ACTIVE GEOMETRIC MODEL: MULTI-COMPARTMENT MODEL-BASED SEGMENTATION & REGISTRATION

A Thesis

Submitted to the Faculty

of

Purdue University

by

Prateep Mukherjee

In Partial Fulfillment of the

Requirements for the Degree

of

Master of Science

August 2013

Purdue University

Indianapolis, Indiana

This work is dedicated to my family and friends.

ACKNOWLEDGMENTS

I am heartily thankful to my supervisor, Dr. Gavriil Tsechpenakis, whose encouragement, guidance and support from the initial to the final level enabled me to develop an understanding of the subject.

I also want to thank Dr. Rajeev Raje and Dr. Mihran Tuceryan for agreeing to be a part of my Thesis Committee.

Thank you to all my friends and well-wishers for their good wishes and support. And most importantly, I would like to thank my family for their unconditional love and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Mukherjee, Prateep. M.S., Purdue University, August 2013. Active Geometric Model: Multi-compartment Model-based Segmentation & Registration. Major Professor: Gavriil Tsechpenakis.

We present a novel, variational and statistical approach for model-based segmentation. Our model generalizes the Chan-Vese model, proposed for concurrent segmentation of multiple objects embedded in the same image domain. We also propose a novel shape descriptor, namely the *M*ulti-*C*ompartment *D*istance *F*unctions or *mcdf*. Our proposed framework for segmentation is two-fold: first, several training samples distributed across various classes are registered onto a common frame of reference; then, we use a variational method similar to Active Shape Models (or ASMs) to generate an *average shape model* and hence use the latter to partition new images. The key advantages of such a framework are: (i) landmark-free automated shape training; (ii) strict shape constrained model to fit test data. Our model can naturally deal with shapes of arbitrary dimension and topology(closed/open curves). We term our model *Active Geometric Model*, since it focuses on segmentation of geometric shapes. We demonstrate the power of the proposed framework in two important medical applications: one for morphology estimation of 3D Motor Neuron compartments, another for thickness estimation of Henle's Fiber Layer in the retina. We also compare the qualitative and quantitative performance of our method with that of several other state-of-the-art segmentation methods.

1 INTRODUCTION

In this thesis, we discuss two important problems in computer vision - concurrent shape/surface registration and segmentation of multiple objects in medical images. This is an important problem in different clinical and pre-clinical domains. We focus on two such applications. The first is estimating morphology of motor neurons (MN) acquired using confocal microscopy in *Drosophila* larva. Motor Neurons constitute the building blocks for motor circuits in the nervous system. Estimating Motor Neuron morphology is the first step to understand connectivity patterns of a normal motor circuit, as well as determining differences between neuronal morphologies in a wild-type and mutant brain. According to recent study [1], 12 different morphological types of MNs have been discovered so far, which can be further subdivided depending on morphological directions.(Fig. 1.2d) The problem of morphology estimation here is challenging due to diverse structural variations of neuronal compartments, namely soma, axon and dendrites, both in and across different classes. The second application aims at estimating Henle's Fiber Layer (HFL) thickness in *Directional Optical Coherence Tomography* (D-OCT). This is important for an optician to detect macular degeneration of the retina.

The challenging problem in both these applications is analyzing shape/surface of multiple interconnected compartments in an image. Our proposed model to solve such a problem is two-fold. First, we align images acquired from different viewpoints or sensors, onto a common frame of reference. This is a classical example of *image registration*. Shape/surface registration is an interesting problem that manifests itself in a number of applications, such as pattern recognition, and image processing. Essentially such a problem consists of aligning two anatomical structures, which are represented as curve, or surface. While it is a difficult problem to solve in general,

for medical images,in particular, it often poses unique challenges, as a result of poor image contrast, speckle noise and highly varying/asymmetric shapes. For example, in many clinical scenarios, images from several modalities may be acquired and the diagnostician's task is to combine or extract this information to draw meaningful conclusions. Under such conditions, need for automated tools becomes necessary in statistical modeling and characterization of samples. Note that if prior knowledge of data is known, surface alignment can be done using *expectation maximization*. In practice however, such prior knowledge of data is generally not available and therefore more advanced sample-driven implicit techniques need to be introduced.

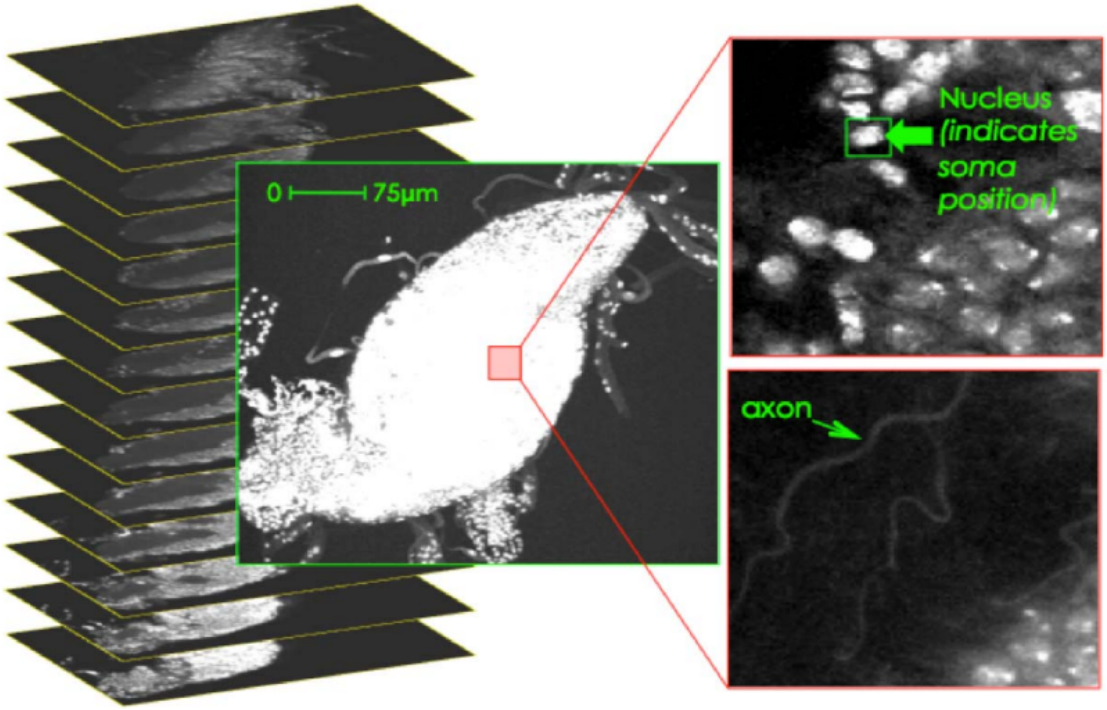
In the second step of our model, we capture the most important modes of variations for each compartment to generate an *average shape model*. The latter is used as a template to segment different compartments in a new image. This constitutes the method of *image segmentation*. Segmentation is a method to extract meaningful areas (or regions) from a given image. Majority of segmentation methods, in literature, is focused on the segmentation of a *single* object (e.g., into foreground and background regions) in an image. In many applications, however, like MRI-CT diagnosis and cell microscopy, concurrent analysis of multiple anatomical structures (e.g. gray matter and white matter) is essential for clinical treatment. This introduces the problem of *multi-compartment image segmentation*. Chan and Vese [2] in 2002, first proposed a multi-phase model for segmentation. Multi-compartment models have since then been used in many medical applications for concurrent segmentation of anatomical structures of interest. Most of the well-known methods in this field use level sets [3] for font propagation. Level sets provide a variational framework to represent contours of multiple objects in the same image domain. The basic idea of the level set method is to represent a contour as the zero level set of a higher dimensional function, called a level set function (LSF), and formulate the motion of the contour as the evolution of the level function. Muti-compartment models based on level sets use multiple LSFs to represent contours of multiple objects. However, this becomes computationally

expensive as the number of objects of interest in an image increases. Also, coupling of multiple level sets allow for overlaps and gaps in image domain, thus destroying topological consistency.

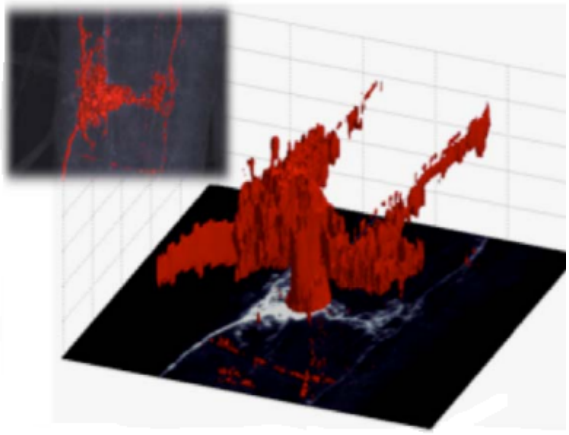
In this thesis, we introduce a novel multi-compartment method for automated registration and segmentation of medical images. Our method improves computational efficiency as well as preserves topology. We apply our method to 3D and 2D medical images. At the end of this chapter, we highlight the contributions made in this work. In chapter 2, we provide an overview of the related work in this area that are related to the applications treated in this thesis. In chapter 3, we present a novel method to represent shapes of anatomical structures. In chapter 4, we propose a variational method, using our novel shape descriptors, to align different images. In chapter 5, we discuss the second step of our framework to concurrently segment multiple objects in an image. Finally, in chapter 6, we discuss the data acquisition techniques for our applications, as well as provide detailed qualitative and quantitative evaluation of our method. In the rest of this chapter, we describe our two applications in brief and throw light on the need for precise registration and segmentation of multiple structures to derive meaningful conclusions.

1.1 Motor Neuron Morphology Estimation

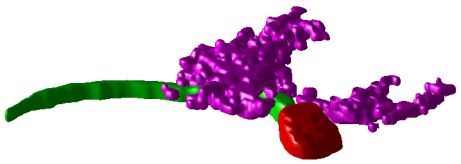
Normal locomotive behavior is fundamentally determined by the precise patterns of motor neuron(MN) connectivity that are dictated by the selective connection of motor axons with muscle targets as well as specific dendritic input from presynaptic neurons. (Fig. 1.1) In the vertebrate spinal cord, Motor Neurons are organized into columns, groups of MNs that target individual muscles are clustered into Motor Neuron pools [4]. Different MN pools elaborate distinct morphology patterns and respond to sensory stimulation with different latencies, demonstrating that the selectivity of synaptic input is directly influenced by the differential patterning of Motor Neuron



(a) Example of the image stacks used to detect and trace morphologically pre-determined motor neurons in part of the *Drosophila* CNS. Left: representative slices of the stack. Middle: the collapsed stack image, i.e. the intensity sum along z-axis. Right: two magnified regions chosen from different depths at the same x-y location, highlighted in red in the collapsed stack image



(b) Data(green channel) in 3D, after basic preprocessing



(c) Desired neuron partitioning into compartments: soma(red),axon(green) and dendrites(magenta)

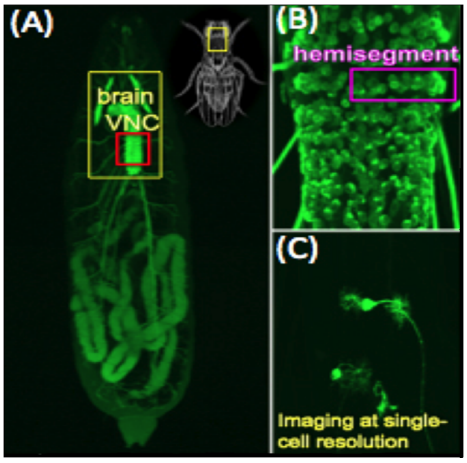
Figure 1.1. Global neuron morphology estimation

structures in the spinal cord [5]. Therefore, determining how different MN subtypes pattern and organize their morphology is a crucial step towards understanding how motor circuits are assembled to control locomotion.

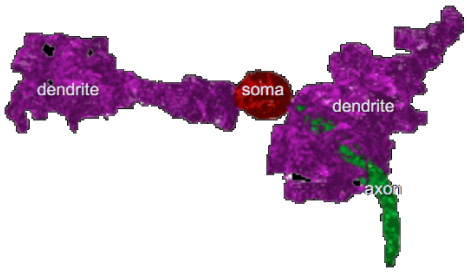
The Drosophila embryonic central nervous system comprises the developing brain and ventral nerve cord(VNC). (Fig. 1.2) The VNC, which can be considered functionally analogous to the mammalian spinal cord, is segmentally reiterated and bilaterally symmetrical with respect to the ventral midline. There are approximately 400 neurons, including an estimated 38 Motor Neurons, within each hemisegment(or half-segment) of the VNC. (Fig. 1.2a) Abdominal hemisegments in the embryo comprise 30 stereotyped body wall muscles, each of which is innervated by one of more of the 38 Motor Neurons. The muscle innervation pattern of individual Motor Neurons is further highly stereotyped making embryonic Motor Neurons uniquely identifiable [6]. During larval development, dendrites of these abdominal Motor Neurons undergo considerable growth and dendritic branching is dramatically increased, likely reflecting extensive changes in synaptic connectivity that are required for more complex larval behaviors such as peristaltic movements required for normal locomotion.

According to the study in [1], the morphological features that uniquely describe individual Motor Neuron subtypes are: relative position between the soma and the CNS center; relative position between the axon and the CNS center; direction and extent of the axon; position of the dendrites along the axon; relative positions of the dendrites and the soma; extent of the dendrites. To calculate these features, and apart from the segmentation of the neuron volume from its surroundings, we need to estimate the positions and shapes of the individual compartments.

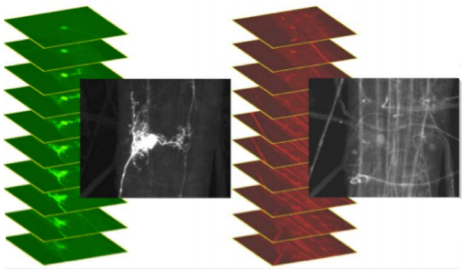
In the next section, we describe another such application, in the domain of optical coherence tomography, where estimation of morphology of anatomical structures is



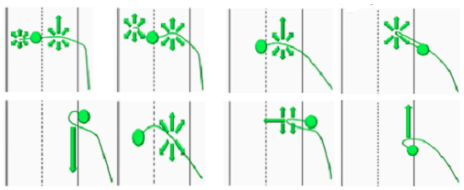
(a) **(A)** Whole larva with expressed with GFP, also showing Central Nervous System(CNS), i.e. brain and Ventral Nerve Cord(VNC). **(B)** Magnification of the VNC in the red box in (A). **(C)** Single MN clones.



(b) Neuron volume along with labeled compartments



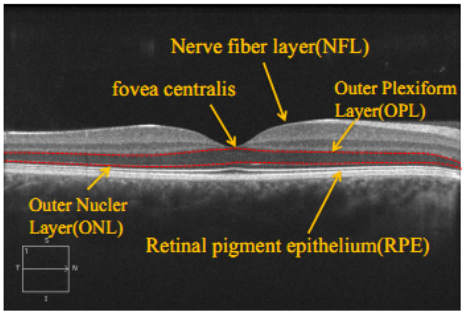
(c) Image stacks, one for each channel(green:neuron images; red:CNS reference images



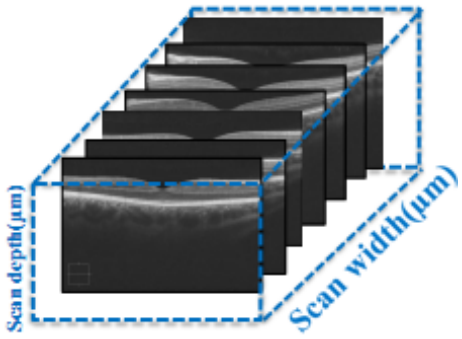
(d) MN morphologies, shown in 2D ground-truth sketches: the arrows indicate the dendrite positions and extends, the disk illustrates the soma, the curved lines show the axons, while the dotted lines represent the VNC midline.

Figure 1.2. Image datasets for modeling individual neurons

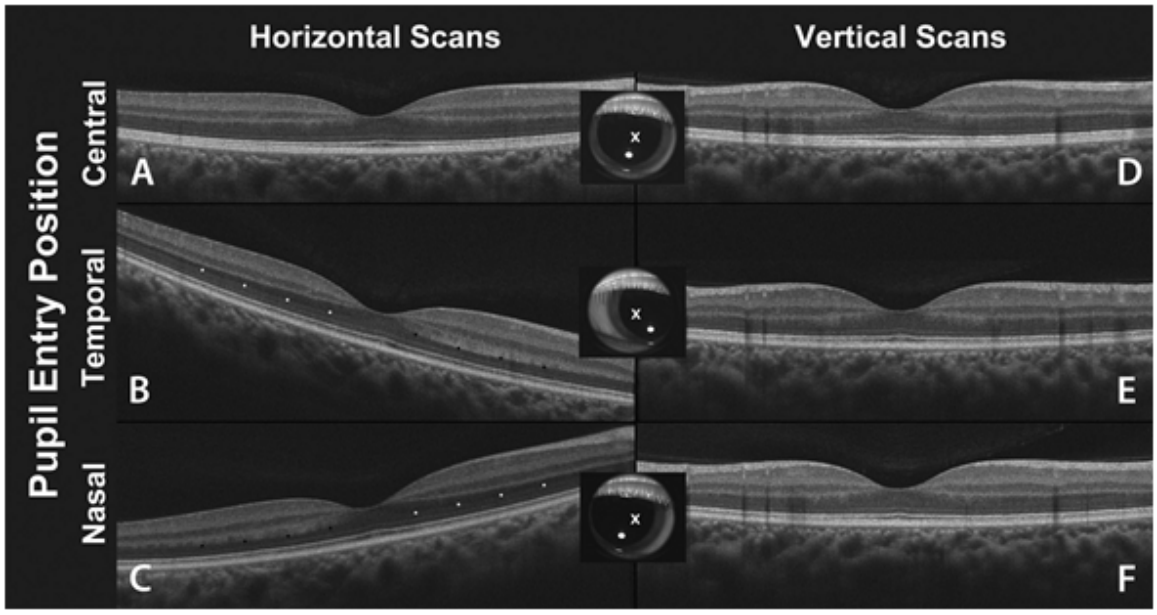
essential for clinical diagnosis. The next application is 2D, that is, it deals with segmenting a layer in a planar image.



(a) Retinal layer with labeled parts



(b) OCT B-scan images showing direction of scanning



(c) Directed OCT images **(A)** Central entry pupil position; **(B)** Temporal displacement of entry beam; **(C)** Nasal displacement of the entry beam; **(D-F)** Vertical scans at each pupil entry position.

Figure 1.3. Image datasets for D-OCT

1.2 Henle's Fiber Layer Thickness Estimation in OCT

Spectral domain optical coherence tomography(SD-OCT) is an ubiquitous technique in opthalmology that utilizes a broadband infrared light source and interferometry to generate depth-resolved reflectivity profiles. By analyzing interference fringes using a reverse Fourier transformation, anatomical details can be acquired with a resolution

of 5 μm. Utilizing horizontal and vertical scanners, cross-sectional and volumetric reconstructions of the retina can be ascertained.

Directional OCT(D-OCT) modifies SD-OCT by acquiring images from multiple pupil positions. The angle of incidence of the imaging beam on the retina is modified to at least three different positions using this technique - consequently tissues that demonstrate reflectivity will only appear hyper-reflective when illuminated from certain positions. Because it is comprised of thin microtubules, one of the directionally reflective tissues in the retina is *Henle's Fiber Layer*(HFL). During standard SD-OCT acquisition, HFL is not routinely discernible as a distinct layer and blends into the adjacent retinal layers.

