Spring 2017

Multiple surface segmentation using novel deep learning and graph based methods

Abhay Shah
University of Iowa

MULTIPLE SURFACE SEGMENTATION USING NOVEL DEEP LEARNING AND GRAPH BASED METHODS

by

Abhay Shah

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

May 2017

Thesis Supervisor: Professor Xiaodong Wu

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

PH.D. THESIS

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

Abhay Shah

has been approved by the Examining Committee for the thesis requirement for the Doctor of Philosophy degree in Electrical and Computer Engineering at the May 2017 graduation.

Thesis Committee: _____
Xiaodong Wu, Thesis Supervisor

_____
Michael D. Abrámoff

_____
Mona K. Garvin

_____
Mathews Jacob

_____
Punam Saha

ACKNOWLEDGEMENTS

Six years of persistence, five years of research, five exciting seasons of playing cricket, six seasons of roller coaster hawkeye football (one of them being a perfect season), tying the knot (Indian term for getting married) with my high school sweetheart, belief in my abilities, consistent support and supervision from my advisor, epitomizes my dream of completing a doctoral research thesis. This has been nothing less than a journey of self actualization and adventure. Now, finally I get the chance of writing my PhD thesis to realize my dream and sincerely acknowledge the help, support, encouragement and guidance I received, without which completing my PhD would never have been a reality.

First and foremost, I would like to thank Dr. Wu for being my research advisor since 2012. He taught me the art of solving a research problem and the thought process required for the same. I learnt persistence, work ethic, precision and the craft of conducting research to attain the set goals from him. His support and guidance, always helped me to stay positive and to never give up during time of failure in research but to treat them as an opportunity to learn from it. He has been a true mentor and I have been fortunate to have had the opportunity to work with him on interesting real world applications/methods and learn many skills and attributes which have been valuable in both my academic and personal life. My sincere gratitude to him for being my doctoral research thesis supervisor.

Next, I would like to thank Dr. Abramoff, Dr. Garvin, Dr. Jacobs and Dr. Saha for serving on my dissertation committee. It has been my honour and pleasure to have you all in my dissertation committee and I sincerly thank you for your invaluable guidance throughout this process, your discussions, ideas, and feedback. I have been fortunate enough to work with most of you during my research work. Many thanks to Dr. Abramoff for working closely with us on interesting research problems. He has the knack of bridging the gap between clinical background and applications with

engineering solutions, and allowed me to learn how to approach such problems. The most memorable thing I imbibed from him, is his positive attitude and the excitement of solving a research problem. I have had the previlege to work with Dr. Garvin and Dr. Jacobs on research projects as well. Working with you, I learnt about the different areas of applications of methods developed in my research work. I am glad, that I could contribute my efforts in these projects.

I have been fortunate to learn from the awesome professors in the ECE department. I sincerely would like to thank them for the same. Gratitude to Cathy Kern and Dina Blanc, our department secretaries for all their help throughout these six years.

I am very grateful to IDX for giving me the opportunity to intern with them over the last two summers. My sincerest gratitude to Meindert, Ryan and Warren, for all the intellectual discussions, training and a great learning experience at IDX. You helped me to learn and apply deep learning to practical problems which has been a pivotal factor in allowing me to integrate the same to my research work.

I would like to thank all my friends here in Iowa City for all the favours, support, encouragement, discussions and good times throughout these years. Your kindness and tolerance have strengthened me every single day.

I would like to dedicate this dissertation to my parents. Thank you, mom and dad for believing in me and always encouraging me to find my own path in life. I can never express in enough words how thankful I am for all your support, sacrifices and love. I love you both and hope this precious milestone in my life, brings you pride. Finally, my wife Akanksha, thank you for being there. I cannot express enough my love and gratitude to you, for putting up with a graudate student and for your uncondional love. You have been an unwavering pillar of support throughout my times of success and failure. You have always motivated me and brightened my life. This honour and success is equally yours.

ABSTRACT

The task of automatically segmenting 3-D surfaces representing object boundaries is important in quantitative analysis of volumetric images, which plays a vital role in numerous biomedical applications. For the diagnosis and management of disease, segmentation of images of organs and tissues is a crucial step for the quantification of medical images. Segmentation finds the boundaries or, limited to the 3-D case, the surfaces, that separate regions, tissues or areas of an image, and it is essential that these boundaries approximate the true boundary, typically by human experts, as closely as possible. Recently, graph-based methods with a global optimization property have been studied and used for various such segmentation applications. Sepecifically, the state-of-the-art graph search (optimal surface segmentation) method has been successfully used for various such biomedical applications. Despite their widespread use for image segmentation, real world medical image segmentation problems often pose difficult challenges, wherein graph based segmentation methods in its purest form may not be able to perform the segmentation task successfully.

This doctoral work has a twofold objective. 1)To identify medical image segmentation problems which are difficult to solve using existing graph based method and develop novel methods by employing graph search as a building block to improve segmentation accuracy and efficiency. 2) To develop a novel multiple surface segmentation strategy using deep learning which is more computationally efficient and generic than the exisiting graph based methods, while eliminating the need for human expert intervention as required in the current surface segmentation methods. This developed method is possibly the first of its kind where the method does not require and human expert designed operations.

To accomplish the objectives of this thesis work, frameworks of graph based and deep learning methods is proposed to achieve the goal by successfully fulfilling the following three aims. First, an efficient, automated and accurate graph based method is developed to segment surfaces which have steep change in surface profiles and abrupt

distance changes between two adjacent surfaces. The developed method is applied and validated on intra-retinal layer segmentation of Spectral Domain Optical Coherence Tomograph (SD-OCT) images of eye with Glaucoma, Age Related Macular Degneration and Pigment Epithelium Detachment. Second, a globally optimal graph based method is developed to attain subvoxel and super resolution accuracy for multiple surface segmentation problem while imposing convex constraints. The developed method was applied to retinal layer segmentation of SD-OCT images of normal eye and vessel walls in Intravascular Ultrasound (IVUS) images. Third, a deep learning based multiple surface segmentation is developed which is more generic, computaionally effieient and eliminates the requirement of human expert interventions (like transformation designs, feature extraction, parameter tuning, constraint modelling etc.) required by existing surface segmentation methods in varying capacities. The developed method was applied to SD-OCT images of normal and diseased eyes, to validate the superior segmentaion performance, computation efficieny and the generic nature of the framework, compared to the state-of-the-art graph search method.

PUBLIC ABSTRACT

For the diagnosis and management of disease, segmentation of images of organs and tissues is a crucial step for the quantification of medical images. Segmentation finds the boundaries/surfaces, that separate regions, tissues or areas of an image, and it is essential that these boundaries approximate the true boundary. Recently, graph-based methods (specifically graph search/optimal surface segmetnation) with a global optimization property have been studied and used for various such biomedical applications. Despite their widespread use for image segmentation, real world medical image segmentation problems often pose difficult challenges, wherein graph based segmentation methods may not be able to perform the segmentation task successfully.

In this doctoral thesis, novel frameworks of graph based and deep learning methods are proposed to accomplish the task of multiple surface segmentation. The devloped methods tackle various challenges posed in real work medical imaging applications where the target surfaces to be segmented are complex due to presence of pathologies. The presented frameworks achieve higher segmentation accuracy compared to graph search methods for such complex surface segmentation problems, and allows for sub-voxel and super resolution accurate surface segmentations. The developed novel deep learning based multiple surface segmentation method provides for a more generic and computationally efficient framework, wherein a single network is capable of inferring on multiple surface segmentations for both normal and diseased cases and thereby, makes the framework different in principal as compared to graph based methods in terms of elimination of human expert interventions (like transformation designs, feature extraction, parameter tuning, constraint modelling etc.). The deep learning based method is possibly the first of its kind where the method does not require and human expert designed operations for surface segmentation applications. The developed methods have been extensively compared to the existing state-of-the-art graph based mehtod and validated on various intra-retinal layer segmentation applications in Optical Coherence Tomography (OCT) images of the eye.

TABLE OF CONTENTS

LIST OF TABLES x

LIST OF FIGURES xii

CHAPTER

1 INTRODUCTION 1

1.1 Thesis Organization 3

2 GRAPH SEARCH AND CONVOLUTION NEURAL NETS 5

2.1 Graph Search (Optimal Surface Segmentation) 5
2.2 Convolution Neural Nets (CNNs) 8
2.3 Related Work for Retinal Layer Segmentation in OCT 12

3 MOTIVATION AND SPECIFIC AIMS 14

3.1 Motivation 14
3.1.1 Segmentation Problems Difficult to Solve Using Graph Search 14
3.1.2 Need for a Generic Surface Segmentation Method Without Requirement of Human Expert Intervention 18
3.2 Specific Aims 20

4 MULTIPLE SURFACE SEGMENTATION USING TRUNCATED CONVEX PRIORS (AIM 1) 23

4.1 Introduction 23
4.2 Single Surface Segmentation Using Truncated Convex Priors (Aim 1.1) 24
4.2.1 Method Design 25
4.2.2 Method Overview 26
4.2.3 Graph Construction 28
4.2.4 Experiment Setting 30
4.2.5 Results 31
4.3 Multiple Surface Segmentation Using Truncated Convex Priors (Aim 1.2) 32
4.3.1 Method Design 35
4.3.2 Method Overview 36
4.3.3 Graph Construction 39
4.3.4 Experiment Setting 41
4.3.5 Results 43

4.4 Segmentation of Surfaces in SD-OCT Volumes with Retinal Pigment Epithelial Detachment (PED) Using Truncated Convex Priors(Aim 1.3) . . . 44
4.4.1 Experiment Setting . . . 46
4.4.2 Results . . . 48
4.5 Discussion . . . 49
4.6 Conclusion . . . 51

5 OPTIMAL MULTIPLE SURFACE SEGMENTATION WITH CONVEX PRIORS IN IRREGULARLY SAMPLED SPACE (AIM 2) . . . 53

5.1 Introduction . . . 53
5.2 Optimal Multiple Surface Segmentation with Convex Priors in Irregularly Sampled Space (Aim 2.1) . . . 57
5.2.1 Problem Formulation and Energy Function . . . 57
5.2.2 Graph Construction . . . 61
5.2.3 Surface Recovery from Minimum s-t cut . . . 68
5.3 Proof of Correctness of Graph Construction (Aim 2.2) . . . 69
5.3.1 Proof for Lemma 1 . . . 69
5.3.2 Proof for Lemma 2 . . . 72
5.4 Validation for Subvoxel Accuracy and Super Resolution Segmentation Accuracy (Aim 2.3) . . . 77
5.4.1 Experiment Setting for SD-OCT Volumes of Normal Eye 77
5.4.2 Segmentation Results for SD-OCT Volumes of Normal Eye 83
5.4.3 Experiment Setting for IVUS Images . . . 89
5.4.4 Segmentation Results of Lumen and Media in IUVS Images for Subvoxel Accuracy . . . 94
5.5 Discussion . . . 100
5.6 Conclusion . . . 101

6 MULTIPLE SURFACE SEGMENTATION USING DEEP LEARNING (AIM 3) . . . 102

6.1 Introduction . . . 102
6.2 Motivation from Graph Search for CNN Based Surface Segmentation 103
6.3 Single Surface Segmentation Using CNNs (Aim 3.1) . . . 105
6.3.1 Method . . . 105
6.3.2 Network Architecture . . . 107
6.3.3 Experiment Objectives . . . 109
6.3.4 Experiment Setting . . . 110
6.3.5 Results . . . 112
6.4 Multiple Surface Segmentation Using CNNs (Aim 3.2) . . . 114
6.4.1 Method . . . 115
6.4.2 Network Architecture . . . 115
6.4.3 Experiments . . . 116

6.4.4 Results 118
6.5 Discussion 119
6.6 Conclusion 123

7 CONCLUSIONS 124

REFERENCES 128

LIST OF TABLES

Table

4.1 Results from truncated convex method and graph search method. The UMSP and UASSD errors, are expressed as mean ± standard deviation in μm. 32

4.2 Parameters used in the experiment. M_x and M_y are truncation factors in x and y directions, M is the truncation factor for the surface separation term of a surface pair, d is the minimum separation between a surface pair. TC - Truncated Convex. 42

4.3 Unsigned mean surface positioning errors (UMSP) (mean ± standard deviation)μm for downsampled version of data. Obsv - Expert manual tracing, TC -Truncated Convex, OSDC - Optimal Surface Detection with Convex priors. 44

4.4 Unsigned mean surface positioning errors (UMSP) (mean ± standard deviation)μm for data in original resolution. Obsv - Expert manual tracing, TC -Truncated Convex. 44

4.5 Parameters used in the experiment. Δ_x and Δ_y are the hard constraints used in x and y directions, M_x and M_y are truncation factors in x and y directions, M is the truncation factor for the surface separation term of a surface pair, d is the minimum separation between a surface pair. 47

4.6 Unsigned mean surface positioning errors (UMSP) (mean ± standard deviation)μm. Obsv - Expert manual tracing. 48

5.1 Summary of inter-column edge weights of the graph construction in Fig. 5.3, based on a quadratic convex function of the form $\psi(k_1 - k_2) = (k_1 - k_2)^2$. 67

5.2 Unsigned mean surface positioning error (UMSP) (mean ± standard deviation) in voxels for subvoxel accuracy. Obsv - Subvoxel accurate expert manual tracings. 85

5.3 Unsigned average symmetric surface distance error (UASSD) (mean ± standard deviation) in μm for subvoxel accuracy. Obsv - Expert manual tracings. 85

5.4 Unsigned average symmetric surface distance error (UASSD) (mean ± standard deviation) in μm for super resolution accuracy. Obsv - Expert manual tracings, η - Down-sampling scale in z direction. 89

5.5 Overview of the proposed and compared method features 91

5.6 Evaluation measures of each method with respect to expert manual tracings. Error measures expressed as mean and (standard deviation). An empty table cell indicates that the method was not applied to Lumen or Media. 96

6.1 UMSP error for single surface on testing set from CNN trained for patch size N × Z. 113

6.2 UMSPE expressed as (mean ± 95% CI) in voxels. RS - Reference Standard. A single CNN was trained to infer on each type of data while 3 different set of parameters were used for the G-OSC method. $N = 32$ was used as the patch size (32×512). 121

LIST OF FIGURES

Figure

2.1 (left) Surface smoothness constraint for two neighboring columns p and q for a surface S_i. The arcs in red show the feasible st-cut for a given surface position $S_i(p)$ for column p. (right) Surface separation constraint for two corresponding columns for adjacent surfaces S_i, S_{i+1}. The arcs in red show the feasible st-cut for a given surface position $S_{i+1}(p)$ for column p. 9

2.2 Surface smoothness constraint for two neighboring columns p and q for a surface S_i. The arc in red shows a st-cut for a given surface position $S_i(p)$ for column p. The smoothness cost for the given cut is $w \times \psi(S_i(p) - S_i(q))$. 9

2.3 Illustration of a simple CNN architecture for image classification into 4 classes. [1] 11

3.1 Steep change in surface smoothness can be seen in SD-OCT image of an eye with severe glaucoma (*left*). Abrupt changes in surface separation between surface 2 (S_2) and surface 3 (S_3) can be seen in SD-OCT image of an eye with AMD (*right*). 15

3.2 Example of an image slice with PED. Expert manual tracings shown as Yellow - Internal Limitng Membrane (ILM), Red - Outer Retinal Pigment Epithelium (ORPE), Blue - Inner Choroid (IC). The PED is indicated by the brown arrow (left). The steep change in surface smoothness of ORPE is indicated by the purple arrow and the abrupt change in the surface separation between ORPE and IC is indicated by the green arrow (right). 16

3.3 Illustrative example of non-equidistant spacing between adjoining graph nodes. (a) A voxel column in a 3D image. (b) Representation of voxel column in (a) as graph nodes in the graph space. (c) Shifted positions of the graph nodes after the deformation is applied in the vertical direction to the voxel column. (d) A voxel column in a 3D image with sampled voxels shown in brown boundary (e) Representation of voxel column in (d) as graph nodes in the graph space. 17

4.1 Example for a linear convex penalty (*left*) and a linear truncated convex penalty (*right*). $S(a)$ and $S(b)$ are surface positions for a pair of neighboring columns a and b. M is the truncation factor (bound on the maximum possible penalty). 24

4.2 (Central B-scans (xz slice) from optic nerve head centered scans from (a) subject with less severe glaucoma, (b) subject with more severe glaucoma. 25

4.3 Example of interval construction for a 2D data cost slice D for surface S with 3 neighboring columns and Z=4. Each voxel is represented as a node. Length of the interval (L) = 2. Consecutive interval ($T'_m s$) construction at each iteration m is shown. 27

4.4 Method pipeline for single surface segmentation. N is the maximum index of possible iterations. 27

4.5 (a) Example graph construction for Case1. Additional arcs added to (a) for (b)(c) Symmetric Cases, (d) Combination Case. Length of interval (L) =3. 29

4.6 Examples of comparison between the manual tracings from an expert independent observer (yellow), the segmentation result obtained using truncated convex method (red) and Original Graph Search (green), respectively. 33

4.7 Segmentation result obtained for slices from different OCT volumes using the proposed method(red) and Original Graph Search (green). The manual tracings were not available for these slices. 34

4.8 Pipeline for our method. $N = Z - 2 - \sum_{i=1}^{\lambda-1} d_{(i,i+1)}$. 37

4.9 Example of interval construction for 2-D data cost slice D_i for surface S_i and D_{i+1} for adjacent surface S_{i+1} with three corresponding columns and Z=4. L =2 and minimum separation constraint $d_{i,i+1} = 1$. 38

4.10 (a)An example of graph construction for inter-surface arcs for base case. Additional arcs added to (a) for (b)(c) Symmetric cases, (d) combination case. Length of interval $L = 3$ 40

4.11 Top two rows show image slices of Type II and bottom row shows image slice of Type I. Yellow - ILM, Red - IRPEDC and Blue - OBM. (a)Expert manual tracing, segmentation using (b)our method, (c)OSDC method. . 43

4.12 Each image in a given column shows the same B-scan from a SD-OCT volume with PED. Yellow - ILM, Red - ORPE and Blue - IC. 50

5.1 Example of a 3×3 voxel grid to demonstrate subvoxel accuracy. Each voxel is represented by a red node in the graph space. (a)Graph nodes with equidistant spacing between them. True subvoxel accurate surface is shown in green. The segmented surface using optimal surface segmentation method with hard constraints is shown in yellow. (b) The displacement field derived from the grid is applied to the central nodes displacing the centers to exploit the information from the partial volume effect shown by brown arrows. The resultant segmentation with the subvoxel accurate graph search is shown in blue. 56

5.2 Example of column structure for irregularly sampled space using mapping function. 60

5.3 Example graph construction of two neighboring columns a and b to demonstrate enforcement of convex surface smoothness constraints in irregularly sampled space. 65

5.4 An example graph for incorporation of surface separation constraint between two corresponding columns is shown. Only the inter-surface edges are shown for clarity. The minimum separation constraint $d_{i,i+1} = 2$. It can be seen that cut C_1 is a feasible cut since the minimum separation constraint is not violated while cut C_2 is infeasible since the minimum separation constraint is violated as $L_a(z' = 1) - L_a(z = 1) < d_{i,i+1}$ 68

5.5 (a)A single B-scan from a SD-OCT volume of a normal eye, (b) Three identified target surfaces S_1, S_2 and S_3. 78

5.6 Experiment design for segmentation of SD-OCT volumes of normal eye with subvoxel accuracy. 81

5.7 Experiment design for segmentation of SD-OCT volumes of normal eye for super resolution accuracy. 81

5.8 Unsigned mean surface positioning errors observed in 25 volumetric OCT images for subvoxel accuracy. 84

5.9 Unsigned average symmetric surface distance errors observed in 25 volumetric OCT images for subvoxel accuracy. 84

5.10 Illustration of results on a single B-scan from input image volume. Red - ILM expert tracing, Green - IRPEDC expert tracing, Dark blue - OBM expert tracing, Yellow - ILM automated segmentation, Magenta - IRPEDC automated segmentation, Light blue - OBM automated segmentation. . 86

5.11 Illustration of results on a single B-scan from input image volume. Red - ILM expert tracing, Green - IRPEDC expert tracing, Dark blue - OBM expert tracing, Yellow - ILM automated segmentation, Magenta - IRPEDC automated segmentation, Light blue - OBM automated segmentation. . . 87

5.12 Unsigned average symmetric surface distance errors observed in 25 volumetric OCT images for super resolution segmentation accuracy. η is the downsampling scale. 90

5.13 (a) A single frame of an IVUS multiframe dataset (b) Expert manual tracings of the Lumen (red) and Media (green). 90

5.14 (a) A single frame of an IVUS multiframe dataset (b) Polar transformation of (a). Red - Lumen, Green - Media. 93

5.15 Jaccard measure (JM) metric for IVUS data. Higher JM indicates a larger overlap of the automated segmentation and manual segmentations. It can be seen that the proposed method has the highest JM among all fully automated methods and is comparable with the semi-automated method P3. 97

5.16 Percentage of area difference (PAD) metric for IVUS data. Lower PAD indicates a smaller segmentation area difference between the automated and manual segmentations. 98

5.17 Hausdroff distance (HD) metric for IVUS data. Lower HD indicates a closer alignment of the automated and manual segmentations. 98

5.18 Qualitative illustrations of lumen and media segmentation using our method. Each image is a single frame of an IVUS multiframe dataset. Red - Lumen expert tracing, Green - Media expert tracing, Yellow - Lumen segmentation (our method), Blue - Media segmentation (our method). 99

6.1 Illustration of local smoothness based on patches for a single B-scan from an OCT image. The ILM surface is shown in red. It can be seen that the target surface is infact a combination of sufficiently sized local surface segments which exhibit various types of local surface smoothness (profiles). 105

6.2 Illustration of patch extraction from a B-scan. The target surface is shown in red. The extracted patch shown on the right with the corresponding surface to be used as target for training CNN on the given patch. 106

6.3 Illustration of data augmentation applied to an input patch. The target surface is shown in red. (a)Extracted patch from a B-scan (b)Translation (c)Rotation (d)Translation and rotation, as applied to (a). 107

6.4 The architecture of the CNN learned in our work for N=32. The numbers along each side of the cuboid indicate the dimensions of the feature maps. The inside cuboid (*green*) represents the convolution kernel and the inside square (*green*) represents the pooling region size. The number of hidden neurons in the fully connected layers are marked aside. IP=Input Patch, CV-L=Convolution Layer, MP-L=Max-Pooling Layer, FC-L=Fully Connected Layer, E=Euclidean Loss Layer. 108

6.5 Illustration of difference in surface profiles on a single Macula B-scan and ONH B-scan. The ILM surface is shown in red. (left) Macula scan. (right) ONH scan. 110

6.6 (a)-(b) Expert manual tracings and segmentation from the proposed method for the ILM surface on the same macular B-scan is shown in *cyan* and *yellow* respectively. (c) 3-D rendering of the segmented surface using the proposed method for a validation macular OCT volume. 113

6.7 Machine generated truth (from graph search method) and segmentation from the proposed method for the ILM surface is shown in *red* and *yellow* respectively. Top row- B-scans in original resolution. Bottom row- Magnified version of the blue boxes to demonstrate superior quality of segmentation from the proposed method. (a)Macular B-scan from test set. (b)-(c) ONH B-scans. Unlike the graph search method where post processing is applied to the output segmentation, no such post processing is applied to the segmentations obtained from the proposed method. 114

6.8 Illustration of difference in surface profiles on a single B-scan. (left) Normal Eye (right) Eye with AMD. S_1 = ILM and S_2 = IRPE, are shown in red. 115

6.9 The architecture of the CNN learned in our work for N=32 and $\lambda = 2$. The numbers along each side of the cuboid indicate the dimensions of the feature maps. The inside cuboid (*green*) represents the convolution kernel and the inside square (*green*) represents the pooling region size. The number of hidden neurons in the fully connected layers are marked on the side. IP=Input Patch, CV-L=Convolution Layer, MP-L=Max-Pooling Layer, FC-L=Fully Connected Layer, E=Euclidean Loss Layer. . 117

6.10 Illustration of 20 chosen feature maps from the first convolution layer for a given patch shown on the left. The ILM and IRPE are shown in red. . 119

6.11 Each row shows the same B-scan from a Normal or AMD OCT volume. (a) CNN-S vs. RS (b) G-OSC vs. RS, for surfaces S_1 =ILM and S_2 = IRPE. RS = Reference Standard, Red = reference standard, Green = Segmentation using proposed method and Blue = Segmentation using G-OSC method. In the 4th row, we had the reference standard reviewed by a fellow-ship trained retinal specialist, who stated that the CNN-S method is closer to the real surface than the reference standard. 120

CHAPTER 1
INTRODUCTION

Medical images have made a great impact on medicine, diagnosis, and treatment. The most important part of image processing is image segmentation. The task of optimally identifying object boundaries and regions is important in segmentation and quantitative analysis of volumetric medical images. Accurate segmentation of medical images is a key step for diagnosis, maintenance and treatment planning of diseases. The segmentation problems discussed and investigated in this doctoral thesis work are identified as surface segmentation problems. Such problems relate to segmentation of the boundary of an organ, tissue or any object of interest within an image.

Popular graph based segmentation approaches: graph search [2] and graph-cut [3] [4] have been extensively used for various such image segmentation applications. These approaches formulate the segmentation problem as an energy minimization problem wherein each energy term models a specific aspect of the target object segmentation. The framework of these approaches provide for a robust and flexible platform to extend these methods to incorporate numerous aspects of the segmentation problem at hand while maintaining the optimality of the solution. The advantages of using such a formulation are: 1) flexible modeling ability to incorporate various terms like likelihoods, neighbour relationships, prior information, context information and surface/object interaction; 2) graph based methods provide a framework for computing efficient solutions for a large variety of image segmentation problems.

Many, real world medical image segmentation problems often pose difficult challenges, wherein graph based segmentation methods in its purest form may not be able to perform the segmentation task successfully. The problems generally arise due to increased complexity in the target surfaces due to presence of pathology. Hence, traditional graph based approaches are not able to model such surfaces with high accuracy. Furthermore, the accuracy of the segmentation in these methods is limited

to the voxel level and thus, does not allow for subvoxel accuracy in segmentation of the surfaces. However, as discussed before these methods provide a very flexible framework for novel incorpoations of prior informations to aid the segmentation of challenging surface segmentation problems. This work, in part, shall exploit the graph search framework in order to segment certain difficult segmentation problems and extend the framework to allow for subvoxel accuracy.

Also, graph based frameworks require human interaction/intervention in various aspects. Human expert designed transforms, features, region of interests, parameter tuning and modelling of target surfaces is required for efficient and accurate segmentation. Failure to design any of the required components reasonably shall result in poor segmentation results, even though the result may be optimal with respect to the model used in the framework. Furthermore, the framework is exteremely slow and requires substantial memory to process large volumetric images in the original resolution. The complexity of memory requirement and processing time exponentially rises with the increase in number of target surfaces. Deep learning on the other hand, where all transformation levels are determined from training data, instead of being designed by experts [5], has been highly successful in a large number of computer vision and image analysis detection tasks, substantially outperforming all classical image analysis techniques [6], and, given the spatial coherence that is characteristic of images, typically implemented as Convolutional Neural Networks (CNN) [6]. In part, this thesis work shall leverage the observations and structure of the graph search framework to design a novel CNN based approach to perfom multiple surface segmentation tasks, without the requirement of any expert intervention.

The goal of this work is twofold. 1) Identification of segmentation problems which are difficult to segment using the traditionally employed graph based segmentation methods (specifically graph search) and develop novel graph based methods which build upon the existing graph based methods for more accurate and efficient segmen-

tation of such problems with validation on medical image segmentation applications. 2) Develop a deep learning based multiple surface segmentation approach as means for a generic segmentation approach which does not require any human expert designed transforms, features etc and at the same time is more efficient, accurate and reduces complexity of memory requirement compared to graph search method.

In this doctoral work, majority of the applications are based on intra-retinal layer segmentations in Spectral Domain Optical Coherence Tomography (SD-OCT) images of the normal eyes and eyes with pathology. However, the developed methods are generic and are readily applicable to other similar surface segmentation applications.

1.1 Thesis Organization

The rest of this thesis is organized as follows.

- **Chaper 2** breifly discusses technical details regarding the popular graph search framework and Convolution Neural Nets (CNNs). Both these respective frameworks form the building blocks of the developed novel methods for multiple surface segmentation. Further, since majority of the applications in this work are related to segmentation of surfaces in SD-OCT volumes, a brief literature review of simultaneous multiple intra-retinal surface segmentation methods is provided.

- **Chaper 3** identifies segmentation problems which are difficult to solve using the exisiting graph search based methods, explores the need for a more genreic method for multiple surface segmetation, describes the motivations and finally presents the aims for this thesis work.

- **Chaper 4** describes the novel graph based method developed for multiple surface segmentation using truncated convex priors. Details about the method, summary of the applications and experiment results are presented for the simultanoeus intraretinal layer segmentation of specific surfaces in Optical Co-

herence Tomography (OCT) images of normal eye, eye with Glaucoma eye with Age related Macular Degeneration (AMD) and eye with Pigment Epithelium Detachment (PED).

- **Chaper 5** provides the details of the novel method developed for multiple surface segmentation in irregularly sampled space. The chapter summarizes the applications and experiment results of the method for multiple surface segmentation of three important intraretinal layers in OCT images of normal eye and analyzes subvoxel and super resolution accuracy of the segmentations. Additionally, this chapter also summarizes the experiment results for application of the method to segment two membrances in Intra Vascular Ultrasound Images (IVUS) images and compares it with other publicly available methods used on the same data.

- **Chaper 6** describes in detail the developed method for multiple surface segmentation using CNNs. Futher, summary of the experiments, results and analysis for single and multiple surface segmentation in OCT images of both normal and diseased eye is provided.

- **Chaper 7** concludes the previous chapters and discusses the possible limitations and the potential future work.

Each chapter in this thesis consists of various notations and formulations. In order to maintain consistency and ease of understanding for the reader, each chapter related to a specific aim of this thesis shall reformulate the problem mathematically. Further, state-of-the-art graph search method [2] has been used in conjunction with other methods to perform analysis and comparisons with the developed methods in this work. The methods developed in this thesis are generically applicable to similar surface segmentation problems and are not limited to medical image segmentaion applications.

CHAPTER 2
GRAPH SEARCH AND CONVOLUTION NEURAL NETS

2.1 Graph Search (Optimal Surface Segmentation)

Optimal net surface segmentation method [7] is the pioneer work which first introduced a graph based framework to segment multiple surfaces in the N-D space ($N \geq 2$) in polynomial time. The framework used a key observation regarding the column structure in digital volumetric images to generate multi-column graphs, where every voxel is transformed to a node in the graph space. The algorithm constructs various edges between the nodes to encode the modelling of the surface and prior information, such that the solution of the same using a minimum st-cut [8], provides the resultant surfaces. Furthermore, the method guarantees global optimality with respect to the constraints employed. The popular optimal surface segmentation method (graph search) [2] was infact developed using the optimal net surface segmentation method and demonstrated the method's application to various surface segmentation problems. The method is currently regarded as the state-of-the-art method for multiple surface segmentation.

Note- In this work the terms graph search and optimal surface segmentation method shall be used interchangebly, due to the popularity of the former term (graph search) but infact both the methods refer to the same method [2].

Graph search methods [2] [9] are an efficient tool employed for optimal surface segmentation of multiple globally optimal surfaces in volumetric datasets. The state-of-the-art method with a global optimization property has been widely used for various medical image segmentation tasks, such as knee bone and cartilage [10] [11], heart [12] [13], airways and vessels tress [14] [15], lungs [16], liver [17], prostate and bladder [18] and retinal surfaces [19] [20] [21]. The method is capable of segmenting both 'terrain like' and 'closed' surfaces with global optimality with a pre-defined set of constraints. This technique has also been used to segment objects and surfaces

with more complex topologies by making use of various prior information [9] [22] and pre-segmentation of the target object. The method is readily extensible to segment multiple terrain like surfaces in N ($N \geq 2$) dimensions, though in this review, we briefly discuss the method for segmenting multiple surfaces in a 3-D image volume represented as a cube.

Consider a volumetric image $I(x, y, z)$ of size $X \times Y \times Z$. A surface is defined as a function $S(x, y)$, where $x \in \mathbf{x} = \{0, 1, ...X - 1\}$, $y \in \mathbf{y} = \{0, 1, ...Y - 1\}$ and $S(x, y) \in \mathbf{z} = \{0, 1, ...Z - 1\}$. Each (x, y)-pair corresponds to a *column* of voxels $\{(I(x, y, z)|z = 0, 1, \ldots, Z - 1\}$, denoted by $col(x, y)$ parallel to the z-axis. We use p and q to denote two neighboring (x, y)-pairs in the image domain $\mathbf{x} \times \mathbf{y}$ and N_s to denote the neighborhood setting of image domain. The function $S(p)$ can be viewed as labeling for $col(p)$ with the label set $\mathbf{z}$ ($S(p) \in \mathbf{z}$). Thus, the surface function $S(p)$ intersects $col(p)$ at a single voxel location. The method seeks to simultaneously find λ ($\lambda \geq 1$) globally optimal surfaces $S_i(x, y)$, $i = 1, 2 \ldots \lambda$ in I subjected to prior informations and certain pre-defined constraints.

The method encodes certain prior informations into the graph which model the target surfaces. The first is a cost function $\sum_{p \in \mathbf{x} \times \mathbf{y}} D_i(S_i(p))$, which measures the total cost of all voxels on a surface S_i. An optimal surface is the surface with the minimum cost among all the feasible surfaces definable in the 3-D volume. Second, and the feasibility of the surface with respect to certain surface smoothness and surface separation constraints.

The surface smoothness $\sum_{(p,q) \in N_s} V_{pq}(S_i(p), S_i(q))$ as shown in Eqn 2.1, controls the feasibilty of surface positions in two neighboring columns. The surface smoothness term herein is modelled as a *hard constraint* and specifies the maximum possible difference between surface positions of two neighboring columns.

$$V_{ab}(S_i(p), S_i(q)) = \begin{cases} \infty, & \text{if } |S_i(p) - S_i(q)| > \Delta_{pq}, \\ 0\ , & \text{otherwise} \end{cases} \tag{2.1}$$

where Δ_{pq} is the hard constraint imposed for surface smoothness.

The surface separation term $H_p(S_i(p), S_{i+1}(p))$ as shown in Eqn 2.2, similarly incorporates a hard constraint for the allowed minimum and maximum separation between two adjacent surfaces.

$$H_p(S_i(p), S_{i+1}(p)) = \begin{cases} \infty, & \text{if } (S_{i+1}(p) - S_i(p)) < \delta_{min}, \\ \infty, & \text{if } (S_{i+1}(p) - S_i(p)) > \delta_{max}, \\ 0\ , & \text{otherwise} \end{cases} \tag{2.2}$$

where δ_{min} and δ_{max} are the allowable minimum and maximum hard constraint. Example of the surface smoothness and surface separation constraints are shown in Fig 2.1.

λ subgraphs $G_1, \ldots G_i, G_{i+1}, \ldots G_\lambda$ are constructed where each voxel in the image is represented by a node in the subgraph G_i for a given surface S_i. The method enables the incorporation of the following prior information. First, the likelihood of a voxel belonging to a given surface S_i where the inverse likelihood for each voxel is encoded in the graph and is termed as the data cost. Second, the smoothness of a given surface S_i which is defined as the maximum allowed jump of any two adjacent voxels on a feasible surface in a given direction and is commonly termed as hard constraints. These constraints are incorporated by adding inter-column edges between nodes of a pair of neighboring columns in graph G_i for surface S_i. Lastly, the separation between a given pair of two adjacent surfaces which is also modelled as a hard constraint. This information is encoded by adding inter-surface edges between

the corresponding columns (columns formed on the same (x, y) pair) of graphs G_i and G_{i+1} for two adjacent surfaces S_i and S_{i+1}.

The graph G for the simultaneous segmentation of λ surfaces is constructed by the union of the λ subgraphs G_i's with the addition of the inter-surface edges as discussed above. A single minimum st-cut is then computed on graph G to obtain the target surface segmentations S_i's.

The hallmark of the method is the allowed flexibility in modelling the surface smoothness and surface separation term as shown in Eqn 2.1 and 2.2 respectively. For instance, the surface smoothness term can be modified as shown in Eqn 2.3 to impose a convex penalty based on the difference in surface position of the two columns [2] , and the co-efficient w can be used as tuning parameter to control the degree of regularization of the surface (provides for a balance between the data cost term and the surface smoothness term). Example of surface smoothness constraint with convex penalty is shown in Fig 2.2. The surface smoothness term can similarly be modified to incorporate prior information [9] [22].

$$V_{ab}(S_i(p), S_i(q)) = w \times \psi(S_i(p) - S_i(q)) \tag{2.3}$$

where $w > 0$ is the weight co-efficent and $\psi(.)$ is a convex function.

2.2 Convolution Neural Nets (CNNs)

Convolution neural netwrork (CNN) [23] [24] [6] has recently gained popularity and has been successfully employed for a variety of computer vision problems. A generic neural network is generally organized into at least three layers of neurons: the input layer, the hidden layer, and the output layer. For a fully connected network, each neuron is connected to the neurons from the preceding layer and takes in as input the values of them. Let $x_1, x_2, \ldots, x_N$ denote the inputs to a neuron z_q. The

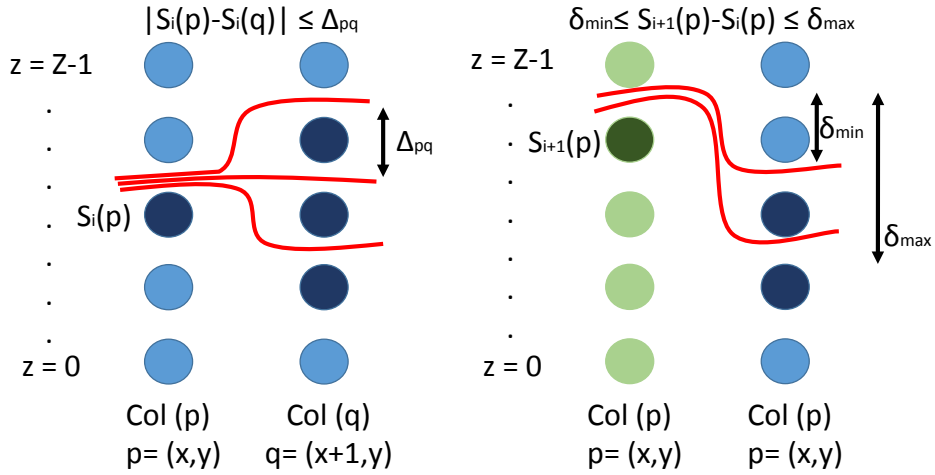


Figure 2.1: (left) Surface smoothness constraint for two neighboring columns p and q for a surface S_i. The arcs in red show the feasible st-cut for a given surface position $S_i(p)$ for column p. (right) Surface separation constraint for two corresponding columns for adjacent surfaces S_i, S_{i+1}. The arcs in red show the feasible st-cut for a given surface position $S_{i+1}(p)$ for column p.

