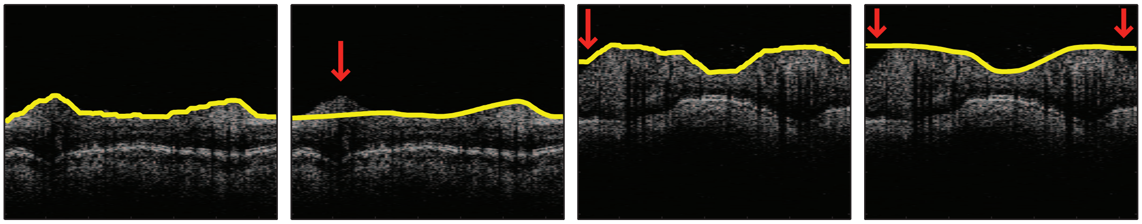

(a) Our method. (b) Stratus method. (c) Our method. (d) Stratus method.

Figure 4.3: Examples of "complete failures" (arrows) for the Stratus OCT-3 method and our corresponding results for the internal limiting membrane published in MICCAI 2006 [40]. (a)–(b) Our method versus the Stratus OCT-3 method for an inner circular scan on a patient with papilledema, (b) is considered a "complete failure." (c)–(d) Our method versus the Stratus OCT-3 method for an inner circular scan on a different patient with papilledema, (c) is considered a "minor failure" and (d) is considered a "complete failure."

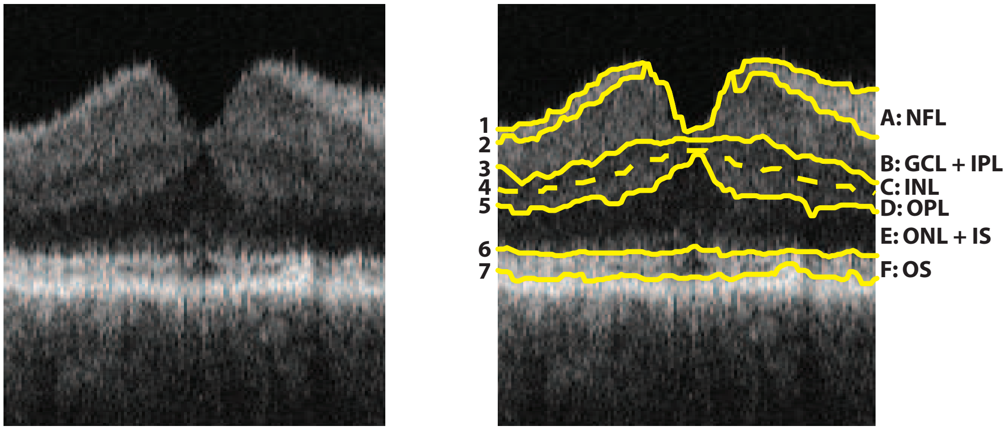
Furthermore, it was the first work to our knowledge that presented a *3-D* approach for the segmentation of such images.

4.2 Macular segmentation overview

After our initial work with peripapillary OCT segmentation, most of our effort related to time-domain OCT segmentation was focused on the intraretinal layer segmentation of macular scans [36–39] using the graph search approach presented in Chapter 3. Recall (see Fig. 2.5 (a–c) and Fig. 2.6 in Chapter 2) that one macular OCT image series (using the fast macular Stratus OCT-3 protocol) consists of six radial linear cross-sectional scans centered at the fovea. In this work, repeated series were acquired (six, if possible) for each eye, so that up to six raw scans existed at each angular location. The overall goal of the segmentation method was to determine surfaces defining the retinal layers on a composite 3-D image derived from the repeated raw scans.

Fig. 4.4 illustrates the surfaces and corresponding retinal layers we desired to find on each 3-D composite image (surface 4 is indicated with a hashed line because it was not found in all variations of our approach). As mentioned in Chapter 2, based on histology and higher-resolution OCT images published in the literature [16], we assumed the surfaces roughly had the following anatomical correspondence: surface 1 corresponded to the vitreo-retinal interface (VRI), surface 2 corresponded to the separation of the retinal nerve fiber layer (NFL) above from the ganglion cell layer (GCL) below, surface 3 corresponded to the separation of the inner plexiform layer (IPL) above from the inner nuclear layer (INL) below, surface 4 corresponded to separation of INL above from the outer plexiform layer (OPL) below, surface 5 corresponded to the separation of the OPL above from the outer nuclear layer (ONL) below, surface 6 corresponded to the junction between the photoreceptor inner and outer segments (IS/OS), and surface 7 corresponded to the separation of the photoreceptor outer segment (OS) from the retinal pigment epithelium (RPE). The corresponding six layers (labeled A though F in Figure 4.4(b)) may well be associated with the following anatomical layers: A) NFL, B) GCL + IPL, C) INL, D) OPL, E) ONL + IS, and F) OS.

There were two stages to the overall approach: I) the creation of a composite 3-D macular image from the raw scans and II) the determination of the surfaces on the 3-D composite image. An overview of the data flow in the segmentation process can be found in Fig. 4.5. While Stage I remained the same in all of our work for the segmentation of macular time-domain images, Stage II was performed with three



(a) (b)

Figure 4.4: Example composite image with labeled intralayer segmentation (top and bottom of images have been cropped to aid in visualization). (a) Composite image. (b) Seven surfaces (labeled 1-7) and six corresponding intralayers (labeled A through F). The anatomical correspondence is our current presumption based on histology and example images from higher-resolution research OCT scanners [16]: A) NFL (nerve fiber layer), B) GCL + IPL (ganglion cell layer and inner plexiform layer), C) INL (inner nuclear layer), D) OPL (outer plexiform layer), D) ONL + IS (outer nuclear layer and photoreceptor inner segments), E) OS (photoreceptor outer segments).

different variations:

1. The segmentation of six total surfaces (surfaces 1, 2, 3, 5, 6, and 7) using only on-surface costs and constant smoothness and interaction constraints [38, 39].
2. The segmentation of seven total surfaces using in-region costs for the interior surfaces (surfaces 2, 3, 4, and 5) and non-varying smoothness and surface interaction constraints [36].
3. The segmentation of seven total surfaces using in-region costs for the interior surfaces and using varying smoothness and surface interaction constraints [37].

In the next few sections, the methods for Stage I and each of the variations for Stage II are described in more detail.

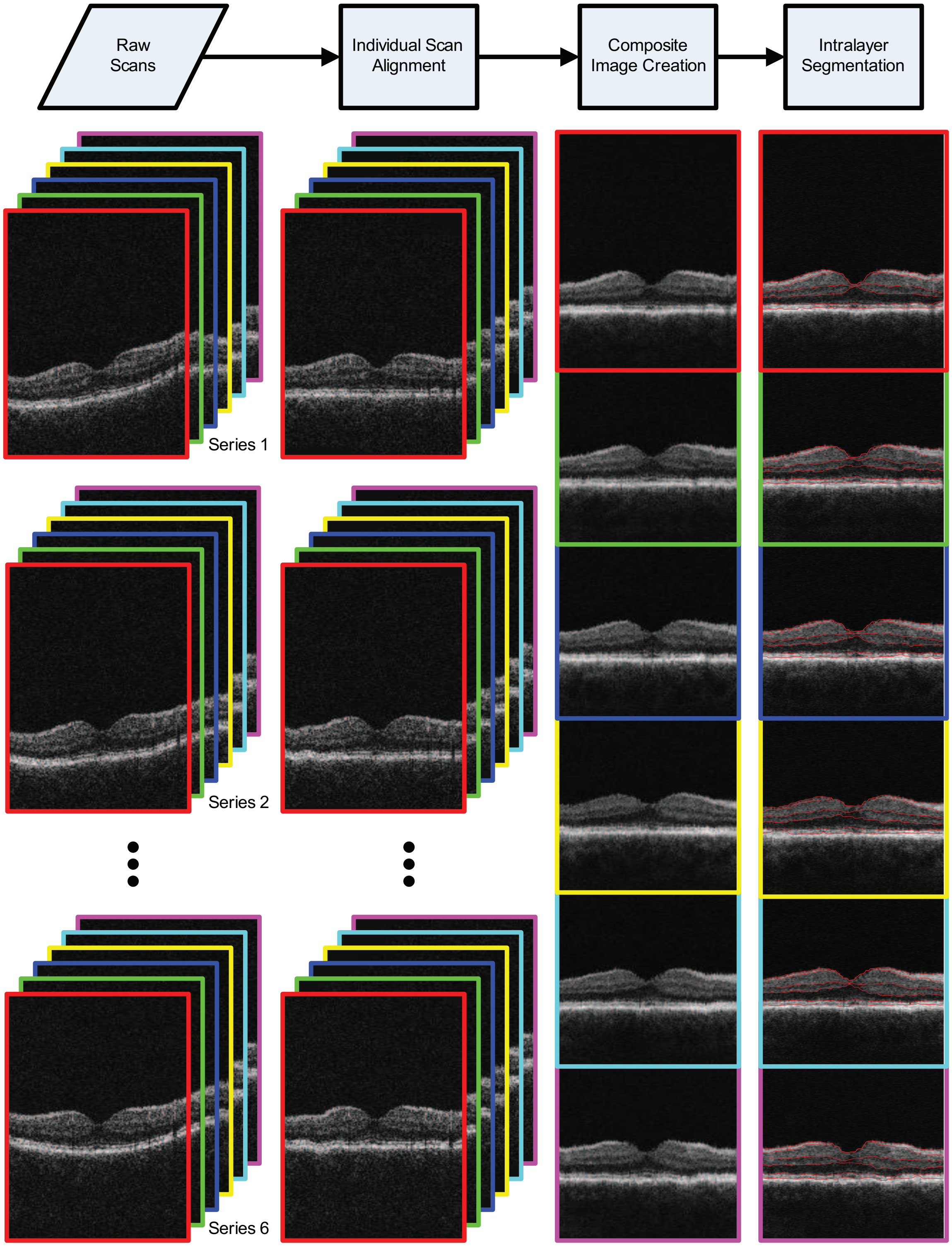


Figure 4.5: Overview of segmentation steps for the data associated with one eye. First, each individual scan was aligned so that the RPE (boundary 7) was approximately horizontal in the image. Second, images from each location were registered and averaged to form a composite image. Finally, the intralayer boundaries were determined using a 3-D graph-search approach. All steps were performed automatically.

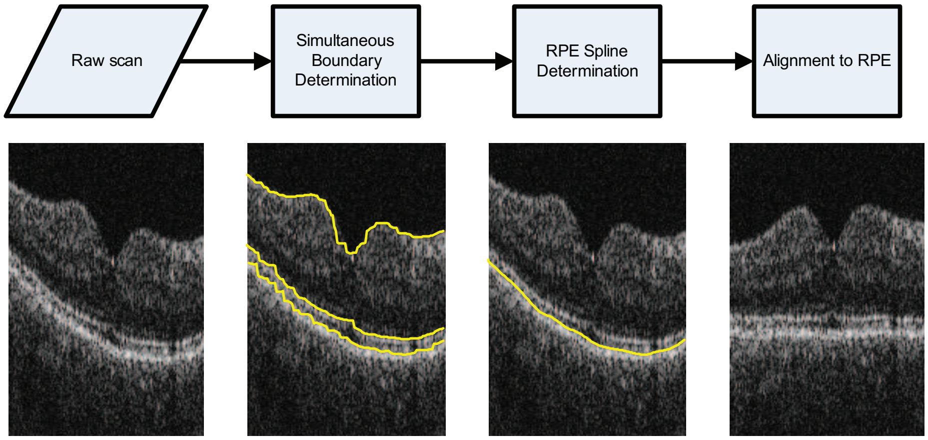
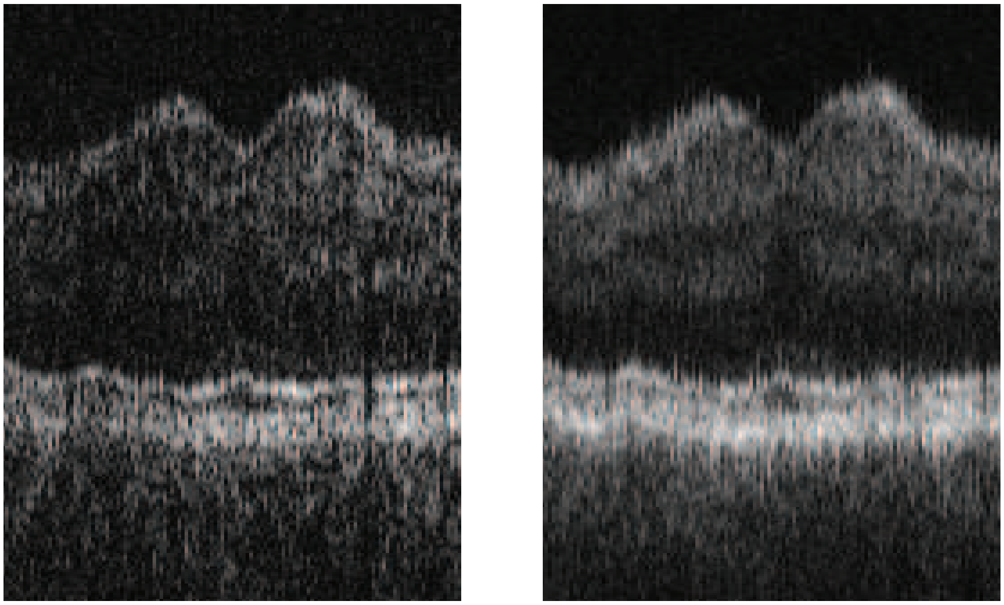


Figure 4.6: Individual scan alignment (top and bottom of images have been cropped to aid in visualization).

4.3 Stage I: Alignment and registration of macular scans

The 3-D composite image associated with each eye was created in two major steps. In the first step (Fig. 4.6), raw scans for a particular angular location (e.g., all the vertical scans) were individually aligned so that that the boundary 7 (the retinal pigment epithelium) appeared approximately straight in the aligned image. The purpose of the alignment was twofold: to aid in the final 3-D segmentation and to allow for better visualization. Each scan was aligned by first finding boundaries 1, 6, and 7 simultaneously using an optimal graph search approach similar to that used during stage II (described in more detail in later sections), but performed in 2-D. To ensure smoothness, a least-squares spline was fit to boundary 7. The columns were then translated so that this spline would be a straight line in the aligned image.

In the second step of this stage, each aligned image was registered to the first image in its location set by exhaustively searching for the best whole-pixel translation (according to the mutual information registration metric [43]) to align each of its columns to the corresponding target image column. The position of boundary 7 determined during the first step was used as a guide to determine the range of



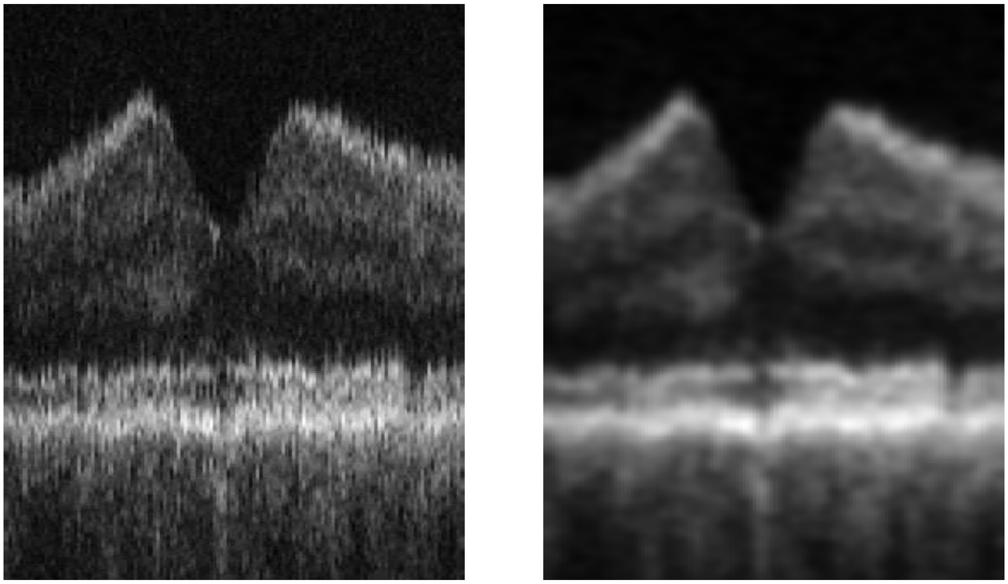
(a) (b)

Figure 4.7: Comparison between an individual scan and a 2-D composite scan (top and bottom of images have been cropped to aid in visualization). (a) Individual scan. (b) Composite scan.

translations to be tested for each column. The registered images in each location set were averaged together to form the composite image for that particular angular location. The purpose of averaging the images was to obtain a representative scan of that location that had a higher signal-to-noise ratio than any of the raw scans. An example of an individual scan and the corresponding 2-D composite scan is shown in Fig. 4.7. The set of 2-D composite images (one for each angular location) formed the 3-D composite image used in the next stage.

4.4 Stage II: Segmentation of each 3-D composite image

In the second stage, the surfaces were found on the 3-D composite image. As a preprocessing step, a speckle-reducing anisotropic diffusion method [41] was applied (Fig. 4.8). Surfaces 1, 6, and 7 were then simultaneously found using the optimal

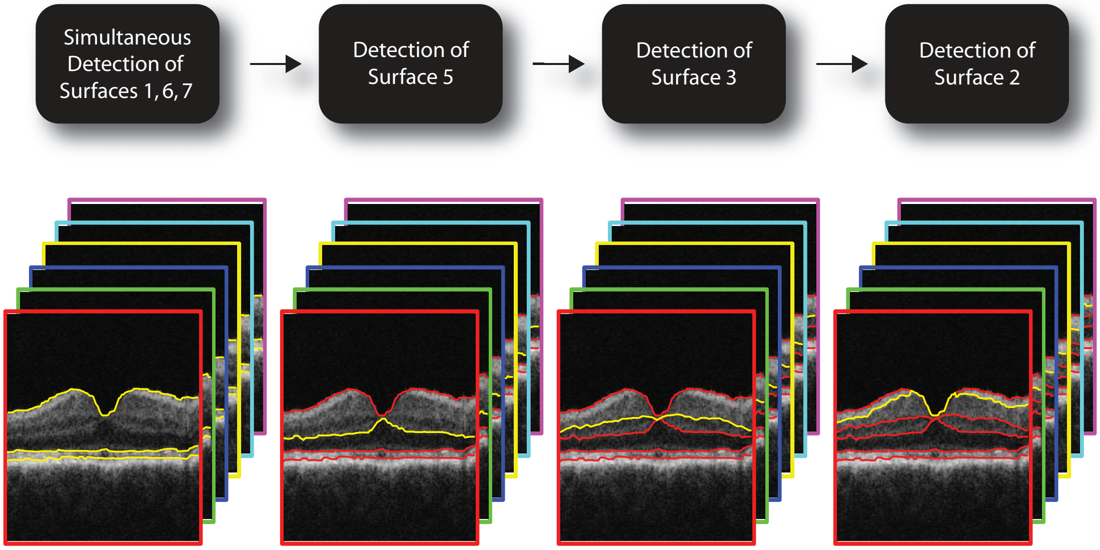


(a) (b)

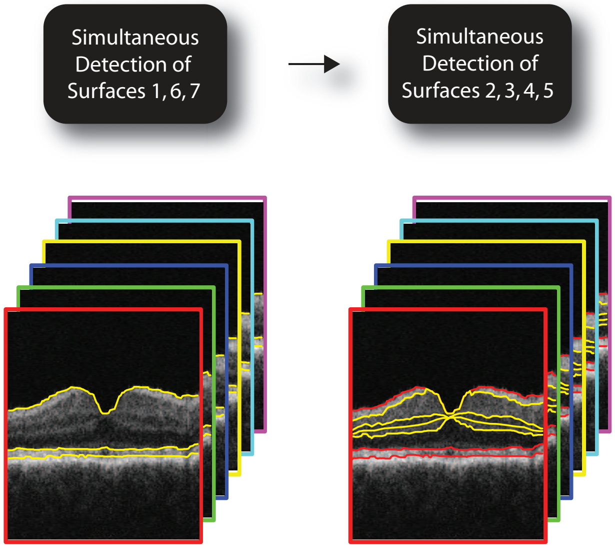
Figure 4.8: Example of using a speckle-reducing anisotropic diffusion (SRAD) method as a preprocessing step (top and bottom of images have been cropped to aid in visualization). (a) Composite scan. (b) Composite scan after application of the SRAD method.

graph search approach. As mentioned previously, the remaining surfaces were then found using one of three variations: 1) sequentially using only on-surface costs and constant constraints [38, 39], 2) simultaneously using only in-region costs and constant constraints [36], or 3) simultaneously using only in-region costs and varying constraints learned from examples. When detected sequentially, surface 4 was not detected. The other surfaces were found in the following order: surface 5, surface 3, and finally, surface 2. Fig. 4.9 illustrates the segmentation order for each variation of the approach.

The nature of the macular scans (Fig. 2.5(a–c)) made it natural to use a discrete cylindrical coordinate system when working with each 3-D composite image (the z-axis coincided with the intersection of the six 2-D composite scans). The coordinates of each voxel could thus be described with the triple (r, θ, z), where r reflected the



(a) Variation 1



(b) Variations 2 and 3

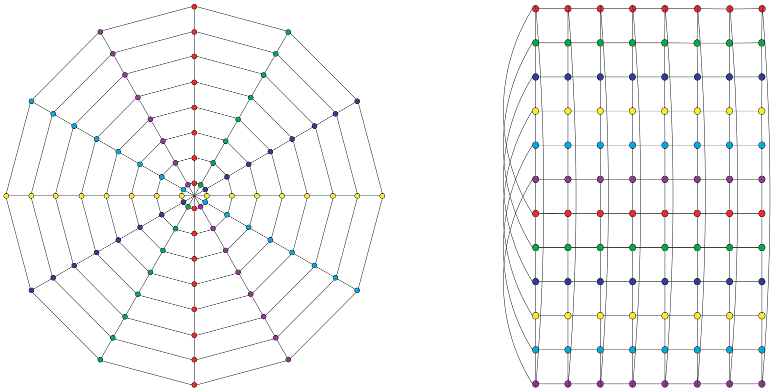
Figure 4.9: Surface segmentation order for each variation of the approach. After finding surfaces 1, 6, and 7 simultaneously, the remaining surfaces were either found sequentially or simultaneously. (a) Sequential detection of interior surfaces when only using on-surface costs in variation 1 [38, 39]. (b) Simultaneous detection of interior surfaces when using in-region costs for interior surfaces in variations 2 and 3 [36, 37].

distance of the voxel from the z-axis, θ reflected the angular location of the voxel (0, 30, 60, 90, 120, 150, 180, 210, 240, 270, 300, or 330), and z reflected the row of the voxel in the corresponding 2-D image. Note that with this coordinate system, voxels in the left half of each 2-D image had a different θ value than those in the right half (for example, for the vertical 2-D scan shown in red in Fig. 2.5(a–c), voxels in the right half of the image had a θ value of 90 while those in the left half had a θ value of 270). Each surface could be defined with a function $f(r, \theta)$, mapping (r, θ) pairs to z-values. The base graph in Fig. 4.10 schematically shows the neighborhood relationship for the columns and the corresponding smoothness constraints when using this coordinate system.

Using this coordinate system only changed the surface feasibility notation described in Chapter 3 slightly. For example, in the constant-constraint case, a surface was considered feasible if [38, 39]:

- θ smoothness constraint: $|f(r, \theta + 30) - f(r, \theta)|$ was less than or equal to Δ_θ for all (r, θ), $\theta \leq 300$.
- Circularity constraint: $|f(r, 0) - f(r, 330)|$ was less than or equal to Δ_θ for all r.
- r smoothness constraint: $|f(r + 1, \theta) - f(r, \theta)|$ was less than or equal to Δ_r for all (r, θ).
- Constraint to connect the left and right halves of the 2-D scans together (a smoothness constraint): $|f(0, \theta) - f(0, \theta + 180)|$ was less than or equal to Δ_r for (r, θ) pairs in which $0 \leq \theta \leq 150$.
- Surface interaction constraint for each pair of surfaces f_1 and f_2: The distance between the two surfaces was at least δ^l voxels and at most δ^u voxels (i.e., $\delta^l \leq f_1(r, \theta) - f_2(r, \theta) \leq \delta^u$ for all (r, θ)).

As described in Chapter 3, the smoothness constraints essentially required the z-values of neighboring surface points (see Fig. 4.10) on a particular surface to be



(a) (b)

Figure 4.10: Schematic view of neighborhood relationship for 3-D macular OCT segmentation. The edges indicate neighborhood connectivity of one "column" of z-values at a (r, θ) pair to another. For each edge shown, smoothness constraints existed between corresponding voxel z-columns for the two (r, θ) pairs connected to the edge. (a) Base graph using cylindrical coordinates. (b) Base graph using unwrapped coordinate system (as was stored in the computer).

within a specified range (given by $\pm\Delta_\theta$ or $\pm\Delta_r$) and the surface interaction constraints required the surface z-values for a particular surface to be within a specified range of the corresponding points on the other surfaces. In the varying constraint case, each of the constraints became a function of (r, θ) location as described for rectangular coordinates in Chapter 3.

4.5 Cost functions

4.5.1 On-surface cost functions (variation 1)

Clearly the defined cost functions were an important component in determining the desired surfaces. In the first variation, the cost function for each surface was constructed from a linear combination of base "intuitive" cost function terms so as to satisfy expected properties of the surface. For example, it was expected that the first surface could be characterized by a combination of the following two properties: 1) the presence of an edge with a dark-to-light transition and 2) the lack of bright voxels above the surface. Correspondingly, the cost function for the first surface was

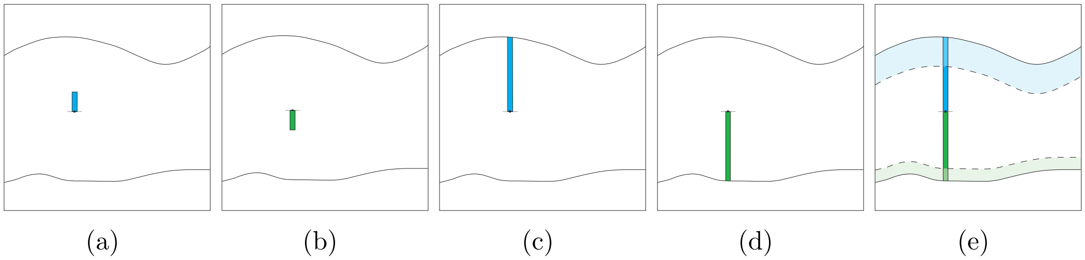


Figure 4.11: Some examples for where the image information comes from in a regional cost function term. Dark borders represent surrounding surfaces (may not be known) of the surface for which the cost function term is being defined. In cases for which an upper or lower surrounding surface does not exist (i.e., the first and last surfaces), the corresponding dark border represents the boundary of the image.

defined as a normalized combination of a signed edge image (to favor the dark-to-light transition) and a cumulative image (created starting at the top of the image so as to discourage the detection of surfaces for which there were many bright pixels above the surface).

The cost functions for all of the surfaces followed the general pattern of having an edge-based term (to either favor a dark-to-light transition or a light-to-dark transition) and one or more localized region-based terms (such as the cumulative image used in the cost function for surface 1). Depending on the prior knowledge of the locations of other surfaces, regional information used in this variation generally was acquired from the locations illustrated in Fig. 4.11. Because both surrounding surfaces of each surface were often not known (surface 2 was the only surface for which the two surrounding surfaces were known) before designing its cost function, it was common to only use regional information from a limited region surrounding the surface (e.g., as in Fig. 4.11(a–b)).

More specifically, each of the surface cost functions was constructed from a normalized combination of a set of the following terms:

- Signed edge term (using Sobel kernel) favoring a dark-to-light transition (used for surfaces 1, 6, and 7).

- Signed edge term (using Sobel kernel) favoring a light-to-dark transition (used for surfaces 2, 3, and 5).
- Summation of pixel intensities in a limited region (Fig. 4.11(a)) above each potential surface voxel to encourage favoring surfaces with dark regions above surface (used for surfaces 6 and 7).
- Negated summation of pixel intensities in a limited region (Fig. 4.11(a)) above each potential surface voxel to encourage favoring surfaces with bright regions above surface (used for surface 2).
- Summation of pixel intensities in a limited region (Fig. 4.11(b)) below each potential surface voxel to encourage favoring surfaces with dark regions below (used for surface 3).
- Negated summation of pixel intensities in a limited region (Fig. 4.11(b)) below each potential surface voxel to encourage favoring surfaces with bright regions below (used for surfaces 6 and 7).
- Cumulative term acquired starting at the top of the image and accumulating downwards (Fig. 4.11(c)) to discourage finding surfaces with bright pixels above the surface (used for surface 1).
- Cumulative term acquired starting from the known boundary below and accumulating upwards (Fig. 4.11(d)) to discourage finding surfaces with bright pixels below the surface (used for surface 5).
- Chan-Vese [42] inspired term that attempted to minimize the intensity variances surrounding the surface. *A-priori* estimated means of the two regions separated by the surface were computed from a region surrounding each known surface (as shown in Fig. 4.11(e) with the lighter intensity region indicated by a dashed line). Because the best use of this term required the prior location of the two surrounding surfaces, only surface 2 used this term.

4.5.2 In-region cost functions (variation 2)

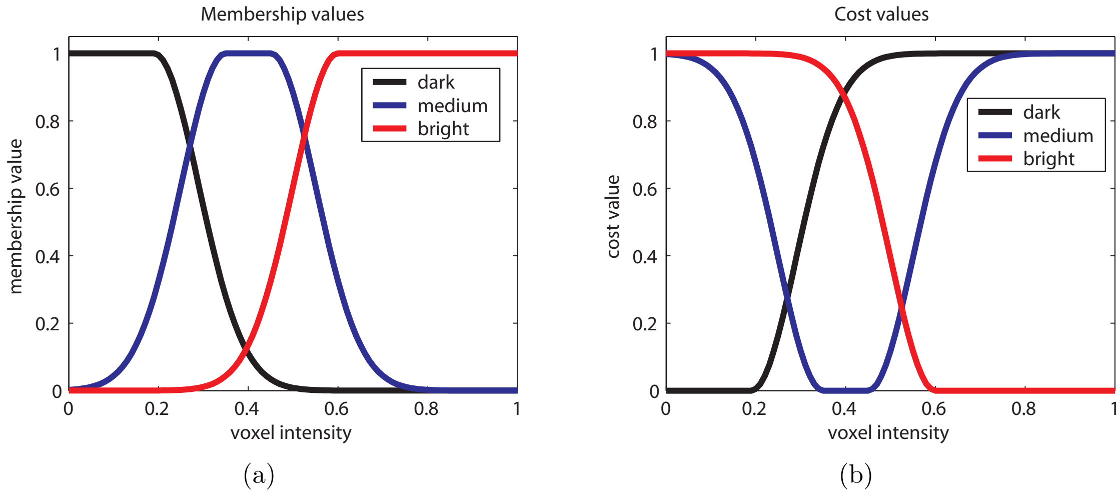
For the work presented at IPMI 2007 [36] and MICCAI 2007 [37], in-region cost functions were used for finding the interior surfaces simultaneously (surfaces 2, 3, 4, and 5). (Surfaces 1, 6, and 7 were simultaneously found using the same on-surface cost functions described in the previous section.)

Motivated by the observation that the intensity of each of the five interior regions could be described as being dark, medium, or bright (region A was bright, region B was medium, region C was dark, region D was medium, region E was dark), the in-region cost values were set based on fuzzy membership functions. Based on Gaussians, each membership function mapped a normalized image intensity value to a value between 0 and 1, with higher values reflecting a greater likelihood of belonging to the particular intensity group. The corresponding cost value was then defined as 1 minus the membership value. Fig. 4.12 shows an example plot of these membership functions and their corresponding cost values. More specifically, the dark membership function, $\text{dark}_{\text{mem}}(x)$, was defined as

$$\text{dark}_{\text{mem}}(x) = \begin{cases} 1 & \text{for } x \leq \Delta d \\ e^{-(x-\Delta d)^2/2\sigma^2} & \text{for } x > \Delta d \end{cases}, \tag{4.4}$$

the medium membership function, $\text{med}_{\text{mem}}(x)$, was defined as

$$\text{med}_{\text{mem}}(x) = \begin{cases} e^{-(x-(c_m-\Delta m))^2/2\sigma^2} & \text{for } x < c_m - \Delta m \\ 1 & \text{for } c_m - \Delta m \leq x \leq c_m + \Delta m \\ e^{-(x-(c_m+\Delta m))^2/2\sigma^2} & \text{for } x > c_m + \Delta m \end{cases}, \tag{4.5}$$



(a) (b)

Figure 4.12: Example dark, medium, and bright membership functions and corresponding cost values.

and the bright membership function, $\text{bright}_{\text{mem}}(x)$, was defined as

$$\text{bright}_{\text{mem}}(x) = \begin{cases} e^{-(x-(1-\Delta b))^2/2\sigma^2} & \text{for } x < 1 - \Delta b \\ 1 & \text{for } x \geq 1 - \Delta b \end{cases} . \tag{4.6}$$

Note that the precise membership functions used could be described by the five parameters Δd, Δm, and Δb, c_m, and σ. To allow for varying membership functions for each image, Δd, c_m, and Δb were estimated from the image data by computing the mean intensity value of regions that were assumed to have a dark, medium, or bright intensity, respectively. The assumed dark region was taken as 50–70 μm above surface 7, the assumed medium region was taken as a 40–60 μm below surface 1 (not including the middle voxels closest to the fovea), and the assumed bright region was taken as 0–24 μm below surface 7. These estimates could be taken because surfaces 1, 6, and 7 had already been determined.

4.6 Incorporation of varying constraints (variation 3)

For each subject, the smoothness and surface interaction constraints were learned in a leave-one-subject-out fashion as described in Chapter 3 (but only in the radial direction as opposed to in both the x and y-directions). Varying surface interaction constraints (thickness constraints) were used only in the simultaneous segmentation of the interior surfaces 2, 3, 4 and 5. The varying constraints were computed for the following surface pairings: 1-2, 2-3, 3-4, 4-5, and 5-6, with cost function modification used to enforce the constraints for surface pairings 1-2 and 5-6.

4.7 Validation methods

4.7.1 Data

Each variation of the algorithm was tested on fast macular scans from 12 subjects with unilateral chronic anterior ischemic optic neuropathy. Note that the unilateral nature of the disease meant that we had data for 24 eyes, 12 of which were affected by optic neuropathy, 12 of which were not. In almost all cases (21/24 eyes), six repeated series were used to create the 3-D composite image for each eye. (Each of the remaining three eyes used fewer than six repeated series to create the 3-D composite image.) The resulting 24 3-D composite images were each comprised of 6 composite 2-D scans (144 total composite 2-D scans) of size 128 × 1024 pixels. The corresponding reported physical width and height of the 2-D raw scans (and thus also the composite scans) was 6 mm × 2 mm, resulting in a pixel size of approximately 50 μm (horizontally) × 2 μm (vertically).

4.7.2 Manual tracing software

Custom-made software for manually tracing boundaries on time-domain Stratus OCT-3 images (Fig. 4.13) was developed. The software allowed for the "manual alignment" of the scans to a particular boundary for better visualization while tracing.