D-OCT allows the distinct visualization of HFL, and consequently allows the underlying outer nuclear layer(ONL), the retinal layer that contains photoreceptor nuclei to become visible and quantifiable. ONL thining is responsible for vision loss in a number of retinal degenerations including age-related macular degeneration(AMD) [7], the most common cause of irreversible vision loss in the United States and Europe. D-OCT imaging of the retina therefore could provide to direct ONL loss, and could provide a predictor of AMD and a biomarker for its progression.

When applying D-OCT for *in vivo* imaging of the human retina, small movements of the head and saccadic eye movements leads to small axial movements of the retina relative to the imaging geometry during image recording. This behavior will in reality lead to decorrelation of the speckle noise from one image to the next when recording a series of B-scans. The B-scans obtained from different positions of the pupil result in central, temporal and nasal alignment of the retinal wall (Fig. 1.3c). Diagnosis of the pathological subjects requires a proper alignment of the different B-scans onto a common frame of reference.

1.3 Thesis Contributions

Three main contributions are made in this thesis:

1. *Multi-Compartment Shape Representation:* A novel model for representing contours of multiple objects of interest in the same image domain is developed. Our proposed shape descriptors, called *multi-compartment distance functions*(MCDF) has three main advantages. First, it uses the powerful level set function for representing boundaries, which provide flexible topological changes and yield contours with no self-intersections. Second, we reduce the computational cost of traditional multi-phase level set methods by decomposing them into fewer level sets. Third, our shape descriptors are invariant to rotation and translation, which makes them ideal for global affine motion correction.

2. *Landmark Free Framework for Training:* Our proposed MCDFs make our training phase completely free of landmarks. This is an important step towards automated feature selection and training. Moreover, our model requires less expert interveneince which significantly saves resources. MCDFs also make our model less prone to human errors and thus provides a robust framework for generating *average shape models*.

3. *ASM with Variational Shape Constraints:* We improve the performance of traditional ASMs by incorporating variational shape constraints. Important advantages of such a model over existing methods are that it preserves topology in the image domain and that the fitted model do not allow overlaps or vacuums.

2 PREVIOUS WORKS

This thesis focusses on two main areas of computer vision - *Registration* and *Segmentation*. In sections 2.1 and 2.2, we describe drawbacks of some of the existing methods used for registration and segmentation in biomedical images. In sections 2.3 and 2.4, we analyze the state-of-the-art methods in MN morphology and HFL thickness estimation. We also throw light some of the drawbacks of these methods, as well as propose possible solutions.

2.1 Image Registration

At its simplest, image registration involves estimating a mapping between a pair of images. One image is assumed to be stationary, called the reference($\mathcal{I}_T$) image, whereas the other, called the source($\mathcal{I}_S$) image, is spatially transformed to match it. In order to transform the source to match the reference, it is necessary to determine a mapping from each voxel position in the reference to a corresponding position in the source. Mathematically, this can be stated as $\mathcal{T}^* = \text{argmin}_{\mathcal{T}}\, C(\mathcal{I}_T, \mathcal{I}_S \circ \mathcal{T}) + \alpha S(\mathcal{T})$, where α is a constant governing the strength of the penalty. The most important criteria of modeling such a problem relies on efficient and robust representation of shape features and choosing a suitable optimization procedure to perform the registration. Such a framework consists of following components (i) shape representation, (ii) transformation and (iii) registration criterion.

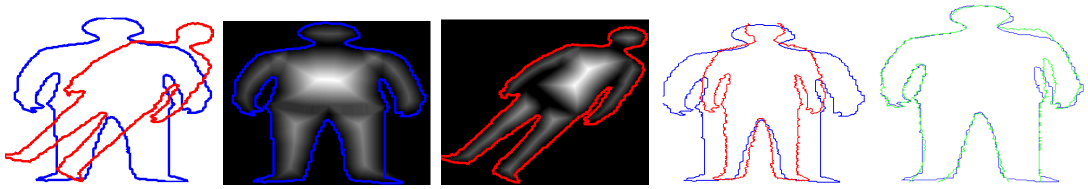
Shape representation is the most important component for registering geometric shapes. The use of point-based snakes [8], Fourier descriptors, level set representations [3] are some well-known approaches to represent shapes as well as their variations. Point clouds [9] are the simplest shape representations used, but they suffer

from being sensitive to the selection of the number of points. In addition, they require a large number of parameters to deal with shape deformations. Interpolation techniques like splines [10], triangulated surfaces [11], deformable models [12] attempt to address these issues through a parametric representation of the shapes or surfaces. In addition, Medial axis [13] and more recently, level-set functions [3] adopt the idea of representing shapes using distance transforms.

Transformation refers to the selected global, local or hierarchical (global-to-local) fitting. Global transformation models apply to an entire shape; these type of transformations are either rigid, similarity-based, affine or perspective. Local transformation models, on the other hand, represent point-wise deformations that deform a shape locally and non-rigidly. A detailed discussion of these is provided in Chapter 4. Thin Plate Splines(TPS), Iterative Closest-Point(ICP) [14], and optical flow [15] are some examples of local transformations. Hierarchical models are also popular since they cover the entire transformation domain using both global and local transformations.

Registration Criterion is the approach used to recover the optimal transformation parameters given a shape representation and a transformation model. Popular approaches in this area can be classified into two sub-categories. The first is to establish geometric feature correspondences and then estimate the transformation parameters using the correspondences [16]. The second is to recover the optimal transformation parameters through optimization of energy functionals [17].

In this thesis, we propose a shape registration method in a variational framework. Our overall approach is depicted in Fig. 2.1. The implicit shape representation is generic, that is it can handle naturally shapes of arbitrary dimension and topology, and can also be used in statistical shape modeling. The representation is also stable and robust to shape perturbations and noise.



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

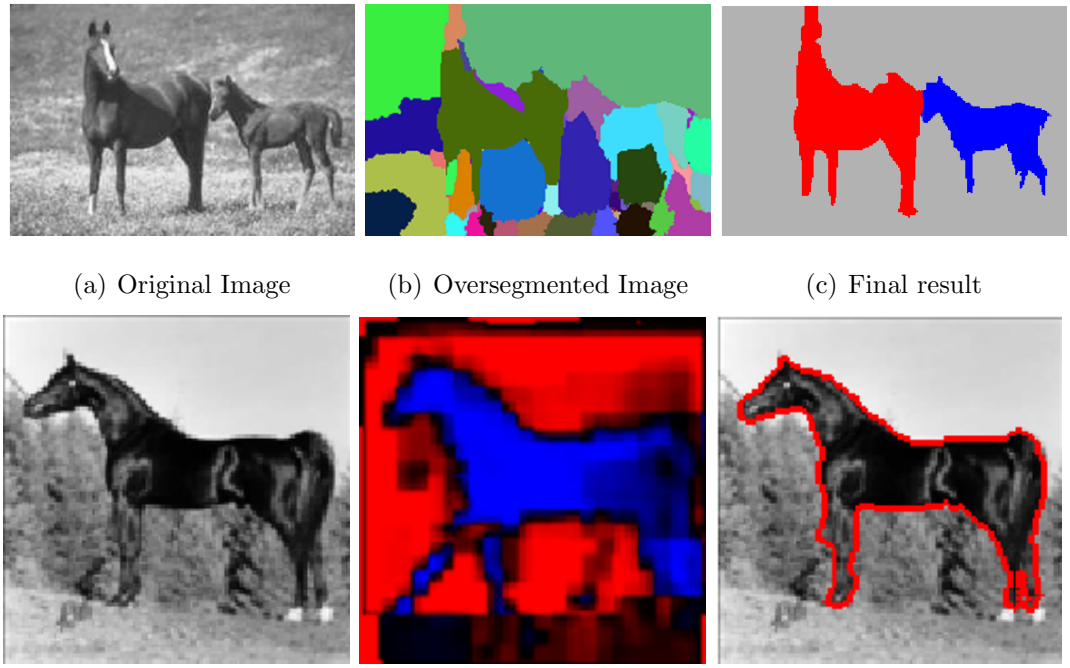
Figure 2.1. (a) Initial condition($\mathcal{I}_T$ in blue, $\mathcal{I}_S$ in red); (b) Distance map of $\mathcal{I}_T$; (c) Distance map of $\mathcal{I}_S$; (d) Global alignment, only zero level sets shown; (e) Result after deformable registration; transformed source shape(in green) overlaid on the target shape(in blue).

In our approach, the optimal global and local transformation parameters are recovered by optimizing Sum-of-Squared-Differences(SSD) criterion. This choice is primarily because the many degrees of freedom in local deformations require input equality to achieve a global optimum. Other suitable criteia might be Normalized Cross-Correlation(NCC) and Mutual Information(MI). Our method is powerful to adapt to any of these optimization criterion, and the choice is driven by the application.

2.2 Image Segmentation

Image Segmentation is the method of clustering pixels into salient image regions, i.e.,regions corresponding to individual surfaces, objects, or natural parts of objects. Segmentation has a variety of interpretations that depends on each application. For example, given an image and some clues of its contents, the goal of segmentation is to decompose the image into superpixels, which are roughly coherent in color and texture. The application of fitting lines to edge points is also a segmentation problem, wherein the goal is to organize a some tokens that belong together because they fit a line. Cosegmentation [18] is another useful application of segmentation methods, where the goal is to find a fundamental matrix to a set of feature points on images which share a common foreground. A property ubiquitous to all of these applications is that pixels, tokens, super-pixels etc that belong together conform to the same

model. Segmentation methods are broadly divided into two categories (i) Bottom-up, and (i) Top-down.



(a) Original Image (b) Oversegmented Image (c) Final result

(d) Sample Training Image (e) Average Shape Model (f) Model fitting

Figure 2.2. Top Row: Bottom-up segmentation; Bottom Row: Top-down segmentation.

Bottom-up segmentation approaches use different image-based criteria and search algorithms to find heterogeneous segments within the image. A common bottom-up approach is to use a graph representation of the image (with the nodes representing pixels or super-pixels) and partition the graph into subsets corresponding to salient image regions. The resulting segmentation takes into account texture, average intensity and boundary properties if image-regions. The segmentation applies successive, recursive coarsening, in which homogeneous segments at a given level are used to form larger homogeneous segments at the next level. In this manner, the image is segmented into fewer and fewer segments. The algorithm also provides a measure of saliency that ranks segments according to their distinctiveness. Uniform segments

that contrast with their surrounding will be highly salient, and will therefore have a very low energy. Examples of bottom-up segmentation are Normalized cuts, Graph-cuts, Mean-shift methods [19]. An example of bottom-up segmentation is shown in Fig. 2.2, where the segmentation result looks quite good.

Top-down segmentation approaches rely on acquired class-specific information, and can only be applied to images from a specific class. These include deformable templates, active shape models(ASM) [20] and active contours [8]. The high variability of shape and appearance of objects can be dealt with by working with image patches(or fragments). These patches are used as shape primitives for the class. The approach can be divided into two stages - training and fitting. In the training stage, a set of informative image patches is constructed from training data to capture possible variations in the shape and appearance of common object parts of a given class. A set of *average features* (i.e., features that are highly likely to be detected in class images compared with non-class images) are derived from this training set. These are used in the segmentation stage to classify a novel input image as well as detect the approximate location and scaling of the corresponding objects.

From the above categorization, it is intuitive that top-down segmentation works better in case of large variations in topology or morphology of the objects. The drawback of top-down methods are that they require considerable manual labor to annotate *landmarks* or distinctive features of objects in training data set. Our work is motivated in alleviating this problem and propose a variational framework for this purpose.

2.3 Related Work in MN Morphology Estimation

In this section, we focus in describing existing work in estimating neuronal structures. To better understand the context of this work, first we introduce the technique to obtain neurons at single-cell resolution.

Previous efforts to visualize morphology of single MNs have relied on dye-backfilling methods such as DiI. DiI is a lipophilic dye that is taken by the axon and diffuses along the cell membrane to reveal neuronal morphology. However, this technique is limited in that neurons can only be labeled one at a time. Furthermore, dye-labeling methods are not compatible with long-term sample preservation, limiting the use of secondary markers. To overcome these limitations, we use mosaic analysis with a repressible cell marker (MARCM) [21], a genetic technique that allows us to label and image individual MNs (Fig. 1.1b). The use of fluorescent proteins used in dying neurons allows for high-resolution in vivo imaging with minimal photobleaching, reduced phototoxicity and enhanced labeling of the neuronal membrane. We discuss the details of MARCM technique in Chapter 6.

In the last decade there has been an increasing interest in the problem of segmenting tubular and tree-like structures such as the neuron dendrites. For example, Zhou et al. [22] use a level set-based approach to segment precisely dendrites from 2-photon microscopy images. Peng et al. [23] reconstruct neurites (axons and dendrites) with fuzzy boundaries using a region-based deformable model initialized by detecting neurite regions and connecting them in a shortest-path manner. In the work of [24], the tubular-like structures are segmented with shape-constrained geodesic active contours, after braching points are manually annotated in dendrites imaged with confocal microscopy, while in [25] a method for automated branching points detection is presented. In [26], the tree structure of the dendrites is estimated from multi-photon microscopic images, using skeletonization and splines. Dima et.al [27] use wavelets to trace the branches boundaries, while Uehara et. al [28] use confocal and 2-photon

microscopy images to show a shape-driven approach where the dendrite branches are approximated with cylinders that serve as topology constraints for a wave propagation method. In the work of [29], dendrite structures are estimated by detecting their skeletons and then applying intensity-based fuzzy c-means clustering, while in [30] the detected skeletons are used in a graph-theoretic approach to provide fine tracing of the dendrite branches. In the recent works of Kaynig et.al [31] neuron boundaries are estimated from electron microscopy data with an energy minimization using graph cuts. Finally, it is worth noting the existence of software packages, such as Neurolucida[1], used in [32] for tracing dendritic structures and calculating their morphology statistics, and NeuronStudio[2].

2.4 Related Work in HFL Thickness Estimation

Images of optical tomography modality are obtained from both healthy and diseased subjects. The latter suffer from drusen related to non-exudative age-related macular degeneration(AMD). All eyes are scanned with the Cirrus OCT instrument by a single experienced technician. Horizontal frame-averaged B-scans acquired with Cirrus at the two extreme horizontal entrance positions are used to measure the contribution of different zones of reflectivity between the external limiting membrane(ELM) and the inner plexiform layer(IPL) in five right eyes. (Fig. 1.3a) Specifically, the side of B-scan that allowed full visualization of the transition to the outer plexiform layer (OPL) is used for this analysis where the inner segment/outer segment (IS/OS) was 1 mm from the foveal center. The distance from the ELM to the posterior edge of the OPL and the distance from the ELM to the edge of the observed hyporeflective zone within this layer is measured along a line perpendicular to IS/OS. The relative contribution of this hyporeflective zone corresponding to HFL is reported as a percentage of the total distance between the ELM and the OPL. Finally, frame-averaged

[1]MBF Bioscience `http://www.mbfscience.com/neurolucida`

[2]Computational Neurobiology and Imaging Center, Mount Sinai School of Medicie, NY: `http://research.mssm.edu/cnic/tools-ns.html`

B-scans obtained from each subject are exported to ImageJ, where a manual tool is used to segment the images collected at each of three pupil entry positions (nasal, central, temporal). Segmented layer data is then processed through custom computing software to determine the thickness between the ELM and the outer edge of the OPL.

In this thesis, we apply our model to the above two applications. Our model is a variational framework built over the Active Shape Models, introduced by Cootes et al [20]. Our goal in the former is to provide a computationally comprehensive description for the entire structure of larval motor neurons, with simultaneous segmentation and labeling of the individual morphological compartments. In the latter application, we train our model on retinal D-OCT images and then build a model to segment the HFL layer. This provides an automatic method for measuring HFL thickness.

3 SHAPE REPRESENTATION

Advances in medical imaging technology have provided the ability to acquire high resolution 3D medical images. Image analysis techniques provide more advanced visualizations and aid in disease diagnosis. For example, automatic extraction of anatomical geometry from a medical image is useful in planning radiation beam therapy to apply maximum radiation close to a tumor while minimizing exposure to surrounding organs. Analysis of diffusion tensor magnetic resonance images of neonatal brains can give information about the early stages of development in brain connectivity.

These examples benefit from a particular tool from medical image analysis known as *statistical shape analysis* which describes the geometric variability of anatomy. Most approaches in shape analysis have used either used linear parametric models of anatomic shape, and thus, linear statistical techniques to analyze the shape variability. Parametric models suffer from being robust to high degree of variability in shapes, like bending and twisting. On the other hand, richer models of shape and richer variations of shape can be achieved with nonlinear and/or nonparametric models. For example, level-set distance functions [3] have shown great promise in representing the interior of anatomic structures and describing shape changes in intuitive terms such as bending.

3.1 The Level Set Method

Level set methods are based in the context of fluid mechanics and provide both a nice framework and efficient practical tools for solving Partial Differential Equations(PDE's). The idea of level-set evolutions is as follows. We consider a family of hypersurfaces $\mathcal{S}(\mathbf{p}, t)$ in $\Re^3$, where $\mathbf{p}$ parametrizes the surface and t is the time, that evolve according to the following PDE:

$$\frac{\partial \mathcal{S}}{\partial t} = \beta \mathcal{N} \tag{3.1}$$

with initial condition $\mathcal{S}(t=0) = \mathcal{S}_0$, where $\mathcal{N}$ is the inward unit normal vector of $\mathcal{S}$, β is a velocity function and $\mathcal{S}_0$ is some initial closed surface. They key idea in [3] is to introduce the function $u : \Re^3 \times \Re \to \Re$ such that

$$u(\mathcal{S}, t) = 0 \quad \forall t$$

By differentiation (and along with $\mathcal{N} = -\frac{\nabla u}{|\nabla u|}$ and Eq. (3.1), we obtain the Hamilton-Jacobi equation:

$$\frac{\partial u}{\partial t} = \beta \left| \nabla u \right|$$

with initial conditions $u(.,0) = u_0(.)$, where u_0 is some initial function $\Re^3 \to \Re$ such that $u_0(\mathcal{S}_0) = 0$.

3.2 The Multi-phase Level-set Method

In numerous practical applications, images consist of numerous objects of interest. For example, in CT and MRI diagnosis, it is necessary to accurately segment white matter and grey matter simultaneously. Using multiple level set functions, multi-phase deformable models can segment multiple objects simultaneously [2, 33]. There are largely two kinds of approaches for multi-phase models. One is to associate each region with a level-set function and add some constraints [33]. The other is to represent regions by the combination of several level set functions [2]. The latter model permits multiple object boundaries, and guarantees no overlaps or vacuums. Moreover, it also substantially reduces the computational cost as the number of objects or compartments grow. However, this approach has two key limitations. Firstly, image-based forces in the multiphase framework are applied to the level set functions rather than to the objects themselves. Thus, it is possible that while the lengths of the level set functions are minimized, it may not be so for the boundaries of the

objects themselves. Secondly, the evolution/optimization can get "stuck" in situations where a pixel needs to acquire a label that can be reached only by changing two level set functions at the same time. The existing evolution strategy cannot resolve these situations, which are also commonly found with an increasing number of objects.

