AUTOMATED NEUROVASCULAR TRACING AND ANALYSIS OF THE

KNIFE-EDGE SCANNING MICROSCOPE INDIA INK DATA SET

A Thesis

by

WENJIE YANG

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,   Yoonsuck Choe
Committee Members,   John Keyser
                     Ergun Akleman
Head of Department,   Dilma Da Silva

December 2014

Major Subject: Computer Science

ABSTRACT

The 3D reconstruction of neurovascular network plays an important role in understanding the functions of the blood vessels in different brain regions. Many techniques have been applied to acquire microscopic neurovascular data. The Knife-Edge Scanning Microscope (KESM) is a physical sectioning microscopy instrument developed by the Brain Network Lab in Texas A&M University which enables imaging of an entire mouse brain at sub-micrometer resolution. With the KESM image data, we can trace the neurovascular structure of the whole mouse brain. For the large neurovascular volume like the KESM data set, complicated tracing algorithm with template matching process is not fast enough. Also, KESM imaging might involve gaps and noise in data when acquiring the large volume of data. To solve these issues, a novel automated neurovascular tracing and data analysis method with less processing time and high accuracy is developed in this thesis.

First, an automated seed point selection algorithm was described in my approach. The seed points on every outer boundary surface of the volume were selected as the start points of tracing. Second, a vector-based tracing method was developed to trace vascular network in 3D space. Third, the properties of the extracted vascular network were analyzed. Finally, the accuracy of the tracing method was evaluated using synthetic data. This approach is expected to help explore the entire vascular network of KESM automatically without human assistance.

DEDICATION

To my family.

ACKNOWLEDGEMENTS

First of all, I would like to thank my advisor, Dr. Yoonsuck Choe, for his constant support and encouragement throughout my research. I am deeply grateful of his earnest guidance and advices during the writing of this thesis. Second, I would like to express my gratitude to Dr. John Keyser and Dr. Ergun Akleman for their valuable comments on my research. Also, I would like to thank Chul Sung, Donghui Han, Wencong Zhang, Han Wang and Yue Zhuo for their great help. From the discussion with them I have benefited a lot for the preparation of my thesis. I also would like to thank my lab mates and friends at Texas A&M University for their support during my research work.

Finally, my deepest gratitude goes to my family for their loving considerations and great confidence in me all through these years.

TABLE OF CONTENTS

Page

LIST OF FIGURES

viii

LIST OF TABLES

1. INTRODUCTION

1.1   Motivation

The neurovascular system is a complex network. Understanding the neurovascular architecture of brain is helpful to understand the function of the blood vessels in different brain regions. To study the neurovascular architecture, many techniques and tools have been developed to acquire microscopic neurovascular data. The Knife-Edge Scanning Microscope (KESM) is a physical sectioning microscopy instrument developed in Brain Network Lab at Texas A&M University. It enables to scan through entire mouse brain tissues at sub-micrometer resolution. By using KESM, we have obtained whole-brain scale neurovascular data from mouse brain stained by India ink [2].

The next step after obtaining raw neurovascular data is to trace the vascular geometric structure in order to reconstruct the vascular network for further analysis. Since the KESM system offers large scale data set at high resolution, a high-performance 3D vascular tracing method is required. For the large volume of 3D neurovascular image, complicated tracing algorithm with template matching process is not fast enough. Also, KESM imaging might involve gaps and noise in data when acquiring the large volume of data. To achieve better performance, the tracing strategy should be robust in data with gaps and noise. In addition, the volume of vasculature data contains fine filament-like structures which should be traced to fully reconstruct the microstructure of the brain.

1.2   Goal of the research

The goal of this research is to develop a fast and robust seed point selection and vascular tracing method to reconstruct the neurovascular structure in the KESM

India ink data set. The tracing method would be able to trace efficiently fine vascular structures that span a large volume with gaps and noise. Also, the research targets towards analyzing the tracing result by collecting the length and the volume statistics of the traced vascular structure.

1.3 Approach

In order to automatically trace the vasculature in a certain unit volume, a series of steps are taken in the tracing method: image preprocessing, seed point selection, vector-based vasculature tracing, bifurcation and curve detection, vascular network exploration and data analysis. (1) Preprocessing is applied to the KESM India ink data including stacking the 2-dimensional image sequence and thresholding the images. The 3-dimensional vascular volume file is created by stacking the 2-dimensional image sequence. The image is converted to binary image using Otsu's thresholding method [13]. (2) Seed point selection is performed on each outer surface of the vascular volume. The seed points are generated by selecting the centroids of vessel cross sectional regions on the six outer surfaces of the cubic volume. Since the seed points are used as the starting points for tracing, the initial tracing direction of each seed point is given based on which surface the seed point is located. (3) A novel vector-based vascular tracing method is applied by using a virtual flat box to observe the vessels and to determine the subsequent points on the medial axis of vessel. The flat box is aligned based on the coordinates and initial tracing direction of each seed point. The centroid of vessel cross sectional region on the predicting window of the flat box is determined as the next point on the centerline. The vessel is traced by estimating the centroids of vessel cross section while traversing the path of the flat box. Also, the predicting window of the flat box can be dynamically expanded to hold vessel cross sectional region with increasing radius. (4) Vessel bifurcation,

2

inner curve, and gaps are also handled in the approach to ensure the robustness of the method. The flat box detects bifurcation point while tracing the vessel and records the branch structure when bifurcation occurs. If a vessel with double turns of different directions exists in the volume, the flat box tracing method detects the position of the double turn and changes tracing directions to trace this kind of vessel. Moreover, due to the imaging method of KESM system, there might be gaps between image slices. The flat box tracing method skips over the small gaps while tracing. (5) Vascular network exploration is performed based on the results from the processes above. To explore the vascular network, breadth first search is applied to record the vascular network architecture in the volume. The tracing process is repeated on X, Y, Z axis directions to make sure there is no missed vessel. (6) The properties of traced vascular are analyzed next. In the tracing process, the areas of vessel cross sectional regions are recorded and the radius of the cross section is estimated based on the area. The length of vessel is computed as the sum of the distance between the two neighboring centerline points. The volume of the vessels are estimated at the same time.

After vascular tracing is complete, the traced vascular structure is overlaid with the original volume data using a visualization software (Paraview). The accuracy of the tracing method as well as the processing time are quantitively analyzed.

1.4   Significance

This thesis presents a novel vector-based vascular tracing method with automated seed point selection. The vascular tracing method reduces the processing time and maintains high accuracy. The method can explore the entire vascular sub-network automatically without human assistance. It can also trace different structures of vessels in the KESM India ink data set that contains image gaps and noise. This

thesis is expected to contribute to the analysis of vascular network in larger brain image data sets.

1.5 Outline of the thesis

This thesis is organized as follows. In chapter 2, background on the KESM system and KESM India ink data set will be introduced. In addition, related work about automatic seed point selection and tracing algorithms will be discussed. In chapter 3, I will explain the methodology of the thesis, including preprocessing, seed point selection, tracing method, and data analysis. In chapter 4, experimental results and quantitive analysis will be presented. Finally, in chapter 5, the discussion about the experimental results, open issues, future works and conclusion will be presented.

2. BACKGROUND AND RELATED WORK

Research about the microstructure of mouse plays an important role in neuroscience as it helps infer the function of the brain and medical conditions. To explore the neuronal and vascular structure of the brain, researchers have developed different techniques and instruments to acquire brain image data. The Knife-Edge Scanning Microscope (KESM) [9] is one of such instruments that was developed in the Brain Networks Laboratory (BNL) at Texas A&M University to acquire high-resolution large-scale brain image data. In the beginning of this chapter, I will briefly review the technique of KESM. By using the KESM, we have acquired mouse brain data stained with different methods, including vasculature data. The KESM data will then be introduced. Once the whole volume brain data is obtained, the next step is to extract the microstructures in the data and to reconstruct the vascular network from the images. Different methods exist for tracing filament-like structures. I will present this part in the related work.

2.1 Background

In this section, the KESM system and the data it collected are briefly introduced.