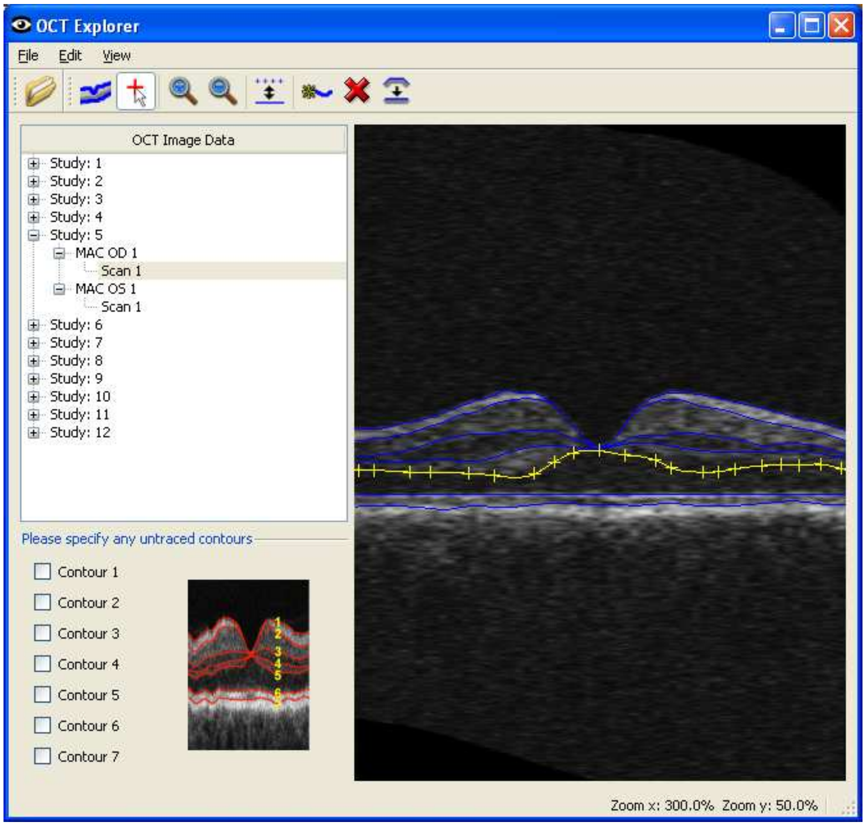


Figure 4.13: Screenshot of manual tracing software developed for tracing macular time-domain OCT scans.

4.7.3 Validation of variation 1

One raw scan from each eye was independently traced by three human experts (using the software illustrated in Fig. 4.13) with the average of the three tracings being used as the reference standard. Experts were masked to the clinical status of the scanned eye. The experts did not attempt to trace borders that they did not consider visible. The algorithmic result on the corresponding composite 2-D scan was converted into the coordinate system of the raw scan (undoing alignment/registration) and the mean signed and unsigned border positioning errors for each border were computed (the middle 30 pixels were not included to exclude the fovea). The signed and unsigned border positioning errors were also computed between the observers in the following ways:

- Observer 1 versus observer 2,
- Observer 2 versus observer 3,
- Observer 1 versus observer 3,
- Observer 1 versus the average of observers 2 and 3,
- Observer 2 versus the average of observers 1 and 3, and
- Observer 3 versus the average of observers 1 and 2.

Standard deviations (over images) were also computed.

In addition, in order to compute an example clinically meaningful measure, the mean thickness (again, not including the middle 30 pixels to exclude the fovea) of layers defined from the first border to each of the remaining borders was computed using the algorithm and each observer. The average thicknesses computed from the three observers was used as a reference standard and the absolute differences between the algorithmic thicknesses and reference thicknesses were computed. Furthermore, absolute thickness differences were computed between the observers using the same cases as was done for computing the border positioning errors.

4.7.4 Validation of variation 2

One raw scan from each eye was independently traced by two human experts with the average of the two tracings being used as the reference standard. As in variation 1, the experts did not attempt to trace borders that were not considered visible. Again, the algorithmic result on the corresponding composite 2-D scan was converted into the coordinate system of the raw scan (inversely transforming the alignment/registration) and the mean and the maximum unsigned border positioning errors for each border were computed (the middle 30 pixels were not included to exclude the fovea). The unsigned border positioning errors were also computed using one observer as a reference standard for the other. For each border, a paired t-test was used to test for significant differences in the computed mean border positioning errors (with Bonferroni correction, p-values $< \frac{0.05}{7} = 0.007$ were considered significant).

4.7.5 Validation of variation 3

The same reference standard was used as in variation 2. The mean unsigned border positioning errors for each border were computed for the method using (piecewise) constant constraints and the approach with varying constraints. However, the middle 30 pixels *were* included in this case. The unsigned border positioning errors were also computed using one observer as a reference standard for the other.

4.8 Results using only on-surface costs (variation 1)

Tables 4.2 and 4.3 summarize the computed border position errors. Table 4.4 and Fig. 4.14 summarize the thickness difference results. The border positioning errors and thickness differences between the algorithm and the reference standard were very similar to those computed between the observers. For example, the algorithm's overall unsigned border positioning error was 6.1 ± 2.9 μm, while the overall observer error averaged 6.9 ± 3.3 μm (ranging from 5.5 ± 3.0 μm for observer 2 versus the average

Table 4.2: Variation 1: Summary of mean unsigned border positioning errors.†

Border	Avg. Obs. vs. Algo.	Obs. 1 vs. Obs. 2	Obs. 2 vs. Obs. 3	Obs. 1 vs. Obs. 3	Obs. 1 vs. Obs. 2,3	Obs. 2 vs. Obs. 1,3	Obs. 3 vs. Obs. 1,2
1	4.4 ± 1.2	4.0 ± 1.1	4.8 ± 1.4	5.9 ± 1.3	4.3 ± 0.9	3.3 ± 1.0	5.0 ± 1.3
2*	7.6 ± 3.1	6.9 ± 4.4	7.5 ± 4.6	5.8 ± 1.2	5.3 ± 2.3	6.5 ± 4.4	5.9 ± 2.3
3*	6.9 ± 2.0	8.2 ± 2.1	8.6 ± 2.6	8.1 ± 3.0	6.9 ± 2.4	7.5 ± 1.7	7.3 ± 2.3
5	7.0 ± 2.9	7.5 ± 3.0	9.1 ± 3.6	9.4 ± 4.4	7.3 ± 3.2	6.8 ± 2.4	8.4 ± 3.9
6	3.0 ± 1.1	4.4 ± 1.6	5.5 ± 1.8	7.8 ± 2.8	5.6 ± 2.1	3.4 ± 1.0	6.4 ± 2.3
7	7.6 ± 2.9	8.5 ± 4.8	8.0 ± 3.0	11.5 ± 4.6	9.2 ± 4.5	6.2 ± 3.1	9.0 ± 3.0
overall	6.1 ± 2.9	7.0 ± 3.5	7.2 ± 3.3	8.2 ± 3.8	6.5 ± 3.2	5.5 ± 3.0	7.0 ± 3.0

† Mean ± SD for 24 scans in μm. For each boundary, differences were not computed for the middle 30 pixels (out of 128) to exclude the fovea.

* Errors were not computed for those scans in which boundary was determined to not be visible by at least one expert.

of observers 1 and 3 to 8.2 ± 3.8 μm for observer 1 versus observer 3). In terms of thickness differences, the smallest errors for both the algorithm and the observers were for the layer defined by surfaces 1 and 6 (2.2 ±1.8 μm for the algorithm and an average of 2.2 ± 1.9 μm for the observers), while the largest errors for the algorithm were for the layer defined by surfaces 1 and 7 (6.2 ± 3.9 μm). (However, note that surface 7 was also the surface displaying the largest bias based on the signed border positioning errors.) The largest thickness differences between the observers were for the layer defined by surfaces 1 and 5 (overall average error of 8.2 ± 4.5 μm, which is larger than the algorithm's error of 4.8 ± 4.7 μm for this layer). Three example results (reflecting the best case, the median case, and the worst case according to the overall unsigned border positioning error) are shown in Fig. 4.15. The mean intraretinal layer segmentation time (after alignment/registration) was 4.1 ± 0.9 minutes (using a Windows XP workstation with a 3.2 GHz Intel Xeon CPU).

4.9 Results using in-region costs (variation 2)

The computed unsigned and maximal border position errors are summarized in Table 4.5. Except for the unsigned border positioning errors for surface 2 (which

Table 4.3: Variation 1: Summary of mean signed border positioning errors.[†]

Border	Avg. Obs. vs. Algo.	Obs. 1 vs. Obs. 2	Obs. 2 vs. Obs. 3	Obs. 1 vs. Obs. 3	Obs. 1 vs. Obs. 2,3	Obs. 2 vs. Obs. 1,3	Obs. 3 vs. Obs. 1,2
1	0.1 ± 2.5	1.4 ± 1.6	3.8 ± 1.9	5.2 ± 1.7	3.3 ± 1.3	1.2 ± 1.6	-4.5 ± 1.6
2[*]	1.4 ± 3.9	-1.4 ± 5.7	-1.4 ± 5.8	-2.7 ± 2.5	-2.0 ± 3.3	0.0 ± 5.6	2.0 ± 3.5
3[*]	-3.9 ± 3.4	-0.1 ± 3.9	-5.0 ± 4.0	-5.0 ± 3.6	-2.5 ± 3.2	-2.5 ± 3.5	5.0 ± 3.3
5	-3.0 ± 4.9	0.2 ± 4.8	-6.9 ± 5.1	-6.7 ± 6.1	-3.3 ± 4.8	-3.5 ± 3.9	6.8 ± 5.1
6	0.9 ± 1.9	2.7 ± 2.1	4.8 ± 2.4	7.5 ± 3.1	5.1 ± 2.4	1.1 ± 1.7	-6.2 ± 2.6
7	-5.6 ± 4.0	6.1 ± 6.9	5.0 ± 6.1	11.1 ± 4.9	8.6 ± 5.2	-0.5 ± 6.0	-8.0 ± 4.3
overall	-1.8 ± 4.4	1.6 ± 5.1	0.2 ± 6.5	1.9 ± 7.8	1.7 ± 5.7	-0.7 ± 4.4	-1.0 ± 6.8

[†] Mean ± SD for 24 scans in μm. For each boundary, differences were not computed for the middle 30 pixels (out of 128) to exclude the fovea.

[*] Errors were not computed for those scans in which boundary was determined to not be visible by at least one expert.

Table 4.4: Variation 1: Summary of mean absolute thickness differences.[†]

Border	Avg. Obs. vs. Algo.	Obs. 1 vs. Obs. 2	Obs. 2 vs. Obs. 3	Obs. 1 vs. Obs. 3	Obs. 1 vs. Obs. 2,3	Obs. 2 vs. Obs. 1,3	Obs. 3 vs. Obs. 1,2
1–2[*]	3.1 ± 3.1	4.4 ± 5.0	7.1 ± 3.7	8.1 ± 3.2	5.8 ± 3.0	4.6 ± 3.7	6.9 ± 3.5
1–3[*]	4.1 ± 3.3	4.1 ± 2.6	8.8 ± 4.2	10.2 ± 3.8	5.9 ± 3.5	4.4 ± 3.2	9.4 ± 3.6
1–5	4.8 ± 4.7	3.5 ± 2.8	10.7 ± 5.0	11.9 ± 6.2	6.6 ± 4.6	5.0 ± 3.1	11.3 ± 5.2
1–6	2.2 ± 1.8	2.2 ± 1.9	2.0 ± 1.4	3.1 ± 2.9	2.4 ± 2.4	1.3 ± 1.0	2.4 ± 2.0
1–7	6.2 ± 3.9	6.4 ± 5.5	5.2 ± 3.9	6.0 ± 4.7	5.9 ± 4.3	5.0 ± 4.2	4.7 ± 3.3

[†] Mean ± SD for 24 scans in μm.. For each boundary, thickness differences were not computed for the middle 30 pixels (out of 128) to exclude the fovea.

[*] Errors were not computed for those scans in which the lower boundary was determined to not be visible by at least one expert.

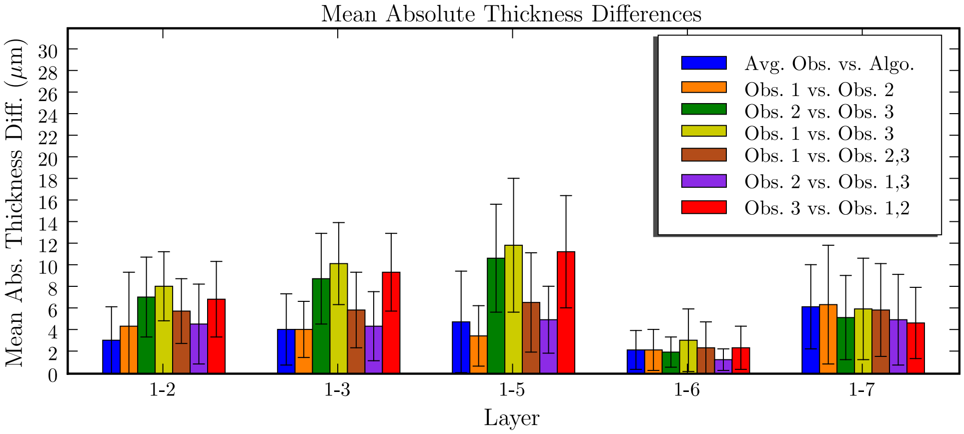
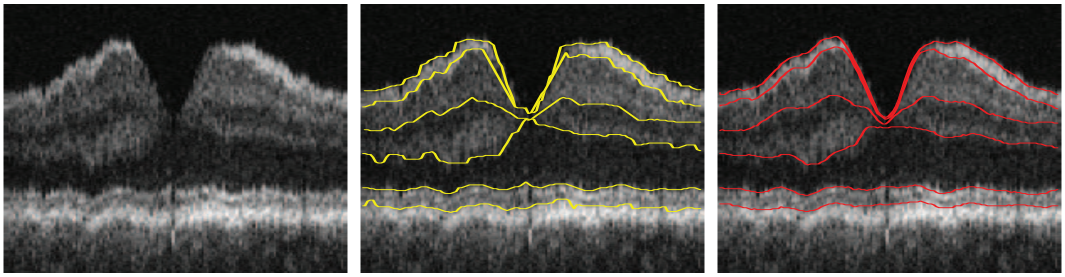


Figure 4.14: Bar chart of mean thickness differences using variation 1 (error bars reflect standard deviations).

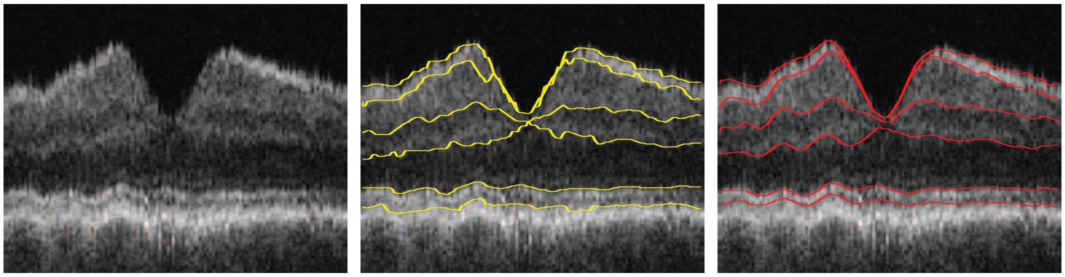
was significantly greater, $p < 0.001$), the computed mean errors for all the surfaces were significantly smaller or not significantly different from that between the human observers ($p < 0.001$ for surface 1; $p = 0.11$ for surface 3; $p = 0.04$ for surface 4; $p = 0.80$ for surface 5; $p < 0.001$ for surface 6; $p = 0.004$ for surface 7). The overall mean (all borders combined) unsigned border positioning error for the algorithm was 7.8 ± 5.0 μm with an overall maximum unsigned border positioning error of 24.7 ± 12.9 μm. This was comparable to the overall mean and maximum border positioning errors computed between the observers which were 8.1 ± 3.6 μm and 22.8 ± 9.2 μm, respectively, and compared well with the true 9–10 μm resolution of the OCT imaging system reported in the literature [44]. An example result is shown in Fig. 4.16.

4.10 Results using varying constraints (variation 3)

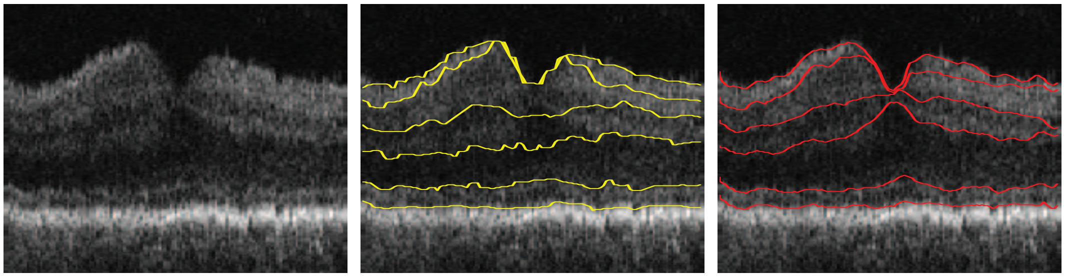
The computed unsigned border positioning errors are summarized in Table 4.6. Introducing the varying constraints significantly decreased the mean unsigned border positioning errors [overall mean error of 7.3 ± 3.7 μm using varying constraints



(a) Best Case (b) Algorithm (c) Ref. Standard



(d) Median Case (e) Algorithm (f) Ref. Standard



(g) Worst Case (h) Algorithm (i) Ref. Standard

Figure 4.15: Three example results reflecting the best, median, and worst performances according to the overall unsigned border positioning error using the first variation. (a) Best case composite image. (b) Best case composite image with segmented borders. (c) Best case composite image with average manual tracing. (d) Median case composite image. (e) Median case composite image with segmented borders. (f) Median case composite image with average manual tracing. (g) Worst case composite image. (h) Worst case composite image with segmented borders. (i) Worst case composite image with average manual tracing.

Table 4.5: Variation 2: Summary of mean unsigned border positioning errors.[†]

	Algorithm vs. Avg. Observer		Observer 1 vs. Observer 2	
Border	Mean	Maximum	Mean	Maximum
1	4.0 ± 1.2	16.9 ± 9.0	5.9 ± 1.3	16.4 ± 5.0
2*	11.2 ± 5.2	37.1 ± 11.9	5.8 ± 1.2	21.5 ± 8.6
3*	10.0 ± 4.7	29.0 ± 9.3	8.4 ± 3.3	26.0 ± 11.8
4*	10.4 ± 5.1	31.4 ± 14.3	7.7 ± 2.1	22.7 ± 6.6
5	9.1 ± 6.5	27.1 ± 13.2	9.4 ± 4.4	28.5 ± 12.5
6	3.5 ± 2.0	13.1 ± 7.5	7.8 ± 2.8	19.3 ± 5.6
7	7.8 ± 2.5	22.5 ± 7.2	11.5 ± 4.6	24.8 ± 5.8

[†] Mean ± SD for 24 in μm. For each boundary, differences were not computed for the middle 30 pixels (out of 128) to exclude the fovea.

[*] Errors were not computed for those scans in which boundary was determined to not be visible by at least one expert.

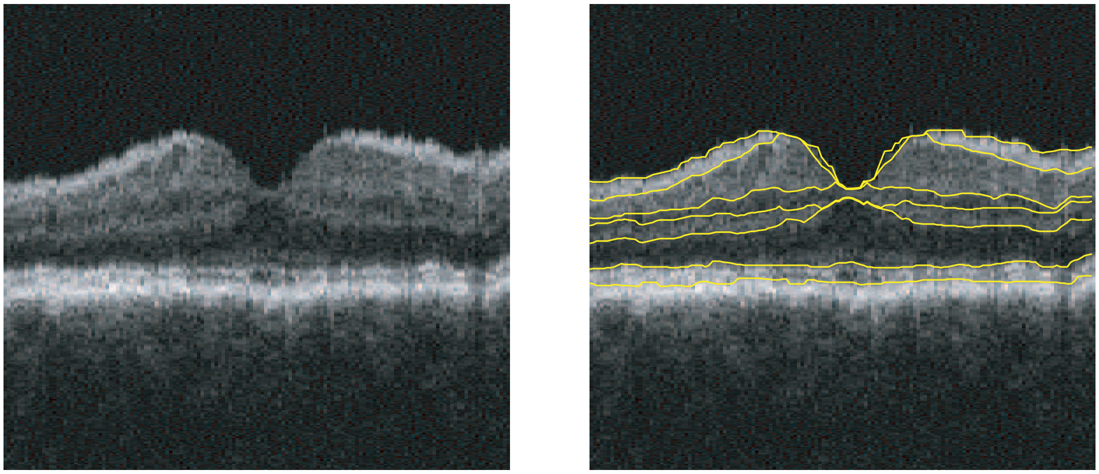

(a) (b)

Figure 4.16: Example result using variation 2 shown on a 2-D scan from one of the 3-D images.

Table 4.6: Variation 3: Summary of mean unsigned border positioning errors[†].

		Algorithm vs. Avg. Observer	
Border	Obs. 1 vs. Obs. 2	Varying S & I	Constant S & I
1	5.9 ± 1.2	4.6 ± 1.6	5.1 ± 2.3
2[‡]	6.3 ± 1.0	10.9 ± 4.3	11.2 ± 5.0
3[‡]	9.1 ± 3.7	8.9 ± 4.4	10.2 ± 4.9
4[‡]	7.6 ± 2.0	8.6 ± 1.9	12.0 ± 4.5
5	9.4 ± 3.9	8.3 ± 2.9	9.7 ± 5.4
6	7.6 ± 2.4	3.0 ± 1.0	3.6 ± 2.5
7	11.4 ± 4.7	7.6 ± 2.2	7.8 ± 2.2
Overall	8.2 ± 3.5	7.3 ± 3.7	8.3 ± 4.9

[†] Mean ± SD for 24 scans in μm.

[‡] Errors were not computed for those scans in which boundary was determined to not be visible by at least one observer.

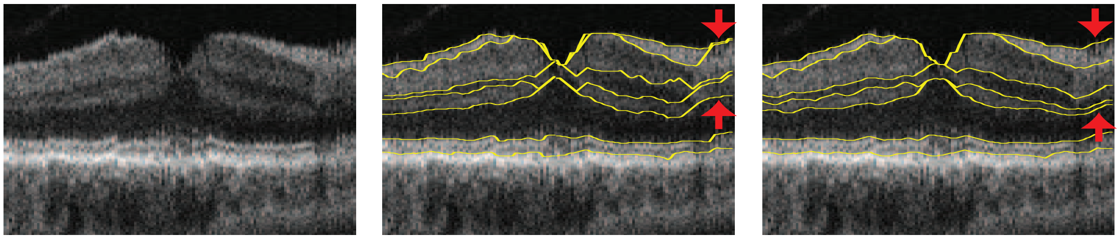

(a) Original (b) Constant S & I (c) Varying S & I

Figure 4.17: Example improvement in segmentation result by using varying smoothness and surface interaction constraints (variation 3) [37].

compared to 8.3 ± 4.9 μm using constant constraints ($p < 0.001$ using paired t-test) and 8.2 ± 3.5 μm for the inter-observer variability ($p = 0.01$)]. Even more noticeable differences occurred in cases in which image information was locally ambiguous. Fig. 4.17 shows an example improvement.

4.11 Discussion

This chapter has presented three different variations for segmenting the interior surfaces of macular time-domain images. When comparing the results to the human experts, all variations performed well (achieving quantitative results similar to that of human experts).

Variations 1 and 2 reflect two different very different choices for the cost functions: variation 1 used only on-surface costs, while variation 2 used only in-region costs. The choice of using *only* in-region costs (instead of having both in-region and on-surface costs) for the segmentation of interior surfaces in variation 2 was to demonstrate the feasibility of such type of costs since they had never been used in an application before. Thus, it was not to suggest that only using in-region costs would be a better approach than only using on-surface costs. In fact, both approaches performed similarly well. (Note, however, that the quantitative results are not directly comparable since a different reference standard was used in the two cases). Nevertheless, a more appropriate cost-function choice would most likely be to find an appropriate combination on on-surface and in-region costs, as is done for the spectral case in the next chapter.

Given a particular cost function, using varying constraints instead of constant constraints will likely improve the segmentation results (e.g., as was indicated by the comparisons made for variation 3 when using in-region cost functions for the interior surfaces). In essence, having varying constraints allows you to more appropriately define what is considered to be a feasible set of surfaces, which is particularly helpful when the image information is locally ambiguous.

To our knowledge, the work summarized in this chapter [36–40] reflects the first reported approach for the *3-D* segmentation of intraretinal layers on OCT images. Admittedly, the sparsely acquired data makes the utility of a 3-D approach for time-domain cases not as noticeable as would be observed with true volumetric data. However, perhaps most importantly, it sets the important stage for being able to segment spectral-domain images in 3-D (see next chapter).

CHAPTER 5
SPECTRAL OCT SEGMENTATION

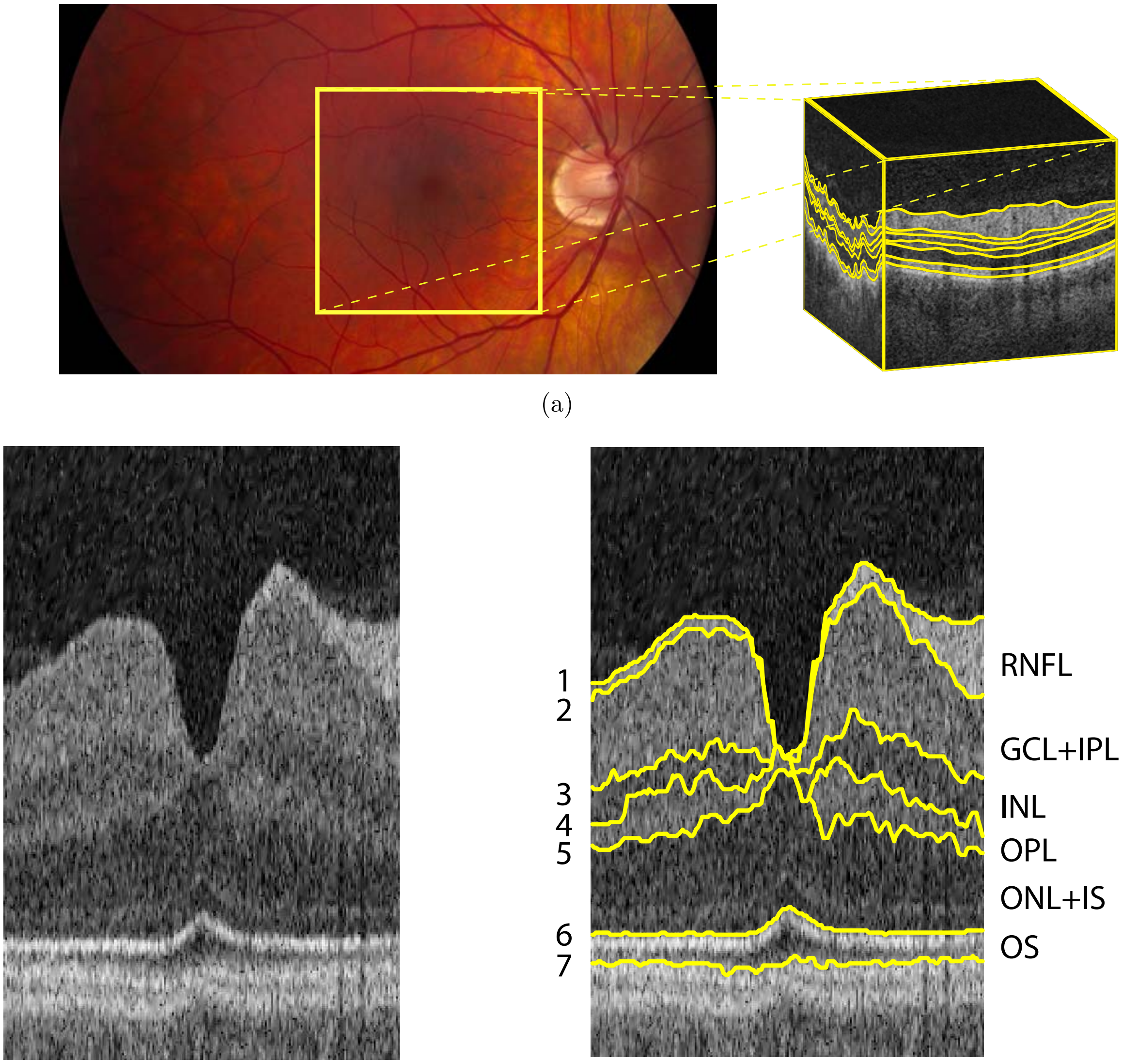
This chapter presents details about the methods, experimental methods, and results used for segmenting macular spectral-domain OCT images.

5.1 Intraretinal layer segmentation of macular scans

Fig. 5.1 illustrates the seven surfaces we desired to find on macular spectral-domain OCT images. Even though the surfaces were the same as those segmented in the time-domain case, the larger coverage of the macular region created a greater variety of layer appearances (see example slices from a macular spectral-domain volumetric image illustrated in Fig. 5.2). An overview of the steps used for segmenting spectral-domain OCT images is illustrated in Fig. 5.3. As in the time-domain case, the images were first flattened so that the surfaces near the RPE became approximately a flat plane. However, three major differences in the spectral flattening step existed from the time-domain case: 1) the preliminary segmentation of the first group of surfaces (surfaces 1, 6, and 7) was performed in 3-D instead of in 2-D, 2) the preliminary segmentation was performed on a downsampled image (by a factor of four) to save time and memory requirements, and 3) a thin-plate spline (with regularization) was used for fitting the found RPE surface instead of using 2-D least-squares splines. After flattening, the intraretinal layer segmentation of all seven surfaces was performed on a truncated full-resolution version of the 3-D image. More details about the flattening and intraretinal layer segmentation steps can be found in the next few sections.

5.1.1 Flattening

Flattening the 3-D OCT volumes served multiple purposes: 1) to allow for smaller image sizes in the segmentation step, 2) to provide for a more consistent shape for segmentation purposes, and 3) to make visualization easier. For example, having

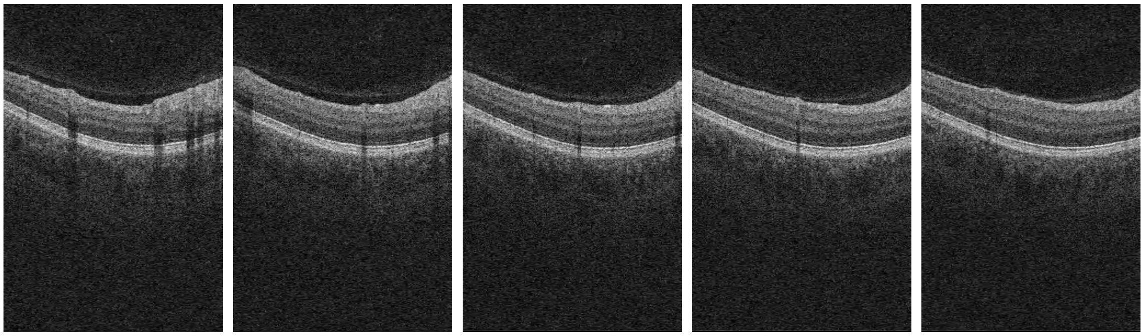


(a)

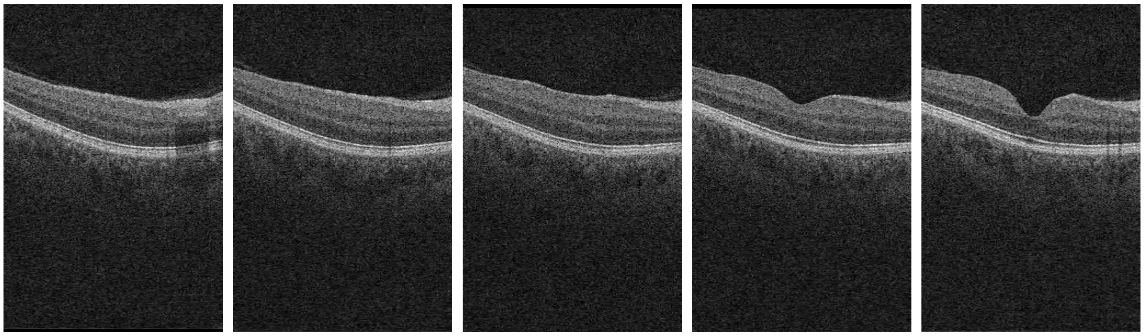
(b)

(c)

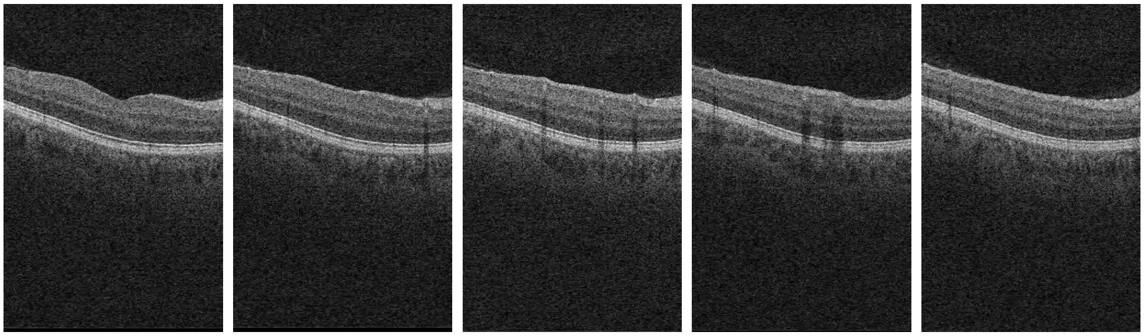
Figure 5.1: Illustration of surfaces to be segmented on macular spectral-domain OCT images. (a) Fundus photograph with schematic OCT volume (size 6 × 6 × 2 mm^3) with surfaces indicated with yellow lines. (b) Example slice (flattened/truncated) from the center of an OCT volume. (c) Seven surfaces (and corresponding six layers) on example slice. (RNFL = retinal 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.)



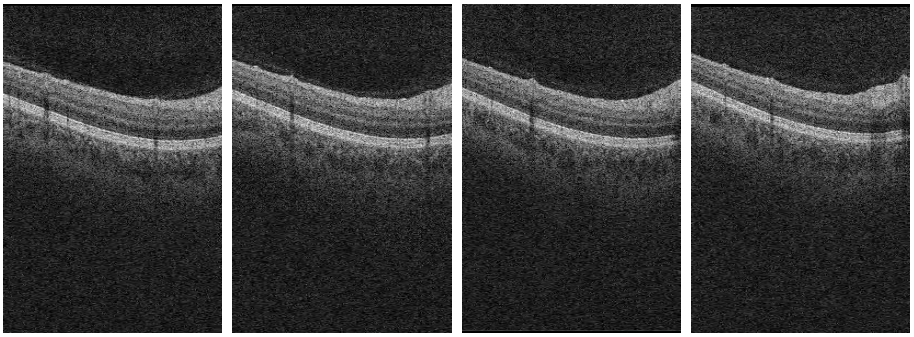
(a) Slice 0 (b) Slice 11 (c) Slice 22 (d) Slice 33 (e) Slice 44



(f) Slice 55 (g) Slice 66 (h) Slice 77 (i) Slice 88 (j) Slice 99



(k) Slice 110 (l) Slice 121 (m) Slice 132 (n) Slice 143 (o) Slice 154



(p) Slice 165 (q) Slice 176 (r) Slice 187 (s) Slice 198

Figure 5.2: Example slices from a macular spectral-domain OCT volume.