In this thesis, we present a multiple object geometry model that (1) guarantees no overlaps or gaps, (2) includes only a few level-set functions independently of the number of regions, (3) can enforce relationships and topological constraints on any or all objects and groups of objects if desired, and (4) the framework can easily be scaled to higher dimensions (4D and so on) without significant changes to the underlying model. In this framework, the evolution of the level set functions representing any number of objects or compartments is recast into the evolution of a fixed, small number of distance functions and an equal number of corresponding label functions.

3.3 Definitions and Related Notations

Let I be an image defined on a domain Ω, $dim(\Omega) = d$ where d = 2 for 2D and d = 3 for 3D. We consider N objects $O_1, O_2, \cdots O_N$, each containing points from the domain $\mathbf{x} \in \Omega$ such that these objects cover the whole domain with no overlaps or vacuums. Formally, $\cup_{i=1}^{N} O_i = \Omega$ and $O_i \cap O_j = \emptyset, \forall i \neq j$.

Object *signed distance functions*, denoted $\mathbf{\Phi}_i$ are commonly used in the level set literature for their desirable numerical properties. These functions are negative inside their respective objects, positive outside, and give the distance to the object boundary at every point $\mathbf{x}$. So, if $E(\mathbf{x}, O_i) = min_{\mathbf{y} \in O_i} \parallel \mathbf{x} - \mathbf{y} \parallel$, then

$$\mathbf{\Phi}_i(\mathbf{x}) = \begin{cases} 0, & \mathbf{x} \in O_i \\ +E(\mathbf{x}, O_i), & \mathbf{x} \in R(O_i) \\ -E(\mathbf{x}, O_i), & \mathbf{x} \in [\Omega - R(O_i)] \end{cases} \tag{3.2}$$

Label function Next, we define the set of *label* functions that describes the local configuration of neighboring objects at $\mathbf{x}$ as follows:

$$\begin{aligned}
\forall \mathbf{x}, \quad & L_0(\mathbf{x}) = i \quad iff \quad \mathbf{\Phi}_1(\mathbf{x}) > 0, \\
& L_1(\mathbf{x}) = \underset{j \neq L_0(\mathbf{x})}{\operatorname{argmin}} \mathbf{\Phi}_j(\mathbf{x}), \\
& L_2(\mathbf{x}) = \underset{j \neq L_0(\mathbf{x}), L_1(\mathbf{x})}{\operatorname{argmin}} \mathbf{\Phi}_j(\mathbf{x}), \\
& \vdots \\
& L_{N-1}(\mathbf{x}) = \underset{j \neq L_k(\mathbf{x}), k=0,\ldots N-2}{\operatorname{argmin}} \mathbf{\Phi}_j(\mathbf{x})
\end{aligned} \tag{3.3}$$

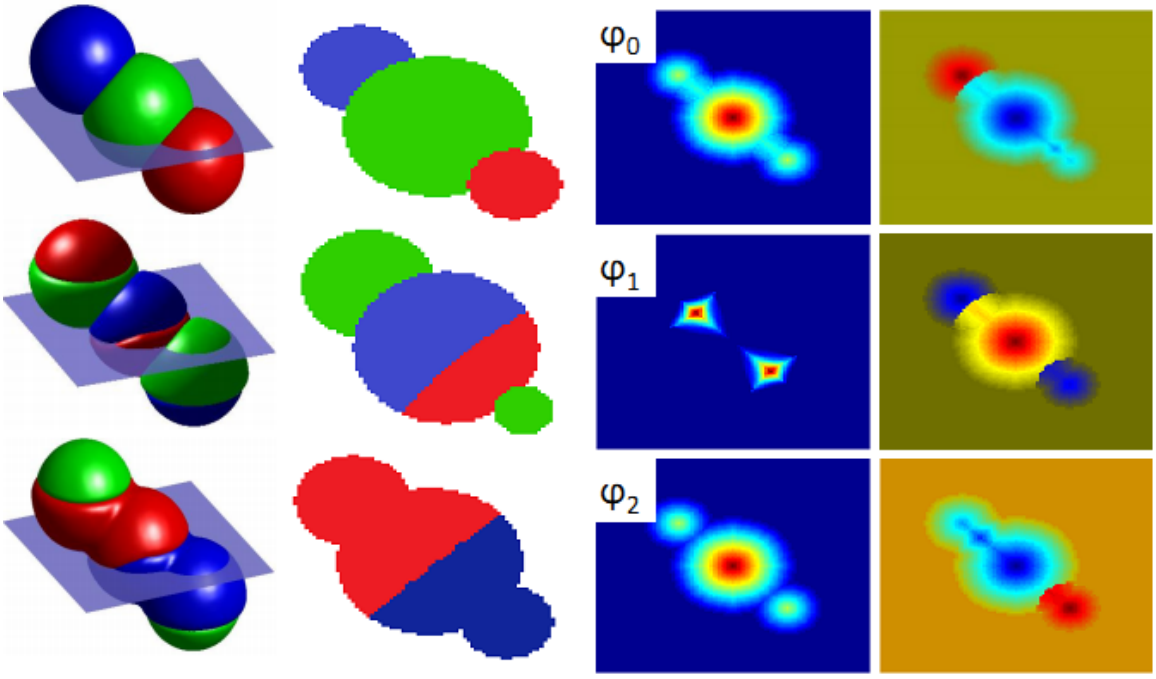
The label functions $L_0, L_1, \cdots L_{N-1}$ give a detailed description of the configuration of the objects. In particular, L_0 is the zero-level label function and the first-level label L_1 identifies the closest neighboring object at each point. The ϕ_j function gives the distance to the boundary of O_j, and hence, minimizing over j yields the nearest object. More generally, $L_k(\mathbf{x}) = i, 1 \leq k \leq N-1$, if and only if O_i is the kth closest neighbor to $\mathbf{x}$.

Distance function Having defined the *label* functions for each object $O_i \in \Omega$, let us define the N distance functions, or zero-order level set functions for each compartment.

$$\begin{aligned}
\forall \mathbf{x}, \quad & \varphi_0(\mathbf{x}) = \mathbf{\Phi}_{L_1}(\mathbf{x}), \\
& \varphi_1(\mathbf{x}) = \mathbf{\Phi}_{L_2}(\mathbf{x}) - \mathbf{\Phi}_{L_1}(\mathbf{x}), \\
& \varphi_2(\mathbf{x}) = \mathbf{\Phi}_{L_3}(\mathbf{x}) - \mathbf{\Phi}_{L_2}(\mathbf{x}), \\
& \vdots \\
& \varphi_{N-2}(\mathbf{x}) = \mathbf{\Phi}_{L_{N-1}}(\mathbf{x}) - \mathbf{\Phi}_{L_{N-2}}(\mathbf{x}),
\end{aligned} \tag{3.4}$$

Here, the particular level set functions that are used in the right hand side of this decomposition are determined by the labels $L_1, L_2, \cdots L_{N-1}$. These distance functions

specify the additional distances one must travel to reach the succession of next-closest neighbors. For example, φ_0 is the distance from $\mathbf{x}$ to its first neighbor and φ_1 is the additional distance that must be traveled to get to the second neighbor, and so on.



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

Figure 3.1. Label and Distance functions for a "toy" example (colors indicate compartment labels) Panels (a-c): top row-L_0, middle row-L_1, bottom row-L_2. (a) Color-labeled compartments(green, red and blue), (b) a cross-section (plane indicated in (a)) of their label functions, and (c) corresponding cross-sections of their distance functionsφ_0, φ_1, φ_2; (d) *mcdf*s for each compartment $i = 1, 2, 3$, from top to bottom respectively.

Fig. 3.1(a-c) shows an example of this label-distance decomposition on a "toy" image. Note that the 3 connected objects are shown in 3D, the label and corresponding distance functions are shown on a horizontal slice from the 3D volume.

Multi-compartment Distance Functions(MCDF) Given all the label and distance functions at a point $\mathbf{x}$, the signed distance function of object O_i can be recovered as follows:

$$\phi_i(\mathbf{x}) = \begin{cases} \varphi_0(\mathbf{x}), & i = L_0(\mathbf{x}) \\ -\varphi_0(\mathbf{x}), & i = L_1(\mathbf{x}) \\ -\varphi_0(\mathbf{x}) - \varphi_1(\mathbf{x}), & i = L_2(\mathbf{x}) \\ -\varphi_0(\mathbf{x}) - \varphi_1(\mathbf{x}) - \varphi_2(\mathbf{x}), & i = L_3(\mathbf{x}) \\ \vdots & \\ \sum_{j=0}^{N-2} -\varphi_j(\mathbf{x}), & i = L_{N-1}(\mathbf{x}) \end{cases} \tag{3.5}$$

It is seen here that in order to recover all the signed distance functions, we need all the label and distance functions. Therefore, so far, there is apparently no advantage to the proposed decomposition, as it does not appear to lead to a compact representation or lower complexity.

However, a key observation here is that only signed distance values near object boundaries are required to accurately carry out geometric level set computations. In fact, topology at $\mathbf{x}$ should be derived from object L_0 to which $\mathbf{x}$ is assigned and the object L_1 that is closest to $\mathbf{x}$. Therefore, the higher-order terms in (3.5) can be removed in order to compute geometric deformations. Accordingly, we simply drop the higher-order terms in (3.5) to get the following approximate signed distance functions:
2D:

$$\phi_i(\mathbf{x}) = \begin{cases} \varphi_0(\mathbf{x}), & i = L_0(\mathbf{x}), \\ -\varphi_0(\mathbf{x}), & i = L_1(\mathbf{x}), \\ -\varphi_0(\mathbf{x}) - \varphi_1(\mathbf{x}), & i \neq L_{0,1}(\mathbf{x}) \end{cases} \tag{3.6}$$

3D:

$$\phi_i(\mathbf{x}) = \begin{cases} \varphi_0(\mathbf{x}), & i = L_0(\mathbf{x}) \\ -\varphi_0(\mathbf{x}), & i = L_1(\mathbf{x}) \\ -\varphi_0(\mathbf{x}) - \varphi_1(\mathbf{x}), & i = L_2(\mathbf{x}) \\ -\varphi_0(\mathbf{x}) - \varphi_1(\mathbf{x}) - \varphi_2(\mathbf{x}), & i \neq L_{0,1,2}(\mathbf{x}) \end{cases} \tag{3.7}$$

The functions $\phi_i(\mathbf{x})$ resulting from this decomposition are positive inside and negative outside each object and are Lipschitz continuous. They are therefore legitimate level set functions despite being only approximations to the true signed distance functions.

It is worth noticing that the number of functions in the approximations for 2D and 3D are different. This is because of the difference in structure of boundaries between objects. In 2D, any boundary between neighboring objects is either a curve segment (between two objects) or a point (where three or more objects meet). In 3D, the boundaries are made of surface patches (between two objects), curve segments (between more than three objects), and points (for four or more objects). Therefore, in order to describe all the possible boundaries, we need to be able to describe all these structures in their respective spaces. The curves are defined with one level set function in 2D, and the points joining them require two functions. Similarly, surfaces in 3D require one function, curves are defined as the intersection of at least two surfaces, and points require three surfaces.

3.4 Training with MCDFs

In this section, we describe the use of the previously defined *mcdf*s in generating discriminative shape descriptors for the training datasets. We apply the above model to D-OCT B-scans (2D) and MN volumes (3D). These are discussed in sequence.

D-OCT Let us consider only 3 compartments in Fig. 1.3(a). The compartments are denoted as follows: portion of the eye above the retina, the retinal layer and portion below the retina. Note that such an assumption is only for the sake of simplicity of the model, and can easily be extended to N compartments for specific applications. First, we initialize each compartment by the zero-level label function L_0. The initialization is done by K-means method, wherein we input some initial seeds and number of compartments(=3). To compute the higher order label functions, we look at pairs

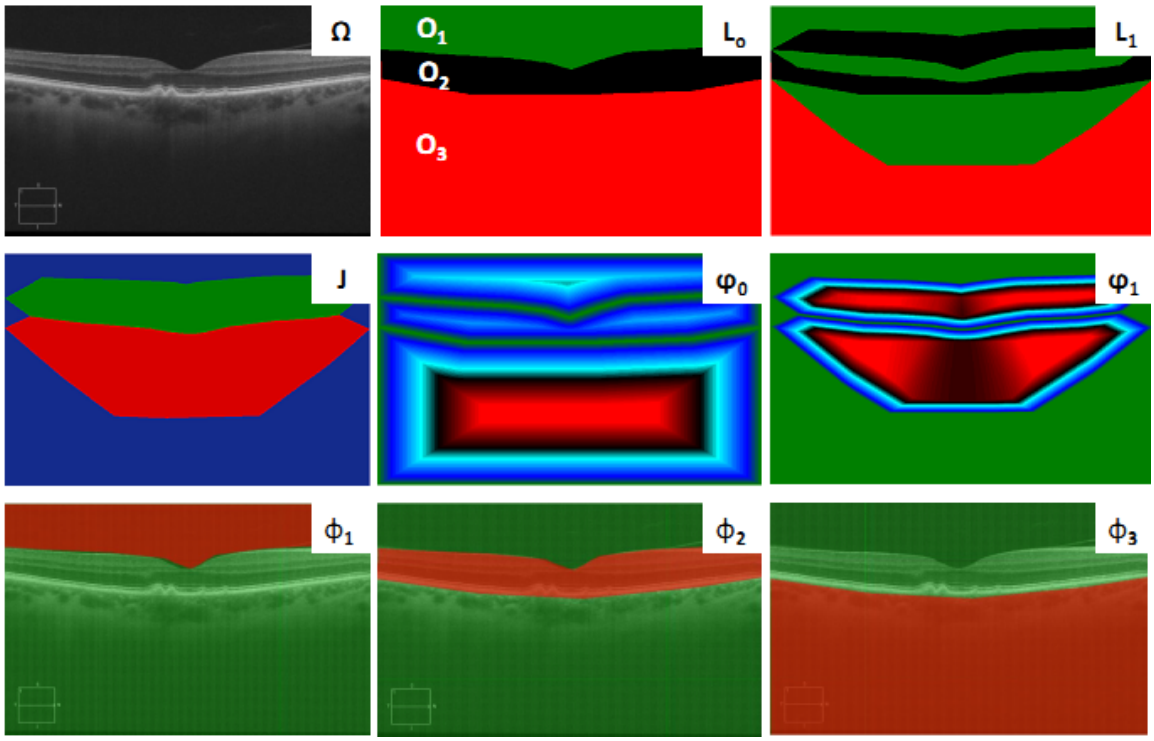


Figure 3.2. *Top:* Original Image and Label functions *Middle: Joint part* and distance functions. *Bottom:* Multi-compartment distance functions for each compartment O_i.

of objects. Consider the shared boundary $b_{ij} = |O_i \cap O_j|$ between two adjacent compartments O_i and O_j, and their union $B_{ij} = |O_i \cup O_j|$. The *joint part* of O_i and O_j is defined as:

$$J_{ij} = \{x \in O_i \cup O_j \quad s.t. \quad min_{y \in b_{ij}} \|x - y\| < min_{z \in B_{ij}} \|x - z\|\}$$

It is to be noted that joint parts of two objects is empty if they have no shared boundary. Next we define L_1 as follows:

$$L_1(x) = j, \quad if \quad x \in O_j \cap J_{ij} \forall i, j, i \neq j$$

With this notation, x is in J_{ij} if $\{L_0(x), L_1(x)\}$ is equal to $\{i, j\}$ or $\{j, i\}$.

Having defined the label functions, we now describe the distance functions as follows:

$$\begin{aligned}\varphi_0 &= \phi_{L_0} \\ &= \sum_{i=1}^{3} max(\mathbf{\Phi}_i, 0), \\ \varphi_1 &= \phi_{L_1} - \phi_{L_0} \\ &= \sum_{i=1}^{3} max(\mathbf{\Phi}_{ij}, 0) - max(\mathbf{\Phi}_i, \mathbf{\Phi}_j, 0)\end{aligned}$$

where $\mathbf{\Phi}_i$ and $\mathbf{\Phi}_{ij}$ are the level set functions (3.2) of O_i, and $O_i \cup O_j$. Finally, the *mcdf*s of each compartment are computed as,

$$\phi_i(x) = \begin{cases} \varphi_0(x), & L_0(x) = i; \\ -\varphi_0(x), & L_0(x) \neq i, L_1(x) = i; \\ -\varphi_0(x) - \varphi_1(x), & \text{otherwise} \end{cases} \tag{3.8}$$

Fig. 3.2 illustrates the label functions and *mcdf*s of each compartment.

The relationship between $\phi_i, i = 1, 2, \cdots N$, and the new functions φ_0, φ_1, L_0, L_1 being set, any curve transformation originally applied to ϕ_1 can be transferred to φ_0, φ_1, L and F. Furthermore it can also be proven that this representation $[\phi_i]$ is invariant to translation and rotation.