2.1.1 Knife-Edge Scanning Microscope

The KESM was initially designed as a brain tissue scanner by Bruce H. McCormick at BNL. It is a unique instrument that combines physical sectioning and imaging so that brain tissues can be scanned at whole-brain scale as well as in sub-micrometer resolution. The structure of the instrument is shown in Figure 2.1.

As shown in the figure, the air-bearing stage is controlled to provide ultra-precise mechanical movement in three-axises with 20 nm precision in x and y axises, and
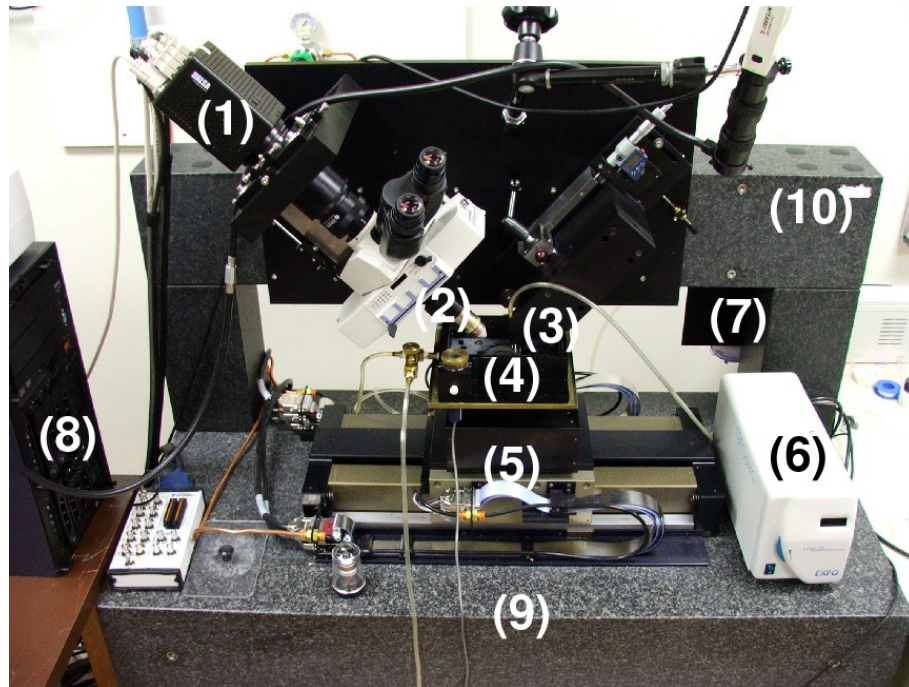
Figure 2.1: A photo of the Knife-Edge Scanning Microscope (KESM). (1) high speed line-scan camera, (2) microscope objective, (3) diamond knife and light collimator, (4) water immersion specimen tank, (5) air-bearing three-axes precision stage, (6) white-light illuminator, (7) water pump (in the back), (8) computer server installed with image acquisition and stage controller boards, (9) granite base, and (10) granite bridge. Adapted from [1].

25 nm in z axis [12], which provides solid registration for KESM imaging. One of the key features in KESM is the custom knife-collimator assembly. It is specially made to serve a dual role: a white-light collimator and a diamond microtome. The diamond knife is rigidly mounted to a massive granite bridge over the three-axis stage. On the other side, a microscope objective is held over the stage. Through the microscope, a high-speed line-scan camera can image the brain tissues at sub-micrometer resolution. A closeup of KESM is shown in Figure 2.2.

The operational principle of KESM is to simultaneously section and image tissue blocks. Details are shown in Figure 2.3. A brain tissue, embedded in hard polymer
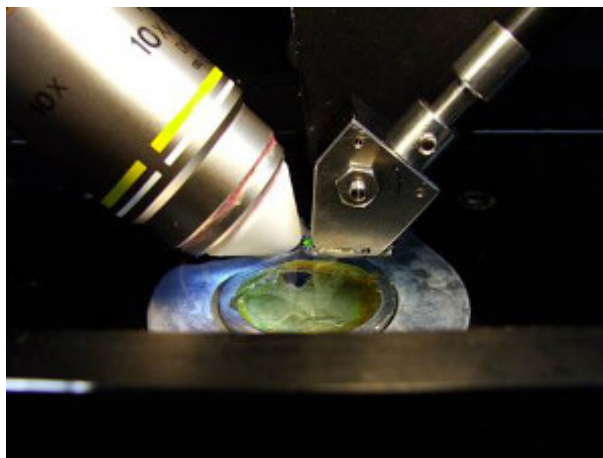
Figure 2.2: Close-up of KESM

resin block, is fixed on the specimen tank immersed under water. The microscope objective and diamond knife are held in place and adjusted to align. Instead of moving the diamond knife, the specimen tank is controlled by the stage to move agains the knife (direction of green arrow), generating a very thin (about 1μ m) piece of tissue at the tip of the knife where distortion can be minimized. Meanwhile, the white-light illumination source provides light through the diamond knife (light path highlighted by white arrow), illuminating the tissue piece on the knife tip. When perfectly aligned, the camera scans the tissue piece through microscope objective and transfers the image signal to the computer server.

Due to limitation of the field of view and width of the knife, the whole tissue block cannot be sectioned in one sweep. Stair-step sectioning [8] was developed to overcome this limitation. The basic idea is to section part of tissue block in a pre-defined order such that the block can be fully sectioned and the tissue block does not bump into the objective. With this technique, imaging time for a $1cm^3$ cube at $0.6\mu m \times 0.7\mu m \times 1.0\mu m$ is about 100 hours [4].
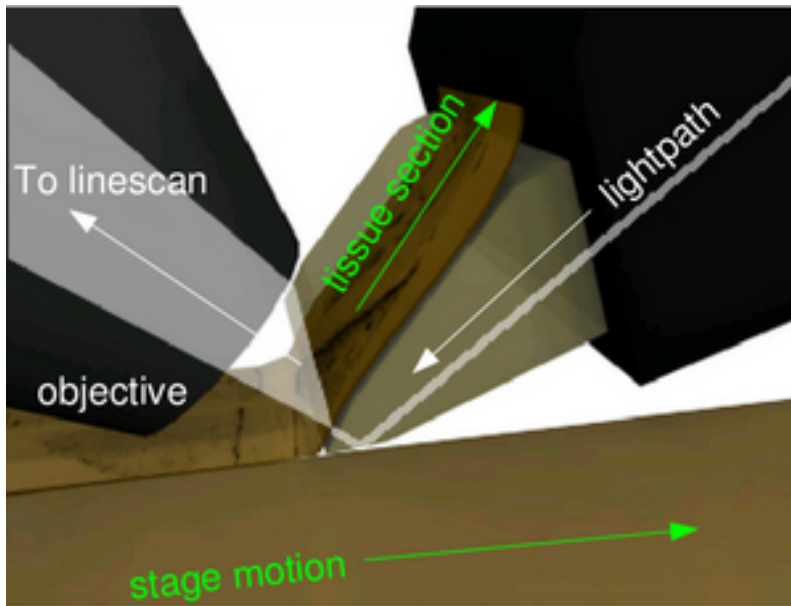
Figure 2.3: Imaging principle

2.1.2   KESM data

By using KESM, we have successfully acquired image data from the mouse brain. To be specific, we obtained mouse brain data stained in Golgi (neuronal morphology) and India ink (vasculature) in 2008 and reported in [4, 5]. In addition, mouse brain stained in Nissl, which reveals the distribution of cell body, was acquired in 2010 [3]. These data provide insights into the brain microstructure at the system-level. The data are also processed and visualized so that they can be accessed online through the KESM Brain Atlas [5].

In my work, I mainly focus on India ink data, which shows the vascular network of the mouse brain. Figure 2.4 shows the vasculature data set.