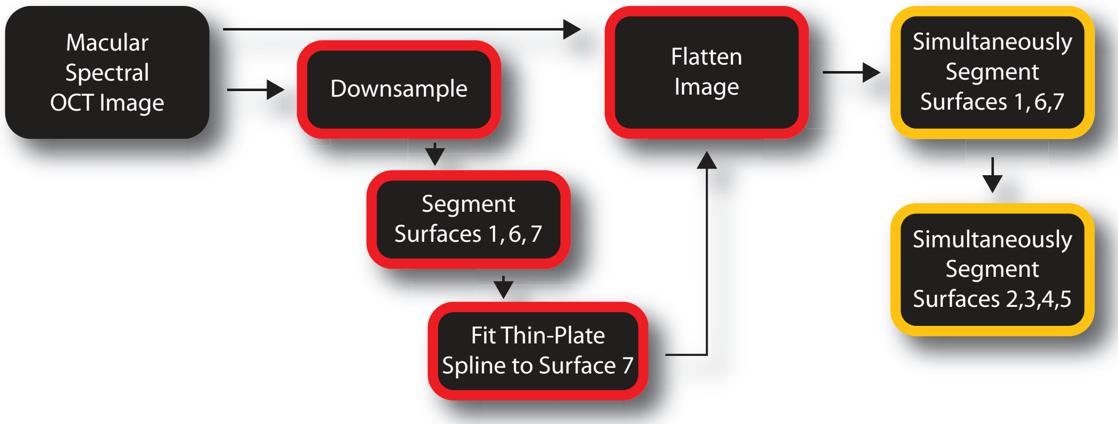
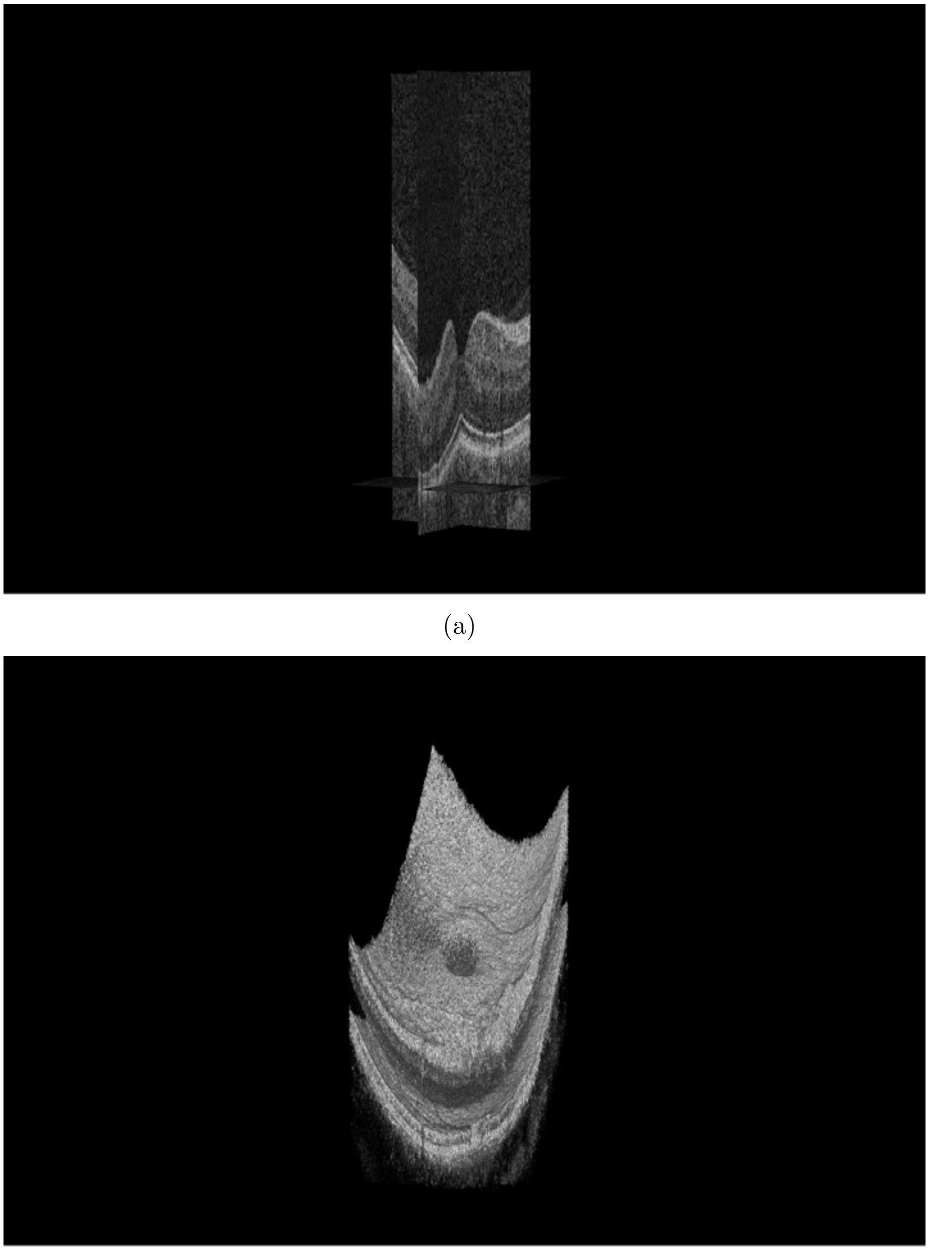


Figure 5.3: Overview of steps used in segmenting macular spectral OCT images.

a consistent shape aided in both visualization and learning appropriate constraints for segmentation. In addition, flattening the image made it possible to truncate the image substantially in the axial (z-direction), thereby reducing the memory and time-requirements of the intraretinal layer segmentation approach. Flattening an image involved the following steps:

- Downsample image (by a factor of 4).
- Simultaneously find surfaces 1, 6, and 7 on downsampled image using the optimal 3-D graph search approach.
- Fit a (regularized) thin-plate spline [45] to an upsampled version of surface 7.
- Translate all the columns of the (full-resolution) image so that the fitted surface becomes a flat plane.
- Truncate the flattened image based on surface locations 1 and 7.

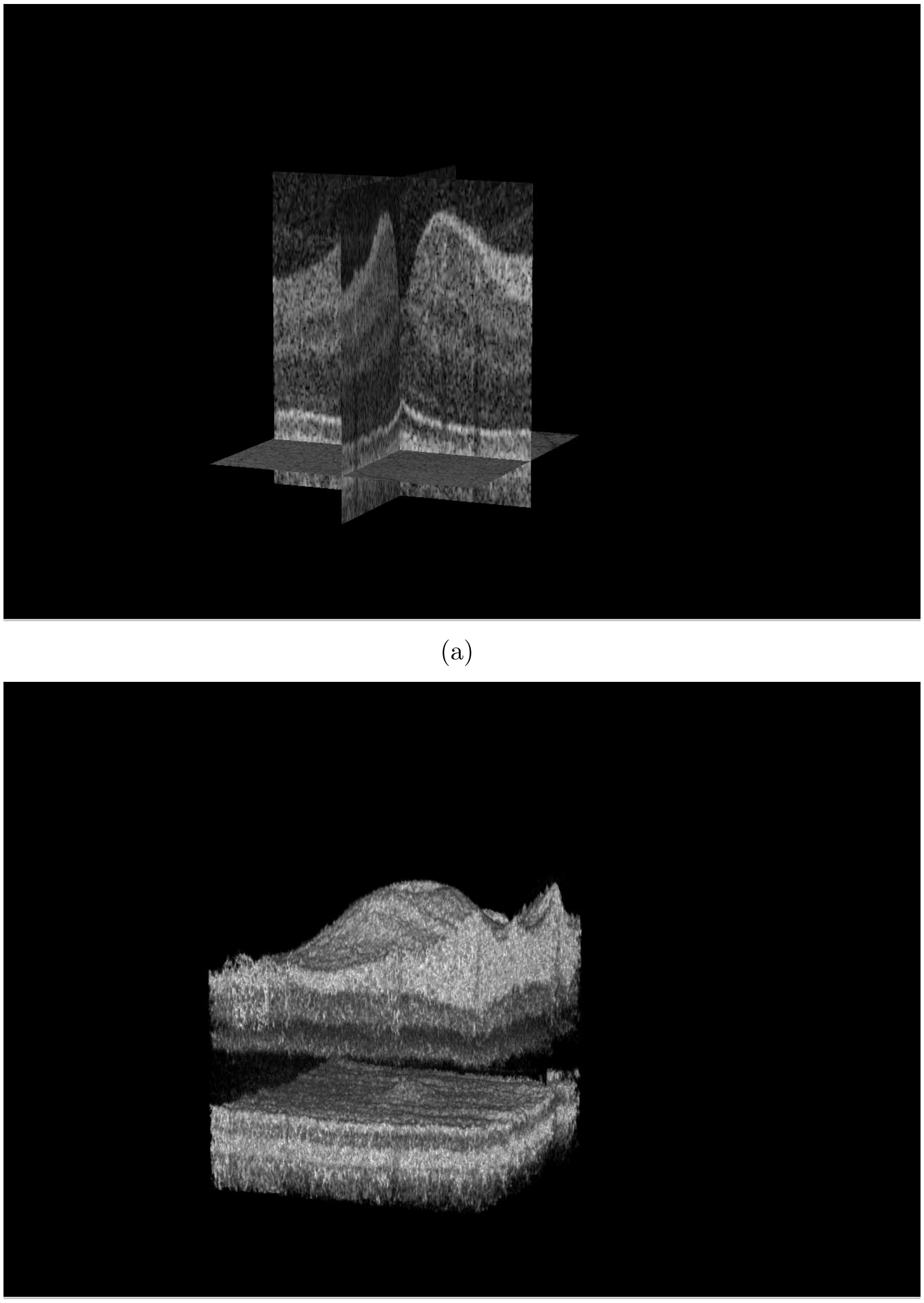
An example image before flattening is illustrated in Fig. 5.4. The same image after flattening is illustrated in Fig. 5.5.



(a)

(b)

Figure 5.4: Rendering of spectral OCT image before flattening. (a) View of three orthogonal planes. (b) Volume rendering.



(a)

(b)

Figure 5.5: Rendering of spectral OCT image after flattening. (a) View of three orthogonal planes. (b) Volume rendering.

5.1.2 Intraretinal layer segmentation

As in the time-domain case, after applying a speckle-reducing anisotropic diffusion method [41], the surfaces on the flattened and truncated volumetric spectral images were segmented in two groups: surfaces 1, 6, and 7 were segmented simultaneously first followed by the simultaneous segmentation of surfaces 2, 3, 4, and 5. The cost functions used for surfaces 1, 6, and 7 were the same as those used in our prior work for the time-domain case (see previous chapter and [38, 39]). Recall that the cost functions for these surfaces were on-surface cost functions comprised of an edge-term (favoring a dark-to-bright transition using a Sobel kernel) and localized regional terms (cumulative term used for surface 1, terms favoring surfaces with a dark region above and bright region below used for surfaces 6 and 7).

The cost functions for the interior surfaces (surfaces 2, 3, 4, and 5) used a combination of on-surface and in-region cost functions. The used on-surface cost functions for the interior surfaces were signed edge terms (using a Sobel kernel) to favor either a bright-to-dark transition (used for surfaces 2, 3, and 5) or a dark-to-bright transition (used for surface 4). The in-region cost terms were based on fuzzy membership functions in dark, medium, or bright intensity classes (similar to what was used for the time-domain case). Based on Gaussians, each membership function mapped a normalized image intensity to a value between 0 and 1, with higher values reflecting a greater likelihood of belonging to the particular intensity group. The corresponding cost value was then defined as 1 minus the membership value. As in the time-domain case, the dark membership function, $\text{dark}_{\text{mem}}(x)$, was defined as

$$\text{dark}_{\text{mem}}(x) = \begin{cases} 1 & \text{for } x \leq \Delta d \\ e^{-(x-\Delta d)^2/2\sigma^2} & \text{for } x > \Delta d \end{cases}, \tag{5.1}$$

the medium membership function, $\mathrm{med}_{\mathrm{mem}}(x)$, was defined as

$$\mathrm{med}_{\mathrm{mem}}(x) = \begin{cases} e^{-(x-(c_m-\Delta m))^2/2\sigma^2} & \text{for } x < c_m - \Delta m \\ 1 & \text{for } c_m - \Delta m \leq x \leq c_m + \Delta m \\ e^{-(x-(c_m+\Delta m))^2/2\sigma^2} & \text{for } x > c_m + \Delta m \end{cases}, \tag{5.2}$$

and the bright membership function, $\mathrm{bright}_{\mathrm{mem}}(x)$, was defined as

$$\mathrm{bright}_{\mathrm{mem}}(x) = \begin{cases} e^{-(x-(1-\Delta b))^2/2\sigma^2} & \text{for } x < 1 - \Delta b \\ 1 & \text{for } x \geq 1 - \Delta b \end{cases}. \tag{5.3}$$

The region between surfaces 1 and 2 used a bright membership function (adjustable by parameter Δb), the region between surfaces 2 and 3 used a medium membership function (adjustable by parameter c_m), the region between surfaces 3 and 4 used a dark membership function (adjustable by parameter Δd_1), the region between surfaces 4 and 5 used a medium membership function (adjustable by parameter c_m), and the region between surfaces 5 and 6 used a dark membership function (adjustable by parameter Δd_2). Note that there were actually two dark parameters, Δd_1 and Δd_2, for the two dark regions. This choice was based on our experience in noticing that the region bounded by surfaces 5 and 6 tended to have smaller intensity values than the region bounded by surfaces 3 and 4.

The parameters Δb, c_m, and Δd_2 were automatically estimated for each volumetric image by computing the mean intensity in regions that were expected to have that related intensity class. These regions could be estimated by utilizing learned thickness constraints. For example, the RNFL (layer defined by surfaces 1 and 2) must contain the voxels between surface 1 and the surface defined by adding the minimum allowed distance between surfaces 1 and 2 to surface 1. The mean intensity in this region could thus be used for estimating (1 - Δb). Table 5.1 provides a summary of the bounding

Table 5.1: Bounding surfaces for estimating in-region cost parameters.[†]

Parameter	Top estimation surface	Bottom estimation surface
Δb	$f_1(x, y)$	$f_1(x, y) + \delta^l_{1-2}(x, y)$
c_m	$f_1(x, y) + \delta^u_{1-2}(x, y)$	$f_1(x, y) + \delta^l_{1-2}(x, y) + \delta^l_{2-3}(x, y)$
Δd_2	$f_6(x, y) - \delta^l_{5-6}(x, y)$	$f_6(x, y)$

[†] (x, y) columns for which the top estimation surface would go below the bottom estimation surface were excluded

surfaces for each region. Columns for which the top surface would overlap with the bottom surface were excluded from the estimation region. Δd_1 was estimated based on finding the relationship between Δd_2 and Δd_1 using training data (i.e., $\Delta d_1 = \Delta d_2 + c$, where c was a constant determined from the training data). The other parameters, σ and Δm were set to reasonable fixed values based on our prior experiments with the time-domain cases.

An appropriate combination of the on-surface and in-region cost functions for these surfaces were experimentally determined based on a training set, as described further in the Experimental methods section (Section 5.2). Similarly, the surface smoothness and surface interaction constraints used for finding the interior surfaces were experimentally determined based on a training set (again, described in more detail in the Experimental methods section).

5.2 Experimental methods

5.2.1 Subject data

Macula-centered 3-D OCT volumes ($200 \times 200 \times 1024$ voxels, $6 \times 6 \times 2$ mm^3) were obtained from the right eyes of 27 normal subjects using three Cirrus$^{\text{TM}}$ HD-OCT machines (provided by Carl Zeiss Meditec, Inc., Dublin, CA, USA). As illustrated in Fig. 5.6, the data were organized into two groups:

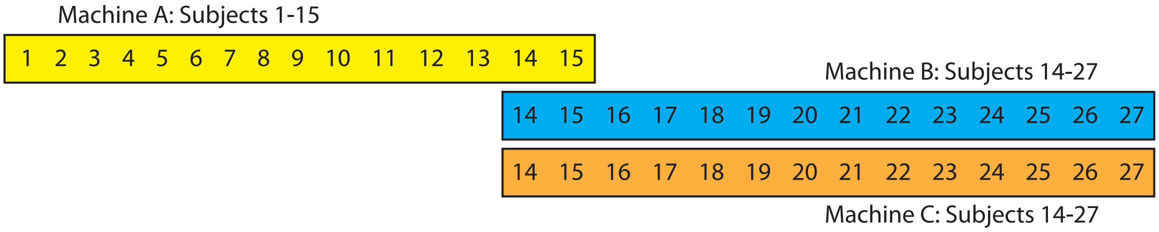


Figure 5.6: Summary of data available for training and testing the macular spectral OCT segmentation approach.

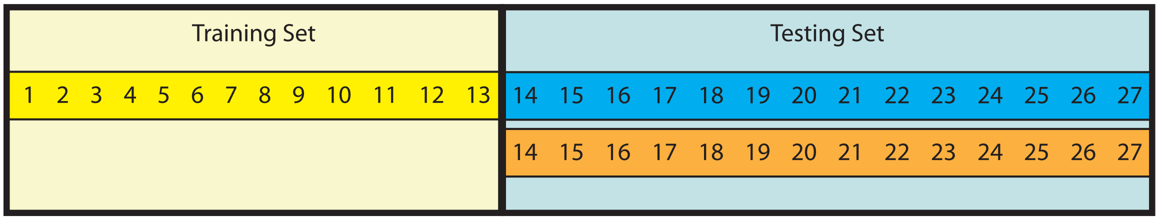


Figure 5.7: Data using for training and testing. The scans of the 13 subjects for which repeated scans were not available were used as a training set. The repeated scans of the 14 other subjects were used as a testing set.

- *Normals with age information:* a 3-D OCT volume of the right eye of each subject 1–15 (mean age of 39.4 ± 8.2 years) was acquired using one machine (machine A).
- *Normals with repeated scans:* a 3-D OCT volume of the right eye of each subject 14–27 was acquired twice using two different machines (machine B and machine C).

Thus, 43 total volumes were acquired ($15 + 14 \times 2$). The volumetric scans of the right eyes of subjects 1–13 (using machine A) were used for training the algorithm and the repeated volumes of the right eyes of subjects 14–27 (using machines B and C) were used for testing the algorithm (Fig. 5.7).

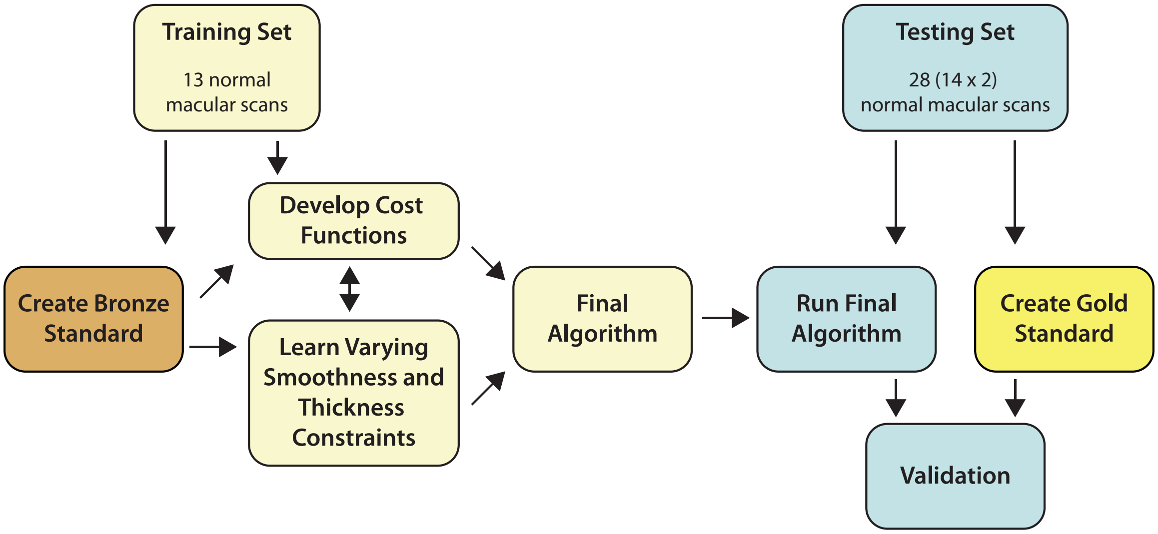


Figure 5.8: Spectral training/testing flowchart.

5.2.2 Reference standards

Fig. 5.8 illustrates the use of reference standards in the training and testing process. Two reference standards were used: a "bronze" standard (using for training) and a "gold" standard (used for testing). The bronze standard involved creating complete edited segmentations for all 13 volumes in the training set (editing performed by one observer). This standard was used for learning the constraints and cost functions for use in the final algorithm. The gold standard was created by having two ophthalmologists trace 10 slices on each volume and then taking the average of these tracings to be used as the reference standard. This standard was used for validation of the approach on the test set. To prevent an unintended bias from influencing the final algorithm, no segmentations were performed on the test set until the training was complete. More details about the creation of the bronze and gold standards are given in the following sections.

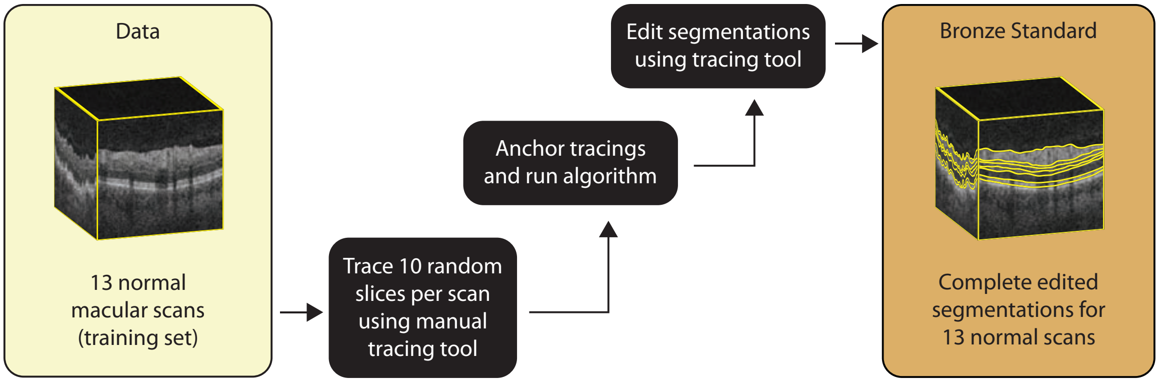


Figure 5.9: Steps used in creating the bronze standard.

5.2.2.1 Bronze standard (used for training)

A summary of the steps used in creating the bronze standard are illustrated in Fig. 5.9. First, each volume was divided into 10 regions (illustrated in Fig. 5.10) and one slice was randomly selected in each region. Each region consisted of 22 slices, expect for the middle two regions (near the fovea), which consisted of 12 slices each. Using custom-designed manual tracing software (as was also used for the time-domain case), the surfaces were manually traced on each of the 10 slices. Although only 10 slices were traced in each volume, all 200 slices were available to provide 3-D context (typically viewed in a cine-like mode). In addition, for better visualization, the software allowed for translating the columns to flatten the image to a manually traced boundary of interest (typically boundary 6 or 7).

After the boundaries were manually defined on each of the selected slices, the volumes were flattened to the RPE by fitting a thin-plate spline to this manually traced surface and then translating the columns so that this surface became a flat plane. Next, using the manual tracings as "anchor points," a preliminary version of the segmentation approach was performed to segment the remaining slices. The "anchoring" was done by modifying the edge-based cost functions to have a very large value for all pixels in the selected slices except for the pixels on the "anchored

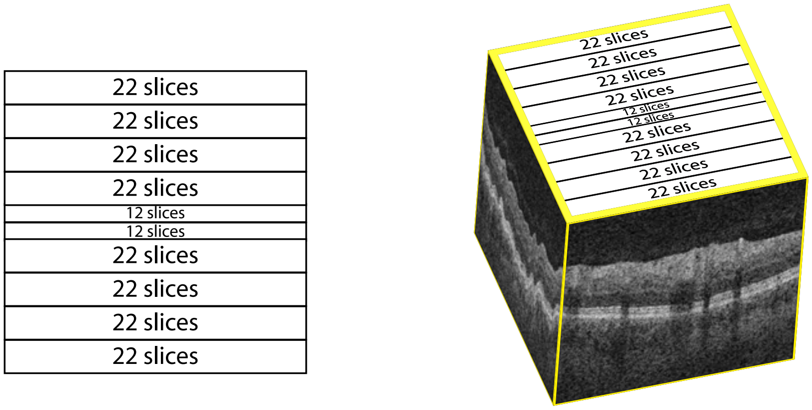


Figure 5.10: Regions for randomly selecting slices for the macular spectral OCT independent standard (one slice was randomly chosen from each region).

surface." Because the constraints had not been learned yet, the smoothness and surface interaction constraints were set to values more permissive than were expected to be necessary. The cost functions for surfaces 1, 6, and 7 were the same as those used in prior work for the time-domain case [36–39]. The cost functions for the interior surfaces (surfaces 2, 3, 4, and 5) were set in a similar fashion as in the time-domain method using regional information only [36], with the exception that the estimation of the dark, medium, and bright intensity classes were set by computing the mean of those actual regions on the traced slices (e.g., the mean bright intensity was estimated by computing the mean value of those pixels between surfaces 1 and 2 on the 10 traced slices).

Finally, the resulting surfaces were reloaded back into the manual tracings application to be edited. For ease of editing, only every 5th point was loaded (a spline was used to interpolate the boundaries between the points). The editing was performed on the flattened images. The final complete edited segmentations defined the bronze standard.

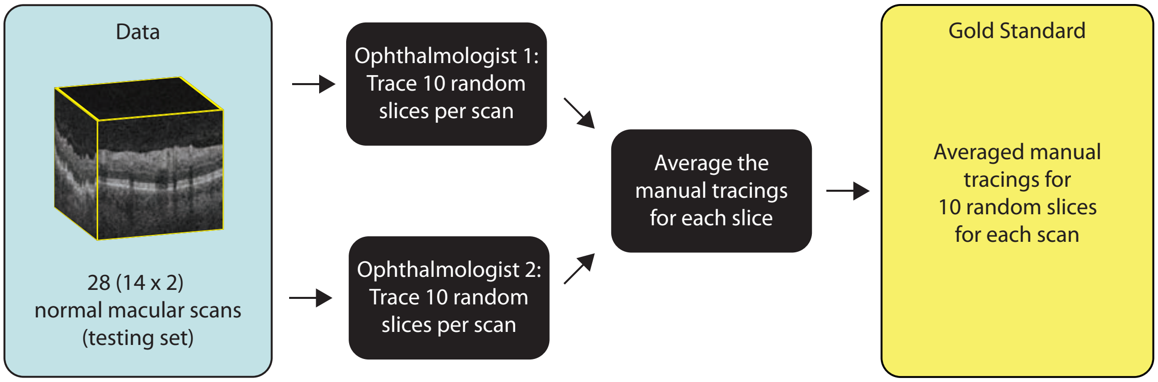


Figure 5.11: Steps used in creating the gold standard.

5.2.2.2 Gold standard (used for testing)

A summary of the steps used in creating the gold standard are illustrated in Fig. 5.9. For each 3-D OCT volume, 10 slices were randomly traced independently by two ophthalmologists. The slices were selected from the same regions as was done for bronze standard (illustrated in Fig. 5.10). The average of these tracings defined the gold standard.

5.2.3 Learning the constraints

Using the complete edited segmentations from the training set (bronze standard), the surface interaction ("thickness") and smoothness constraints for the interior surfaces were learned as described in a general fashion in Section 3.4.1 of Chapter 3. The smoothness constraints were learned in the x-direction and y-direction separately.

5.2.4 Learning cost function parameters

As mentioned in the previous section, most of the in-region parameters were determined automatically in a subject-specific fashion during the segmentation method itself by using the learned thickness constraints to determine the regions of interest for estimation. However, the relationship between the two dark parameters were estimated based on the training data. In particular, it was assumed that this relationship could be described as $\Delta d_1 = \Delta d_2 + c$. Then the constant c was estimated by taking the difference of the mean intensity of the two dark regions based on the 10×13 manually traced slices (before segmentation/editing) in the training set.

To learn an appropriate combination of the on-surface and in-region cost functions for the simultaneous segmentation of surfaces 2, 3, 4, and 5, first the overall cost function was described using one parameter (α) to indicate the relative weight of the on-surface and in-region cost terms ($0 \leq \alpha \leq 1$):

$$C_{\{f_2(x,y),f_3(x,y),\ldots,f_5(x,y)\}} = \alpha \sum_{i=2}^{5} C_{f_i(x,y)} + (1-\alpha) \sum_{i=1}^{5} C_{R_i} \ , \tag{5.4}$$

where

$$C_{f_i(x,y)} = \sum_{\{(x,y,z)|z=f_i(x,y)\}} c_{\text{surf}_i}(x,y,z) \ , \tag{5.5}$$

and

$$C_{R_i} = \sum_{(x,y,z)\in R_i} c_{\text{reg}_i}(x,y,z) \ . \tag{5.6}$$

With this description, $\alpha = 1$ corresponded to only using on-surface cost functions and $\alpha = 0$ corresponded to only using in-region cost terms. Various values of α were tested on images in the training set. For each value of α, the mean unsigned border positioning error from the bronze standard for each surface was computed. Values of α in increments of 0.1 were first tested, followed by using smaller increments as deemed necessary for obtaining the best value.

Next, in order to allow for a different weighting factor for each surface, the cost function for the interior surfaces was written as:

$$C_{\{f_2(x,y),\ldots,f_5(x,y)\}} = c_2 C_{f_2(x,y)} + c_3 C_{f_3(x,y)} + c_4 C_{f_4(x,y)} + c_5 C_{f_5(x,y)} + \sum_{i=1}^{5} C_{R_i} \ , \tag{5.7}$$

where c_2, c_3, c_4, and c_5 were constant weighting factors used for the on-surface cost terms. Note that in this formulation the weighting factor for the in-region costs was set to 1 so that only the on-surface weights would change. Letting α_i be the best value of α for surface i, the weighting c_i for each on-surface cost function was set as follows:

$$c_i = \begin{cases} \alpha_i/(1-\alpha_i) & \text{if } 1-\alpha_i > 0 \\ C_{\text{large}} & \text{if } 1-\alpha_i = 0 \end{cases}, \tag{5.8}$$

where C_{large} is a relatively large constant value. In this formulation, the desired ratio (determined from α_i) of the on-surface costs to the in-region costs was preserved for each surface. However, it could also change the relative "importance" of each surface with respect to the others if the values of α varied greatly amongst the surfaces. Thus, this cost function formulation was then run on the cases in the training set to ensure that it in fact produced a smaller overall error. (If not, then only one value of α would be used as in Eq. 5.4.) This defined the "final" cost function to be used in the test set.

5.2.5 Validation on the test set

5.2.5.1 Border positioning errors

After training was complete (learning the constraints and cost functions), the final algorithm was run on all images in the test set. Signed and unsigned border positioning differences were computed for the following cases (over the 10 selected slices for each volume):

- ophthalmologist 1 versus ophthalmologist 2,
- the algorithm versus the average of ophthalmologists 1 and 2 (gold standard),
- the algorithm versus ophthalmologist 1, and
- the algorithm versus ophthalmologist 2.

In order to compare the interobserver variability of the ophthalmologists with the errors of the algorithm, differences in the "errors" (e.g., comparing the border positioning differences between the two ophthalmologists to the border positioning differences between algorithm and the gold standard) for each border were fitted using a univariate analysis of variance (ANOVA) model. The models were fitted using the GLM (general linear models) procedure from the Statistical Analysis System (SAS) version 9.1. Resulting p-values were adjusted using the SIMULATE option for pairwise comparisons among least squares method-specific means that were estimated using the GLM procedure. This adjustment approach provided family-wise error rate protection by computing adjusted p-values based on simulated data. The type I error rate was specified to be 0.05.

5.2.5.2 Thickness reproducibility

For each volume, individual layer thicknesses (on the 10 selected slices) were computed based on the following surface-definition methods:

- the algorithm,
- the manual tracings of ophthalmologist 1,
- the manual tracings of ophthalmologist 2, and
- the average tracings from ophthalmologist 1 and ophthalmologist 2 (gold standard).

Next, for each subject, the absolute thickness difference between machine B and machine C were computed for each of the surface-definition methods above. A multivariate analysis of variance (MANOVA) model was fitted using the GLM (general linear models) procedure from the Statistical Analysis System (SAS) version 9.1 to compare the vector of layer thickness difference measurements for each of the surface-definition methods above (algorithm, ophthalmologist 1, ophthalmologist 2, average ophthalmologist). To identify any specific differences, a univariate analysis of variance (ANOVA) model was fit for each layer. As before, resulting p-values were then adjusted using the SIMULATE option for pairwise comparisons among least squares method-specific means that were estimated using the GLM procedure and the type I error rate was specified to be 0.05.

5.3 Results

5.3.1 Training set

5.3.1.1 Learned constraints

Fig. 5.12 illustrates the resulting learned thickness (surface interaction) constraints. For reference, the computed mean thicknesses are also shown in the figure. A visualization of the thickness standard deviation values are illustrated in Fig. 5.13. The learned smoothness constraints (the maximum allowed decrease of z-value in the x-direction, the maximum allowed increase of z-value in the x-direction, the maximum allowed decrease of z-value in the y-direction, and the maximum allowed increase of z-value in the y-direction) are illustrated in Fig. 5.14. As a point of clarification, "increasing" or "decreasing" refers to increasing or decreasing the z-value when moving from one column to the next (e.g., moving from column (x,y_i) to column $(x+1,\ y_i)$ for the x-direction constraints) in a coordinate system in which $z = 0$ was at the *top* of the images. However, it may be more intuitive to think of $z = 0$ as being located at the bottom of the images. If this were the case, the "increase" and "decrease" labels of the images in Fig. 5.14 would just have to be swapped.

5.3.1.2 Cost function combination

Fig. 5.15 summarizes the computed unsigned positioning errors for each tested value of α. As mentioned previously, first values of α in increments of 0.1 were tested. For additional precision in determining the best value of α for each surface, the following additional values of α were also tested: 0.83, 0.85, 0.88, 0.93, 0.95, and 0.98. Table 5.2 summarizes the resulting best values of α for each surface. The last column of this table provides the ratio of α_i to $1-\alpha_i$, which was used as the on-surface cost weighting term.