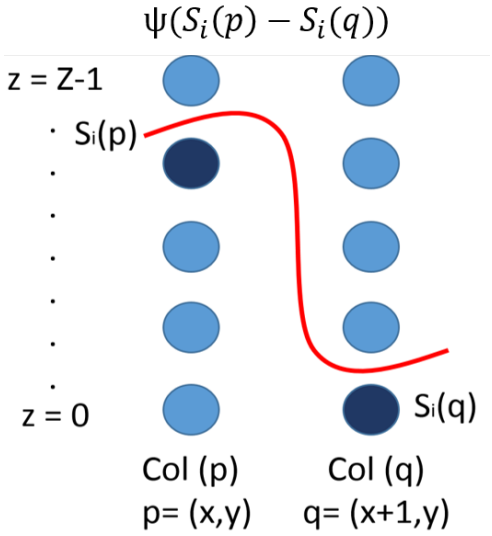


Figure 2.2: Surface smoothness constraint for two neighboring columns p and q for a surface S_i. The arc in red shows a st-cut for a given surface position $S_i(p)$ for column p. The smoothness cost for the given cut is $w \times \psi(S_i(p) - S_i(q))$.

intermediate output at the respective neuron is computed as :

$$a_q = \sum_{n=1}^{n=N} w_{qn} x_n + b_q \tag{2.4}$$

where w_{qn} is a weight parameter and b_q is a bias term associated with the ouput z_q defined as $z_q = f(a_q)$ and $f(.)$ is the activation function (introduces non-linearity). Given a set of training examples, the objective of the training process is to learn the parameters w_{qn}, b_q so as to minimize some objective or loss function. The standard approach to learning these parameters is the error backpropagation algorithm [24].

A CNN [6] is a multilayer, hierarchical neural network but bears at least three principal factors different from a generic neural network: local receptive fields, weight sharing and spatial pooling layers. The network topology exploits the stationary nature of natural images by learning features using locally connected networks. CNN employs a local receptive field rather than a global one, which is accomplished by capturing local structure of image through constraining each neuron to depend only on a spatially local neighborhood of the neurons in the proceeding layer. Moreover, weights are shared across different neurons in the same layer, which can be translated to evaluating the same filter over all local windows of the input image. A feature in the image at some location can be calculated by convolving the feature detector and the image at that location, hence such a layer in the network is termed as a Convolution Layer [6].

Spatial pooling [6] in CNN is to divide the image into an array of blocks and then evaluate a pooling function over the responses in each block. The goal of pooling is to reduce the dimensionality of the convolutional responses and enforce a translational invariance (to a small degree) into the model. In the case of max pooling [6], the response for each block is taken to be the maximum value over all response values within the block. A typical CNN consists of multiple layers, alternating between convolution and pooling. Compared with shallow CNN architectures, deep CNN

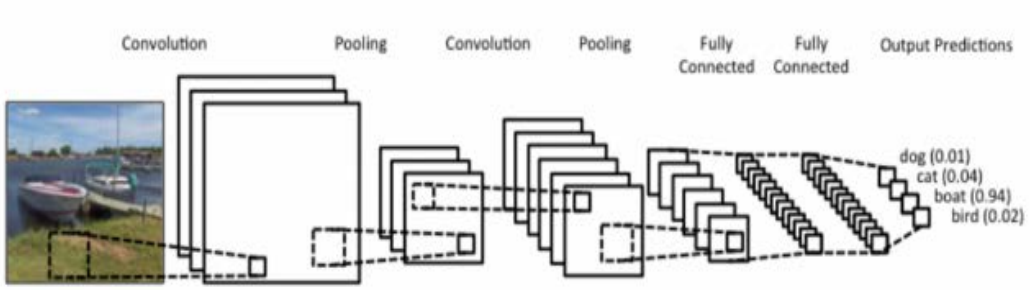


Figure 2.3: Illustration of a simple CNN architecture for image classification into 4 classes. [1]

has more hidden layers. Lower layers which are defined as the ones closer to the input construct low-level convolutional filters and provide low-level encoding of the input image. In contrast, higher layers learn more and more complicated structures. In CNN, stride length is used to specify the number of pixels with which the local receptive field is moved to the right from left (or down from top). An example of a CNN architecture for image classification is shown in Fig 2.3. The layers used in the CNN architecture for this work are described in detail in Section 6.3.2.

Some examples of successfuly application of CNNs in computer vision are image classification and object recognition [6] [25] [26] [27], text recognition [28] and action recognition [29]. In medical imaging analayis applications, CNNs have also been successfully applied for body part recognition [30], brain tumor segmentation [31], lung nodule classification [32], multiple sclerosis lesion segmentation [33], pulmonary embolism detection [34], glaucoma detection [35], gland detection [36], mammogram segmentation [37] and diabetic retinopathy detection [38]. However, most of these applications are where the CNNs have been used for classification, i.e. assigining a class to each pixel/voxel in the image or assigning a class/label to the entire image.

2.3 Related Work for Retinal Layer Segmentation in OCT

Spectral Domain Optical Coherence Tomography (SD-OCT) [39] [40] is widely used in imaging of the retina because of its ability to extract cross-sectional information of the retina. Automated methods for layer segmentation in OCT volumes have certain advatages compared to manually tracings the retinal layers, such as improved time-efficiency, and low intra-/inter-observer variability.

Intensity-based surface segmentation methods had been developed which utilize correlation between adjacent A-scans [41], iterative thresholding technique [42] using prior information from vessels and segmentation using a Canny edge detector [43] [44]. Such methods suffer from limitations based of incomplete incorporation of textural information and are therefore sensitive to noise, blood vessel shadows, and motion artifacts.

Active contour based methods [45] [46] [47], shape model based method [48] [49] and machine-learning-based methods [50] [51] [52] are other common algorithms employed for segmentation of surfaces in OCT volumes. However, these methods are not fully automated, require prior information to design various term within the respective models, does not provide flexibility for different applications and are primarily designed for target specific applications.

Graph based methods were developed which transform each voxel of the image into a node in the graph and then employs various graph based algorithms to segment the target surfaces based on the formulation. Traditional graph algorithms like dynamic programing and Dijakstra's method for finding the shortest path have been applied to segment retinal surfaces in OCT B-scans [53] [54] [55] [56]. However, these methods also require careful design of transformations, initlaization of starting points and does not exploit the full 3-D information.

Finally, the set of the methods based on the state-of-the-art graph search (optimal

surface segmentation) [2] [7] were developed and applied to numerous multiple surface segmetnation applications realted to retinal layers in OCT volumes. The method was first applied to simultaneously detect multiple retinal surfaces in SD-OCT macular scans [19]. However, the major practical limitation of the method was the time complexity which grows exponentially with the size of the image and the number of target surfaces. To overcome the limitation, the multi-resolution graph search technique [20] was developed to accelerate the processing time and was applied to SD-OCT optic nerve head (ONH) scans for multiple surface segmentation. Further, machine-learning based cost generation for graph seach [57] [58] was developed and was applied to surface segmentation of OCT images of humans, mice and canines. The framework was further extended to incorporate prior information of the surface profiles [9] [22]. By adding the prior information, the graph-theoretic algorithm has added the ability to entirely consider both global and local optimization and achieves better robustness.

CHAPTER 3
MOTIVATION AND SPECIFIC AIMS

3.1 Motivation

The popular graph search method has been extensively used for various surface segmentation application. However, the method may fail or may not be efficient enough in certain real world applications. The applications discussed in this thesis work are primarily medical image segmentation application, although the developed methods can be generically applied to any similar segmentation problem. Here, such identified segmentation problem and the motivation for development of methods to solve such problems is enumerated in Section 3.1.1. Furthemore, the method requires human expert designed data cost, transformations, smoothness constraints etc. Therefore, the need for a generic method to eliminate human expert intervention for surface segmentation problems is discussed in Section 3.1.2.

3.1.1 Segmentation Problems Difficult to Solve Using Graph Search

First, the methods may fail when abrupt changes are present in target surface topology. Graph search based methods [2] [7] may have a problem in cases with presence of steep surface smoothness changes and abrupt surface separation (distance) changes between a pair of interacting surfaces. Some examples in real world applications are spectral-domain optical coherence tomography (SD-OCT) volumes of subjects with severe glaucoma [59], dursen due to age-related macular degeneration (AMD) [60] (Fig. 3.1 shows an example) and Pigment Epthithelium Detachement (PED) [61] (Fig. 3.2 shows an example) in the eye. The graph search method as discussed in Section 2.1 uses hard smoothness constraints that are a constant in each direction to specify the maximum allowed "jump" of any two adjacent voxels on a feasible surface. For multiple surface detection, the method makes use of maximum and minimum constant constraints for the surface separation to specify the maxi-

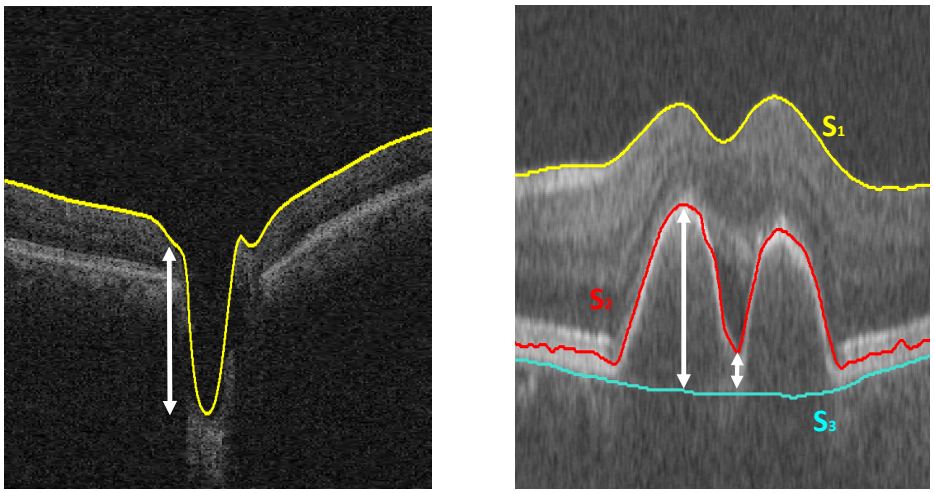


Figure 3.1: Steep change in surface smoothness can be seen in SD-OCT image of an eye with severe glaucoma (*left*). Abrupt changes in surface separation between surface 2 (S_2) and surface 3 (S_3) can be seen in SD-OCT image of an eye with AMD (*right*).

mum and minimum allowed distances between a pair of surfaces. This does not allow for flexibility in constraining surfaces. Steep jumps in surface positions of neighboring and corresponding graph columns can only be captured using high values of the smoothness and separation constraints, which results in unconstrained surfaces.

One possible way to solve this problem is to use varying feasibility constraints learned from a training set, as reported in [19]. However, this method cannot penalize the deviation inside the allowed constraints. Methods employing trained hard and soft constraints [22] [9], use prior terms to penalize local changes in surface smoothness and surface separation. The prior term requires learning and may give inaccurate results when there is plenty of variation within the data. It may also result in over smoothing of the surface since it is not discontinuity preserving. Moreover this method is more time intensive than tradional graph search method.

Furthermore, approaches using multiple resolution technique [62] for reduction of time and memory consumption, requires image flattening and identification of

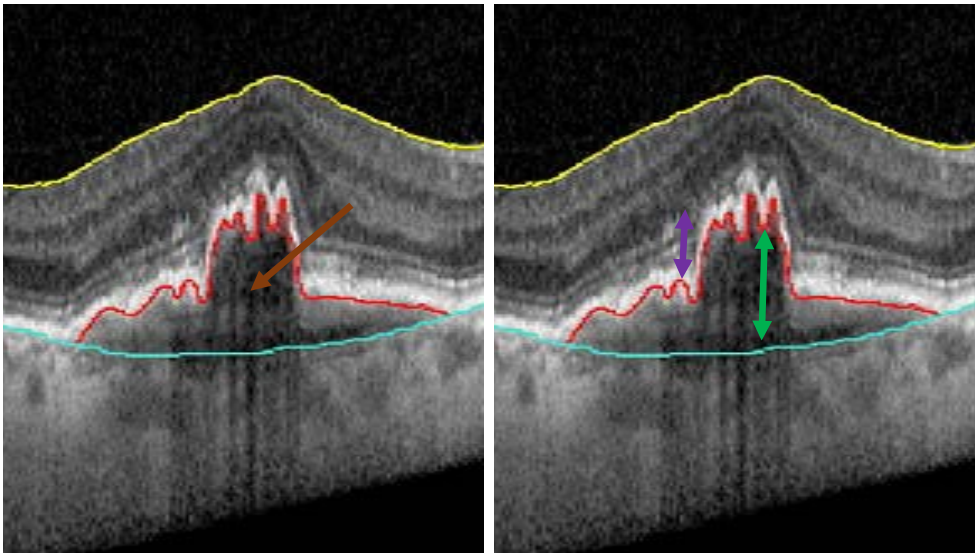


Figure 3.2: Example of an image slice with PED. Expert manual tracings shown as Yellow - Internal Limitng Membrane (ILM), Red - Outer Retinal Pigment Epithelium (ORPE), Blue - Inner Choroid (IC). The PED is indicated by the brown arrow (left). The steep change in surface smoothness of ORPE is indicated by the purple arrow and the abrupt change in the surface separation between ORPE and IC is indicated by the green arrow (right).

region of interest at each step. Identifying a region of interest for cases with abrupt surface smoothness or separation changes due to presence of pathological objects is difficult and may result in suboptimal results. Therefore, incorporating a prior which preserves these discontinuities (does not over penalize these abrupt changes in surface topology) shall provide for more flexibility and accuracy for segmentation of such surface topologies. This problem is addressed by developing a novel graph based method while incorporating a truncated convex prior (similar to a saturation function) for modelling the surface smoothness and surface separation constraints.

Second, inability of the graph search method to achieve subvoxel accuracy while using convex priors (convex penalty for surface smoothness/separation) by performing segmentation tasks in an irregularly sampled space. Within the graph search framework Ref (2.1), the graph is constructed such that the graph nodes correspond to the

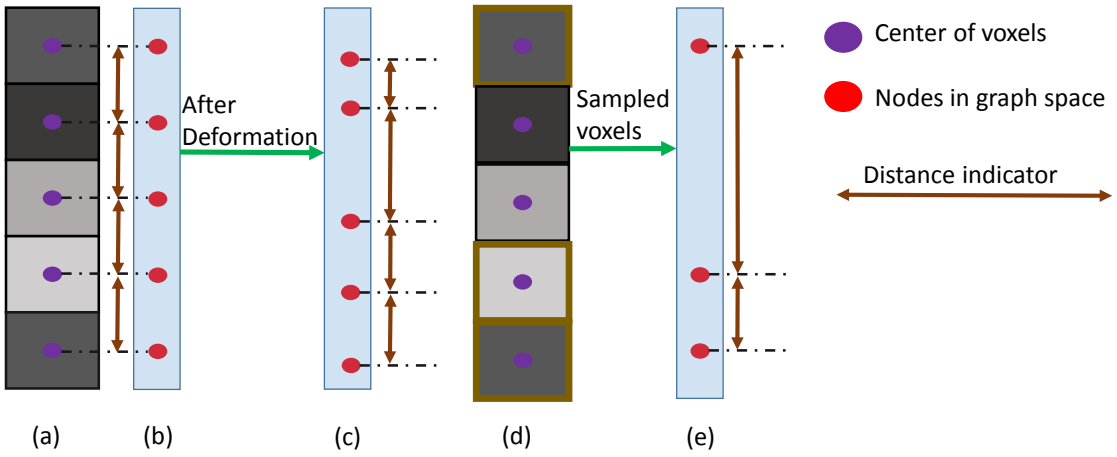


Figure 3.3: Illustrative example of non-equidistant spacing between adjoining graph nodes. (a) A voxel column in a 3D image. (b) Representation of voxel column in (a) as graph nodes in the graph space. (c) Shifted positions of the graph nodes after the deformation is applied in the vertical direction to the voxel column. (d) A voxel column in a 3D image with sampled voxels shown in brown boundary (e) Representation of voxel column in (d) as graph nodes in the graph space.

center of evenly distributed voxels. A volumetric image data is typically represented as an orthogonal matrix of intensities. Volumetric images are obtained by discretizing into voxels the continuous intensity function uniformly sampled by sensors, resulting in partial volume effects [63] [64]. Partial volume effects contain additional information that can be exploited using graph techniques to achieve subvoxel accuracy for surface segmentation by applying a deformation [65]. The deformation may be applied using a displacement field directly obtained from the volumetric image data. Specifically, such a deformation shall result in non-equidistant spacing between the adjoining nodes in the graph due to shifts in the centers of voxels because of the deformation as shown in Fig 3.3(b-c). For the purpose of terrain like surface segmentation in 3-D volumes using graph search methods, this may be considered equivalent to a generalized case of a cube volume formed by voxels whose centers are irregularly spaced along the z dimension.

The traditional graph search method [2] is not capable of segmenting surfaces with subvoxel accuracy since the deformed volume comprises of voxels with irregular spacing between the adjoining centers of the voxels. To address this problem, the subvoxel accurate graph search method [65] was developed to simultaneously segment multiple surfaces in a volumetric image by constructing a graph with non-equidistant spacing between the adjoining nodes. The method first creates the graph using the conventional optimal surface segmentation method [2], then deforms it using a displacement field and finally adjusts the inter-column edges and inter-surface edges to incorporate the modification of these constraints due to the displacement in the center of the voxels. The method employs hard surface smoothness and surface separation constraints. Thus, incorporating a convex prior to the graph search based framework to achieve subvoxel accuracy of segmentations is of significance to provide for more accurate and flexible modeling of target segmentations. Furthermore, the notion of irregularly sampled space can be further used for creating subvolumes (possibly region of interests) from a given volume by irregularly sampling voxels along each column in the z dimension, which shall result in non-equidistant spacing between the adjoining voxel centers as shown in Fig 3.3(d-e). Therefore, to address this segmentation problem a new graph based method is developed to incorporate convex priors for multiple surface segmentation in irregularly sampled space.

3.1.2 Need for a Generic Surface Segmentation Method Without Requirement of Human Expert Intervention

The exisiting surface segmentation approaches as discussed in Section 2.3 including the state-of-the-art graph search approach, offer certain flexibities in terms of using the methods for various application. However, none of the methods are generic enough or completely independent of expert based design of transformation, features, parameters and constraints. In this doctoral work, all the applications have been

compared to the graph search method which has proven to be superior to the rest of the methods discussed for the use in surface segmentation applications. Therefore, the limitations of the existing methods discussed in this section is majorly related to the graph search method, however most of them are also valid to the other methods in certain respects as well.

First, the graph search method is principally dependent on the design of the data cost term, surface smoothness term and the surface separation term. Failure, in resaonable design of any of these terms shall result in poor segmentation results, even though the solution shall be optimal with respect to the objective function. This attribute, transfers the key responsibility of the design of these term to users. A good design of these parameters is therefore also dependent on the kind of application. For example, the data cost term may be designed in a different manner for a normal case compared to a diseased case or the surface smoothness term is different for surfaces in normal case compared to a case with glaucoma where certain surface profiles are more complex. Such a design, results in application specific design of terms and thus, does not offer a generic design which is applicable to majority of the applications related to a certain body part or tissue.

Second, the surface constraints also require tuning of certain parameters. For example, the weight co-efficient discussed in Section 2.1 for Eqn 2.3. These parameters may be tuned using a grid search approach, however the tuning becomes very difficult in cases where the target surfaces are complex due to presence of a pathology. Furthermore, in such cases, even though the parameter is tuned for the entire dataset, there generally exists parameters specific to each volume which may provide better segmentation results. Therefore, it is difficult to tune the parameters for such applications and to model surface constraints which are generic enough to encompass the entire dataset.

Third, the traditional graph search method [2] is limited by inefficiencies with

respect to the memory requirement and processing time. The former stems from the fact that the entire volume has to be loaded into the memory in terms of nodes and edges in the underlying graph which increases with the increase of target surfaces. The inefficency in processing time is caused by time bound relationship with respect to the number of nodes and edges in the graph for solving the *st*-min cut problem using the max flow algorithm. Therefore, previously, segmenting OCT volumes in original resolution used to be a time intensive process. In order to tackle this challenge, the multi-resolution graph search [20] framework was developed which extracted a region of interest iteratively at each downsampled scale, thus making the method computationally efficient. However, the region of extraction scheme and the order of the surfaces to be segmented requires expert intervention. Furthermore, the schemes are also application specific.

Therefore, evidently, there is a need for a method which is generic enough and eliminates the aspect of expert human intervention with respect to designing of the various attribures required by the segmentation method. For example, a method which is able to segment surfaces in both normal and diseased cases with the very same parameters, does not require the design of data terms and surface constraints by the users, does not require expert tuning of parameters and at the same time is efficient in terms of memory requirement and processing time. To address, this problem, the popular deep learning framework using Convolution Neural Nets (CNN) are employed to learn the essential transformations, features and paramters from training data without any human intervention. The design of the framework is done by making certain important observations within the graph search framework and then leveraging them to construct a framework with a close relationship with the graph search method.

3.2 Specific Aims

In particular, the thesis has the following aims:

- **Aim 1: Multiple surface segmentation with truncated convex priors.**

  In this aim, we will seek to model the surface smoothness penalty and the surface separation penalty in the graph search framework as a truncated convex function to avoid over penalizing the sharp changes in the surface topologies. Specifically, we propose the following:

  - **Aim 1.1:** Develop single surface segmentation algorithm using truncated convex prior: Applied to segmentation of the Internal Limiting Membrane (ILM) in Spectral Domain Optical Coherence Tomography (SD-OCT) volumes with a deep cup.
  - **Aim 1.2:** Develop a multiple surface segmentation approach with truncated convex priors and validate on SD-OCT datasets of normal eyes and eyes with Age Related Macular Degeneration (AMD).
  - **Aim 1.3:** Simultaneous segmentation of ILM, Outer Retinal Pigment Epithelium (ORPE) and Inner Choroid (IC) in SD-OCT volumes with Retinal Pigment Epithelial Detachment (PED) using truncated convex priors.

- **Aim 2: Optimal multiple surface segmentation with convex priors in irregularly sampled space.**

  In this aim, we propose the generalization of the graph search method with convex priors framework to be applicable in irregularly sampled space wherein the nodes within the columns of the constructed graph are allowed to have non-equidistant spacing between them. Specifically, we propose the following:

  - **Aim 2.1:** Develop the method for Optimal multiple surface segmentation with convex priors in irregularly sampled space.
  - **Aim 2.2:** Prove the correctness of the graph construction with respect to global optimality of the solution.

  - **Aim 2.3:** Validation of the method for subvoxel accuracy and super resolution segmentation accuracy.

- **Aim 3: Multiple surface segmentation using deep learning.** In this aim, we propose to design a novel framework using deep learning to accomplish the task of multiple surface segmentation, with the goal of attaining higer accuracy and computational efficiency comapred to graph search method. The framework shall also eleiminate the requirement of any human expert intervention. The framework shall be gneric enough such that the same trained network is able to infer on multiple surface segmentaion for both normal and diseases cases, unlike graph search which requires two different set of optimized parameters and human expert designed transformations and strategies. Specifically, we propose the following:

  - **Aim 3.1:** Develop the deep learning based method for single surface segmentation and validate on SD-OCT volumes of normal eye.
  - **Aim 3.2:** Develop the deep learning based method for multiple surface segmentation and validate on SD-OCT volumes of normal eye and eye with AMD.

CHAPTER 4
MULTIPLE SURFACE SEGMENTATION USING TRUNCATED CONVEX PRIORS (AIM 1)

4.1 Introduction

The **objective of Aim 1** is to overcome the challenges posed by real world applications, wherein the target surfaces have steep change in surface smoothness or/and abrupt change in surface separation between two surfaces. Such surfaces are difficult to segment using the traditional graph based methods as discussed in Section 3.1.

The goal of this aim is to develop an automated graph based segmentation algorithm by employing a truncated convex penalty for surface smoothness and surface separation to capture the steep and abrupt changes in surface profiles and surface interactions. The method also ensures the enforcement of a minimum separation between a surface pair. A truncated convex penalty is discontinuity preserving having a bound on the largest possible penalty. An example of a truncated convex penalty function is shown in Fig. 4.1. This avoids over penalizing sharp changes in surface smoothness and surface separation. The main idea is to take advantage of a local search technique. The advantage of using such a method is two folds: it allows for enforcement of truncated convex penalty, and it is significantly faster than graph search methods for larger data volumes. A local range search method explores a large search space by considering a range of labels. A label may be interpreted as the surface position for a column in a 3-D volume. It gives a solution by iteratively computing a maximum flow for a subset of the voxel domain. The key idea is at each iteration the best labeling is selected by either retaining its current label or changing to a new label while maintaining the truncated convex constraint. The range expansion method [66] allows for the incorporation of a truncated convex penalty. Using this method as a building block, we develop techniques for single surface and multiple surface segmentation with truncated convex penalties for surface smoothness and

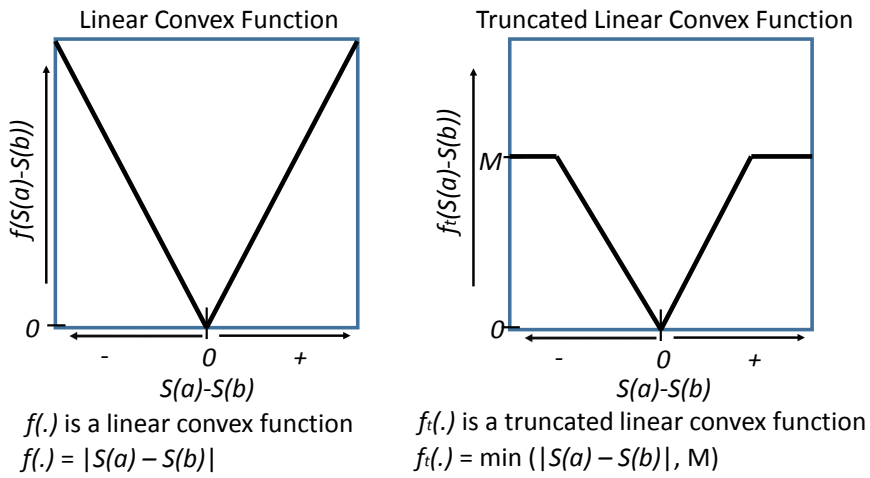


Figure 4.1: Example for a linear convex penalty (*left*) and a linear truncated convex penalty (*right*). $S(a)$ and $S(b)$ are surface positions for a pair of neighboring columns a and b. M is the truncation factor (bound on the maximum possible penalty).

surface separation. The method is further evaluated and validated by segmenting specific surfaces in normal OCT data and OCT data with pathologies like Glaucoma, AMD and PED.

4.2 Single Surface Segmentation Using Truncated Convex Priors (Aim 1.1)

Spectral-domain optical coherence tomography (SD-OCT) is used clinically for the diagnosis and management of glaucoma, with the internal limiting membrane (ILM) being one important structure of interest as it reflects the upper bounding surface of the optic cup. Glaucoma is the second leading cause of blindness in the developed world, and is characterized by gradual cupping of the optic nerve head (ONH) and visual field loss [59]. The hallmark of severe glaucoma is the deep cupping of the optic nerve, which is visible in cross-sectional ONH-centered SD-OCT volumes by an increased steepness of the internal limiting membrane in the central cup region. In such cases of severe glaucoma, the boundary of the optic cup can become very steep and difficult to segment as illustrated in Fig 4.2.

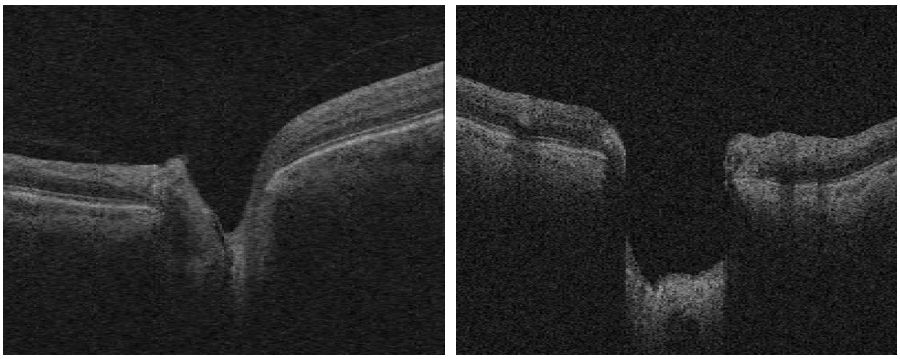


Figure 4.2: (Central B-scans (xz slice) from optic nerve head centered scans from (a) subject with less severe glaucoma, (b) subject with more severe glaucoma.

4.2.1 Method Design

Consider a volumetric image $I(x, y, z)$ of size $X \times Y \times Z$. A surface is defined as $S(x, y)$, where $x \in \mathbf{x} = \{0, 1, ...X - 1\}$, $y \in \mathbf{y} = \{0, 1, ...Y - 1\}$ and $S(x, y) \in \mathbf{z} = \{0, 1, ...Z - 1\}$, thus intersecting each column at a single voxel location. Each (x, y) pair denoted as $col(x, y)$ forms a column parallel to the z-axis. We use N_s to denote the neighborhood setting of image domain.

The goal of single surface segmentation problem is to seek the "best" surface $S(x, y)$ in I, which is transformed into an energy minimization problem. The function $S(x, y)$ can be viewed as labeling with the label set $\mathbf{z}$. Thus, the surface segmentation problem, in fact, is a multiple labeling problem. To simplify the notation, we use a and b to denote two neighboring (x, y) pair in the image domain $\mathbf{x} \times \mathbf{y}$.

Energy function - The energy function $E(S)$ for the surface S takes the following form:

$$E(S) = \sum_{a \in \mathbf{x} \times \mathbf{y}} D(S(a)) + \sum_{(a,b) \in N_s} V_{ab}(S(a), S(b)) \tag{4.1}$$

The data cost term $D(S(a))$ measures the total inverse probability of all voxels

on a surface S, while the surface smoothness term $\sum_{(a,b)\in N} V_{ab}(S(a), S(b))$ measures the extent to which S is not piecewise smooth.

It is known that the energy function $E(S)$ can be minimized by minimum st-cuts [8] when $V_{ab}(.)$ is convex [67]. Since, a convex $V_{ab}(.)$ may over smooth the surface S, its highly desirable to make use of a truncated convex $V_{ab}(.)$ to preserve discontinuity of the target surface. Convexity encourages smoothness while truncation helps preserve the surface discontinuity. A truncated convex function takes the following form:

$$V_{ab}(S(a), S(b)) = w_{ab}\ (min(f(S(a), S(b)), M)) \tag{4.2}$$

where $f(.)$ is a convex function, $M > 0$ is the truncation factor, $w_{ab} \geq 0$.

4.2.2 Method Overview

A local range search technique like the range expansion algorithm [66] is applied to enforce the truncated convex penalties. Since the algorithm is iterative in nature, a constant interval length L is defined to determine the subset of consecutive labels at each iteration. An interval is defined as $T \subset \mathbf{z}$ where $\mathbf{z} = \{0, 1, ...Z-1\}$. An interval of length L at iteration m is defined as $T_m = [m, min(m + L - 1, Z - 1)]$, where $m = 0, 1, ...Z - 2$. At each such interval T_m a graph G_m is constructed using each voxel from $I(x, y, z)$ as a node where $x \in \mathbf{x}$, $y \in \mathbf{y}$ and $z \in \mathbf{T_m}$. Example of consecutive interval construction is shown in Fig 4.3. The basic idea of the algorithm is to iteratively compute the minimum st-cut on each of the graphs G_m constructed for each consecutive interval. At each interval, the best label is selected by either retaining the label from the previous graph G_{m-1} solution or by changing to a new label belonging to the current interval. The iterations are continued until all the defined intervals have been iterated over, at the end of which the total energy $E(S)$ of the segmented surface S is compared to the previous solution. If the energy is found

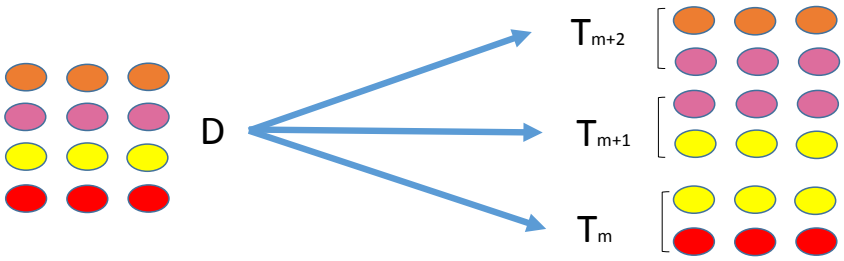


Figure 4.3: Example of interval construction for a 2D data cost slice D for surface S with 3 neighboring columns and Z=4. Each voxel is represented as a node. Length of the interval (L) = 2. Consecutive interval ($T'_m s$) construction at each iteration m is shown.

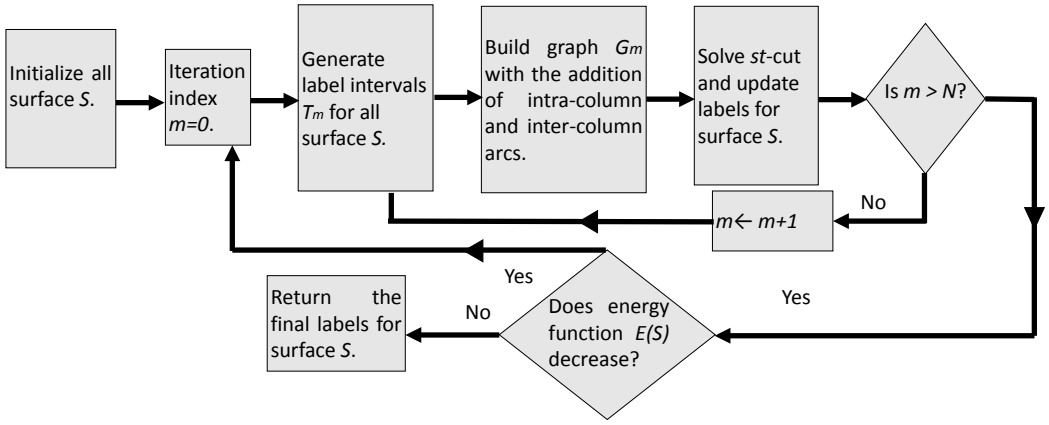


Figure 4.4: Method pipeline for single surface segmentation. N is the maximum index of possible iterations.

to have decreased, the entire defined label intervals are iterated over again, initialized with the current obtained solution. The method terminates when $E(S)$ converges. The method pipeline is shown in Fig 4.4.

4.2.3 Graph Construction

The graph G_m is constructed at each iteration m for corresponding subvolume $\mathbf{x} \times \mathbf{y} \times T_m$. Herein, each voxel in the subvolume is represented by a node in the graph. Let us assume at an iteration m the set of labels given for $col(a)$ in the cost volume D be $T_m = [q_a, q_a + 1, ...q_a + L\text{-}1]$. q_a hence represents a node in G_m for $col(a)$ with label q $((x, y) = a, z = q)$.

To encode the data cost and ensure monotonicity of the target surface, intra-column arcs are added. For all $k \in [0, L-2]$, we add a weighted arc with $D(q_a + k)$ weight from node $q_a + k$ to $q_a + k + 1$ and an arc with $+\infty$ weight in the opposite direction. We also add a weighted arc with $D(q_a + L\text{-}1)$ weight from node $q_a + L\text{-}1$ to the terminal node t. Furthermore, the labeling for a given column can either retain its current labeling or can switch to a label in the interval. Denote the current labeling for $col(a)$ at the start of iteration m as $S_{m-1}(a)$. If $S_{m-1}(a) \notin T_m$, we add a *label retaining* arc with $D(S_{m-1}(a))$ weight from source node s to node q_a allowing $col(a)$ to retain its current label or add an arc with $+\infty$ weight from source node s to node q_a otherwise. This ensures that any finite st-cut shall always cut each column only once and the arc weights encode the data term of the energy function $E(S)$.

The next task is to incorporate inter column arcs to ensure truncated convex penalty for surface smoothness. Let $col(a)$ and $col(b)$ be two neighboring columns. The basic idea is to add arcs between the two columns at a given interval to model the convex function being used. Such arcs are added by making similar use of discrete equivalent of second derivatives of the convex function [9] [66]. Using this construction it is proved in [7] that the total weight of the arcs that are cut between two neighboring columns equals to a penalty represented by the convex function plus a constant of approximation.

The last task is to enforce truncation of the convex penalty. As discussed earlier, at each iteration the previous labeling may be retained or a new label could be chosen

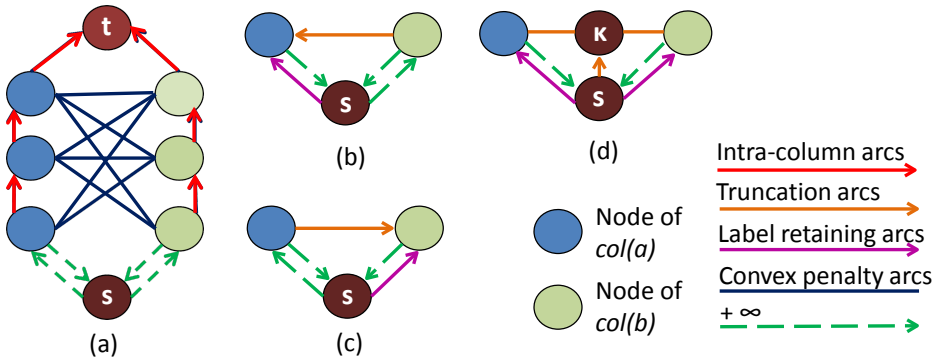


Figure 4.5: (a) Example graph construction for Case1. Additional arcs added to (a) for (b)(c) Symmetric Cases, (d) Combination Case. Length of interval (L) =3.

from the current interval. We distinguish four cases for a pair of neighboring columns $col(a)$ and $col(b)$ as follows:

Base Case: $S_{m-1}(a) \in T_m$ and $S_{m-1}(b) \in T_m$.

Symmetric Cases: $S_{m-1}(a) \notin T_m$ and $S_{m-1}(b) \in T_m$, $S_{m-1}(a) \in T_m$ and $S_{m-1}(b) \notin T_m$.

Combination Case: $S_{m-1}(a) \notin T_m$ and $S_{m-1}(b) \notin T_m$.

No further arcs are added to the base case. For the symmetric cases we add a *truncation arc* with $w_{ab}(M + f(L)/2)$ from node q_a to q_b in both the directions to the graph of the base case. For the combination case, we add a new node κ to the graph of the base case. The truncation arcs with weight $w_{ab}(M + f(L)/2)$ are added from node κ to q_a and from node κ to q_b. An additional truncation arc is added from source node s to node κ with a weight setting of $V_{ab}(S_{m-1}(a), S_{m-1}(b)) + w_{ab}f(L)/2$. The graph construction for all the cases are shown in Fig 4.5.

At each iteration, the graph is solved by computing a maximum flow which minimizes the energy function $E(S)$ approximately. The labeling $S_m(a)$ either retain its old label $S_{m-1}(a)$ or changes to a label belonging to T_m. A complete set of iterations is

defined as a set of iterations where constructed graphs at each T_m, $m = 0, 1, ...Z-2$ is solved. The set of iterations are continued till the energy cannot be further minimized thereby yielding the final labeling $S(\cdot)$.