The data shown below was processed to show a three-dimensional view. The KESM data is basically a large 3D image stack consisting of 2D images. The three-dimensional visualization methods help us better explore the KESM data.
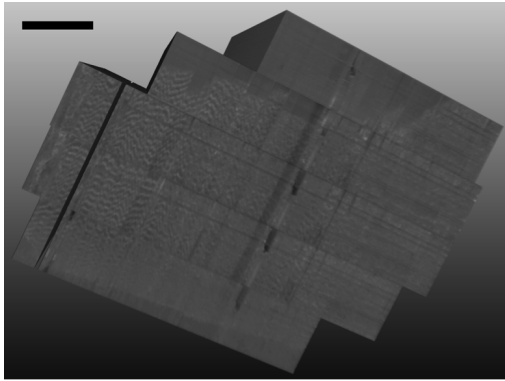
(a) Raw data volume(bar = 1.44 mm)
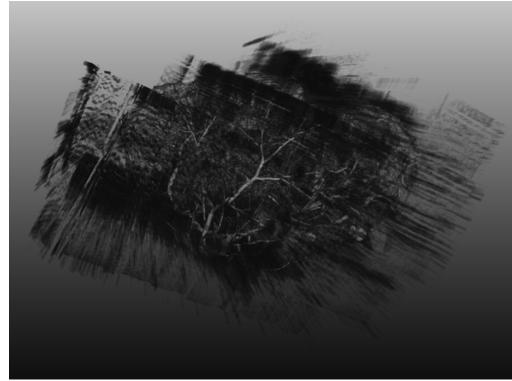
(b) Initial thresholding of (a)
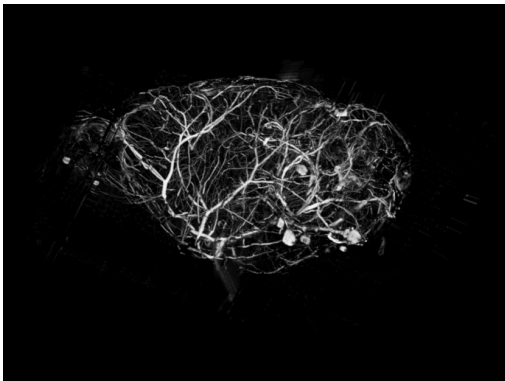
(c) Sagittal view.
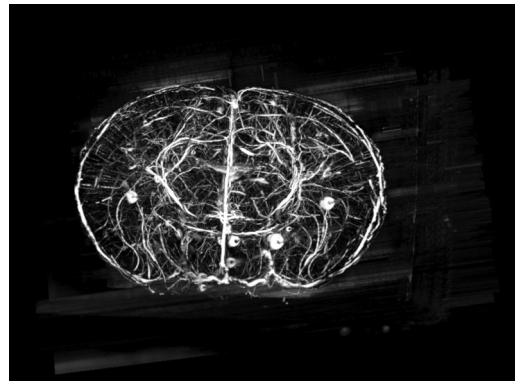
(d) Coronal view.
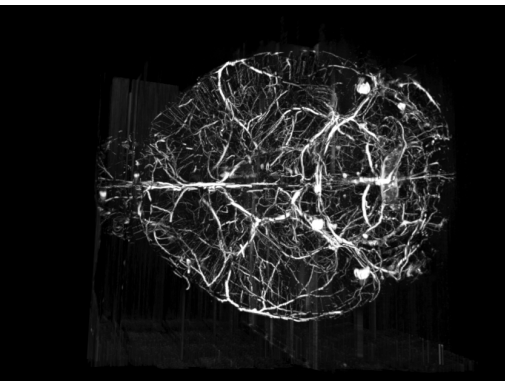
(e) Horizontal view.
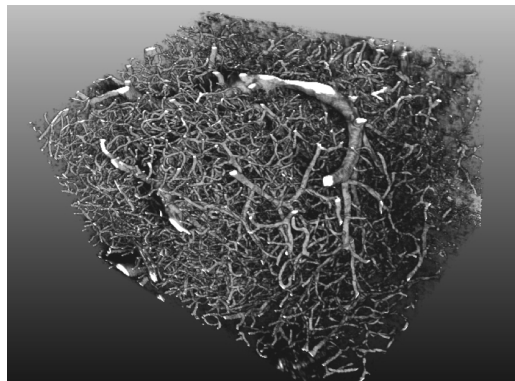
(f) Close-up (1.5 mm-wide block).

Figure 2.4: Different visualization of KESM vasculature data. (a) A raw data block in a sagittal view. (b) is generated from (a) where the data block was lightly thresholded. The boundary of the raw data and the content starts to appear. (c)-(e) are fully thresholded version of (a) but in different views(sagittal, coronal, and horizontal). The whole brain shape can be seen. (f) Intricate details within a 1.5 mm-wide block. Adapted from [4].

2.2   Related work

After obtaining the image data across the whole brain, it is necessary to extract the complete microstructure of the data and reconstruct the vascular network from the raw data. Different techniques have been developed to trace the structure of the vasculature data. Most of the methods include seed point selection as the first step toward tracing. In this section, I will review the related works on seed point selection and tracing.

2.2.1   Automatic seed point selection

For tracking tasks in medical image processing, seed points are defined as the initial position for tracking. Given the seed point within the boundary of vasculature, we can extract the structure by region growing based on some predefined criteria [15]. In the simplest form, the seed point requires manual selection to choose the proper position. However, to make the tracking automatic we can develop an automated seed point selection algorithm to determine seed point.

Xiao et al.[18] determined seed points based on the idea of gradient and Hessian matrix. Before selecting the seed points, they tried to suppress the background and highlight vascular structure in the image by using ridge enhancement. Next, the seed points were determined as the local maximum points in the gray level of the image within the boundary of the vasculature. A local maximum point's intensity was brighter than other points in the local area. Its gradient should be equal to zero and its Hessian matrix should be negative. Since there were still a lot of noise with low intensity values in the image, the extracted seed points were refined by an intensity threshold, where points with lower intensity values were discarded.

Shan et al.[17] developed an automatic seed point selection method for region growing algorithm in breast ultrasound image segmentation. First, they used speckle

10

reduction method to reduce noise and preserve edges. They then iteratively selected thresholds based on the histogram and breast lesion's spatial characteristics. They deleted the boundary-connected region and ranked the remaining region. The remaining region was scored by considering the area, variance, and distance to the center. The region with the highest score was chosen to be the lesion. Finally, they calculated the center of the region as the seed point. They also considered the cases where the lesion shape was irregular.

Zhang et al.[19] proposed a seed point selection method in extracting neurite structures in microscopy images of neurons. The method consists of two steps. In the first step, the initial points are detected by searching the local maxima. Instead of searching all the points, they only focus on a set of grid lines. Pixels on the lines are processed by a low-pass filter to remove high-frequency components and to ensure the correctness of local maxima. However, there is a possibility that redundant points are selected due to noise and image artifacts in the first step. In the second step, candidate points are selected by the criteria of signal-to-noise ratio. Points whose intensity values are greater than a threshold are considered reliable. The method ensures the completeness of the seed points while one issue is that redundant points are generated. Additional steps are required to take care of the redundant tracing.

In summary, seed point selection methods in the related work are developed for different data sets which is not suitable for KESM vascular data set. The data set obtained from KESM brain atlas in this research is already preprocessed. Thus, the vessel cross sectional region in the data set has clear boundary. In addition, the method of finding local maximum point might result in redundant seed points which needs to be removed before using them in the next vascular tracing step. Also, the computation of gradient and Hessian matrix in finding local maxima method is relatively time consuming.