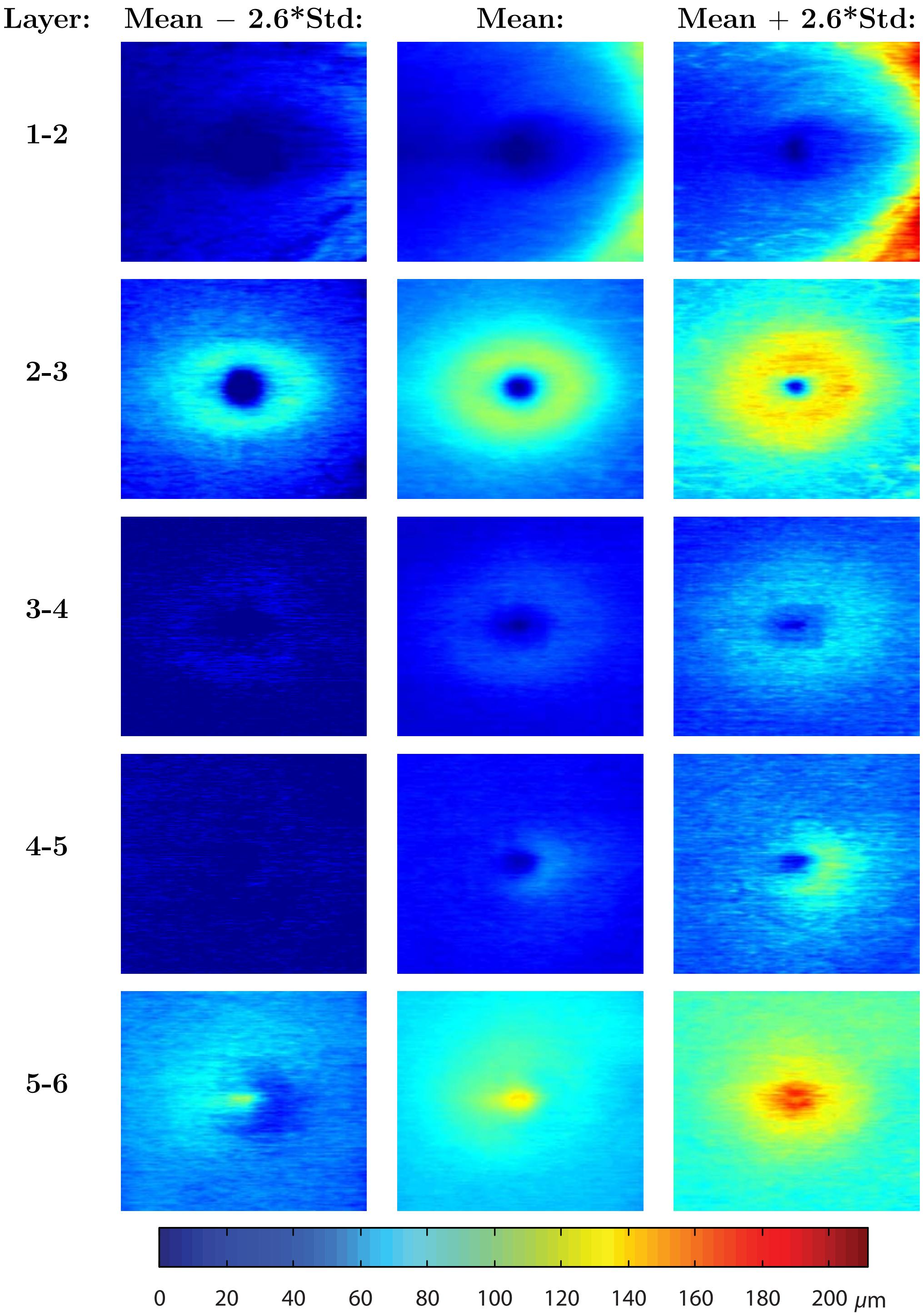
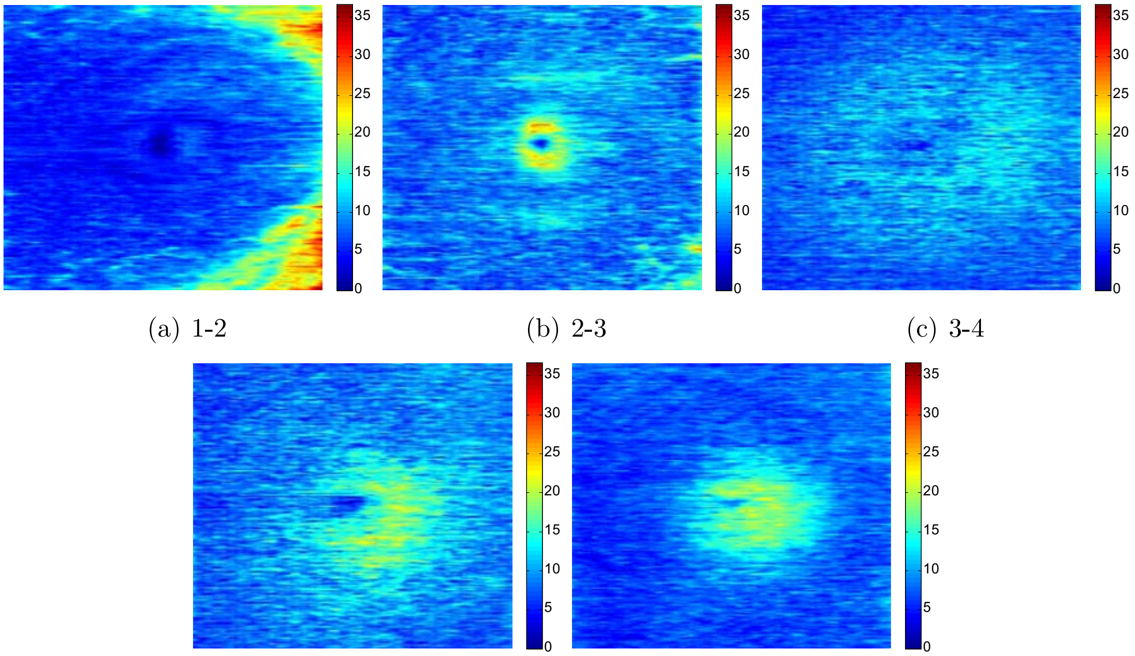


Figure 5.12: Visualization of learned thickness constraints. The minimum and maximum thickness constraints are illustrated in the second and fourth columns, respectively, while the mean thicknesses are illustrated in the third column. Each image is oriented so that the temporal side is on the left and the nasal side is on the right.



(a) 1-2 (b) 2-3 (c) 3-4

(d) 4-5 (e) 5-6

Figure 5.13: Visualization of thickness standard deviation values in microns.

Table 5.2: Best values of α for each surface.

Surface	α_i	$1-\alpha_i$	$\frac{\alpha_i}{1-\alpha_i}$
2	0.95	0.05	19.0
3	0.80	0.20	4.0
4	0.83	0.17	4.9
5	0.83	0.17	4.9

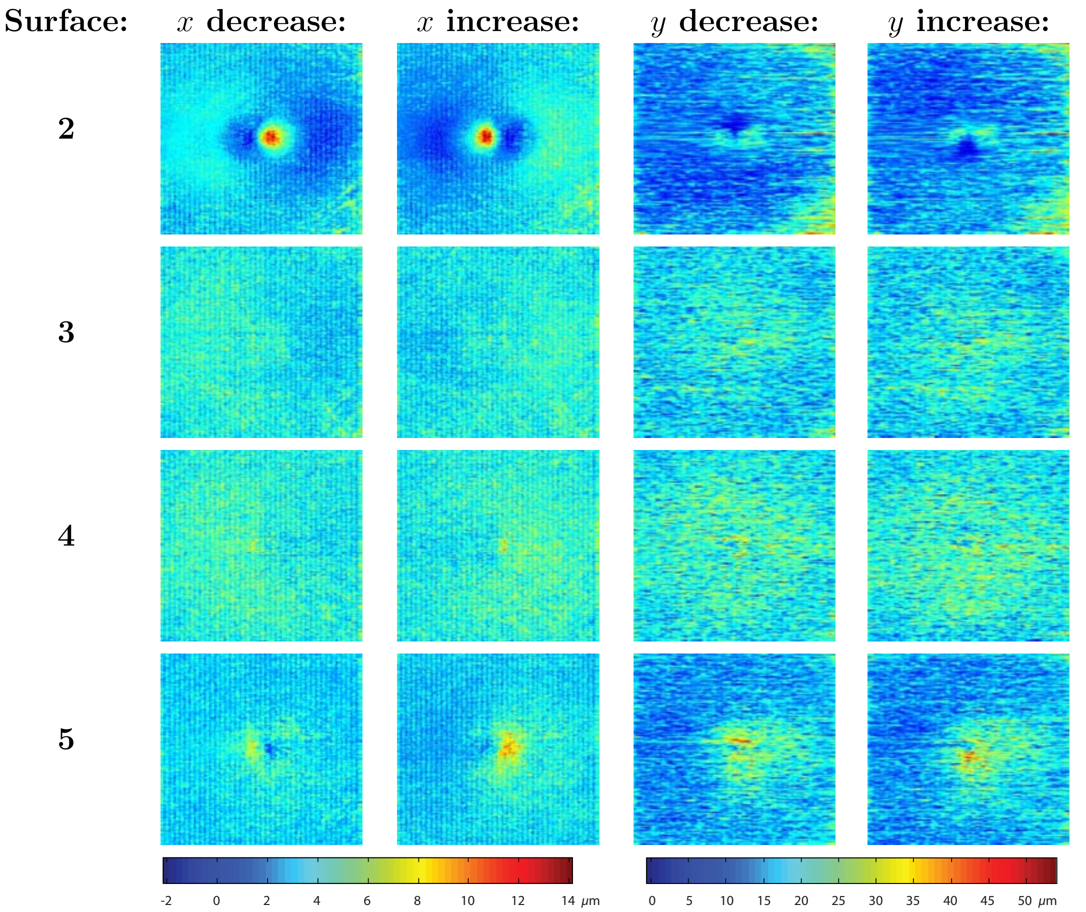


Figure 5.14: Learned smoothness constraints.

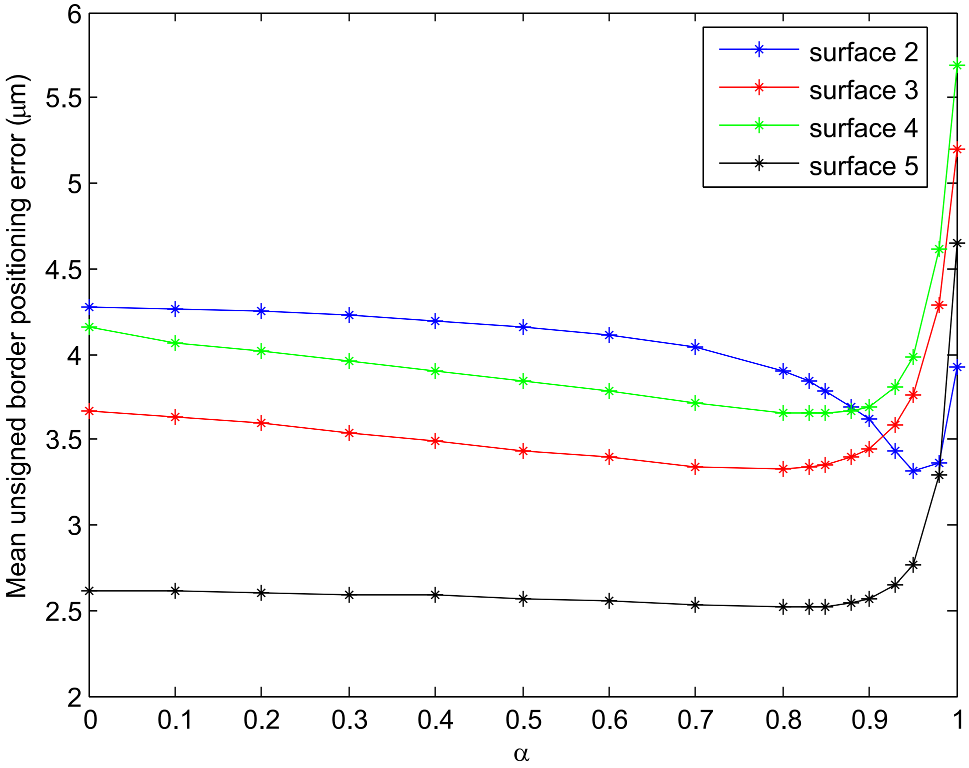


Figure 5.15: Unsigned border positioning error results for different combinations of edge and regional information on training set.

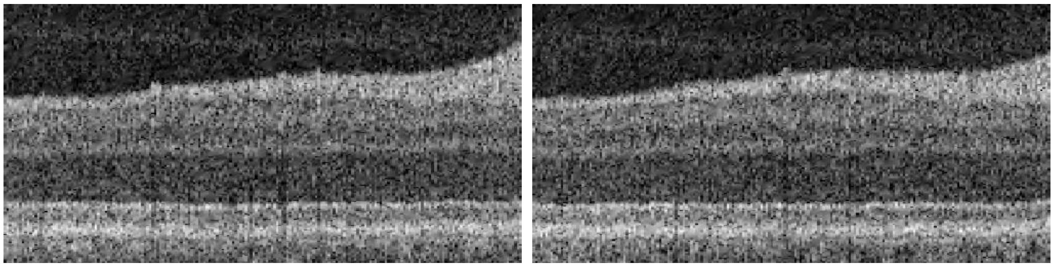
5.3.2 Test set

5.3.2.1 Example results

Qualitatively, the algorithm performed very well overall, with only minor local inaccuracies. The one exception might be surface 7, as it would sometimes "flip" between two probable surfaces in a localized region (this will be discussed more in the discussion). Fig. 5.16 and Fig. 5.17 illustrate the slices/results on the volume with the smallest overall unsigned border positioning error. (Slices without resulting surfaces are shown in Fig. 5.16 and the flattened slices with resulting surfaces are shown in Fig. 5.17). Similarly, Fig. 5.18 and Fig. 5.19 illustrate the slices/results on the volume with the median overall unsigned border positioning error and Fig. 5.20 and Fig. 5.21 illustrate the slices/results on the volume with the largest overall unsigned border positioning error.

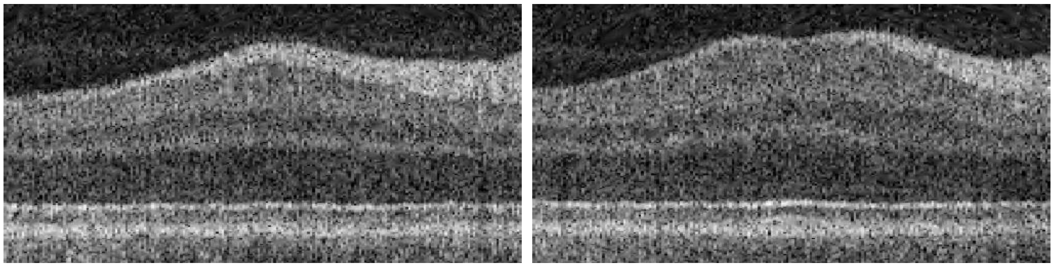
5.3.2.2 Border positioning errors

The resulting unsigned and signed border positioning errors are illustrated in Table 5.3 and Table 5.4, respectively. The overall mean unsigned border positioning error was 5.71 ± 1.98 μm for the comparison between the two ophthalmologists, 5.69 ± 2.41 μm for the comparison between the algorithm and the average ophthalmologist, 6.73 ± 2.48 μm for the comparison of the algorithm versus ophthalmologist 1, and 5.89 ± 2.74 μm for the comparison of the algorithm versus ophthalmologist 2. When looking at each border individually (using separate ANOVA models and adjusting the p-values to take into account the multiple comparisons), the unsigned errors of the algorithm (when compared to the gold standard defined as the average tracings of the two ophthalmologists) were significantly smaller than those between the two ophthalmologists for border 1 ($p < 0.0001$), border 2 ($p = 0.0112$), and border 3 ($p < 0.0001$). The unsigned errors of the algorithm were significantly larger than those between the two ophthalmologists for border 4 ($p = 0.0294$) and border 7



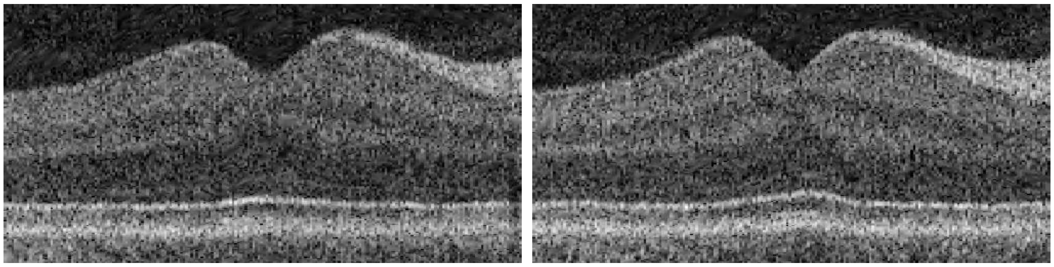
(a) Slice 15

(b) Slice 28



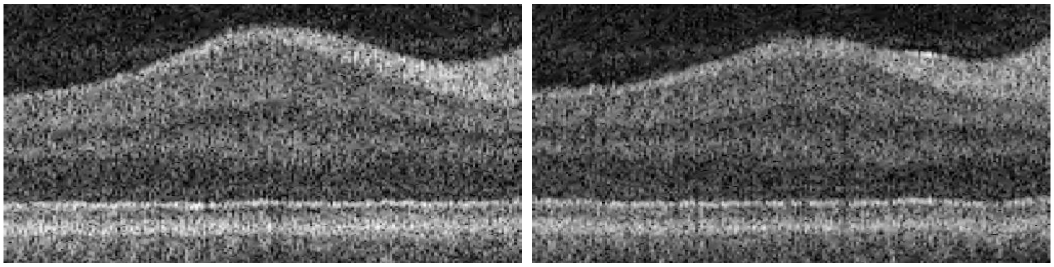
(c) Slice 58

(d) Slice 83



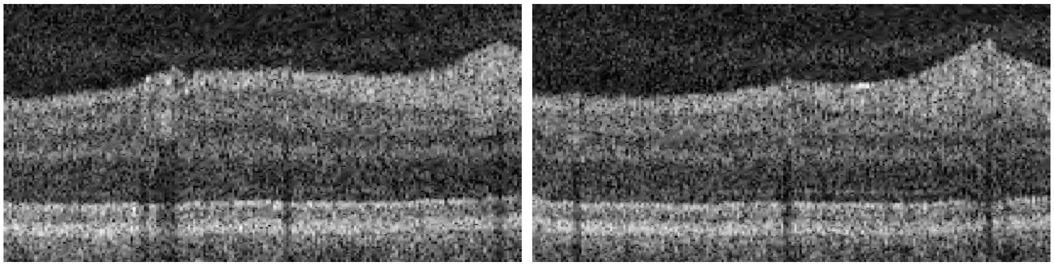
(e) Slice 92

(f) Slice 109



(g) Slice 132

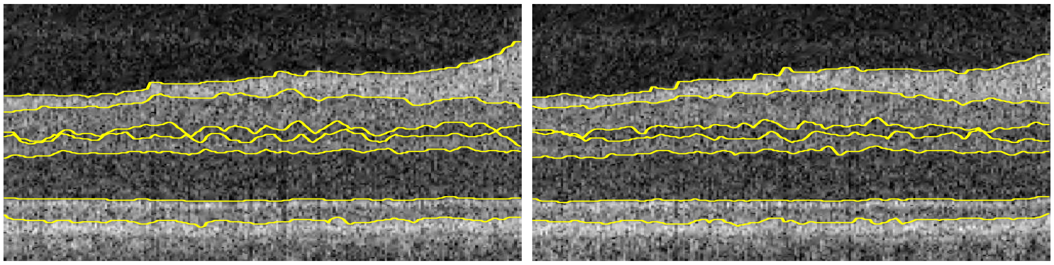
(h) Slice 142



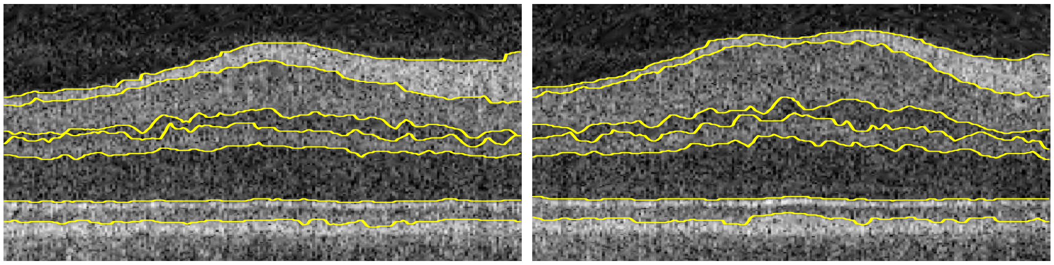
(i) Slice 165

(j) Slice 190

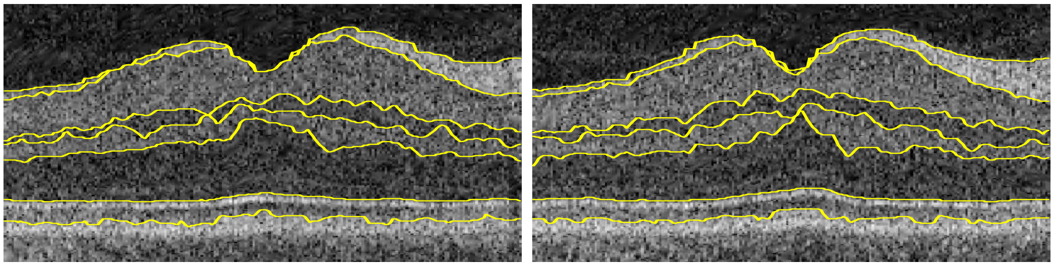
Figure 5.16: The 10 slices used for validation for the OCT volume with the smallest overall mean unsigned border positioning error.



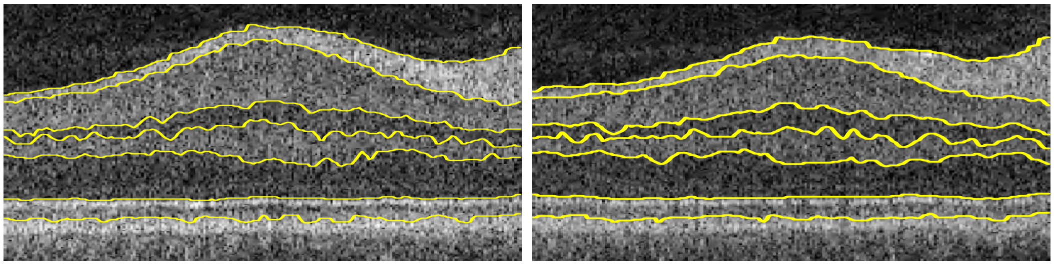
(a) Slice 15 (b) Slice 28



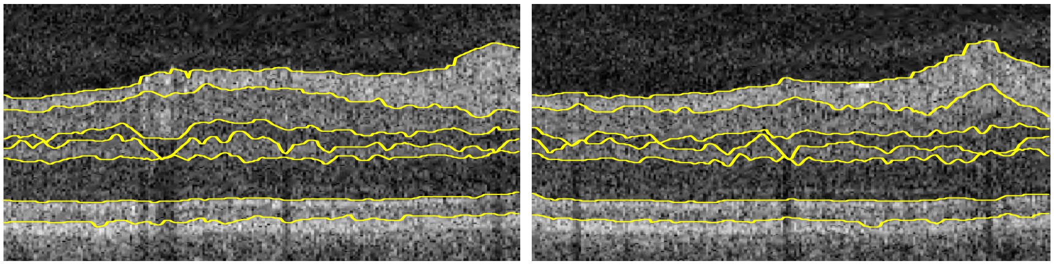
(c) Slice 58 (d) Slice 83



(e) Slice 92 (f) Slice 109

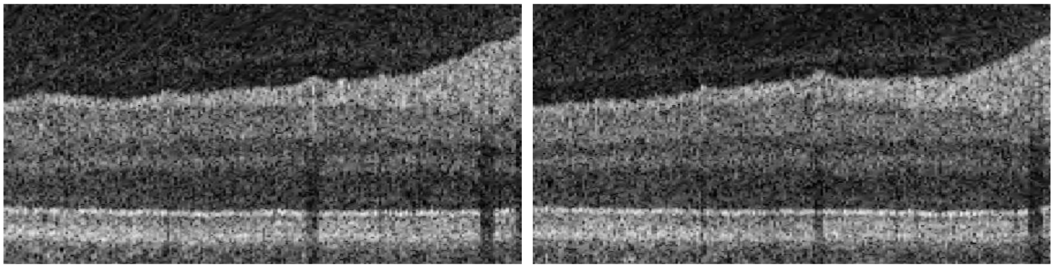


(g) Slice 132 (h) Slice 142

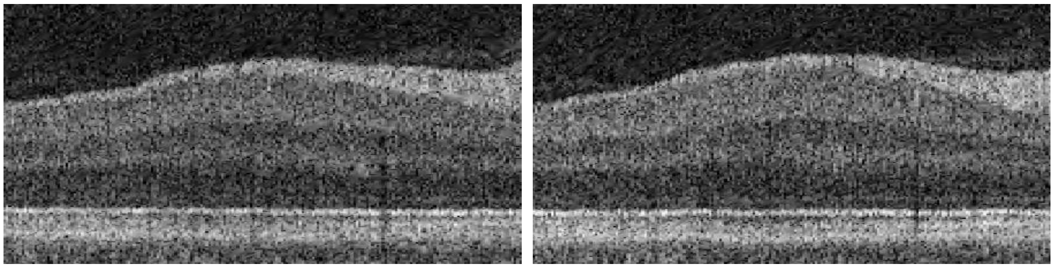


(i) Slice 165 (j) Slice 190

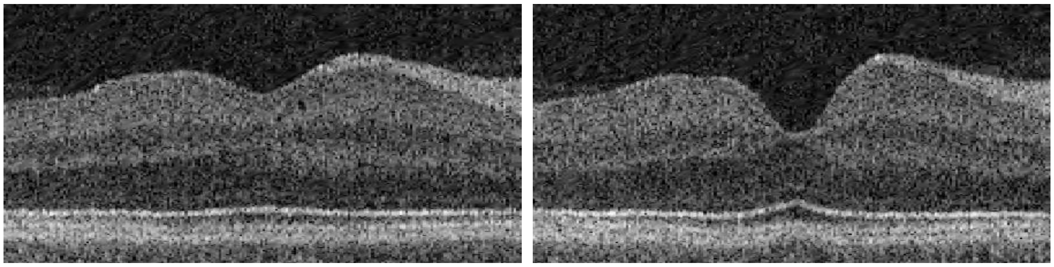
Figure 5.17: Example 7-surface 3-D segmentation results shown on the 10 slices used for validation (best case according to overall mean unsigned positioning error).



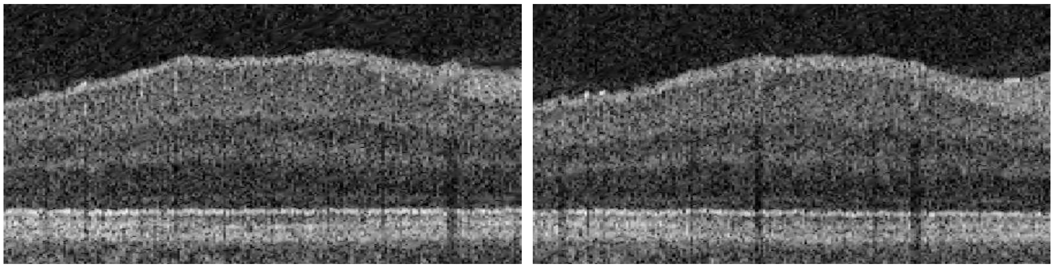
(a) Slice 18 (b) Slice 25



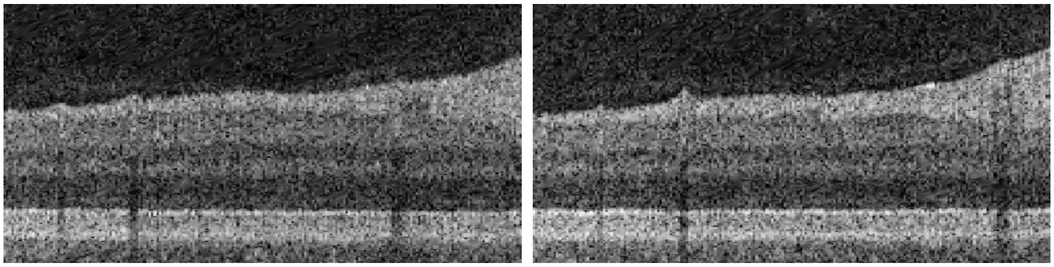
(c) Slice 58 (d) Slice 68



(e) Slice 91 (f) Slice 101

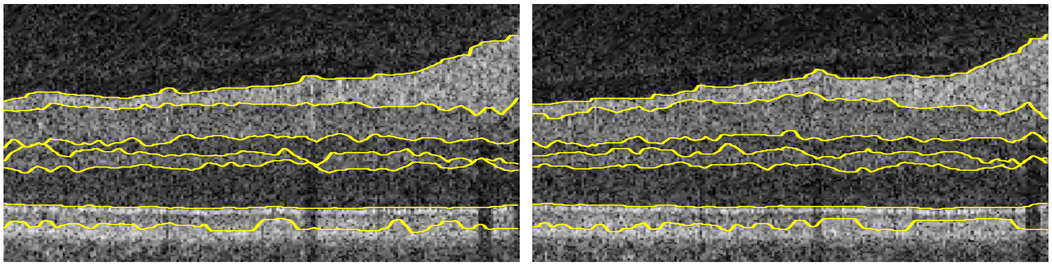


(g) Slice 127 (h) Slice 141

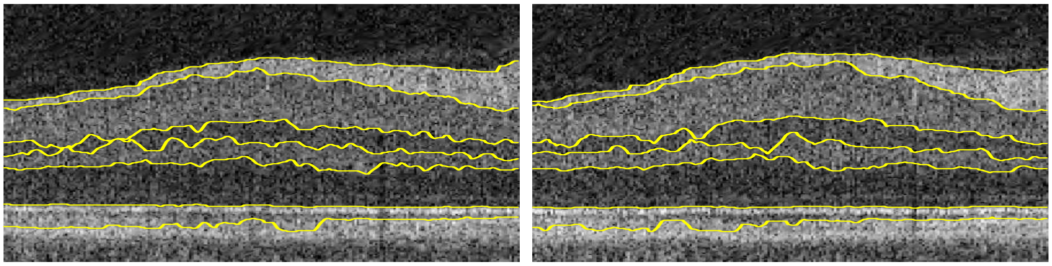


(i) Slice 174 (j) Slice 181

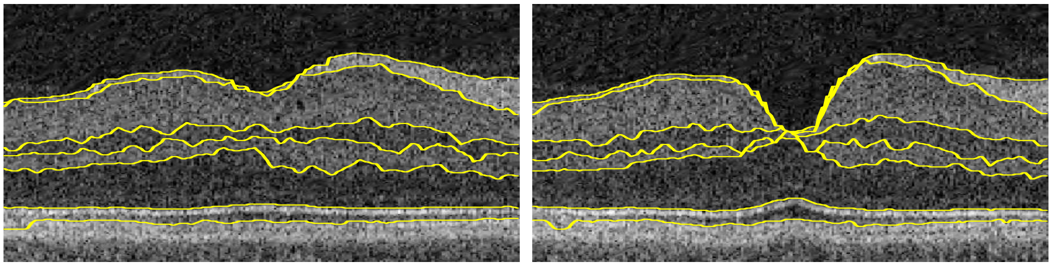
Figure 5.18: The 10 slices used for validation for the OCT volume with the median overall mean unsigned border positioning error.



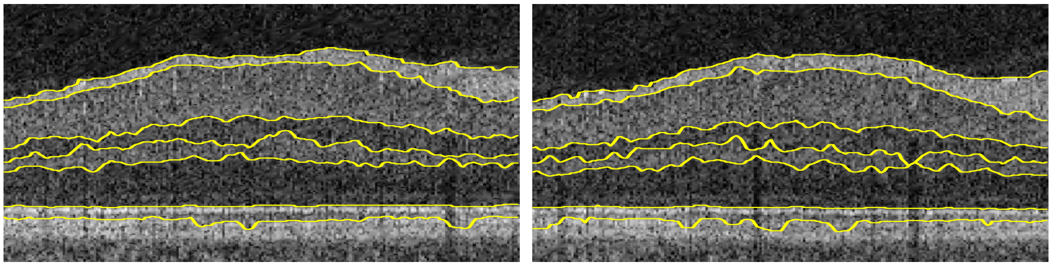
(a) Slice 18 (b) Slice 25



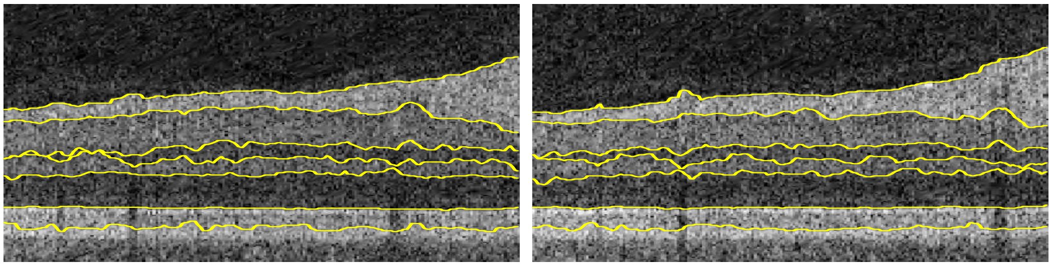
(c) Slice 58 (d) Slice 68



(e) Slice 91 (f) Slice 101

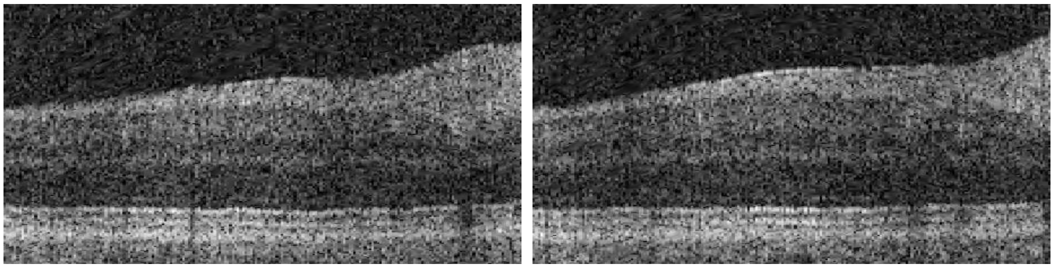


(g) Slice 127 (h) Slice 141

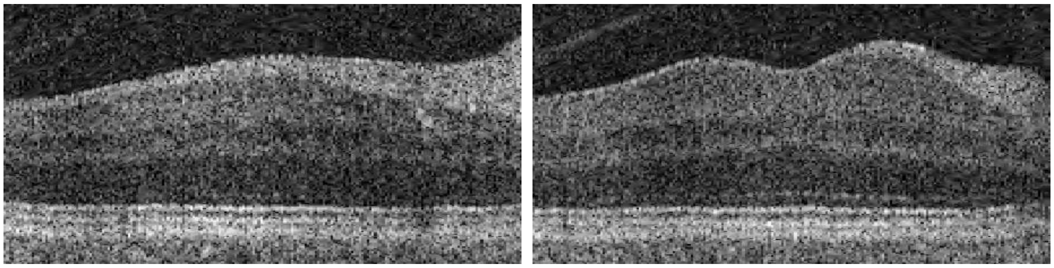


(i) Slice 174 (j) Slice 181

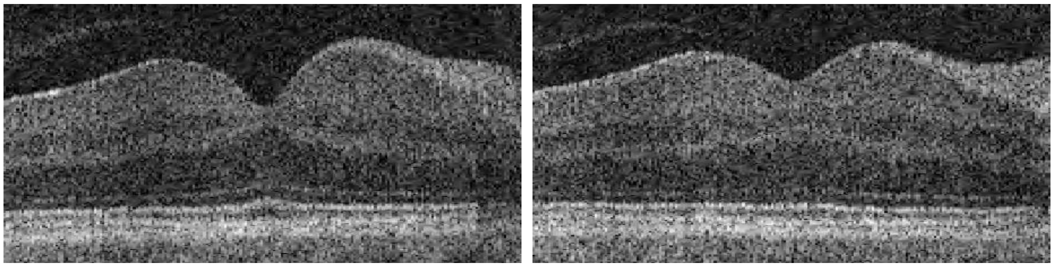
Figure 5.19: Example 7-surface 3-D segmentation results shown on the 10 slices used for validation (median case according to overall mean unsigned positioning error).