Lemma 1 (Invariance) *The mcdf shape descriptors in (3.8) are invariant to rotation and translation.*

Proof Paragios et al [34] showed that $\mathbf{\Phi}(\mathbf{x})$ (Eq. 3.2) is invariant to rotation and translation. We prove the above lemma by showing that ϕ_i for compartment O_1 corresponds to $\mathbf{\Phi}(\mathbf{x})$. It can be proved similarly, without loss of generality, for other compartments in Ω. Let, $\mathbf{x} = (x, y)$ be a point in $\mathcal{I}_T$, which is transformed to $\hat{\mathbf{x}}$ in $\mathcal{I}_S$ after rotating $\mathcal{I}_T$ by θ angle and translating by a vector (T_x, T_y). Thus,

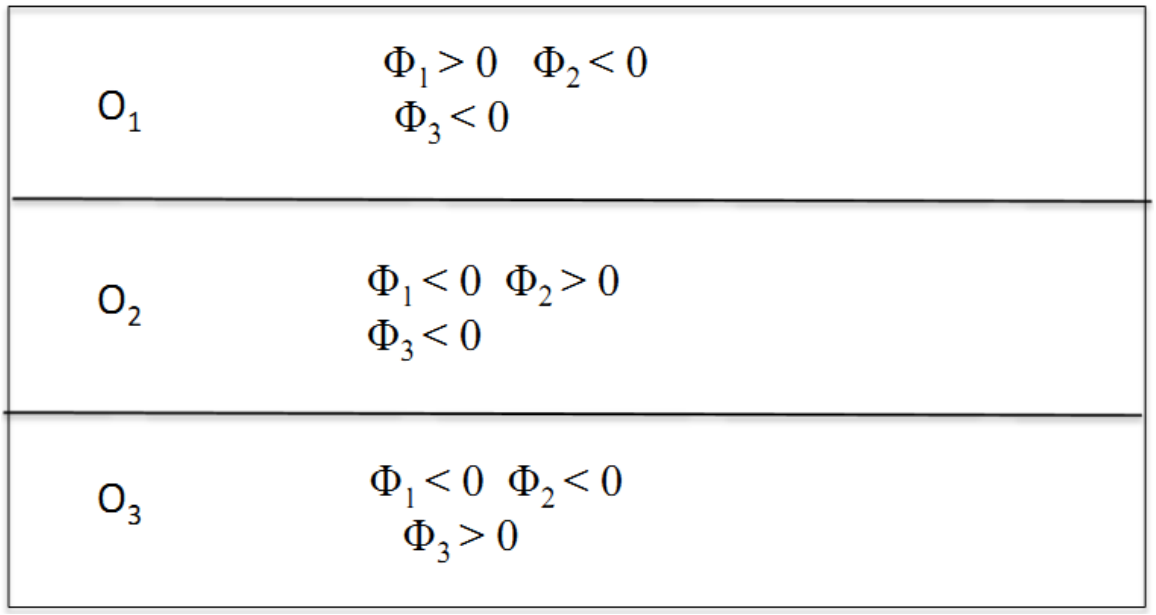


Figure 3.3. Sketch of the three compartments and their distance functions.

$$A(x,y) = \begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix} = s \begin{pmatrix} cos(\theta) & sin(\theta) \\ -sin(\theta) & cos(\theta) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} T_x \\ T_y \end{pmatrix} \tag{3.9}$$

The use of inverse transformation between $\mathcal{I}_S$ and $\mathcal{I}_T$ for $\mathbf{x}$ leads to the following equation:

$$\begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix} = \begin{pmatrix} xcos(-\theta) & ysin(-\theta) - T_x \\ -xsin(-\theta) & ycos(-\theta) - T_y \end{pmatrix} \tag{3.10}$$

Case 1 ($L_0(\mathbf{x}) = 1$) *In this case,* $\mathbf{x}$ *lies in* O_1*. Therefore,* $\varphi_0(\mathbf{x}) = \mathbf{\Phi}_1(\mathbf{x})$*, since only* $\mathbf{\Phi}_i$ *is positive, while* $\mathbf{\Phi}_j, j \neq i$ *are all negative. Therefore,*

$$\phi_1(\mathbf{x}) = \mathbf{\Phi}_1(\mathbf{x}) = min_{\mathbf{y} \in O_1} \|\mathbf{x} - \mathbf{y}\| . \tag{3.11}$$

Case 2 ($L_1(\mathbf{x}) = 1$) *In this case,* $\mathbf{x}$ *lies in the closest neighbor to* O_1*, which is* O_2*. All other distance functions get cancelled in* ϕ_0*. Therefore,*

$$\phi_1(\mathbf{x}) = \mathbf{\Phi}_2(\mathbf{x}) = min_{\mathbf{y}\in O_2} \|\mathbf{x}-\mathbf{y}\| \tag{3.12}$$

Case 3 ($L_{0,1}(\mathbf{x}) \neq 1$) *In this case,* $\mathbf{x}$ *lies in* O_3 *(Fig. 3.3). Therefore,* $\phi_1(\mathbf{x}) = -\varphi_0(\mathbf{x}) - \varphi_1(\mathbf{x})$. *It can be easily shown that in* φ_0, *the only term remaining is* $\mathbf{\Phi}_3$. *Similarly, in* φ_1, *the term remaining is* $\mathbf{\Phi}_{23} - \mathbf{\Phi}_3$. *Therefore,*

$$\begin{aligned}\phi_1(\mathbf{x}) &= -\mathbf{\Phi}_3(\mathbf{x}) - (\mathbf{\Phi}_{23}(\mathbf{x}) - \mathbf{\Phi}_3(\mathbf{x})) \\ &= -\mathbf{\Phi}_{23}(\mathbf{x}) = -min_{\mathbf{y}\in\overline{O_2\cup O_3}} \|\mathbf{x}-\mathbf{y}\|\end{aligned} \tag{3.13}$$

Hence we prove the *mcdf* shape descriptors are affine transformation invariant. ■

Thus, it is clear that *mcdf* descriptors for compartments in image are invariant to affine transformations. Also, these descriptors describe the topological characterestics of the compartments, both locally and globally. Therefore, they provide robust shape cues for image registration.

Motor Neurons Without loss of generality, let us consider a volume that consists of three sub-volumes (compartments) $\{O_i | i = 1, 2, 3\}$, with i denoting the part labels, which, in our application, correspond to soma, axon and dendrite of a neuron with a single dendrite. If Ω is the 3D image domain and $\mathbf{x}$ denotes location in Cartesian coordinates, $\forall \mathbf{x} \in \Omega$ three label functions are defined as:

$$\begin{aligned}&L_0 : L_0(\mathbf{x}) = i, \quad if \quad \mathbf{x} \in O_i; \\ &L_1 : L_1(\mathbf{x}) = j, \quad \text{if } O_j \text{ is first closest neighbor to } \mathbf{x}; \\ &L_2 : L_2(\mathbf{x}) = k, \quad \text{if } O_k \text{ is second closest neighbor to } \mathbf{x};\end{aligned}$$

Then the distance functions can be computed as follows:

$$\varphi_0(\mathbf{x}) = \sum_{i=1}^{3} max(\mathbf{\Phi}_i, 0)$$
$$\varphi_1(\mathbf{x}) = \sum_{i,j=1,i\neq j}^{3} \{max(\mathbf{\Phi}_{ij}, 0) - max(\mathbf{\Phi}_i, 0)\}$$
$$\varphi_2(\mathbf{x}) = \sum_{i,j,k=1,i\neq j\neq k}^{3} \{max(\mathbf{\Phi}_{ijk}, 0) - max(\mathbf{\Phi}_{ij}, \mathbf{\Phi}_{jk}, \mathbf{\Phi}_{ik}, 0)\}$$

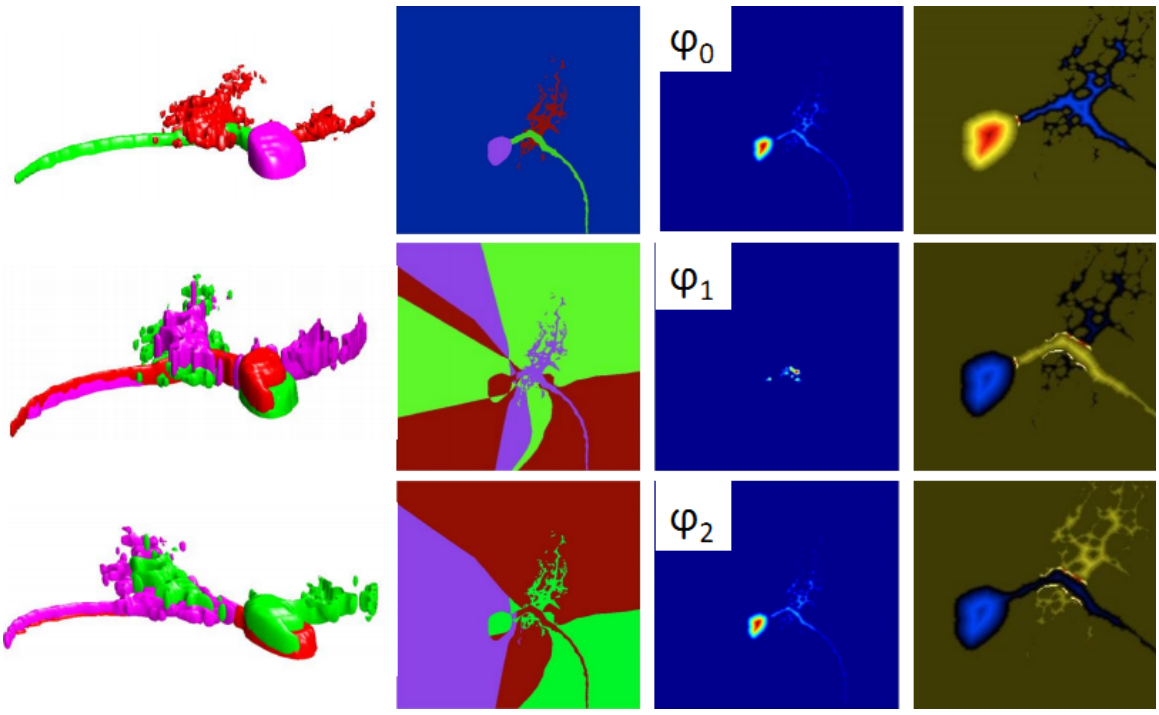
where $\mathbf{\Phi}_i$, $\mathbf{\Phi}_{ij}$ and $\mathbf{\Phi}_{ijk}$ are the distance transforms of the compartment O_i, $O_i \cup O_j$, and $O_i \cup O_j \cup O_k$.

Then *mcdf* of the i-th compartment is defined as:

$$\phi_i(\mathbf{x}) = \begin{cases} \varphi_0(\mathbf{x}), & L_0(\mathbf{x}) = i; \\ -\varphi_0(\mathbf{x}), & L_1(\mathbf{x}) = i; \\ -\varphi_0(\mathbf{x}) - \varphi_1(\mathbf{x}), & L_2(\mathbf{x}) = i; \\ -\varphi_0(\mathbf{x}) - \varphi_1(\mathbf{x}) - \varphi_2(\mathbf{x}), & L_{0,1,2}(\mathbf{x}) \neq i \end{cases}$$

Fig. 3.4 illustrate the multi-compartment label and distance functions for a neuronal volume. In Fig. 3.4(d), it is seen that topology of the shared boundaries of each compartment can be captured using level set functions. Notice that transformation of $\Omega \rightarrow \Re$ recasts this problem as joint multi-modal registration of each compartment. Such transformation is very suitable to track moving interfaces.

Note that, our shape descriptors are completely non-parametric. This avoids the task of training our model for suitable parameter configurations. Multi compartment distance functions provide topologicial features in images, like MN, where local shape features are indiscernible. This representation provides a feature space in which objective functions that are optimized using a gradient descent method can be conveniently used. One can prove that the gradient of the embedding distance function is a unit vector in the normal direction of the shape, and the representation satisfies



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

Figure 3.4. Label and distance functions of a neuron(colors indicate compartment labels). Panels (a)–(c): top row -L_0, middle row-L_1, bottom row-L_2. (a) Color-labeled volumes, (b) a cross-section of each case, and (c) corresponding cross-sections of the distance functions φ_0,φ_1, φ_2; (d) *mcdf*s for the three compartments - top: $\phi_{\bar{S}}$ (soma); middle: $\phi_{\bar{A}}$(axon); bottom: $\phi_{\bar{D}}$ (dendrite)

a sufficient condition for the convergence of gradient descent methods, which require continuous first derivatives. Furthermore, the use of the implicit representation facilitates the imposition of smoothness constraints since one would like to align the original structures as well as their clones that are positioned coherently in the image.

In Chapter 4, we show the proposed representation of shape and topology augments even relatively simple optimization procedures, like Sum-of-Squared-Distance (SSD). One reason for this is the inherent convex curvature of *mcdf*s. The convex subspace of level set functions guarantee a global minimum, which may not be unique. In this application, we seek a global alignment of each compartment so as to get an

average contour model. For registering D-OCT images, we incorporate local deformations along with global motions for estimation of registration parameters. Thus, our proposed framework is *variational*, which couples global transformation along side pixel-wise local deformations.

One concern associated with the level set distance functions, in general, is its efficiency in registration, since it has one dimension higher than the original shape. However, we alleviate this concern by using only a narrow band around the shape in the embedding space as the sample domain for registration. This significantly speeds up the execution, while producing comparable results to that using the entire image.

4 REGISTRATION METHODS

We introduced the concept of registration briefly in Chapter 2. In this chapter, we analyze in details the different methods for medical image registration. Registration, however, is an ubiquitous method that is applied to numerous domains, like satellite imaging, military automatic target recognition(ATR), remote sensing(multispectral classification, environmental monitoring, change detection), image mosaicing, weather forecasting, creating super-resolution images, integrating information into geographic information system(GIS), compiling and analyzing images and data from satellites. The primary goal in all of these applications is to project multiple data onto a common frame of reference for improved visualization and analysis. In medical imaging, registration is a powerful tool for clinicians to better visualize medical data. Image registration is the process of overlaying two or more images of the same scene taken at different times, from different viewpoints or by different sensors.

As mentioned earlier, registration is used for alignment of diverse variety of images, acquired from different modalities. Nevertheless, the majority of the registration methods consists of the following four steps.

- Salient and distinctive objects (closed-boundary regions, edges, contours, line intersections, corners etc.) are manually or, preferably, automatically detected. For further processing, these features can be represented by their point representatives (centers of gravity, line endings), which are called control points(CPs) in the literature.

- Next, the correspondence between the features detected in the reference image and those detected in the source image is established. Various feature descrip-

tors and similarity measures along with spatial relationships among the features are used for that purpose.

- The type and parameters of the mapping functions, aligning the source image with the reference image, are estimated. The parameters of the mapping functions are computed by means of the established feature correspondence.
- The source image is transformed by means of the mapping functions. Image values in non-integer coordinates are computed by interpolation technique.

Estimating the transform model depends on the application at hand as it requires a priori known information about the acquisition process and expected image degradations. Hence, depending on data acquisition techniques and nature of anatomical structures of interest, image registration methods vary in their model formulation.

4.1 Types of Models in Image Registration

There are three kinds of registration models mostly used in medical imaging. We describe these in brief.

The first form of registration is rigid registration. Transforming an object rigidly implies applying a transformation to it such that its shape remains unchanged. In order to transform the source to match the reference, it is necessary to determine a mapping from each voxel position in the reference to a corresponding position in the source. A rigid-body transformation in 3D is defined by 6 degrees of freedom(DOF), namely 3 translation and 3 rotation parameters. For each point $\mathbf{x} = [x_1, x_2, x_3]$ in an image, an affine mapping can be defined onto the co-ordinates of another space $\mathbf{y} = [y_1, y_2, y_3]$, whose relationship is expressed as a simple matrix multiplication $\mathbf{y} = \mathbf{M}\mathbf{x}$, where $\mathbf{M}$ is an appropriately constructed matrix from the transformation parameters. For rigid transformations specially, this is written as $\mathbf{y} = \mathbf{R}\mathbf{x} + \mathbf{T}$. Here, $\mathbf{R}$ is the rotation

matrix and $\mathbf{T}$ is the translation vector. Their parameters are as follows:

- Translations: If a point $\mathbf{x}$ is to be translated by $\mathbf{T}$ units, then the transformation is simply written as $\mathbf{y} = \mathbf{x} + \mathbf{T}$. In 3D, translation is mathematically written as $\mathbf{T} = [T_x, T_y, T_z, 1]^\top$ with each component T_i along the i^{th} axis.

- Rotations: In 2D, rotation is defined by a single angle. Consider a point at co-ordinate (x_1, x_2) on a two dimensional plane. A rotation of this point to new co-ordinate (y_1, y_2) by θ radians around the origin, can be generated by the transformation $y_1 = \cos(\theta)x_1 + \sin(\theta)x_2$ and $y_2 = -\sin(\theta)x_1 + \cos(\theta)x_2$. Thus, the rotation matrix in 2D can be written as:

$$\mathbf{R} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \tag{4.1}$$

  In 3D, there are 3 orthogonal planes that an object can be rotated in. These planes of rotation are normally expressed as being around the axes. Rotations of θ_1 radians about X-axis, θ_2 radians about Y, and θ_3 about Z-axis, are combined as shown in the equation below:

$$\mathbf{R} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & c_1 & s_1 & 0 \\ 0 & -s_1 & c_1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c_2 & 0 & s_2 & 0 \\ 0 & 1 & 0 & 0 \\ -s_2 & 0 & c_2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c_3 & s_3 & 0 & 0 \\ -s_3 & c_3 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{4.2}$$

  where $\forall i \in (1, 2, 3)$, c_i, s_i denote the cosine and sine of angles $\theta_1, \theta_2, \theta_3$ respectively. For notational convenience, hereafter first matrix(θ_1) is denoted as $\mathbf{R}_1$, second(θ_2) as $\mathbf{R}_2$ and third(θ_3) as $\mathbf{R}_3$.

Going a step further, non-rigid registration methods aim for an alignment at sub-pixel level. Non-rigid registration is a two-step process: first, the global motion is corrected using a rigid or affine transformation. The globally aligned image then becomes the starting estimate for a second stage, where the local motion is further modeled a local objective function. The combined model $\mathbf{M}$ is as follows:

$$\mathbf{M}(\mathbf{x}) = \mathbf{M}_{global}(\mathbf{x}) + \mathbf{M}_{local}(\mathbf{x})$$

In addition to the 6 DOFs used in rigid registration, a scale parameter($\mathbf{s}$) is also introduced in $\mathbf{M}$. In 3D, the scaling matrix($\mathbf{S}$) is written as follows:

$$\mathbf{S} = \begin{bmatrix} s_x & 0 & 0 & 0 \\ 0 & s_y & 0 & 0 \\ 0 & 0 & s_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

where, s_x, s_y, s_z are the scaling factors about the principal axes. The objective function in such a setup similarly consists of two terms: a global similarity term, which aligns the source image with approximate correspondence onto the reference, and a local deformation term, which forces pixel-wise correspondence. Mathematically,

$$E = E_{sim}(\mathcal{I}_T, \mathbf{M}(\mathcal{I}_S(\mathbf{x}))) + \alpha E_{def}(\mathcal{I}_T, \mathbf{M}(\mathcal{I}_S(\mathbf{x} + \mathbf{u}))) \tag{4.3}$$

Here, $\mathbf{u}$ is the local deformation term. The local deformation field $\mathbf{u}$ is defined in the image plane and has different values for adjacent pixels in non-rigid objects. α is a constant that balances the contributions of the two terms(global motion, local deformations). The interpretation of this criterion is simple; registration errors caused by the use of the rigid transformation are corrected using the local deformation field. Local deformations increase significantly the complexity of the model. A simple way to decrease this complexity is to constrain the search space. The local deformation field is to be computed only in the vicinity of the source shape. This hierarchical framework improves performance of the registration process. At the same time, the

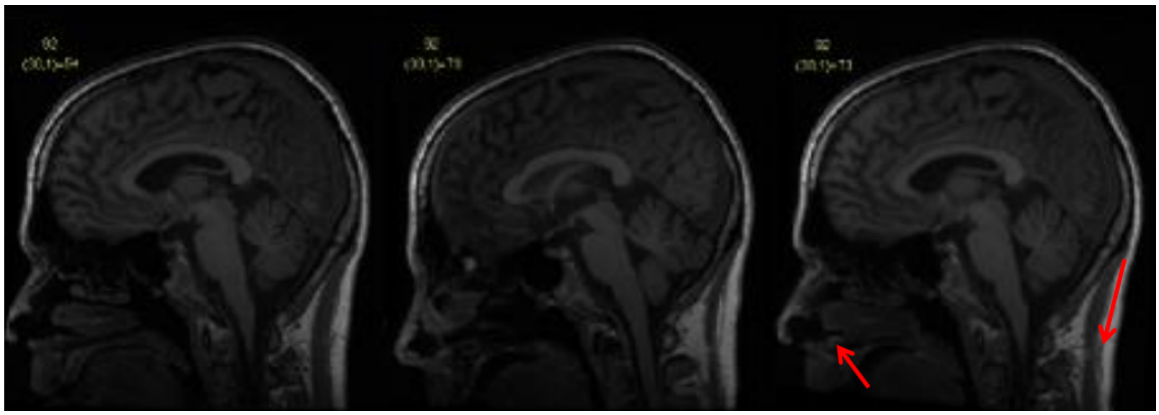
role of the local registration field is to deal with local deformations and non-rigid parts. Therefore, local deformations are computed in a small band defined in the vicinity of the shape. Hence, (4.3) is modified as follows:

$$E = N(.)E_{sim}(\mathcal{I}_T, \mathbf{M}(\mathcal{I}_S(\mathbf{x}))) + \alpha N(.)E_{def}(\mathcal{I}_T, \mathbf{M}(\mathcal{I}_S(\mathbf{x}+\mathbf{u}))) \tag{4.4}$$

Here, N(.) is a binary function that ensures transformation of only those pixels in the neighborhood of the reference compartment.