4.2.4 Experiment Setting

The data used for the testing consisted of 10 optic nerve head (ONH) OCT volumes from 10 patients diagnosed with glaucoma, acquired on spectral-domain Cirrus (Carl Zeiss Meditec, Inc., Dublin CA). Each dataset had dimensions $6mm \times 6mm \times 2mm$ ($200 \times 200 \times 1024$ voxels). Manual tracings from an independent expert was obtained from 10 slices, selected randomly (from 10 sections of 20 slices each) in each dataset.

The data term is computed using the cost function of the image. In this study, for each SD-OCT volume, the cost image is obtained by a linear combination of a 3-D Sobel filters ($3 \times 3 \times 3$) for enhancing the vertical edges around the neural canal opening (NCO). Another 3-D Sobel ($3 \times 3 \times 3$) filter was used to extract the horizontal boundary information of retinal layers for dark-to-bright transitions along the vertical direction. Then, a dynamical parameter of readjusting the cost image contrast can be assessed by computing the mean background intensity-value at the top small corner regions, in which retinal tissue is usually absent. To this step, the cost function design is completed. To avoid finding surfaces other than the ILM, a previous version of the multiresolution graph-theoretic approach [62] was applied for each SD-OCT volume, and the original input cost images were modified such that the locations of the surfaces other than the ILM had a very high cost.

For comparison, the graph search algorithm [2] [19] was also used to segment the same input cost image. Considering to accelerate the computational time of this method, we down-sampled the input cost image by 4 for each SD-OCT volume first. We then pre-segmented the ILM in image low resolution domain. After up-sampling the pre-segmentation result from the low-resolution into the original resolution, the region of interest (ROI) of the input cost image was defined by a narrow band covering

the 3-D pre-segmented ILM at the center surface and 40 voxels height at both top and bottom directions. Then, the graph search algorithm was performed one more time on the refined cost image, and the final ILM segmentation was achieved.

The proposed method with truncated convex prior was used to segment the input cost image in the original resolution. For the smoothness constraint we used the truncated linear metric. The algorithm parameters used in our work were, w_{ab} =1.5, M=20 and L=2.

The segmentation accuracy of both methods were estimated using 1) unsigned mean surface positioning error (UMSP) given by Equation (4.3) and 2) average symmetric symmetric surface distance (ASSD) given by Equation (4.4).

$$UMSP(S_i) = \frac{\sum_{a \in \mathbf{x} \times \mathbf{y}} |S_i^{man}(a) - S_i^{auto}(a)|}{|S_i^{man}|} \tag{4.3}$$

where $S_i^{man}(a)$ and $S_i^{auto}(a)$ is the surface position for the i^{th} surface for column a given by expert manual tracings and automated segmentation respectively.

$$ASSD(S_i) = \frac{\sum_{a \in \mathbf{x} \times \mathbf{y}} r(S_i^{man}(a), S_i^{auto}) + \sum_{a \in \mathbf{x} \times \mathbf{y}} r(S_i^{auto}(a), S_i^{man})}{|S_i^{man}| + |S_i^{auto}|} \tag{4.4}$$

where $r(x, A)$ gives the shortest distance between a point x and any point on surface A. The ASSD measures how close the segmented surface is from the ground truth. A paired t-test was used to compare the segmentation results. A p-value of 0.05 was considered significant.

4.2.5 Results

The UMSP error for our proposed method was significantly lower ($p < 0.04$) from the graph search method. The ASSD for our proposed method was lower ($p < 0.07$) from the graph search method. Results from our proposed method is significantly

better on the cup slices (for which the algorithm was developed). Since only a few slices from each dataset contained the central cup region, for which the ground truth was available; we also calculated the above mentioned errors on only those slices containing the optic cup for each dataset. We found that UMSP error for our proposed method was significantly lower ($p < 0.03$) from the graph search method. The ASSD for our proposed method was significantly lower ($p < 0.04$) from the graph search method. The percentage improvement in the mean errors of our proposed method over the graph search method was also calculated and it demonstrates a significant improvement. The results of both the methods are summarized in Table 4.1. Illustration of the results are shown in Fig 4.6 and Fig 4.7.

Table 4.1: Results from truncated convex method and graph search method. The UMSP and UASSD errors, are expressed as mean ± standard deviation in μm.

	All slices		Cup slices	
	UMSP	ASSD	UMSP	ASSD
Truncated Convex	6.35 ± 1.35	3.40 ± 0.17	11.56 ± 4.97	4.37 ± 0.82
Graph search	7.69 ± 3.23	3.86 ± 0.54	15.79 ± 10.95	6.05 ± 3.49
% Improvement	17.17	11.92	26.78	27.7

4.3 Multiple Surface Segmentation Using Truncated Convex Priors (Aim 1.2)

Multiple surface segmentation with mutual interaction between surface pairs is a challenging task in medical image analysis. The process is time intensive for large volumetric data. The problem becomes further complicated with the presence of sharp discontinuities in a surface topology or abrupt change in surface separation between a pair of surfaces or the presence of both as can be seen in Fig 3.1, especially in datasets with presence of pathology. As discussed earlier, traditional graph search methods [2] [9] may over smooth the target surfaces and may not be able to segment

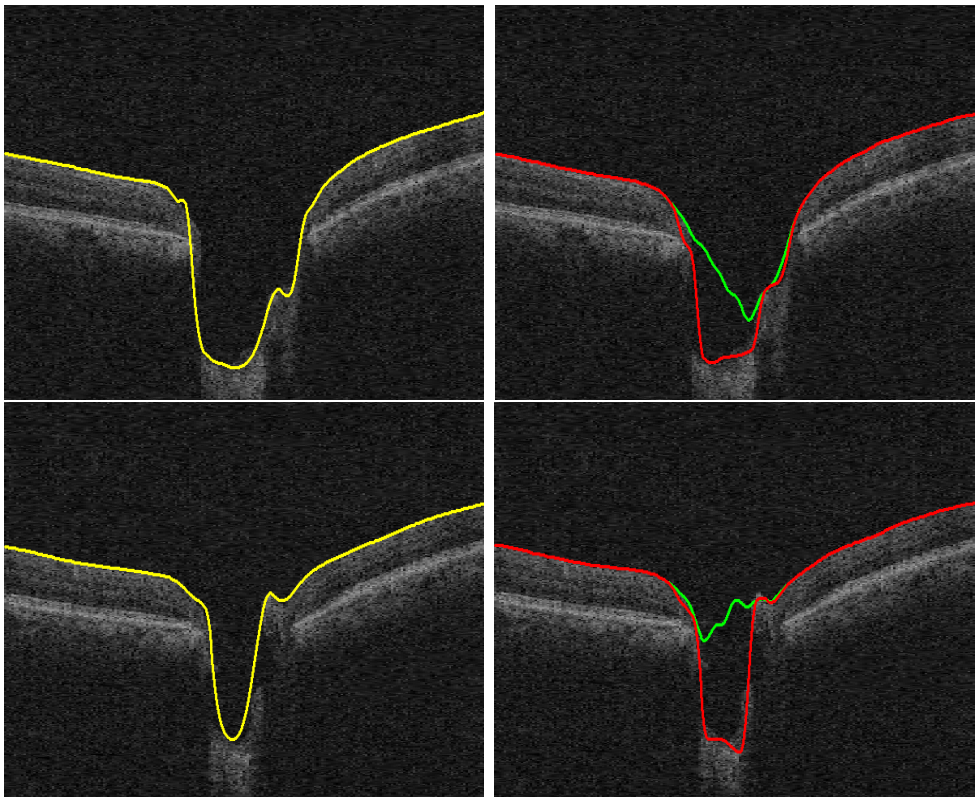


Figure 4.6: Examples of comparison between the manual tracings from an expert independent observer (yellow), the segmentation result obtained using truncated convex method (red) and Original Graph Search (green), respectively.

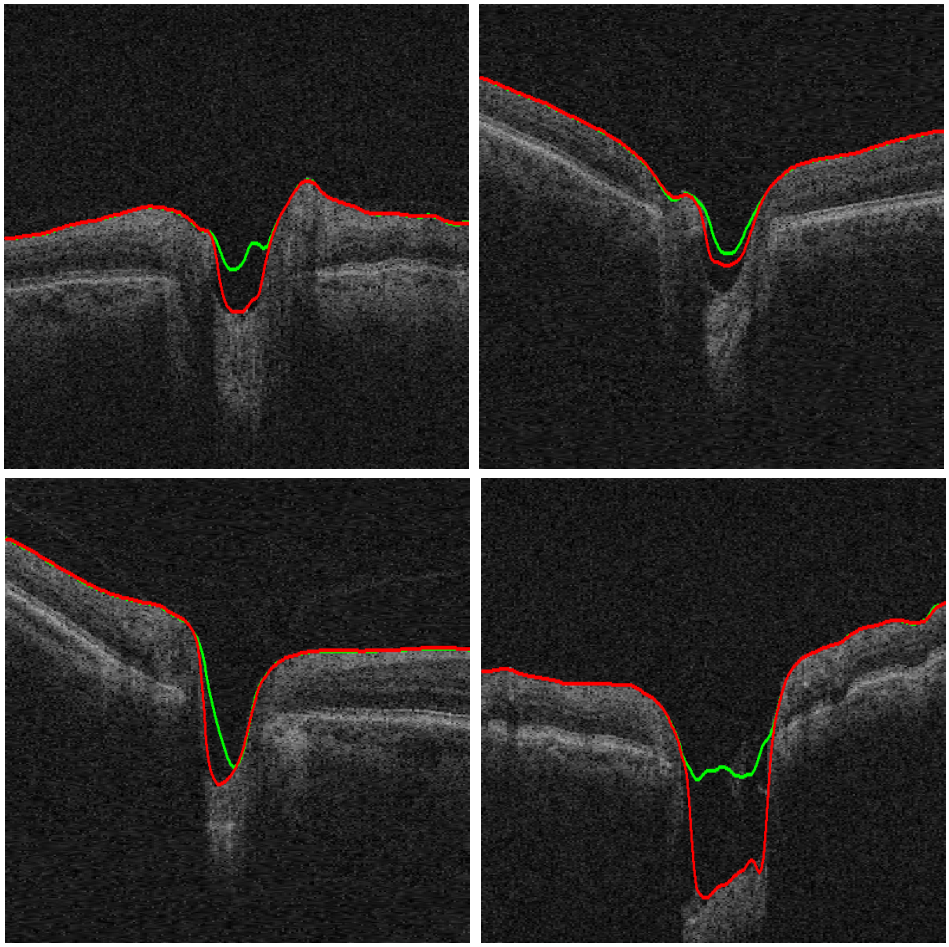


Figure 4.7: Segmentation result obtained for slices from different OCT volumes using the proposed method(red) and Original Graph Search (green). The manual tracings were not available for these slices.

the target surfaces with sufficient accuracy for such cases. To address this challenge, an automated method was developed to segment multiple surfaces simultaneously using truncated convex priors to handle this problem by building upon the framework for single surface segmentation using truncated convex priors. The method is also capable of enforcing a minimum separation between a pair of surfaces. Furthermore, it is sufficiently fast to segment large volumetric data in original resolution compared to the traditional graph search methods. The proposed method was validated on simultaneous intraretinal layer segmentation of optical coherence tomography images of normal eye and eyes effected by severe drusen due to age related macular degeneration.

4.3.1 Method Design

For ease of understanding, we define the problem of multiple surface segmentation with truncated convex priors in a similar manner as described in Section 4.2.1. Consider a volumetric image $I(x, y, z)$ of size $X \times Y \times Z$. A surface is defined as a function $S(x, y)$, where $x \in \mathbf{x} = \{0, 1, ...X - 1\}$, $y \in \mathbf{y} = \{0, 1, ...Y - 1\}$ and $S(x, y) \in \mathbf{z} = \{0, 1, ...Z - 1\}$. Each (x, y)-pair corresponds to a *column* of voxels $\{(I(x, y, z)|z = 0, 1, \ldots, Z - 1\}$, denoted by $col(x, y)$. We use a and b to denote two neighboring (x, y)-pairs in the image domain $\mathbf{x} \times \mathbf{y}$ and N_s to denote the neighborhood setting of image domain. The function $S(a)$ can be viewed as labeling for a with the label set $\mathbf{z}$ ($S(a) \in \mathbf{z}$). For simultaneously segmenting $\lambda(\lambda \geq 2)$ distinct but interrelated surfaces, the goal of the problem is to seek the "best" surfaces $S_i(a)$, where $i = 1, 2, ...\lambda$ in I with minimum separation $d_{i,i+1}$ where $i = 1, 2, ...\lambda - 1$ between each adjacent pair of surfaces S_i and S_{i+1}.

Energy Function - The problem is transformed into an energy minimization problem. The energy function $E(S)$ takes the following form :

$$E(S) = \sum_{i=1}^{\lambda}(\sum_{a \in \mathbf{x} \times \mathbf{y}} D_i(S_i(a)) + \sum_{(a,b) \in N_s} V_{ab}(S_i(a), S_i(b))) + \sum_{i=1}^{\lambda-1} \sum_{a \in \mathbf{x} \times \mathbf{y}} H_a(S_i(a), S_{i+1}(a)) \tag{4.5}$$

The data cost term $\sum_{a \in \mathbf{x} \times \mathbf{y}} D_i(S_i(a))$ measures the total cost of all voxels on a surface S_i, while the surface smoothness term $\sum_{(a,b) \in N_s} V_{ab}(S_i(a), S_i(b))$ as shown in Eqn 4.2, models the surface smoothness term as a truncated convex function as described in Section 4.2.1. The surface separation term $H_a(S_i(a), S_{i+1}(a))$ incorporates a truncated convex penalty for the separation between two adjacent surfaces, and ensures a minimum separation between them, which takes the following form :

$$H_a(S_i(a), S_{i+1}(a)) = \begin{cases} \infty, & \text{if } (S_{i+1}(a) - S_i(a)) < d_{i,i+1}, \\ w_a \min(f(S_{i+1}(a) - S_i(a)), M_{i,i+1}), & \text{otherwise} \end{cases} \tag{4.6}$$

where $f(.)$ is a convex function, $M_{i,i+1} > 0$ is the truncation factor, and $w_a \geq 0$.

4.3.2 Method Overview

The method is developed by extending the single surface segmentation with truncated convex prior method. Therefore, similarly the proposed method is iterative in nature. The pipeline for our method is shown in Fig. 4.8. At each iteration, it searches a small subset of the solution space defined by a label interval. A constant interval length L is defined to determine the subset of consecutive labels to be considered for all λ surfaces at each iteration.

An interval of consecutive labels for surface S_i is defined as $T_i \subset \mathbf{z}$ where $\mathbf{z} =$

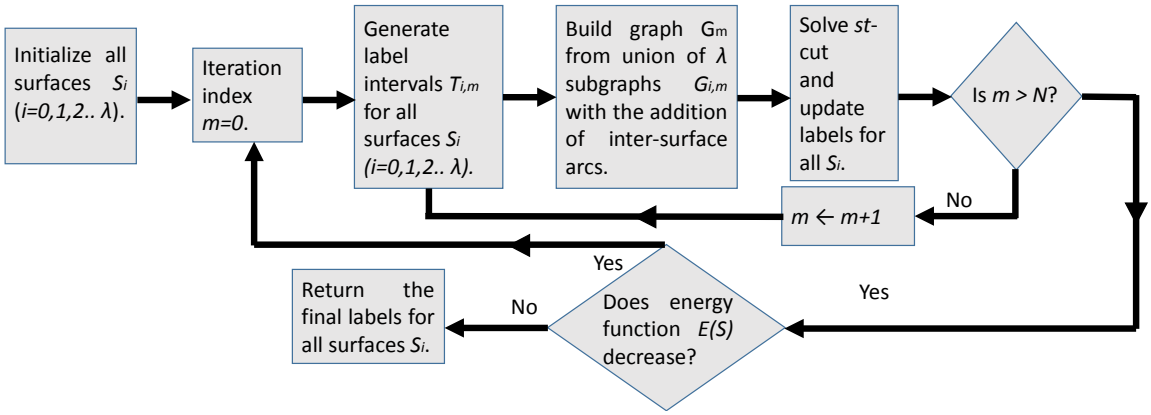


Figure 4.8: Pipeline for our method. $N = Z - 2 - \sum_{i=1}^{\lambda-1} d_{(i,i+1)}$.

$\{0, 1, ...Z-1\}$. Denote $T_{i,m}$ shown in Eqn.(4.7) as the label interval for S_i at iteration m, where $m = 0, 1, ...Z - 2 - \sum_{i=1}^{\lambda-1} d_{(i,i+1)}$. For each pair of adjacent surfaces S_i and S_{i+1}, interval $T_{i+1,m}$ is displaced by $d_{i,i+1}$ from $T_{i,m}$ to ensure the minimum separation constraint between S_i and S_{i+1}. An example of interval construction is shown in Fig 4.9.

$$T_{i,m} = \{l \mid m + \sum_{j=1}^{i} d_{(j-1,j)} \leq l \leq \min(m + \sum_{j=1}^{i} d_{(j-1,j)} + (L-1),\ Z-1)\} \quad (4.7)$$

The search problem is reduced to compute a minimum st-cut in a properly constructed graph in the subvolume generated at a given interval. The key for the graph construction is to encode the truncated convex priors. For each surface S_i, a subgraph $G_{i,m}$ (where m is the index for iteration and i is the index of the sought surface) is then constructed using the technique for the single surface detection as discussed in Section 4.9. This incorporates all the intra-column arcs for surface monotonicity for data cost volume $D_{i,m}$ (for searching S_i at iteration m) and inter-column arcs for surface smoothness (truncated convex penalty) to search a single surface S_i. The graph G_m for the simultaneous search of all λ surfaces at the iteration m consists of

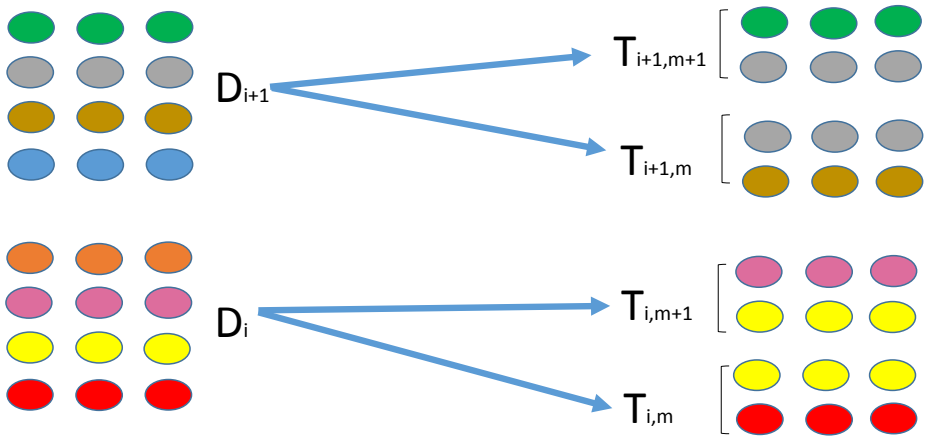


Figure 4.9: Example of interval construction for 2-D data cost slice D_i for surface S_i and D_{i+1} for adjacent surface S_{i+1} with three corresponding columns and Z=4. L =2 and minimum separation constraint $d_{i,i+1} = 1$.

the union of those λ subgraphs $G_{i,m}$'s. In addition, we add inter-surface arcs between the corresponding columns (columns for the same (x,y) pair) of graph nodes in the data cost volumes $D_{i,m}$ and $D_{i+1,m}$ in the input image I. The inter-surface arcs incorporate the truncated convex penalty for the point-wise surface distance changes between two surfaces. Similar to the single surface segmentation method, four cases are distinguished on whether the segmented surfaces from previous iteration are in the current search space for the same pair of columns of two adjoining surfaces. In all cases the convex penalty is enforced by adding arcs each with a weight of the second derivative of the convex function $f(.)$. The truncation and minimum separation constraints are incorporated by additional arcs from case to case. The graph G_m is then iteratively solved by computing a maximum flow which minimizes the energy function $E(S)$ (Eqn. (4.5)). The method terminates while $E(S)$ converges, thus resulting in the final labeling for λ target surfaces.

4.3.3 Graph Construction

In iteration m, we search for each surface S_i in the sub-volume $\mathbf{x}\times\mathbf{y}\times T_{i,m}$ of I. Let us assume at iteration m, the set of labels given by Eq.(4.7) for corresponding columns $col(a,i)$(resp., $col(a,i+1)$) in $D_{i,m}$(resp., $D_{i+1,m}$) is $T_{i,m} = [q_{a,i}, q_{a,i}+1, \ldots q_{a,i}+L-1]$ (resp., $T_{i+1,m} = [q_{a,i+1}, q_{a,i+1}+1, \ldots q_{a,i+1}+L-1]$), i.e., $T_{i,m}$(resp., $T_{i+1,m}$) includes all possible surface positions that S_i(resp., S_{i+1}) can change into at iteration m. For ease of notation we refer each node in the graph with its corresponding label. Denote the initial surface position of S_i on column $col(a)$ at the beginning of the iteration m as $S_{i,m-1}(a)$. At each iteration m, a labeling can either retain its current label $S_{i,m-1}(a)$ or can be changed to a label belonging to the interval $T_{i,m}$. A label retaining arc as described in Section 4.2.1, exists from node s to node $q_{a,i}$ with weight $D_i(S_{i,m-1}(a))$ when $S_{i,m-1}(a) \notin T_{i,m}$.

The task now is to add inter-surface arcs between corresponding columns of sub-graphs $G'_{i,m}s$ to enforce the minimum separation and the truncated convex penalty for surface separation between a pair of adjacent surfaces. The convex penalty is enforced by adding arcs in a similar manner as the single surface detection method with following difference. For all $k = k'$ except when $k = k' = 0$, we put in an arc of weight $+\infty$ from node $q_{a,i} + k$ to $q_{a,i+1} + k'$. Therefore, no st-cut shall be possible when $S_{i+1,m}(a) - S_{i,m}(a) < d_{i,i+1}$ within the interval at iteration m, thus ensuring the minimum separation constraints are not violated within the interval. An example is shown in Fig. 4.10(a).

The final task is to add the truncation arcs to truncate the convex penalty while ensuring the minimum separation constraint is preserved. As discussed earlier, four cases are distinguished for a pair of corresponding columns as follows:

Base Case: $S_{i,m-1}(a) \in T_{i,m}$ and $S_{i+1,m-1}(a) \in T_{i+1,m}$.

Symmetric Cases: $S_{i,m-1}(a) \notin T_{i,m}$ and $S_{i+1,m-1}(a) \in T_{i+1,m}$, $S_{i,m-1}(a) \in T_{i,m}$ and $S_{i+1,m-1}(a) \notin T_{i+1,m}$.

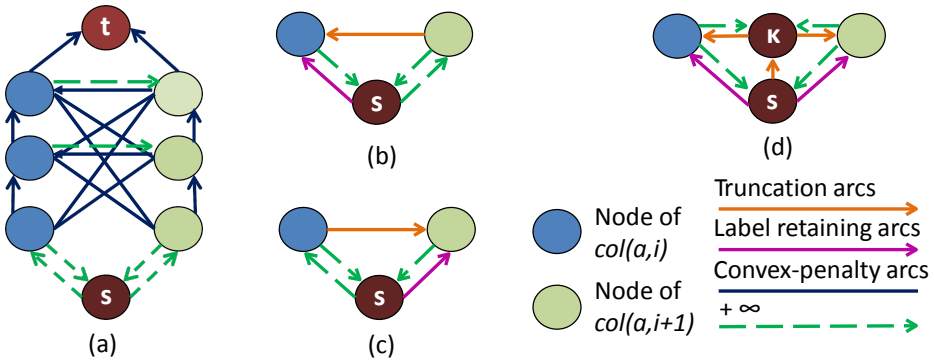


Figure 4.10: (a)An example of graph construction for inter-surface arcs for base case. Additional arcs added to (a) for (b)(c) Symmetric cases, (d) combination case. Length of interval $L = 3$

Combination Case: $S_{i,m-1}(a) \notin T_{i,m}$ and $S_{i+1,m-1}(a) \notin T_{i+1,m}$.

For the first symmetric case when $S_{i,m-1}(a) \notin T_{i,m}$ and $S_{i+1,m-1}(a) \in T_{i+1,m}$ (Case 1), we additionally introduce to the construction shown in Fig. 4.10(a); a truncation arc from node $q_{a,i+1}$ to node $q_{a,i}$ whose weight is $w_a M + \frac{w_a}{2} f(L)$, if $(q_{a,i+1} - S_{i,m-1}(a)) \geq d_{i,i+1}$ and is $+\infty$ otherwise, to encode the truncated penalty and the minimum separation constraint. Note that any finite st-cut including the label retaining arc must also include the truncation arc (Fig. 4.10(b)), hence enforcing the truncated convex penalty with possible overestimation. For the case when $S_{i,m-1}(a) \in T_{i,m}$ and $S_{i+1,m-1}(a) \notin T_{i+1,m}$ (Case 2), we symmetrically add arcs as discussed for Case 1 (Fig. 4.10(c)).

For the combination case we add the arcs of the symmetric cases to the graph of base case while introducing a new node κ. An additional truncation arc is added from source node s to κ with a weight of $H_a(S_{i,m-1}(a), S_{i+1,m-1}(a))$ (Fig. 4.10(d)). Note that any finite st-cut including both the label retaining arcs $((s, q_{a,i}), (s, q_{a,i+1}))$, must also include the truncation arc (s, κ), hence enforcing the truncated convex penalty. This completes the graph construction.

At each iteration, the graph is solved by computing a maximum flow which minimizes the energy function approximately. The set of iterations are continued till the energy cannot be further minimized thereby yielding the final labeling $S_i(\cdot)$.

4.3.4 Experiment Setting

The experiment compares segmentation accuracy of the proposed method (truncated convex prior) and the optimal surface detection method using convex prior without truncation (OSDC) [9]. 20 SD-OCT scans of normal eyes (Type I), 20 SD-OCT scans of eyes with AMD (Type II) and their respective expert manual tracings were obtained from the publicly available repository of datasets Ref. [68]. The 3-D volumes ($1000 \times 100 \times 512$ voxels with voxel size $6.54 \times 67 \times 3.23$ μm^3) for our study were randomly selected from the repository. Segmenting the surfaces simultaneously using OSDC [9] in original resolution is not efficient enough for large data volumes. To make fair comparisons, we first downsample the image by a factor of four in the x direction to reduce the computation time. The datasets were segmented in both their original resolution and down-sampled version by our method to demonstrate the performance and capacity of our method for large clinical datasets. For cases where segmentation was done in the down-sampled version, the resulting segmentation was up-sampled to original resolution for comparison purposes.

The three surfaces simultaneously segmented in this study are S_1-Internal Limiting Membrane (ILM), S_2-inner aspect of retinal pigment epithelium drusen complex (IRPEDC), S_3-outer aspect of Bruch's membrane (OBM) as shown in Fig. 3.1. Comparison was done by calculating the unsigned mean surface positioning errors (UMSP) as described in Equation (4.3) as absolute distances between the computed surfaces and the expert manual tracings in each column of the image. Statistical significance of observed differences was determined by paired Student t-tests for which p value of 0.05 were considered significant.

The data cost volumes (data cost term) were generated (computed) as follows.

First, a $11 \times 11 \times 11$ Gaussian filter with a standard deviation of 11 was applied. To detect S_1 and S_3, a 3-D Sobel filter ($3 \times 3 \times 3$) emphasizing the vertical edges for the dark to bright and bright to dark transitions respectively were applied. To detect S_2, we apply the following operations to each slice of the volume. Edges are extracted using a high pass filter; image is normalized to range from 0 to 1; a binary mask is generated for the region containing S_2 by thresholding of 0.5 and finally mask is applied to the data cost volume for S_1. Parameters are reported for downsampled version of the datasets and are summarized in Table 4.2.

Table 4.2: Parameters used in the experiment. M_x and M_y are truncation factors in x and y directions, M is the truncation factor for the surface separation term of a surface pair, d is the minimum separation between a surface pair. TC - Truncated Convex.

		TC			OSDC	TC	
Surface	Dataset	M_x	M_y	Surface pair	d	M	d
S_1	Type I	30	5	S_1-S_2	30	15	30
S_2	Type I	30	5	S_2-S_3	3	3	3
S_3	Type I	10	2				
S_1	Type II	30	5	S_1-S_2	20	10	20
S_2	Type II	10	5	S_2-S_3	3	5	3
S_3	Type II	5	2				

For both the methods, we use a linear convex function $f(x) = |x|$. For our method, an interval length $L = 2$ was used and surface S_1 (resp., S_2, S_3) was initialized as 0 (resp., $d_{1,2}$, $d_{1,2}$+$d_{2,3}$). The parameters and the weight coefficients (w_{ab}, w_a) were experimentally determined by testing on a similar group of datasets (with the same data size) obtained from the same repository [68] for best results.

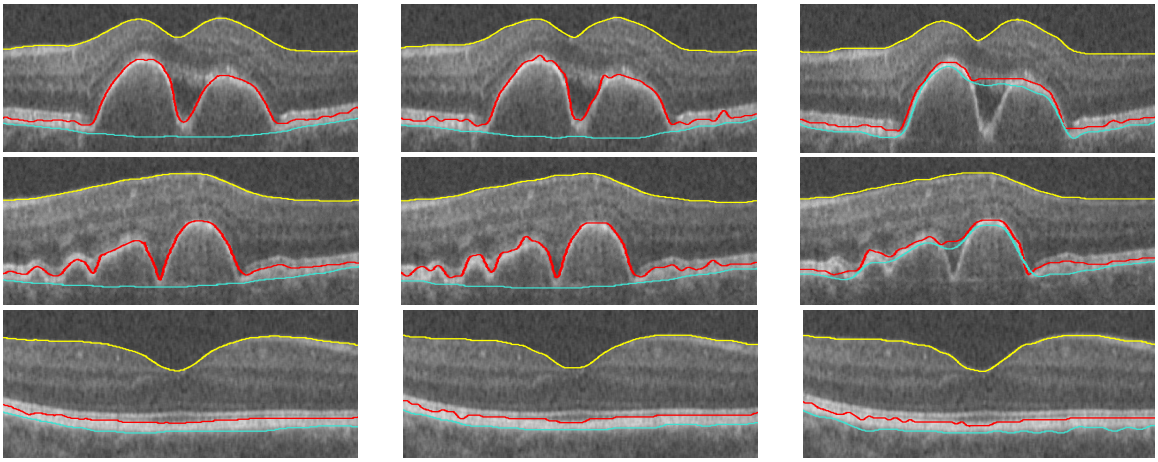


Figure 4.11: Top two rows show image slices of Type II and bottom row shows image slice of Type I. Yellow - ILM, Red - IRPEDC and Blue - OBM. (a)Expert manual tracing, segmentation using (b)our method, (c)OSDC method.

4.3.5 Results

Illustrative results of our proposed method and the OSDC for downsampled data can be seen in Fig. 4.11. Quantitative comparison between our method and OSDC is summarized in Table 4.3 and Table 4.4. For the downsampled version of the datasets, our method produced significantly lower UMSP for surfaces S_1 ($p < 0.05$), S_2 ($p < 0.03$) and S_3 ($p < 0.002$) in Type II datasets. In type I datasets, our method significantly lowered UMSP for surface S_3 ($p < 0.05$). Comparisons were also made between the segmentations using our method in original resolution and OSDC in downsampled version. Our method significantly improved the UMSP for S_1 ($p < 0.001$), S_2 ($p < 0.006$) and S_3 ($p < 0.001$) in both types of the datasets.

For the downsampled version of the datasets, our method with average computation time of 539 seconds is much faster than OSDC method with average computation time of 3671 seconds. Average computation time using our method was 3394 seconds for datasets in original resolution.

Table 4.3: Unsigned mean surface positioning errors (UMSP) (mean ± standard deviation)μm for downsampled version of data. Obsv - Expert manual tracing, TC -Truncated Convex, OSDC - Optimal Surface Detection with Convex priors.

	Data in downsampled resolution			
	Normal eye (Type I)		Eye with AMD (Type II)	
Surface	TC vs. Obsv	OSDC vs. Obsv	TC vs. Obsv	OSDC vs. Obsv
S_1	3.62 ± 0.23	3.67 ± 0.30	3.95 ± 0.72	4.24 ± 0.56
S_2	5.56 ± 2.13	5.77 ± 2.41	6.86 ± 2.04	8.06 ± 2.79
S_3	3.69 ± 0.70	3.98 ± 0.60	4.56 ± 1.40	11.65 ± 8.72
Overall	4.29 ± 1.02	4.47 ± 1.10	5.12 ±1.39	7.98 ± 4.02

Table 4.4: Unsigned mean surface positioning errors (UMSP) (mean ± standard deviation)μm for data in original resolution. Obsv - Expert manual tracing, TC -Truncated Convex.

	Data in original resolution	
	Type I	Type II
Surface	TC vs. Obsv	TC vs. Obsv
S_1	1.99 ± 0.36	2.07 ± 0.38
S_2	4.72 ± 1.68	6.49 ± 2.46
S_3	2.95 ± 0.41	3.64 ± 0.62
Overall	3.32 ± 0.82	4.06 ± 1.15

4.4 Segmentation of Surfaces in SD-OCT Volumes with Retinal Pigment Epithelial Detachment (PED) Using Truncated Convex Priors(Aim 1.3)

Retinal pigment epithelial detachment (PED) is a prominent feature of age-related macular degeneration (AMD), which is the leading cause of severe vision loss in people over age 60 [61] [69]. Initial studies have demonstrated that PEDs may evolve over time and patients diagnosed with PED associated with AMD frequently have co-existing choroidal neovascularization (CNV), or have a higher risk of developing CNV, which can eventually cause severe visual acuity loss [61] [70]. In SD-OCT images, the

Retinal Pigment Epithelium (RPE) appears as a bright layer, and PED appears as a localized, relatively pronounced dome-shaped elevation of the RPE layer, as shown in Fig 3.2. Automated segmentation of Outer Retinal Pigment Epithelium (ORPE) and Inner Choroid (IC) as shown in Fig 3.2, are of significance for the identification of PED as the segmentation information, the morphological and optical features of each individual layer and their difference from normal ones can be analyzed, which can improve the understanding of the disease progression and diagnosis in a clinical setting.

Tradional graph search approaches usually experience difficulty when additional structures exist because of pathology such as retinal pigment epithelial detachment (PED) as discussed in Section 3.1.1. In these cases, layer segmentation becomes challenging due to the following two reasons. First, the surface terrains may have sharp changes in smoothness. Secondly, there may be abrupt changes in the surface distance between a pair of adjacent surfaces. Therefore, the prior knowledge about morphological and optical features used for normal retinal image segmentation may not be valid. Several methods have been reported for segmentation of OCT images with PED. Penha et al. [71] utilized the software on the commercially available Cirrus SD-OCT in conjunction with proposed method of Gregori *et al.* [72] to detect the RPE and subsequently the PED. Ding *et al.* [73] detected the top and bottom surfaces of the retina as constraints for subretinal and sub-RPE fluid detection. Quellec *et al.* [74] segmented 11 surfaces in OCT images with fluid-associated abnormalities. Sun *et al.* [75] proposed an automated framework to segment PED by effectively combining the multi-scale graph search, shape-constrained graph cut and mathematical morphology algorithm.

However, for all the works reported in [71] [73] [74], no evaluation of layer segmentation accuracy was given. Shi *et al.* [76] segmented 11 retinal surfaces in OCT images with PED by utilizing the multi-resolution graph search method [62]. However, due

to the usage of loose constraints, the RPE floor is corrected as a post processing step at places where it goes below the estimated normal RPE floor. Therefore the method may have a problem in more severe cases to define the region of interest required for the multi-resolution approach. In this aim, we simultaneously segment three PED related surfaces: Internal Limiting Membrane (ILM), Outer Retinal Pigment Epithelium (ORPE) and Inner Choroid (IC) in SD-OCT images with presence of PED using the developed method for multiple surface segmentation with truncated convex priors.

4.4.1 Experiment Setting

The experiment compares segmentation accuracy of the truncated convex method (TC) [77], graph search method with hard constraints (GSH) [2] and graph search with convex constraints without truncation (GSC) [9] on SD-OCT volumes with presence of PED. The study inlcuded 22 SD-OCT (Heidelberg Spectralis) volume cubes from 22 eyes of 22 subjects (one eye per patient) with the presence of submacular vascularized/fibrovascular PED. Each OCT volume scan consisted of a macular cube of $1024 \times 37 \times 496$ voxels with an average physical size of $5.76 \times 4.44 \times 1.92 mm^3$. The three surfaces segmented in this study are S_1 - ILM, S_2 - ORPE and S_3 - IC as shown in Fig 3.2. Comparison was done by calculating the unsigned mean surface positioning errors (UMSP) as described in Equation (4.3) as absolute distances between the computed surfaces and the expert manual tracings in each column of the image. Statistical significance of observed differences was determined by paired Student t-tests for which p value of 0.05 were considered significant.

The data cost volumes were generated as follows. First, a $11 \times 11 \times 11$ Gaussian filter with a standard deviation of 11 was applied. To detect S_1 and S_3, a 3-D Sobel filter ($3 \times 3 \times 3$) emphasizing the vertical edges for the dark to bright and bright to dark transitions respectively were applied. To detect S_2, we apply the following operations to each slice of the volume. Edges are extracted using a high pass filter; image is normalized to range from 0 to 1; a binary mask is generated for the region

containing S_2 by thresholding of 0.5 and finally mask is applied to the data cost volume for S_1. As a pre-processing step, the surface S_2 for the first B-scan of each SD-OCT volume is segmented using the TC method. The segmented surface is then used to flatten [62] the entire SD-OCT volume. The rationale behind the same is that in the given data, the detachment is not present in the first few B-scans and therefore, the segmentation can be used as a reference for image flattening. The methods in this study were applied to the flattened cost volumes and the resultant segmentation was mapped back to the original unflattend image. The parameters used in this study are summarized in Table 4.5 .

Table 4.5: Parameters used in the experiment. Δ_x and Δ_y are the hard constraints used in x and y directions, M_x and M_y are truncation factors in x and y directions, M is the truncation factor for the surface separation term of a surface pair, d is the minimum separation between a surface pair.

	GSH		TC			GSH	GSC	TC	
Surface	Δ_x	Δ_y	M_x	M_y	Surface pair	d	d	M	d
S_1	2	2	30	5	S_1-S_2	45	45	8	45
S_2	3	2	30	5	S_2-S_3	0	0	3	0
S_3	1	1	15	15					

For the TC and GSC method, we used a a linear convex function $f(x) = |x|$. For the TC method, an interval length $L = 2$ was used and surface S_1 (resp., S_2, S_3) was initialized as 0 (resp., $d_{1,2}$, $d_{1,2}$+$d_{2,3}$). The parameters and weight co-efficients used in each of the methods were experimentally determined by testing the data with a variety of different parameters to achieve the best results with respect to each method.

4.4.2 Results

Quantitative comparison between the TC, GSH and GSC methods are summarized in Table 4.6. The TC method significantly lowered the UMSP error for surfaces S_1 ($p < 0.05$), S_2 ($p < 0.01$) and S_3 ($p < 0.001$) compared to the GSH and GSC method. The results also show the substantial improvement in segmentation accuracy of surface S_3 which is generally the most difficult to segment. The results also show that the GSH method is more accurate than GSC method for surface S_2, since the convex penalty between surfaces S_2 and S_3 is overpenalizing the surface distance for the GSC method. However, this is difficult to control since the minimum separation between those two surface is 0.