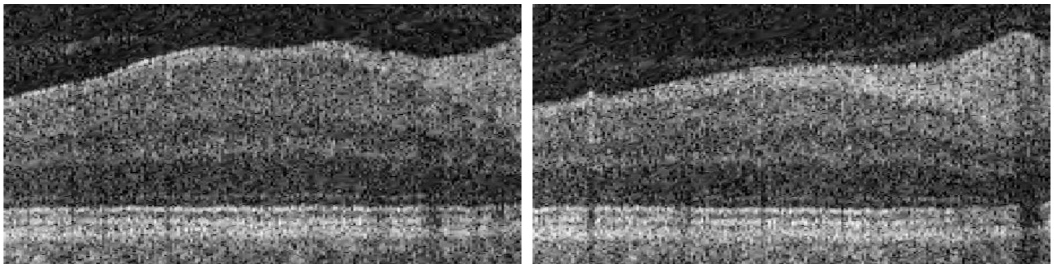
(a) Slice 18 (b) Slice 35



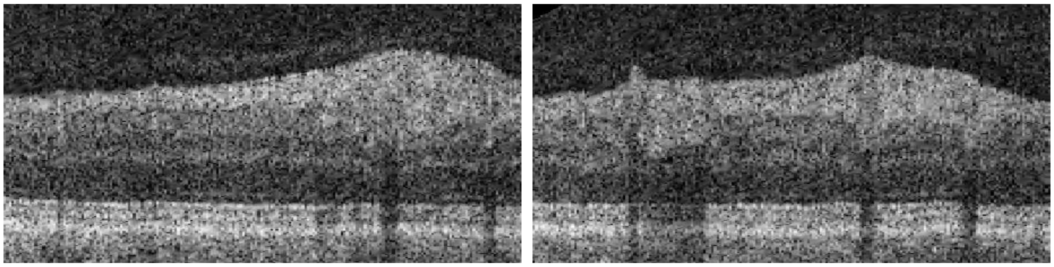
(c) Slice 46 (d) Slice 83



(e) Slice 96 (f) Slice 108

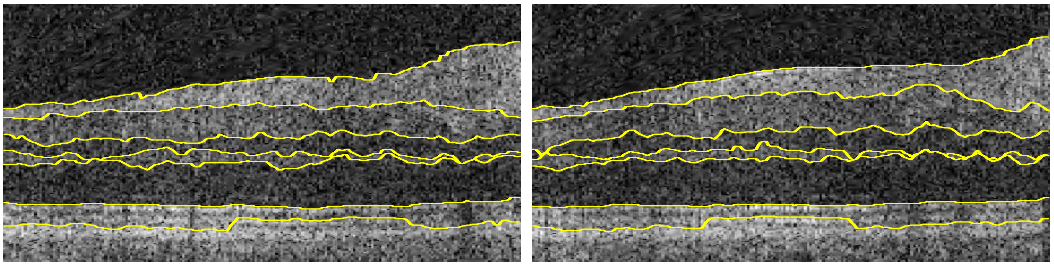


(g) Slice 125 (h) Slice 150

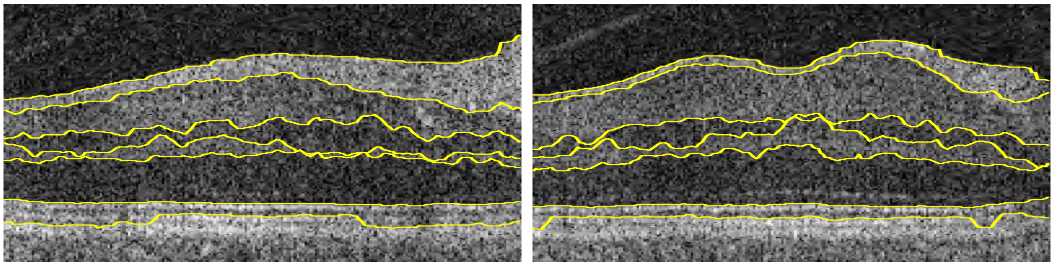


(i) Slice 171 (j) Slice 189

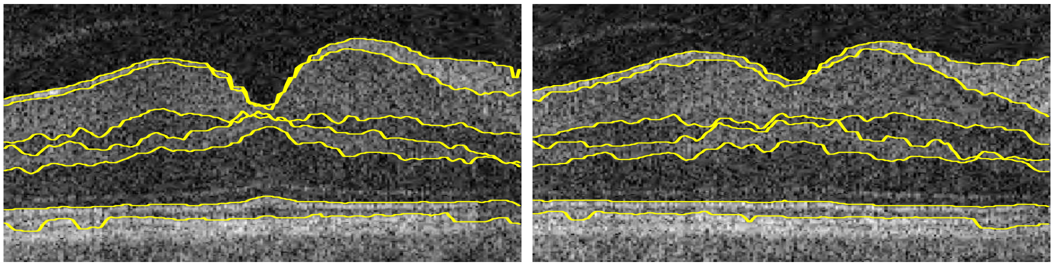
Figure 5.20: The 10 slices used for validation for the OCT volume with the largest overall mean unsigned border positioning error.



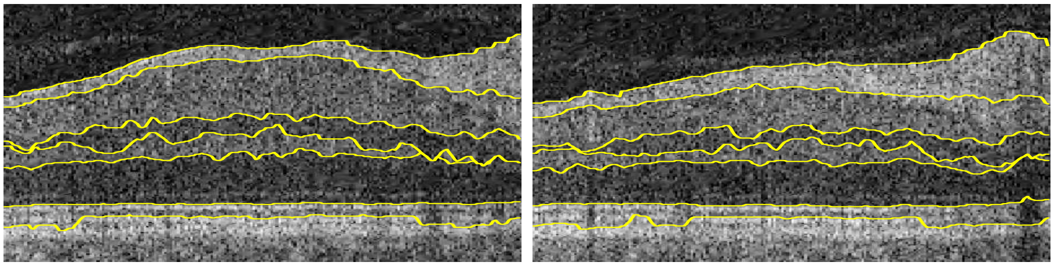
(a) Slice 18 (b) Slice 35



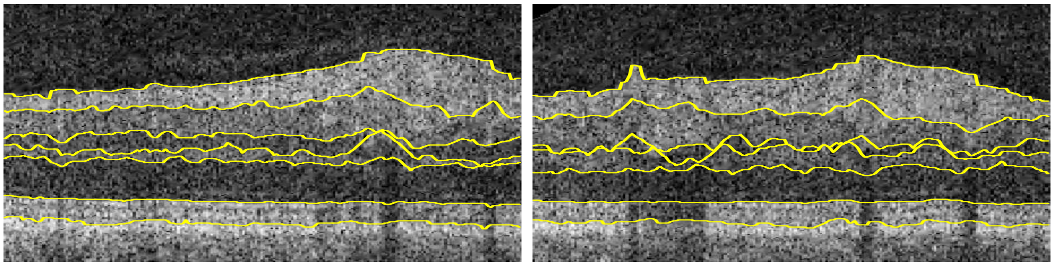
(c) Slice 46 (d) Slice 83



(e) Slice 96 (f) Slice 108



(g) Slice 125 (h) Slice 150



(i) Slice 171 (j) Slice 189

Figure 5.21: Example 7-surface 3-D segmentation results shown on the 10 slices used for validation (worst case according to overall mean unsigned positioning error).

Table 5.3: Summary of mean unsigned border positioning errors† for test set.

Border	Obs. 1 vs. Obs. 2	Algo. vs. Avg. Obs.	Algo. vs. Obs. 1	Algo. vs. Obs. 2
1	3.48 ± 0.41	2.85 ± 0.32	3.70 ± 0.44	2.93 ± 0.34
2	5.49 ± 0.90	4.98 ± 1.24	6.28 ± 1.43	5.09 ± 1.25
3	6.68 ± 1.19	7.25 ± 1.10	8.30 ± 1.42	7.54 ± 1.18
4	7.06 ± 1.41	7.79 ± 0.74	9.59 ± 1.25	7.35 ± 0.54
5	6.16 ± 1.10	5.18 ± 0.82	6.50 ± 1.24	5.46 ± 0.75
6	3.99 ± 2.28	3.30 ± 1.60	4.28 ± 1.93	3.09 ± 2.03
7	7.11 ± 2.07	8.47 ± 2.29	8.48 ± 1.74	9.77 ± 2.76
overall	5.71 ± 1.98	5.69 ± 2.41	6.73 ± 2.48	5.89 ± 2.74

† Mean ± SD in μm.

Table 5.4: Summary of mean signed border positioning errors† for test set.

Border	Obs. 1 vs. Obs. 2	Algo. vs. Avg. Obs.	Algo. vs. Obs. 1	Algo. vs. Obs. 2
1	0.20 ± 1.02	-1.52 ± 0.45	-1.62 ± 0.78	-1.42 ± 0.56
2	-0.34 ± 1.52	-2.04 ± 1.55	-1.87 ± 1.72	-2.21 ± 1.73
3	0.08 ± 2.12	5.02 ± 1.41	4.98 ± 1.78	5.06 ± 1.75
4	3.74 ± 2.47	-3.53 ± 2.18	-5.40 ± 3.07	-1.66 ± 1.76
5	1.88 ± 2.03	-1.19 ± 1.66	-2.14 ± 2.25	-0.25 ± 1.59
6	1.10 ± 2.57	-1.55 ± 1.78	-2.10 ± 2.09	-1.00 ± 2.29
7	4.10 ± 3.75	6.25 ± 3.37	4.20 ± 3.64	8.30 ± 4.07

† Mean ± SD in μm.

($p = 0.0377$). The unsigned border positioning errors for border 3 and border 6 were not significantly different between the algorithm and the two ophthalmologists. Mean signed errors with a magnitude greater than 4 μm (approximately 2 pixels) were present in all comparisons for surface 7, in all comparisons involving the algorithm for surface 3, and in the comparison of the algorithm versus ophthalmologist 1 for surface 4.

Table 5.5: Summary of absolute thickness differences between repeated scans † for test set.

Layer	Algo.	Obs. 1	Obs. 2	Avg. Obs.
1-2	1.13 ± 0.88	1.34 ± 1.08	1.58 ± 1.50	1.07 ± 1.12
2-3	0.62 ± 0.51	2.18 ± 1.13	2.23 ± 1.25	1.26 ± 1.37
3-4	1.47 ± 0.86	1.64 ± 1.29	1.70 ± 1.15	1.20 ± 1.10
4-5	1.73 ± 1.23	1.93 ± 2.47	1.44 ± 1.42	1.38 ± 1.53
5-6	0.88 ± 0.70	3.27 ± 2.81	1.82 ± 2.51	1.94 ± 1.86
6-7	1.89 ± 1.45	2.17 ± 2.32	3.45 ± 3.02	2.46 ± 2.37
overall	1.28 ± 1.06	2.10 ± 2.01	2.04 ± 2.00	1.55 ± 1.65

† Mean ± SD in μm.

5.3.2.3 Thickness reproducibility

A summary of the computed absolute thickness differences between machines for each thickness definition method (based on layer segmentations of the algorithm, ophthalmologist 1, ophthalmologist 2, and those of both ophthalmologists combined) are shown in Table 5.5. The overall absolute thickness difference between the two machines (combining all layers) based on the algorithmic results averaged 1.28 ± 1.06 μm, while the overall thickness differences between the two machines averaged 2.10 ± 2.01 μm based on the manual tracings of ophthalmologist 1, 2.04 ± 2.00 μm based on the manual tracings of ophthalmologist 2, and 1.55 ± 1.65 μm based on the average manual tracings of the two ophthalmologists.

The fitted MANOVA model found a significant difference between the methods ($p < 0.0001$, Wilks' Lambda statistic), so individual (for each layer) ANOVA models were fitted. The adjusted p-values are summarized in Table 5.6. Thus, the algorithm demonstrated significantly smaller (more reproducible) absolute thickness differences between repeated scans for layers 2-3 ($p = 0.0004$) and 5-6 ($p = 0.0063$) when compared to ophthalmologist 1 and significantly smaller absolute thickness differences for layer 2-3 ($p = 0.0002$) when compared to ophthalmologist 2. None of the other comparisons of absolute layer thickness differences were significant.

Table 5.6: Adjusted p-values for fitted ANOVA models[†] of absolute thickness differences between repeated scans.

Layer	Obs. 1 vs. Obs. 2	Algo. vs. Avg. Obs.	Algo. vs. Obs. 1	Algo. vs. Obs. 2
1-2	0.9463	0.9984	0.9058	0.6124
2-3	0.9990	0.2841	0.0004	0.0002
3-4	0.9977	0.8506	0.9519	0.8924
4-5	0.8085	0.9232	0.9835	0.9523
5-6	0.1644	0.4207	0.0063	0.5161
6-7	0.3343	0.8702	0.9801	0.1742

[†] An ANOVA model was fit for each layer.

5.4 Discussion

5.4.1 Accuracy and reproducibility

With respect to accuracy, the unsigned border positioning errors were comparable to that of the two ophthalmologists (overall mean error of 5.69 ± 2.41 μm for the algorithm versus the average of the ophthalmologists and 5.71 ± 1.98 μm for differences between the two ophthalmologists). In addition, the thickness reproducibility between repeated scans was also comparable when comparing the algorithm with the two ophthalmologists.

It is important to note that the time required for each ophthalmologist to manually trace the scans was substantial – a minimum of 30 minutes were required to trace the 10 slices of each volume, but many volumes took longer than 30 minutes to trace. Thus, it would not be feasible to have an ophthalmologist trace all 200 slices in order to perform a thickness assessment (requiring approximately 10 hours of devoted tracing time). In addition, even if an ophthalmologist (or other expert) were willing to trace more slices, it is likely that the accuracy could suffer with the additional slices to trace. Thus, having an automated approach is necessary if quantitative assessments based on the surfaces are desired.

For both the algorithm and the ophthalmologist, the position of surface 7 was sometimes difficult to define. This was because two boundaries (interfaces) were sometimes visible, both of which could be feasibility defined as the correct surface and it was not clear which boundary was the appropriate one. Thus, the algorithm would sometimes find a surface that would transition between both boundaries within each slice; whereas each ophthalmologist would stick to one boundary for a particular slice, but would sometimes pick the other boundary on the repeated scan (or in other slices within the volume). This double-boundary was not noticed on the time-domain scans for which the cost function was based (and it wasn't present in all slices of the spectral scans). Thus, modifying the cost function for this surface would be something to explore. In addition, it would be useful to obtain a better understanding of these surfaces through histological comparisons.

5.4.2 Reducing memory and time requirements

The large amount of data in spectral OCT datasets make the time and memory requirements more of an issue when compared to the time-domain case. In fact, the methods had to be implemented to run on a 64-bit operating system so that up to approximately 10 GB of RAM could be used during the segmentation process. Furthermore, the time requirements were substantial (requiring hours of processing time per volume). However, there are multiple ways the memory and time requirements could be reduced.

For example, one way of reducing the required processing time would be to try a different *s*-*t* cut approach in the graph search (in this chapter we used an approach based on that developed by Boykov [35]). In fact, we have performed some preliminary experiments that indicate that substantial amounts of time could be saved by just changing this approach (e.g., for a case with a small number of slices, a reduction in running time from 2.5 hours to 5 minutes was achieved by changing from Boykov's method to a push-relabel method). Using a parallel implementation would be another way of speeding things up.

In addition, there were a number of redundancies in the approach. Better utilization of some of the redundancies would be another way of making the approach faster. For example, after the preliminary segmentation of surfaces 1, 6, and 7 was performed on the downsampled volume, this information was only used for flattening the image, but not used in the later segmentation steps. It may make sense to reuse this preliminary segmentation in order to reduce the size of the graph required in later steps (e.g., by defining a region of interest based on the preliminary result for the full-resolution segmentation). Interestingly, it was the segmentation of the first set of surfaces (1, 6, and 7) that took the most time (the segmentation of the interior surfaces took minutes instead of hours), so utilizing the preliminary segmentation would be expected to provide a substantial savings. Furthermore, reducing the size of the graph in this way would save on memory requirements as well. It may also be wise to try a more general multiresolution approach in order to save on time and memory requirements (e.g., by performing the segmentation on more than just the two resolutions used in this work). Thus, even though the current implementation was relatively slow, our preliminary experience indicates that with some of the above modifications, the segmentation time could reasonably be expected to run in minutes instead of hours.

5.4.3 Applicability to optic nerve head 3-D OCT segmentation

Although the work presented in this chapter has focused on segmenting macular images, many of the developed methods are also applicable to segmenting spectral OCT images centered at the optic nerve head as well. The typical region of the fundus covered by a optic nerve head scan is illustrated in Fig. 5.22, with example slices from one volume provided in Fig. 5.23. In addition to segmenting the layers in the peripapillary region surrounding the optic disc, one can also segment structures of the optic nerve head on these scans.

The presence of the optic nerve head does present a different set of challenges

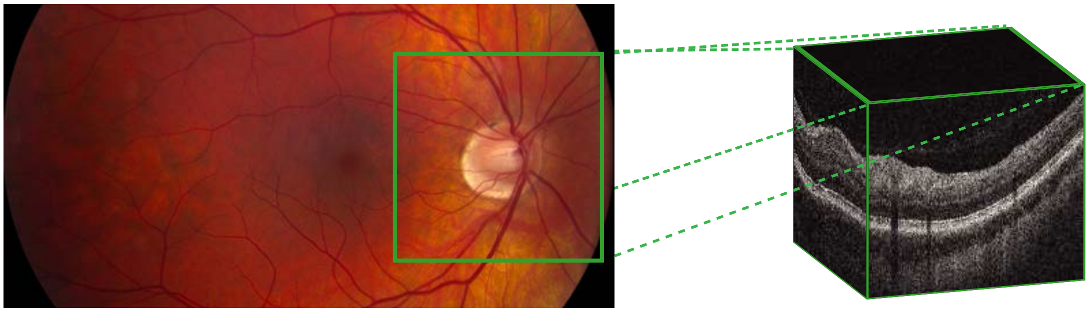
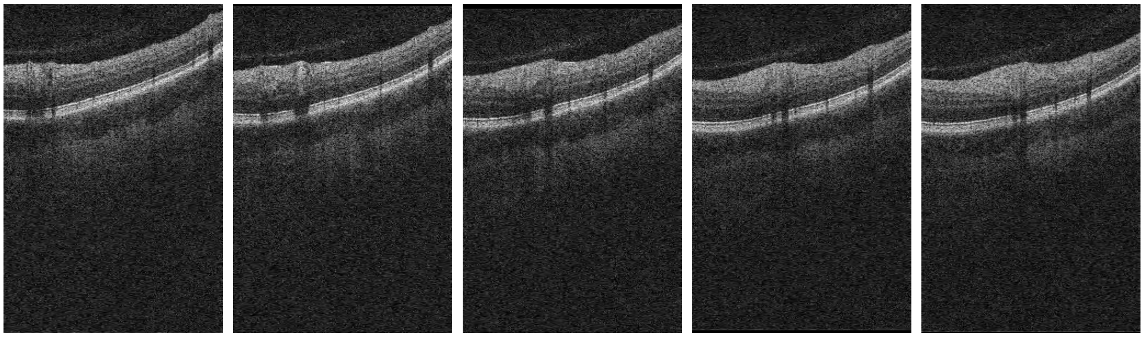


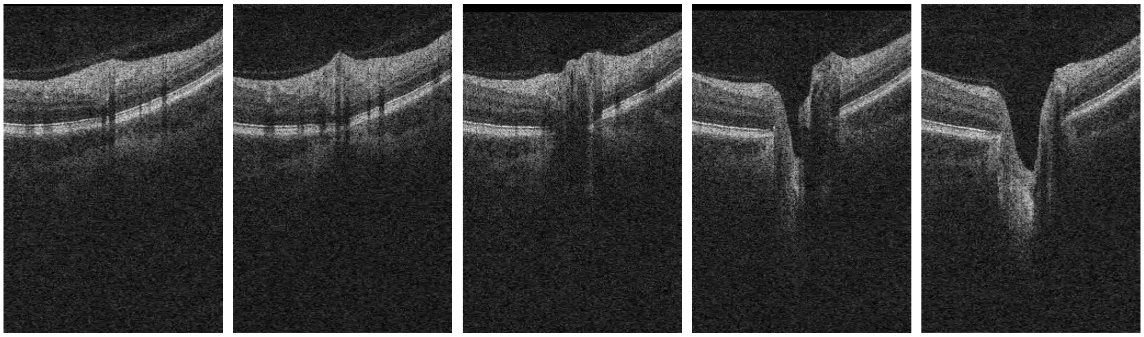
Figure 5.22: Fundus photograph with schematic OCT volume (size 6 × 6 × 2 mm^3) centered at optic nerve head.

as the layers “disappear” in this region. Although there are multiple ways to approach segmenting these images, here we present one possible solution as illustrated in Fig. 5.24. Like in the macular case, the first major step is to flatten the image. The method is the same as in the macular case, with the exception that a “don’t care” region is specified based on an estimate of the optic disc region when segmenting the first three surfaces on the downsampled volume. Note that this estimate does not have to be accurate (i.e., if the images are centered properly, a fixed-size rectangle or circular region would work fine) because the goal is only to obtain a set of points for fitting the thin-plate spline (points inside of the “don’t care” region will not be used). Alternatively, one could obtain this estimate from the segmentation of a projection image. Based on the segmentation points, a thin-plate spline (with regularization) is fit to the points found on surface 7 (RPE) outside of the optic disc region. The flattening of this surface then defines a reference “RPE plane.” Because the thin-plate spline is defined across the entire volume, this reference plane also exists in the optic nerve head region, even though the RPE is not present there.

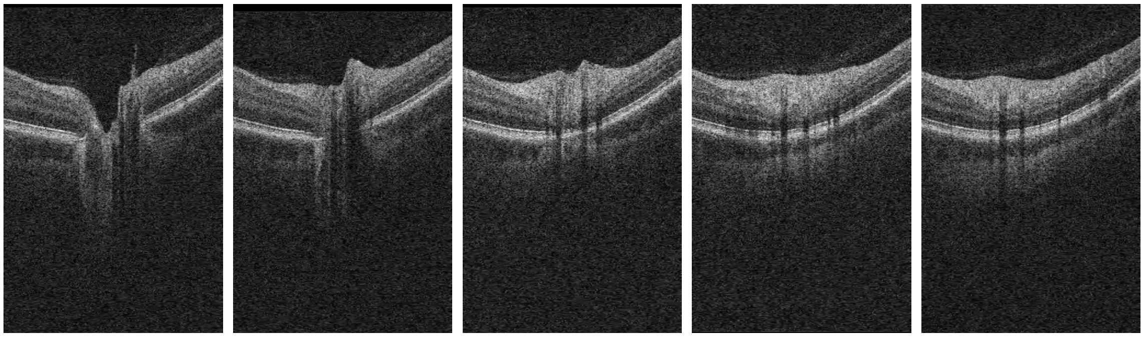
Once the image has been flattening, the next step would be to more accurately estimate the optic disc and cup margins at the level of the RPE plane. One way of doing this would be to perform the segmentation on a 2-D projection image at the level of the RPE. The projection image can be obtained by computing the mean



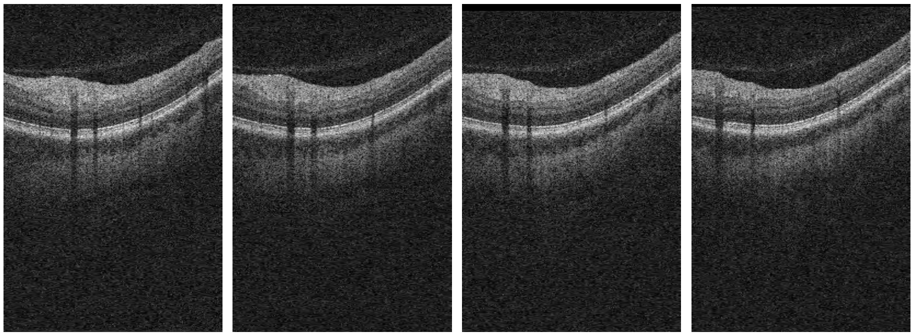
(a) Slice 0 (b) Slice 11 (c) Slice 22 (d) Slice 33 (e) Slice 44



(f) Slice 55 (g) Slice 66 (h) Slice 77 (i) Slice 88 (j) Slice 99



(k) Slice 110 (l) Slice 121 (m) Slice 132 (n) Slice 143 (o) Slice 154



(p) Slice 165 (q) Slice 176 (r) Slice 187 (s) Slice 198

Figure 5.23: Example slices from a spectral-domain OCT volume centered at the optic nerve head.

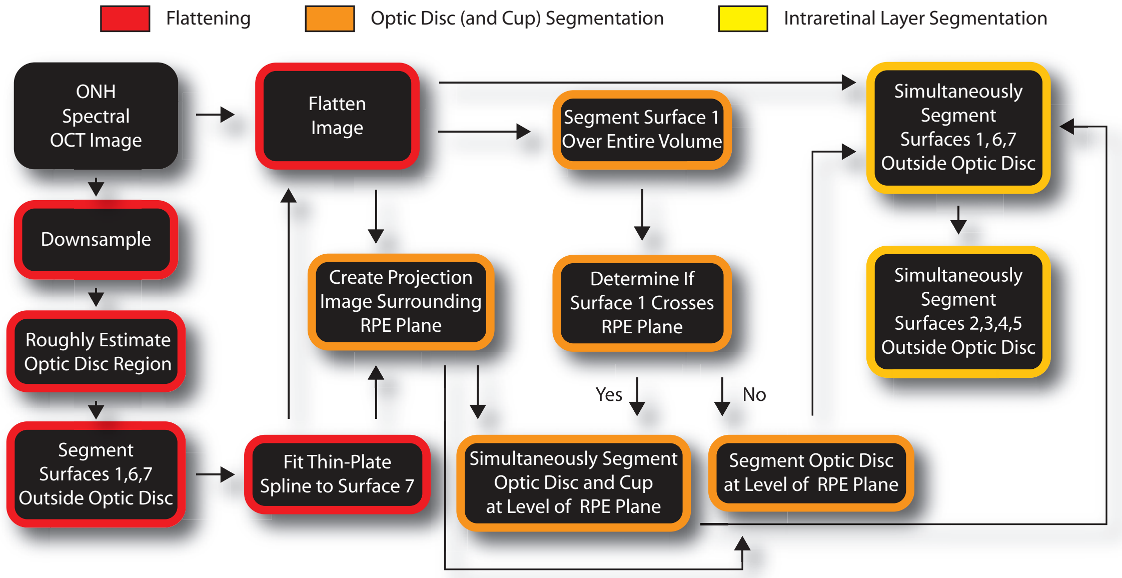


Figure 5.24: A possible set of steps for segmenting ONH spectral OCT images.

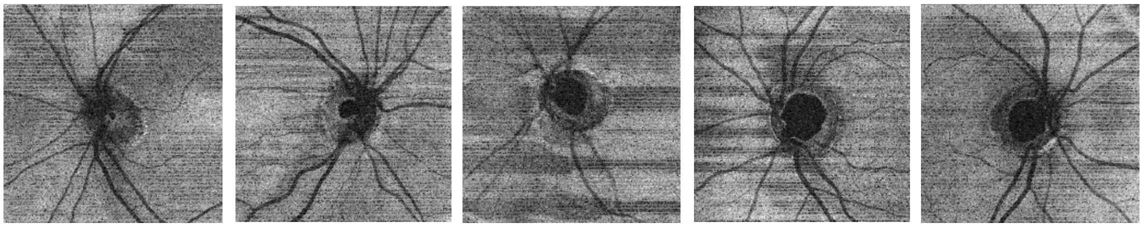
projection of intensity values from a small number of slices surrounding the RPE plane. If the cup margin crosses this plane, both the disc and the cup margin can be segmented on the projection image; otherwise only the estimate optic disc margin can be segmented. Note that the optic disc margin would only be an estimate as we would actually be segmenting the neural canal opening. Because the cup margin can also be defined based on the location of the first surface (ILM), the choice of whether to find two boundaries or one could be made based on whether the first surface crosses the plane. Thus, before segmenting the projection image, one might consider also finding an estimate of the first surface for purposes of determining its maximum depth; however, one could also try both options and then pick the best one (based on the cost function or other metrics). One way of segmenting the boundaries in the projection images would be to use an optimal graph search approach on an "unwrapped" image in polar coordinates.

As a feasibility study for this approach, we flattened spectral volumes and segmented the RPE-plane projection images of 10 eyes (from 9 subjects). A signed edge-based term (favoring a dark-to-bright transition in the radial direction) was used as the cost function for both surfaces and they were determined simultaneously using a 2-D version of the graph search. In one case, the cup did not cross the RPE plane, and so we used a single-boundary approach for that case. Fig. 5.25 illustrates all of the results. While the results are clearly not perfect, they are very promising. One way of improving them might be to take into account the vessel locations (see next section for segmentation approach). In addition, note that future work would be necessary for determining how closely this estimate of the optic disc margin corresponds to the optic disc observed on fundus photographs.

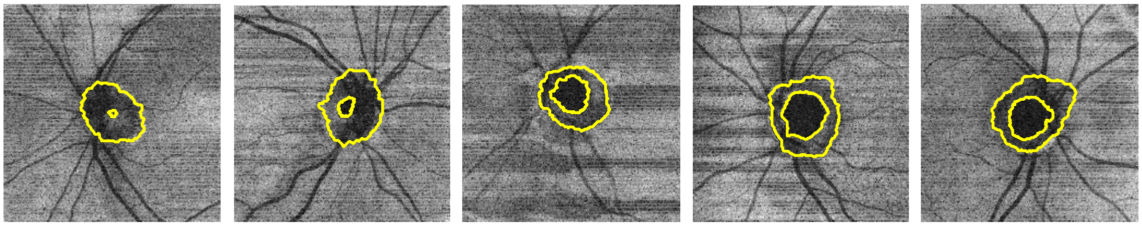
Once a better estimate of the optic disc has been determined, the layers can be segmented in the peripapillary region by setting the region inside of the optic disc as being a “don’t care” region in the segmentation approach. In this way, the segmentation can be performed similarly as in the macular case. Of course, a different set of constraints would likely be needed. In addition, more work may need to be done to decide if changes in the cost functions would be necessary. Fig. 5.26 illustrates an example of the resulting outputs of the overall approach.

5.4.4 Applicability to vessel segmentation

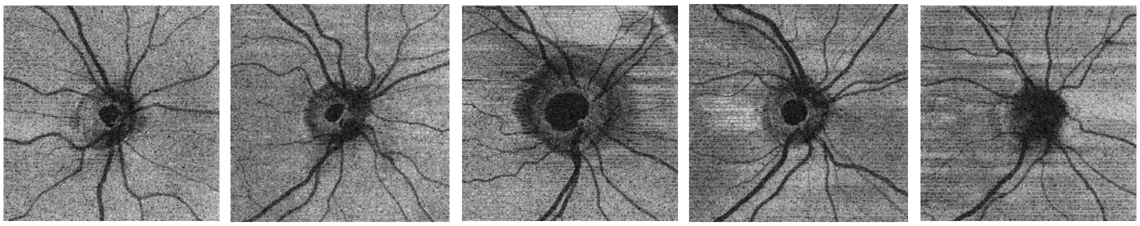
In addition to allowing the estimation of the cup and rim margins at the level of the RPE plane (see section above), utilizing a “smart” projection method can also allow for better vessel segmentations. For example, based on the segmentation approach developed in this chapter, in the work by Niemeijer and Garvin *et al.* [46], projection images were formed by identifying an OCT layer in which the vascular shadows had the highest contrast. This was based on the observation that the absorption of light by the blood causes shadows to appear below the position of the vessels (see Fig. 5.27(a)). The layer was delineated by surface 6 and an additional surface 8 that was specially segmented for this work. This layer was chosen to provide a better contrast



(a) Subj. P1, OS (b) Subj. P2, OD (c) Subj. P3, OS (d) Subj. P4, OS (e) Subj. P4, OD



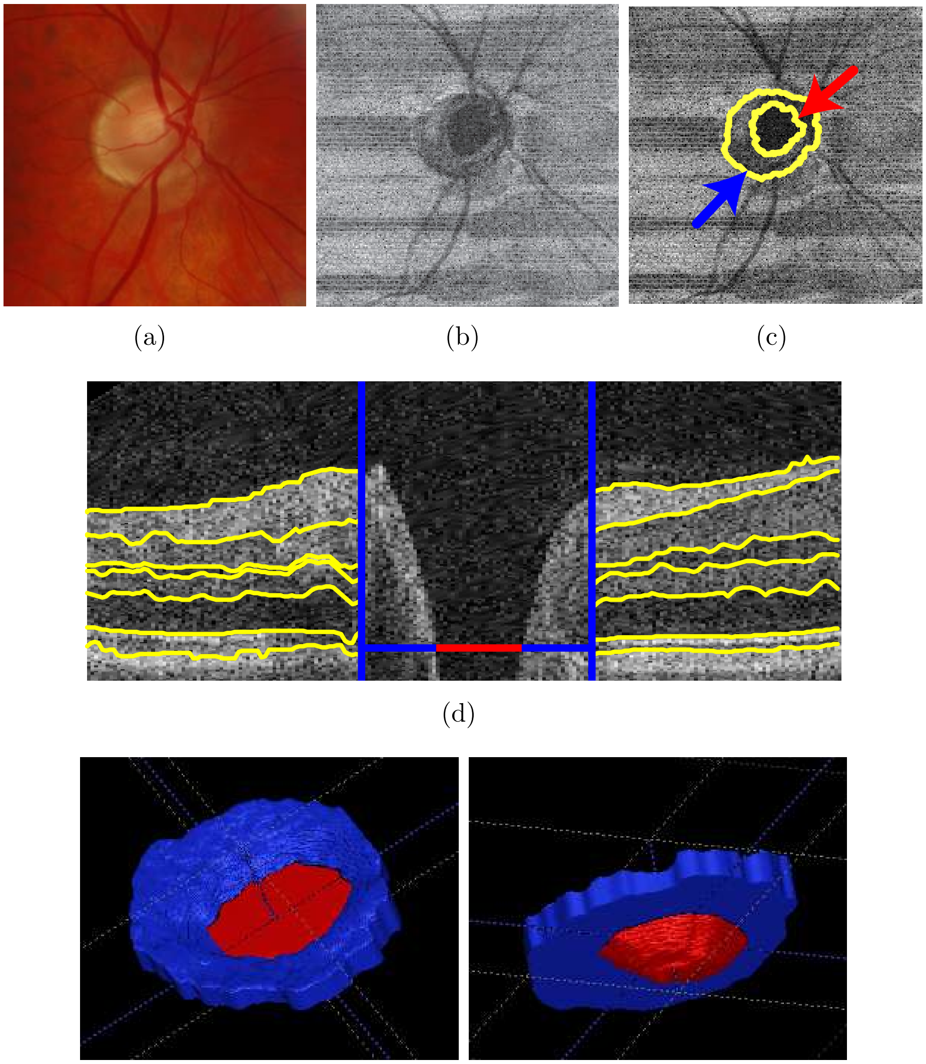
(f) Subj. P1, OS (g) Subj. P2, OD (h) Subj. P3, OS (i) Subj. P4, OS (j) Subj. P4, OD



(k) Subj. P5, OD (l) Subj. P6, OD (m) Subj. P7, OD (n) Subj. P8, OD (o) Subj. P9, OD

(p) Subj. P5, OD (q) Subj. P6, OD (r) Subj. P7, OD (s) Subj. P8, OD (t) Subj. P9, OD

Figure 5.25: Example segmentation of estimated optic disc and cup boundaries at the level of the RPE plane.