2.2.2   Tracing algorithms

Han et al.[7] developed a Maximum Intensity Projection (MIP) based 3D vector tracing algorithm. First in 2D tracing, they implemented a moving window method to select the fiber direction. They also include "Cubic Tangential Trace Spline" in the tracing method to make it robust to branching. Then, making use of available 2D tracing method, the algorithm extended the approach into a 3D tracing method. First, at each seed point they determined the three local axis length by boundary detection. After obtaining the three local axis lengths, a local volume was generated. The local MIP along the longest axis length had little information about vessel direction. Therefore, only the two other MIPs were used to determine fiber direction [6]. A multi-scale filter applying the Hessian matrix was used for fiber direction detection in the 2D image. When maximum intensity of the next segment was less than a threshold or the next candidate point came across previous trace results, the tracing terminated [7]. An overview of the method is shown in Figure 2.5
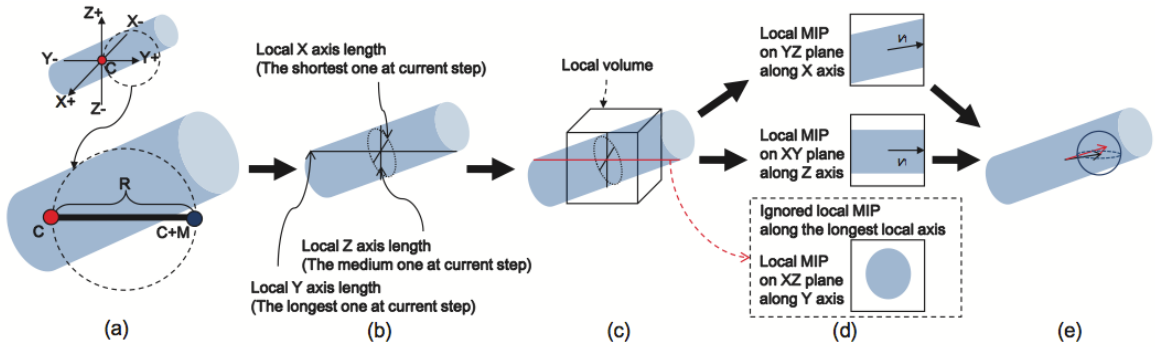
Figure 2.5: Overview of the tracing steps. (a) Direction detection in $y+$ axis. (b) Caculating three local axis lengths to adjust trace point. (c) Based on medium axix length, local volume is estimated. (d) Local MIP and 2D tracing results. (e) Evaluation and adjustment on 3D vessel directio using the Hessian matrix. Adapted from [6]

Palagyi et al.[14] extracted the centerline of vasculature in medical images using a fast 3D curve-thinning algorithm. The curve-thinning method was an iterative object reduction method preserving line-end points which can directly produce one-voxel-wide centerlines. In the curve-thinning algorithm, one iteration was decomposed into six sub-iterations aligned with the six main orthogonal directions in 3D. In each sub-iteration, there were two phases: (1) the border points of the actual type that were "simple"[1] and not line-end were marked potential deletable, (2) in the re-checking step, the potential point was deleted if it remained simple and not line-end after removing some previous visited points.

Mayerich et al.[10] presented a predictor-corrector algorithm to trace filament structures in volume data. They extracted the centerline of filaments by determining the path in a particle called tracer. To predict next estimated tracer position, they sampled a series of vectors that laid within a solid angle of the previous vector. For each vector, cross-sectional image orthogonal to the vector was taken. Then they chose the vector with the cross-sectional image that most closely matched the contrast of template to be the next tracing direction.

In summary, Han et al.'s method is not efficient enough for tracing large vascular volume. In the method, two 2D MIPs are processed with complicated computation in each iteration. The processing time of a $128 \times 128 \times 128$ volume was about 200 seconds [4]. Palagyi et al.'s method cannot trace the vessel with small gaps in vascular data like KESM. Mayerich et al.'s method involves template matching process which is time consuming when tracing large vascular volume. Thus, a novel vascular tracing with high efficiency and accuracy is necessary. Also, in all case, seed point selection was done manually.

---

[1]A voxel is called a simple point if its deletion does not alter the topology of the structure.

3. METHODS

In this chapter, the methodology of this thesis is described. There are mainly eight steps for conducting the experiment: preprocessing, seed point selection, automated vector tracing, bifurcation detection, vessel curve detection, avoiding gaps, vascular network exploration, and data analysis. Each step is discussed in detail in the following sections.

3.1 Preprocessing of the KESM India ink data

The vascular data in this research was obtained from the KESM brain atlas [5]. The raw images of the KESM India ink data were acquired from the coronal slice of a mouse brain perfused with India ink to label the vascular network [11]. The vascular image from the KESM brain atlas is preprocessed from the raw image with image inversion and Gamma correction to increase the image quality [5].

3.1.1 Stacking image sequence to 3D volume file

In order to perform vector tracing in 3D, we need to create a 3D vasculature data set using the images downloaded from the KESM brain atlas. In the mouse brain image sequence, we cropped the area with the same coordinates for every slice then stacked the cropped 2D image sequence to a small 3D volume file. Figure 3.1 illustrates an example of stacked 3D India ink vascular volume file.
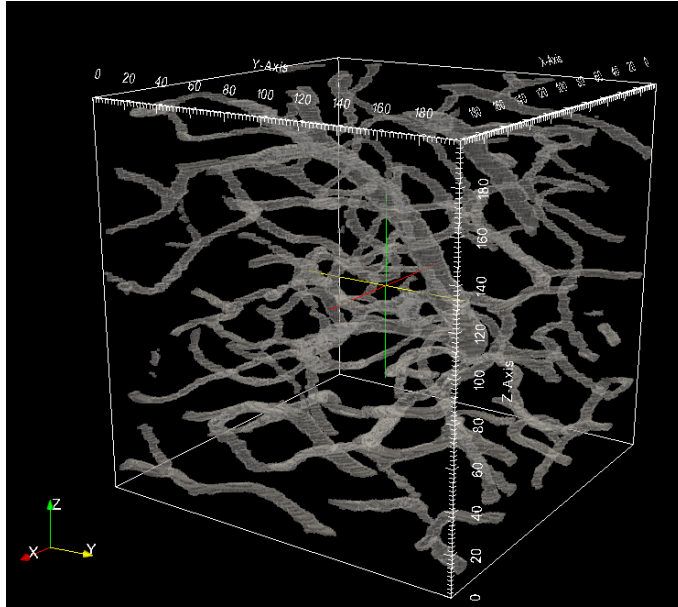
Figure 3.1: Stacked 3D vascular volume data set

3.1.2 Deciding the image threshold

In seed point selection and automated vector tracing, an image threshold is needed to segment the vessel cross sectional region and the background based on the intensity value of the pixels. To obtain the image threshold, Otsu's thresholding method was applied [13]. Otsu's method automatically searches for the optimal threshold that minimizes the intra-class variance of the thresholded black and white pixels. Since the vascular image data is bimodal, after computing the gray level histogram of the image Otsu's method finds a value of threshold which lies between the two peaks of histogram such that variance of both classes are minimized. Based on the optimal threshold obtained by the Otsu's method, the image is converted to binary image. Figure 3.2 shows an original vessel cross sectional image and the binary image after thresholding. In this image, the result of image threshold is 129.
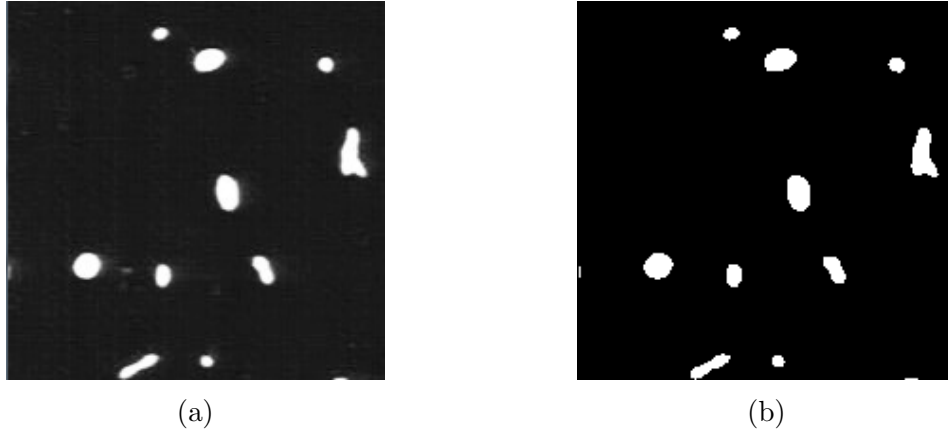
Figure 3.2: Deciding the image threshold. (a) The original vessel cross sectional image. (b) The image after thresholding.

3.2 Seed point selection

Once the 3-dimensional vasculature volume is generated, the six outer surfaces of the volume data cube are obtained. On each surface, we define the centroid of each vessel cross sectional region as the seed point for tracing within the volume. The surface image of the vascular volume is viewed as a 2-dimensional image represented by x and y coordinates.