Table 4.6: Unsigned mean surface positioning errors (UMSP) (mean ± standard deviation)μm. Obsv - Expert manual tracing.

Surface	GSH vs. Obsv	GSC vs. Obsv	TC vs. Obsv
S_1	4.02 ± 0.96	3.48 ± 0.77	3.28 ± 0.69
S_2	11.02 ± 3.25	8.93 ± 2.39	7.34 ± 1.78
S_3	13.69 ± 5.49	15.24 ± 7.31	8.16 ± 2.59

Illustrative results are presented in Fig. 4.12. Herein, all rows in a given column shows the same B-scan from a SD-OCT volume with PED. The first row shows the expert mannual tracings, the second, third and fourth row shows the automated segmentation results for the GSH, GSC and TC methods. The illustrative results demonstrates the qualitative superior performance of the TC method. Fig. 4.12 also demonstrates that the GSH and GSC method struggle to accurately segment surfaces S_2 and S_3. This is because with the hard constraints the GSH method over smoothes the steep changes in the surface profiles while lack of regularization (even though constrained) causes surface S_3 to become irregular as compared to the manual tracing

for surface S_3. It can also be observed that the GSC method provies a reasonable segmentation for surface S_2 but the over penalizes the surface distance between surfaces S_2 and S_3, resulting in a mutual attraction between the surfaces. The TC method, on the other hand, exploits the truncation property and imposes a convex penalty where the surface separation is small between surfaces S_2 and S_3 and truncates it when the separation is bigger, thus providing a better modelling of the surface profiles which is reflected in the results shown.

4.5 Discussion

The developed truncated convex method works better than the graph search method with hard constraints and the graph search method with convex constraints because the smoothness constraint can incorporate the traditional constraints, as well as, ensures that there is no over-penalizing the steep difference in labeling using the truncation factor. The results shown in Section 4.2.5, demonstrates that the truncated convex method allows for the sharp change in the surface smoothness of the ILM surface while the graph search method with hard constraints oversmoothes the surface in the cup region. Consequently, increasing the hard constraint to allow for capturing of the steep change in the surface profile shall result in an uncosntrained surface segmentation which shall also be susceptible to noise. The results shown in Section 4.3.5, show that the truncated convex method is more accurate for multiple surface segmentation in the presence of pathology and results in superior segmentation compared to graph search with convex constraints which tends to over penalize the steep change in surface smoothness and surface separation. The applicability of the approach can also be seen from the results shown in Section 4.4.2, which clearly demonstrate that the truncated convex method results in better segmentation accuracy compared to the GSH and GSC methods. In addition, the approach does not require multiple-resolution processing and is much faster in segmenting SD-OCT volumes in original resolution. Infact, for the segmentation of surfaces in SD-OCT

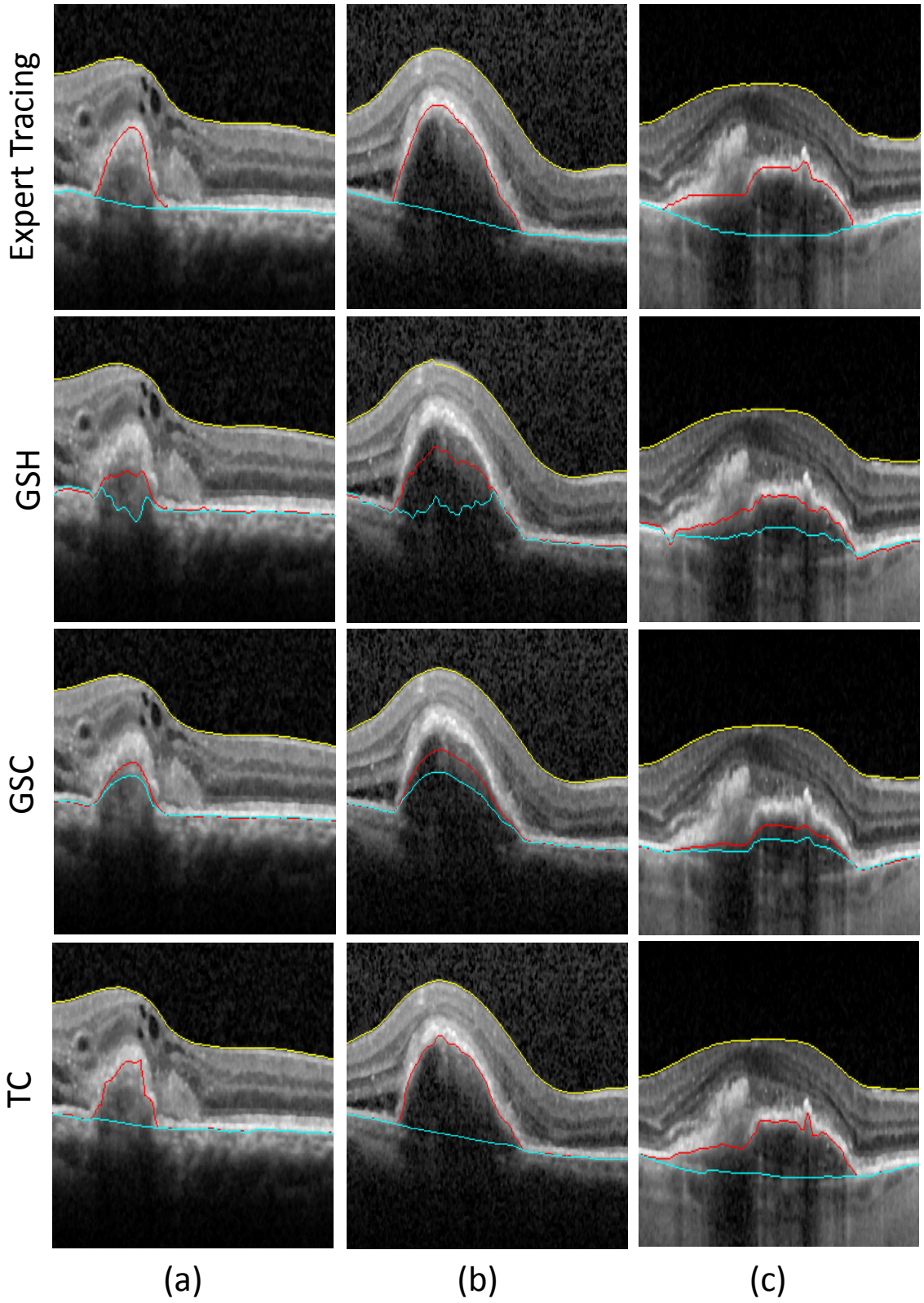


Figure 4.12: Each image in a given column shows the same B-scan from a SD-OCT volume with PED. Yellow - ILM, Red - ORPE and Blue - IC.

volumes with presence of PED, the ORPE surface can be further constrained by also segmenting the inner aspect of the retinal pigment epithelium and imposing a convex penalty between the two. This shall cause for a mutual attraction between the two surfaces and therefore, pull the ORPE surface towards more accurate segmentation.

However, a few limitations are noticed in this work. First, the method does not guarantee a globally optimal solution since a truncated convex function is submodular in nature and hence is optimized using an approximate algorithm. Second, the method introduces two extra parameters (truncation factor and interval length) compared to the graph search based methods. Thus, with the increase in parameters, the tuning of these parameters become more complex compared to graph search based methods. Therefore, more advanced automated techniques may be used for training of the parameters. Third, by introducing the truncation, in theory, thus results in unconstraining the surface profile after the truncation factor limit, which may cause complications in certain applications. However, most of the graph search methods, rely heavily on generation of a good enough data cost term and such a data cost term shall partially limit the surface from becoming unconstrained.

4.6 Conclusion

The focus of this chapter was on the development of a novel mutilpe surface segmentation method which segments surface with a presence of steep change in surface smoothness and/or sharp changes in surface distance between two surfaces, as in the cases of SD-OCT images with presence of pathology. In this work, the method was developed, applied to various such segmentation problems and demonstrated the superior performance of the developed method in terms of both segmentation accuracy and running time, compared to graph search method with hard constraints and convex constraints. The results also demonstrate that the developed method is more efficient and capable of segmenting large OCT volumes in original resolution without using a multiple resolution approach. The results clearly show the improved

performance of the proposed approach and applicability in future studies and clinical applications.

CHAPTER 5
OPTIMAL MULTIPLE SURFACE SEGMENTATION WITH CONVEX PRIORS IN IRREGULARLY SAMPLED SPACE (AIM 2)

5.1 Introduction

The **objective of Aim 2** is the development of an optimal graph based mutiple surface segmentation method with convex priors in irregularly sampled space. In this aim the extension of the framework presented in Ref. [65] to incorporate convex priors to achieve subvoxel accuracy which requires to model the surface smoothness constraints in the irregularly sampled space. The proposed method is a generalization of the graph based optimal surface segmentation with convex priors in the regularly sampled space (equidistant spacing between voxel centers). Consequently, the graph constructed in the regularly sampled space forms a special case in the irregularly sampled space framework.

The proposed method shall have a potential to handle various interesting problems and is evaluated in the following aspects. First, the method can be used to achieve subvoxel accuracy while using convex constraints, thus providing flexibility in modeling the target surfaces to achieve higher accuracy. Second, the method may provide for an avenue to perform super resolution segmentation with sufficient accuracy, i.e achieve adequate segmentation accuracy by operating in the downsampled version of data as compared to segmenting the data in the original resolution, thus providing a prospective alternative for more cost effective imaging with lower resolution image acquisition hardware. Last, typically the design of the region of interest in the multi-resolution graph search method [62] defines a region of interest with a constant length for each column (like a uniformly sized band based on the pre-segmentation result). The proposed method shall provide for a framework to allow segmentation of surfaces with non-uniformly sized region of interests and region of interest which may comprise of union of multiple disjoint region of interests for the same target segmentation. Such a flexibility, may allow fast segmentation of the target surfaces

in the original resolution itself and hence could be used as a potential tool by the multi-resolution graph search method.

The segmentation problem is transformed into an energy minimization problem [2] [67] [4]. A graph is then constructed, wherein the graph nodes correspond to the center of evenly distributed voxels (equidistant spacing between adjoining nodes). Edges are added between the nodes in the graph to correctly encode the different terms in the energy function. The energy function can then be minimized using a minimum *s-t* cut [2] [8]. The resultant minimum *s-t* cut corresponds to the surface position of the target surface in the voxel grid. The method requires appropriate encoding of the following three types of energy terms as described in Section 2.1 into the graph construction. A data term (also commonly known as the data cost term) which measures the inverse likelihood of all voxels on a surface, a surface smoothness term (surface smoothness constraint) which specifies the regularity of the target surfaces and finally, a surface separation term (surface separation constraint) which governs the feasible distance between two adjacent surfaces.

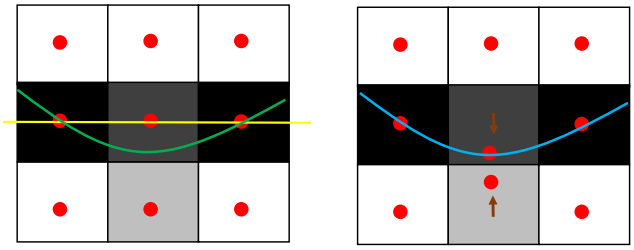
Volumetric images are obtained by discretizing the continuous intensity function uniformly sampled by sensors, resulting in partial volume effects [63] [64]. Partial volume effects are inherent in images as voxels 'combine' partial information from various features (such as tissues) of the imaged object. The spatial resolution in images is limited by the detector/sensor design and by the reconstruction process, which results in 3-D image blurring introduced by the finite spatial resolution of the imaging sysytem [78]. Mathematically, the finite resolution effect is described by a 3-D convolution operation, where the image is formed by the convolution of the actual source with the 3-D point spread function of the imaging system, which causes spillover between regions. The signal intensity in each voxel is the mean of signal intensities of the underlying tissues included in that voxel. Even if the imaging system has perfect spatial resolution, there is still some partial volume effect

because of image sampling [78]. The partial volume effects can be ignored if the intensity or the gradient is measured at the center of each voxel to assign the costs and the graph is created with equidistant nodes. However, since volumetric data is typically represented as an orthogonal matrix of intensities, the surface segmentation cannot achieve a precision greater than a single unit voxel, i.e. the distance between two adjoining nodes in the graph space. Accuracy higher than a single unit voxel (subvoxel accuracy) can be attained by exploiting partial volume effects in the image volumes [65] [79] which leads to non-equidistant spacing between the adjoining graph nodes resulting in an irregularly sampled space.

The ignored partial volume information can be utilized by computing a displacement field directly from the volumetric data [65] to identify the subvoxel accurate location of the centers within each voxel, thus requiring a generalized construction of the graph with non-equidistant spacing between orthogonal adjoining nodes (irregularly sampled space). Increased subvoxel segmentation accuracy attained by exploiting the partial volume effects has the potential for better diagnosis and treatment of disease. Or, instead, equivalent segmentation accuracy with lower resolution image acquisition hardware allows for more cost-effective imaging.

The optimal surface segmentation technique employing the different types of smoothness constraints as discussed above is not capable of segmenting surfaces with subvoxel accuracy in a volume which requires segmentation in a grid comprising of non uniformly sampled voxels where the spacing between the orthogonally adjoining nodes is non-equidistant.

To address this problem, a subvoxel accurate graph search method [65] was developed to simultaneously segment multiple surfaces in a volumetric image by exploiting the additional partial volume information in the voxels. A displacement field is computed from the original volumetric data. The method first creates the graph using the conventional optimal surface segmentation method [2], then deforms it using a dis-



(a) (b)

Figure 5.1: Example of a 3×3 voxel grid to demonstrate subvoxel accuracy. Each voxel is represented by a red node in the graph space. (a)Graph nodes with equidistant spacing between them. True subvoxel accurate surface is shown in green. The segmented surface using optimal surface segmentation method with hard constraints is shown in yellow. (b) The displacement field derived from the grid is applied to the central nodes displacing the centers to exploit the information from the partial volume effect shown by brown arrows. The resultant segmentation with the subvoxel accurate graph search is shown in blue.

placement field and finally adjusts the inter-column edges and inter-surface edges to incorporate the modification of these constraints. Specifically, such a deformation results in non-equidistant spacing between the adjoining nodes which can be considered equivalent to a generalized case of a cube volume formed by non-uniform sampling along the z dimension for the purposes of 3-D surface segmentation. The method demonstrated achievement of subvoxel accuracy compared to the traditionally used optimal surface segmentation method [2]. An example is shown in Fig. 5.1. However, the method employs hard surface smoothness which does not allow flexibility in constraining surfaces. Specifically, our approach was previously not capable of incorporating a convex surface smoothness constraint in the graph with non-equidistant spacing between adjoining nodes.

The main novelty of this work is is extension of the framework presented in Ref. [65] to incorporate convex surface smoothness constraints for multiple surface segmentation in irregularly sampled space. The proposed method is a generalization of the graph based optimal surface segmentation with convex priors [9] in the

regularly sampled space. Consequently, the graph constructed in the regularly sampled space forms a special case in the irregularly sampled space framework where the spacing between the adjoining nodes is set to be a constant (equidistant). Usage of a convex prior allows for incorporation of many different prior information in the graph framework as discussed previously while attaining subvoxel accuracy. Unlike the subvoxel accurate graph search method [65], the proposed method does not require a two step process to create the graph by the conventional method and then readjust the edges, but instead provides a one step function to add edges between nodes from two neighboring columns to incorporate the convex prior. The method provides the globally optimal solution by directly solving the problem in the irregularly sampled space which fundamentally distinguishes the approach from the local adjustments made to the segmentation in the regularly sampled space as reported in [79].

In the following sections, explaination of the formulation for the optimal surface segmentation method in the regularly sampled space, formulation and description of the novel graph construction to incorporate the convex smoothness constraints in the irregularly sampled space is presented. Next, the evaluation is performed on the spectral domain optical coherence tomography (SD-OCT) volumes of the retina and intravascular multi-frame ultrasound (IVUS) image datasets for validation and applicability of the method to demonstrate subvoxel and super resolution segmentation accuracy compared to optimal surface segmentation method with convex priors in regularly sampled space [9].

5.2 Optimal Multiple Surface Segmentation with Convex Priors in Irregularly Sampled Space (Aim 2.1)

5.2.1 Problem Formulation and Energy Function

The problem formulation for the widely used optimal surface segmentation methods [2] [7] [9] is described as follows. Consider a volume $I(x, y, z)$ of size $X \times Y \times Z$.

A surface is defined as a function $S(x, y)$, where $x \in \mathbf{x} = \{0, 1, ...X - 1\}$, $y \in \mathbf{y}$ $=\{0, 1, ...Y - 1\}$ and $S(x, y) \in \mathbf{z} = \{0, 1, ...Z - 1\}$. It is worth noting that the center of voxels are uniformly sampled. Each (x, y)-pair corresponds to a voxel *column* $\{(I(x, y, z)|z = 0, 1, \ldots, Z-1\}$. We use a and b to denote two columns corresponding to two neighbouring (x, y)-pairs in the domain $\mathbf{x} \times \mathbf{y}$ and N_s to denote the neighbourhood setting of image domain. The function $S(a)$ can be viewed as labeling for a with the label set $\mathbf{z}$ ($S(a) \in \mathbf{z}$). For simultaneously segmenting $\lambda(\lambda \geq 2)$ distinct but interrelated surfaces, the goal of the problem is to seek the globally optimal surfaces $S_i(a)$, where $i = 1, 2, ...\lambda$ in I with minimum separation $d_{j,j+1}$ where $j = 1, 2, ...\lambda - 1$ between each adjacent pair of surfaces S_j and S_{j+1}.

The problem is transformed into an energy minimization problem. The energy function $E(S)$ takes the following form as shown in Eqn. (5.1):

$$\begin{aligned} E(S) = &\sum_{i=1}^{\lambda}\Big(\sum_{a\in\mathbf{x}\times\mathbf{y}} D_i(S_i(a)) + \sum_{(a,b)\in N_s} V_{ab}(S_i(a), S_i(b))\Big) \\ &+ \sum_{i=1}^{\lambda-1}\sum_{a\in\mathbf{x}\times\mathbf{y}} H_a(S_{i+1}(a), S_i(a)) \end{aligned} \tag{5.1}$$

The data cost term $\sum_{a\in\mathbf{x}\times\mathbf{y}} D_i(S_i(a))$ measures the total cost of all voxels on a surface S_i. The surface smoothness term $\sum_{(a,b)\in N_s} V_{ab}(S_i(a), S_i(b))$ constrains the connectivity of a surface in 3-D and regularizes the surface. Intuitively, this defines how rigid the surface is. The surface separation term $H_a(S_i(a), S_{i+1}(a))$ constrains the distance of surface S_i to S_{i+1}. The energy function is appropriately encoded in a graph. A minimum s-t cut is then computed on the graph to get the solutions for the target surfaces S_i's.

Typically graph construction is done with equidistant spacing between the adjoining nodes (regularly sampled space). The main novelty of this work is to allow for optimal surface segmentation in the irregularly sampled space with convex surface

smoothness constraints by allowing non-equidistant spacing between the nodes.

We formulate the multiple surface segmentation problem in a similar manner for the irregularly sampled space. Consider a volume $\tilde{I}(x, y, \tilde{z})$ where $x \in \mathbf{x} = \{0, 1, ...X-1\}$, $y \in \mathbf{y} = \{0, 1, ...Y-1\}$ and $\tilde{z} \in \mathbb{R}$. Each (x, y)-pair corresponds to a *column* $\{(\tilde{I}(x, y, \tilde{z})|\tilde{z} \in \mathbb{R}$, denoted by $col(x, y)$. Assume each $col(x, y)$ has exactly Z elements obtained by sampling strictly in the increasing order along the $\tilde{z}$ direction, resulting in the volume $I(x, y, z)$ of size $X \times Y \times Z$, where $x \in \mathbf{x} = \{0, 1, ...X-1\}$, $y \in \mathbf{y} = \{0, 1, ...Y-1\}$ and $z \in \mathbf{z} = \{0, 1, ...Z-1\}$, thus possibly allowing for non-equidistant spacing between two adjacent elements in the column. As discussed previously a and b are used to denote two neighbouring columns.

We define a mapping function for each column a as $L_a : \{0, 1, ...Z-1\} \rightarrow \mathbb{R}$ which maps the index of sampled points in $I(a, z)$ to $\tilde{I}(a, \tilde{z})$. For example, $L_a(i)$ denotes the $\tilde{z}$ coordinate of the i+1-th sample along column a, and $L_a(i+1) > L_a(i)$ because of the strictly increasing order of sampling along column a. An example is shown in Fig. 5.2. Further, a surface function for column a is defined as $S(a)$, where $S(a) \in \mathbf{z} = \{0, 1, ...Z-1\}$. The function $L_a(S(a))$ can be viewed as labeling for surface $S(a)$ with the label set $\mathbb{R}$ ($L_a(S(a)) \in \mathbb{R}$). For simultaneously segmenting λ ($\lambda \geq 2$) surfaces, the goal of the problem is to seek the labeling for surfaces $L_a(S_i(a))$ where $i = 1, 2 \ldots \lambda$ in I with minimum separation $d_{j,j+1}$ where $j = 1, 2, \ldots \lambda - 1$ between adjacent pair of surfaces. It is to be noted, that the surfaces are ordered, i.e, $L_a(S_{i+1}(a)) \geq L_a(S_i(a))$.

The corresponding energy function for this formulation is shown in Equation 5.2:

$$\begin{aligned} E(S) = & \sum_{i=1}^{\lambda} \Big(\sum_{a \in \mathbf{x} \times \mathbf{y}} D_i(L_a(S_i(a))) \\ & + \sum_{(a,b) \in N_s} V_{ab}(L_a(S_i(a)), L_b(S_i(b))) \\ & + \sum_{i=1}^{\lambda-1} \sum_{a \in \mathbf{x} \times \mathbf{y}} H_a(L_a(S_{i+1}(a)), L_a(S_i(a))) \end{aligned} \tag{5.2}$$

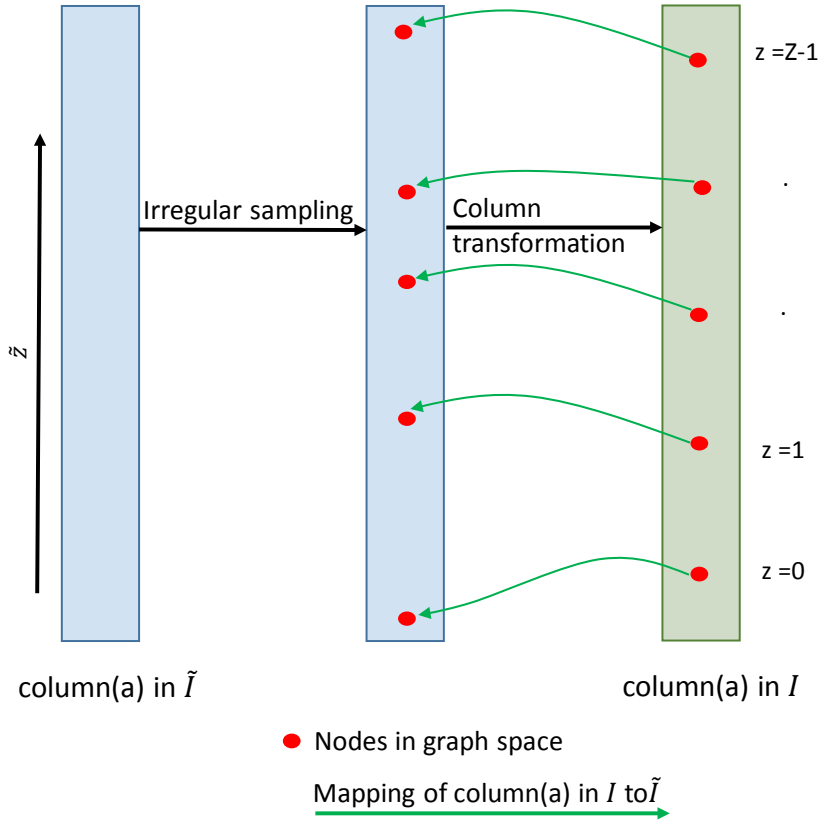


Figure 5.2: Example of column structure for irregularly sampled space using mapping function.

Herein, the surface smoothness term is modelled as a convex function as shown in Equation (5.3).

$$V_{ab}(L_a(S_i(a)), L_b(S_i(b))) = \psi(L_a(S_i(a)) - L_b(S_i(b))) \tag{5.3}$$

where, $\psi(.)$ is a convex function, and without loss of generality, we assume that $\psi(0) = 0$ [7].

The surface separation term is modelled as a hard constraint for enforcing the minimum separation between a pair of surfaces as shown in Equation (5.4).

$$H_a(L_a(S_{i+1}(a)), L_a(S_i(a))) = \begin{cases} \infty, & \text{if } L_a(S_{i+1}(a)) - L_a(S_i(a)) < d_{i,i+1} \\ 0, & \text{otherwise} \end{cases} \tag{5.4}$$

where $d_{i,i+1}$ is the minimum separation between a pair of adjacent surfaces. The method is also capable of incorporating a convex surface separation penalty while enforcing a minimum separation constraint in the irregularly sampled space using the same framework.

5.2.2 Graph Construction

The high level idea of the graph construction for our method is similar to the traditionally used optimal surface segmentation methods. For each surface S_i, a subgraph G_i is constructed. Herein, the intra-column edges are added to enforce surface monotonicity and encode the data term for cost volume D_i (for searching S_i). Inter-column edges are added between a pair of neighbouring columns a and b to enforce the surface smoothness penalty term $V_{ab}(.)$.

The graph G for the simultaneous search of all λ surfaces consists of the union of those λ subgraphs G_i's. Furthermore, inter-surface edges are added between the

corresponding columns of subgraphs G_i and G_{i+1} to incorporate the surface separation term for surface distance changes between two surfaces. A pair of columns with respect to the same (x, y)-pairs in the domain $\mathbf{x} \times \mathbf{y}$ of subgraphs G_i, G_{i+1} for two adjacent surfaces is defined as corresponding columns. The graph G is then solved by computing a maximum flow which minimizes the energy function $E(S)$ (Equation. (5.2)). The positions of the λ target surfaces are obtained by mapping the resultant solution to $\mathbb{R}$ space using the mapping function $L_a(.)$.

The graph is constructed using the cost volumes generated for λ surfaces from volume $I(x, y, z)$. Each element in the cost volume D_i to search S_i is represented by a node $n_i(a, z)$ $(z \in \mathbf{z})$ in G_i. The following edges are added to incorporate the different energy terms:

5.2.2.1 Intra-column Edges

To ensure the monotonicity of the target surfaces (i.e., the target surface intersects each column exactly one time) and encode the data cost term; intra-column edges are added to each subgraph G_i as described in Ref. [2]. Along every column a for surface S_i, each node $n_i(a, z)(z > 0)$ has a directed edge with $+\infty$ weight to the node immediately below it and an edge with $D_i(L_a(z-1))$ weight in the opposite direction. Additionally, an edge with $+\infty$ weight is added from the source node s to each node $n_i(a, 0)$ and an edge with $D_i(L_a(Z-1))$ weight is added from node $n_i(a, Z-1)$ to the terminal node t.

Any s-t cut with finite cost contains only one of the finite weight edges $D_i(L_a(.))$ for each column a, thus enforcing surface monotonicity. This is because, if any s-t cut included more than one finite weight edges, then by construction it must include at least one infinite weight edge thereby making its cost infinite. Therefore, any finite s-t cut shall intersect each column exactly one time.

5.2.2.2 Inter-column Edges

Inter-column arcs are added between pairs of neighbouring columns a and b to each subgraph G_i to encode the surface smoothness term. The incorporation of a convex smoothness term is presented here. Denote a function operator $f(r_1, r_2)$ as shown in Equation (5.5).

$$f(r_1, r_2) = \begin{cases} 0\ , & \text{if } r_1 < r_2 \\ \psi(r_1 - r_2), & \text{otherwise} \end{cases} \tag{5.5}$$

where $\psi(.)$ is a convex function.

A general weight setting function $g(.)$ is used for the inter-column edges between two neighboring columns. The following inter-column edges are added :

For all $k_1 \in [0, Z-1]$ and $k_2 \in [1, Z-1]$, a directed edge with weight setting $g(k_1, k_2)$ as shown in Equation (5.6) is added from node $n_i(a, k_1)$ to node $n_i(b, k_2)$. Additionally, a directed edge is added from node $n_i(a, k_1)$ to terminal node t with weight setting $g(k_1, Z)$.

$$\begin{aligned} g(k_1, k_2) = & f(L_a(k_1), L_b(k_2 - 1)) \\ & - f(L_a(k_1 - 1), L_b(k_2 - 1)) - f(L_a(k_1), L_b(k_2)) \\ & + f(L_a(k_1 - 1), L_b(k_2)) \end{aligned} \tag{5.6}$$

Where, if $k_1 = 0$, then $k_1 - 1 \notin \mathbf{z}$, therefore $f(L_a(k_1 - 1), L_b(k_2 - 1)) = f(L_a(k_1 - 1), L_b(k_2)) = 0$ and if $k_2 = Z$, then $k_2 \notin \mathbf{z}$, therefore $f(L_a(k_1), L_b(k_2)) = f(L_a(k_1 - 1), L_b(k_2)) = 0$.

In a similar manner, for all $k_1 \in [0, Z-1]$ and $k_2 \in [1, Z-1]$, edges are constructed from nodes $n_i(b, k_1)$ to nodes $n_i(a, k_2)$ with weight setting $g(k_1, k_2)$ as shown in Equation (5.7). Additionally a directed edge is added from node $n_i(b, k_1)$ to terminal node

t with weight setting $g(k_1, Z)$.

$$\begin{aligned} g(k_1, k_2) = & f(L_b(k_1), L_a(k_2 - 1)) \\ & - f(L_b(k_1 - 1), L_a(k_2 - 1)) - f(L_b(k_1), L_a(k_2)) \\ & + f(L_b(k_1 - 1), L_a(k_2)) \end{aligned} \tag{5.7}$$

It should be noted that weight setting function $g(k_1, k_2)$ in Equation (5.7) is similar to Equation (5.6) with only the column mapping function $L_a(.)$ and $L_b(.)$ interchanged. Also, in practice we only add edges with edge weight greater than zero in the graph.

Lemma 1: For any k_1 and k_2, the function $g(k_1, k_2)$ is non-negative.

Lemma 2: In any finite s-t cut C, the total weight of the edges between any two adjacent columns a and b (denoted by $C_{a,b}$) equals to the surface smoothness cost of the resulting surface S_i with $S_i(a) = k_1$ and $S_i(b) = k_2$, which is $\psi(L_a(k_1) - L_b(k_2))$, where $\psi(.)$ is a convex function.

Proof of Lemma 1 and Lemma 2 are presented in Section 5.3.

Example of a graph construction of two neighbouring columns a and b for a given surface with enforcement of convex surface smoothness constraint is shown in Fig. 5.3. Herein, an edge from $n_i(a, k_1)$ to node $n_i(b, k_2)$ is denoted as $E_i(a_{k_1}, b_{k_2})$ for the i-th surface. For clarity, an edge $E_i(a_{k_1}, b_{k_2})$ is denoted as Type I if $k_2 > k_1$, as Type II if $k_2 = k_1$ and as Type III if $k_2 < k_1$. The respective edge weights in the graph are summarized in Table 5.1. The convex function used in the example is a quadratic, taking the form $\psi(k_1 - k_2) = (k_1 - k_2)^2$.

The following can be verified from the example shown Fig. 5.3:

- The correct cost of cut $C_1 = (21 - 12)^2 = 81$. It can be verified that the inter-column edges contributing to the cost of cut C_1 are Type I edges $E(a_2, b_3)$ and

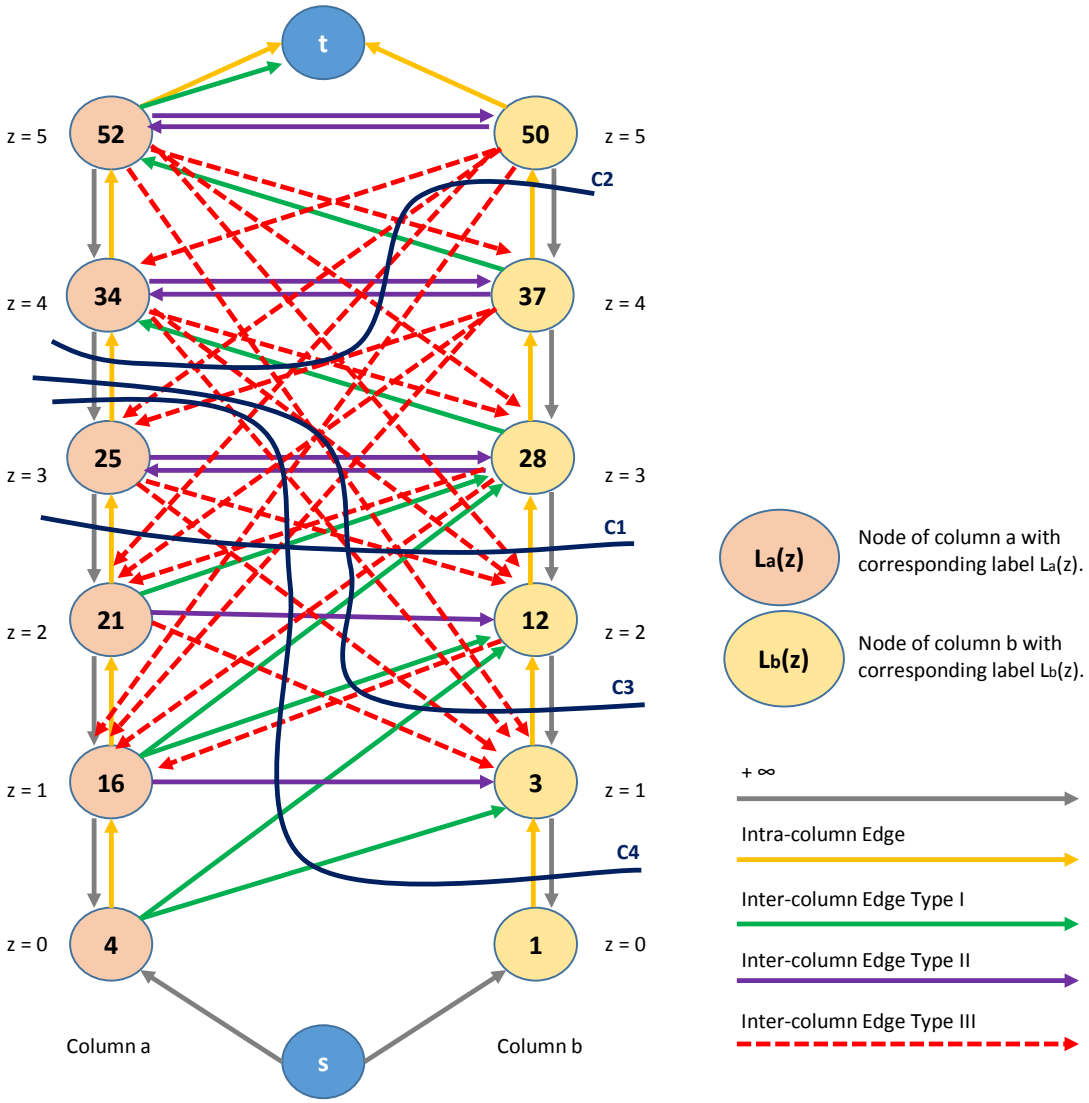


Figure 5.3: Example graph construction of two neighboring columns a and b to demonstrate enforcement of convex surface smoothness constraints in irregularly sampled space.

$E(a_1, b_3)$. Summing the edge weights from Table 5.1, cost of cut $C_1 = 65+16 = 81$.

- The correct cost of cut $C_2 = (25-37)^2 = 144$. It can be verified that the inter-column edges contributing to the cost of cut C_2 are Type I edges $E(b_4, a_5)$, $E(b_3, a_4)$ and Type II edge $E(b_4, a_4)$. Summing the edge weights from Table 5.1, cost of cut $C_2 = 9 + 9 + 126 = 144$.

- The correct cost of cut $C_3 = (25-3)^2 = 484$. It can be verified that the inter-column edges contributing to the cost of cut C_3 are Type I edges $E(a_0, b_2)$, $E(a_1, b_2)$, $E(a_1, b_3)$, $E(a_2, b_3)$, Type II edges $E(a_3, b_3)$, $E(a_2, b_2)$ and Type III edge $E(a_3, b_2)$. Summing the edge weights from Table 5.1, cost of cut $C_3 = 1 + 152 + 16 + 65 + 88 + 90 + 72 = 484$.

- The correct cost of cut $C_4 = (25-1)^2 = 576$. It can be verified that the inter-column edges contributing to the cost of cut C_4 are Type I edges $E(a_0, b_1)$, $E(a_0, b_2)$, $E(a_1, b_2)$, $E(a_1, b_3)$, $E(a_2, b_3)$, Type II edges $E(a_3, b_3)$, $E(a_2, b_2)$, $E(a_1, b_1)$ and Type III edges $E(a_3, b_2)$, $E(a_3, b_1)$, $E(a_2, b_1)$. Summing the edge weights from Table 5.1, cost of cut $C_4 = 8+1+152+16+65+88+90+48+72+16+20 = 576$.

5.2.2.3 Inter-surface Edges

The surface separation term $H_a(.)$ between two adjacent surfaces is enforced by adding edges in a similar manner as described in Ref. [65] from column a in subgraph G_i to corresponding column a in subgraph G_{i+1}. Along every column a in G_i, each node $n_i(a, z)$ has a directed edge with $+\infty$ weight to the node $n_{i+1}(a, z')$, $(z' \in \mathbf{z}, L_a(z') - L_a(z) \geq d_{i,i+1}, L_a(z'-1) - L_a(z) < d_{i,i+1})$. Additionally an edge with $+\infty$ weight is added from node $n_i(a, z)$ to the terminal node t if $L_a(Z-1) - L_a(z) < d_{i,i+1}$.

It can be verified, that no finite s-t cut is possible when $L_a(z') - L_a(z) < d_{i,i+1}$,

Table 5.1: Summary of inter-column edge weights of the graph construction in Fig. 5.3, based on a quadratic convex function of the form $\psi(k_1 - k_2) = (k_1 - k_2)^2$.