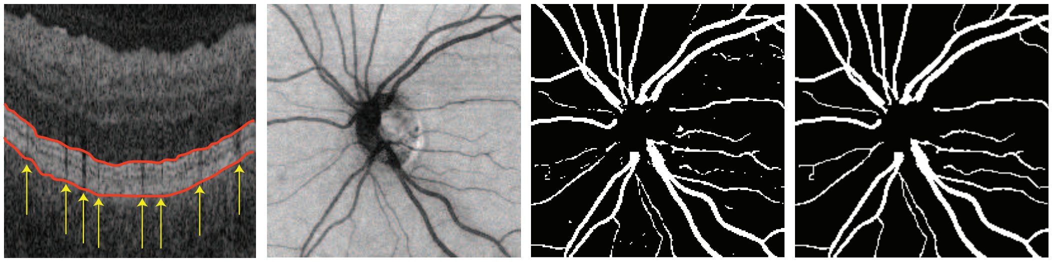


(a) (b) (c)

(d)

(e) (f)

Figure 5.26: Segmentation of optic nerve head landmarks and retinal layers. (a) Fundus photo of the optic disc. (b) En-face projection of OCT image. (c) 3-D segmentation (yellow lines) of RPE opening (blue arrow) and cup margin (red arrow). (d) 3-D segmentation of layers, RPE reference plane (red), RPE opening (blue). (e) Volumetric rendering of pre-RPE rim and post-RPE cup indices (seen from anterior). (f) Volumetric rendering of pre-RPE rim and post-RPE cup indices (seen from posterior).



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

Figure 5.27: Example of spectral 3-D OCT vessel segmentation. (a) The vessel shadows indicate the position of the vasculature. Also indicated in red are the slice intersections of the two surfaces that delineate the subvolume in which vessels are segmented. (b) 2-D projection image extracted from the projected subvolume of the spectral 3-D OCT volume. (c) Automatic vessel segmentation. (d) Vessel segmentation after postprocessing – removing disconnected pieces and connecting large segments.

in the projection image over that provided by using a naive projection method which utilized all voxels in the image for creating the projection. Next, a classification approach (based on kNN classification with k=31 and using Gaussian filter derivatives as features) was used to label all pixels as vessel or non-vessel (Fig. 5.27(c–d)). When training the method on a set of 5 3-D scans and testing on a different set of 10 optic nerve centered 3-D OCT scans, we successfully identified the retinal vessels with high sensitivity and specificity as determined using Receiver Operating Characteristics (ROC) analysis (area under the curve of 0.97). Furthermore, as illustrated in Fig. 5.28, better results were achieved when using the smart projection method than with the naive projection method (naive method had an area under the curve of 0.939). This work has also been extended to deal with segmenting the vessels on macular OCT scans [47].

Having a vessel segmentation opens the door to an array of additional image analysis possibilities. For example, one might consider using the vessels to appropriately modify the cost functions and/or smoothness constraints for better layer segmentations. This is because of the different appearance of OCT columns in which a vessel

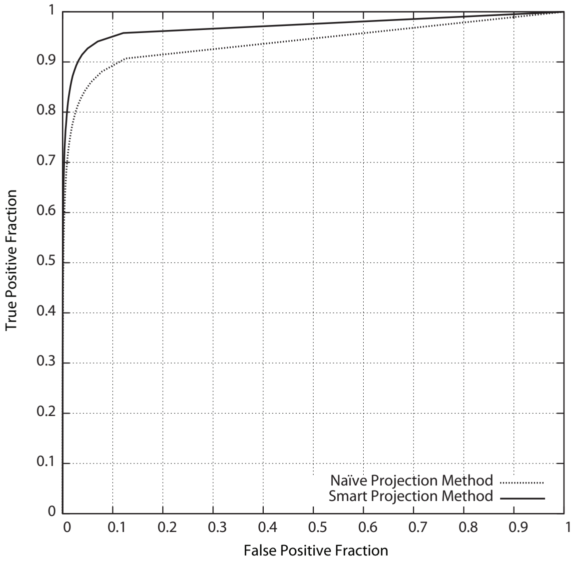


Figure 5.28: The resulting ROC curves of the vessel segmentation results when using naive and smart projection methods [46]. The area under the curve is 0.939 and 0.970 respectively.

is present (e.g., the presence of shadows in the outer layers). In addition, one might consider “removing” the vessel regions when computing global layer properties such as thickness. Additional uses for the vessel segmentation (e.g., registration with fundus photographs) will be discussed in Chapter 7.

CHAPTER 6
MACULAR INTRARETINAL LAYER THICKNESS ASSESSMENT

6.1 Assessment in AION subjects using time-domain OCT

6.1.1 Motivation

In optic neuropathies, time-domain peripapillary OCT scans are commonly used clinically to measure nerve fiber loss surrounding the optic disc. However, especially in acute optic neuropathies (such as in acute anterior ischemic optic neuropathy), swelling may occur in this region, thus potentially masking nerve fiber loss until the swelling has regressed. In addition, it has been shown that in optic neuropathies such as glaucoma and AION, the thickness of the RNFL cannot decrease more than a certain amount, beyond which little change can be measured even though the visual field sensitivity becomes further reduced [48–50]. Furthermore, with increasing damage, the visual field variability increases substantially, making such functional measurements less reliable [51–54]. Thus, there is need for alternative methods for measuring damage with a larger dynamic range and increased reproducibility.

Measuring thickness changes in the macula may offer advantages in this regard because of its larger distance to a (potentially swollen) optic disc, the larger number of ganglion cells per retinal surface area, and the corresponding potential for a larger dynamic range with ganglion cell loss. Also, determining structure-function relationships may be easier in the macula because visual field testing results apply directly to the macular region “under” it. The retinal nerve fiber and ganglion cell layer, together with the inner plexiform layer and nuclear layer, make up the inner retina, the thickness of which it or its sublayers is of the greatest interest. However, current commercial scanners only provide thickness measurements for the total retina in the macula.

As a preliminary study in demonstrating the value of using an intraretinal segmentation approach, in this section, we compare inner and outer retinal thickness

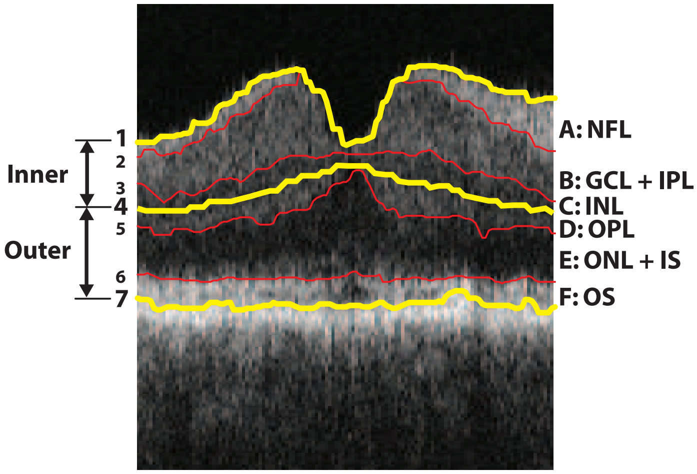


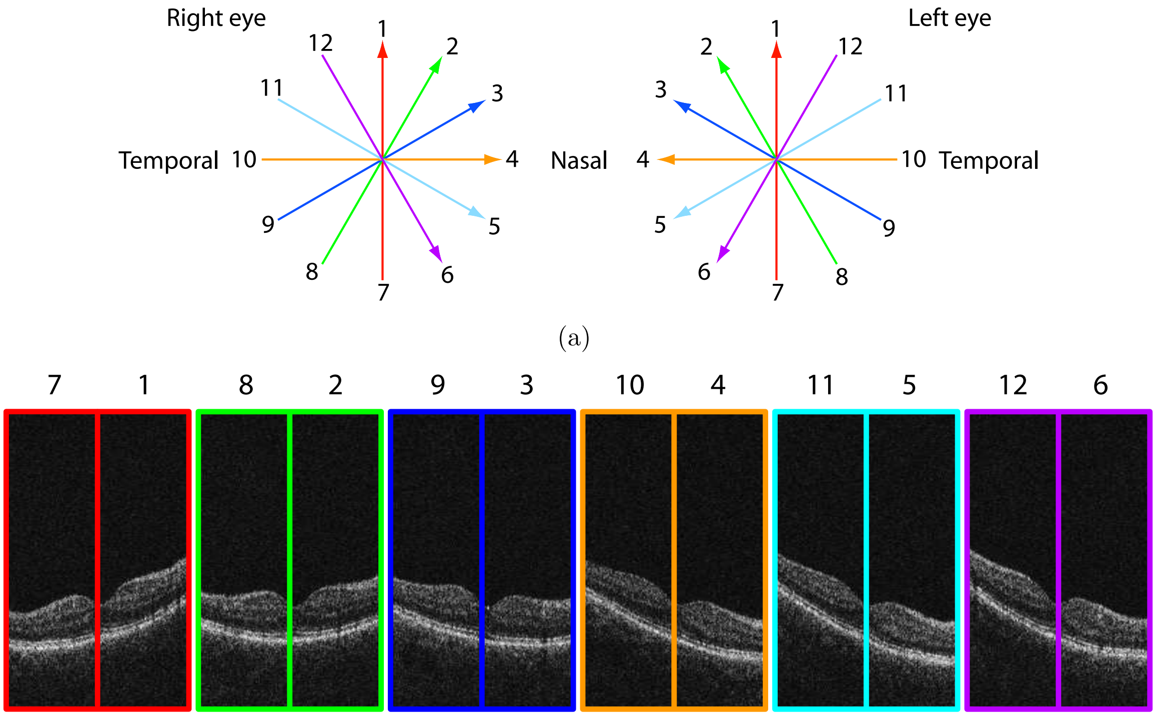
Figure 6.1: Illustration of inner and outer retinal layers on time-domain composite OCT slice. Surface 1 to surface 4 delineated the inner layer, while surface 4 to surface 7 delineated the outer layer.

values of the affected and unaffected eyes of patients with unilateral AION.

6.1.2 Methods

Using our intraretinal layer segmentation results from variation 3 of the time-domain segmentation approach described in Chapter 4 and presented at MIICAI 2007 [37] in 12 patients with unilateral anterior ischemic optic neuropathy, we computed the thickness of the inner retinal layer containing the ganglion cells and their axons (from surface 1 to surface 4 in Fig. 6.1: NFL + GCL + IPL + INL) and the thickness of the remaining outer layers (from surface 4 to surface 7 in Fig. 6.1). We also computed the total retinal thickness. Because the damage in the affected eye often only occurs in a limited region, the thicknesses were computed both over the entire macular region and in selected regions (but excluding the fovea in both cases).

More precisely, for the segmental analysis, each OCT scan was first divided into 12 segments at 30-degree increments, as illustrated in Fig. 6.2. This was done by dividing each of the radial slices into two segments (one for the right half of the scan and one for the left half of the scan as shown in Fig. 6.2(b)). Next, based on



(a)

(b)

Figure 6.2: Illustration of 12 segments shown on OCT data. Note that the OCT data is shown from a clinician's perspective. (a) Division of six scan lines into 12 segments. (b) Twelve segments indicated on OCT data.

visual field sensitivity data (and blinded to the segmentation results and thickness values), an ophthalmologist chose two segments for which he noted a visual field defect. Fig. 6.3 illustrates the corresponding segments on the visual field data from one subject. Note that visual field data is shown from the patient's perspective, while OCT data is typically presented from the clinician's perspective (thus, the nasal-temporal designations are on opposite sides). Also remember that because of the optics of the eye, the superior (top) part of the visual field corresponds to inferior OCT segments and the temporal visual field corresponds to nasal OCT segments. For 10 of the 12 patients, 24-2 visual field threshold testing results were available. For the other two patients, one had 10-2 visual field threshold testing results available and the other only had the results from a kinetic perimetry test.

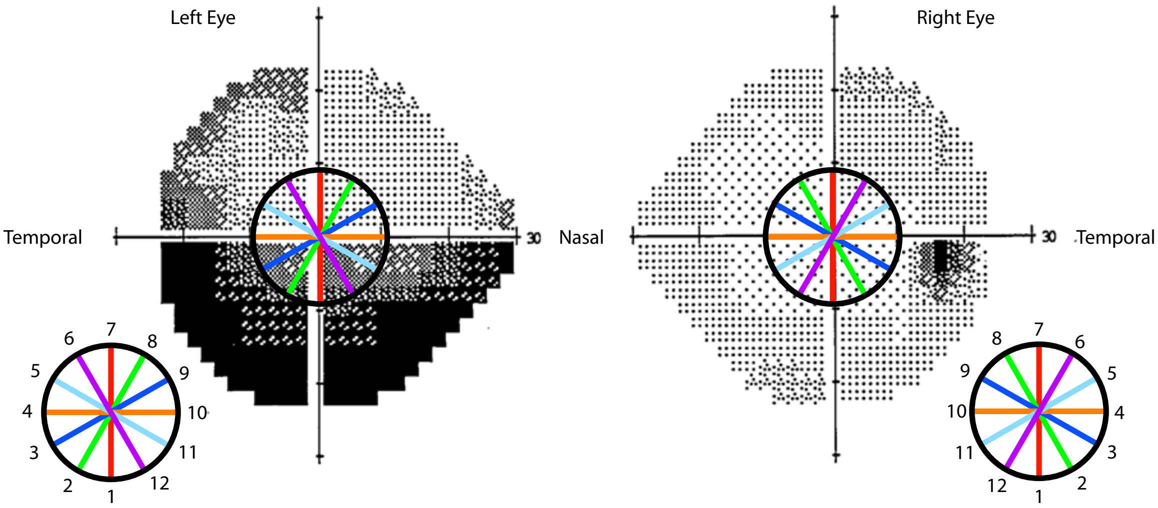


Figure 6.3: Illustration of 12 segments shown on visual field data from one unilateral AION subject. In general, darker regions show areas of visual field defects. In this example, the left eye was the affected AION eye. Note that the visual field data is shown from the patient's perspective.

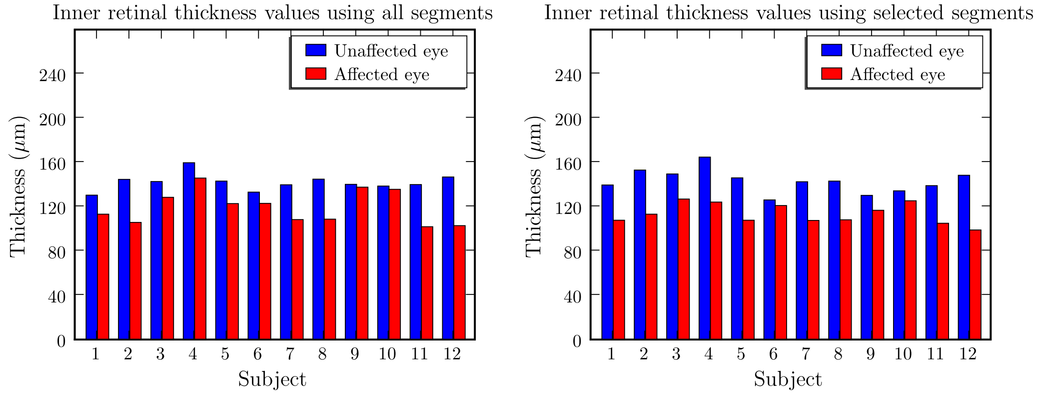
For each thickness approach (overall and segmental), thickness differences between the unaffected and affected eyes were compared using a paired t-test. In addition, for each (r, θ) scan location (using a right-eye orientation from a clinician's perspective), the thickness value was averaged over all the unaffected eyes to give an estimate of an undiseased thickness value for that location. Next, for each diseased eye, the differences between the averaged undiseased values and the diseased-eye thickness values were computed (using a right-eye orientation), and displayed as an interpolated color-coded map (in the orientation of the affected eye). Each difference was computed as the averaged undiseased value *minus* the diseased value so that positive values indicated thinning (and were displayed as warmer colors). This allowed for a localized visual comparison between the visual field data and the thickness difference results.

6.1.3 Results

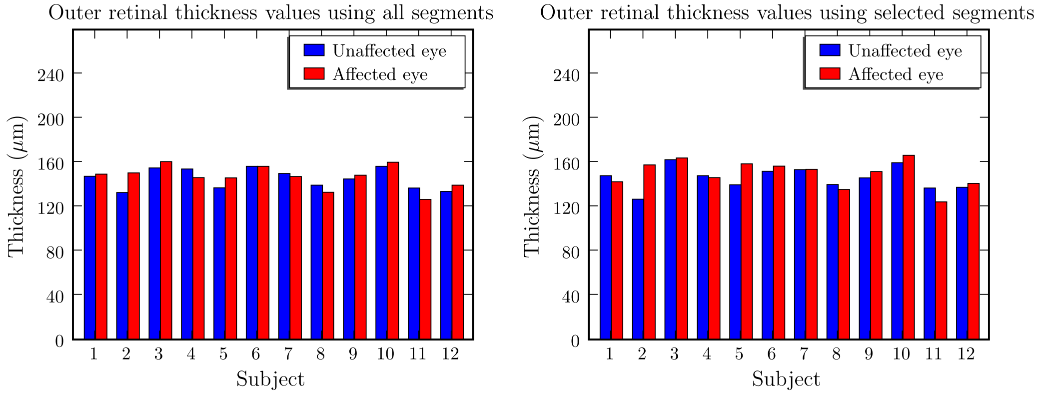
A summary of the resulting thickness values for the affected and unaffected eyes is shown in Fig. 6.4. When considering all segments (not just the selected ones), the inner retinal layer thickness for the affected eye was 22.5 μm smaller on average than for the unaffected eye (or about 16%, statistically significant using paired t-test, $p < 0.001$), while the outer retinal thickness was 1.7 μm larger on average (not statistically significant, $p = 0.48$). The total retinal thickness for the affected eye over all segments was 20.8 μm smaller on average than for the unaffected eye (statistically significant, $p = 0.001$).

In the selected segments, the inner retinal layer thickness for the affected eye was 29.5 μm smaller on average than for the unaffected eye (or about 21%, statistically significant using paired t-test, $p < 0.001$), while the outer retinal thickness was 4.0 μm larger on average (not statistically significant, $p = 0.25$). The total retinal thickness for the affected eye in the selected segments was 25.5 μm smaller on average than for the unaffected eye (statistically significant, $p < 0.001$).

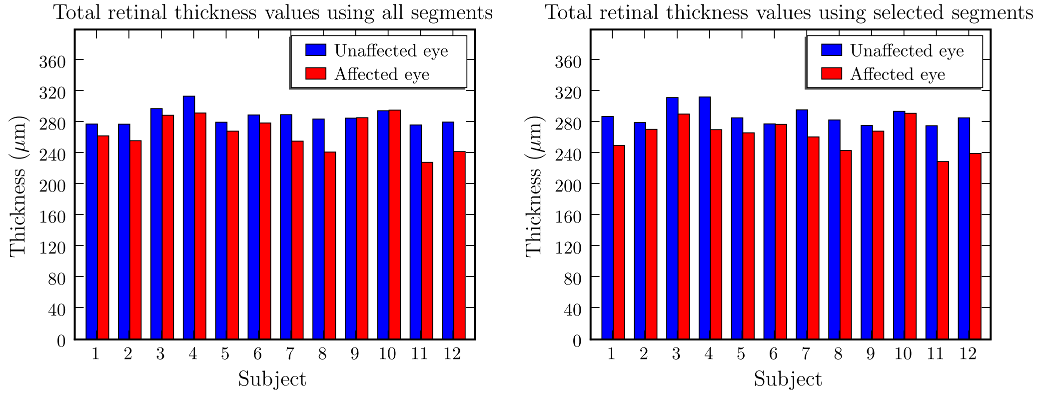
The visual field and inner/outer thickness difference results for the affected eye on all 12 patients are summarized in Fig. 6.5. While the thickness difference plots were computed in a right-eye orientation (from a clinician's perspective), here the thickness plots for the cases in which the left eye was affected are flipped to a left-eye orientation. Because the visual field data is shown from the patient's perspective, one can directly compare the visual field data and the thickness difference plots along the left-right axis (i.e., the different perspectives – clinician versus patient – take care of the fact that the temporal visual field corresponds to the nasal portion of the retina and vice-versa). However, because of the optics of the eye, the upper portion of the visual field should be compared to the lower portion of the thickness difference plots. This correspondence is indicated in the upper left panel of Fig. 6.5. The red circles depict correspondence of the retinal area imaged by a 10 degree macular OCT scan



(a) Inner retina, all segments (b) Inner retina, selected segments



(c) Outer retina, all segments (d) Outer retina, selected segments



(e) Total retina, all segments (f) Total retina, selected segments

Figure 6.4: Summary of thickness values based on our intraretinal layer segmentation approach presented at MICCAI 2007 [37]. (a) Thickness values for the inner retina over all 12 segments. (b) Thickness values for the inner retina over selected segments. (c) Thickness values for the outer retina over all 12 segments. (d) Thickness values for the outer retina over selected segments. (e) Thickness values for the total retina over all 12 segments. (f) Thickness values for the total retina over selected segments. Thickness differences between the affected and unaffected eyes were largest on average for the inner retinal layer over the selected segments.

with the retinal region subjected to visual field testing. Dark areas of the visual field represent a decrease in visual sensitivity.

Fig. 6.6 provides an example of a subject with a lower altitudinal visual field defect in the affected left eye. The visual field sensitivities for both eyes are shown, along with the vertical OCT scan for both eyes (segmented inner layer indicated in yellow). Note the thinning of the inner layer on the affected eye. This thinning is also evident in the line plot of the thickness values for the vertical slice of the affected eye as given by the orange line in Fig. 6.6(d). This line is superimposed on the same inner retinal layer thickness values from undiseased eyes (mean ± 2 SD) indicated by the black and dashed lines.

6.1.4 Discussion

As expected, the inner retina of the affected eye demonstrated significant thinning, while the outer retina did not. This helps to support the notion that focusing on the properties of the inner retinal layers may provide more discriminative information than that of the total retina. This would especially be true in cases of confounding factors causing an increased thickness of the outer retinal layers. In addition, it is important to note that even larger differences between the affected and unaffected eyes were achieved using the segmental approach, further supporting this notion and also suggesting the importance of looking at regions locally.

The color-coded thickness difference maps of the inner retina displayed a strikingly locally specific correspondence between the inner retinal layer thinning and the location of visual sensitivity loss. Such a local correspondence was not present in the outer retina, providing additional support for the hypothesis that local layer-specific retinal thickness is associated with visual field sensitivity. Note that this correspondence was noted even though time-domain OCT images were used in the analysis. It would be expected that even better correspondences would be achieved using more densely acquired spectral-domain OCT images.

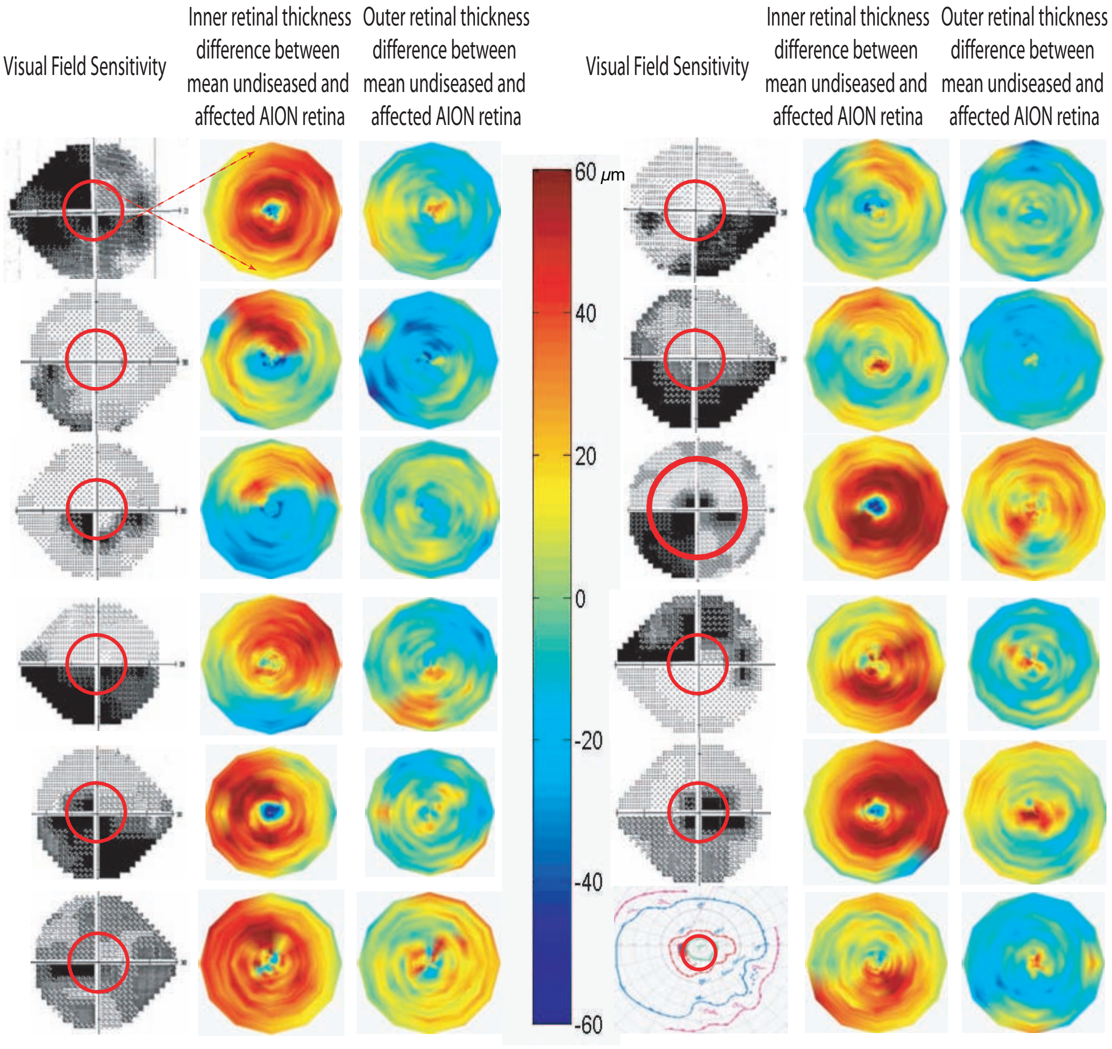
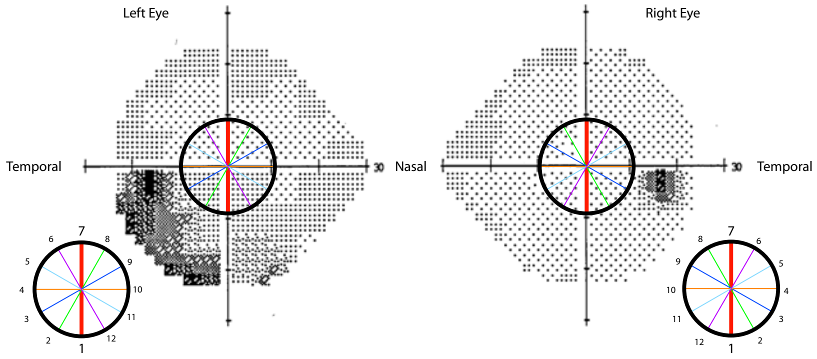
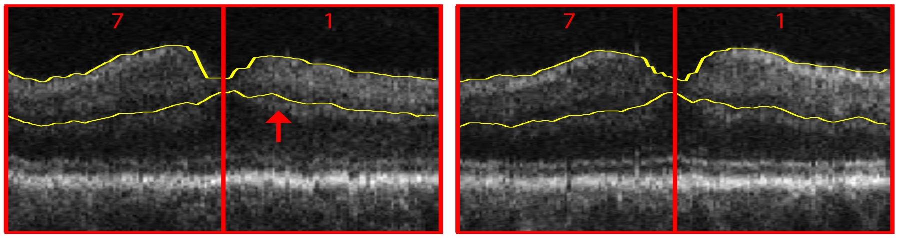


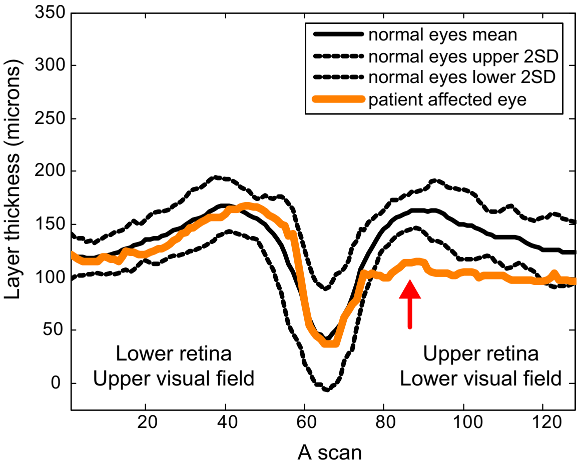
Figure 6.5: Summary of visual field and inner/outer retinal layer thickness difference values for the AION-affected eye on all 12 patients. The correspondence scheme between the visual field data and the color-coded OCT thickness difference maps is indicated in the top left panel (upper visual field corresponds to lower retina). Red areas on the color-coded OCT thickness difference plots represent thinning compared to the average thickness observed in undiseased eyes; calculated as mean local layer thickness observed in the set of 12 undiseased eyes *minus* the local layer thickness observed in the AION diseased eye.



(a) 24-2 threshold test



(b) Inner layer on affected left eye (c) Inner layer on unaffected right eye



(d) Line scan of inner layer thickness

Figure 6.6: Example of localized thinning of the inner retina in patient with AION of the left eye. (a) Visual field threshold test results (24-2) for both the left and right eyes. Note the visual field defect in the left eye. The circled region indicates the coverage area of the fast macular OCT scan (10 degrees). (b-c) Vertical OCT slice for the left and right eyes, with segmented inner layers indicated in yellow and thinning indicated with a red arrow. (d) Line plot of inner retinal thickness on affected eye, with mean ± 2 standard deviations of the unaffected thickness (over all patients) also plotted as a reference.

6.2 Assessment in normal subjects using spectral-domain OCT

6.2.1 Motivation

The more densely acquired spectral OCT data provides even more localized information over that which is possible with time-domain OCT. As illustrated in the previous section with AION subjects, it is likely that utilizing localized information will be important in assessing disease status. For example, local layer thicknesses differences may indicate areas of damage more readily than when averaged out with neighboring undamaged regions. A first step in such an assessment is to define the normal layer thicknesses. In this section, we provide a pilot study analyzing the thicknesses of six layers in normal subjects based on spectral-domain OCT. Part of this work was presented at ARVO 2008 [55].

6.2.2 Methods

We applied the developed 3-D spectral segmentation approach presented in the previous chapter to macula-centered 3-D OCT volumes (200 × 200 × 1024 voxels, 6 × 6 × 2 mm^3) of normal subjects in two groups (data provided by Carl Zeiss Meditec, Inc.):

- 15 normal subjects with a mean age of 39.4 ± 8.2 years (data from machine A as described in the previous chapter).
- 27 normal subjects comprised of the 15 normal subjects above plus 12 additional normal subjects (additional subjects from machine C as described in the previous chapter). Age data was not available for the additional subjects.

Both groups were analyzed because the data from the first group had the advantage of having the age data available, while the second group had the advantage of having a larger number of subjects.

After automatically segmenting seven surfaces on all images, the mean and standard deviation of the thickness of each of the six layers were computed for each (x, y)

location in the en-face plane. The results were displayed as color-coded maps in a right-eye orientation (from the clinician's perspective so that the temporal side corresponded to the left half of the image and the nasal side corresponded to the right half of the image). In addition, temporal-nasal thickness differences were computed and tested for significance using a paired t-test (with Bonferroni correction, p-values less than $\frac{0.05}{6} = 0.008$ were considered significant).

6.2.3 Results

Fig. 6.7 summarizes the resulting thickness and thickness variability maps from the 15 normal subjects with age data available, while Fig. 6.8 summarizes the resulting thickness and thickness variability maps from all 27 normal subjects. The resulting temporal, nasal, and temporal-nasal thicknesses are summarized in Tables 6.1 and 6.2 for the 15-subject group and the 27-subject group, respectively. In both cases, the retinal nerve fiber layer was significantly thicker nasally (by slightly less than 30 μm), as expected ($p < 0.001$). However, other layers showed significant differences as well. For example, the ganglion cell layer (plus the inner plexiform layer) as well as the outer nuclear layer (plus photoreceptor inner segments) were significantly thicker temporally ($p \leq 0.002$). However, the differences were small ($< 4\mu$m). When analyzing all 27 subjects, the inner nuclear layer was significantly thicker nasally ($p < 0.001$), but again, the mean difference was small (approximately 5.4 μm).

6.2.4 Discussion

While the nasally thicker RNFL was expected, some of the other temporal-nasal thickness differences were not. Because the differences were small, the clinical relevance still needs to be investigated. It may also be useful to look at the quantitative differences at an even more localized level.

Perhaps more importantly, note that the above approach sets the groundwork for having a normal atlas. Such an atlas can then be used for localized comparisons with

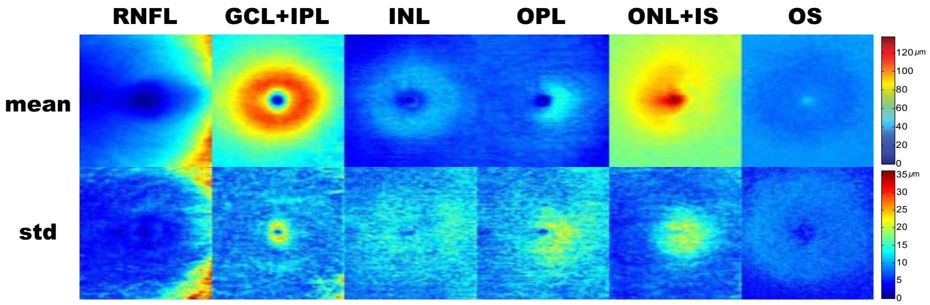


Figure 6.7: Thickness and thickness variability maps of six macular intraretinal layers from 15 normal subjects.