3.2.1 Vessel cross sectional region detection

To calculate the centroid of vessel cross sectional region, the connected component labeling algorithm is applied [16]. The algorithm scans all the pixels in the surface image and groups the connected pixels with intensity value greater than the threshold obtained by Otsu's method in section 3.1.2 to detect the vessel cross sectional regions. Connected component refers to the pixels within the boundary of one vessel. There are mainly two steps in the algorithm. Firstly, the algorithm scans the eight neighboring pixels around one pixel and record the connected pixels. Secondly, after scanning all the pixels in a image, the algorithm groups the connected pixels.

These connected pixel groups represent the vessel cross sectional regions. Figure 3.3 shows the vessel cross sectional regions detected by the connected component labeling algorithm.

| (a) | (b) |
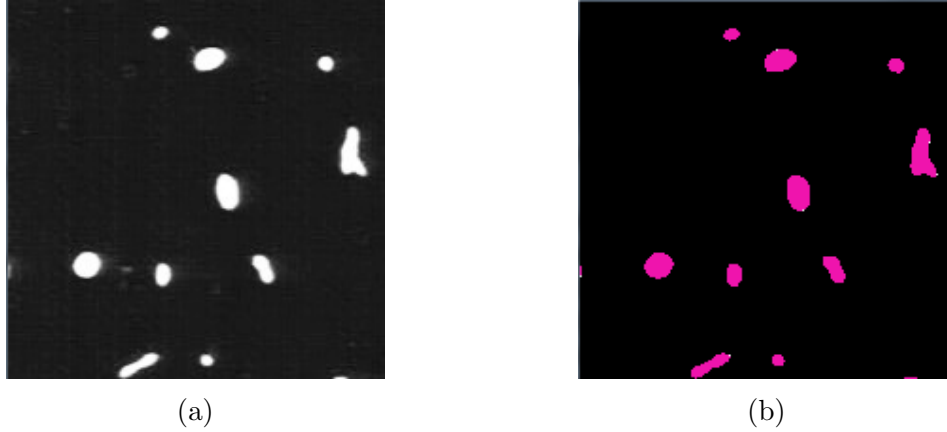
Figure 3.3: The vessel cross sectional region detection. (a) The original vessel cross sectional image. (b) The pink regions mark the detected vessel cross sectional regions.

3.2.2   Centroid detection

After grouping the pixels to determine the vessel cross sectional regions, the next step is to find the centroids of all the regions as seed points. In the KESM India ink data, most of the cross sectional regions are circular or ellipsoidal. To find the centroid of the vessel cross sectional region, a minimum rectangular window bounding the region is used. It is determined by the two coordinates in the group, which are (X_{min}, Y_{min}) and (X_{max}, Y_{max}).

For the most circular cross sectional regions, the seed point can be considered as the center of the minimum rectangular window, i.e. $((X_{min} + X_{max})/2, (Y_{min} + Y_{max})/2)$. However, there are also some cross sectional regions that have an irregular shape. For the irregular shaped cross sectional regions, it is necessary to avoid

17

selecting a point outside of the vessel boundary as the seed point. Therefore, the coordinate of the seed point is selected by the following rules:

1. The x axis coordinate of the seed point is selected as the center of the left and the right boundary of the region as in the common case, i.e. $X_{seedpoint} = ((X_{min} + X_{max})/2)$.

2. Given the x axis coordinate of seed point $X_{seedpoint}$, the pixels with the same x coordinate, $x = X_{seedpoint}$, are scanned from top to bottom in the rectangular window. Then the connected pixels greater than the image threshold are grouped into vertical line segments. The middle point of the longest connected line segment is selected as the seed point.

Figure 3.4 shows the seed point detection process with regular and irregular shapes.

(a)   (b)
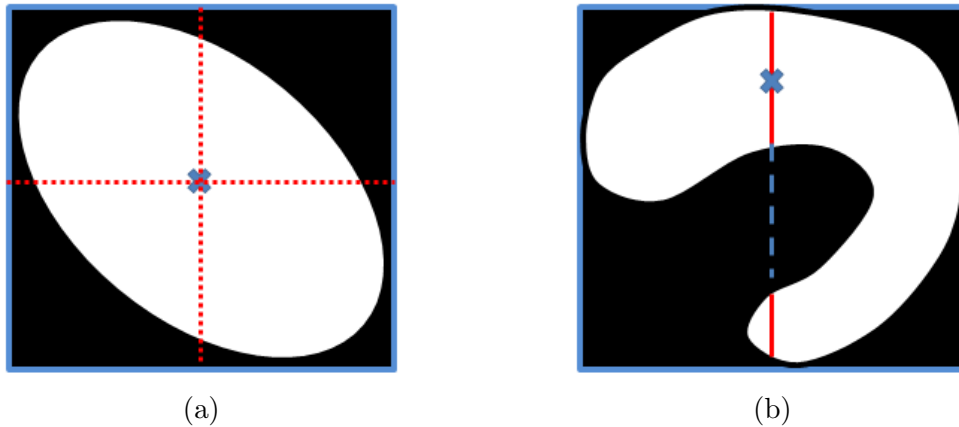
Figure 3.4: Seed point detection in vessel cross sections with ellipsoided (a) and irregular (b) shapes.

Figure 3.5 shows the seed points selected on one of the volume surfaces.

<div align="center">(a)   (b)</div>
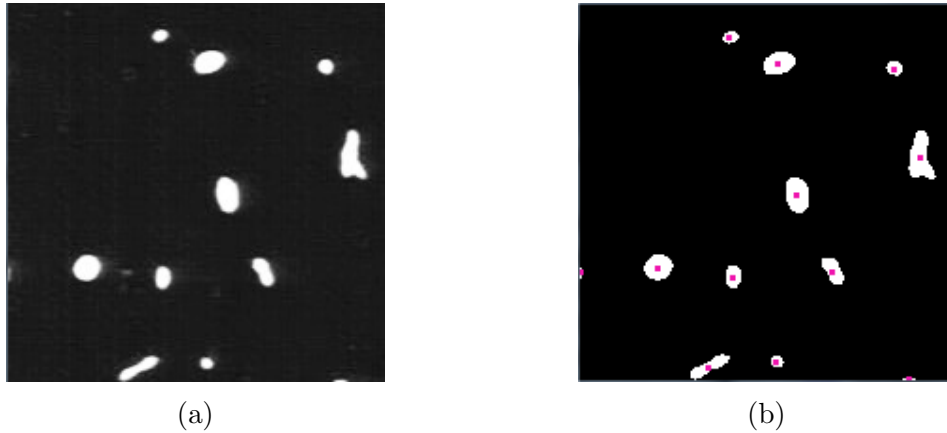
Figure 3.5: Centroid detection. (a) Original vessel cross sectional image. (b) The pink dots mark the centroids of the cross sectional region. These centroids are selected as seed points.

Each seed point is also given an initial tracing direction based on which surface of the cubic volume it is located. The initial tracing directions are illustrated in Figure 3.6.

3.3 Vector tracing method

The vector tracing method starts from a set of seed points and traces the medial axes of the vasculature from the initial points recursively until predefined termination criteria are satisfied. Instead of exploring every pixel in the whole volume, the vector tracing method only explores the pixels close to the vascular structure and analyzes local image properties. Therefore, the vector tracing method is particularly appropriate when processing speed is an important factor. In data sets like the KESM india ink data set, we can make use of vector tracing to achieve high-speed processing.
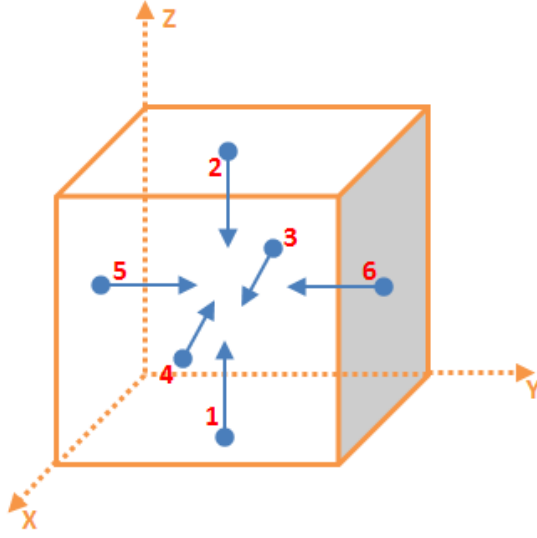
Figure 3.6: The initial tracing directions of seed points. The initial tracing directions of seed points on the bottom and top surface (1, 2) are z+ and z- respectively. The initial tracing directions of seed points on the back and front surface (3, 4) are x+ and x- respectively. The initial direction of seed points on the left and right surface (5, 6) are y+ and y- respectively.