In recent years, a new type of registration scheme has been used for many applications - this is known as articulated registration. This type of registration refers to aligning rigid objects with non-rigid connected parts. It is used for applications such as biometric in radiographs as well as anthropometry and pose estimation [35]. In such a setup, multiple rigid segments are interconnected at joints which deform elastically to maintain surface continuity. For example, Fig. 4.1(c) illustrates a typical example of articulated registration. The rigid parts, like bones, are interconnected by deformable tissue. The goal is to align these anatomical structures, such that the rigid parts are not allowed to deform, but the joints undergo some degree of deformation. The main workhorse for solving articulated registration problems are landmark based approaches. Although such techniques work reasonably well, the selection and position of control points have critical impact on the performance of these models. For applications such as those in Fig. 4.1(c), the anatomical points-of-interest, such as junction of fingers, serve as suitable landmarks which can be used as input for articulated registration framework.

As mentioned earlier, distinctive features in the reference image are chosen as controls for initialization. In our method, we formulate a variational registration model using MCDF descriptors. A major advantage of such a formulation is that convergence of such a model using a suitable optimization criterion gives us the global optimal registration parameters. The distance functions used in the process are Euclidean,



(a) Rigid Registration in Sagittal T1-wighted MRI (Arrows indicate regions of transformation)

(b) Non-rigid Registration of Femur in Adult Male (Arrows indicate regions of deformation)

(c) Articulated Registration of Hands in Radiography (Joints indicated by red circles)

Figure 4.1. Examples of Rigid, Non-rigid and Articulated registration *Left:* Reference; *Middle:* Source; *Right:* Registered Image

which has a convex curvature. Another important advantage is that our model is non-parametric, which requires minimal tuning of parameters.

4.2 Image Similarity Criteria

An important issue in this context is choosing a suitable similarity criterion E_{sim}. Any similarity criterion that preserves convexity is best suited for this purpose. The following are some candidates:

1. Sum-of-Squared-Distances (SSD): SSD is an algorithm for measuring similarity between images. It works by taking absolute difference between each pixel in the original image and the corresponding pixel in the transformed image. These differences are summed, using $\mathcal{L}^1$ norm to create a simple metric of image similarity. Mathematically, this can be stated as follows:

$$E_{SSD}(\mathcal{I}_1, \mathcal{I}_2) = \iint\limits_{(x,y)\in\Omega} (\mathcal{I}_1(x,y) - \mathcal{I}_2(x,y))^2 \quad d\Omega \tag{4.5}$$

2. Normalized Cross Correlation (NCC): Cross-correlation (CC) is a measure of similarity between two waveforms in signal processing. In image processing, cross-correlation between two images is defined as follows:

$$E_{NCC}(I_1, I_2) = \frac{1}{|\mathcal{I}_1|} \iint\limits_{(x,y)\in Omega} \frac{(\mathcal{I}_1 \times \mathcal{I}_2)}{\sigma_{\mathcal{I}_1} \times \sigma_{\mathcal{I}_2}} \quad d\Omega \tag{4.6}$$

   where, $\sigma_{\mathcal{I}_1}$ and $\sigma_{\mathcal{I}_2}$ are the standard deviations of $\mathcal{I}_1$ and $\mathcal{I}_2$ respectively. In vector notation, this can be written as follows:

$$E_{NCC}(\mathcal{I}_1, \mathcal{I}_2) = \left\langle \frac{\mathcal{I}_1}{\|\mathcal{I}_1\|}, \frac{\mathcal{I}_2}{\|\mathcal{I}_2\|} \right\rangle \tag{4.7}$$

3. Mutual Information (MI): Mutual information is an information-theoretic criterion for measuring the global statistical dependency of its two input random

variables. Mathematically, if l_1 and l_2 denote intensity (or distance value) random variables in the image domains of I_1 and I_2 respectively, and $\mathcal{H}$ represents the differential entropy, then the similarity function is defined as follows:

$$E_{MI}(\mathcal{I}_1, I_2) = \mathcal{H}\left[p^{\mathcal{I}_1}(l_1)\right] + \mathcal{H}\left[p^{\mathcal{I}_2}(l_2)\right] - \mathcal{H}\left[p^{\mathcal{I}_1,\mathcal{I}_2}(l_1, l_2)\right] \tag{4.8}$$

Here, p^{I_1} is the intensity probability density function (p.d.f.) in domain of $\mathcal{I}_1$, p^{I_2} is the intensity p.d.f. in domain of $\mathcal{I}_2$ and $p^{\mathcal{I}_1,\mathcal{I}_2}$ is their joint distribution. Differential entropy $\mathcal{H}$ can written as $\mathcal{H}\left[p^{\mathcal{I}}(l)\right] = -\int_{\Omega} p^{\mathcal{I}}(l) \log p^{\mathcal{I}}(l) dl$.

For our applications, we use SSD and NCC criteria as our similarity function. MI is usually used in registration of images acquired from different modalities. In the next section, we describe in details our method of global registration in 3D Neuronal volume. Note that, we only seek a coarse alignment as we want to preserve the morphological variance in compartments of MN belonging to different subtypes.

4.3 Motor Neuron Volume Registration

Using the MCDF shape descriptors defined in Eq. (3.14), we seek a method to register every source volume $\mathcal{S}$ with a reference volume $\mathcal{R}$. Therefore, we need a transformation $\mathcal{F}$ that considers the neuron volume as a deformable set of rigid compartments, to preserve the compartment shape variations. By the latter, we mean that the image volume, as a whole, deforms. However, the compartments, namely soma, axon and dendrites are rigid and thus not allowed to deform. The deformations in the set are due to the changes in joint angle positions. For example, in the hand example (Fig. 4.1c) the compartments, namely phalanges, metacarpels etc. are rigid, although at the joints, consisting of connective tissue, some degree of deformations is allowed. Thus in summary, $\mathcal{F}$ is a set of parameters that correspond to locally rigid registrations: each compartment is registered as a rigid object (translation and rotation),

but the relative topology among compartments can change while preserving their pair-wise adjacency properties. The objective is the minimization of the energy,

$$E(\mathcal{F}) = \iint_{\Omega} \sum_{i \in \{\bar{S},\bar{A},\bar{D}\}} \xi\big(\phi_i^{\mathcal{R}}(\mathbf{x}), \mathcal{F}[\phi_i^{\mathcal{S}}(\mathbf{x})]\big)\big(\phi_i^{\mathcal{R}}(\mathbf{x}) - \mathcal{F}[\phi_i^{\mathcal{S}}(\mathbf{x})]\big)^2 d\mathbf{x} \tag{4.9}$$

where $\{\bar{S}, \bar{A}, \bar{D}\}$ denote the compartment labels *soma*, *axon* and *dendrite*, respectively, and $\phi_i^{\mathcal{R}}$ and $\phi_i^{\mathcal{S}}$ are the *MCDFs* of the i_{th} compartment in the reference $\mathcal{R}$ and source $\mathcal{S}$ volumes, as defined in Eq. (3.14). Here, ξ is the binary function, analogous to $N(.)$ in Eq. (4.4). Given a similarity threshold ε, for any two compartment *MCDF*s, ϕ_1 and ϕ_2,

$$\xi\big(\phi_1, \phi_2\big) = \begin{cases} 0, & \|\phi_1 - \phi_2\| > \varepsilon \\ 1, & \text{otherwise} \end{cases} \tag{4.10}$$

We solve the minimization of Eq. (4.9) using the standard gradient descent Newton procedure in the transformation parameter space. Specifically, if $\mathcal{F} = \{\mathbf{t}, \mathbf{r}\}$ is the set of translation $\mathbf{t} = \{t\}$ and rotation $\mathbf{r} = \{r\}$ parameters, respectively, then in our approach we calculate,

$$\Delta\mathbf{t} = 2 \iint_{\Omega} \sum_{i \in \{\bar{S},\bar{A},\bar{D}\}} \xi(\phi_i^{\mathcal{R}}, \phi_i^{\mathcal{S}}) \frac{\partial \phi_i^{\mathcal{R}}}{\partial \mathbf{x}} \big(\phi_i^{\mathcal{R}} - \mathcal{F}[\phi_i^{\mathcal{S}}]\big) \tag{4.11}$$

$$\Delta\mathbf{r} = 2 \iint_{\Omega} \sum_{i \in \{\bar{S},\bar{A},\bar{D}\}} \xi\big(\phi_i^{\mathcal{R}}, \phi_i^{\mathcal{S}}\big) \big(\nabla \phi_i^{\mathcal{S}} . \frac{\partial \mathcal{F}}{\partial r}\big) \big(\phi_i^{\mathcal{R}} - \mathcal{F}[\phi_i^{\mathcal{S}}]\big) \tag{4.12}$$

Then, given these PDEs in Eq. (4.11, 4.12), we achieve the minimization of the energy in Eq. (4.9) based on the scheme in Fig. 4.2.

Fig. 4.3 shows the neuron multi-compartment volume alignment during the minimization of Eq. 4.9. It is worthwhile to mention that the compartments are not registered pixel-wise. This is due to the reason that our goal here is to preserve the topological variations of the neuronal compartments, alongside getting a coarse alignment of the whole volume. In Chapter 5, where we build the Average Shape Model

input: $\{\phi^{\mathcal{R}}_{\bar{S}}, \phi^{\mathcal{R}}_{\bar{A}}, \phi^{\mathcal{R}}_{\bar{D}}\}, \{\phi^{\mathcal{S}}_{\bar{S}}, \phi^{\mathcal{S}}_{\bar{A}}, \phi^{\mathcal{S}}_{\bar{D}}\}$, $\mathcal{F}_0 = \{\mathbf{t}_0, \mathbf{r_0}\}$

Step 1: compute $E(\mathcal{F}_0)$ from Eq. (4.9)

Step 2: for each compartment, $t \leftarrow t_0 + \Delta\mathbf{t}$ (Eq. (4.11))

Step 3: for each compartment, $r \leftarrow r_0 + \Delta\mathbf{r}$ (Eq. (4.12))

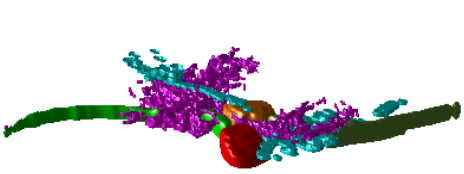
Step 4: set $\mathbf{t} = \{t\}$, $\mathbf{r} = \{r\}$, $\mathcal{F} = \{\mathbf{t}, \mathbf{r}\}$

Step 5: compute $E(\mathcal{F})$ from Eq. (4.9)

Step 6: If $|E(\mathcal{F}) - E(\mathcal{F}_0)| > \epsilon$: *terminate*

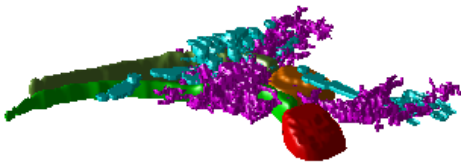
else: $\mathcal{F}_0 \leftarrow \mathcal{F}$ and go back to step 2

Figure 4.2. Minimization of our objective function in Eq. (4.9).

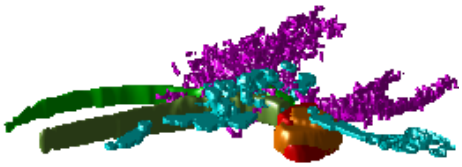


(a) Iteration #: 3

(b) Iteration #: 13



(c) Iteration #: 22

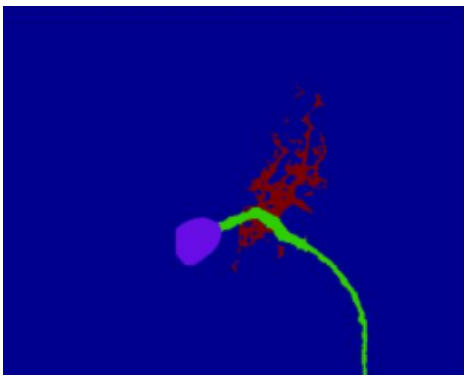


(d) Iteration #: 28

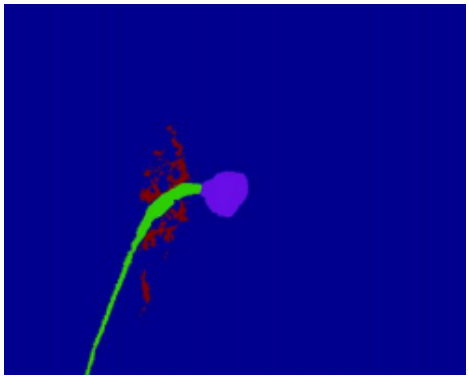
Figure 4.3. In our model, the multi-compartment sample volume registration (training phases) is solved as a *deformable set of pair-wise adjacent rigid parts*. The figure illustrates intermediate iterations during minimization of the objective in Eq. (4.9). Reference($\mathcal{R}$) and source($\mathcal{S}$) compartments are shown in $\{orange, dark\ green, cyan\}$ and $\{red, light\ green, magenta\}$ respectively.

for ASM, we capture these variations using Principal Component Analysis (PCA) in the higher dimensional MCDF subspace.

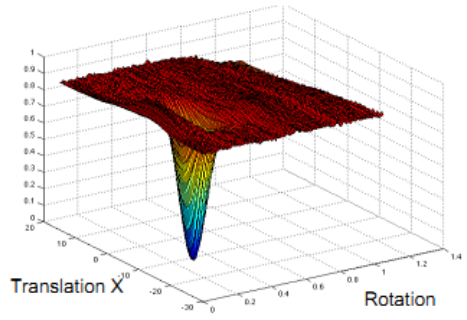
In order to perform a study on the performance of our global registration technique, we take the 2D similarity transformation model on two horizontal slices from reference($\mathcal{R}$)



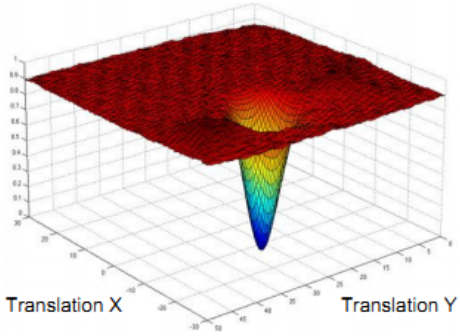
(a) 2D slice from Reference($\mathcal{R}$) volume



(b) 2D slice from Source($\mathcal{S}$) volume



(c) Translation in x and r unknown



(d) Translation in x and y unknown

Figure 4.4. Empirical validation of convexity

and source($\mathcal{S}$) volumes, with three parameters: translations in x and y directions, and the 2D rotation angle. Then, we constrain the unknown parameter space in two dimensions and empirically evaluate the form of the global registration objective function Eq. (4.9). In Fig. 4.4, we have studied the following two cases:

1. translation in x and rotation angle r unknown (Fig. 4.4c)
2. translations in x, y directions unknown (Fig. 4.4d),

In each case, we quantized the search space using a uniform sampling of 100 elements for all unknown parameters. Translations in (x, y) were in the range of $[-30, 30] \times [-30, 30]$ and rotation in $[-\frac{\pi}{3}, \frac{\pi}{3}]$. Then, we performed an exhaustive grid search over the space of two unknown parameters, by considering all possible combinations derived from sampling, while the other parameter was fixed. The resulting objective functional as defined in Eq. 4.9, as shown in Fig. 4.4(c-d) is smooth

and exhibits a single global minimum. This indicates the well-behaved optimization criterion with smooth convergence properties.

Next, we discuss another registration method, Non-Rigid or Deformable registration, where we use MCDF shape descriptors to acquire a global alignment as well as local pixel-wise deformation. We show application of our method on Directed OCT data.

4.4 D-OCT B-scan Registration

Enhancing the signal-to-noise ration (SNR) in OCT B-scans is an important step for clinical treatment. When applying OCT for *in vivo* imaging of the human retina, small movements of the head and pupils leads to axial movements of the retina relative to the imaging geometry. In the case of directed OCT, this problem becomes more challenging due to changes in the beam entry position. The natural solution to such a problem is a pixel-wise registration of the different orientations of images. Thomadsen et al. [36] proposed a mean-shift based registration algorithm with a regularization term to correct vertical and horizontal shifts between corresponding columns of two retinal scans. Their proposed method primarily corrects global shifts between the reference and source images. However, to attain a better visualization of the Henle's Fiber Layer, it is necessary to project images in all the different orientations onto a common frame of reference or orientation. (Fig. 1.3) The preferred orientation in this case is the horizontal direction, or central entry position of the infrared beam.
As discussed before, our goal here is to first find a global alignment of the three orientations of retinal scans onto a common frame of reference. Next, we use this coarse grained image for local deformations to obtain a pixel-wise mapping. To augment our registration framework, we first align the *foveal center* (Fig. 1.3) of the images. This provides a good initialization to our registration framework.

Let, $\mathcal{F}$ be the set of transformation parameters. $\mathcal{F}$ consists of two *translation* (t_x and t_y) parameters, one *rotation* (θ) and one *scale* (s) parameter. The objective, similar to Eq. (4.4), is the minimization of the following energy:

$$\begin{aligned} E(\mathcal{F}) = \quad & \alpha \iint_{\Omega} \sum_{c} \xi\big(\phi_c^{\mathcal{R}}, \phi_c^{\mathcal{S}}\big)\big(s\phi_c^{\mathcal{R}} - \mathcal{F}[\phi_c^{\mathcal{S}}]\big)^2 d\Omega \\ & +(1-\alpha) \iint_{\Omega} \sum_{c} \xi\big(\phi_c^{\mathcal{R}}, \phi_c^{\mathcal{S}}\big)\big(s\phi_c^{\mathcal{R}} - \mathcal{F}[\phi_c^{\mathcal{S}}(\mathbf{x}+\mathbf{u})]\big)^2 d\Omega \end{aligned} \tag{4.13}$$

where, c represents each compartment in (O_1, O_2, O_3) (Fig. 3.2). ξ is the narrow-band binary function as defined in Eq. (4.10). The interpretation of this criterion is that, registration errors caused by the use of the rigid transformation are corrected using the local deformation field($\mathbf{u}$).

We solve Eq. (4.13) using the standard Quasi-Newton gradient descent method in the transformation parameter space. We obtain the updates for translation($\mathbf{t}$), rotation(θ) and scale(s) parameters using the following partial differential equations.