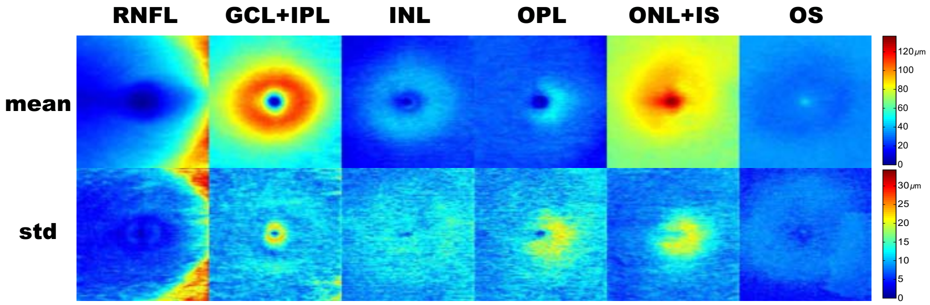


Figure 6.8: Thickness and thickness variability maps of six macular intraretinal layers from 27 normal subjects.

Table 6.1: Temporal-nasal mean thickness summary for 15 normal subjects[†] in μm.

Layer	Mean temporal	Mean nasal	Mean difference	p-value[‡]
1-2 (RNFL)	23.2 ± 2.0	49.9 ± 5.8	-26.7 ± 5.2	$p < 0.001^*$
2-3 (GCL+IPL)	73.0 ± 3.6	70.3 ± 3.9	2.7 ± 2.3	$p < 0.001^*$
3-4 (INL)	23.7 ± 2.7	29.0 ± 5.9	-5.3 ± 7.8	$p = 0.020$
4-5 (OPL)	26.9 ± 2.4	28.3 ± 5.6	-1.3 ± 6.3	$p = 0.424$
5-6 (ONL+IS)	79.1 ± 5.1	75.2 ± 5.8	3.8 ± 3.5	$p < 0.001^*$
6-7 (OS)	33.4 ± 3.7	33.4 ± 3.0	-0.1 ± 2.1	$p = 0.913$

[†] Only subjects in which age data was available were included.

[‡] Each p-value computed using paired t-test. With Bonferroni correction, p-values less than $\frac{0.05}{6} = 0.008$ were considered significant.

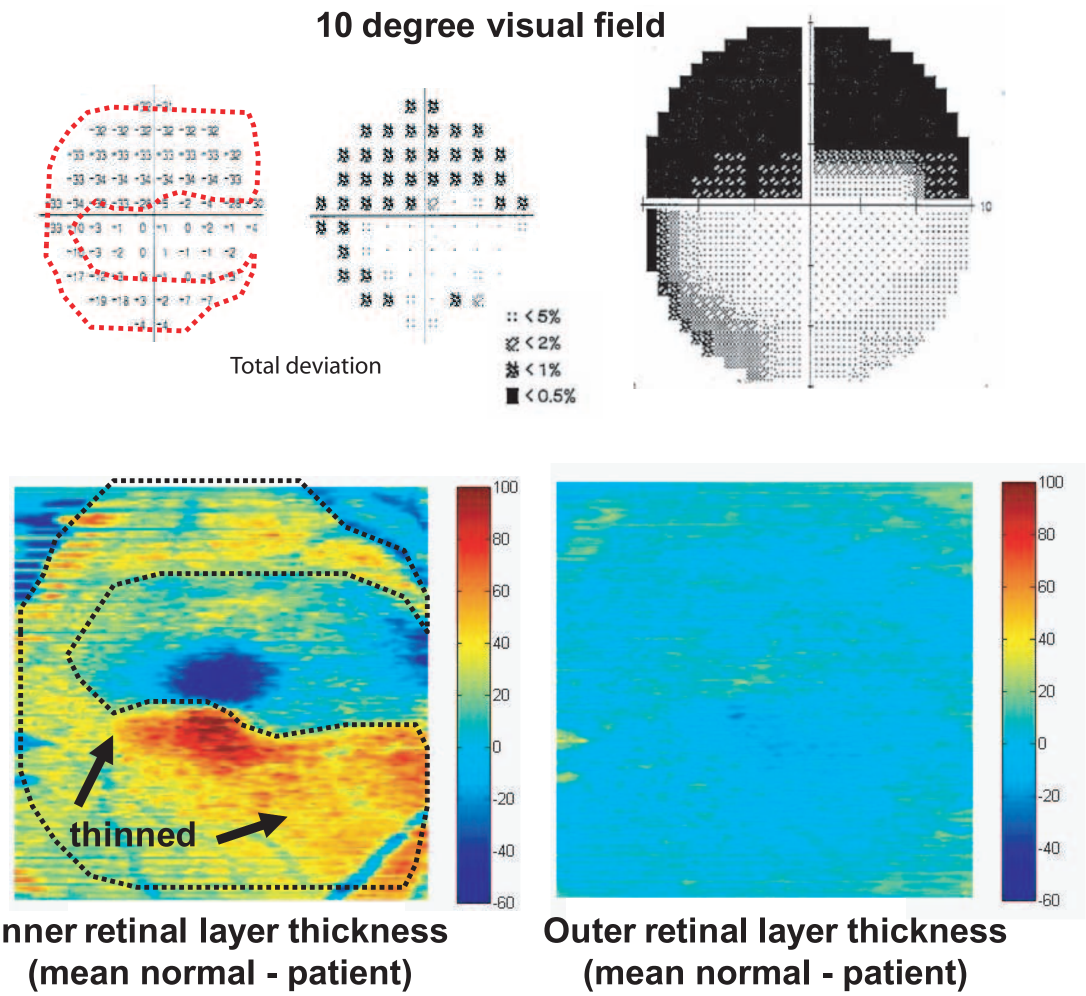
Table 6.2: Temporal-nasal mean thickness summary for 27 normal subjects[†] in μm.

Layer	Mean temporal	Mean nasal	Mean difference	p-value[‡]
1-2 (RNFL)	23.5 ± 2.3	52.5 ± 6.8	-29.0 ± 5.6	$p < 0.001^*$
2-3 (GCL+IPL)	72.1 ± 5.3	69.9 ± 5.3	2.2 ± 2.4	$p < 0.001^*$
3-4 (INL)	23.6 ± 2.9	29.1 ± 4.9	-5.4 ± 6.4	$p < 0.001^*$
4-5 (OPL)	26.9 ± 2.6	26.8 ± 5.3	0.1 ± 5.2	$p = 0.908$
5-6 (ONL+IS)	80.6 ± 6.1	78.1 ± 6.8	2.5 ± 3.9	$p = 0.002^*$
6-7 (OS)	33.5 ± 3.0	33.1 ± 2.8	0.5 ± 2.8	$p = 0.377$

[†] All subjects were included (including those for which age was unknown).

[‡] Each p-value computed using paired t-test. With Bonferroni correction, p-values less than $\frac{0.05}{6} = 0.008$ were considered significant.

patient data. For example, Fig. 6.9 illustrates the use of normal thickness data to produce thickness difference plots of the inner and outer layers for a patient with primary open angle glaucoma. After segmenting an OCT volume of the patient's eye into the inner and outer retinal layers, color-coded difference plots were produced by subtracting the glaucoma patient's thickness value from the normal value for each (x, y) location. In this way, positive values indicated thinning (illustrated in the lower portion of the figure). In addition, the visual field sensitivity results (using a Humphrey 10-2 visual field testing protocol which covers the central 10 degrees) are illustrated in the upper portion of the figure. In particular, the gray scale map (upper right portion of figure) shows a visual sensitivity loss that was worst in the top part of the visual field (dark areas), but also shows some loss in the inferior field. A plot of the visual field sensitivity difference from normal is also shown in the upper left corner with the abnormal area with decrease in sensitivity surrounded by a red dotted line. The statistical probability plot of the same visual field data is shown in the upper middle portion of the figure. Note the high spatial correlation between the thinned inner layer in the inferior (and to some degree in the superior) macular regions and the corresponding superior (and inferior) areas of the visual field defects. However, there was no such thinning in the outer retina which is known not to be



Inner retinal layer thickness (mean normal - patient)

Outer retinal layer thickness (mean normal - patient)

Figure 6.9: Spectral-domain macular OCT and visual field correlation in patient with glaucoma. Note the superior visual field defect encroaching on central fixation and the corresponding inner retinal thinning of ganglion cell layer in the inferior macula.

affected in most glaucomatous damage.

CHAPTER 7
GENERAL DISCUSSION AND FUTURE DIRECTIONS

The methods developed for this thesis have a number of potential applications and future directions. While some of the general extensions to the graph search approach described in Chapter 3 are applicable to a broad range of segmentation tasks; we will concentrate on ophthalmic applications in this chapter. This rest of this chapter is organized as follows. First, we present a sampling of some additional image analysis directions related to optical coherence tomography, concentrating on the ones related to layer segmentation. Finally, we present how these tasks may be applied to clinical applications.

7.1 Additional image analysis directions

7.1.1 Layer properties

While thickness has been shown to be an important property of retinal layers, other properties may be useful as well. For example, one may also consider computing a variety texture-based properties (such as those as described in [56–58]). For example, some of the properties could include:

- Thickness values of each layer (and group of layers),
- Mean and variance of the (normalized) intensity of each of layer,
- Average regional response to Gaussian filters of various orders, scales and orientations [58],
- Average regional response to Gabor filters of different preferred orientations and spatial frequencies [59, 60], and
- Other traditionally used textural features including variance and the co-occurrence matrix features proposed by Haralick [61].

Other properties of interest could include shape-based terms (e.g., measuring the slope of a surface of interest) and binary terms such as the presence/absence of a vessel shadow.

Spectral OCT also makes it possible to quantify the properties as a more local level. Thus, one could consider measuring each property in specified regions of interest instead of globally. Such a local approach will especially be important when regions of the eye are affected differently (e.g., see thickness changes between affected and unaffected eyes of subjects with unilateral AION in Chapter 6).

7.1.2 Registration

We have already mentioned how vessels can be segmented on OCT images in Chapter 5 and in [46, 47]. Recall that such segmentations are performed on a 2-D projection of a layer near the RPE plane. Because approaches already exist for the segmentation of the vasculature in fundus photographs [62], the vessel information can be used to register fundus photographs and OCT volumes. Such a registration could then be used in a variety of applications such as for obtaining a better understanding of the correspondence of clinically relevant features in both modalities. In addition, the OCT vessel information could also be used to register OCT volumes to one another. This would not only allow the creation of an atlas (see next section), but also allow for a more consistent "placement" of learned varying smoothness and thickness constraints for use in the intraretinal layer segmentation approach.

7.1.3 Atlas creation

The creation of a 3-D atlas which contains information about normal layer properties and typical inter-subject variations is an important step in being able to assess differences from normal in patient data. A preliminary example of this use was presented with the glaucoma patient at the end of Chapter 6. While the mean thickness maps (and associated standard deviations) of the layers presented in Chapter 6 can be considered a very basic atlas, normalizing the data to a common retinal coordinate system (with the help of registration approaches presented above) would likely reduce the inter-subject variation, enabling better detection of abnormalities in patient data. Having properties besides thickness (see Section 7.1.1) may also be important. Furthermore, it would be useful to create separate atlases for individual groups based on age, gender, and race.

7.2 Disease-specific applications

7.2.1 Glaucoma and AION

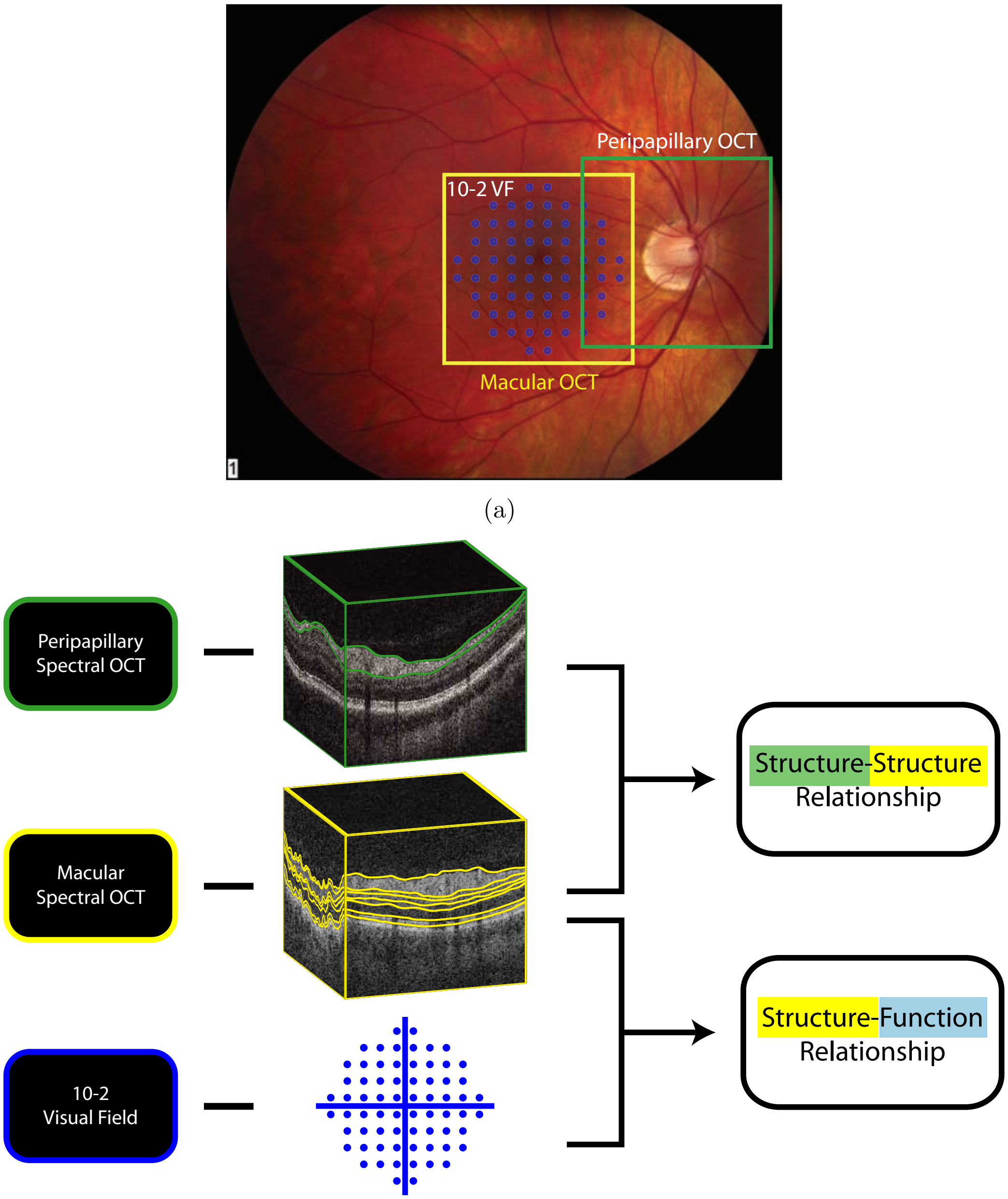
Glaucoma is a progressive, but treatable optic neuropathy characterized clinically by optic nerve cupping with corresponding visual field deficits, with or without elevated intraocular pressure (IOP). As mentioned previously (see Chapter 6), the current methods for assessing glaucomatous damage leave much to be desired in terms of reproducibility and dynamic range. For example, visual field testing results (the gold standard functional measurement) have a large degree of variability, especially with increasing amounts of damage [51–54]. The assessment of stereo optic disc photographs by human experts (including looking at such measurements as the cup-to-disc ratio) is subjective and exhibits considerable inter-observer and intra-observer variability [63, 64]. Retinal nerve fiber thickness measurements reach a minimum thickness even as the visual field sensitivity worsens [48–50]. Thus, there is a definite need for better methods for monitoring progression that are less subjective, more reproducible, and/or have a larger dynamic range. This has been one of the motivations for focusing on segmenting the macula in this work (again, see previous chapter).

To test our hypothesis that the macular region does in fact contain important information for obtaining measurements that are more reproducible and have a larger dynamic range than existing techniques, further studies need to be performed. As a first step, one could study and verify the expected correspondence between properties of layers of the macula and their expected corresponding locations in the peripapillary regions. This would involve first involve acquiring both macular and peripapillary OCT scans for glaucoma (and other optic neuropathies such as AION) and normal subjects (Fig. 7.1). The macular scans would be segmented into six layers whereas only the retinal nerve fiber would need to be segmented on the peripapillary scans. After segmenting the scans, properties of interest could be extracted (focusing on thickness, but also possibly examining other features as well), and the relationships explored in corresponding regions, at both a global and regional level. Fig. 7.2 illustrates such a relationship with regional comparisons (1a vs. α; 2 vs. α; 4 vs. α; 1b vs. β; 3 vs. β; and 5 vs. β) and global comparisons ([1a+2+4+6] vs. α and [1b+3+5+6] vs. β). Fig. 7.3 illustrates possible adjustments to these regions.

In addition to looking at the structure-structure relationship above, one could also examine the structure-function relationship in the macula at a local level. More specifically, a Humphrey 10-2 visual field test could be performed to result in 68 test locations (separated by 2 degrees) covering a 10 degree radius of the visual field centered at the fovea. Note that the visual sensitivity test measures sensitivity in decibels of light attenuation needed to reach threshold perception. A direct correspondence can then be established between the regional macular OCT data (called macular columns) and the 68 visual field test locations, as illustrated in Fig. 7.4. This will then allow for examining localized associations between the visual field data and columnar properties. In addition to using standard statistical approaches (such as multivariate modeling), a classifier approach could also be developed to examine the relationships.

7.2.2 CNV and DME

Choroidal neovascularization (CNV) and diabetic macular edema (DME) exist in the most severe complications of age-related macular degeneration and diabetic retinopathy, respectively. Both involve an abnormal growth and increased permeability of blood vessels and cause abnormalities to form in the retina. These abnormalities can be referred to as symptomatic exudate-associated derangements or SEADs. Example SEADs in time-domain OCT images from AMD are illustrated in Fig. 7.5.



(a)

(b)

Figure 7.1: Illustration of data used for assessing structure-structure and structure-function relationships in glaucoma. (a) Schematic illustration of macular OCT, peripapillary OCT, and 10-2 visual field test locations on a fundus photograph. (a) Schematic of segmented peripapillary OCT scan, segmented macular OCT scan, and 10-2 visual field test, with corresponding structure-structure and structure-function relationships indicated.

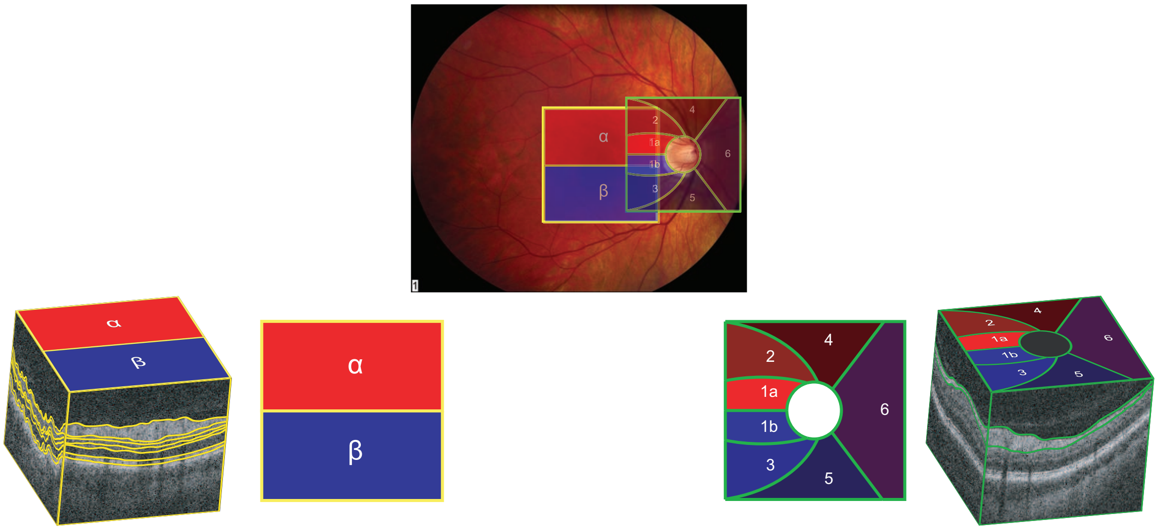
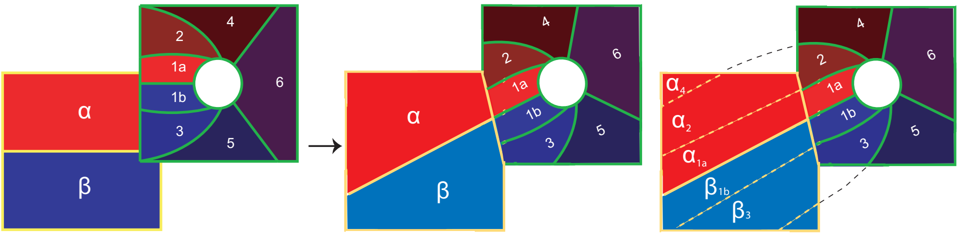


Figure 7.2: Finding relationships between the macular and peripapillary regions.



(a)

(b)

Figure 7.3: Example adjustment of the labeled regions so that the regions match the underlying anatomy – a larger than realistic angular difference used for demonstrating the principle. (a) Schematic of the proposed adjustment. (b) The macular region could also be subdivided as shown.

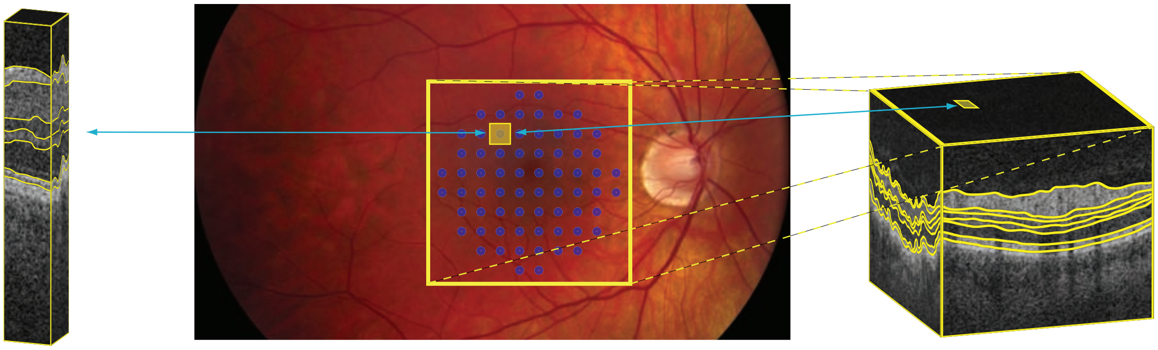


Figure 7.4: Example macular column and associated Humphrey 10-2 VF sensitivity test location.

In AMD, anti-VEGF injections have been used successfully to treat patients with CNV; however, the injections do incure risks of serious ocular events and are expensive. Furthermore, each patient can respond differently. Thus, it would be desirable to develop a patient-specific dosing strategy with the minimum number of injections required. Quantifying the SEAD properties (such as volume and texture) and properties of the surrounding layers would be one important step in being able to determine an optimal dosing strategy. Such a quantification would also be important in assessing treatment strategies for DME as well.

In order to do this, it will become necessary to segment both the SEADs and the layers in these images. The presence of the SEADs will require the development of additional segmentation approaches. One possibility would be to segment the SEADs first using a classifier-based approach, followed by the layer segmentation using the SEAD locations as “don’t care” regions. An example preliminary segmentation in SEADs in AMD using an interactive level-set approach is illustrated in Fig. 7.6, with a DME example illustrated in Fig. 7.7.

With segmented SEADs and layers of an OCT image at a given time, one can then imagine computing a number of local properties of the SEADs (e.g., volume, footprint area, texture, etc.) and the layers (see Section 7.1.1). Features could also be computed on a corresponding fundus photograph and registered together using the retinal vasculature, providing multiple derived features for each time point (see Fig. 7.8). Functional information, such as that resulting from ETDRS visual acuity and contrast sensitivity tests could also be performed.

Such features could thus be acquired based on the tests at multiple time points during a treatment scheme and then used to help learn the relationship between the features and response to treatment. In particular, it would be of interest to learn both 1) which features change over time and 2) which features are associated with outcome parameters (e.g., total retinal thickness changes, decrease in numbers/size of SEADs,

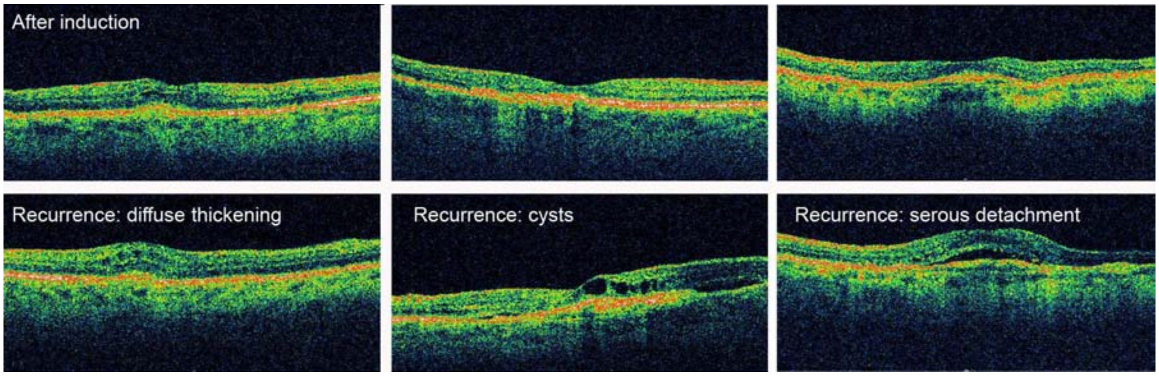


Figure 7.5: CNV recurrence in AMD after successful treatment induction. (CNV = choroidal neovascularization, SEAD = symptomatic exudate-associated derangement.)

number of treatments necessary in a time period). Fig. 7.9 illustrates a proposed scheme for acquiring the data in order to provide data for learning such relationships. In this example scheme, there is an “induction” phase of approximately 3 months in which data are regularly acquired more frequently than the treatments (this data can be useful in looking at short-term changes). This is followed by a longer period of time in which the data are only acquired when treatments are given (with treatment time decided by a clinician).

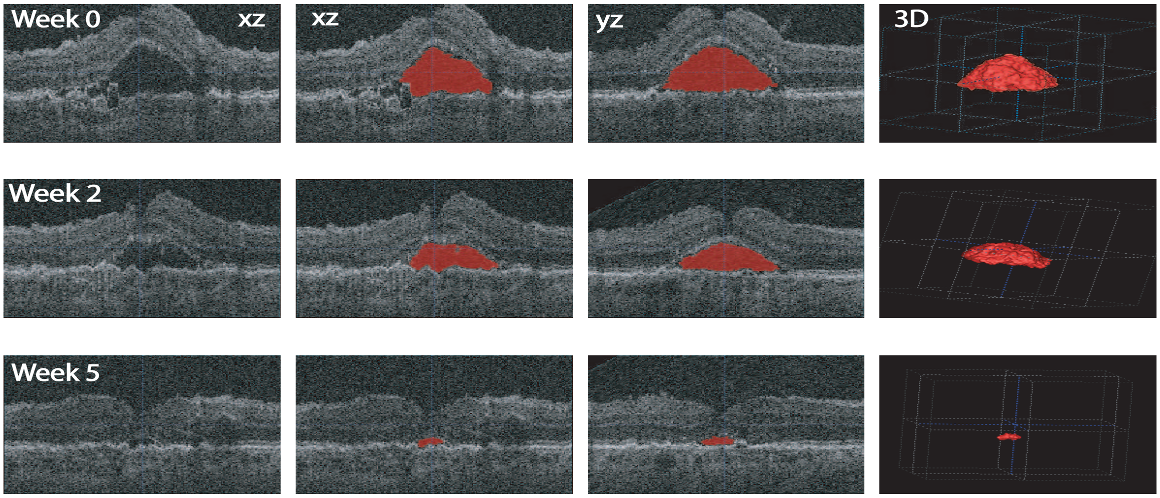


Figure 7.6: Example SEAD response to treatment in AMD.

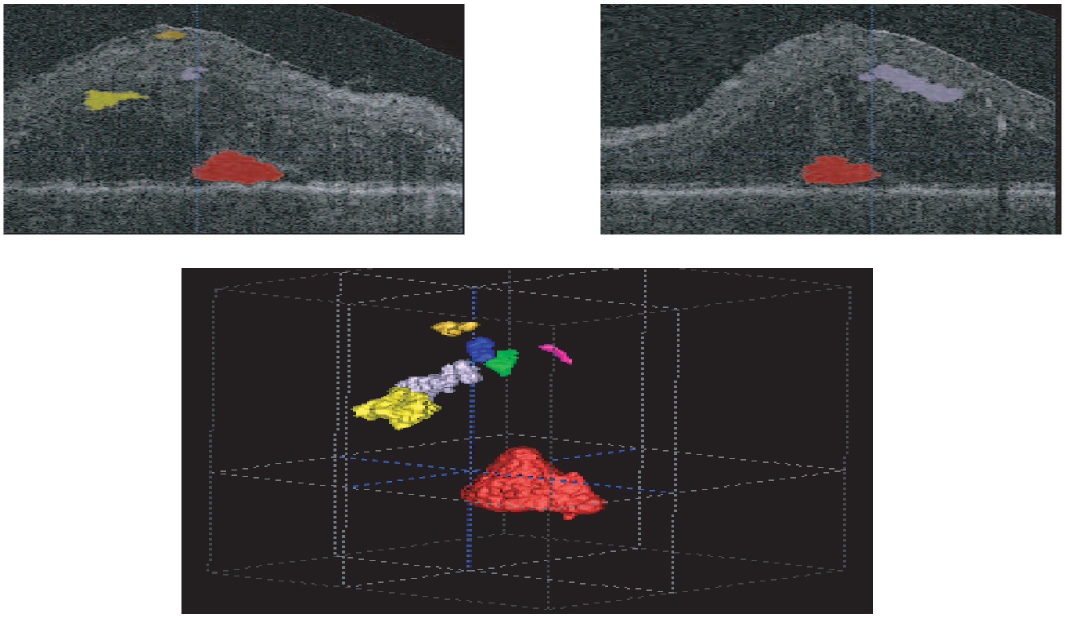


Figure 7.7: Example SEAD segmentation in DME.

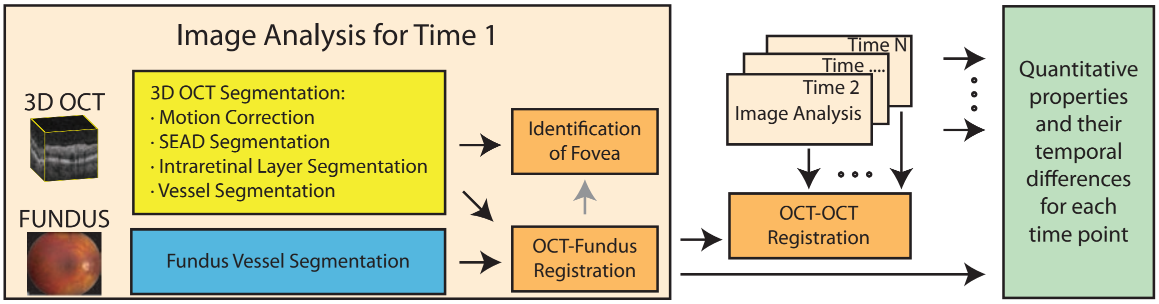


Figure 7.8: Steps leading to quantitative analysis in OCT images.

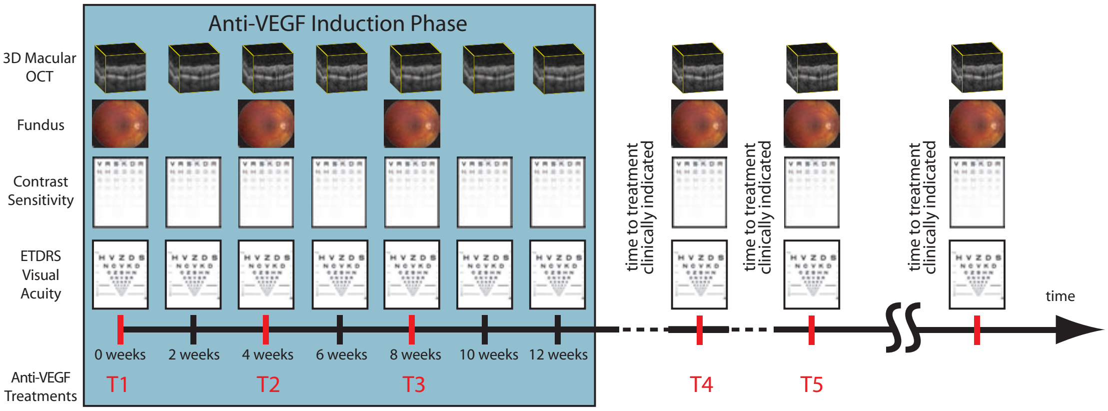


Figure 7.9: Predicting response to anti-VEGF treatment in CNV patients.

CHAPTER 8
CONCLUSIONS

To summarize, first recall the specific aims for this thesis:

- **Aim 1:** Develop a method for the incorporation of regional information and varying constraints into the optimal 3-D graph search.
- **Aim 2:** Develop a method for learning cost functions and varying constraints from examples for use in the 3-D graph search.
- **Aim 3:** Develop and validate a method for the 3-D segmentation of intraretinal layers from OCT images.
- **Aim 4:** Use the intraretinal layer segmentation approach to (a) help determine which macular layers show significant changes in thickness in unilateral anterior ischemic optic neuropathy (AION) subjects and (b) perform pilot analysis of localized thickness and thickness variability of individual layers in normal subjects.

The developed method for the incorporation of regional information and varying constraints into the optimal 3-D graph search (Aim 1) was described for the general case in Chapter 3. While this development was motivated by needing to segment OCT images, the method remains general so that it can be applicable to other segmentation tasks as well. A method for learning appropriate constraints from examples (second part of Aim 2) was also described in this chapter, with Chapter 4 illustrating how such an approach can improve segmentation results. Chapters 4 and 5 described the developed methods and validation of an automated method for the 3-D segmentation of intraretinal layers in time-domain (Chapter 4) and spectral-domain (Chapter 5) OCT images (Aim 3), achieving results comparable to that of human observers. Chapter 5 also included an approach for learning in-region cost terms and determining an appropriate combination of on-surface and in-region cost terms from examples (first part of Aim 2). Finally, in Chapter 6, the developed segmentation approach was applied to illustrate how the inner layer thins in anterior ischemic optic neuropathy, while the outer layer does not (Aim 4a) and to perform an pilot analysis of localized thickness and thickness variability of individual layers in normal subjects (Aim 4b).

In addition, this thesis has provided potential future directions for this work. The previous chapter (Chapter 7) has just touched on some of the possible ophthalmic applications for utilizing an automated 3-D intraretinal layer segmentation approach (and other associated image analysis tools). Thus, such an approach shows great promise in further contributing to the ophthalmology community.

APPENDIX
GRAPH THEORY BACKGROUND

This appendix reviews some of the relevant graph principles and algorithms related to transforming segmentation problems into finding a minimum-cost path in a graph (frequently used for 2-D segmentation tasks) or finding a maximum flow (minimum cut) in a graph (used as the final step in the 3-D segmentation approach used in this thesis). After presenting definitions, example algorithms are presented. Note that much of the material is based on that found in the *Introduction to Algorithms* textbook by Cormen, Leiserson, Rivest, and Stein [65] (algorithms reused with permission courtesy of The MIT Press).