3.3.1 The moving flat box tracing method

The vector-based tracing method in this thesis applies a moving virtual flat box region which is defined as "moving flat box". The basic idea of the algorithm is to trace the medial axis of the tube-like vessel by estimating the centroids of cross section while traversing the path of the moving flat box.

The height of the moving flat box is the step size in tracing, while the width and length of the moving flat box are initially set empirically. The height of the box is much smaller than the width or the length, thus it is "flat" like a pizza box. At the beginning of tracing, the two surfaces of the moving flat box that intersected with the vascular are square. The tracing of a vessel starts from the seed point. The initial position of the moving flat box is determined by the seed point. The moving

flat box is placed in order to let the seed point be the center of one of the square surfaces of the flat box (origin window). The position of the predicting window of the flat box is based on the initial tracing direction as shown in Figure 3.7.
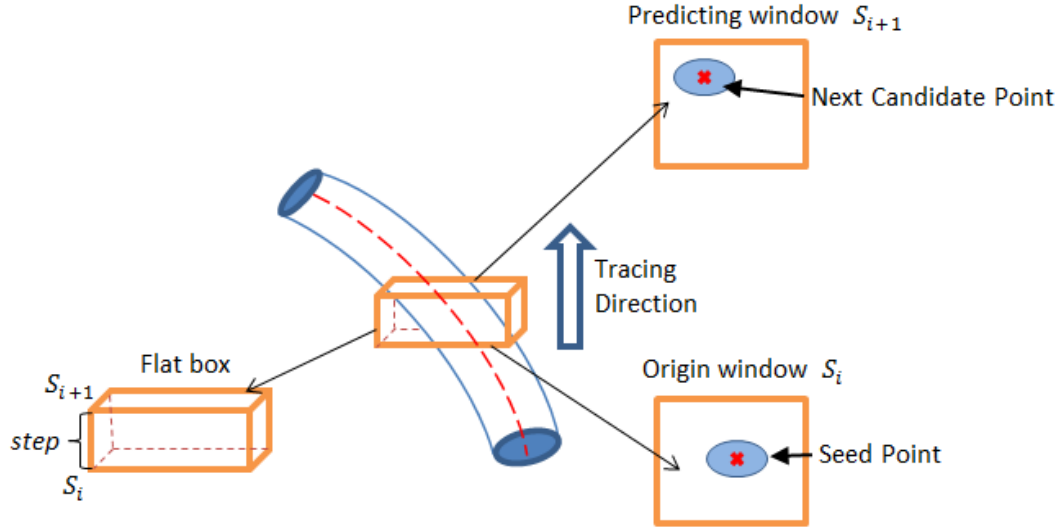
Figure 3.7: Placing the flat box for tracing. The S_i surface of the flat box is placed based on the position of the seed point. The predicting window is on the opposite surface S_{i+1} along the tracing direction. The centroid of the cross section in the predicting window is determined to be the next candidate point in the centerline. The distance between S_i and S_{i+1} is the tracing step.

Given the tracing direction, the vessel cross sectional regions on the predicting window of the flat box can be determined. With the connected component labeling method discussed in section 3.2.1, the centroids of the vessel cross section regions on the predicting window can be found. By connecting the seed point and the next candidate centerline point, a vector $\vec{v}$ is generated. The vector $\vec{v}$ represents the moving trajectory of vessel at the current position of the flat box as shown in Figure 3.8.
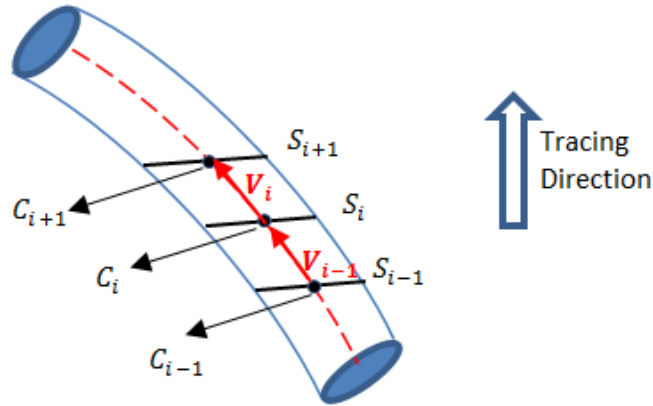
Figure 3.8: The tracing vectors. The tracing vector V_i is obtained by connecting the candidate centerline points on S_i and S_{i+1}.

The flat box is placed based on the predicted centerline points recursively to trace the whole structure of a vessel. While tracing, the boundary of the vessel cross sections and the detected centerline points are recorded.

3.3.2 Dynamic flat box

Considering the complicated structure of the KESM India ink data, the size of the moving flat box in the algorithm should be determined dynamically to ensure the accuracy of tracing. In certain cases, for example the vasculature inflects significantly at a point or the diameter of the vasculature is greater than the width of the flat box , the flat box might not capture a complete cross section of the vasculature. In such cases, a fixed-size flat box cannot guarantee the prediction of the next centroid of the vascular cross section. This section describes the implementation of the dynamic adjustment of the flat box size.

The dynamic flat box expands in four directions based on whether the vasculature intersects the boundary of the flat box. The four directions the box can expand are up, down, left and right corresponding to the four edges of square surface in the

flat box. If the vasculature intersects an edge in the flat box, then the flat box expands a certain length towards the corresponding direction. Figure 3.9 shows how the dynamic flat box expands to adjust the predicting window size. The expansion repeats until there is no intersection between the observed vessel and the flat box.
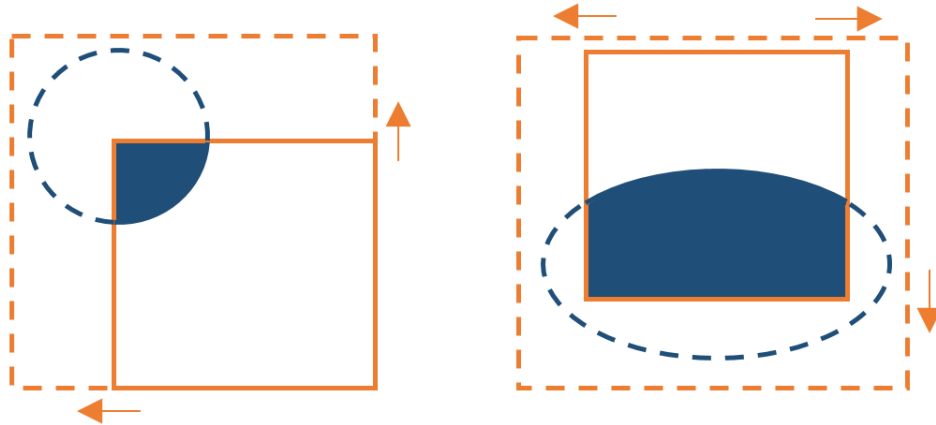
Figure 3.9: The expanding direction of the dynamic flat box.

Note that in the dynamic flat box expanding rules the predicting window of the flat box might contain more than one group of white pixels that belong to a different vessel than the one currently being traced, especially when the volume is packed with many vessels. Since the next candidate centroid should be located in the same vessel with the previous centroid, the expanding process terminates when one complete vessel cross section is found on the predicting window of the flat box. At this time, even though there are vessels intersecting the predicting window boundary, only the complete vessel cross section is considered.

When there are more than one complete cross sectional regions on the predicting window, the next candidate region is determined to be the one closest to the previous

identified point on the centerline. In order to calculate the distance between the identified point on S_n and candidate cross sectional regions on S_{n+1}, first I compute the centroids of candidate cross sectional regions. Then the distances between these centroids and the identified point on S_n are computed. The distance of the points is computed in 2-dimensional coordinate space. The cross sectional region with the minimum distance is selected as the next candidate cross sectional region and the centroid of which becomes the next candidate point on the centerline of the vessel. In the expanding process, if the the edge of the flat box reaches the the edge of the volume, the flat box stops expanding.