Translation

$$\begin{aligned} \Delta\mathbf{t} = \quad & 2\alpha \iint_{\Omega} \sum_{c} \xi\big(\phi_c^{\mathcal{R}}, \phi_c^{\mathcal{S}}\big) \frac{\partial \phi_c^{\mathcal{S}}}{\partial \mathbf{x}} E_{sim} \\ & +2(1-\alpha) \iint_{\Omega} \sum_{c} \xi\big(\phi_c^{\mathcal{R}}, \phi_c^{\mathcal{S}}\big) \frac{\partial \phi_c^{\mathcal{S}}}{\partial \mathbf{x}} E_{deform} \end{aligned} \tag{4.14}$$

Rotation

$$\begin{aligned} \Delta\theta = \quad & 2\alpha \iint_{\Omega} \sum_{c} \xi\big(\phi_c^{\mathcal{R}}, \phi_c^{\mathcal{S}}\big)\big(\nabla\phi_c^{I_s}, \frac{\partial \mathcal{F}}{\partial \theta}\big) E_{sim} \\ & +2(1-\alpha) \iint_{\Omega} \sum_{c} \xi\big(\phi_c^{\mathcal{R}}, \phi_c^{\mathcal{S}}\big)\big(\nabla\phi_c^{\mathcal{S}}, \frac{\partial \mathcal{F}}{\partial \theta}\big) E_{deform} \end{aligned} \tag{4.15}$$

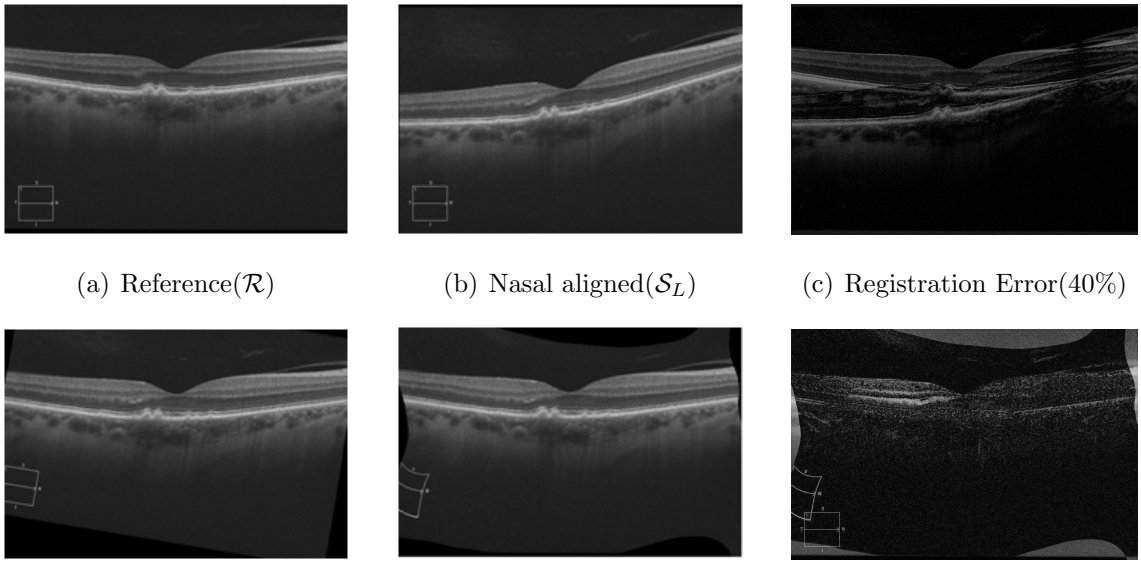
Scale Factor

$$\begin{aligned} \Delta s = \quad & 2\alpha \iint_{\Omega} \sum_{c} \xi\big(\phi_c^{\mathcal{R}}, \phi_c^{\mathcal{S}}\big)\big(\phi_c^{\mathcal{S}} - \frac{\partial \phi_c^{\mathcal{R}}}{\partial \mathbf{x}} R_\theta(\mathbf{x})\big) E_{sim} \\ & +2(1-\alpha) \iint_{\Omega} \sum_{c} \xi\big(\phi_c^{\mathcal{R}}, \phi_c^{\mathcal{S}}\big)\big(\phi_c^{\mathcal{S}} - \frac{\partial \phi_c^{\mathcal{R}}}{\partial \mathbf{x}} R_\theta(\mathbf{x})\big) E_{deform} \end{aligned} \tag{4.16}$$

Deformation

$$\Delta\mathbf{u} = \sum_{c} 2\big(1-\alpha\big)\xi\big(\phi_c^{\mathcal{R}}, \phi_c^{\mathcal{S}}\big)\nabla_x\phi_c^{\mathcal{S}}\big(\mathbf{x}+\mathbf{u}\big)E_{deform} \tag{4.17}$$

where $E_{sim} = \big(s\phi_c^{\mathcal{R}} - \mathcal{F}[\phi_c^{\mathcal{S}}]\big)$ is the global alignment term and $E_{deform} = \big(s\phi_c^{\mathcal{R}} - \mathcal{F}[\phi_c^{\mathcal{S}}(\mathbf{x}+\mathbf{u})]\big)$ denotes the local deformation field at each pixel $\mathbf{x}$. R_θ denotes the first derivative of the 2D Rotation matrix (Eq. 4.1). Note that, the deformation parameter $\mathbf{u}$ is computed at each pixel of every iteration. The interpretation of Eq. (4.13) is simple. The first term aims at finding a pixel-wise distance correspondences according to a rigid transformation. The second term seeks at correcting the correspondences in the pixel level using a local deformation field over the existing global model. This analogy is pictorially shown in Fig. 4.5.



(a) Reference($\mathcal{R}$) (b) Nasal aligned($\mathcal{S}_L$) (c) Registration Error(40%)

(d) Global Registration (e) Global-to-Local Registration (f) Registration Error(10%)

Figure 4.5. Global and Global-to-Local Registration

Fig. 4.3 and Fig. 4.5 show that our proposed method is robust and is able to align compartments with pixel-wise intensity correspondences. Our registration framework

aligns images or volumes from different viewpoints onto a common reference frame. The next step is to capture the principal components of shape variances of all these objects. This is described in Chapter 5.

In summary, in this chapter we formulated a variational level set framework for registration. Our model is non-parametric, implicitly convex and hence guarantees convergence inside the narrow band of the zero-order level set function. We illustrated performance of our model in 2D as well as 3D data. Qualitative results indicate our method is robust, invariant to affine transformations, and can successfully correct affine motion in varied scenarios. Quantitative analysis and comparisons are provided in Chapter 6.

5 MODEL-BASED SEGMENTATION

In this chapter, we discuss in details a popular *Top-down* model-based segmentation method, the *Active Shape Models* (ASM), and its applications to image segmentation.

5.1 Active Shape Models: Modeling and Background

Introduced first by Cootes et al. [20], ASMs are similar to Active Contour Models (or "Snakes" [8]), in that the shape boundary is modeled by a vector of control points, or landmarks, which are updated at each iteration conforming to a local profile in the image domain. A landmark represents a distinguishable point, like a feature, present in most of the training images under consideration. Shapes are represented as a set of these landmarks, which are aligned first in all of the training data. The mean shape is the mean of the aligned training shapes. The ASM starts the search for landmarks from the mean shape. It then repeats the following two steps until convergence (i) updates the location of landmarks by template matching of the image texture around each point (ii) conform the tentative shape to a global shape model. The individual template matches are unreliable and the shape model pools the results of the weak template matches to form a stronger overall classifier. Hence, there are two types of submodel that make up the ASM overall.

1. *The profile models* (one for each landmark at each pyramid level) are used to locate the approximate position of each landmark by template matching. Classical ASM forms a fixed-length normalized gradient vector (called the *profile*) by sampling the image along the normal direction to the shape boundary at each landmark. During training on manually landmarked faces, at each landmark we calculate the mean profile vector $\bar{\mathbf{g}}$ and the profile covariance matrix $\mathbf{S_g}$.

During searching, we displace the landmark along the normal line to the pixel whose profile $\mathbf{g}$ has lowest Mahalanobis distance from the mean profile $\bar{\mathbf{g}}$:

$$Mahalanobis\ distance = (\mathbf{g} - \bar{\mathbf{g}})^T \mathbf{S}_{\mathbf{g}}^{-1} (\mathbf{g} - \bar{\mathbf{g}}) \tag{5.1}$$

2. *The shape model* specifies allowable motions of landmarks. It generates a shape $\hat{\mathbf{x}}$ which is defined as follows:

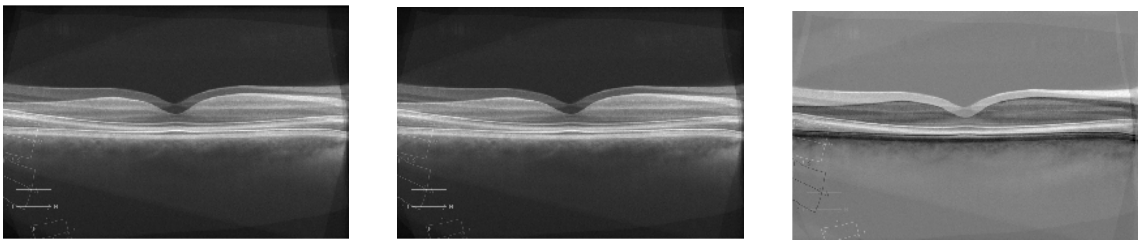
$$\hat{\mathbf{x}} = \bar{\mathbf{x}} + \mathbf{\Phi}\mathbf{b} \tag{5.2}$$

where $\bar{\mathbf{x}}$ is the mean shape, $\mathbf{b}$ is a parameter vector, and $\mathbf{\Phi}$ is the matrix of selected eigenvectors of the covariance matrix $\mathbf{S}_s$ of the points of the aligned training shapes. We model the variations in the training set using principal component analysis (PCA). We can generate various shapes with Eq. 5.2 by varying the vector parameter $\mathbf{b}$.

Extensions of the ASM model [37] propose an implicit representation of the shape boundary, like level sets. Level set functions (LSF), as discussed before, are ideal to account for prior knowledge of shapes. Thus, we represent shapes as zero-order LSF and update LSFs at each iteration following the ASM framework. In the next section, we describe the method of modeling prior knowledge in the level set space.

5.2 AGM: Modeling prior Knowledge in Level Sets

Let us consider a training set C_i of N registered curves or surfaces. Also, let the distance transform used to represent C_i as a LSF be ϕ_i. The next step is the construction of the shape model, using the aligned contours. In order to create an invariant representation, we first normalize the training set ϕ_i. Subtraction of the mean, obtained by averaging over ϕ_i's, is a common selection. However, simple averaging would not provide a distance function. To overcome this limitation, we use a variational method, seeking to estimate the distance function ($\phi_{\mathcal{M}}$) that minimizes:



(a) Mean Image (b) Principal mode(+σ) (c) Principal mode(-σ)

Figure 5.1. Major modes of variations in ASM

$$E(\phi_{\mathcal{M}}) = \sum_{i=1}^{N} \int_{\Omega} \left(\phi_i - \phi_{\mathcal{M}}\right)^2 d\Omega, \;\; SUBJECT\ TO : |\nabla \phi_{\mathcal{M}}|^2 = 1 \tag{5.3}$$

Eq. (5.3) can be optimized through a gradient descent method:

$$\frac{d\phi_{\mathcal{M}}}{dt} = \sum_{i=1}^{N} \left(\phi_i - \phi_{\mathcal{M}}\right)$$

while $\phi_{\mathcal{M}}$ is projected to the space of the distance functions. The two steps alternate until convergence. As in Eq. (5.2), once the samples ϕ_i are centered with respect to $\phi_{\mathcal{M}}$, the most important modes of variations are recovered through PCA:

$$\psi(\boldsymbol{\lambda}) = \phi_{\mathcal{M}} + \sum_{i=1}^{m} \lambda_i U_i$$

where m is the number of retained modes of variation, U_i are these modes (eigenvectors), and λ_i are the corresponding eigenvalues. ψ is hereafter called as the *average shape model*. Fig. 5.1 show an example of our approach in D-OCT images. Moreover, the implicit representation of the surfaces make the modeling phase entirely automatic.

Next, we formulate the evolution of of a LSF $\phi(\mathbf{x})$ towards the average shape model, under the influence of an affine transformation matrix $A(\mathbf{x}) = s\mathbf{R}(\mathbf{x}) + \mathbf{T}$. Here, s is the scale factor, $\mathbf{R}$ is the rotation matrix following Eq. (4.1) in 2D, and Eq. (4.2) in 3D, and $\mathbf{T}$ is the translation vector. For computational speed, we estimate the

prior within the vicinity of the zero-order level set representation(ψ). The objective for evolving ϕ can be written as:

$$E(\phi, A, \boldsymbol{\lambda}) = \int_{\Omega} \xi(\phi)\big(s\phi - [\phi_{\mathcal{M}}(A) - \sum_{j=1}^{m} \lambda_j U_j(A)]\big)^2 d\Omega \tag{5.4}$$

Eq. (5.4) can be optimized using gradient descent Quasi-Newton method for its parameters. The equation of evolution of ϕ is given as:

$\boldsymbol{\phi}$:

$$\frac{d\phi}{dt} = -2\int_{\Omega} \xi(\phi)\big(s\phi - \phi_{\mathcal{M}}(A)\big) - \int_{\Omega} \xi(\phi)\big(s\phi - \phi_{\mathcal{M}}(A)\big)^2$$

Differentiation with respect to the modes weights give us a close form of the optimal parameters by solving the linear system $U\lambda = b$ where:

$\boldsymbol{\lambda}$:

$$\begin{cases} \boldsymbol{U}(j) = \int_{\Omega} \xi(\phi) U_i(A) U_j(A) d\Omega \\ \boldsymbol{b} = \int_{\Omega} \xi(\phi)\big(s\phi - \phi_{\mathcal{M}}(A)\big) d\Omega \end{cases}$$

where $\boldsymbol{U}$ is a $m \times m$ positive definite matrix. Finally, the energy with respect to the transformation parameters is minimized using calculus of variations:

Scale Factor:

$$\frac{ds}{dt} = 2\int_{\Omega} \big(s\phi - \phi_p(A)\big)\big(- \phi + \nabla\phi_p(A).\frac{\partial A}{\partial s}\big) d\Omega$$

Rotation:

$$\frac{dR}{dt} = 2\int_{\Omega} \big(s\phi - \phi_p(A)\big)\big(\nabla\phi_p(A)\ .\ \frac{\partial A}{\partial R}\big) d\Omega\ ,\ R \in \{\mathbf{R}_1, \mathbf{R}_2, \mathbf{R}_3\}$$

Translation:

$$\frac{dT}{dt} = 2\int_{\Omega} \big(s\phi - \phi_p(A)\big)\big(\nabla\phi_p(A)\ .\ \frac{\partial A}{\partial T}\big) d\Omega\ ,\ T \in \{\mathbf{T}_x, \mathbf{T}_y, \mathbf{T}_z\}$$

where, $\phi_p(A) = [\phi_{\mathcal{M}}(A) - \sum_{j=1}^{m} \lambda_j U_j(A)]$ represents variations of the mean shape $\phi_{\mathcal{M}}$.

Eq. (5.4) provides a variational framework for level set evolution. Note that, the above model is automated and requires no manually positioned landmarks. The use

of LSFs add numerous advantages to the model, and also restricts gaps or overlaps of the contour boundaries. Also since the distance functions are convex, the resulting subspace has convex curvature. Thus, optimization using gradient descent methods provide a global optimal segmentation result. Most importantly, the above model is flexible and can be generalized to multiple compartments, as in our case. We term this variant of Active Shape Models as *Active Geometric Models* (or AGM). Henceforth, we replace the term ASM with AGM. In sections 5.3 and 5.4, we generalize the above model for multimodal segmentation, and demonstrate its performance on D-OCT and MN compartment partitioning respectively.

5.3 AGM: application in 2D (HFL Thickness estimation)

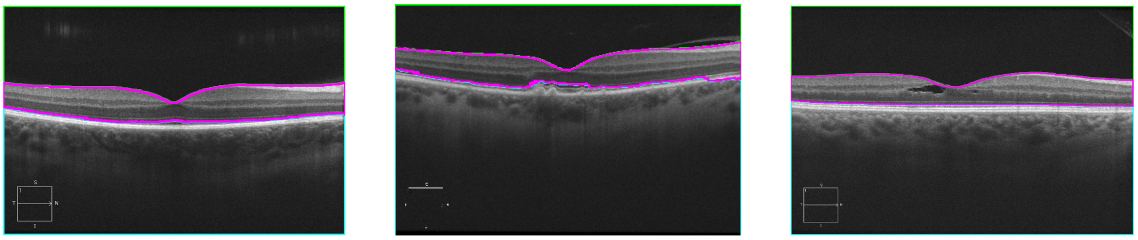
As discussed earlier, the image domain is divided into 3 compartments ($O_i, i \in (1, 2, 3)$). We represent each compartment by its *mcdf*, ϕ_i. The variations of each compartment is modeled independently. Thus,

$$\psi_i = \phi_i^{\mathcal{M}} + \sum_{j=1}^{m} \lambda_j^i U_j^i,\ i \in (O_1, O_2, O_3)$$

where, $\phi_i^{\mathcal{M}}$ is the mean shape of O_i, computed from Eq. (5.3). λ_j^i are the linear weight factors, and U_j^i are the eigenvectors corresponding to the m major modes of variations for O_i. In this case, the first two modes of variations represent the major part of the class (90%), while the third one (9%) accounts for non-symmetric properties of the retinal wall. Notice that, each *mcdf* ϕ_i can be decomposed into the distance functions φ_0, and φ_1, according to Eq. (3.8). Hence, it is not required to store LSF for each compartment separately. This gives faster convergence as well as less memory overhead.

Finally, we update ϕ_i at each iteration using the following objective.

$$E(\phi, A, \boldsymbol{\lambda}) = \int_{\Omega} \sum_{i=1}^{N} \xi(\phi_i)\big(s\phi_i - \phi_i^p\big)^2 d\Omega$$



(a) JC_0382_OD (b) ThDrusen_OD (c) Ahrens.Melinda_OD

Figure 5.2. Three cases of HFL Thickness estimation. Left: Normal eye; Center, Right: Pathological eye. HFL boundary is represented by magenta color, O_1 and O_3 compartments are denoted by green and blue colors respectively.

where, $\phi_i^p = \phi_i^{\mathcal{M}} + \sum_{j=1}^{m} \lambda_j^i U_j^i$. The model parameters are obtained using a similar gradient descent method, as described before.

To obtain better results, we update our model on a coarse-to-fine resolution. Initially, all the images are scaled by a factor $\frac{1}{2}$. After convergence was reached on one level, calculations were performed at the next finer level with doubled resolution of the original size. The distance transform subspace was also appropriately scaled. A few results are shown in Fig. 5.2. Detailed analysis and comparisons with other current methods are provided in Chapter 6.

Qualitative results illustrate that our proposed method works robustly, both for normal and pathological cases. Notice that the shape boundaries are disjoint and also cater to bending and twisting.

5.4 AGM: application in 3D (Motor Neuron morphology estimation)

Let, each training set consist of the aligned set $< \phi_S^n, \phi_A^n, \phi_D^n >$, $n = 1, \cdots N$, that is, the *mcdf*s of soma, axon and dendrites. As discussed before, we normalize the training set by computing the mean, or *average shape*, over all the training *mcdf*s.

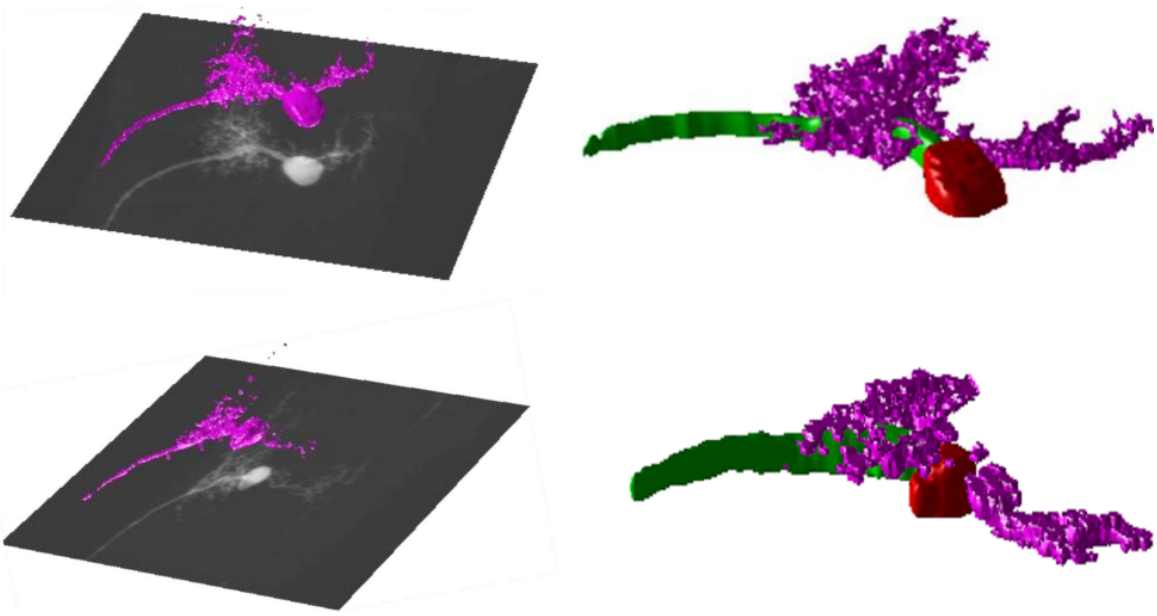


Figure 5.3. Neuron volume partitioning using AGM; rows correspond to two different datasets. Left column: collapsed-stack neuron images, with the volumes illustrating the intensity in 3D. Right column: resulting compartment sub-volumes.