Edge	Type	Weight	Edge	Type	Weight
$E(a_0, b_1)$	I	8	$E(b_2, a_1)$	III	64
$E(a_0, b_2)$	I	1	$E(b_3, a_1)$	III	368
$E(a_1, b_1)$	II	48	$E(b_3, a_2)$	III	95
$E(a_1, b_2)$	I	152	$E(b_3, a_3)$	II	40
$E(a_1, b_3)$	I	16	$E(b_3, a_4)$	I	9
$E(a_2, b_1)$	III	20	$E(b_4, a_1)$	III	216
$E(a_2, b_2)$	II	90	$E(b_4, a_2)$	III	90
$E(a_2, b_3)$	I	65	$E(b_4, a_3)$	III	72
$E(a_3, b_1)$	III	16	$E(b_4, a_4)$	II	126
$E(a_3, b_2)$	III	72	$E(b_4, a_5)$	I	9
$E(a_3, b_3)$	II	88	$E(b_5, a_1)$	III	312
$E(a_4, b_1)$	III	36	$E(b_5, a_2)$	III	130
$E(a_4, b_2)$	III	162	$E(b_5, a_3)$	III	104
$E(a_4, b_3)$	III	279	$E(b_5, a_4)$	III	234
$E(a_4, b_4)$	II	36	$E(b_5, a_5)$	II	247
$E(a_5, b_1)$	III	72			
$E(a_5, b_2)$	III	324			
$E(a_5, b_3)$	III	576			
$E(a_5, b_4)$	III	315			
$E(a_5, b_5)$	II	221			
$E(a_5, b_6)$	I	4			

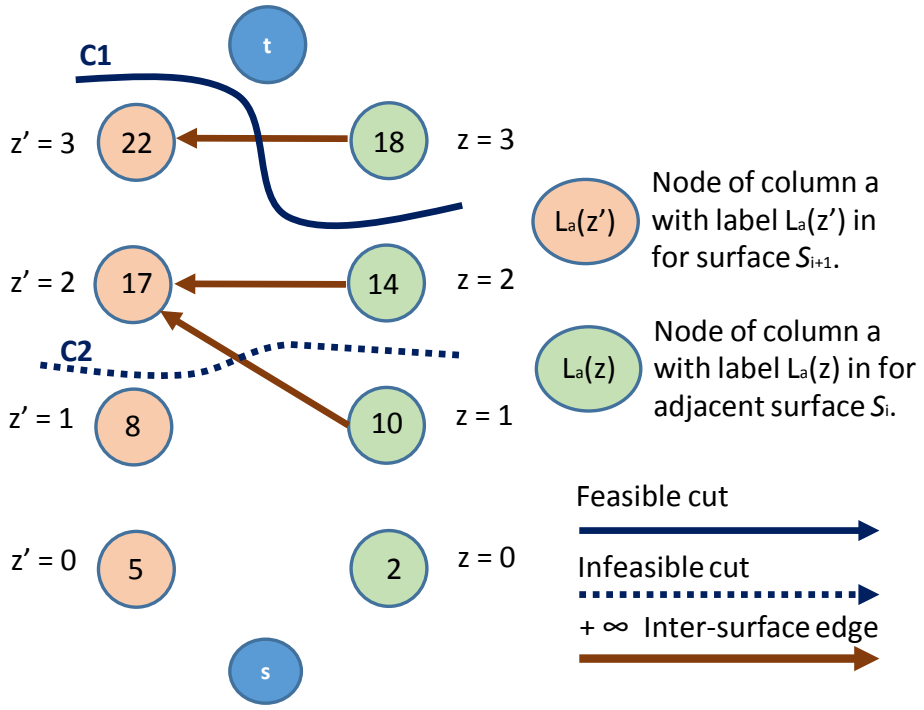


Figure 5.4: An example graph for incorporation of surface separation constraint between two corresponding columns is shown. Only the inter-surface edges are shown for clarity. The minimum separation constraint $d_{i,i+1} = 2$. It can be seen that cut C_1 is a feasible cut since the minimum separation constraint is not violated while cut C_2 is infeasible since the minimum separation constraint is violated as $L_a(z' = 1) - L_a(z = 1) < d_{i,i+1}$

since by construction an inter-surface edge of $+\infty$ weight will be cut, thus making the cost infinite. An example of a graph construction for two corresponding columns of adjacent pair of surfaces with enforcement of the surface separation constraint is shown in Fig. 5.4.

Thus the surface separation term $H_a(.)$ is correctly encoded in graph G.

5.2.3 Surface Recovery from Minimum s-t cut

The minimum s-t cut in the graph uniquely defines optimal λ surfaces S_i where $i = 1, 2 \ldots \lambda$. For a given surface S_i, the surface location for each $col(x, y) \in \mathbf{z}$, where $x \in \mathbf{x}$ and $y \in \mathbf{y}$ is given by the minimum s-t cut. The final surface positions for each

column a is recovered by applying the mapping function $L_a : \{0, 1, ...Z-1\} \rightarrow \mathbb{R}$, where $a \in \mathbf{x} \times \mathbf{y}$, thereby yielding the resultant surface positions for each column $L_a(S_i(a)) \in \tilde{z}$, where $\tilde{z} \in \mathbb{R}$.

5.3 Proof of Correctness of Graph Construction (Aim 2.2)

5.3.1 Proof for Lemma 1

Lemma 1: For any k_1 and k_2, the function $g(k_1, k_2)$ is non-negative.

Proof: Let us consider the function $g(k_1, k_2)$ for edges from column a to neighboring column b as shown in Equation (5.6). We need to prove that $g(k_1, k_2) \geq 0$. The reader should recall because of the strictly increasing order of sampling, $L_a(k_1) > L_a(k_1 - 1)$ and $L_b(k_2) > L_b(k_2 - 1)$. $\psi(\cdot)$ is a convex function with $\psi(0) = 0$. The proof is presented in a case-by-case basis.

Case 1: $L_a(k_1) < L_b(k_2 - 1)$

Thus, $L_a(k_1 - 1) < L_b(k_2 - 1)$. As $L_b(k_2) > L_b(k_2 - 1)$, we have $L_a(k_1) < L_b(k_2)$ and $L_a(k_1 - 1) < L_b(k_2)$. Since $f(r_1, r_2) = 0$ if $r_1 < r_2$. It is straightforward to verify that $g(k_1, k_2) = 0$, in Equation (5.6).

Case 2: $L_a(k_1) \geq L_b(k_2 - 1)$ and $L_a(k_1) < L_b(k_2)$

In this case, as $L_a(k_1) > L_a(k_1 - 1)$, we have $L_a(k_1 - 1) < L_b(k_2)$. Thus, $g(k_1, k_2)$ takes the following form in Equation (5.8).

$$g(k_1, k_2) = f(L_a(k_1), L_b(k_2 - 1)) - f(L_a(k_1 - 1), L_b(k_2 - 1)) \tag{5.8}$$

If $L_a(k_1 - 1) < L_b(k_2 - 1)$, then $g(k_1, k_2) = f(L_a(k_1), L_b(k_2 - 1)) = \psi(L_a(k_1) - L_b(k_2 - 1))$. Thus, $g(k_1, k_2) \geq 0$ as $\psi(L_a(k_1) - L_b(k_2 - 1)) \geq 0$ with $L_a(k_1) \geq L_b(k_2 - 1)$.

If $L_a(k_1 - 1) < L_b(k_2 - 1)$, then $g(k_1, k_2) = \psi(L_a(k_1) - L_b(k_2 - 1)) - \psi(L_a(k_1 - 1) - L_b(k_2 - 1))$. We know that $L_a(k_1) - L_b(k_2 - 1) > L_a(k_1 - 1) - L_b(k_2 - 1) > 0$.

Thus, $g(k_1, k_2) > 0$ as $\psi(0) = 0$.

Therefore, in this case $g(k_1, k_2) > 0$.

Case 3: $L_a(k_1) \geq L_b(k_2)$

In this case, $L_a(k_1) > L_b(k_2-1)$ as $L_b(k_2) > L_b(k_2-1)$. We distinguish three subcases: 1) $L_a(k_1 - 1) < L_b(k_2 - 1)$, 2) $L_a(k_1 - 1) < L_b(k_2)$ and 2) $L_a(k_1 - 1) \geq L_b(k_2 - 1)$, and 3) $L_a(k_1 - 1) \geq L_b(k_2)$.

Subcase 1): If $L_a(k_1 - 1) < L_b(k_2 - 1)$, then

$$\begin{aligned} g(k_1, k_2) &= f(L_a(k_1), L_b(k_2 - 1)) - f(L_a(k_1), L_b(k_2)) \\ &= \psi(L_a(k_1) - L_b(k_2 - 1)) - \psi(L_a(k_1) - L_b(k_2)) \end{aligned}$$

Since $L_b(k_2 - 1) < L_b(k_2)$, we have $L_a(k_1) - L_b(k_2 - 1) > L_a(k_1) - L_b(k_2)$. Thus, $g(k_1, k_2) > 0$ as $\psi(0) = 0$.

Subcase 2): If $L_a(k_1 - 1) < L_b(k_2)$ and $L_a(k_1 - 1) \geq L_b(k_2 - 1)$, then $g(k_1, k_2)$ takes the form shown in Equation (5.9) as $L_a(k_1) \geq L_b(k_2) > L_a(k_1 - 1) \geq L_b(k_2 - 1)$.

$$\begin{aligned} g(k_1, k_2) =& f(L_a(k_1), L_b(k_2 - 1)) \\ & - f(L_a(k_1 - 1), L_b(k_2 - 1)) - f(L_a(k_1), L_b(k_2)) \\ & = \psi(L_a(k_1) - L_b(k_2 - 1)) \\ & - \psi(L_a(k_1 - 1) - L_b(k_2 - 1)) - \psi(L_a(k_1) - L_b(k_2)) \end{aligned} \tag{5.9}$$

Let $L_a(k_1) - L_b(k_2) = \delta_1$, $L_b(k_2) - L_a(k_1 - 1) = \delta_2$ and $L_a(k_1 - 1) - L_b(k_2 - 1) = \delta_3$, where $\delta_1 \geq 0$, $\delta_2 > 0$ and $\delta_3 \geq 0$.

Rewriting Equation (5.9) and substituting these values, we get the following expression expression,

$$\begin{aligned} g(k_1,k_2) &= \psi(L_a(k_1) - L_b(k_2-1)) \\ &\quad - \psi(L_a(k_1-1) - L_b(k_2-1)) - \psi(L_a(k_1) - L_b(k_2)) \\ &= \psi(\delta_1+\delta_2+\delta_3) - \psi(\delta_3) - \psi(\delta_1) \end{aligned}$$

It can be verified that $g(k_1,k_2) > 0$ as $\psi(\cdot)$ is convex.

Subcase 3): If $L_a(k_1-1) \geq L_b(k_2)$, then $L_a(k_1)-L_b(k_2-1) > 0$, $L_a(k_1-1)-L_b(k_2) \geq 0$, $L_a(k_1-1) - L_b(k_2-1) > 0$, and $L_a(k_1) - L_b(k_2) > 0$. Hence,

$$\begin{aligned} g(k_1,k_2) &= \psi(L_a(k_1) - L_b(k_2-1)) \\ &\quad - \psi(L_a(k_1-1) - L_b(k_2-1)) - \psi(L_a(k_1) - L_b(k_2)) \\ &\quad + \psi(L_a(k_1-1) - L_b(k_2)). \end{aligned}$$

In this subcase, let $L_a(k_1) - L_a(k_1-1) = \delta_1$, $L_a(k_1-1) - L_b(k_2) = \delta_2$ and $L_b(k_2) - L_b(k_2-1) = \delta_3$, where $\delta_1 > 0$, $\delta_2 \geq 0$ and $\delta_3 > 0$. Substituting this in the expression for $g(k_1,k_2)$, we get

$$g(k_1,k_2) = \psi(\delta_1+\delta_2+\delta_3) - \psi(\delta_2+\delta_3) - \psi(\delta_1+\delta_2) + \psi(\delta_2).$$

Let us first consider the case, $\delta_2 = 0$, we get the following expression,

$$g(k_1,k_2) = \psi(\delta_1+\delta_3) - \psi(\delta_3) - \psi(\delta_1)$$

It can be verified that $g(k_1,k_2) > 0$ as $\psi(\cdot)$ is convex.

Next, consider the case when $\delta_2 > 0$. It can be observed that $\delta_1+\delta_2+\delta_3 > \delta_1+\delta_2 > \delta_2$. Therefore, $\delta_1+\delta_2$ can be expressed as, $\delta_1+\delta_2 = \lambda_1\delta_2 + (1-\lambda_1)(\delta_1+\delta_2+\delta_3)$.

Solving for λ_1, we get $\lambda_1 = \frac{\delta_3}{\delta_1+\delta_3}$.

Similarly, it can be observed that $\delta_1+\delta_2+\delta_3 > \delta_2+\delta_3 > \delta_2$ and $\delta_2+\delta_3$ can be expressed as, $\delta_2+\delta_3 = \lambda_2\delta_2 + (1-\lambda_2)(\delta_1+\delta_2+\delta_3)$ where $\lambda_2 = \frac{\delta_1}{\delta_1+\delta_3}$.

From the definition of a convex function, and adding the above two expressions, we get the following,

$\psi(\delta_1 + \delta_2) + \psi(\delta_2 + \delta_3) \leq (\lambda_1 + \lambda_2)\psi(\delta_2) + (2 - \lambda_1 - \lambda_2)\psi(\delta_1 + \delta_2 + \delta_3)$.

Substituting the value of λ_1 and λ_2, we get $\psi(\delta_1 + \delta_2) + \psi(\delta_2 + \delta_3) \leq \psi(\delta_2) + \psi(\delta_1 + \delta_2 + \delta_3)$. Therefore it can be verified that $g(k_1, k_2) \geq 0$.

Thus, through these exhaustive cases, it is shown that for any k_1 and k_2, the function $g(k_1, k_2) \geq 0$ or in other words is non-negative.

5.3.2 Proof for Lemma 2

Lemma 2: In any finite s-t cut C, the total weight of the edges between any two adjacent columns a and b (denoted by $C_{a,b}$) equals to the surface smoothness cost of the resulting surface S_i with $S_i(a) = k_1$ and $S_i(b) = k_2$, which is $\psi(L_a(k_1) - L_b(k_2))$, where $\psi(.)$ is a convex function.

Proof: Denote an edge from $n_i(a, k_1)$ to node $n_i(b, k_2)$ as $E_i(a_{k_1}, b_{k_2})$ for the i-th surface. Assume $k_1 \geq k_2$. Proof for the case when $k_2 \geq k_1$ can be done in a similar manner by interchanging the notations for column a and column b. To show: cost of cut $C_{a,b} = \psi(L_a(k_1) - L_b(k_2))$.

We start by observing such a s-t cut $C_{a,b}$ will consist of only the following inter-column edges: $\{E_i(a_m, b_n)\ ,\ \ 0 \leq m \leq k_1,\ \ k_2 + 1 \leq n \leq Z\}$

Summing up the weights of the above edges using Equation 5.6, we obtain the following expression:

$$
\begin{aligned}
C_{a,b} =& g(k_1, Z) + g(k_1, Z-1) + g(k_1, Z-2) \\
&+ \ldots + g(k_1, k_2+1) \\
&+ g(k_1-1, Z) + g(k_1-1, Z-1) + g(k_1-1, Z-2) \\
&+ \ldots + g(k_1-1, k_2+1) \\
&. \\
&. \\
&. \\
&+ g(0, Z) + g(0, Z-1) + g(0, Z-2) \\
&+ \ldots + g(0, k_2+1)
\end{aligned}
\tag{5.10}
$$

Let us first evaluate part of Equation (5.10) for k, where $0 \leq k \leq k_1$ as shown below:

$g(k, Z) + g(k, Z-1) + g(k, Z-2) + \ldots + g(k, k_2+1)$

$= f(L_a(k), L_b(Z-1)) - f(L_a(k-1), L_b(Z-1))$
$-f(L_a(k), L_b(Z)) + f(L_a(k-1), L_b(Z))$
$+f(L_a(k), L_b(Z-2)) - f(L_a(k-1), L_b(Z-2))$
$-f(L_a(k), L_b(Z-1)) + f(L_a(k-1), L_b(Z-1))$
$+f(L_a(k), L_b(Z-3)) - f(L_a(k-1), L_b(Z-3))$
$-f(L_a(k), L_b(Z-2)) + f(L_a(k-1), L_b(Z-2))$
.
.
.
$+f(L_a(k), L_b(k_2)) - f(L_a(k-1), L_b(k_2))$
$-f(L_a(k), L_b(k_2+1)) + f(L_a(k-1), L_b(k_2+1))$

$$= f(L_a(k), L_b(k_2)) - f(L_a(k-1), L_b(k_2))$$
$$-f(L_a(k), L_b(Z)) + f(L_a(k-1), L_b(Z))$$

As described in Section 5.2.2.2,

$$\begin{aligned} & f(L_a(k), L_b(Z)) = 0, \;\; f(L_a(k-1), L_b(Z)) = 0 \\ & (\because Z \notin \mathbf{z}) \\ =& f(L_a(k), L_b(k_2)) - f(L_a(k-1), L_b(k_2)) \end{aligned} \tag{5.11}$$

By simplifying Equation (5.10) using Equation (5.11), it follows that:

$$\begin{aligned} C_{a,b} = & f(L_a(k_1), L_b(k_2)) - f(L_a(k_1-1), L_b(k_2)) \\ & + f(L_a(k_1-1), L_b(k_2)) - f(L_a(k_1-2), L_b(k_2)) \\ & . \\ & . \\ & . \\ & + f(L_a(1), L_b(k_2)) - f(L_a(0), L_b(k_2)) \\ & + f(L_a(0), L_b(k_2)) - f(L_a(-1), L_b(k_2)) \end{aligned} \tag{5.12}$$

$$\begin{aligned} =& f(L_a(k_1), L_b(k_2)) - f(L_a(-1), L_b(k_2)) \\ & f(L_a(-1), L_b(k_2)) = 0, \;\; (\because -1 \notin \mathbf{z}) \\ =& \psi(L_a(k_1) - L_b(k_2)) \end{aligned}$$

Therefore, for this case it is shown that cost of cut $C_{a,b} = \psi(L_a(k_1) - L_b(k_2))$.

In a similar manner when $k_2 \geq k_1$, the s-t cut $C_{b,a}$ will consist of the following inter-column edges: $\{E_i(b_m, a_n)\ ,\ \ 0 \leq m \leq k_2,\ \ k_1 + 1 \leq n \leq Z\}$.

Summing up the weights of the above edges using Equation 5.7, we obtain the expression shown in Eqn 5.13.

$$\begin{aligned} C_{b,a} =& g(k_2, Z) + g(k_2, Z-1) + g(k_2, Z-2) \\ & + \ldots + g(k_2, k_1+1) \\ & g(k_2-1, Z) + g(k_2-1, Z-1) + g(k_2-1, Z-2) \\ & + \ldots + g(k_2-1, k_1+1) \\ & . \\ & . \\ & . \\ & g(0, Z) + g(0, Z-1) + g(0, Z-2) \\ & + \ldots + g(0, k_1+1) \end{aligned} \tag{5.13}$$

Similar to the previous case, let us first evaluate part of Equation (5.13) for k, where $0 \leq k \leq k_2$ as shown below:

$g(k, Z) + g(k, Z-1) + g(k, Z-2) + \ldots + g(k, k_1+1)$

$$\begin{aligned} &= f(L_b(k), L_a(Z-1)) - f(L_b(k-1), L_a(Z-1)) \\ &- f(L_b(k), L_a(Z)) + f(L_b(k-1), L_a(Z)) \\ &+ f(L_b(k), L_a(Z-2)) - f(L_b(k-1), L_a(Z-2)) \\ &- f(L_b(k), L_a(Z-1)) + f(L_b(k-1), L_a(Z-1)) \\ &+ f(L_b(k), L_a(Z-3)) - f(L_b(k-1), L_a(Z-3)) \\ &- f(L_b(k), L_a(Z-2)) + f(L_b(k-1), L_a(Z-2)) \end{aligned}$$

.

.

.

$$+f(L_b(k), L_a(k_1)) - f(L_b(k-1), L_a(k_1))$$
$$-f(L_b(k), L_a(k_1+1)) + f(L_b(k-1), L_a(k_1+1))$$

$$= f(L_b(k), L_a(k_1)) - f(L_b(k-1), L_a(k_1))$$
$$-f(L_b(k), L_a(Z)) + f(L_b(k-1), L_a(Z))$$

As described in Section 5.2.2.2,

$$\begin{aligned} & f(L_b(k), L_a(Z)) = 0, \ f(L_b(k-1), L_a(Z)) = 0 (\because Z \notin \mathbf{z}) \\ =& f(L_b(k), L_a(k_1)) - f(L_b(k-1), L_a(k_1)) \end{aligned} \tag{5.14}$$

By simplifying Equation (5.13) using Equation (5.14), it follows that:

$$\begin{aligned} C_{b,a} = & f(L_b(k_2), L_a(k_1)) - f(L_b(k_2-1), L_a(k_1)) \\ & + f(L_b(k_2-1), L_a(k_1)) - f(L_b(k_2-2), L_a(k_1)) \\ & . \\ & . \\ & . \\ & + f(L_b(1), L_a(k_1)) - f(L_b(0), L_a(k_1)) \\ & + f(L_b(0), L_a(k_1)) - f(L_b(-1), L_a(k_1)) \end{aligned} \tag{5.15}$$

$$\begin{aligned} =& f(L_b(k_2), L_a(k_1)) - f(L_b(-1), L_a(k_1)) \\ & f(L_b(-1), L_a(k_1)) = 0, \ (\because -1 \notin \mathbf{z}) \\ =& \psi(L_b(k_2) - L_a(k_1)) \end{aligned}$$

Therefore, for this case it is shown that cost of cut $C_{b,a} = \psi(L_b(k_2) - L_a(k_1))$. Thus the surface smoothness term $V_{ab}(.)$ is correctly encoded in graph G.

5.4 Validation for Subvoxel Accuracy and Super Resolution Segmentation Accuracy (Aim 2.3)

In this section, we present the application of our method on Spectral Domain Optical Coherence Tomography (SD-OCT) volumes to segment multiple surfaces simultaneously with subvoxel and super resolution segmentation accuracy. The proposed method was also applied to Intravascular Ultrasound (IVUS) images for Lumen and Media segmentation with subvoxel accuracy. The experiment conducted on the SD-OCT volumes have a twofold objective. The first experiment is designed to demonstrate multiple surface segmentation with convex priors while achieving subvoxel accuracy. The second experiment is used to show that the proposed method has the potential to perform super resolution segmentation with sufficient accuracy, i.e achieve adequate segmentation accuracy by operating in the downsampled version of data as compared to segmenting the data in the original resolution with convex surface smoothness constraints.

5.4.1 Experiment Setting for SD-OCT Volumes of Normal Eye

To demonstrate utility of the method in simultaneous segmentation of multiple surfaces, three surfaces were simultaneously segmented in this study. The surfaces are S_1- Internal Limiting Membrane (ILM), S_2- Inner Aspect of Retinal Pigment Epithelium Drusen Complex (IRPEDC) and S_3- Outer Aspect of Bruch Membrane (OBM) as shown in Fig. 5.5.

As discussed earlier Abrámoff *et al.* [65] exploited the additional information contained in the partial volume effect by generalizing the graph by applying a deformation field (allowing non-equidistant spacing between nodes) to achieve subvoxel accuracy. The method used hard smoothness constraints to model the surface smoothness term. Our method is directly applicable to such cases and allows for usage of a convex smoothness penalty for surface smoothness. In other words, our method can be used

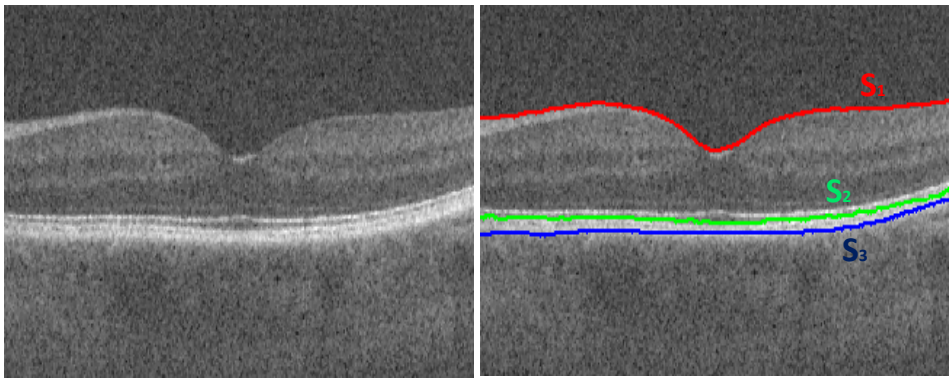


Figure 5.5: (a)A single B-scan from a SD-OCT volume of a normal eye, (b) Three identified target surfaces S_1, S_2 and S_3.

to achieve subvoxel accuracy or super resolution accuracy with a convex smoothness penalty term. In this experiment we compare the segmentation accuracy of the proposed method in the irregularly sampled space to the optimal surface segmentation method using convex smoothness constraints in the regularly sampled space (OSCS) [9].

5.4.1.1 Data

30 SD-OCT volumes of normal eyes and their respective expert manual tracings were obtained from the publicly available repository of datasets Ref. [68]. The 3-D volumes ($1000 \times 100 \times 512$ voxels with voxel size $6.54 \times 67 \times 3.23$ μm^3) for our study were randomly selected from the repository. 5 SD-OCT volumes were used for training of the algorithm and cost function image parameters while the remaining 25 SD-OCT volumes were used for testing. The obtained expert manual tracings were marked with equidistant voxel centers. Thus, for fair comparison the original image volumes were down-sampled to create"input volume data". The target surfaces were then mapped from high resolution to its location in the down-sampled resolution to generate "subvoxel accurate expert manual tracings".

5.4.1.2 Workflow

The original SD-OCT volumes first undergo pre-processing which involves the application of a $10 \times 10 \times 10$ median filter followed by a $10 \times 10 \times 10$ Gaussian filter with a standard deviation of 7 to denoise the original data. The resulting volumes are then down-sampled by a factor of 4 in the x direction, followed by a down-sampling by a factor of η in the z direction, resulting in input volume data of size $250 \times 100 \times \frac{512}{\eta}$ voxels. Further cost function image volumes $D_{i,\eta}$, $(i = 1, 2, 3)$ are generated for each target surface at scale η using the input volume data.

Experiment for Subvoxel Accuracy - The down-sampling factor $\eta = 4$ is chosen for this experiment. The cost function image volumes are simultaneously segmented using the OSCS method to obtain the segmentation for comparison with the proposed method. Thereafter, a shift of evenly distributed voxels to deformed space is computed using gradient vector flow (GVF) [80] as described in Section 5.4.1.4 on the input volume data. The deformation field is then applied to the cost function image volumes to obtain $D'_{i,\eta=4}$, $(i = 1, 2, 3)$ and the shifted position of each voxel center is recorded. More details regarding the application of the deformation field on the cost function image volume can be found in Ref. [65]. Finally, the deformed cost function image volumes $D'_{i,\eta=4}$, $(i = 1, 2, 3)$ are segmented using the proposed method with non-equidistant spacing of the voxel centers based on the shifted voxel centers.

The generated input volume data is used to evaluate segmentation accuracy of the two methods with respect to the subvoxel accurate expert manual tracings. For fair and robust analysis, the deformation obtained from the GVF was applied to the automated segmentations obtained from the OSCS method, resulting in deformed OSCS (DOSCS) segmentations. The quantitative analysis was conducted by comparing the segmentations obtained from the OSCS method, proposed method and DOSCS seg-

mentations with the subvoxel accurate expert manual tracings. The corresponding workflow of the experiment is shown in Fig. 5.6.

Experiment for Super Resolution Accuracy - In this experiment we down-sample the data to four different scales at $\eta = 2, 4, 6, 8$. The cost function image volume at the original scale ($\eta = 1$) is simultaneously segmented using the OSCS method to obtain the segmentation for comparison with the proposed method. The proposed method is applied to input volume data at the down-sampled scales. The shift of evenly distributed voxels to deformed space at each down-sampling scale is computed using gradient vector flow (GVF) [80] on the input volume data at each scale η. The deformation field is then applied to the cost function image volumes at their respective scales and the shifted position of each voxel center is recorded. Finally, the deformed cost function image volumes $D'_{i,\eta}$, ($i = 1, 2, 3$ and $\eta = 2, 4, 6, 8$) at each scale n are segmented using the proposed method with non-equidistant spacing of the voxel centers based on the shifted voxel centers due to the applied deformation.

The quantitative analysis was conducted by comparing the segmentations obtained from the OSCS method at the original scale and the proposed method at different down-sampling scales with the expert manual tracings. The corresponding workflow of the experiment is shown in Fig. 5.7.

5.4.1.3 Cost Function Design

The cost function image volumes encode the data term shown in Equation (5.2). To detect surfaces S_1 and S_3, a 3-D Sobel filter of size $5 \times 5 \times 5$ voxels is applied to generate cost volumes D_1 and D_3 wherein the vertical edges for dark to bright transitions and bright to dark transitions are emphasized. To detect surface S_2, a machine learning based approach is adopted to generate the cost volume for the same. For each voxel $I(x, y, z)$ in an image slice of the input volume data, a 11×11 window centered at voxel $I(x, y, z)$ is used to generate a feature vector comprising of the

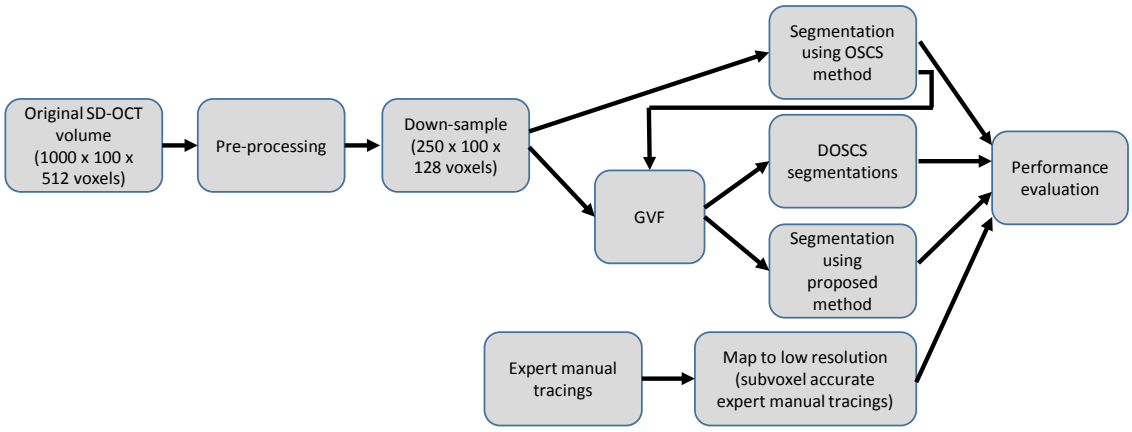


Figure 5.6: Experiment design for segmentation of SD-OCT volumes of normal eye with subvoxel accuracy.

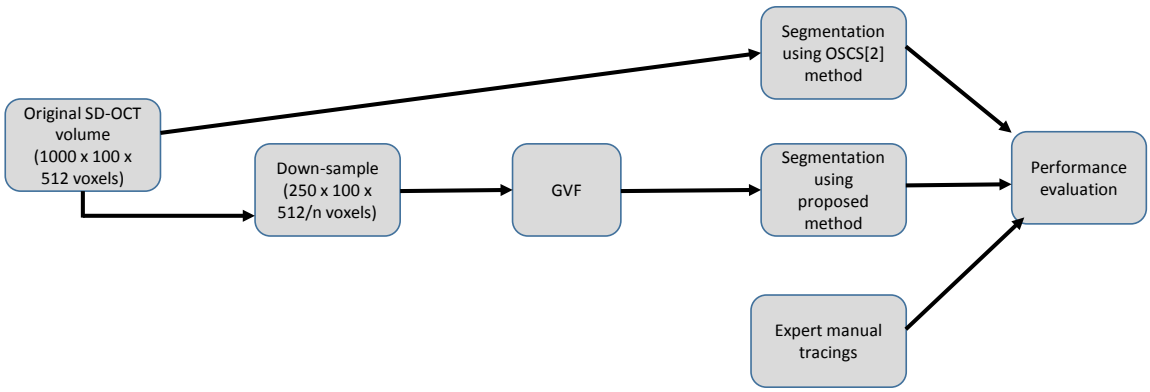


Figure 5.7: Experiment design for segmentation of SD-OCT volumes of normal eye for super resolution accuracy.

intensity values of each voxel in the given window, thus resulting in 121 features. A random forest classifier [81] with 10 trees is then trained on voxels of the training set input image volumes to learn the probability maps which indicate the likelihood of voxel belonging to the surface of interest with respect to the expert manual tracings. The trained classifier is then applied to each voxel of the test set resulting in a probability map $D_2'(x, y, z)$. Finally, cost volume D_2 to detect S_2 is generated by assigning $D_2(x, y, z) = (1 - D_2'(x, y, z)) \times 255$ as voxel intensity.

5.4.1.4 Gradient Vector Field

A gradient vector field (GVF) [80] is a feature preserving diffusion of the gradient in a given image volume. In this study, GVF is used as a deformation field $F(x, y, z)$ obtained directly from the input volume data acting on the center of each voxel (x, y, z) to shift the evenly distributed voxels to the deformed space. The voxel centers are thus displaced towards the regions where salient transitions of image properties are more likely to occur. The shift of the centers of the voxels is given by Equation (5.16).

$$(x', y', z') = (x, y, z) + \lambda F(x, y, z) \tag{5.16}$$

where λ is a normalization factor. The displacement of each voxel center is confined to the same voxel. Therefore, $F(x, y, z)$ is normalized such that the maximum deformation is equal to half of the voxel size δ. The normalization factor takes the following form as show in Equation (5.17).

$$\lambda = \frac{\delta}{2 \times max_{(x,y,z)\in(X,Y,Z)}||F(x, y, z)||} \tag{5.17}$$

5.4.1.5 Parameter Setting

The same parameters were used for segmenting the three surface by the OSCS method and the proposed method at each scale n. A linear (convex) function, $\psi(k_1 -$

$k_2) = |k_1 - k_2|$ was used to model the surface smoothness term $V_{ab}(.)$. The surface separation term $H_a(.)$ is modelled as a hard constraint for enforcing the minimum separation between a pair of surfaces. The minimum separation parameters used are $d_{1,2} = 15$ and $d_{2,3} = 1$ for $\eta = 4$ in the experiment for subvoxel accuracy. The minimum separation parameters used in the experiment for super resolution accuracy at different scales are: $d_{1,2} = 60$ and $d_{2,3} = 4$ for $\eta = 1$, $d_{1,2} = 30$ and $d_{2,3} = 2$ for $\eta = 2$, $d_{1,2} = 15$ and $d_{2,3} = 1$ for $\eta = 4$, $d_{1,2} = 10$ and $d_{2,3} = 0.8$ for $\eta = 6$, $d_{1,2} = 7$ and $d_{2,3} = 0.5$ for $\eta = 8$.

5.4.2 Segmentation Results for SD-OCT Volumes of Normal Eye

The segmentation accuracy was estimated using unsigned mean surface positioning error (UMSP) and unsigned average symmetric surface distance error (UASSD). The UMSP error for a surface was computed by averaging the vertical difference between the subvoxel accurate manual tracings and the automated segmentations for all the columns in the input volume data. The UASSD error for a surface was calculated by averaging the closest distance between all surface points of the automated segmentation and those of the expert manual tracings in the physical space. Statistical significance of the observed differences was determined using paired t-test for which p values of 0.05 were considered significant.

Results for Subvoxel Accuracy - The USMP errors are summarized in Table 5.2 and the UASSD errors are summarized in Table 5.3. Fig. 5.8 and 5.9 show the performance comparison of the proposed method, the OSCS method [9] and the DOSCS segmentations. Our method produced significantly lower UMSP and UASSD errors for S_1 ($p<0.01$), S_2 ($p<0.01$) and S_3 ($p<0.001$) compared to the OSCS method and the DOSCS segmentations.

Qualitatively the algorithm produced very good and consistent segmentations. The qualitative illustrations are shown in Fig. 5.10 and 5.11. In both the illustrations,

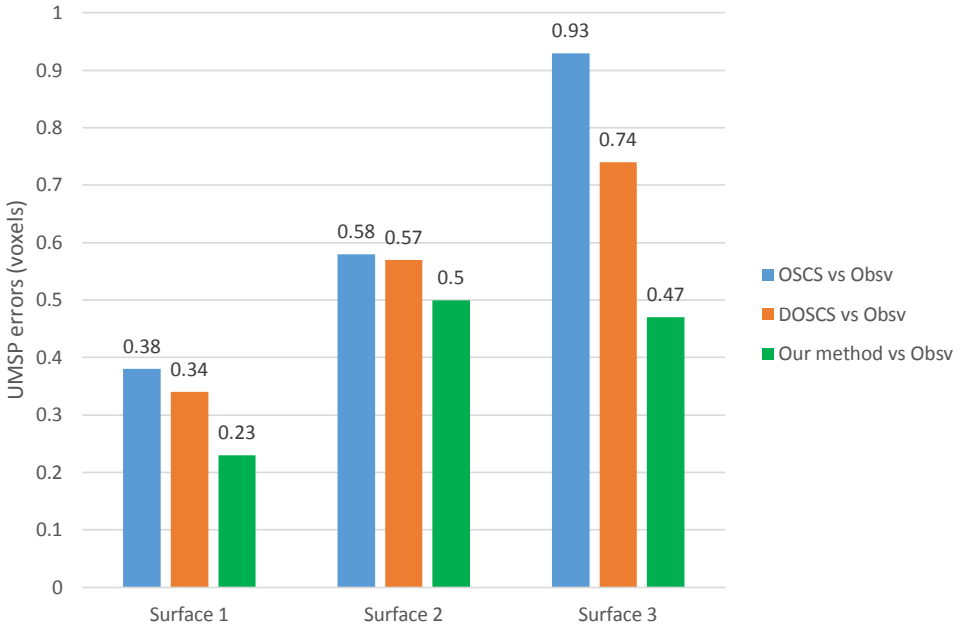


Figure 5.8: Unsigned mean surface positioning errors observed in 25 volumetric OCT images for subvoxel accuracy.

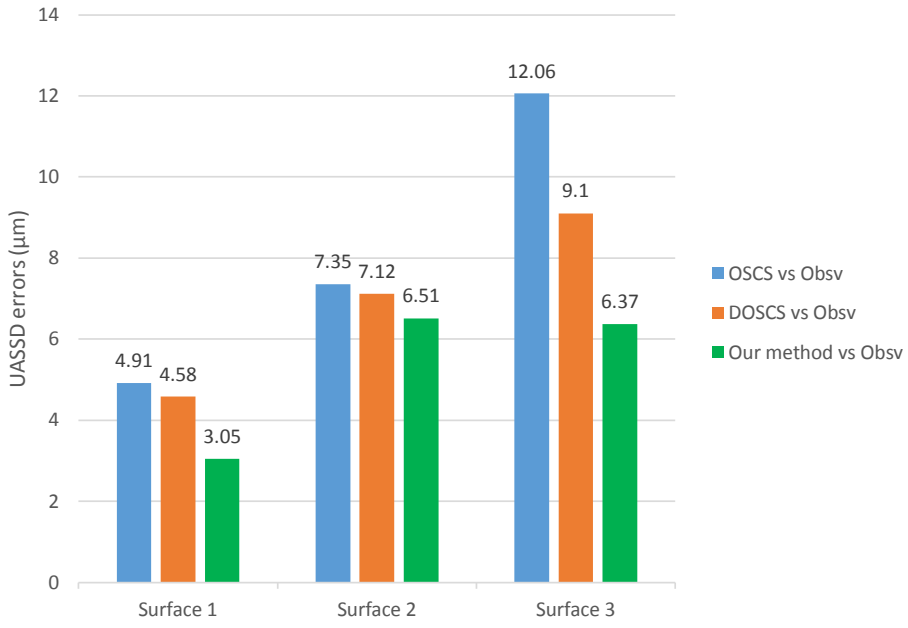


Figure 5.9: Unsigned average symmetric surface distance errors observed in 25 volumetric OCT images for subvoxel accuracy.