The vector $\vec{v}$ is generated the same way in the section 3.3.1 once the next candidate point on the centerline is determined by the dynamically expanding flat box. Also, the area of the selected cross sectional region is recorded for further data analysis.

3.3.3 Flat box exploratory directions

The moving flat box exploratory directions include $x+$, x-, $y+$, y-, $z+$ and z-axis directions. The initial tracing directions are determined based on the location of the seed points.

If the moving flat box only traces in one direction, there might be vessels in the volume that cannot be traced. In this case, using a flat box with another tracing direction can ensure that vessel is traced as shown in Figure 3.10. Thus, when the moving flat box cannot capture vascular data, using the moving flat box with a different moving direction on x, y or z axis directions to trace can make sure there is no missed vessels in the volume.
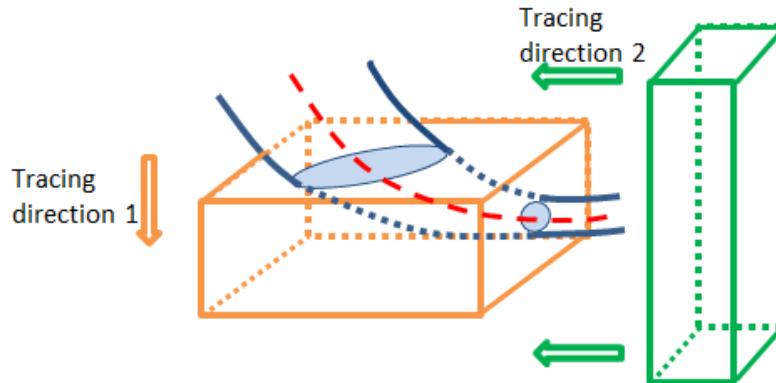
Figure 3.10: The flat box traced from different directions to ensure no vessel is missed. If the vessel cannot be traced by tracing in direction 1, it can be traced by another tracing direction. In this case, using a flat box with tracing direction 2 can solve the problem.

3.3.4   Bifurcation detection

Bifurcation of vasculature is common in the India ink data set. Thus, there would be significant improvement on the accuracy of tracing if the bifurcation can be handled correctly. This section discusses the method of detecting vessel bifurcations. Typical bifurcations are shown in Figure 3.12a.

In the process of tracing with the moving flat box, the area of each vessel cross section is recorded. The radius of the vasculature at each step can be estimated by the corresponding cross section area. By studying the India ink data set, I found that most of the time the bifurcations are located at the part of the vessel with sudden increase in diameter, which means there is a relatively large vessel cross section at the bifurcation point as shown in Figure 3.11. If the vessel radius keeps increasing as flat box moving along the vessel and then decreased at a certain point, then it is likely that bifurcation occurred at that part of the vessel.

| (a) | (b) | (c) |
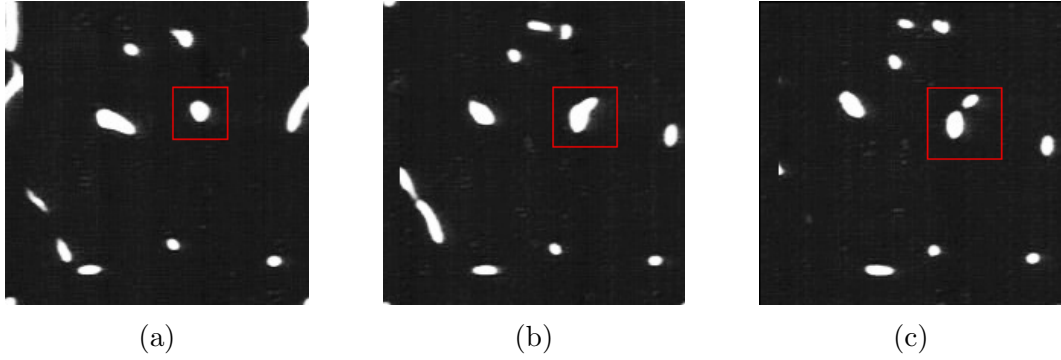
Figure 3.11: The change of the vessel cross sectional regions at a bifurcation point. The regions highlighted with red lines in (a), (b) and (c) show a vessel splitting into two which is a bifurcation.

Let r_i be the radius of vessel cross section on slice S_i. Let d_i be the difference between two consecutive radius of vessel cross sections in the sequence while moving the flat box, i.e. $d_i = r_{i+1} - r_i$. As the flat box moves along the medial axis of the vessel, the sequence of d_i is computed. With two thresholds to compare with d_i, the two indexes of slices are determined where the vessel radius begins to increase and decrease respectively. Let S_{start} be the slice where the radius begins to increase and S_{end} the slice where the radius begins to decrease. Then the slice where the bifurcation point is located is computed as the medial slice of S_{start} and S_{end}: $S_{bifurcation} = (S_{start} + S_{end})/2$. The bifurcation point $p_{bifurcation}$ is then determined by the nearest determined centerline point with the slice $S_{bifurcation}$.

The bifurcation point is inserted into a queue containing possible bifurcation points of the current vessel. If a sub-vessel can be traced from this predicted bifurcation point, the sub-vessel is recorded to connect with the main vessel from the bifurcation point.

Next, the centroid of the sub-vessel cross section region needs to be identified in order to connect with the bifurcation point. Given the bifurcation point and the

26

flat box bounding the cross section at that point, I place the flat box to capture the vascular cross section on the slice of S_{end}. Thus, on that slice the bifurcation cross section should be captured as shown in Figure 3.12b. The square window of the predicting window of the flat box keeps expanding on the slice until it holds two bifurcation cross sectional regions. The cross sectional region without the detected point on the original main vessel is the sub-vessel cross sectional region. The centroid of the sub-vessel cross sectional region is then computed in a similar way as the predicting centroid in the main vessel.

(a)   (b)
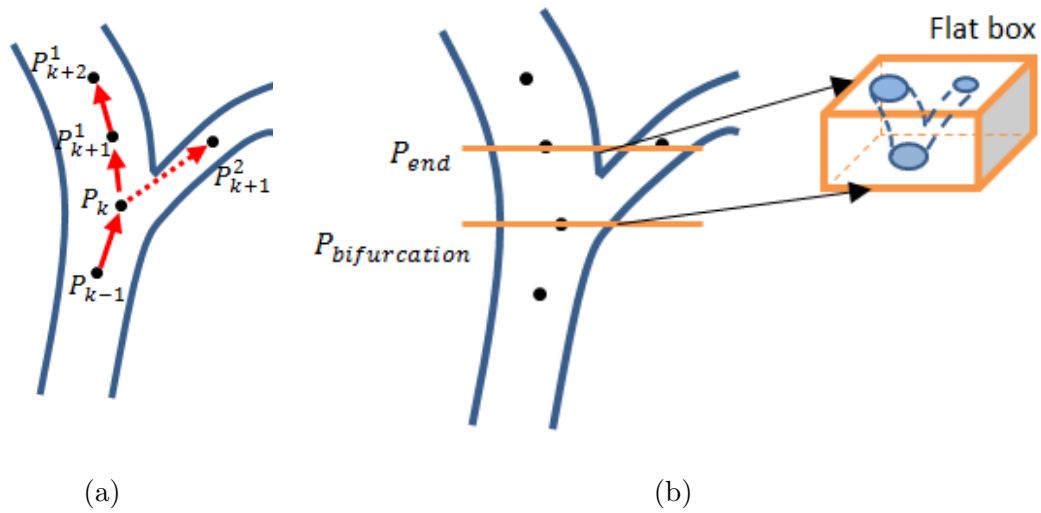
Figure 3.12: Bifurcation detection in tracing. (a) The vessel splits to two branches at the bifurcation point P_k. (b) The flat box is placed at the bifurcation point to predict the start point of sub-vessel centerline.