Mathematically, for each compartment $i = \{\bar{S}, \bar{A}, \bar{D}\}$, we estimate the *mcdf* ϕ_i^M that minimizes the following objective,

$$E(\phi_i^M) = \sum_{i=1}^{N} \int_{\Omega} [\phi_i^n(\mathbf{x}) - \phi_i^M(\mathbf{x})]^2 d\mathbf{x}$$

$$s.t. \left|\nabla \phi_i^M\right|^2 = 1, \quad i \in (\bar{S}, \bar{A}, \bar{D})$$

Once the samples are normalized, the m most important modes of variations U are recovered through PCA,

$$\psi_i = \phi_i^M + \sum_{j=1}^{m} \lambda_j U_j, \;\; i \in (\bar{S}, \bar{A}, \bar{D})$$

Fig. 5.3 shows the neuron partitioning results of our Active Geometric Model. We provide more results and analysis of other subtypes in Chapter 6. These results show that our model is able to capture variances in the training data, and successfully evolves the mean shape contour to fit new test data.

In summary, in this chapter we provided a variant of ASMs using level sets as shape parameters. We also extended our model for multi-object segmentation scenarios. It is to be stressed here that our model is completely non-parametric and hence minimal manual intervention is needed for training. The use of level set functions augments our method and achieves faster convergence. Also, our method requires less memory overhead for multi-compartment partitions, and hence can easily be scaled to data in higher dimensions.

6 RESULTS

In this chapter, we provide detailed results and analysis of our proposed method with applications to *Henle's Fiber Layer* (HFL) thickness and *Motor Neuron* (MN) morphology estimation. We broadly divide this chapter into four sections. In section 6.1, we describe data acquisition methods for imaging Motor Neurons in *Drosophila*, as well as retinal B-scan imaging using Optical Coherence Tomography (OCT). In section 6.2, we provide both qualitative and quantitative of our registration framework. In section 6.3, we benchmark our AGM model against explicit ASM and multi-compartment level set evolution (MCGM). Finally, in section 6.4, we analyze the computational throughput of our method compared to those of ASM, MCGM and Snakes.

6.1 Data Acquisition

Due in large part to the simplicity and morphological stereotype of its neuromuscular system, *Drosophila* has served as an invaluable model in the study of motor circuit formation. Abdominal hemisegments in the embryo and larva comprise 30 highly stereotyped body wall muscles, each of which is innervated by one or more of an estimated 38 unique motor neurons. These motor neurons can be subdivided into distinct classes based on their association with specific nerve branches, and these subclasses can be further refined by morphological directions.

Our dataset consists of *12 distinct morphological classes* of motor neurons that elaborate morphologically similar dendrites and target functionally related muscle subsets. Despite morphological similarities, the origin and axonal projection patterns of these motor neurons have been well characterized, and provides reliable foundation for the

unambiguous identification of single motor neurons. The dendritic morphologies of individual motor neurons are typically difficult to resolve because of the vast number of different cell types in the larval CNS. To analyze dendritic arborization patterns of motor neurons in the larval CNS in detail, we used the mosaic analysis with a repressible cell marker *(MARCM)* [21] system to genetically label single motor neuron clones with a membrane-targeted GFP. Although labeled MNs generated by MARCM can be imaged live in the intact animal, muscle contraction by larva hinders the acquisition of confocal images through consecutive z positions. The brain and VNC are therefore exposed by dissection and the tissue is fixed with formaldehyde before immunostaining with antibodies directed against mCD8 and a secondary marker, Fasciclin II (FasII). FasII labels axon fascicles that divide the VNC into distinct territories and provides a frame of reference in which to map the relative positions of the MN soma and dendrites [6]. The entire morphology of single MNs is then imaged with laser scanning confocal microscopy that produces the image stacks to be analyzed. (Fig. 1.2). Here we use only the green channel for the estimation of the neuron morphology.

Healthy normal volunteers and patients with drusen related to non-exudative age-related macular degeneration (AMD) were imaged. All normal volunteers had visual acuity of at least 20/40, spherical error within ± 6.0 diopters (D), and no evidence of any other ocular diseases. The patients with non-exudative AMD were enrolled from the outpatient departments of Bascom Palmer Eye Institute. All eyes were scanned with the Cirrus OCT instrument by a single experienced technician. 200×200 raster scan patterns were used, each covering a 6×6 mm area on the retina with a homogeneous sampling grid of 200 horizontal B-scans with 200 A-scans per B-scan. The A-scans are separated by 30 microns in both the horizontal and vertical directions. Each participant had both eyes scanned. The scans were centered on the fovea based on the participant's fixation.

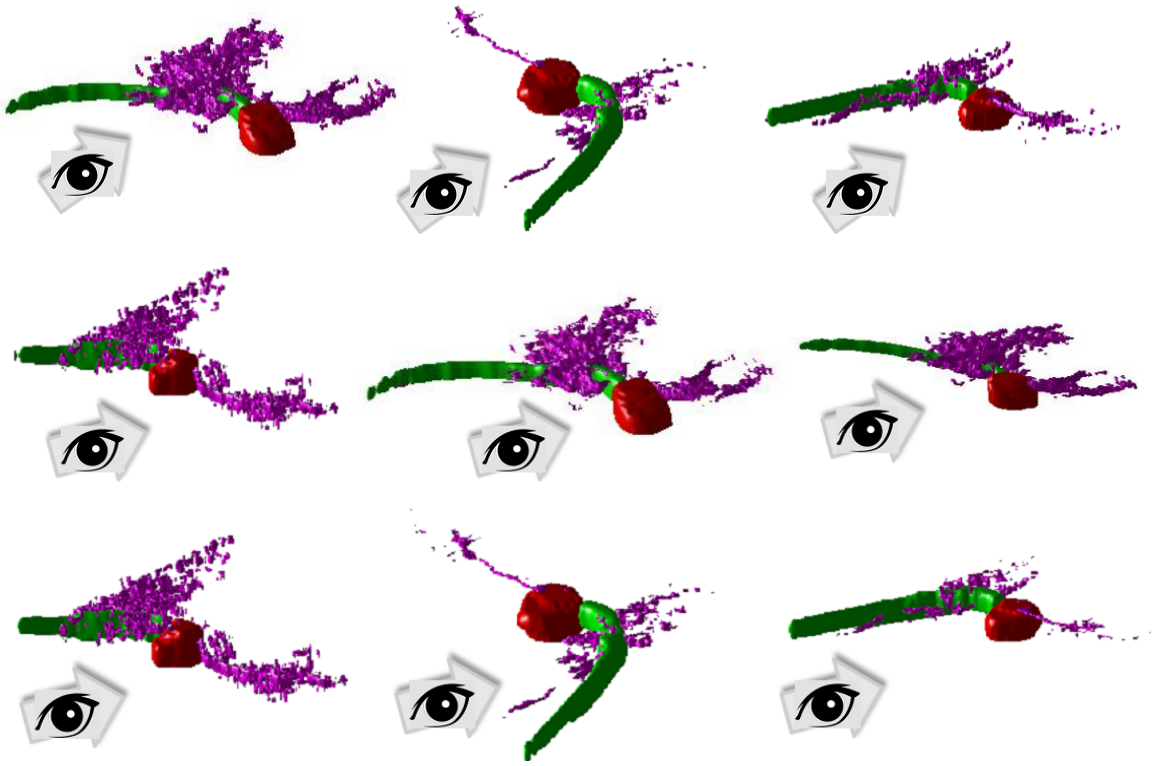


Figure 6.1. 3D Neuronal volumes. *Left:* Reference/Static($\mathcal{I}_R$); *Center:* Source/Moving($\mathcal{I}_S$); *Right:* Registered volume. The arrows indicate the view angle in which results are visualized. Red compartment indicates soma, green indicates axon and magenta indicates dendrites.

6.2 Image Registration

In this section, we illustrate performance of our model both for *rigid/global* registration, applied to align MN compartments, and *non-rigid/local* registration, applied to aligning retinal layer in D-OCT B-scans. These are described in order.

We applied our described variational registration method to register 12 pairs of labeled MN volumes. These neurons were chosen from three different subtypes - MN1, MN15/16 and MN9. The neuronal volumes were intensity normalized to within 0 and 255. Thereafter, the reference and source volumes were smoothened using a Gaussian filter of window size $10 \times 10 \times 10$. Fig. 6.1 show three such results. It is evident that our method successfully corrects global motion between reference and source vol-

umes. We remind the reader here that our registration framework, at this point, aims only for a global alignment between the compartments, so as to preserved individual topological characteristics of different subtypes.

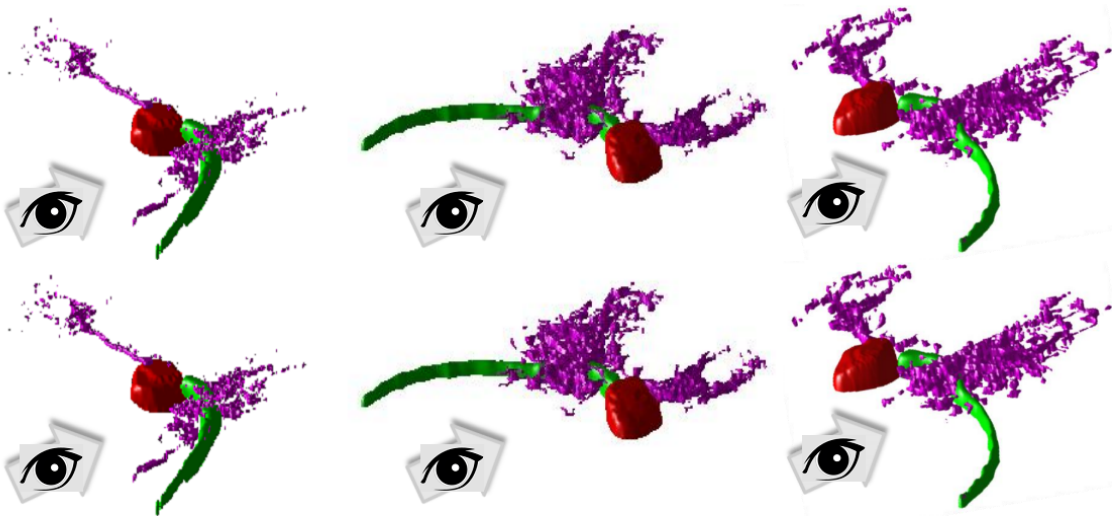
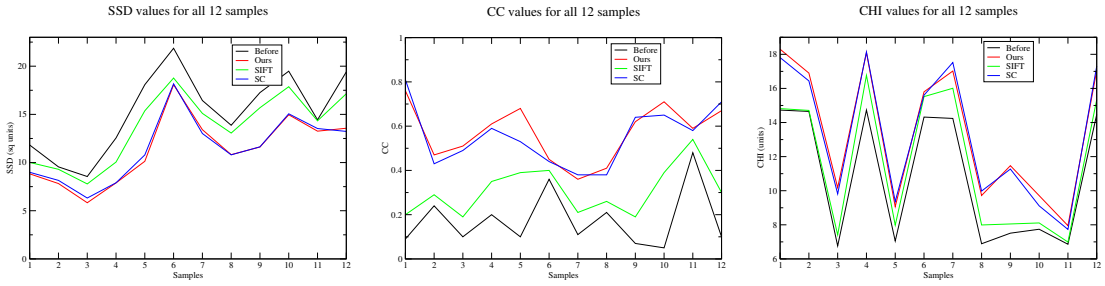


Figure 6.2. 3D Neuronal volumes registered using SC(*Top*) and SIFT(*Bottom*) as landmark points. We show output on the third pair from Fig. 6.1. Arrows indicate view angle. Red compartment denotes soma. green denotes axon and magenta denotes dendrites.

To emphasize the robustness of implicit shape representation in registration, we compare our method with state-of-the-art rigid registration methods using Shape Context (SC) [38] and SIFT [39] feature descriptors. SC descriptors are closer to level set representation in theory as they too work in the distance transform subspace. SIFT feature descriptors, on the other hand, account for edges, intensity distributions, texture and other features to generate the feature vector. Fig. 6.2 illustrates rigid registration using SC and SIFT feature vectors. Qualitatively, it is seen that registration using SC descriptors perform better than SIFT. While the results show efficiency of our framework, we perform detailed statistical analysis to verify our performance measure. A variety of similarity measures have been proposed to measure registration accuracy. For our experiments, we use Sum-of-Squared-Differences(SSD),

Correlation-Coefficient (CC) and Chi-squared distance (CHI). SSD and CC metrics have been defined in chapter 4. Chi-squared distance measure is defined as $\chi^2(X,Y) = \sum_{i,j} \frac{(X_{i,j}-Y_{i,j})^2}{X_{i,j}+Y_{i,j}}$ where X and Y are the normalized images, i and j correspond to pixel(i,j) of the image space. Fig. 6.3 proves that our proposed method works at par with SC based method, even out-performing SC in certain cases.



(a) SSD (b) CC (c) CHI

Figure 6.3. Mean metric values for 12 pairs of neuronal volumes are compared using our method, and rigid registration using Shape Context(SC) and SIFT feature matches.

Next, we demonstrate our method for the purpose of non-rigid registration of retinal layer in directional B-scans. The image pairs used in all the experiments were first preprocessed using a Gaussian filter with $W/5 \times H/5$, where W and H are the width and height of the reference image. α was set to 0.25, which gives a smooth deformation field. In fig. 6.4, we show two reference$(\mathcal{I}_R)$ images, one from a normal eye and another from a pathological eye. Two sets of source images, one from temporal displacement$(\mathcal{I}_L)$ and another in nasal displacement$(\mathcal{I}_R)$ are used for registration. Fig. 6.5 also show comparisons of our method with optical-flow based DEMONS algorithm and Free Form Deformation(FFD). Table 6.1 provides the quantitative analysis. It shows distributions of the pixel mismatches over Ω wrt different optimization criteria. Thus, if $\mathcal{I}$, $(|\mathcal{I}| = W \times H)$ is the indicator variable for pixels which differ in absolute intensity in Ω, then reported value $= \frac{|\mathcal{I}\neq 0|}{|\Omega|} \times 100$. Average improvement

using our method is $\sim 50\%$. More importantly, these results show that our method is robust to different optimization criteria.

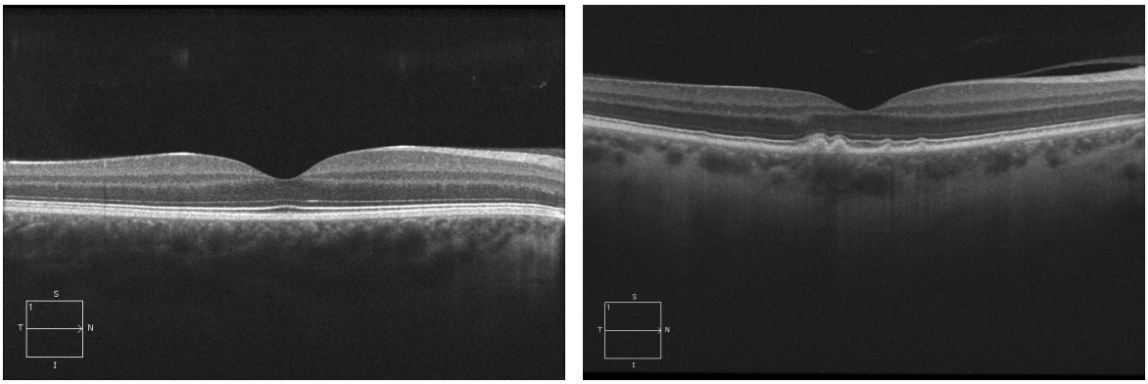
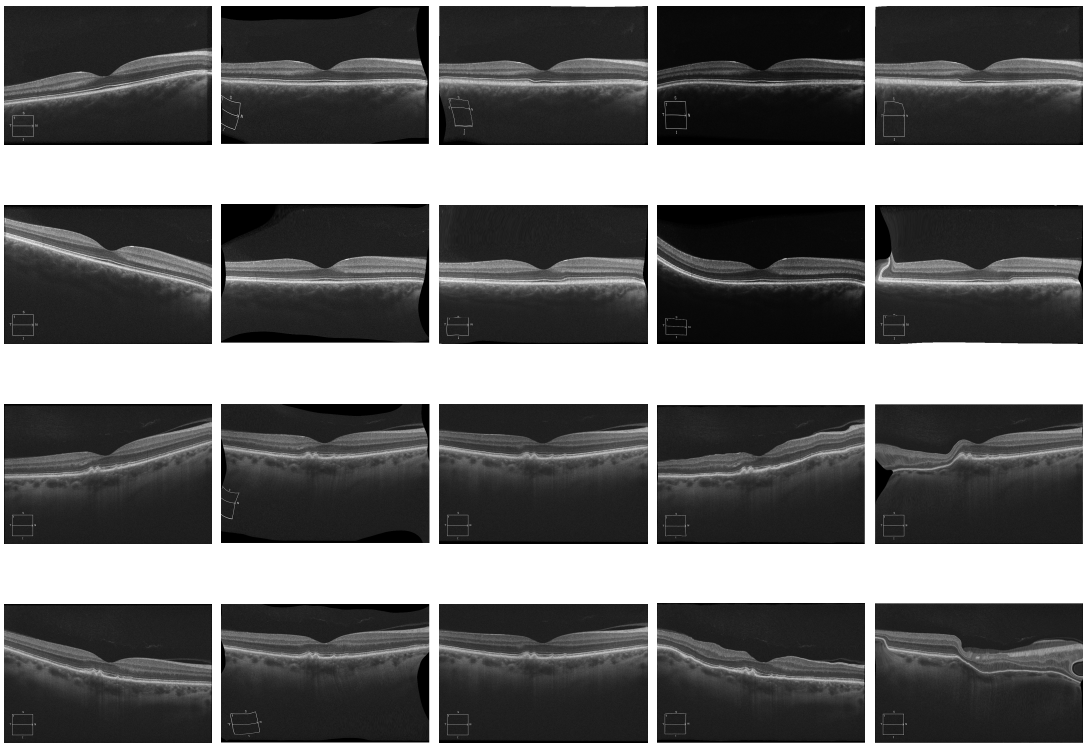


Figure 6.4. Two horizontal reference images in *central* entry position.(*Left*) Normal eye; (*Right*) Pathological eye.