A.1 Basic definitions and notation

A graph $G = (V, E)$ consists of a set of vertices V and a set of edges E. Each edge e in the edge set E is an ordered pair of vertices from V ($E \subseteq V \times V$). In a directed graph, the order of the listed vertices in the pair matters and the second listed vertex in the pair is defined as being adjacent to the first vertex. For example, if (u, v) is an edge in a directed graph, v is adjacent to u; however, u is not adjacent to v unless (v, u) is also an edge in the graph. Note that this contrasts with an undirected graph in which the order of the vertices does not matter, making the adjacency relationship symmetric. Thus, if (u, v) is an edge in an undirected graph (the same edge as (v, u)), u would be adjacent to v, and v would also be adjacent to u.

In an edge-weighted graph (also known as just a weighted graph), each edge also has a weight associated with it. The weight of an edge (u, v) will be referred to as $w(u, v)$. An example of a directed graph consisting of the vertices $\{a, b, c, d, e, f\}$ and edges $\{(a, c), (a, d), (c, e), (d, e), (d, b), (d, f)\}$ is shown in Fig. A.1(a). A weighted version of this graph is shown in Fig. A.1(b).

In a vertex-weighted graph, each vertex has a weight associated with it. The weight of of a vertex v will be referred to as $w(v)$. Note that in many image segmen-

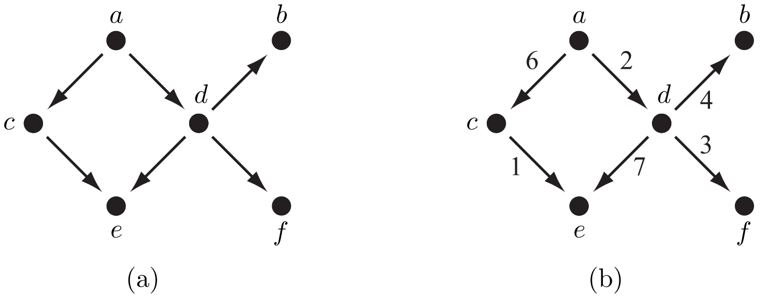


Figure A.1: Unweighted and weighted directed graphs. (a) Directed graph with vertex set $V = \{a, b, c, d, e, f\}$ and edge set $E = \{(a, c), (a, d), (c, e), (d, e), (d, b), (d, f)\}$. (b) Directed graph as in (a), but with example edge weights $w(a, c) = 6$, $w(a, d) = 2$, $w(c, e) = 1$, $w(d, e) = 7$, $w(d, b) = 4$, $w(d, f) = 3$.

tation approaches, these "weights" are also often referred to as "costs."

A.1.1 Paths

A path $p = \langle v_0, v_1, \ldots, v_{k-1}, v_k \rangle$ through a graph can be represented as a sequence of vertices in which each vertex in the sequence (other than the first) is adjacent to the previous vertex in the sequence. In addition to the vertices $v_0, v_1, \ldots, v_{k-1}, v_k$, the path also contains the edges $(v_0, v_1), (v_1, v_2), \ldots, (v_{k-1}, v_k)$ [66]. If the graph is edge-weighted, each path also has a cost associated with it, defined as the sum of its edge weights:

$$\mathrm{cost}(p) = \sum_{i=1}^{k} w(v_{i-1}, v_i) \ , \tag{A.1}$$

where $\langle v_0, v_1, \ldots, v_{k-1}, v_k \rangle$ are the vertices in the path. For example, one path in the simple graph of Fig. A.1(b) is $\langle a, c, e \rangle$, with $cost(p) = w(a, c) + w(c, e) = 7$.

In case of a vertex-weighted graph, the cost of a path can be written as:

$$\mathrm{cost}(p) = \sum_{i=0}^{k} w(v_i) \ , \tag{A.2}$$

and in the case of both an edge-weighted and a vertex-weighted graph, the cost of a path can be written as:

$$\text{cost}(p) = \sum_{i=0}^{k} w(v_i) + \sum_{i=1}^{k} w(v_{i-1}, v_i) \ . \tag{A.3}$$

A cycle in a directed graph is a path that has at least one edge for which the first and last vertices are the same. If a graph has no cycles, it is acyclic.

A.1.2 Flow networks

As described in [67], a flow network is a special type of directed weighted graph. In particular, in a flow network:

1. The weight of each edge $(u, v) \in E$ is nonnegative. We will refer to these weights as *capacities* and define a capacity function $c(u, v)$ such that $c(u, v) = w(u, v)$ for $(u, v) \in E$ and $(c, v) = 0$ for $(u, v) \notin E$. (Note that the capacity function is defined for all (u, v) pairs, even when $(u, v) \notin E$.)
2. There are two special vertices: a source s and a sink t. Furthermore, every vertex $v \in V$ lies on some path from the source to the sink. Thus, the graph is connected and $|E| \geq |V| - 1$.

Intuitively, one can think of a flow network as modeling a system in which material is produced at a source and “flows” through conduits (edges) until it reaches the sink. The capacities reflect the maximum amount of material that can flow across each conduit. Fig. A.2 illustrates an example flow network.

A particular flow can be described by a function $f(u, v)$ in which the following three properties are satisfied:

- The flow $f(u, v)$ between two vertices is no more than the capacity of flow between those two vertices. (For all $u, v \in V$, $f(u, v) \leq c(u, v)$.)
- The flow between two vertices in one direction is the negative of the flow in the opposite direction. (For all $u, v \in V$, $f(u, v) = -f(v, u)$.)

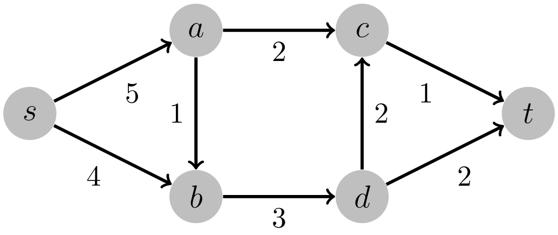


Figure A.2: Example flow network with capacities indicated.

- Flow is conserved. (For all $u \in V - \{s, t\}, \sum_{v \in V} f(u, v) = 0$, meaning that except for the source and sink, the total flow out of each vertex is 0. Because of the second property above, this also means that the total flow into each vertex (other than the source or sink) is also 0. Another interpretation of the conservative property is that the positive flow entering each vertex must equal the positive flow leaving the vertex.)

Note that although capacities must be positive, the flow between vertices can be positive, negative, or zero. The total flow out of the source is called the value of a flow f, $|f|$:

$$|f| = \sum_{v \in V} f(s, v) \; . \tag{A.4}$$

Given a flow f, the residual capacity $c_f(u, v)$ for each pair of vertices is given by

$$c_f(u, v) = c(u, v) - f(u, v) \; , \tag{A.5}$$

which reflects the additional amount of flow that the can flow between vertices u and v. Then the residual network of the network flow $G = (V, E)$ is a graph $G_f(V, E_f)$ where

$$E_f = \{(u, v) \in V \times V : c_f(u, v) > 0\} \; . \tag{A.6}$$

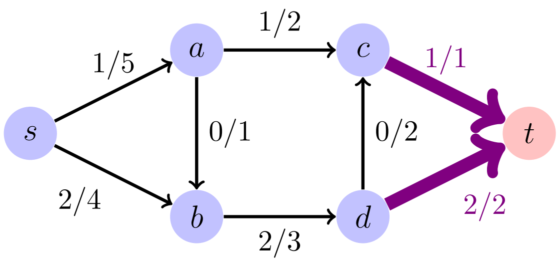


Figure A.3: Example maximum flow and corresponding minimum cut. Flows are indicated in the form $f(u, v)$ / $c(u, v)$ (flow over capacity), with only positive flows between vertices being indicated. The value of the maximum flow is 3 (note that other choices of flows would also produce this maximum value). In this example, the vertices in the source set S of the minimum cut are $\{s, a, b, c, d\}$ (indicated in blue), while the sink set only consists of the sink vertex t (indicated in red). The capacity of the cut is 3 (directed edges from S to T are indicated in purple).

Cuts are another important concept in flow networks. In particular, a cut (S, T) of a flow network is a partition of the vertices V into two disjoint sets: S and T $(T = V - S)$ with the requirement that $s \in S$ and $t \in T$. The capacity $c(S, T)$ of such a cut is defined as the summation of the capacities of the edges from vertices in S to vertices in T:

$$c(S,T) = \sum_{u \in S} \sum_{v \in T} c(u, v) \ . \tag{A.7}$$

A minimum cut of a flow network is defined as a cut with the minimum capacity. Fig. A.3 illustrates an example maximum flow and corresponding minimum cut for the flow network in Fig. A.2.

A.1.3 Closed sets

A closed set is a subset of vertices in a graph such that no directed edges leave the set. Given a vertex-weighted graph and a closed set, the cost of a closed set can be defined as the summation of all vertices belonging to the set.

A.2 Finding a minimum-cost path in a DAG

As shown in [66] and [68], the minimum cost path between a start vertex s and each other vertex in a directed acyclic graph can be found in $\Theta(V+E)$ time by "relaxing" edges according to a topological sort of the vertices (Algorithm A.1). The process of "relaxing" an edge (u,v) involves testing whether the current estimate of the shortest path through v can be improved by going through (u,v), as shown in Algorithm A.2. In other words, if $d[v]$ represents the cost of the current estimate of the minimum cost path from s to v, $d[v]$ is updated to $d[u]+w(u,v)$ if $d[u]+w(u,v) < d[v]$ [66]. The actual path is saved by also storing the predecessor of each vertex v in $\pi[v]$ (also updated during the relaxation of an edge). Initially, all the predecessors are set to NIL and the current estimates of the shortest cost paths are set to ∞, as shown in Algorithm A.3.

```
Dag-Shortest-Paths($G, w, s$)
topologically sort the vertices of $G$
Initialize-Single-Source($G, s$)
for each vertex $u$, taken in topologically sorted order
    do for each vertex $v \in Adj[u]$
        do Relax($u, v, w$)
```

Algorithm A.1: Dag-Shortest-Paths(G, w, s). Copied with permission from [66].

A.3 Finding a maximum flow (minimum cut) in a flow network

The goal of the maximum-flow problem is to find a flow f of maximum value in a flow network G with source s and sink t. Alternatively, this can be thought of as a finding a minimum cut in the flow network. This is because of the max-flow min-cut

```
RELAX($u, v, w$)
if $d[v] > d[u] + w(u, v)$
    then $d[v] \leftarrow d[u] + w(u, v)$
         $\pi[v] \leftarrow u$
```

Algorithm A.2: Relax(u, v, w). Copied with permission from [66].

```
INITIALIZE-SINGLE-SOURCE($G, s$)
for each vertex $v \in V[G]$
    do $d[v] \leftarrow \infty$
       $\pi[v] \leftarrow$ NIL
$d[s] \leftarrow 0$
```

Algorithm A.3: Initialize-Single-Source(G, s). Copied with permission from [66].

theorem which says that the value of the maximum flow is equal to the capacity of the minimum cut [67]. Note that given a maximum flow (and corresponding residual network G_f), a corresponding minimum cut can be found by defining the cut (S, T) as follows:

$$S = \{v \in V : \text{there exists a path from } s \text{ to } v \text{ in } G_f\}\ , \tag{A.8}$$

and $T = V - S$.

Multiple algorithm exist for finding a maximum flow (minimum cut) in a network flow graph. Here, we present two common methods, each of which serves as a basis for a number of more specific algorithms. The first general method, the Ford-Fulkerson method (Algorithm A.4), involves repeatedly finding paths from the source to sink for which additional material can flow. Such paths are called "augmenting paths" and are defined as simple paths that exist from s to t in the residual network G_f. Once no more augmenting paths can be found, the maximum flow has been found and the algorithm terminates. Note that the graph cut method of Boykov *et al.* [69] (designed with image segmentation tasks in mind) is a variation of an “augmenting path” method.

```
FORD-FULKERSON-METHOD($G, s, t$)
initialize flow $f$ to 0
while there exists an augmenting path $p$
    do augment flow $f$ along $p$
return $f$
```

Algorithm A.4: FORD-FULKERSON-METHOD(G, s, t). Copied with permission from [67].

Algorithm A.5 provides a more detailed version of a basic Ford-Fulkerson algorithm. The actual running time of such an algorithm depends on how the augmenting path is determined. For example, if a breadth-first search is used to finding the augmenting path p in line 4 (i.e., the chosen augmenting path is the shortest path from s to t in the residual network, where each edge has unit weight), then the algorithm runs in $O(VE^2)$ time (the Edmonds-Karp algorithm) [67].

A second general method for solving maximum-flow problems, the “push-relabel” method, involves working with the vertices of the graph at a more local level. Furthermore, instead of maintaining a proper flow function throughout the algorithm, push-relabel algorithms relax the “flow conservation” requirement so that instead of requiring that the total positive flow into a vertex has to equal the total positive flow out of a vertex, it just requires that the flow in must be larger than or equal to the

```
FORD-FULKERSON($G, s, t$)
for each edge $(u, v) \in E[G]$
    do $f[u, v] \leftarrow 0$
       $f[v, u] \leftarrow 0$
while there exists a path $p$ from $s$ to $t$ in the residual network $G_f$
    do $c_f(p) \leftarrow \min\{c_f(u, v) : (u, v) \text{ is in } p\}$
       for each edge $(u, v)$ in $p$
           do $f[u, v] \leftarrow f[u, v] + c_f(p)$
              $f[u, v] \leftarrow -f[u, v]$
```

Algorithm A.5: FORD-FULKERSON(G, s, t). Copied with permission from [67].

flow out ($\sum_{v \in V} f(v, u) \geq 0$ for all vertices $u \in V - \{s\}$). This relaxed version of a flow is called a "preflow." The other properties of a flow (i.e., obeying capacity constraints and requiring the flow in one direction between two vertices to be the negative of the flow in the opposite direction) are still maintained in a preflow. The excess flow $e(u)$ into u is defined as

$$e(u) = \sum_{v \in V} f(v, u) . \tag{A.9}$$

Push-relabel algorithms also maintain what is known as a "height function," in which an integer value intuitively representing a height is assigned to each vertex. To be a valid height function, a function h must satisfy the following properties:

- $h(s) = |V|$,
- $h(t) = 0$, and
- $h(u) \leq h(v) + 1$ for every residual edge $(u, v) \in E_f$.

Algorithm A.6 provides a generic version of a push-relabel algorithm. After initialization (Algorithm A.7), a sequence of "push" (Algorithm A.8) and "relabel" (Algorithm A.9) operations are performed to modify a height function h and preflow f associated with the vertices. The "push" operations essentially work to modify the preflow by "pushing" excess flow from vertices to neighboring vertices, while the "relabel" operations work to modify the height function (still maintaining a proper height function). At termination, no excess flows exist and the preflow thus reflects a proper flow. Furthermore, it reflects a *maximum* flow because of the fact that with a proper height function and preflow (as is maintained in the algorithm), no path exists in the residual network G_f from the source s to the sink t [67]. The general algorithm can run in $O(V^2E)$ time, but with appropriate data structures and careful choice of the order of the push and relabel operations, faster running times are possible. More details and further analysis of the push-relabel method can be found in [67].

```
GENERIC-PUSH-RELABEL($G$)
INITIALIZE-PREFLOW($G, s$)
while there exists an applicable push or relabel operation
    do select an applicable push or relabel operation and perform it
```

Algorithm A.6: GENERIC-PUSH-RELABEL(G). Copied with permission from [67].

A.4 Finding a minimum-cost closed set

As shown in [6, 70, 71], an optimal closed set can be found in a directed vertex-weighted graph $G = (V, E)$ by computing a minimum s-t cut in a related graph $G_{st} = (V_{st}, E_{st})$. In particular, for finding a minimum-cost closed set, G_{st} is constructed as follows:

- Define the vertex set V_{st} by adding a source node s and sink node t to the vertex set of G ($V_{st} = V \bigcup \{s, t\}$).
- Let V^+ denote the set of vertices in G with a nonnegative cost. Let V^- denote the set of vertices with a negative cost. Define the edge set E_{st} by adding

```
INITIALIZE-PREFLOW$(G, s)$
for each vertex $u \in V[G]$
    do $h[u] \leftarrow 0$
       $e[u] \leftarrow 0$
for each edge $(u, v) \in E[G]$
    do $f[u, v] \leftarrow 0$
       $f[v, u] \leftarrow 0$
$h[s] \leftarrow |V[G]|$
for each vertex $u \in Adj[s]$
    do $f[s, u] \leftarrow c(s, u)$
       $f[u, s] \leftarrow -c(s, u)$
       $e[u] \leftarrow c(s, u)$
       $e[s] \leftarrow e[s] - c(s, u)$
```

Algorithm A.7: INITIALIZE-PREFLOW(G, s). Copied with permission from [67].

```
PUSH$(u, v)$
▷ Applies when: u is overflowing, $c_f(u, v) > 0$, and $h[u] = h[v] + 1$.
▷ Action: Push $d_f(u, v) = \min(e[u], c_f(u, v))$ units of flow from $u$ to $v$.
$d_f(u, v) \leftarrow \min(e[u], c_f(u, v))$
$f[u, v] \leftarrow f[u, v] + d_f(u, v)$
$f[v, u] \leftarrow -f[u, v]$
$e[u] \leftarrow e[u] - d_f(u, v)$
$e[v] \leftarrow e[v] + d_f(u, v)$
```

Algorithm A.8: PUSH(u, v). Copied with permission from [67].

```
RELABEL$(u)$
▷ Applies when: $u$ is overflowing and for all $v \in V$ such that $(u, v) \in E_f$,
                we have $h[u] \le h[v]$.
▷ Action: Increase the height of $u$.
$h[u] \leftarrow 1 + \min\{h[v] : (u, v) \in E_f\}$
```

Algorithm A.9: RELABEL(u). Copied with permission from [67].

new edges sets E_{s-} and E_{s+} to the edge set E ($E_{st} = E \bigcup E_{s-} \bigcup E_{t+}$). E_{s-} is defined by adding edges from the source to each of the vertices in V^- ($E_{s-} = \{(s, v) \mid v \in V^-\}$). Each edge $(s, v) \in E_{s-}$ is given a a capacity of $-w(v)$. Similarly, E_{t+} is defined by adding edges from each of the vertices in V^+ to the sink t ($E_{t+} = \{(v, t) \mid v \in V^+\}$). Each edge $(v, t) \in E_{t+}$ is given a capacity of $w(v)$. A capacity of infinity is assigned to each edge in the original edge set E.

With this construction, the source set (without the source s) of the minimum cut corresponds to the minimum-cost closed set.

REFERENCES

[1] D. Huang, E. A. Swanson, C. P. Lin, J. S. Schuman, W. G. Stinson, W. Chang, M. R. Hee, T. Flotte, K. Gregory, and C. A. Puliafito, “Optical coherence tomography,” *Science*, vol. 254, no. 5035, pp. 1178–1181, Nov. 1991.

[2] M. Wojtkowski, R. Leitgeb, A. Kowalczyk, T. Bajraszewski, and A. F. Fercher, “In vivo human retinal imaging by Fourier domain optical coherence tomography,” *Journal of Biomedical Optics*, vol. 7, no. 3, pp. 457–463, July 2002.

[3] D. R. Thedens, D. J. Skorton, and S. R. Fleagle, “Methods of graph searching for border detection in image sequences with applications to cardiac magnetic resonance imaging,” *IEEE Transactions on Medical Imaging*, vol. 14, no. 1, pp. 42–55, 1995.

[4] J. S. Suri, K. Liu, S. Singh, S. N. Laxinarayan, X. Zeng, and L. Reden, “Shape recovery algorithms using level sets in 2-D/3-D medical imagery: A state-of-the-art review,” *IEEE Transactions on Information Technology in Biomedicine*, vol. 6, no. 1, pp. 8–28, 2002.

[5] T. McInerney and D. Terzoppoulos, “Deformable models in medical image analysis: A survey,” *Medical Image Analysis*, vol. 1, no. 2, pp. 91–108, 1996.

[6] K. Li, X. Wu, D. Z. Chen, and M. Sonka, “Optimal surface segmentation in volumetric images – a graph-theoretic approach,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 1, pp. 119–134, 2006.

[7] X. Wu and D. Z. Chen, “Optimal net surface problems with applications,” in *Proceedings of the 29th International Colloquium on Automata, Languages, and Programming (ICALP), LNCS 2380.* Springer-Verlag, 2002, pp. 1029–1042.

[8] The Eye Diseases Prevalence Research Group, “Causes and prevalence of visual impairment among adults in the United States,” *Archives of Ophthalmology*, vol. 122, no. 4, pp. 477–485, Apr. 2004.

[9] Prevent Blindness America and the National Eye Institute, “Vision problems in the U.S.” 2002. [Online]. Available: http://www.nei.nih.gov/eyedata/

[10] T. A. Ciulla, C. D. Regillo, and A. Harris, Eds., *Retina and Optic Nerve Imaging.* Philadelphia, PA: Lippincott Williams and Wilkins, 2003.

[11] B. Hermann, E. J. Fernández, A. Unterhuber, H. Sattmann, A. F. Fercher, W. Drexler, P. M. Prieto, and P. Artal, “Adaptive-optics ultrahigh-resolution

optical coherence tomography," *Optics Letters*, vol. 29, no. 18, pp. 2142–2144, Sept. 2004.

[12] W. Drexler, U. Morgner, R. K. Ghanta, F. X. Kärtner, J. S. Schuman, and J. G. Fujimoto, "Ultrahigh-resolution ophthalmic optical coherence tomography." *Nature Medicine*, vol. 7, no. 4, pp. 502–507, Apr. 2001.

[13] L. A. Remington, *Clinical Anatomoy of the Visual System*, 2nd ed. St. Louis, MO: Elsevier Inc., 2005.

[14] H. Kolb, "How the retina works," *American Scientist*, vol. 91, no. 1, pp. 28–35, 2003.

[15] H. Kolb, E. Fernandez, R. Nelson, and B. W. Jones, "Webvision: The organization of the retina and visual system," 2005. [Online]. Available: http://webvision.med.utah.edu/

[16] G. Wollstein, L. A. Paunescu, T. H. Ko, J. G. Fujimoto, A. Kowalevicz, I. Hartl, S. Beaton, H. Ishikawa, C. Mattox, O. Singh, J. Duker, W. Drexler, and J. S. Schuman, "Ultrahigh-resolution optical coherence tomography in glaucoma," *Ophthalmology*, vol. 112, no. 2, pp. 229–237, 2005.

[17] D. Koozekanani, K. Boyer, and C. Roberts, "Retinal thickness measurements in optical coherence tomography using a Markov boundary model," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, vol. 2, June 2000, pp. 363–370.

[18] ——, "Retinal thickness measurements from optical coherence tomography using a Markov boundary model," *IEEE Transactions on Medical Imaging*, vol. 20, no. 9, pp. 900–916, 2001.

[19] H. Ishikawa, D. M. Stein, G. Wollstein, S. Beaton, J. G. Fujimoto, and J. S. Schuman, "Macular segmentation with optical coherence tomography," *Investigative Ophthalmology & Visual Science*, vol. 46, no. 6, pp. 2012–2017, June 2005.

[20] A. Chan, J. S. Duker, H. Ishikawa, T. H. Ko, J. S. Schuman, and J. G. Fujimoto, "Quantification of photoreceptor layer thickness in normal eyes using optical coherence tomography." *Retina*, vol. 26, no. 6, pp. 655–660, 2006.

[21] D. Cabrera Fernández, "Delineating fluid-filled region boundaries in optical coherence tomography images of the retina." *IEEE Transactions on Medical Imaging*, vol. 24, no. 8, pp. 929–945, Aug. 2005.

[22] D. Cabrera Fernández, H. M. Salinas, and C. A. Puliafito, “Automated detection of retinal layer structures on optical coherence tomography images,” *Optics Express*, vol. 13, no. 25, pp. 10 200–10 216, 2005.

[23] M. Shahidi, Z. Wang, and R. Zelkha, “Quantitative thickness measurement of retinal layers imaged by optical coherence tomography,” *American Journal of Ophthalmology*, vol. 139, no. 6, pp. 1056–1061, June 2005.

[24] M. Baroni, P. Fortunato, and A. L. Torre, “Towards quantitative analysis of retinal features in optical coherence tomography,” *Medical Engineering & Physics*, vol. 29, no. 4, pp. 432–441, May 2007.

[25] R. Malladi, J. A. Sethian, and B. C. Vemuri, “Shape modeling with front propagation: A level set approach,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 2, pp. 158–175, 1995.

[26] J. A. Sethian, *Level Set Methods and Fast Marching Methods*, 2nd ed. New York, NY: Cambridge University Press, 1999.

[27] N. Paragios and R. Deriche, “Geodesic active regions and level set methods for supervised texture segmentation,” *International Journal of Computer Vision*, vol. 46, no. 3, pp. 223–247, 2002.

[28] S. Osher and R. Fedkiw, *Level Set Methods and Dynamic Implicit Surfaces*. New York, NY: Springer-Verlag, 2003.

[29] R. Whitaker, “Isosurfaces and level sets,” in *Insight into Images*. Wellesey, MA: A K Peters, 2004, ch. 8, pp. 193–217.

[30] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, “Active shape models — their training and application,” *Computer Vision and Image Understanding*, vol. 61, no. 1, pp. 38–59, 1995.

[31] T. F. Cootes, G. J. Edwards, and C. J. Taylor, “Active appearance models,” in *Proceedings of the 5th European Conference on Computer Vision (ECCV ’98) – Volume II*. London, UK: Springer-Verlag, 1998, pp. 484–498.

[32] T. Cootes, G. Edwards, and C. Taylor, “Active appearance models,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 681–685, June 2001.

[33] S. Mitchell, J. Bosch, B. Lelieveldt, R. van der Geest, J. Reiber, and M. Sonka, “3-D active appearance models: segmentation of cardiac MR and ultrasound

images," *IEEE Transactions on Medical Imaging*, vol. 21, no. 9, pp. 1167–1178, Sept. 2002.

[34] R. J. Frank, "Optimal surface detection using multi-dimensional graph search: Applications to intravascular ultrasound," Master's thesis, The University of Iowa, 1996.

[35] Y. Y. Boykov and M.-P. Jolly, "Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images," in *Proceedings of the Eighth IEEE International Conference on Computer Vision (ICCV)*, vol. 1. IEEE Computer Society, 2001, pp. 105–112.

[36] M. Haeker, X. Wu, M. D. Abràmoff, R. Kardon, and M. Sonka, "Incorporation of regional information in optimal 3-D graph search with application for intraretinal layer segmentation of optical coherence tomography images," in *Information Processing in Medical Imaging (IPMI)*, ser. Lecture Notes in Computer Science, vol. 4584. Springer, 2007, pp. 607–618.

[37] M. Haeker, M. D. Abràmoff, X. Wu, R. Kardon, and M. Sonka, "Use of varying constraints in optimal 3-D graph search for segmentation of macular optical coherence tomography images," in *Proceedings of the 10th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI 2007)*, ser. Lecture Notes in Computer Science, vol. 4791. Springer-Verlag, 2007, pp. 244–251.

[38] M. Haeker, M. Sonka, R. Kardon, V. A. Shah, X. Wu, and M. D. Abràmoff, "Automated segmentation of intraretinal layers from macular optical coherence tomography images," in *Proceedings of SPIE Medical Imaging 2007: Image Processing*, vol. 6512. SPIE, 2007.

[39] M. K. Garvin, M. D. Abràmoff, R. Kardon, S. R. Russell, X. Wu, and M. Sonka, "Intraretinal layer segmentation of macular optical coherence tomography images using optimal 3-D graph search," *IEEE Transactions on Medical Imaging*, 2008, accepted.

[40] M. Haeker, M. D. Abràmoff, R. Kardon, and M. Sonka, "Segmentation of the surfaces of the retinal layer from OCT images," in *Proceedings of the 9th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI 2006), Part I*, ser. Lecture Notes in Computer Science, vol. 4190. Springer-Verlag, 2006, pp. 800–807.

[41] Y. Yu and S. T. Acton, "Speckle reducing anisotropic diffusion," *IEEE Transactions on Image Processing*, vol. 11, no. 11, pp. 1260–1270, 2002.

[42] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 266–277, 2001.

[43] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Transactions on Medical Imaging*, vol. 16, no. 2, pp. 187–198, Apr. 1997.

[44] R. A. Costa, M. Skaf, L. A. S. Melo, D. Calucci, J. A. Cardillo, J. C. Castro, D. Huang, and M. Wojtkowski, "Retinal assessment using optical coherence tomography," *Progress in Retinal and Eye Research*, vol. 25, no. 3, pp. 325–353, May 2006.

[45] G. Donato and S. Belongie, "Approximate thin plate spline mappings," in *Proceedings of the 7th European Conference on Computer Vision (ECCV 2002), Part III, LNCS 2352*, A. Heyden, G. Sparr, M. Nielsen, and P. Johansen, Eds. Springer-Verlag Berlin Heidelberg, 2002, pp. 21–31.

[46] M. Niemeijer, M. K. Garvin, B. van Ginneken, M. Sonka, and M. D. Abràmoff, "Vessel segmentation in 3D spectral OCT scans of the retina," in *Proceedings of SPIE Medical Imaging 2008: Image Processing*, vol. 6914. SPIE, 2008.

[47] M. Niemeijer, M. Sonka, M. K. Garvin, B. van Ginneken, and M. D. Abramoff, "Automated segmentation of the retinal vasculature in 3D optical coherence tomography images," *Investigative Ophthalmology & Visual Science*, pp. E–Abstract 1832, 2008.

[48] D. C. Hood, S. C. Anderson, M. Wall, and R. H. Kardon, "Structure versus function in glaucoma: An application of a linear model," *Investigative Ophthalmology & Visual Science*, vol. 48, no. 8, pp. 3662–3668, 2007.

[49] D. C. Hood, S. C. Anderson, J. Rouleau, A. S. Wenick, L. K. Grover, M. M. Behrens, J. G. Odel, A. G. Lee, and R. H. Kardon, "Retinal nerve fiber structure versus visual field function in patients with ischemic optic neuropathy: A test of a linear model," *Ophthalmology*, vol. 115, no. 5, pp. 904–910, May 2008.

[50] D. C. Hood and R. H. Kardon, "A framework for comparing structural and functional measures of glaucomatous damage," *Progress in Retinal and Eye Research*, vol. 26, no. 6, pp. 688–710, 2007.

[51] A. Heijl, A. Lindgren, and G. Lindgren, "Test-retest variability in glaucomatous visual fields," *American Journal of Ophthalmology*, vol. 108, no. 2, pp. 130–135, Aug. 1989.

[52] D. B. Henson, S. Chaudry, P. H. Artes, E. B. Faragher, and A. Ansons, "Response variability in the visual field: comparison of optic neuritis, glaucoma, ocular hypertension, and normal eyes," *Investigative Ophthalmology & Visual Science*, vol. 41, no. 2, pp. 417–421, Feb. 2000.

[53] M. Wall, C. A. Johnson, K. E. Kutzko, R. Nguyen, C. Brito, and J. L. Keltner, "Long- and short-term variability of automated perimetry results in patients with optic neuritis and healthy subjects," *Archives of Ophthalmology*, vol. 116, no. 1, pp. 53–61, Jan. 1998.

[54] E. B. Werner, B. Petrig, T. Krupin, and K. I. Bishop, "Variability of automated visual fields in clinically stable glaucoma patients," *Investigative Ophthalmology & Visual Science*, vol. 30, no. 6, pp. 1083–1089, June 1989.

[55] M. K. Garvin, M. Sonka, R. H. Kardon, X. Wu, Y. H. Kwon, S. R. Russell, and M. D. Abramoff, "Three-dimensional analysis of SD OCT: Thickness assessment of six macular layers in normal subjects," *Investigative Ophthalmology & Visual Science*, pp. E–Abstract 1879, 2008.

[56] M. Sonka, V. Hlavac, and R. Boyle, *Image Processing, Analysis, and Machine Vision*, 3rd ed. Toronto, Canada: Thomson Engineering, 2007, (1st edition Chapman and Hall, London, 1993; 2nd edition PWS Pacific Grove, CA, 1997).

[57] Y. Xu, M. Sonka, G. McLennan, J. Guo, and E. A. Hoffman, "MDCT-based 3-D texture classification of emphysema and early smoking related lung pathologies," *IEEE Transactions on Medical Imaging*, vol. 25, no. 4, pp. 464–475, Apr. 2006.

[58] M. D. Abràmoff, W. L. M. Alward, E. C. Greenlee, L. Shuba, C. Y. Kim, J. H. Fingert, and Y. H. Kwon, "Automated segmentation of the optic disc from stereo color photographs using physiologically plausible features," *Investigative Ophthalmology & Visual Science*, vol. 48, no. 4, pp. 1665–1673, Apr. 2007.

[59] S. Grigorescu, N. Petkov, and P. Kruizinga, "Comparison of texture features based on Gabor filters," *IEEE Transactions on Image Processing*, vol. 11, no. 10, pp. 1160–1167, 2002.

[60] Z. Qian, D. N. Metaxas, and L. Axel, "Extraction and tracking of MRI tagging sheets using a 3D Gabor filter bank," in *28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS 2006)*, Aug. 2006, pp. 711–714.

[61] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 3, no. 6, pp. 610–621, 1973.

[62] M. Niemeijer, J. Staal, B. van Ginneken, M. Loog, and M. Abràmoff, "Comparative study of retinal vessel segmentation methods on a new publicly available database," in *Proceedings of SPIE Medical Imaging 2004: Image Processing*, vol. 5370. SPIE, 2004, pp. 648–656.

[63] Y. H. Kwon, Y. I. Kim, M. L. Pereira, P. R. Montague, M. B. Zimmerman, and W. L. Alward, "Rate of optic disc cup progression in treated primary open-angle glaucoma," *Journal of Glaucoma*, vol. 12, no. 5, pp. 409–416, 2003.

[64] M. D. Abrámoff, W. L. Alward, E. C. Greenlee, L. Shuba, C. Y. Kim, J. H. Fingert, and Y. H. Kwon, "Automated segmentation of the optic disc from stereo color photographs using physiologically plausible features." *Investigative Ophthalmology & Visual Science*, vol. 48, pp. 1665–1673, 2007.

[65] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*, 2nd ed. Cambridge, MA: The MIT Press, 2001.

[66] ——, "Single-source shortest paths," in *Introduction to Algorithms*, 2nd ed. Cambridge, MA: The MIT Press, 2001, ch. 24, pp. 580–619.

[67] ——, "Maximum flow," in *Introduction to Algorithms*, 2nd ed. Cambridge, MA: The MIT Press, 2001, ch. 26, pp. 643–698.

[68] E. L. Lawler, *Combinatorial Optimization: Networks and Matroids.* New York, NY: Holt, Rinehart, and Winston, 1976.

[69] Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 9, pp. 1124–1137, 2004.

[70] J.-C. Picard, "Maximal closure of a graph and applications to combinatorial problems," *Management Science*, vol. 22, no. 11, pp. 1268–1272, 1976.

[71] D. S. Hochbaum, "A new - old algorithm for minimum-cut and maximum-flow in closure graphs," *Networks*, vol. 37, no. 4, pp. 171–193, 2001.