Table 5.2: Unsigned mean surface positioning error (UMSP) (mean ± standard deviation) in voxels for subvoxel accuracy. Obsv - Subvoxel accurate expert manual tracings.

Surface	OSCS vs. Obsv	DOSCS vs. Obsv	Our method vs. Obsv
S_1	0.38 ± 0.05	0.34 ± 0.05	0.23 ± 0.04
S_2	0.58 ± 0.37	0.57 ± 0.36	0.50 ± 0.32
S_3	0.93 ± 0.47	0.74 ± 0.45	0.47 ± 0.43
Overall	0.63 ± 0.30	0.55 ± 0.29	0.40 ± 0.26

Table 5.3: Unsigned average symmetric surface distance error (UASSD) (mean ± standard deviation) in μm for subvoxel accuracy. Obsv - Expert manual tracings.

Surface	OSCS vs. Obsv	DOSCS vs. Obsv	Our method vs. Obsv
S_1	4.91 ± 0.63	4.58 ± 0.73	3.05 ± 0.55
S_2	7.35 ± 3.91	7.12 ± 3.76	6.51 ± 3.61
S_3	12.06 ± 5.03	9.10 ± 4.97	6.37 ± 4.77
Overall	8.11 ± 3.19	6.93 ± 3.15	5.31 ± 2.98

the first row shows a single B-scan of the input image volume with the subvoxel accurate expert manual tracings and the automated segmentations. The second, third and fourth row show the magnification of the black boxes in the first row corresponding to the surfaces ILM, IRPEDC and OBM respectively. The first column shows the subvoxel accurate expert manual tracings. The second column shows the subvoxel accurate expert manual tracings vs OSCS segmentation. The third column shows the subvoxel accurate expert manual tracings vs DOSCS segmentation. The fourth column shows the subvoxel accurate expert manual tracings vs segmentation from our method.

It can be seen from Fig. 5.10 and 5.11 that the proposed method yields more accurate segmentations compared to the OSCS method and the DOSCS segmentations

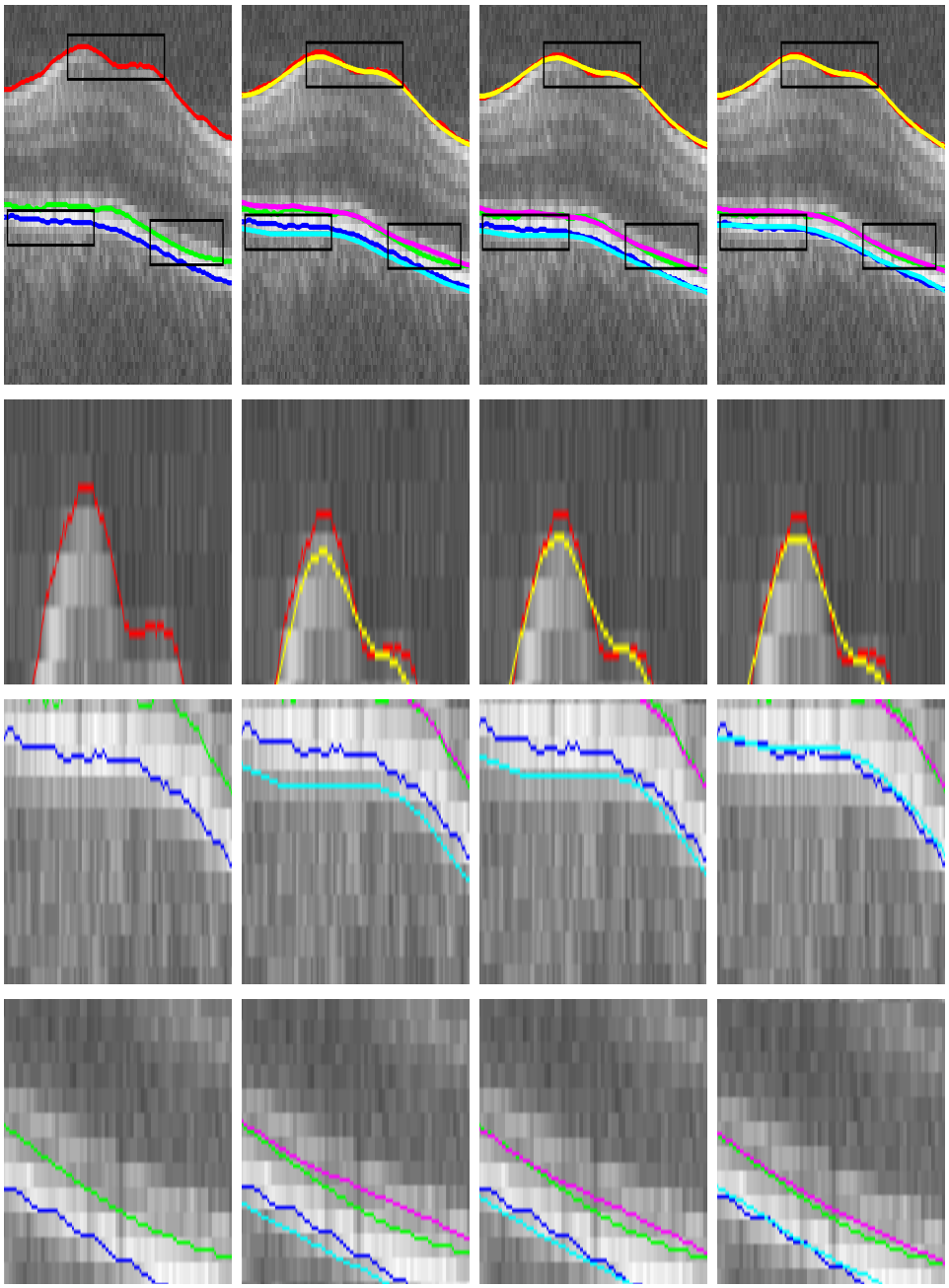


Figure 5.10: Illustration of results on a single B-scan from input image volume. Red - ILM expert tracing, Green - IRPEDC expert tracing, Dark blue - OBM expert tracing, Yellow - ILM automated segmentation, Magenta - IRPEDC automated segmentation, Light blue - OBM automated segmentation.

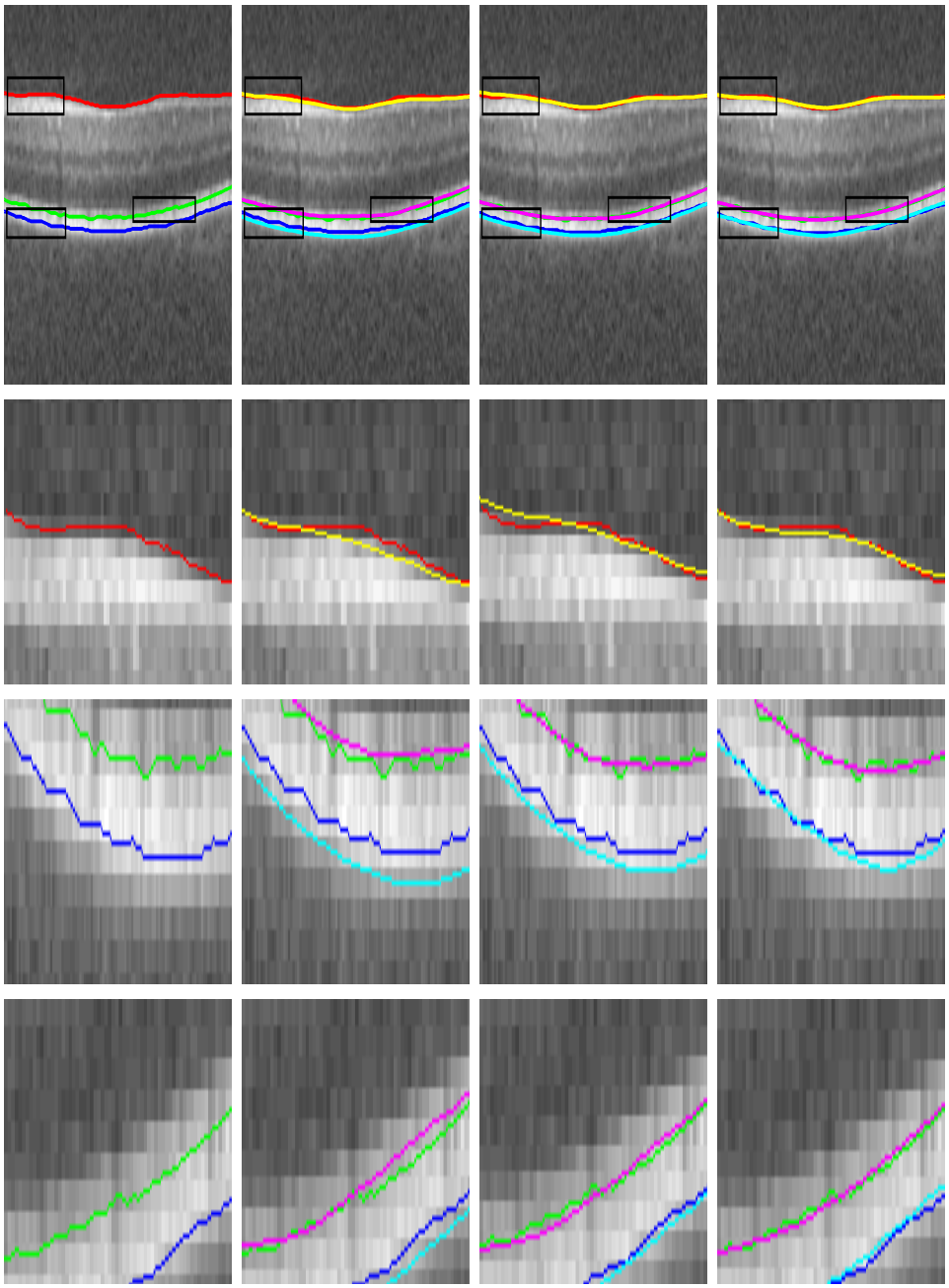


Figure 5.11: Illustration of results on a single B-scan from input image volume. Red - ILM expert tracing, Green - IRPEDC expert tracing, Dark blue - OBM expert tracing, Yellow - ILM automated segmentation, Magenta - IRPEDC automated segmentation, Light blue - OBM automated segmentation.

for Surface 2 and Surface 3, while from Fig. 5.11, it can be seen that the proposed method also yields more accurate segmentation for Surface 1. Furthermore, the second, third and fourth rows clearly demonstrates that the proposed method yields a much higher subvoxel accuracy for the segmentations. It can be seen from the last row in Fig. 5.10 that even after applying the deformation to the OSCS segmentations, the DOSCS segmentation do not achieve the globally optimum solution obtained by using the proposed method with subvoxel accuracy. This is because all nodes encode potential surface locations more precisely when the globally optimum solution is computed in the irregularly sampled graph space by utilizing the information from the partial volume effect.

Results for Super Resolution Accuracy - The UASSD errors are summarized in Table 5.4. The results obtained from the proposed method at different down-sampling scales ($\eta = 2, 4, 6, 8$) was compared to the segmentation obtained from the OSCS method at the original scale ($\eta = 1$) to determine the relative accuracy of the proposed method in the down-sampled resolutions. The performance of the proposed method compared with the OSCS method for super resolution segmentation accuracy is shown in Fig. 5.12. There was no significant difference between the UASSD errors produced by the proposed method at $\eta = 2$ and the OSCS method at original scale ($\eta = 1$) for S_1 (p>0.05), S_2 (p>0.05) and S_3 (p>0.05). There was no significant difference between the UASSD errors produced by the proposed method at $\eta = 4$ and the OSCS method at original scale ($\eta = 1$) for S_1 (p>0.05) and S_3 (p>0.05) while there was a significant difference for S_2 (p<0.05). There was a significant difference observed in between the UASSD errors produced by the proposed method at $\eta = 6, 8$ and the OSCS method at o riginal scale ($\eta = 1$) for S_1 (p<0.05), S_2 (p<0.05) and S_3 (p<0.05).

In other words, the proposed method achieves adequate accuracy for down-sampled resolutions at $\eta = 2, 4$ for the segmented surfaces when compared to the segmentation

accuracy by the OSCS method in the original scale resolution. For lower resolutions at scales $\eta = 6, 8$, the proposed method is not able to provide adequate super resolution accuracy. Therefore, the method has potential for super resolution accuracy at down-sampled scales of $\eta = 2, 4$.

Table 5.4: Unsigned average symmetric surface distance error (UASSD) (mean $\pm$ standard deviation) in μm for super resolution accuracy. Obsv - Expert manual tracings, η - Down-sampling scale in z direction.

Surface	S_1	S_2	S_3
OSCS vs. Obsv at $\eta = 1$	2.53 ± 0.36	5.88 ± 3.28	5.81 ± 4.25
Our method vs. Obsv at $\eta = 2$	2.69 ± 0.41	6.18 ± 3.39	6.07 ± 4.41
Our method vs. Obsv at $\eta = 4$	3.05 ± 0.55	6.51 ± 3.61	6.37 ± 4.77
Our method vs. Obsv at $\eta = 6$	3.79 ± 0.94	7.58 ± 4.22	7.32 ± 5.27
Our method vs. Obsv at $\eta = 8$	5.45 ± 1.71	9.15 ± 5.23	9.21 ± 6.46

5.4.3 Experiment Setting for IVUS Images

To study the applicability of the proposed method in a broader range of image segmentation tasks, segmentation of lumen and media with subvoxel accuracy was performed in Intravascular Ultrasound (IVUS) images as shown in Fig. 5.13.

Atherosclerosis, a disease of the vessel wall, is the major cause of cardiovascular diseases such as heart attack or stroke [82]. Early atherosclerosis results in remodelling, thus retaining the lumen despite plaque accumulation [83]. Atherosclerosis plaque is located between lumen and media that can be identified in IVUS images. Automated IVUS segmentation of lumen and media is of substantial clinical interest and contributes to clinical diagnosis and assessment of plaque [84].

In this experiment we compare the segmentation accuracy of the lumen and media using the proposed method with the complete set of methods used in the standardized evaluation of IVUS image segmentation [84]. The compared methods are namely, P1

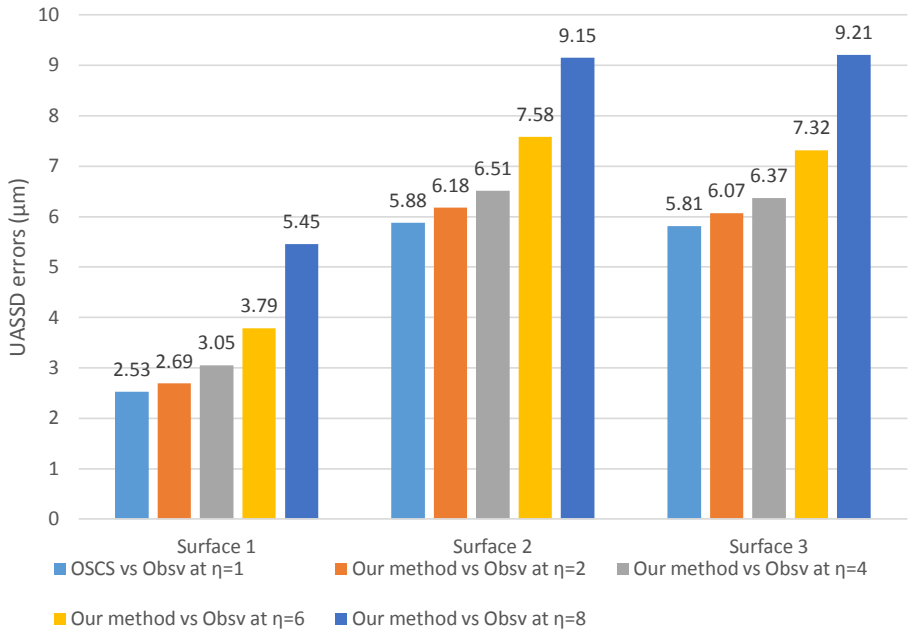
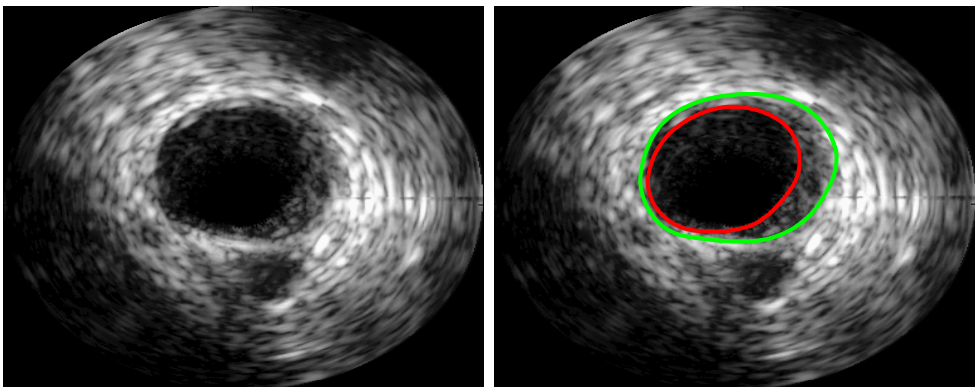


Figure 5.12: Unsigned average symmetric surface distance errors observed in 25 volumetric OCT images for super resolution segmentation accuracy. η is the downsampling scale.



(a) (b)

Figure 5.13: (a) A single frame of an IVUS multiframe dataset (b) Expert manual tracings of the Lumen (red) and Media (green).

- Shape driven segmentation based on linear projections [85], P2 - geodesic active contour based segmentation [86], P3 - Expectation maximization based method [87] [88], P4 - graph search based method [89], P5 - Binary classification of distinguishing between lumen and non-lumen regions based on multi-scale Stacked Sequential learning scheme [90], P6 - Detection of Media border by holistic interpretation of the IVUS image (HoliMAb) [91], P7 - Lumen segmentation based on a Bayesian approach [92], P8 - Sequential detection [93]. Overview of the proposed method and each method's feature [84], including whether the algorithm was applied to lumen and/or media, whether the segmentation was done in 2-D or 3-D and whether the method was semi-automated or fully automated is shown in Table 5.5.

Table 5.5: Overview of the proposed and compared method features

Method	Category	Automation	2-D/3-D
P1	Lumen and Media	Semi	2-D
P2	Lumen	Semi	2-D
P3	Lumen and Media	Semi	2-D
P4	Lumen and Media	Fully	3-D
P5	Lumen	Fully	3-D
P6	Media	Fully	2-D
P7	Lumen	Semi	2-D
P8	Lumen and Media	Fully	2-D
Our method	Lumen and Media	Fully	3-D

5.4.3.1 Data

The data used for this experiment was obtained from the standardized evaluation of IVUS image segmentation [84] database. In this experiment Dataset B as denoted in Ref. [84] was used. The data comprises of a set of 435 images with a size of 384 $\times$ 384 pixels extracted from in vivo pullbacks of human coronary arteries from 10

patients. The r espective expert manual tracings (subvoxel accurate) of lumen and media for the images were also obtained from the reference database. The dataset contains 10 multi-frame datasets, in which 3D context from a full pullback is provided. Each dataset comprises of between 20 and 50 gated frames extracted from the full pullback at the end-diastolic cardiac phase. Further, the obtained data comprised of two groups - training and testing set. Approximately one fourth of the images in the dataset were grouped in the training set and the remaining were grouped as the testing set, to assure fair evaluation of the algorithms with respect to the expert manual tracings. The experiment with the proposed method was conducted in conformance with the directives provided for the IVUS challenge [84].

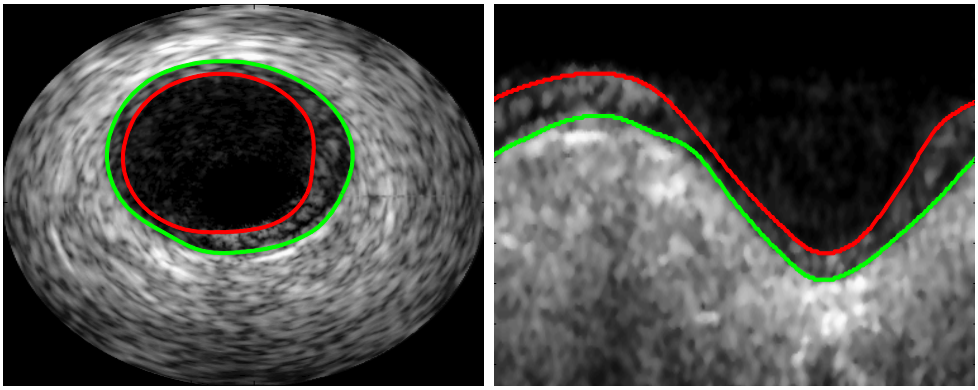
5.4.3.2 Workflow

Each slice of the volumes in the dataset is first converted into a polar coordinate image as shown in Fig 5.14. The generated "polar image volumes" undergo the application of a $7 \times 7 \times 7$ Gaussian filter with a standard deviation of 4 for denoising. Next, cost function image volumes D_{lumen} and D_{media} are generated for the lumen and media respectively. Further the GVF as discussed in Section 5.4.1.4 is computed on the polar image volumes. The deformation field is then applied to cost function image volumes and the shifted position of the voxel centers are recorded. The deformed cost function image volumes D'_{lumen} and D'_{media} are then segmented using the proposed method. Finally the resulting segmentations are mapped back to the original co-ordinate system.

5.4.3.3 Cost Function Design

To detect the lumen and media, a machine learning approach is adopted. For each pixel of the polar image in the training set, a total of 148 features were generated. The following operators are applied in order to generate the features:

- 16 features are generated by applying a set of 16 Gabor filters to the image



(a) (b)

Figure 5.14: (a) A single frame of an IVUS multiframe dataset (b) Polar transformation of (a). Red - Lumen, Green - Media.

according to the following kernel shown in Equation (5.18).

$$G(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-0.5\times((\frac{x}{\sigma_x})^2+(\frac{y}{\sigma_y})^2)+i2\pi(Ux+Vy)} \tag{5.18}$$

The parameters U and V (scaling and orientation) used are U = (0.0442, 0.0884, 0.1768, 0.3536), V = $(0, \pi/4, \pi/2, 3\pi/4)$, $\sigma_x = 0.5622$U and $\sigma_y = 0.4524$U .

- 2 features are generated by applying a 3 × 3 Sobel kernel to the image in the x and y directions.

- 6 features are generated by computing the mean value (m), standard deviation (s) and the ratio $\frac{m}{s}$ of pixel intensities in a sliding window of size 1 × 10 pixels in the x and y directions.

- 2 features defined as shadow (Sh) and relative shadow (Sr) related to the cumulative gray level of the image are generated as shown in the following Equations (5.19),(5.20).

$$Sh(x,y) = \frac{1}{N_r N_c} \sum_{y_s=y}^{N_r} BI(x, y_s) \tag{5.19}$$

$$Sr(x,y) = \frac{1}{N_r N_c} \sum_{y_s=y}^{N_r} y_s BI(x, y_s) \tag{5.20}$$

where $BI(x,y)$ is a binary image obtained by thresholding the image with a thresholding value = 14 and (N_r, N_c) are the image dimensions.

- 1 feature is generated by computing the local binary pattern [94].
- 121 features are generated by using a 11 $\times$ 11 window in a similar manner as discussed in Section 5.4.1.3.

Using the expert manual tracings for the training set two separate random forest classifiers [81] for lumen and media with 10 trees are trained on all the pixels of the images in the training set to learn the probability maps which indicate the likelihood of a pixel belonging to lumen or media respectively. The trained classifiers are then applied to each pixel of the testing set to obtain the two cost function images D_{lumen}, D_{media} for lumen and media in a similar manner as discussed in Section 5.4.1.3.

5.4.3.4 Parameter Setting

A linear (convex) function, $\psi(k_1 - k_2) = |k_1 - k_2|$ was used to model the surface smoothness term $V_{ab}(.)$. The surface separation term $H_a(.)$ is modelled as a hard constraint for enforcing the minimum separation between the lumen and media with $d_{lumen,media} = 2$.

5.4.4 Segmentation Results of Lumen and Media in IUVS Images for Subvoxel Accuracy

The quantitative analysis was carried out by comparing the segmentations obtained by our method with the expert manual tracings (subvoxel accurate). Three evaluation measures were used to quantify the accuracy of the segmentations. We

compare the evaluation measures obtained using our method with the measures of methods (P1-P8) reported in Ref. [84]. The measures used are:

Jaccard Measure (JM) - Quantifies how much the segmented area overlaps with the manual delineated area as shown in Equation (5.21):

$$JM(R_{auto}, R_{man}) = \frac{|R_{auto} \cap R_{man}|}{|R_{auto} \cup R_{man}|} \tag{5.21}$$

where R_{auto} and R_{man} are two vessel regions defined by the manual annotated contour C_{man} and of the automated segmented outline C_{auto} respectively.

Percentage of Area Difference (PAD) - Computes the segmentation area difference as shown in Equation (5.22) :

$$PAD = \frac{|A_{auto} - A_{man}|}{A_{man}} \tag{5.22}$$

where A_{auto} and A_{man} are the vessel areas for the automatic and manual contours respectively.

Hausdroff Distance (HD) - Computes locally the distance between the manual and automated contours as shown in Equation (5.23).

$$HD(C_{auto}, C_{man}) = max_{p \in C_{auto}}\{max_{q \in C_{man}}[d(p, q)]\} \tag{5.23}$$

where p and q are points of the curves C_{auto} and C_{man}, respectively, and $d(p, q)$ is the Euclidean distance.

The quantitative results are summarized in Table 5.6. The comparative performance of the proposed method is shown in Figs. 5.15, 5.16 and 5.17. The results demonstrate that our method performs better than methods P1, P2, P4, P5, P6, P8 and is comparable to methods P3 and P7 with respect to segmentation error

measures for lumen and media. Our method segments both the lumen and media simultaneously while method P7 segments the lumen only. Furthermore, our method is fully automated while methods P3 and P7 are semi-automated. Finally, methods P3 and P7 perform slice by slice segmentation in 2-D while our method performs the segmentation in 3-D and not slice by slice.

Qualitative results are shown in Fig 5.18. The illustration demonstrates that our method produced very good segmentation of the lumen and media. It can also be seen from the illustration that the segmentations from our method are consistent for varied topologies of the lumen and media. Constructing the graph with the shifted voxel centers provides a more accurate encoding of the lumen and media surface positions due to the application of the GVF by adaptively changing the regional node density so that it is higher in regions where the target surface is expected to pass through. Employing a subvoxel accuracy approach allows the segmentation to obtain a greater precision with respect to the subvoxel accurate expert tracings.

Table 5.6: Evaluation measures of each method with respect to expert manual tracings. Error measures expressed as mean and (standard deviation). An empty table cell indicates that the method was not applied to Lumen or Media.

Participant	Lumen			Media		
	JM	PAD	HD	JM	PAD	HD
P1	0.81 (0.12)	0.14 (0.13)	0.47 (0.39)	0.76 (0.13)	0.21 (0.16)	0.64 (0.48)
P2	0.83 (0.08)	0.14 (0.12)	0.51 (0.25)			
P3	**0.88 (0.05)**	**0.06 (0.05)**	**0.34 (0.14)**	**0.91 (0.04)**	**0.05 (0.04)**	**0.31 (0.12)**
P4	0.77 (0.09)	0.15 (0.12)	0.47 (0.22)	0.74 (0.17)	0.23 (0.19)	0.76 (0.48)
P5	0.79 (0.08)	0.16 (0.09)	0.46 (0.30)			
P6				0.84 (0.10)	0.12 (0.12)	0.57 (0.39)
P7	0.84 (0.08)	0.11 (0.12)	0.38 (0.26)			
P8	0.81 (0.09)	0.11 (0.11)	0.42 (0.22)	0.79 (0.11)	0.19 (0.19)	0.60 (0.28)
Our Method	**0.86 (0.04)**	**0.09 (0.03)**	**0.37 (0.14)**	**0.90 (0.03)**	**0.07 (0.03)**	**0.43 (0.12)**

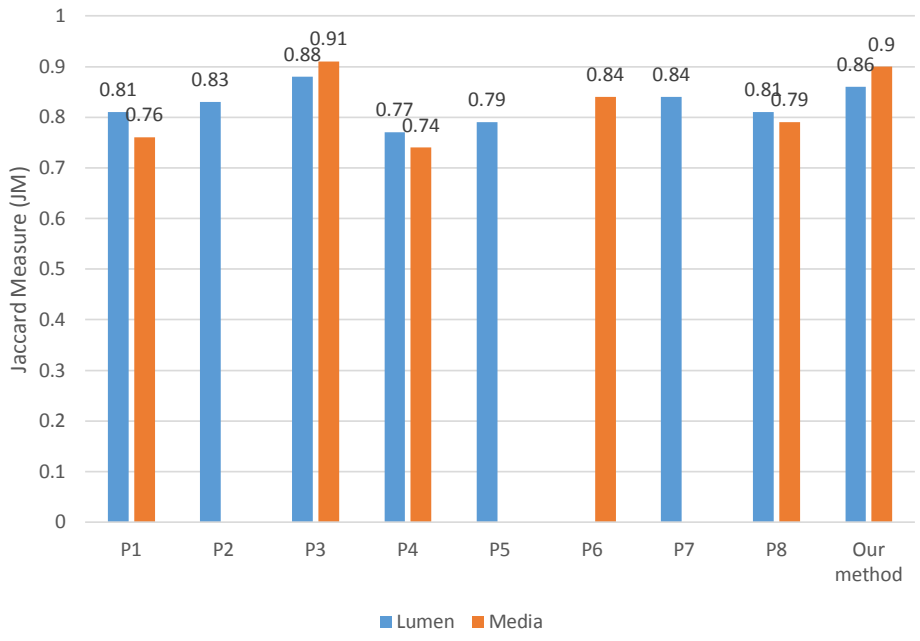


Figure 5.15: Jaccard measure (JM) metric for IVUS data. Higher JM indicates a larger overlap of the automated segmentation and manual segmentations. It can be seen that the proposed method has the highest JM among all fully automated methods and is comparable with the semi-automated method P3.

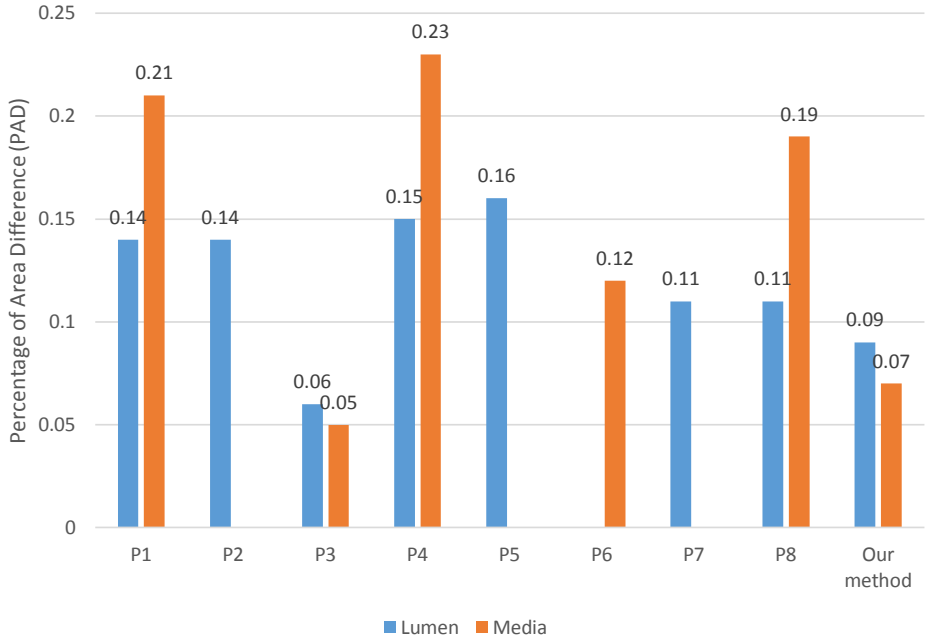


Figure 5.16: Percentage of area difference (PAD) metric for IVUS data. Lower PAD indicates a smaller segmentation area difference between the automated and manual segmentations.

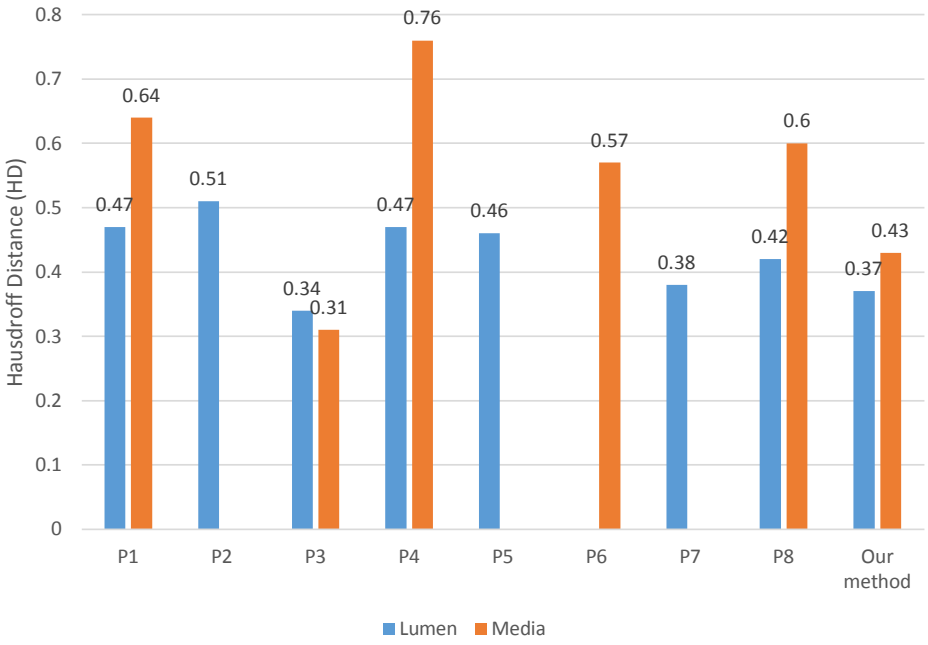


Figure 5.17: Hausdroff distance (HD) metric for IVUS data. Lower HD indicates a closer alignment of the automated and manual segmentations.

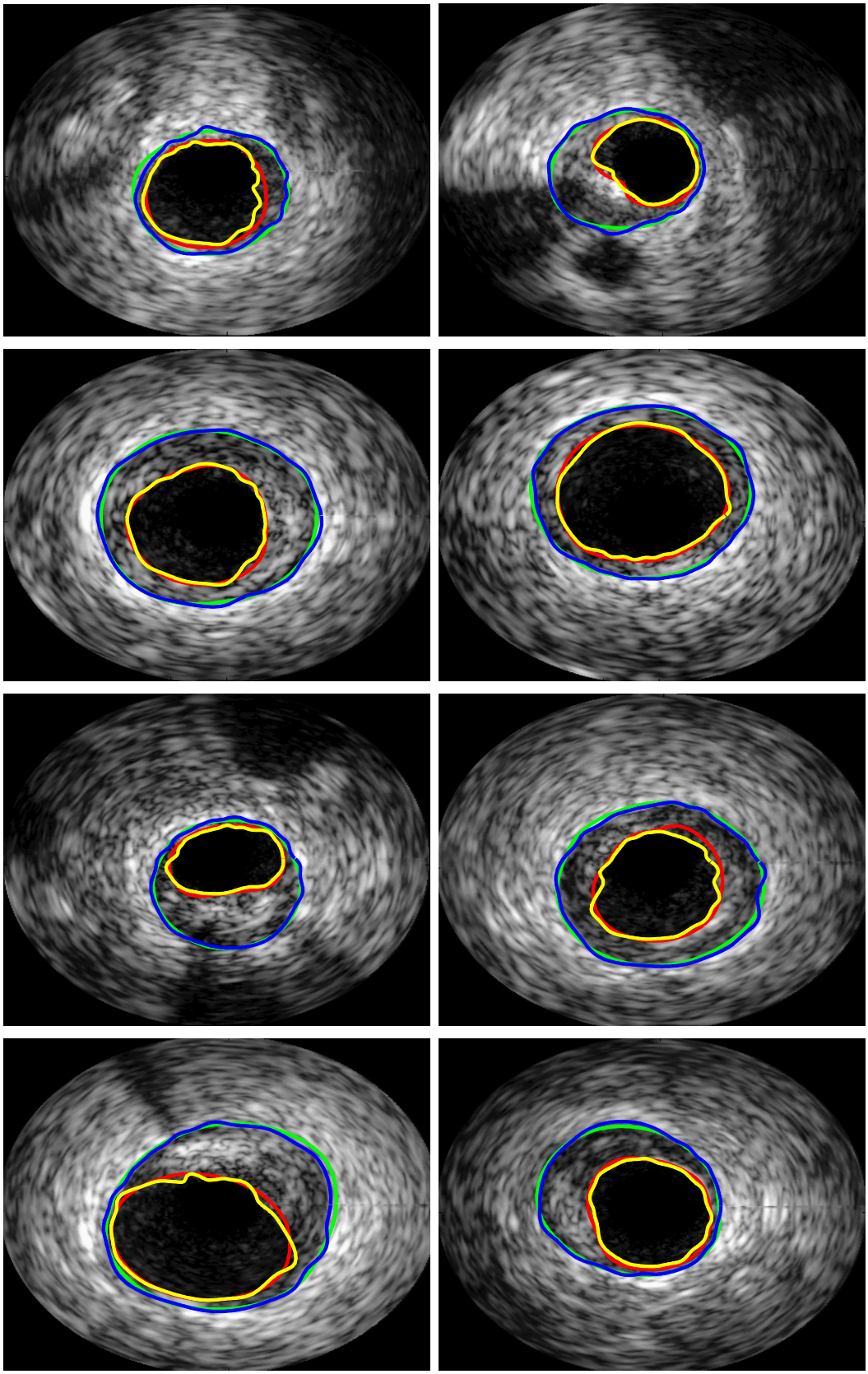


Figure 5.18: Qualitative illustrations of lumen and media segmentation using our method. Each image is a single frame of an IVUS multiframe dataset. Red - Lumen expert tracing, Green - Media expert tracing, Yellow - Lumen segmentation (our method), Blue - Media segmentation (our method).

5.5 Discussion

A novel approach for segmentation of multiple surfaces with convex priors in irregularly sampled space (non-equidistant spacing between orthogonal adjoining nodes) was proposed. The method advances the graph based segmentation framework in several important ways. First, the proposed energy function incorporates a convex surface smoothness penalty in irregularly sampled space through a convex function. Second, the approach allows simultaneous segmentation of multiple surfaces in the irregularly sampled space with the enforcement of a minimum separation constraint. Third, the method guarantees global optimality. Lastly, the proposed method demonstrates utility in achieving subvoxel and super resolution segmentation accuracy while employing a convex penalty to model surface smoothness. To the best of our knowledge, this is the first method that fulfills these three aims at the same time. The hallmark of the proposed method is the ability to perform the segmentation task in an irregularly sampled space which generalizes the optimal surface segmentation framework.

The proposed method is also capable of incorporating convex surface separation penalty while enforcing a minimum separation in the irregularly sampled space. The incorporation of such a penalty would involve modifying the surface separation term in the proposed energy function to impose a convex function based penalty when the minimum separation constraint is not violated. The graph construction to enforce such a penalty can be done using the same framework of the proposed method for enforcing the surface smoothness constraint.