(a) ORIG (b) SSD (c) NCC (d) DEMONS (e) FFD

Figure 6.5. Results of our method using SSD and NCC criteria. Also shown are results obtained using DEMONS and FFD. *Rows 1,2:* temporal orientation of normal and pathlogical eyes. *Rows 3,4:* nasal orientation of normal and pathological eyes respectively.

Table 6.1
Percentage of pixels differing in Absolute Intensities after registration using our method(criteria: SSD and NCC), DEMONS and FFD. Distributions are listed in %.

Type	Orientation	% of pixels differing in Intensity					
		Before Registration	**Our Method**			**DEMONS**	**FFD**
			SSD	NCC	Average		
Normal	*Temporal*	26.75	**10.12**	19.02	14.57	18.60	13.14
Normal	*Nasal*	39.47	**14.44**	25.11	19.78	21.81	26.56
Pathological	*Temporal*	44.56	20.03	**19.26**	19.65	42.09	41.34
Pathological	*Nasal*	47.81	**19.37**	20.69	20.03	43.24	45.61

6.3 Image Segmentation

In this section, we analyze the segmentation results of our model in Motor Neurons and Henle's Fiber Layer thickness estimation. We illustrate both qualititative as well as quantitative results of our method. For quantitative measurements, we validate the accuracy of our method pixel/voxel-wise, by defining as error percentage (%) of misclassified pixels with respect to the compartment area (2D) and volume (3D),

$$error = \frac{|ground\ truth - estimated|}{ground\ truth} \tag{6.1}$$

In the application of Motor Neurons, we segmented neuronal volumes from subtypes of the aforementioned 12 pairs. In Table 6.2, we report results on 6 of these cases. Here, we report *[average, worst-case]* errors over all compartments of final segmentation result.

Table 6.2
Numerical comparisons of our method with ASM and MCGMs.

Subtype	Ours	ASM	MCGM
MN1-lb	[7.4,11.3]	[9.4,16.9]	[15.1,47.9]
MN9-lb	[5.3,12.1]	[13.8, 24.6]	[56.3,77.2]
MN15/16-lb	[10.8,18.7]	[15.4,25.0]	[80.1,91.3]
ISNb	[15.9,25.9]	[25.5,46.9]	[35.1,67.9]
Average Error	**[9.9,17.0]**	**[16.1,28.4]**	**[46.7,71.1]**

Fig. 6.6 shows the qualitative analysis on four neuronal volumes. As anticipated, both the qualitative and quantitative results indicate the segmentation of the soma yields lower errors, due to the small shape variation, and relatively constant morphology. On the other hand, the intensity inhomogenities and random shapes of the dendrites make their accurate segmentation more challenging.

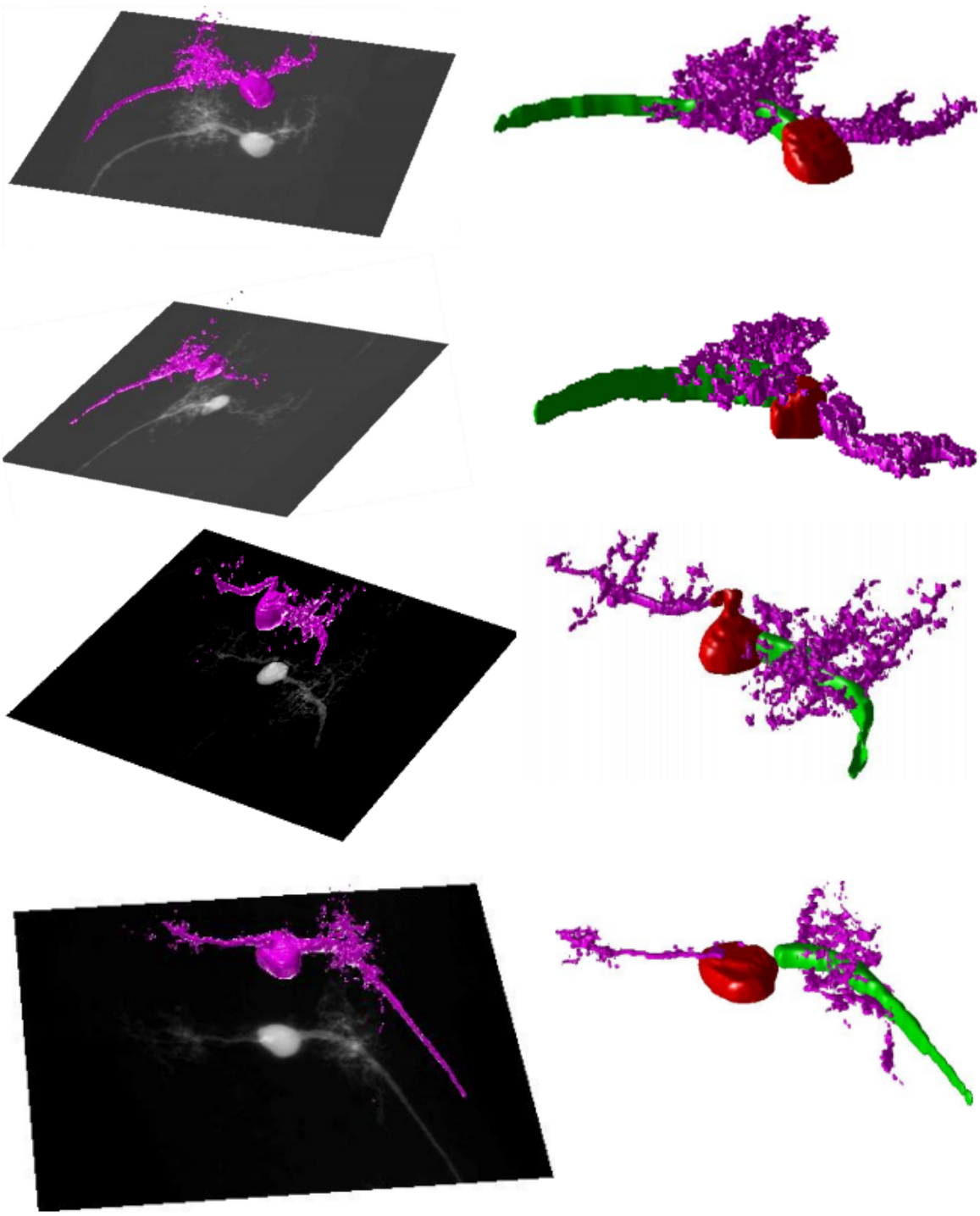


Figure 6.6. Results of our segmentation result on neuronal volumes from different subtypes. Rows 1: MN1-lb, 2: MN9-lb, 3: MN15/16-lb, 4: ISNb. Red colored compartment denotes soma, green color denotes axon, and magenta denotes dendrites.

In the application of HFL, we segmented 10 different retinal B-scans from normal and pathological eyes. Fig. 6.7 shows the qualitative results of our method compared with *Active Contour Models* (ACM) or Snakes [8], a popular segmentation method based on contour evolution. However, unlike AGM, the contour of each compartment is represented by a set of control points which evolve during each iteration.

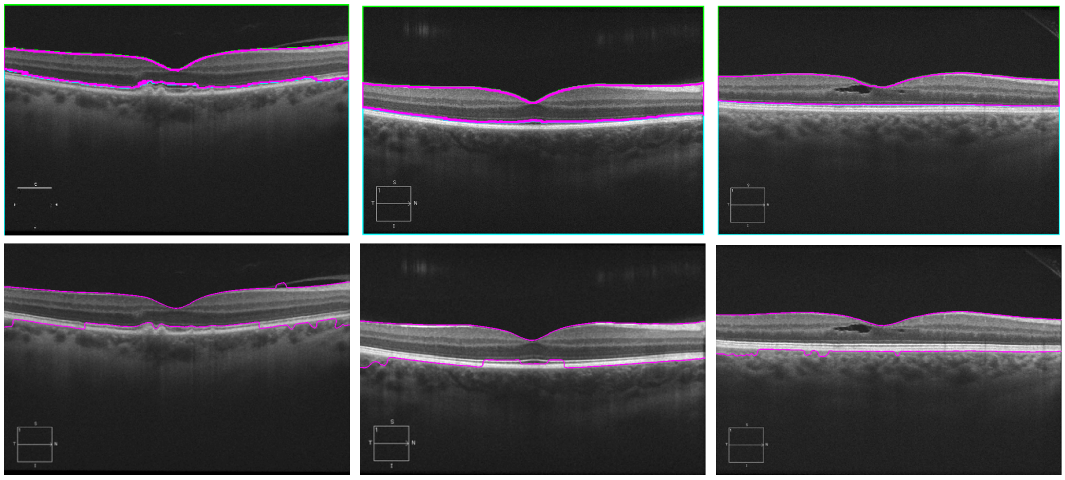


Figure 6.7. Results of our segmentation method(*Top*) and Snakes(*Bottom*) on D-OCT B-scans. Magenta colored compartment denotes the HFL.

Table 6.3 provides quantitative analysis of our method on normal and pathlogical eyes. It is to be noted here AGM converges faster than Snakes without compromising segmentation accuracy.

Table 6.3
Numerical comparisons of our method with Snakes. We measure the average error, defined in Eq. 6.1, in normal and pathological subjects.

Type	AGM	Snakes
Normal	4.3	7.1
Pathological	9.6	15.2
Average Error	6.9	11.1

6.4 Time Analysis

In this section, we compare the computational speed of our proposed AGM model with respect to other methods when applied to rigid and non-rigid registration. We performed all our experiments on an Intel Xeon Quad-core processor, 3.2 *GHz* and 12 GB *RAM*. Table 6.4 shows the convergence throughput of AGM against those of SC and SIFT based methods. It can be seen that our proposed method converges at a faster rate than ICP based algorithms on average. This confirms our claim of using level set functions, rather than feature points, as shape descriptors for better registration accuracy.

Table 6.4
Performance of AGM with respect to SC and SIFT based methods. Time reported in seconds.

DataSet	AGM	SC	SIFT	Volume Size(in pixels)
1	50 [s]	50.9 [s]	51.7 [s]	$1024 \times 1024 \times 21$
2	41.5 [s]	42.2 [s]	44.1 [s]	$1024 \times 1024 \times 14$
3	56 [s]	58 [s]	61 [s]	$1024 \times 1024 \times 25$
4	43 [s]	44.3 [s]	45.8 [s]	$1024 \times 1024 \times 15$
Average	**47.6 [s]**	48.5 [s]	50.7 [s]	

Next, we measure our performance in respect to non-rigid registration of retinal B-scan images. Here, we compare our method with DEMONS and FFD based implementations. Table 6.5 shows the results.

Table 6.5
Performance of AGM with respect to DEMONS and FFD methods. Time reported in seconds.

Type	AGM	DEMONS	FFD
Normal	14 [s]	12 [s]	18 [s]
Pathological	16 [s]	17 [s]	21 [s]
Average	15 [s]	14.5 [s]	19.5 [s]

7 SUMMARY

As part of my thesis work we have developed a *landmark* free model for concurrent image registration and segmentation of multiple objects in an image domain. This is, as shown in two applications, an important component for medical research in different domains. Our method is scalable, topology preserving and computationally fast as compared to other existing methods in this area. Moreover, the use of level set functions empowers our model with numerous advantages, like bending or twisting of shape contours. Also, the variational optimization method used in our work is robust, parameter free and guarantees convergence due to its convex curvature. With respect to segmentation, our method has strict shape constraints, which prevent overlaps or gaps in our result. Finally, our multi-compartment shape descriptors provides an accurate implicit representation of object morphologies, without any expert intervention.

The future work in this area would be to use more training data to generate our shape model, in order to capture variances of other subtypes. In our method, we focus on the Top-down model for segmentation. However, an interesting research problem would be to merge the Bottom-up and Top-down approaches in image segmentation, with respect to multi-phase segmentation. Thus, our model could be adapted to account for low-level cues in the image rather than generating a complex average shape model to encompass all possible variations. Parallelization of our method and/or performance on GPU hardware would be another future objective.

LIST OF REFERENCES

LIST OF REFERENCES

[1] M.D. Kim, Y. Wen, and Y.N. Jan. Patterning and organization of motor neuron dendrites in the drosophila larva. *Developmental Biology*, 336:213–221, 2009.

[2] L. A. Vese and T. F. Chan. A Multiphase Level Set Framework for Image Segmentation Using the Mumford and Shah Model. *International Journal of Computer Vision*, 50:271–293, 2002.

[3] S. Osher and J. A. Sethian. Fronts Propagating with Curvature Dependent Speed: Algorithms Based on Hamilton-Jacobi Formulations. *Journal of Computational Physics*, 79, 1988.

[4] L. Landmesser. The distribution of motor neurones supplying chick hind limb muscles. *Journal of Physiology*, 284:371–389, 1978.

[5] E. Vrieseling and S. Arber. Target-induced transcriptional control of dendritic patterning and connectivity in motor neurons by the ets gene pea3. *Cell*, 127(7):1439–1452, 2006.

[6] M. Landgraf, T. Bossing, G. M. Technau, and M. Bate. The origin, location, and projections of the embryonic abdominal motorneurons of drosophila. *Journal of Neuroscience*, 17(24):9642–9655, 1997.

[7] DM Gass. *Stereoscopic Atlas of Macular Diseases*. V Mosby Company, 1987.

[8] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *IEEE International Conference in Computer Vision*, pages 261–268, 1987.

[9] C. Goodall. Procrustes methods in the statistical analysis of shape. *Royal Statistical Society*, 53(2):285–339, 1991.

[10] D. Rueckert, L. Sonoda, C. Hayes, D. Hill, M Leach, and D Hawkes. Nonrigid registration using free-form deformations: Application to breast mr images. *IEEE Transactions on Medical Imaging*, 18:712–721, 1999.

[11] K.S. Arun, T.S. Huang, and S.D. Blostein. Least-squares fitting of two 3-d point sets. *IEEE Transactions on Medical Imaging*, 9:698–700, 1987.

[12] J. Montagnat and H. Delingette. A hybrid framework for surface registration and deformable models. *Computer Vision and Pattern Recognition, CVPR'97*, C-21:269–281, 1972.

[13] F. Leymarie and B. Kimia. The shock scaffold for representing 3-d shape. *IWVF*, pages 1061–1075, 1992.

[14] S. Rusinkiewicz and M. Levoy. Efficient variants of the icp algorithm. *International Conference on 3D Digital Imaging and Modeling*, pages 145–152, 2001.

[15] C. Chefd'Hotel, G. Hermosillo, and O. Faugeras. A variational approach to multi-modal image matching. *IEEE workshop in Variational and Level Set Methods*, pages 21–28, 2001.

[16] A. E. Johnson and M. Hebert. Recognizing Objects by Matching Oriented Points. In *Computer Vision and Pattern Recognition*, pages 684–689, 1997.

[17] H. Chui and A. Rangarajan. A New Algorithm for Non-Rigid Point Matching. In *Computer Vision and Pattern Recognition*, volume 2, pages 2044–2051, 2000.

[18] L. Mukherjee, V. Singh, and C. R. Dyer. Half-integrality based algorithms for cosegmentation of images. In *Computer Vision and Pattern Recognition*, pages 2028–2035, 2009.

[19] D. Comaniciu and P. Meer. Mean Shift: A Robust Approach Toward Feature Space Analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24:603–619, 2002.

[20] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham. Active Shape Models-Their Training and Application. *Computer Vision and Image Understanding*, 61:38–59, 1995.

[21] T. Lee and L. Luo. Mosaic analysis with a repressible cell marker (marcm) for drosophila neural development. *Trends in Neurosciences*, 24(5):251 – 254, 2001.

[22] W. Zhou, H. Li, and X. Zhou. *3D Dendrite Reconstruction and Spine Identification*. 2008.

[23] H. Peng, Z. Ruan, D. Atasoy, and S. Sternson. Automatic reconstruction of 3d neuron structures using a graph-augmented deformable model. *Bioinformatics*, 26(12):i38–i46, 2010.

[24] S. Schmitt, J. F. Evers, C. Duch, M. Scholz, and K. Obermayer. New methods for the computer-assisted 3-d reconstruction of neurons from confocal image stacks. *NeuroImage*, 23(4):1283–1298, 2004.

[25] Y. Al-Kofahi, N. Dowell-Mesfin, C. Pace, W. Shain, J. N. Turner, and B. Roysam. Improved detection of branching points in algorithms for automated neuron tracing from 3d confocal images. *Cytometry Part A*, 73A(1):36–43, 2008.

[26] S. Urban, S.M. OMalley, B. Walsh, A. Santamara-Pang, P. Saggau, C. Colbert, and I.A. Kakadiaris. Automatic reconstruction of dendrite morphology from optical section stacks. In ReinhardR. Beichel and Milan Sonka, editors, *Computer Vision Approaches to Medical Image Analysis*, volume 4241 of *Lecture Notes in Computer Science*, pages 190–201. Springer Berlin Heidelberg, 2006.

[27] A. Dima, M. Scholz, and K. Obermayer. Automatic segmentation and skeletonization of neurons from confocal microscopy images based on the 3-d wavelet transform. *Image Processing, IEEE Transactions on*, 11(7):790–801, Jul.

[28] C. Uehara, C. M. Colbert, P. Saggau, and I. A. Kakadiaris. Towards automatic reconstruction of dendrite morphology from live neurons. In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2004.

[29] Q. Li, X. Zhou, Z. Deng, M. Baron, M. A. Teylan, Y. Kim, and S. T. C. Wong. A Novel Surface-Based Geometric Approach for 3D Dendritic Spine Detection from Multi-Photon Excitation Microscopy Images. In *IEEE International Symposium on Biomedical Imaging*, pages 1255–1258, 2009.

[30] X. Yuan, J. T. Trachtenberg, S. M. Potter, and B. Roysam. Mdl constrained 3-d grayscale skeletonization algorithm for automated extraction of dendrites and spines from fluorescence confocal images. *Neuroinformatics*, 7:213–232, 2009.

[31] V. Kaynig, T. Fuchs, and J.M. Buhmnann. Neuron geometry extraction by perceptual grouping in ssTEM Images. In *Proceedings IEEE Conf on Computer Vision and Pattern Recognition(CVPR)*, pages 2902–2909, 2010.

[32] A.M. Seggio, A. Narayanaswamy, B. Rosyam, and D.M. Thompson. Self-aligned schwann cell monolayers demonstrate an inhereant ability to direct neurite outgrowth,. *J. Neural Eng.*, 046001, 2010.

[33] H. Zhao, T. Chan, B. Merriman, and S. Osher. A Variational Level Set Approach to Multiphase Motion. 1996.

[34] N. Paragios, M. Rousson, and V. Ramesh. Non-rigid registration using distance functions. *Computer Vision and Image Understanding*, 89:142–165, 2003.

[35] S. N. Le, J. Karlekar, and A. C. Fang. Articulated registration of 3D human geometry to X-ray image. In *Image Processing, IEEE International Conference*, pages 1108–1111, 2008.

[36] J. Thomadsen, U. Christensen, W. Soliman, and B. Sander. Enhancing the signal-to-noise ratio in ophthalmic optical coherence tomography by image registration—method and clinical examples. *Journal of Biomedical Optics*, 12, 2007.

[37] M. Rousson, N. Paragios, and R. Deriche. *Implicit Active Shape Models for 3D Segmentation in MR Imaging*. 2004.

[38] S. J. Belongie, J. Malik, and J. Puzicha. Shape matching and object recognition using shape contexts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24:509–522, 2002.

[39] D. G. Lowe. Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, 60:91–110, 2004.