The method can be used in conjunction with the method proposed by Abrámoff et. al [65] to incorporate prior information using trained hard and soft constraints [22] to achieve subvoxel accuracy. Furthermore, the method can also be incorporated in the image segmentation framework using truncated convex priors [77] [95] to achieve subvoxel accuracy by constructing the convex part of the graph in the irregularly

sampled space, thus providing a potential use for generic modelling of variety of surface constraints to achieve subvoxel accuracy.

The global optimality of the proposed method is evident from the illustration in Fig. 5.10, and shows that segmentation performed in the irregularly sampled space based on the displacement of the voxel centers to correctly encode the partial volume information is more accurate compared to the segmentation performed without any use of partial volume information. The results on SD-OCT volumes of the retina show that subvoxel precision is achieved and that segmentation accuracy compared to the OSCS method and DOSCS segmentations is superior. The results on IVUS images demonstrates that the method achieves high accuracy with respect to subvoxel accurate expert tracings as compared to the methods reported in the IVUS challenge, [84] while being fully automated and performing segmentation in 3-D. Our approach is obviously not limited to these two modalities, for which the experiments were conducted.

5.6 Conclusion

In this chapter, the focus was on Aim 2 as proposed for this thesis work. Herein, the develoepd general framework for simultaneous segmentation of multiple surfaces in the irregularly sampled space with convex priors for achievement of subvoxel and super resolution segmentation accuracy was presented. An edge-weighted graph representation was presented and a globally optimal solution with respect to the employed objective function was achieved by solving a maximum flow problem. The surface smoothness and surface separation constraints provide a flexible means for modelling various inherent properties and interrelations of the desired surfaces in an irregularly sampled grid space. The experiments show the applicability of the method and the potential use for various subvoxel and super resolution segmetnation applications. In principle, this work generalizes the framework of the graph search method. The method is readily extensible to higher dimensions.

CHAPTER 6
MULTIPLE SURFACE SEGMENTATION USING DEEP LEARNING (AIM 3)

6.1 Introduction

Convolutional neural networks (CNN) are deep learning architectures [24] [6] [96]. A CNN is a multilayer perceptron and mainly consists of convolutional layers and subsampling layers. The network topology exploits the stationary nature of natural images by learning features using locally connected networks. Convolutional layers are used to learn small feature detectors based on patches randomly sampled from a large image. A feature in the image at some location can be calculated by convolving the feature detector and the image at that location. A subsampling layer is used to reduce the number of features in order to reduce the computational complexity, to introduce invariance properties, and to reduce the chances of overfitting. It summarizes the statistics of a feature over a region in the image.

The current state-of-the-art methods for surface segmentation require expert designed transforms (generation of data cost), constraints and prior information (surface smoothness and surface separation constraints), and tuning of parameters (coeffecients of various energy terms). Furthermore, the these methods are time consuming and require substntial memory depending on the volume size and the number of target surfaces. The primary **objective of Aim 3** is to develop a deep learning (CNN) based method which eliminates any human expert requirement and at the same time is comptationally efficient both in terms of processing time and memory requirement.

Alternatively, most of the applications discussed in Section 2.2 with respect to image segmentation are exmaples where CNNs are used to identify pixels or voxels as belonging to a certain class, called classification or detection as the output of standard CNN applications and not to identify boundaries in the images, i.e. surface segmentation. The surface segmentation in an image would require the CNN to directly infer upon the surface positions of the target surfaces.

In this work a CNN based method was developed to directly learn the surface positions from the input image volume. The key observations which lead to development and application of this method is inspired by the graph search method by making using of the column structure and exploiting the piece-wise smoothness of surfaces.

6.2 Motivation from Graph Search for CNN Based Surface Segmentation

In order to tackle the challenge of surface segmentation using CNNs, two key questions need to be answered. First, since most of the CNN based methods have been used for classification or detection, how can a boundary be detected using a CNN? More precisely, what is the representation of a surface such that a CNN can directly learn the surface positions? Second, how can the CNN learn the constraints implicitly? We answer these questions by exploiting the column structure used in graph search methods by representing the consecutive target surface positions for given columns of an input image as a vector. For example, m consecutive target surface positions are represented as a m-D vector, which may be interpreted as a point in the m-D space, while maintaining a strict order with respect to the consecutiveness of the target surface positions. The ordering of the target surface positions partially encapsulates the smoothness of the surface. Thereafter, the error (loss) function utilized in the CNN to back propagate the error in the network, is chosen as a Euclidean loss function as shown in Equation (6.2), wherein the network adjusts the weights of the various non linear transformations within the network to minimize the euclidean distance between the CNN output and the target surface positions in the m-D space.

Another important observation with respect to surface smoothness is that in the currently used graph search methods for segmentation of surfaces in OCT volumes, the surface smoothness is piecewise in nature or in other words the surface smoothness penalty (cost) enforced in these methods is the sum of the surface smoothness

penalty ascertained using the difference of two consecutive surface positions. Therefore, it is evident, the expert defined transformation is independently not sufficient to get the resultant segmentation but requires an expert designed application specific smoothing term (regularization with respect to the surface profile) to attain accurate segmentations. However, a CNN should be expected to also learn the different smoothness profiles of the target surface. Since, the smoothness is piecewise as can be observed from the energy function shown in Equation 5.1, it should be sufficient for the CNN to learn the different local surface profiles for individual segments of the surface because the resultant surface is a combination of these segments. An example of surface segments representing local smoothness is shown in Fig.6.1. Hence, the CNN is trained on individual patches of the image with segments of the target surface. The hallmark of this method is learning of local surface profiles, which indeed allows the method to be generic in nature and does not require expert designed transformation for different types of scan like (macula or Optic Nerve Head or normal or disease) as required by the traditionally used methods for retinal layer segmentation.

In order to ensure the CNN learns the different local surface profiles, various data augmentations are introduced during training. Furthermore, experiments were conducted as discussed in Section 6.3.4, to discern the size of the local surface segment sufficient to learn the surface profile locally with high accuracy, thus allowing reasonably accurate resulting target surface segmentation formed by the combination of the local segments. The generation of data cost functions (expert designed transformations) as required by the graph based methods may also be interpreted as the features the CNN learns to infer upon the target surface positions.

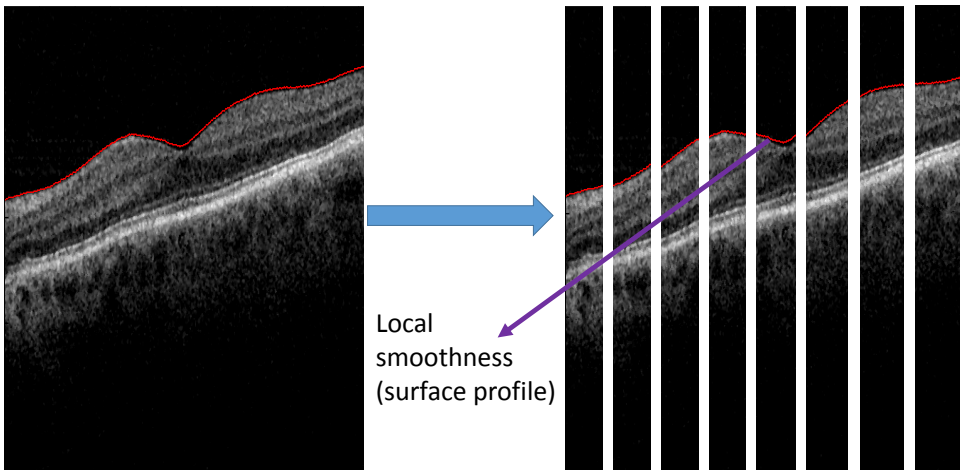


Figure 6.1: Illustration of local smoothness based on patches for a single B-scan from an OCT image. The ILM surface is shown in red. It can be seen that the target surface is infact a combination of sufficiently sized local surface segments which exhibit various types of local surface smoothness (profiles).

6.3 Single Surface Segmentation Using CNNs (Aim 3.1)

6.3.1 Method

Consider an OCT image $I(x, y, z)$ of size $X \times Y \times Z$. A surface is defined as $S(x, y)$, where $x \in \mathbf{x} = \{0, 1, ...X - 1\}$, $y \in \mathbf{y} = \{0, 1, ...Y - 1\}$ and $S(x, y) \in \mathbf{z} = \{0, 1, ...Z - 1\}$. Each (x, y) pair forms a voxel column parallel to the z-axis, wherein the surface $S(x, y)$ intersects each column at a single voxel location. In this work, we present a slice by slice segmentation of an OCT volume. Patches are extracted from B-scans with the target machine generated truth. A patch $P(x_1, z)$ is of size $N \times Z$, where $x_1 \in \mathbf{x_1} = \{0, 1, ...N - 1\}$, $z \in \mathbf{z} = \{0, 1, ...Z - 1\}$ and N is an even number. The target surface to be learnt from P is $\overline{S}(x_2) \in \mathbf{z} = \{0, 1, ...Z - 1\}$ where $x_2 \in \mathbf{x_2} = \{\frac{N}{4}, \frac{N}{4} + 1...\frac{3N}{4} - 1\}$. Essentially, the target surface to be learnt is the surface locations for the middle $\frac{N}{2}$ consecutive columns in P. An example of patch extraction with the target surface is shown in Fig.6.2.

After the extraction of the patches, data augmentation is performed by a com-

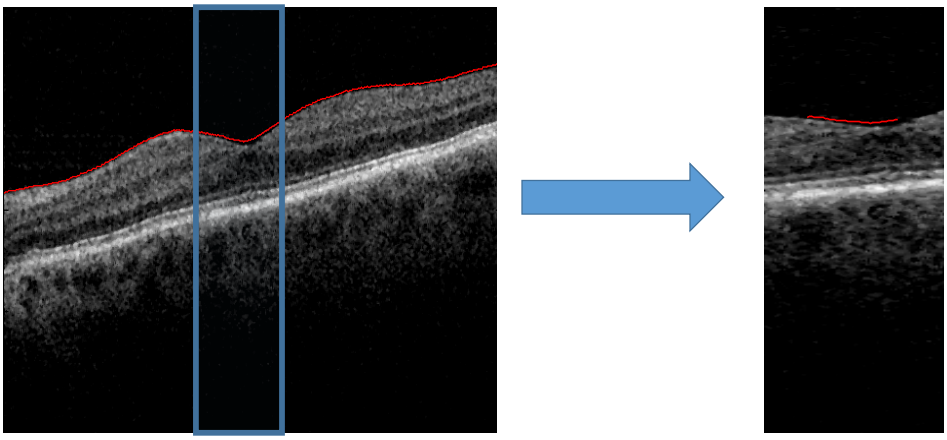


Figure 6.2: Illustration of patch extraction from a B-scan. The target surface is shown in red. The extracted patch shown on the right with the corresponding surface to be used as target for training CNN on the given patch.

bination of translation and rotation to the target surface profile with respect to the image patch. Data augmentation is necessary to increase variance and ensuring that various different types of local profiles with given texture information is learnt by the CNN, thus allowing the CNN to infer on every possible surface profile. For each training patch, three additional training patches were created as described below.

- A random translation value was chosen between -250 and 250 such that the translation was within the range of the patch size. The training patch and the corresponding machine generated truth were translated accordingly.
- A random rotation value was chosen between -55 degrees and 55 degrees. The training patch and the corresponding machine generated truth were rotated accordingly by that amount.
- A combination of the rotation and translation was used to generate another patch.

Special care was taken to extend the generated patches to the original patch size by padding zeros. Examples of data augmentation on patches is shown in Fig.6.3.

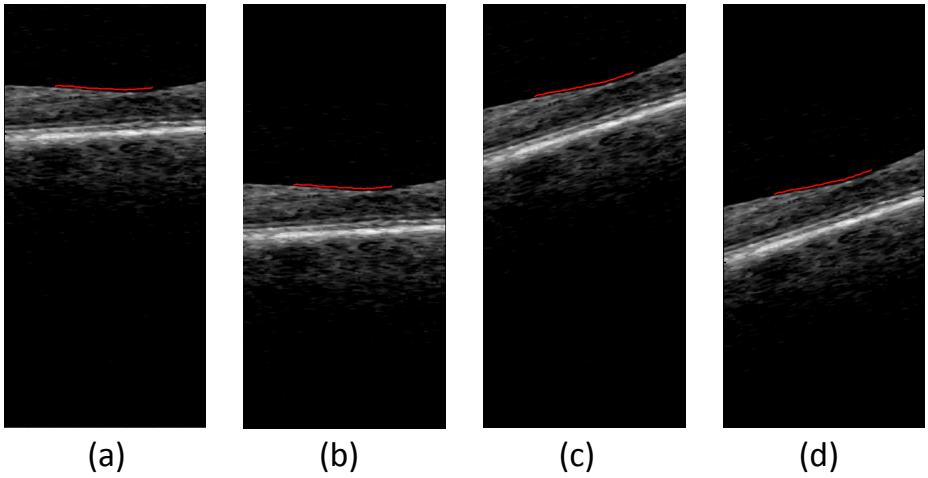


Figure 6.3: Illustration of data augmentation applied to an input patch. The target surface is shown in red. (a)Extracted patch from a B-scan (b)Translation (c)Rotation (d)Translation and rotation, as applied to (a).

6.3.2 Network Architecture

The CNN architecture used in our work is shown in Fig.6.4. The architecture was used to train on patches with N=32. The CNN contains three convolution layers [6], each of which is followed by a max-pooling layer [6] with stride length of two. Thereafter, it is followed by two fully connected layers [6] where the last fully connected layer represents the final output of the middle $\frac{N}{2}$ surface positions for P. From the input patch to the final fully connected layer, the sizes of the feature maps keep decreasing, which helps remove the potential redundant information in P and obtains discriminative features required for inferring the output surface positions for P. Lastly, a Euclidean loss function as shown in Equation (6.2) is utilized to compute the error between CNN output and machine generated truth of P for back propagation during the training phase.

The network starts from a convolution layer, which convolves the input with a number of convolution kernels and yields corresponding number of output feature maps. The convolution operation between an input f and a convolution kernel h is

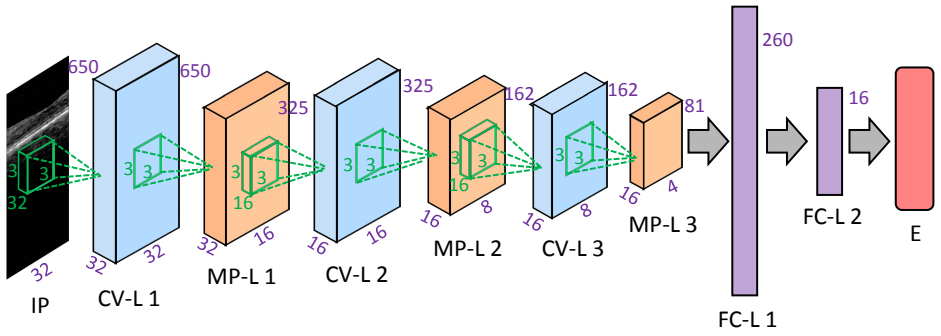


Figure 6.4: The architecture of the CNN learned in our work for N=32. The numbers along each side of the cuboid indicate the dimensions of the feature maps. The inside cuboid (*green*) represents the convolution kernel and the inside square (*green*) represents the pooling region size. The number of hidden neurons in the fully connected layers are marked aside. IP=Input Patch, CV-L=Convolution Layer, MP-L=Max-Pooling Layer, FC-L=Fully Connected Layer, E=Euclidean Loss Layer.

shown in Equation (6.1).

$$s = max(0, \sum_{c} f_c * h_c + b) \tag{6.1}$$

where f_c and h_c is the *c-th* slice from the feature map and the convolution kernel respectively, b is the scalar bias and $*$ is the convolution operation.

We use the rectified linear unit (ReLu) [6] non-linearity in Equation (6.1) for each convolution to perform the non-linear transformations. The ReLu is expressed as $s = \max(0, r)$, where r is the convolution output. Following each convolution layer, a max-pooling layer is introduced, which operates independently on each feature map produced by the previous convolution layer and downsamples each feature map according to the chosen stride size along both the height and width by employing a max operation.

After, three convolution and max-pooling layers, fully connected layers are employed wherein each neuron in the fully connected layer have full connections to all activations in the previous layer. Finally, a Euclidean loss function (used for regress-

ing to real-valued labels) is employed to compute the error on the CNN output.

Unsigned mean surface positioning error [97] is the commonly used error metric for evaluation of surface segmentation accuracy in OCT images. The Euclidean loss function as shown in Equation (6.2), infact computes sum of the squared unsigned surface positioning error between the $\frac{N}{2}$ surface position of the CNN output and the machine generated truth for P, thereby reflecting the exact squared surface positioning error to be used for back propagation during training the network. In other words, the loss function can be interpreted as enforcing a quadratic penalty on the exact difference in the surface positions between the CNN output and the truth.

$$E = \sum_{k_1=0}^{k_1=\frac{N}{2}-1} (\overline{a}_{k_1} - a_{k_1})^2 \tag{6.2}$$

where $\overline{a}_{k_1}$ and a_{k_1} is the k_1-th surface position of the machine generated truth and CNN output respectively for a given P.

6.3.3 Experiment Objectives

In order to train a CNN for surface segmentation, ample amount of OCT data with expert tracings of the target surface is required. However, even when sufficient image data is available, human expert, such as retinal specialists or glaucoma specialists, reference surfaces are extremely expensive, labour intensive, and subject to intra- and interobserver variability. Therefore for this experiment we produced machine generated truth. Experiments using human expert generated truth was carried for mutiple surface segmentation as presented in Section 6.4.3. The popular graph search method [2] was employed for segmenting the OCT data to obtain reference surface tracings to be used by the CNN during the training process. We shall refer to these segmentations as *machine generated truth*. Although, the machine generated truth may not be highly accurate for some data volumes, they are accurate enough for the majority of the data such that the CNN can successful learn the surface segmentations.

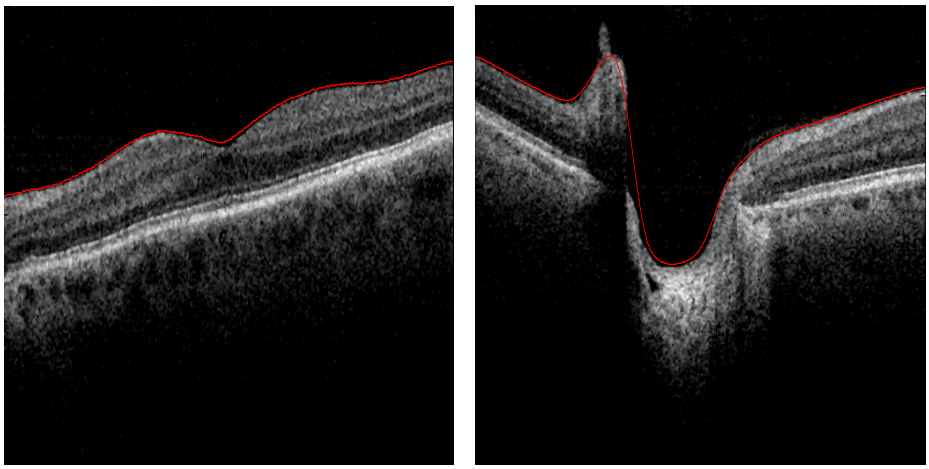


Figure 6.5: Illustration of difference in surface profiles on a single Macula B-scan and ONH B-scan. The ILM surface is shown in red. (left) Macula scan. (right) ONH scan.

The objectives of the experiment are- 1) To demonstrate applicability of the method for segmenting a single suface. 2)To demonstrate that a CNN can learn the target surface with high accuracy with reasonably accurate machine generated annotations and infact yields more accurate segmentations than the graph search method [2] for OCT cases where the method is not accurate enough. 3) The method is invariant with respect to different surface profiles as shall be shown qualitatively by training the networks on Macular scans and infering on Optic Nerve Head Scans. The target surface profile in OCTs vary for numerous reasons, one of them being, type of scans (Macula or Optic Nerve Head). An example is shown in Fig.6.5 to illustrate the difference in surface profile of the Internal Limiting Membrane (ILM) for a Macula and Optic Nerve Head (ONH) B-scan.

6.3.4 Experiment Setting

To demonstrate our method, we segment the Internal Limiting Membrane (ILM) in the OCT volumes as shown in Fig.6.5. 800 SD-OCT scans of normal eyes and eyes with Glaucoma were obtained from an anonymized population study. The Topcon

1000 Spectral Domain OCT scanner was used. The data consisted of both macular and optic nerve head (ONH) centered 3-D volumes of size ($512 \times 128 \times 650$) voxels with voxel size $11.72 \times 3.5 \times 46.88$ μm^3. The data was divided randomly into 700 volumes for training, 90 volumes for testing and 10 volumes for validation. Human expert manual tracings was obtained for 6 B-scans, selected randomly (from 6 sections of 20 B-scans each) in each validation volume. Thus, expert manual tracings for the ILM were only available for 6 B-scans in each of the 10 validation OCT volumes and therefore we employed the multi-resolution graph search approach [20] to obtain machine generated truth over the whole volume image to be used in the training and testing of our method.

The volumes were denoised by applying a median filter of size $5 \times 5 \times 5$. Next the data is normalized with the resultant voxel intensity range varying from -1 to 1. Thereafter, patches of size $N \times 650$ with their respective machine generated truth for the middle $\frac{N}{2}$ consecutive surface positions is extracted using data augmentation as discussed in Section 6.3.1, for both the training and testing volumes, resulting in a training set of $800,000$ and testing set of $90,000$ unique patches. We trained 4 network architectures, similar to the network architecture shown in Fig.6.4 for an input patch size with $N = 48, 32, 16, 8$ to discern and analyze the trade off between accuracy and patch size to be used for sufficiently learning the target surface positions with local smoothness in order to produce highly accurate resultant surface segmentations. The unsigned mean surface positioning error (UMSP) as shown in Eqn (6.3) was used to evaluate the accuracy of the CNN output surface positions with respect to the machine generated truth for the patches in the test set.

$$UMSP = \frac{1}{M} \sum_{k_2=0}^{k_2=M-1} \frac{1}{N/2} \sum_{k_1=0}^{k_1=\frac{N}{2}-1} abs(\overline{a}_{k_1}^{k_2} - a_{k_1}^{k_2}) \tag{6.3}$$

where M is the number of patches, $\overline{a}_{k_1}^{k_2}$ and $a_{k_1}^{k_2}$ is the k_1-th surface position of the

machine generated truth and CNN output respectively for the k_2-th patch.

The best CNN was selected out of the 4 trained networks which had the least UMSP error on the test set and was applied to the 10 validation volumes. Herein, the complete surface for each validation volume was segmented by creating $\frac{512}{N/2}$ patches from each B-scan. The UMSP was computed for each validation volume in a similar manner as shown in Equation (6.3).

In our study we used the NVIDIA GTX Titan X GPU. The publicly available deep learning toolkit CAFFE [98] was used as the CNN implementation in this work.

6.3.5 Results

Quantitative results of the four trained CNNs with $N = 48, 32, 16, 8$ is shown in Table 6.1. The testing UMSP error from CNN-3 and CNN-4 are comparable and therefore, we select CNN-3 as the final CNN for validation, since it has a smaller patch size and slightly outperforms CNN-4. The UMSP error for the surface segmentation obtained from CNN-3 on the validation volumes was also compared to the expert manual tracings. The validation UMSP error for CNN-3 was 0.95 ± 0.06 voxels. Some illustrative results of the surface segmentations from the proposed method using CNN-3 on validation volumes is shown in Fig. 6.6 and it can be observed that the proposed method yields good, consistent and qualitatively similar segmentations to the expert manual tracings. On closer analysis of some B-scans as shown in Fig. 6.7(a), it can be observed that the proposed method segmentations are of superior quality compared to machine generated truth for cases where the machine generated truth is not accurate enough in test set.

Furthermore, the potential and strength of the proposed method is more evident when the trained CNN was applied on the ONH scans where the ILM surface profile is more complex (Fig. 6.5). Although, the CNN was trained on macular scans, the superior quality of the segmentations produced by the proposed method on the ONH scans can be seen in Fig. 6.7(b)(c).

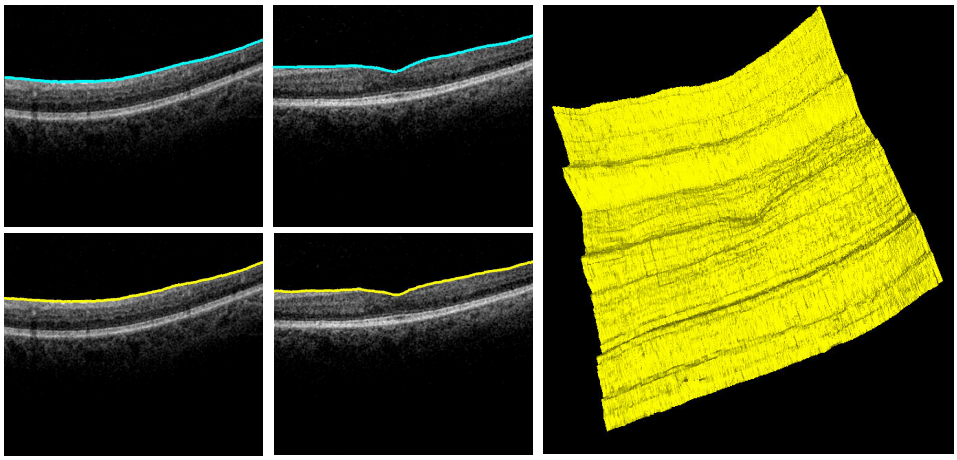


Figure 6.6: (a)-(b) Expert manual tracings and segmentation from the proposed method for the ILM surface on the same macular B-scan is shown in *cyan* and *yellow* respectively. (c) 3-D rendering of the segmented surface using the proposed method for a validation macular OCT volume.

Table 6.1: UMSP error for single surface on testing set from CNN trained for patch size N × Z.

CNN no.	Patch size	No. of surface positions	Testing UMSP error (voxels)
CNN-1	8 × 650	4	2.76
CNN-2	16 × 650	8	1.64
CNN-3	32 × 650	16	1.26
CNN-4	48 × 650	24	1.32

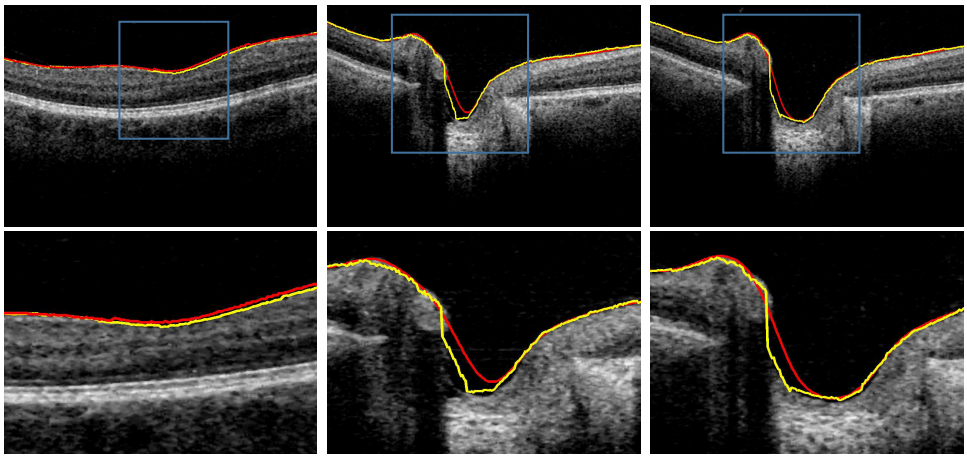


Figure 6.7: Machine generated truth (from graph search method) and segmentation from the proposed method for the ILM surface is shown in *red* and *yellow* respectively. Top row- B-scans in original resolution. Bottom row- Magnified version of the blue boxes to demonstrate superior quality of segmentation from the proposed method. (a)Macular B-scan from test set. (b)-(c) ONH B-scans. Unlike the graph search method where post processing is applied to the output segmentation, no such post processing is applied to the segmentations obtained from the proposed method.

6.4 Multiple Surface Segmentation Using CNNs (Aim 3.2)

The tissue boundaries in OCTs vary by presence and severity of disease. An example is shown in Fig.6.8(a)(b) to illustrate the difference in profile for the Internal Limiting Membrane (ILM) and Inner Retinal Pigment Epithelium (IRPE) in a normal eye and in an eye with AMD.

The method for multiple surface segmentation is very similar to the single surface case and is as follows. Similar to the single surface case, for detecting λ surfaces, the surface positions are represented as a m_2-D vector, where $\lambda = \{1, 2, \dots \lambda\}$, $m_2 = \lambda \times m_1$ and m_1 consecutive surface postions for a surface index i ($i \in \lambda$) are given by $\{((i-1) \times m_1) + 1, ((i-1) \times m_1) + 2, \dots ((i-1) \times m_1) + m_1\}$ index elements in the m_2-D vector.

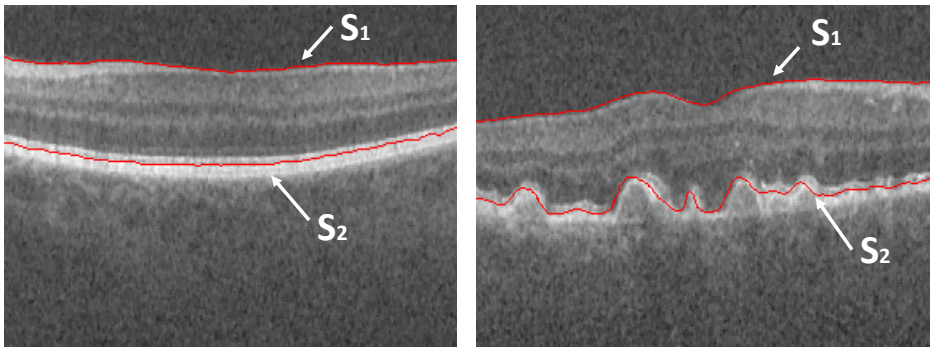


Figure 6.8: Illustration of difference in surface profiles on a single B-scan. (left) Normal Eye (right) Eye with AMD. S_1 = ILM and S_2 = IRPE, are shown in red.

6.4.1 Method

Consider a volumetric image $I(x, y, z)$ of size $X \times Y \times Z$. A surface is defined as $S(x, y)$, where $x \in \mathbf{x} = \{0, 1, ...X - 1\}$, $y \in \mathbf{y} = \{0, 1, ...Y - 1\}$ and $S(x, y) \in \mathbf{z} = \{0, 1, ...Z-1\}$. Each (x, y) pair forms a voxel column parallel to the z-axis, wherein the surface $S(x, y)$ intersects each column at a single voxel location. For simultaneously segmenting $\lambda(\lambda \geq 2)$ surfaces, the goal of the CNN is to learn the surface postions $S_i(x, y)$ $(i \in \lambda)$ for columns formed by each (x, y) pair. In this work, we present a slice by slice segmentation of a 3-D volumetric image applied on OCT volumes. Patches are extracted from B-scans with the target Reference Standard (RS). A patch $P(x_1, z)$ is of size $N \times Z$, where $x_1 \in \mathbf{x_1} = \{0, 1, ...N - 1\}$, $z \in \mathbf{z} = \{0, 1, ...Z - 1\}$ and N is a multiple of 4. The target surfaces S_i's to be learnt simultaneously from P is $\overline{S_i}(x_2) \in \mathbf{z} = \{0, 1, ...Z-1\}$, where $x_2 \in \mathbf{x_2} = \{\frac{N}{4}, \frac{N}{4}+1...\frac{3N}{4}-1\}$. Essentially, the target surfaces to be learnt is the surface locations for the middle $\frac{N}{2}$ consecutive columns in P. Then data augmentation is performed as described before.

6.4.2 Network Architecture

For segmenting λ surfaces simultaneously, the CNN learns λ surfaces for each patch. In our work, $\lambda = 2$ surfaces are segmented simultaneously and therefore for

each patch, the CNN learns two surfaces. The CNN architecture used in our work is shown in Fig.6.9, employed for $\lambda = 2$ and patches with $N = 32$. The architecture was used to train on patches with N=32. The CNN contains three convolution layers [6], each of which is followed by a max-pooling layer [6] with stride length of two. Thereafter, it is followed by two fully connected layers [6], where the last fully connected layer represents the final output of the middle $\frac{N}{2}$ surface positions for 2 target surfaces in P. We use the rectified linear unit (ReLu) [6] non-linearity for each convolution to perform the non-linear transformations. Lastly, a Euclidean loss function (used for regressing to real-valued labels) as shown in Equation (6.4) is utilized to compute the error between CNN outputs and RS of S_i's ($i \in \lambda$) within P for back propagation during the training phase. Unsigned mean surface positioning error (UMSP) as shown in as shown in Eqn (6.3) is one of the commonly used error metric for evaluation of surface segmentation accuracy. The Euclidean loss function (E), essentially computes sum of the squared unsigned surface positioning error over the $\frac{N}{2}$ consecutive surface position for S_i's of the CNN output and the RS for P, thereby reflecting the exact squared surface positioning error. to be used for back propagation during training the network.

$$E = \sum_{i=1}^{i=\lambda} \sum_{k_1=0}^{k_1=\frac{N}{2}-1} (\overline{a}^i_{k_1} - a^i_{k_2})^2 \tag{6.4}$$

where $k_2 = ((i-1) \times N/2) + k_1$, $\overline{a}^i_{k_1}$ and $a^i_{k_2}$ is the k_1-th surface position of reference standard and CNN output respectively, for surface S_i in a given P.

6.4.3 Experiments

The experiments compare segmentation accuracy of the proposed CNN based method (CNN-S) and the G-OSC method [9]. The two surfaces simultaneously segmented in this study are S_1-ILM and S_2-IRPE as shown in Fig. 6.8. 115 OCT scans of normal eyes, 269 OCT scans of eyes with AMD and their respective reference

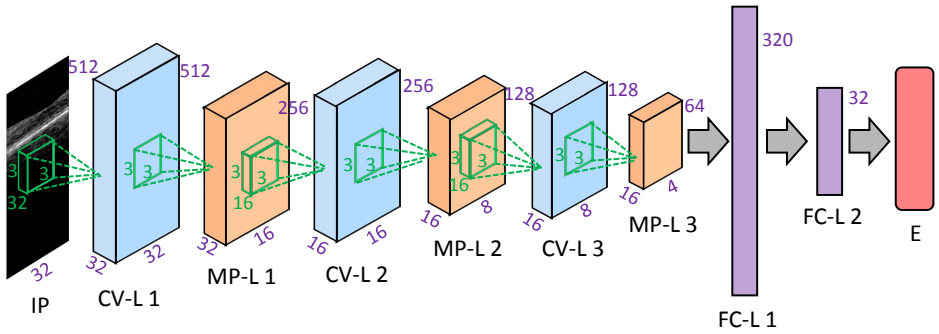


Figure 6.9: The architecture of the CNN learned in our work for N=32 and $\lambda = 2$. The numbers along each side of the cuboid indicate the dimensions of the feature maps. The inside cuboid (*green*) represents the convolution kernel and the inside square (*green*) represents the pooling region size. The number of hidden neurons in the fully connected layers are marked on the side. IP=Input Patch, CV-L=Convolution Layer, MP-L=Max-Pooling Layer, FC-L=Fully Connected Layer, E=Euclidean Loss Layer.

standards (RS) (created by a single expert with aid of the DOCTRAP software [99]) were obtained from the publicly available repository [68]. The 3-D volume size was $1000 \times 100 \times 512$ voxels with voxel size $6.54 \times 67 \times 3.23$ μm^3. The data volumes were divided into a training set (79 normal and 187 AMD), a testing set (16 normal and 62 AMD) and a validation set (20 normal and 20 AMD). The volumes were denoised by applying a median filter of size $5 \times 5 \times 5$ and normalized with the resultant voxel intensity varying from -1 to 1. Thereafter, patches of size $N \times 512$ with their respective RS for the middle $\frac{N}{2}$ consecutive surface positions for S_1 and S_2 is extracted using data augmentation, for training and testing volumes, resulting in a training set of $340,000$ and testing set of $70,000$ patches. In our work, we use $N = 32$. The UMSP, calculated as the average of absolute differences between the computed surface and reference standard for each voxel column in the image, was used to evaluate the accuracy. The complete surfaces for each validation volume were segmented using the CNN-S method by creating $\frac{1016}{N/2}$ patches from each B-scan where each B-scan was zero padded with 8 voxel columns at each extremity. Statistical significance of the observed differences was determined by paired Student t-tests with a p value of 0.05

was considered significant. In our study we used a single NVIDIA Titan X GPU for training the CNN. The publicly available deep learning toolkit CAFFE [98] was used as the CNN implementation. The validation of the G-OSC and CNN-S method were carried out on a on a Linux workstation (3.4 GHz, 16 GB memory). A single CNN was trained to segment both the normal and AMD OCT scans. For a comprehensive comparison, three experiments were performed with the G-OSC method. The first experiment (G-OSC 1) involved segmenting the surfaces in both normal and AMD OCT scans using a single set of optimized parameters. The second (G-OSC 2) and third (G-OSC 3) experiment involved segmenting the normal and AMD OCT scans with different set of expert-designed, optimized parameters, respectively.

6.4.4 Results

The UMSP of the CNN-S method on the testing patches for S_1 was 1.02 ± 0.66 voxels and for S_2 was 1.73 ± 0.91 voxels. The quantitative comparisons between the proposed CNN-S method and the G-OSC method on the validation volumes is summarized in Table 6.2. For the entire validation data, the proposed method produced a significantly lower UMSPE for surfaces S_1 ($p < 0.01$) and S_2 ($p < 0.01$), compared to the segmentation results of G-OSC 1, G-OSC 2 and G-OSC 3. Illustrative results of segmentations from the CNN-S, G-OSC 2 and G-OSC 3 methods on validation volumes are shown in Fig. 6.11. Herein, the illustrations for G-OSC 2 and G-OSC 3 is shown. The CNN-S method yields consistent and qualitatively superior segmentations with respect to the G-OSC method. On closer analysis of some B-scans in the validation data, the CNN-S method produceda high quality segmentation for a few cases where the RS was not accurate enough as verifed by an retinal specialist, (4th row in Fig. 6.11). The features extracted from the first convolution layer is shown in Fig. 6.10. The CNN required 17 days to train on the GPU. The CNN-S method with average computation time of 94.34 seconds (95.35 MB memory) is much faster than G-OSC with average computation time of 2837.46 seconds (6.87 GB memory).

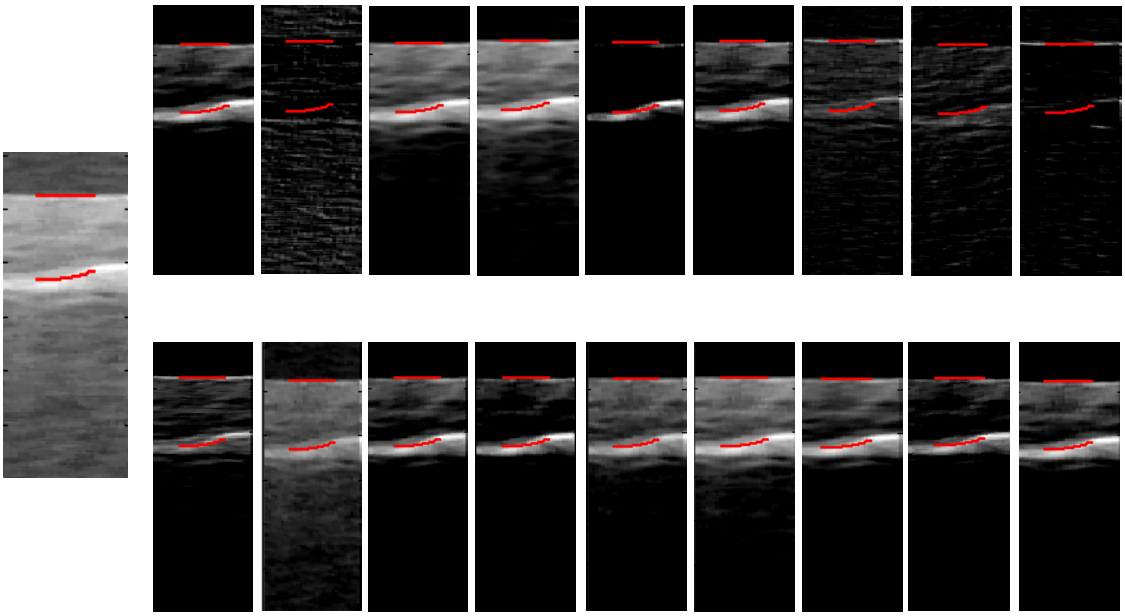


Figure 6.10: Illustration of 20 chosen feature maps from the first convolution layer for a given patch shown on the left. The ILM and IRPE are shown in red.

6.5 Discussion

The results from both the experiments show that a CNN based approach can be used for segmentation of surfaces in volumetric images, thus eliminating the requirement of human intervention for expert design of various transforms. The first experiment for single surface segmentation leverages commonly used surface segmentation method [20] to obtain machine generated truth for a large amount of data, thus allowing training of the CNN. The results also show that our CNN based method can learn the segmentations with sufficient accuracy and in fact results in qualitatively better segmentations for cases where the machine generated surface segmentation is not accurate enough. The developed method demonstrates that even if human denoted labels/truth is unavailable, a CNN can sufficiently learn the target based on a reliable accurate enough automated/semi-automated method.

The second experiment for multiple surface segmentation using CNN demonstrates

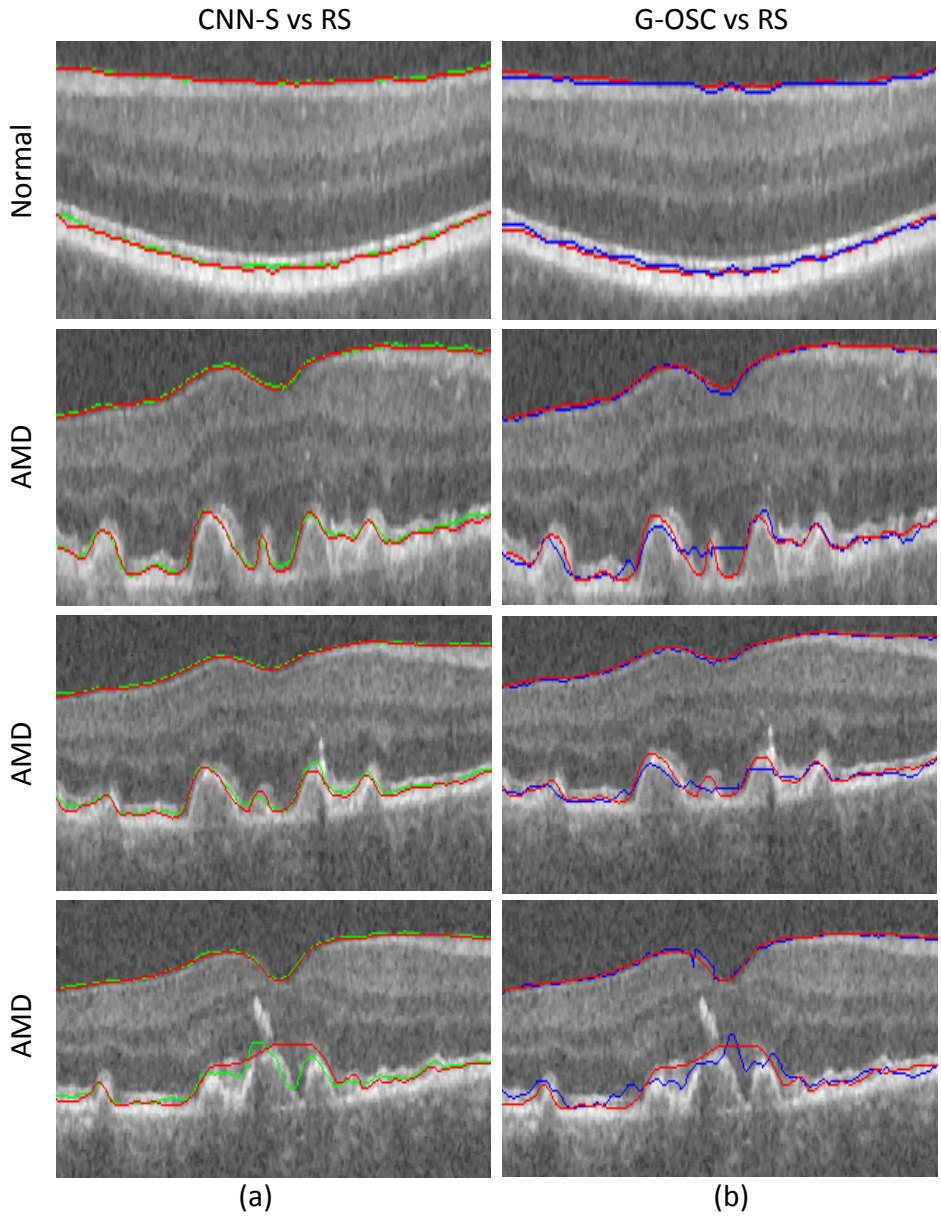


Figure 6.11: Each row shows the same B-scan from a Normal or AMD OCT volume. (a) CNN-S vs. RS (b) G-OSC vs. RS, for surfaces S_1 =ILM and S_2 = IRPE. RS = Reference Standard, Red = reference standard, Green = Segmentation using proposed method and Blue = Segmentation using G-OSC method. In the 4th row, we had the reference standard reviewed by a fellow-ship trained retinal specialist, who stated that the CNN-S method is closer to the real surface than the reference standard.

Table 6.2: UMSPE expressed as (mean ± 95% CI) in voxels. RS - Reference Standard. A single CNN was trained to infer on each type of data while 3 different set of parameters were used for the G-OSC method. $N = 32$ was used as the patch size (32×512).

	Normal and AMD		Normal		AMD	
Surface	G-OSC 1 vs. RS	CNN-S vs. RS	G-OSC 2 vs. RS	CNN-S vs. RS	G-OSC 3 vs. RS	CNN-S vs.RS
S_1	1.45 ± 0.19	0.98 ± 0.08	1.19 ± 0.05	0.89 ± 0.07	1.37 ± 0.22	1.06 ± 0.11
S_2	3.17 ± 0.43	1.56 ± 0.15	1.41 ± 0.11	1.28 ± 0.10	2.88 ± 0.54	1.83 ± 0.26
Overall	2.31 ± 0.29	1.27 ± 0.13	1.31 ± 0.07	1.08 ± 0.08	2.13 ± 0.39	1.44 ± 0.19

learning of the surface segmentation from expert manual tracings. The results show that segmentations by CN-S are better than those from the G-OSC method, while eliminating the requirement of expert designed transforms or adjustments.

The key aspect of the method is the elimination of expert intervention, efficiencies in terms of memory requirement and processing times and the generic nature of the method compared to traditional graph search approaches. This is possible because learning local smoothness allows for a much more generic way of learning the segmentation so that it is robust across a variety of target surface profiles as shown in the illustrations in Fig. 6.6 and 6.11. The experiments show that the developed method used a single CNN to learn various local surface profiles for macular scans in the first case and both normal and AMD data in the second case. The strength and genericness of the method can be seen from Fig 6.7 where the CNN learnt from surface profiles of macular scans but was still able to provide superior qualitative segmentations compared to the multi-resolution graph search method on ONH scans. In the second experiment, comparison to G-OSC 1 shows that the CNN-S methods outperforms the G-OSC method. If the parameters are tuned specifically for each type of disease by using expert prior knowledge while using the G-OSC method, as in the cases of G-OSC 2 and G-OSC 3, a common practice in medical image analysis, the results depict that the CNN-S method still results in superior performance. Therefore, the method has potential for applicability for in clinical setting, where a

single network may be trained for segmenting OCT volumes of various types, instead of mutiple versions of graph search based method to tackle each type of OCT volume.

Another advantage of the proposed method can be realized with respect to the computation time for obtaining the surface segmentations. The graph search method [2] requires a computation time of minutes to hours for obtaining the segmentation of a OCT volume, mainly because the methods requires processing of the entire volume at once. The improved graph based method [20] tackled this problem by using an expert designed multi-resolution scheme to reduce the computation time to minutes. Our proposed CNN based method is patch based and inference on each patch is computed independently of the other, thereby allowing the use of sophisticated hardware like GPUs to process the entire set of patches for a given volume within seconds. The second experiment showed that the inference using the developed method is also much faster than G-OSC method and requires less memory. In fact we plan to parallelize the segmentation for multiple patches, thereby further reducing the computation time, allowing more effetcive interactive use.

The smoothness of a surface at a voxel column level can also be imposed within the loss function of the CNN by adding to the current loss function another term that sums the squared difference of the difference between surface positions of the neighboring columns given by the CNN output and reference segmentation. Such an incorporation to the loss function shall potentially make the CNN more robust, accurate and closely mimic the smoothness terms used in the state-of-the-art methods [47] [2] [20]. The method is readily extensible to 3-D segmentation by employing a 3-D CNN architecture with corresponding representation of the surface for the CNN to train on 3-D patches wgile employing 3-D convolutions.

There are several issues with the developed approach. First, the current work trains on data from images obtained from one kind of a scanner, however it is possible that the trained CNN may not produce consistent segmentations on images

obtained from a different kind of scanner due to the difference in textural and spatial information. Such a challenge may be tackled by creating a training set that spans across images from different types of scanner or by training another CNN which refines the output of the proposed CNN for the given different type of scanner. Second, a drawback of any such learning approach in medical imaging is the limited amount of available training data. Third, the method may require long time to train the network depending on the available processing hardware.

6.6 Conclusion

In this chapter, the focus was to accomplish Aim 3 of this thesis work by developing a CNN based method for segmentation of surfaces in volumetric images without any human expert intervention and implicitly learned surface smoothness. The experiments demonstrated the performance and potential of the developed method through application on normal OCT volumes to segment the ILM surface as well as on normal and AMD OCT volumes to segment the ILM and IRPE surface. The single surface segmentation experiment results show high quantitative accuracy in segmentation of macula scans, while also yielding higher quality segmentation for ONH scans, even though the CNN was trained on macular scans only. The multiple surface segmetnation experiment results show superior segmentation accuracy, lower processing time and memory requirements compared to the G-OSC method. To the best of our knowledge, this is the first method of its kind that does not require any human intervention for surface segmentation.

CHAPTER 7
CONCLUSIONS

In this dissertation, novel multiple surface segmentation methods using graph search based and deep learning based approaches were developed and validated on a variety of intra-retinal layer segmentation applications in SD-OCT volumes. The development of the methods were motivated from real world segmentation problems wherein, the traditional graph search methods are not efficient enough and may have difficulty in solving the segmentation problem. Furthermore, the need for a generic method for multiple surface segmentation which does not require human expert intervention was discussed and a CNN based method was proposed to segment surfaces which is both efficient and more genreic as compared to tradional graph search methods. Herein, a breif review of each of those methods and possible related future works are discussed.

The multiple surface segmentation with truncated convex priors is motivated from real world applications as in the case of segmenting surfaces in SD-OCT volumes with AMD, PED and glaucoma; wherein sharp changes in surface smoothness and abrupt changes in surface separation between two mutually coupled surfaces may exist because of the presence of a pathology. Traditional graph search based methods are inefficent and may find difficulty in segmenting such surfaces becuase the surface constraints generally over smooth the sharp surface jumps or over penalize the abrupt changes in surface separations. The proposed method with truncated convex priors, truncates the convex penalty by a pre-defined truncation factor. Thus, allowing for segmentation of complex surfaces and acts as a disconinuity preserving method. The method was validated on a variety of such surface segmentation applications and the experiment results shows the impoved performance and segmentation accuracy of the method as compared to grpah search methods. The method is also capable of segmenting a large volumetric image in the original resolution because of the iterative

nature of the approach based on creating small subgraphs at each iteration and hence, does not require a multiple resolution approach.

The proposed method for truncated convex priors, may be extended and used to constrain two closely related surfaces. For example, if the inner and outer aspect of the retinal pigment epithelial (RPE) in diseased cases follows a similar surface profile (the layer thickness is constant), then while segmenting other complex surfaces simultaneously with the two RPE layers, a convex penalty without truncation can be imposed between the two surfaces and the other compex surfaces may be modeled with the truncated convex priors. Also, similar to the multiple resolution approach, the image may be segmented in the downsampled version to create the surface initialization for the upsampled scale, which may allow the method to be more acurate and achieve faster convergence.

The optimal surface segmentation with convex priors method for irrgularly sampled space is developed for practical purposes of subvoxel and super resolution accurate segmentations. The current graph search methods are not capable of segmentations with subvoxel and super resolution accuracy, due to the inherenet nature of the construction and encoding of the underlying graph. Herein, nodes in the graph represent the center of voxels which does not explot the partial volume effects in the images, which can be utilized to achieve subvoxel and super resolution segmentation accuracy. Thus, the method optimal surface segmentation method for irregularly sampled space is proposed in this work. The developed method was extensively, validated for subvoxel and super resolution accuracy applications. The proof of correctenss of the graph construction with respect to the global optimality of the solution was also presented. The method, infact generalizes the traditional graph search method, which is a special case of the developed method.

For future work, it will be interesting to integrate the approach with the truncated convex method by modelling the convex part of the graph using the developed

method, thus possibly further increasing the segmentation accuracy of the method for surface segmentations, especially for disease cases. Furthermore, region of interest is a prominent tool employed for various image segmentation application. The goal of the region of interest is to reduce the search space for a segmentation solution. Graph search based methods utilized for multiple surface segmentation of retinal surfaces in OCT volumes, extensively use region of interests to decrease the size of the solution space and to provide for a more strict solution space such that the target surface segmentation is more accurate. Generally, the design of the region of interest in these methods defines a region with a constant length for each column in the graph (like a uniformly sized band based on the pre-segmentation result). The developed method can be applied to simultaneously segment multiple surfaces in irregularly shaped and possibly disjoint region of interests. Once the region of interest is generated, the graph is built only inside the region of interest using the proposed method to search the target segmentations, irrespective of the shape and location of the region of interest. This shall allow for more flexibility and usage of machine learning techniques to design the region of interests.

From the developed methods for Aim 1 and Aim 2 of this thesis work, it is clear that even though graph based segmentation methods are flexible and robust, they suffer from major drawbacks. The methods have to be redesigned for different kind of applications which requires expert design of the various terms in the energy function of the graphs. The design of these terms, require expert defined transforms, feature extractions, constraint design and parameter tuning. The multi-resoltion method may also require expert defined schemes of the order of segmentation of the various surfaces. Thus, the multiple surface segmentation method using deep learning, developed in this work is motivated from the need of a generic method for surface segmentation which does not require expert human intervention. The method was developed by exploiting certain aspects of the graph search method,

specifically the usage of column structure. The developed CNN based method learns the required transformations from the training data. The method was validated on normal and AMD OCT volumes for single and multiple surface segmetnation applications. The experiments clearly demonstrated that the method is infact more generic and accurate than graph search. The local patch based surface profile learning, allows the method to be faster and also more generic. The experiments showed that the developed method is efficient in terms of memory requirement and processing time. The developed method, has the potential of training one universal network which can segment the retinal layer in OCTs of any type in clinical settings. However, the method is limited by the amount of data available.

Future works may include incorporation of surface smoothness and surface separation constraints as used in graph search within the loss function used in the training of the CNN to more closely mimic the graph search framework. Recently, U-Net [100] has gained a lot of popularity for biomedical image segmentation where the voxels are classifed as belonging to a class. The U-Net could be used in conjunction with the developed method by first training the U-Net and then combining the solution of the U-Net to the patch used in training of the CNN. Such a combination may result in more robustness and accurate segmentations. Further, another area of future work may be to extend the developed method to not only segment the surfaces but also classify the OCT volumes into various classes like disease and no disease.

In this doctoral dissertation novel multiple surface segmentation methods were developed and validated extensively for a variety of surface segmetnation problems. The developed methods are readily extendable to higher dimensions and are not limited to the image modalities or segmentaion applications discussed in this work.

REFERENCES

[1] http://www.wildml.com/2015/11/understanding-convolutional-neural-networks-for-nlp/.

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

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

[4] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 23, no. 11, pp. 1222–1239, Nov 2001.

[5] K. Fukushima, "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position," *Biological cybernetics*, vol. 36, no. 4, pp. 193–202, 1980.

[6] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.

[7] X. Wu and D. Z. Chen, "Optimal net surface problems with applications," in *Automata, Languages and Programming*. Springer, 2002, pp. 1029–1042.

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

[9] Q. Song, J. Bai, M. K. Garvin, M. Sonka, J. M. Buatti, and X. Wu, "Optimal multiple surface segmentation with shape and context priors," *Medical Imaging, IEEE Transactions on*, vol. 32, no. 2, pp. 376–386, 2013.

[10] Y. Yin, X. Zhang, R. Williams, X. Wu, D. D. Anderson, and M. Sonka, "Logismoslayered optimal graph image segmentation of multiple objects and surfaces: cartilage segmentation in the knee joint," *Medical Imaging, IEEE Transactions on*, vol. 29, no. 12, pp. 2023–2037, 2010.

[11] S. Kashyap, Y. Yin, and M. Sonka, "Automated analysis of cartilage morphology," in *Biomedical Imaging (ISBI), 2013 IEEE 10th International Symposium on*. IEEE, 2013, pp. 1300–1303.

[12] X. Wu, X. Dou, A. Wahle, and M. Sonka, "Region detection by minimizing intraclass variance with geometric constraints, global optimality, and efficient approximation," *Medical Imaging, IEEE Transactions on*, vol. 30, no. 3, pp. 814–827, 2011.

[13] H. Zhang, A. K. Abiose, D. Gupta, D. N. Campbell, J. B. Martins, M. Sonka, and A. Wahle, "Novel indices for left-ventricular dyssynchrony characterization based on highly automated segmentation from real-time 3-d echocardiography," *Ultrasound in medicine & biology*, vol. 39, no. 1, pp. 72–88, 2013.

[14] X. Liu, D. Z. Chen, M. H. Tawhai, X. Wu, E. A. Hoffman, and M. Sonka, "Optimal graph search based segmentation of airway tree double surfaces across bifurcations," *Medical Imaging, IEEE Transactions on*, vol. 32, no. 3, pp. 493–510, 2013.

[15] C. Bauer, M. A. Krueger, W. J. Lamm, B. J. Smith, R. W. Glenny, and R. R. Beichel, "Airway tree segmentation in serial block-face cryomicrotome images of rat lungs," *Biomedical Engineering, IEEE Transactions on*, vol. 61, no. 1, pp. 119–130, 2014.

[16] S. Sun, M. Sonka, and R. R. Beichel, "Lung segmentation refinement based on optimal surface finding utilizing a hybrid desktop/virtual reality user interface," *Computerized Medical Imaging and Graphics*, vol. 37, no. 1, pp. 15–27, 2013.

[17] X. Zhang, J. Tian, K. Deng, Y. Wu, and X. Li, "Automatic liver segmentation using a statistical shape model with optimal surface detection," *Biomedical Engineering, IEEE Transactions on*, vol. 57, no. 10, pp. 2622–2626, 2010.

[18] Q. Song, Y. Liu, Y. Liu, P. K. Saha, M. Sonka, and X. Wu, "Graph search with appearance and shape information for 3-d prostate and bladder segmentation," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2010*. Springer, 2010, pp. 172–180.

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

[20] K. Lee, M. Niemeijer, M. K. Garvin, Y. H. Kwon, M. Sonka, and M. D. Abramoff, "Segmentation of the optic disc in 3-d oct scans of the optic nerve head," *Medical Imaging, IEEE Transactions on*, vol. 29, no. 1, pp. 159–168, 2010.

[21] M. D. Abràmoff, M. K. Garvin, and M. Sonka, "Retinal imaging and image analysis," *Biomedical Engineering, IEEE Reviews in*, vol. 3, pp. 169–208, 2010.

[22] P. A. Dufour, L. Ceklic, H. Abdillahi, S. Schroder, S. De Dzanet, U. Wolf-Schnurrbusch, and J. Kowal, "Graph-based multi-surface segmentation of oct data using trained hard and soft constraints," *Medical Imaging, IEEE Transactions on*, vol. 32, no. 3, pp. 531–543, 2013.

[23] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Backpropagation applied to handwritten zip code recognition," *Neural computation*, vol. 1, no. 4, pp. 541–551, 1989.

[24] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[25] F. Ning, D. Delhomme, Y. LeCun, F. Piano, L. Bottou, and P. E. Barbano, "Toward automatic phenotyping of developing embryos from videos," *IEEE Transactions on Image Processing*, vol. 14, no. 9, pp. 1360–1371, 2005.

[26] P. Sermanet and Y. LeCun, "Traffic sign recognition with multi-scale convolutional networks," in *Neural Networks (IJCNN), The 2011 International Joint Conference on*. IEEE, 2011, pp. 2809–2813.

[27] D. C. Ciresan, U. Meier, J. Masci, L. Maria Gambardella, and J. Schmidhuber, "Flexible, high performance convolutional neural networks for image classification," in *IJCAI Proceedings-International Joint Conference on Artificial Intelligence*, vol. 22, no. 1. Barcelona, Spain, 2011, p. 1237.

[28] T. Wang, D. J. Wu, A. Coates, and A. Y. Ng, "End-to-end text recognition with convolutional neural networks," in *Pattern Recognition (ICPR), 2012 21st International Conference on*. IEEE, 2012, pp. 3304–3308.

[29] S. Ji, W. Xu, M. Yang, and K. Yu, "3d convolutional neural networks for human action recognition," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 1, pp. 221–231, 2013.

[30] Z. Yan, Y. Zhan, Z. Peng, S. Liao, Y. Shinagawa, D. N. Metaxas, and X. S. Zhou, "Bodypart recognition using multi-stage deep learning," in *International Conference on Information Processing in Medical Imaging*. Springer, 2015, pp. 449–461.

[31] M. B. G. Challenge, "Multimodal brain tumor segmentation benchmark: Change detection," http://braintumorsegmentation.org/, accessed November 5, 2016.

[32] W. Shen, M. Zhou, F. Yang, C. Yang, and J. Tian, "Multi-scale convolutional neural networks for lung nodule classification," in *International Conference on Information Processing in Medical Imaging*. Springer, 2015, pp. 588–599.

[33] T. Brosch, Y. Yoo, L. Y. Tang, D. K. Li, A. Traboulsee, and R. Tam, "Deep convolutional encoder networks for multiple sclerosis lesion segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2015, pp. 3–11.

[34] N. Tajbakhsh, M. B. Gotway, and J. Liang, "Computer-aided pulmonary embolism detection using a novel vessel-aligned multi-planar image representation and convolutional neural networks," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2015, pp. 62–69.

[35] X. Chen, Y. Xu, S. Yan, D. W. K. Wong, T. Y. Wong, and J. Liu, "Automatic feature learning for glaucoma detection based on deep learning," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2015, pp. 669–677.

[36] J. Wang, J. D. MacKenzie, R. Ramachandran, and D. Z. Chen, "Detection of glands and villi by collaboration of domain knowledge and deep learning," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2015, pp. 20–27.

[37] N. Dhungel, G. Carneiro, and A. P. Bradley, "Deep learning and structured prediction for the segmentation of mass in mammograms," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2015, pp. 605–612.

[38] Kaggle, "Diabetic retinopathy detection," http://www.kaggle.com/c/diabetic-retinopathy-detection/, accessed July 15, 2016.

[39] M. Wojtkowski, T. Bajraszewski, I. Gorczyńska, P. Targowski, A. Kowalczyk, W. Wasilewski, and C. Radzewicz, "Ophthalmic imaging by spectral optical coherence tomography," *American journal of ophthalmology*, vol. 138, no. 3, pp. 412–419, 2004.

[40] Z. Yaqoob, J. Wu, and C. Yang, "Spectral domain optical coherence tomography: a better oct imaging strategy," *Biotechniques*, vol. 39, 2005.

[41] A. M. Bagci, M. Shahidi, R. Ansari, M. Blair, N. P. Blair, and R. Zelkha, "Thickness profiles of retinal layers by optical coherence tomography image segmentation," *American journal of ophthalmology*, vol. 146, no. 5, pp. 679–687, 2008.

[42] T. Fabritius, S. Makita, M. Miura, R. Myllylä, and Y. Yasuno, "Automated segmentation of the macula by optical coherence tomography," *Optics express*, vol. 17, no. 18, pp. 15 659–15 669, 2009.

[43] S. Lu, C. Y.-l. Cheung, J. Liu, J. H. Lim, C. K.-s. Leung, and T. Y. Wong, "Automated layer segmentation of optical coherence tomography images," *IEEE Transactions on biomedical engineering*, vol. 57, no. 10, pp. 2605–2608, 2010.

[44] J. Canny, "A computational approach to edge detection," *IEEE Transactions on pattern analysis and machine intelligence*, no. 6, pp. 679–698, 1986.

[45] M. Mujat, R. C. Chan, B. Cense, B. H. Park, C. Joo, T. Akkin, T. C. Chen, and J. F. de Boer, "Retinal nerve fiber layer thickness map determined from optical coherence tomography images," *Optics Express*, vol. 13, no. 23, pp. 9480–9491, 2005.

[46] A. Mishra, A. Wong, K. Bizheva, and D. A. Clausi, "Intra-retinal layer segmentation in optical coherence tomography images," *Optics express*, vol. 17, no. 26, pp. 23 719–23 728, 2009.

[47] A. Yazdanpanah, G. Hamarneh, B. Smith, and M. Sarunic, "Intra-retinal layer segmentation in optical coherence tomography using an active contour approach," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2009, pp. 649–656.

[48] V. Kajić, B. Považay, B. Hermann, B. Hofer, D. Marshall, P. L. Rosin, and W. Drexler, "Robust segmentation of intraretinal layers in the normal human fovea using a novel statistical model based on texture and shape analysis," *Optics express*, vol. 18, no. 14, pp. 14 730–14 744, 2010.

[49] F. Rathke, S. Schmidt, and C. Schnörr, "Order preserving and shape prior constrained intra-retinal layer segmentation in optical coherence tomography," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2011, pp. 370–377.

[50] R. J. Zawadzki, A. R. Fuller, D. F. Wiley, B. Hamann, S. S. Choi, and J. S. Werner, "Adaptation of a support vector machine algorithm for segmentation and visualization of retinal structures in volumetric optical coherence tomography data sets," *Journal of biomedical optics*, vol. 12, no. 4, pp. 041 206–041 206, 2007.

[51] F. Rossant, I. Ghorbel, I. Bloch, M. Paques, and S. Tick, "Automated segmentation of retinal layers in oct imaging and derived ophthalmic measures," in *Biomedical Imaging: From Nano to Macro, 2009. ISBI'09. IEEE International Symposium on*. IEEE, 2009, pp. 1370–1373.

[52] K. Vermeer, J. Van der Schoot, H. Lemij, and J. De Boer, "Automated segmentation by pixel classification of retinal layers in ophthalmic oct images," *Biomedical optics express*, vol. 2, no. 6, pp. 1743–1756, 2011.

[53] Q. Yang, C. A. Reisman, Z. Wang, Y. Fukuma, M. Hangai, N. Yoshimura, A. Tomidokoro, M. Araie, A. S. Raza, D. C. Hood *et al.*, "Automated layer segmentation of macular oct images using dual-scale gradient information," *Optics express*, vol. 18, no. 20, pp. 21 293–21 307, 2010.

[54] S. J. Chiu, X. T. Li, P. Nicholas, C. A. Toth, J. A. Izatt, and S. Farsiu, "Automatic segmentation of seven retinal layers in sdoct images congruent with expert manual segmentation," *Optics express*, vol. 18, no. 18, pp. 19 413–19 428, 2010.

[55] S. J. Chiu, J. A. Izatt, R. V. O'Connell, K. P. Winter, C. A. Toth, and S. Farsiu, "Validated automatic segmentation of amd pathology including drusen and geographic atrophy in sd-oct images," *Investigative ophthalmology & visual science*, vol. 53, no. 1, pp. 53–61, 2012.

[56] J. Tian, B. Varga, G. M. Somfai, W.-H. Lee, W. E. Smiddy, and D. C. DeBuc, "Real-time automatic segmentation of optical coherence tomography volume data of the macular region," *PloS one*, vol. 10, no. 8, p. e0133908, 2015.

[57] B. J. Antony, M. D. Abràmoff, M. M. Harper, W. Jeong, E. H. Sohn, Y. H. Kwon, R. Kardon, and M. K. Garvin, "A combined machine-learning and graph-based framework for the segmentation of retinal surfaces in sd-oct volumes," *Biomedical optics express*, vol. 4, no. 12, pp. 2712–2728, 2013.

[58] B. J. Antony, M. D. Abràmoff, M. Sonka, Y. H. Kwon, and M. K. Garvin, "Incorporation of texture-based features in optimal graph-theoretic approach with application to the 3d segmentation of intraretinal surfaces in sd-oct volumes," in *SPIE Medical Imaging*. International Society for Optics and Photonics, 2012, pp. 83 141G–83 141G.

[59] T. JM, S. A, K. J, R. RM, Q. HA, and J. J, "Racial variations in the prevalence of primary open-angle glaucoma: The baltimore eye survey," *JAMA*, vol. 266, no. 3, pp. 369–374, 1991.

[60] N. M. Bressler, "Age-related macular degeneration is the leading cause of blindness..." *Jama*, vol. 291, no. 15, pp. 1900–1901, 2004.

[61] S. Zayit-Soudry, I. Moroz, and A. Loewenstein, "Retinal pigment epithelial detachment," *Survey of ophthalmology*, vol. 52, no. 3, pp. 227–243, 2007.

[62] K. Lee, M. Niemeijer, M. K. Garvin, Y. H. Kwon, M. Sonka, and M. D. Abràmoff, "Segmentation of the optic disc in 3-d oct scans of the optic nerve head," *Medical Imaging, IEEE Transactions on*, vol. 29, no. 1, pp. 159–168, 2010.

[63] C. Shannon, "Communication in the presence of noise," *Proceedings of the IRE*, vol. 37, no. 1, pp. 10–21, Jan 1949.

[64] A. Trujillo-Pino, K. Krissian, M. Alemán-Flores, and D. Santana-Cedrés, "Accurate subpixel edge location based on partial area effect," *Image and Vision Computing*, vol. 31, no. 1, pp. 72–90, 2013.

[65] M. D. Abràmoff, X. Wu, K. Lee, and L. Tang, "Subvoxel accurate graph search using non-euclidean graph space," *PloS one*, vol. 9, no. 10, p. e107763, 2014.

[66] M. P. Kumar, O. Veksler, and P. H. Torr, "Improved moves for truncated convex models," *J. Mach. Learn. Res.*, vol. 12, pp. 31–67, Feb. 2011.

[67] H. Ishikawa, "Exact optimization for markov random fields with convex priors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, pp. 1333–1336, 2003.

[68] S. Farsiu, S. J. Chiu, R. V. O'Connell, F. A. Folgar, E. Yuan, J. A. Izatt, and C. A. Toth, "Quantitative classification of eyes with and without intermediate age-related macular degeneration using optical coherence tomography," *Ophthalmology*, vol. 121, no. 1, pp. 162–172, 2014.

[69] S. Mrejen *et al.*, "Multimodal imaging of pigment epithelial detachment: a guide to evaluation," *Retina*, vol. 33, no. 9, pp. 1735–1762, 2013.

[70] P. A. Keane, P. J. Patel, S. Liakopoulos, F. M. Heussen, S. R. Sadda, and A. Tufail, "Evaluation of age-related macular degeneration with optical coherence tomography," *Survey of ophthalmology*, vol. 57, no. 5, pp. 389–414, 2012.

[71] F. M. Penha, P. J. Rosenfeld, G. Gregori, M. Falcão, Z. Yehoshua, F. Wang, and W. J. Feuer, "Quantitative imaging of retinal pigment epithelial detachments using spectral-domain optical coherence tomography," *American journal of ophthalmology*, vol. 153, no. 3, pp. 515–523, 2012.

[72] G. Gregori, F. Wang, P. J. Rosenfeld, Z. Yehoshua, N. Z. Gregori, B. J. Lujan, C. A. Puliafito, and W. J. Feuer, "Spectral domain optical coherence tomography imaging of drusen in nonexudative age-related macular degeneration," *Ophthalmology*, vol. 118, no. 7, pp. 1373–1379, 2011.

[73] W. Ding, M. Young, S. Bourgault, S. Lee, D. A. Albiani, A. W. Kirker, F. Forooghian, M. V. Sarunic, A. B. Merkur, and M. F. Beg, "Automatic detection of subretinal fluid and sub-retinal pigment epithelium fluid in optical coherence tomography images," in *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE*. IEEE, 2013, pp. 7388–7391.

[74] G. Quellec, K. Lee, M. Dolejsi, M. K. Garvin, M. D. Abràmoff, and M. Sonka, "Three-dimensional analysis of retinal layer texture: identification of fluid-filled regions in sd-oct of the macula," *Medical Imaging, IEEE Transactions on*, vol. 29, no. 6, pp. 1321–1330, 2010.

[75] Z. Sun, H. Chen, F. Shi, L. Wang, W. Zhu, D. Xiang, C. Yan, L. Li, and X. Chen, "An automated framework for 3d serous pigment epithelium detachment segmentation in sd-oct images," *Scientific reports*, vol. 6, 2016.

[76] F. Shi, X. Chen, H. Zhao, W. Zhu, D. Xiang, E. Gao, M. Sonka, and H. Chen, "Automated 3-d retinal layer segmentation of macular optical coherence tomography images with serous pigment epithelial detachments," *Medical Imaging, IEEE Transactions on*, vol. 34, no. 2, pp. 441–452, 2015.

[77] A. Shah, J. Bai, Z. Hu, S. Sadda, and X. Wu, "Multiple surface segmentation using truncated convex priors," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015*. Springer, 2015, pp. 97–104.

[78] M. Soret, S. L. Bacharach, and I. Buvat, "Partial-volume effect in pet tumor imaging," *Journal of Nuclear Medicine*, vol. 48, no. 6, pp. 932–945, 2007.

[79] F. Malmberg, J. Lindblad, N. Sladoje, and I. Nyström, "A graph-based framework for sub-pixel image segmentation," *Theoretical Computer Science*, vol. 412, no. 15, pp. 1338–1349, 2011.

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

[81] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

[82] J. Frostegård, "Sle, atherosclerosis and cardiovascular disease," *Journal of internal medicine*, vol. 257, no. 6, pp. 485–495, 2005.

[83] S. Glagov, E. Weisenberg, C. K. Zarins, R. Stankunavicius, and G. J. Kolettis, "Compensatory enlargement of human atherosclerotic coronary arteries," *New England Journal of Medicine*, vol. 316, no. 22, pp. 1371–1375, 1987.

[84] S. Balocco, C. Gatta, F. Ciompi, A. Wahle, P. Radeva, S. Carlier, G. Unal, E. Sanidas, J. Mauri, X. Carillo *et al.*, "Standardized evaluation methodology and reference database for evaluating ivus image segmentation," *Computerized Medical Imaging and Graphics*, vol. 38, no. 2, pp. 70–90, 2014.

[85] G. Unal, S. Bucher, S. Carlier, G. Slabaugh, T. Fang, and K. Tanaka, "Shape-driven segmentation of the arterial wall in intravascular ultrasound images," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 12, no. 3, pp. 335–347, 2008.

[86] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," *International journal of computer vision*, vol. 22, no. 1, pp. 61–79, 1997.

[87] M.-H. R. Cardinal, J. Meunier, G. Soulez, R. L. Maurice, É. Therasse, and G. Cloutier, "Intravascular ultrasound image segmentation: a three-dimensional fast-marching method based on gray level distributions," *IEEE transactions on medical imaging*, vol. 25, no. 5, pp. 590–601, 2006.

[88] M.-H. R. Cardinal, G. Soulez, J.-C. Tardif, J. Meunier, and G. Cloutier, "Fast-marching segmentation of three-dimensional intravascular ultrasound images: a pre-and post-intervention study," *Medical physics*, vol. 37, no. 7, pp. 3633–3647, 2010.

[89] R. Downe, A. Wahle, T. Kovarnik, H. Skalicka, J. Lopez, J. Horak, and M. Sonka, "Segmentation of intravascular ultrasound images using graph search and a novel cost function," in *Proc. 2nd MICCAI workshop on computer vision for intravascular and intracardiac imaging*, 2008, pp. 71–9.

[90] C. Gatta, E. Puertas, and O. Pujol, "Multi-scale stacked sequential learning," *Pattern Recognition*, vol. 44, no. 10, pp. 2414–2426, 2011.

[91] F. Ciompi, O. Pujol, C. Gatta, M. Alberti, S. Balocco, X. Carrillo, J. Mauri-Ferre, and P. Radeva, "Holimab: A holistic approach for media–adventitia border detection in intravascular ultrasound," *Medical image analysis*, vol. 16, no. 6, pp. 1085–1100, 2012.

[92] E. G. Mendizabal-Ruiz, M. Rivera, and I. A. Kakadiaris, "Segmentation of the luminal border in intravascular ultrasound b-mode images using a probabilistic approach," *Medical image analysis*, vol. 17, no. 6, pp. 649–670, 2013.

[93] C. V. Bourantas, F. G. Kalatzis, M. I. Papafaklis, D. I. Fotiadis, A. C. Tweddel, I. C. Kourtis, C. S. Katsouras, and L. K. Michalis, "Angiocare: An automated system for fast three-dimensional coronary reconstruction by integrating angiographic and intracoronary ultrasound data," *Catheterization and Cardiovascular Interventions*, vol. 72, no. 2, pp. 166–175, 2008.

[94] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, no. 7, pp. 971–987, 2002.

[95] A. Shah, J.-K. Wang, M. K. Garvin, M. Sonka, and X. Wu, "Automated surface segmentation of internal limiting membrane in spectral-domain optical coherence tomography volumes with a deep cup using a 3-d range expansion approach," in *Biomedical Imaging (ISBI), 2014 IEEE 11th International Symposium on*. IEEE, 2014, pp. 1405–1408.

[96] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," in *European conference on computer vision*. Springer, 2014, pp. 818–833.

[97] 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*, vol. 27, no. 10, pp. 1495–1505, 2008.

[98] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, "Caffe: Convolutional architecture for fast feature embedding," *arXiv preprint arXiv:1408.5093*, 2014.

[99] N. Jain, S. Farsiu, A. A. Khanifar, S. Bearelly, R. T. Smith, J. A. Izatt, and C. A. Toth, "Quantitative comparison of drusen segmented on sd-oct versus drusen delineated on color fundus photographs," *Investigative ophthalmology & visual science*, vol. 51, no. 10, pp. 4875–4883, 2010.

[100] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2015, pp. 234–241.