In certain cases, the square window of the flat box might contain more than two groups of white pixels i.e., two objects. The sub-vessel group is determined by the minimum distance of the bifurcation point and each group, where the distance of

27

bifurcation point $p_{bifurcation}$ and group g_i is defined as.

$$Distance_i = \min \sqrt{(x - x_i)^2 + (y - y_i)^2}, p_{bifurcation} = (x, y), \forall (x_i, y_i) \in g_i$$

Besides the group that contains the identified point on the centerline, the group with the minimum $Distance_i$ is selected as the sub-vessel cross section. Similarly, the centroid of the sub-vessel cross section is determined as the first detected centerline point on the sub-vessel.

After the bifurcation detection process, the relation between the bifurcation point and the first centerline point on the sub-vessel is determined. Then the flat box can start from the initial point on the sub-vessel to trace the whole sub-vessel.

3.3.5 Vessel curve detection

In the process of tracing, there is a special case where vessel curve detection is necessary. The vessel curve in this section means the am abrupt direction change along the vessel which often appear as a letter "U". An example of vessel curve is shown in Figure 3.13a. In the India ink data, it is hard to trace the vessel curve with the moving flat box while tracing in a single direction. In order to trace the entire vasculature, we need to take care of vessel curves.

To be specific, if we place a flat box from the initial point, the flat box would stop tracing when it reaches the edge of the turn. The flat box assumes that there is no vessel structure in the initial tracing direction and the rest of the vessel structure would be missed by exploring in only one direction. Likely, tracing from the seed point on the other end of the vessel would lead to the same problem as shown in Figure 3.13b, where the dashed part never gets traced. Thus, it is necessary to figure out a way to deal with this kind of curve in the vasculature.

One way to solve the problem is backtracking and changing the flat box tracing

(a)

(b)
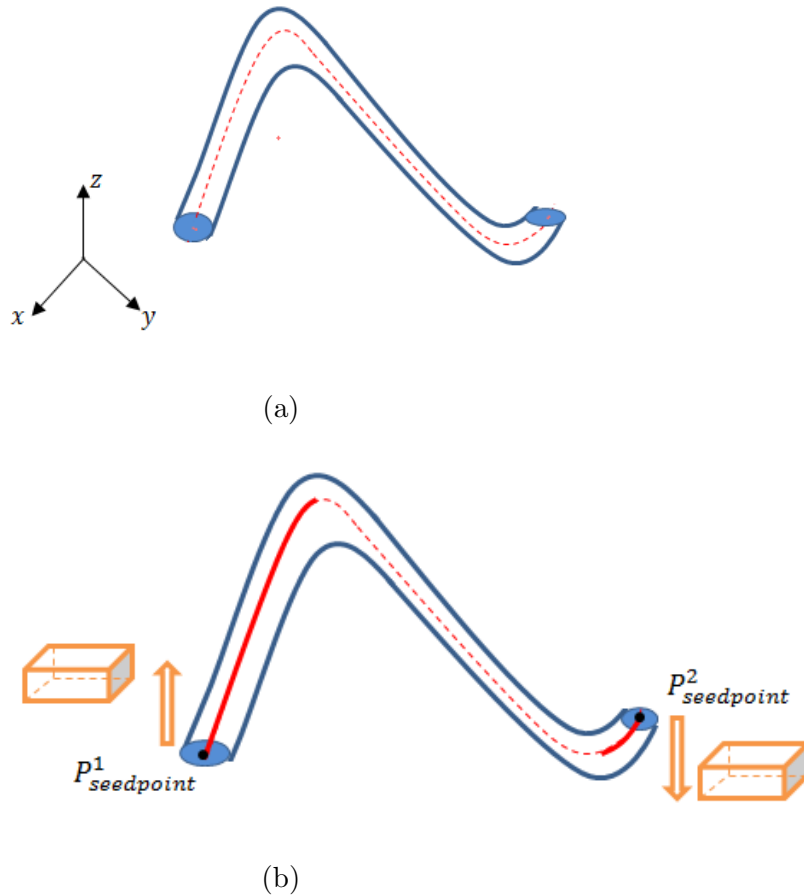
Figure 3.13: Vessel curve detection. (a) The vessel curve is an abrupt change of direction along the vessel. (b) The seed points $P^1_{seedpoint}$, $P^2_{seedpoint}$ with the initial tracing direction can only trace part of the vessel curve. To solve the issue, the tracing direction of the flat box has to be changed while tracing the curve.

direction afterwards. When the flat box reaches the apparent end of a vessel, it is necessary to check if there is a U-turn. To check this, a seed point is inserted at certain steps from the end of the vessel and a different initial direction is given based on the previous trend of vessel as shown in Figure 3.14. The changed direction is determined based on the following rules:

1. If the flat box moves along the z-axis, then check the moving direction in x-axis.

The changed direction should be consistent with the direction of the previous vessel in x-axis.

2. If the flat box moves along the x-axis, then check the moving direction in y-axis. The changed direction should be consistent with the direction of the previous vessel in y-axis.

3. If the flat box moves along the y-axis, then check the moving direction in z-axis. The changed direction should be consistent with the direction of the previous vessel in z-axis.
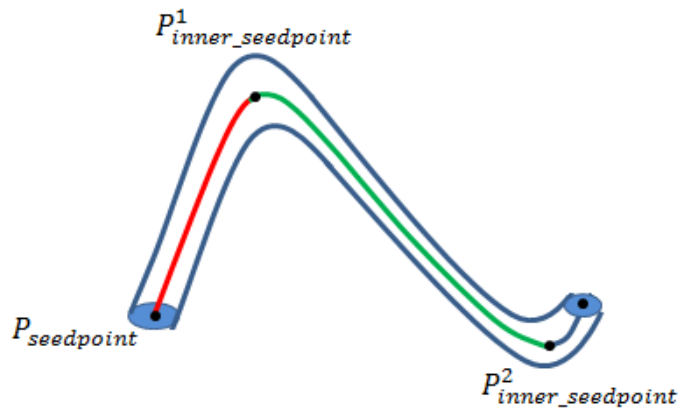
Figure 3.14: Insert seed point to trace vessel curve.

The only concern with the backtracking method is that it might trace back the already processed part of the same vessel. By using the changed direction rules above, we can avoid the issue of tracing back the already processed part. An integrated tracing result for the vessel is generated by connecting each part of the tracing result along the vessel curve.

3.3.6   Avoiding gaps

Due to the issues like imaging errors and staining artifacts in the KESM system, there might be some gaps between the vascular slices in the volume data set. In a volume with gaps, the moving flat box might terminate tracing when it reaches the slice where the gaps exist. This would result in incomplete tracing. To overcome this problem, the tracing method should avoid terminating at the gaps while traversing the moving flat box.

Let R_{end}^{v1} be the last vessel cross section region before the gap. In the tracing method, the moving flat box checks if there is a gap between slices when the predicting window cannot capture any vessel pixel. In such case, the predicting window on the moving flat box moves a conservative step (about 2 voxels) further along the tracing direction to check whether there is any vessel pixels. If one vessel cross section can be captured by the predicting window of the moving flat box, the vessel cross section region is recorded as R_{start}^{v2}. If more than one vessel cross sections are captured, the vessel cross section with the minimum distance to the previously identified centerline point is considered as R_{start}^{v2}. The distance between a point and a connected group region is computed as the distance between the point and the centroid of the group in 2-dimensional space. After that, the similarity of the two parts of the vessel should be ensured so that it would not connect points in different vessels. The similarity of vessels is estimated by the radius of cross section regions R_{end}^{v1} and R_{start}^{v2}. If the radius difference between region R_{end}^{v1} and R_{start}^{v2} is under a predefined threshold, it is decided to be similar. By connecting the two centroids of R_{end}^{v1} and R_{start}^{v2} over the gap, the tracing ignores small gaps in vascular volume.

3.3.7   Termination criteria

The vector tracing method must terminate under certain criteria so that the time consumption is limited and the correctness of tracing result is guaranteed. The tracing method terminates under the following conditions:

1. If the next candidate point predicted by the moving flat box reaches the boundary of the volume and the previously traced point on the centerline does not lie on the edge of the volume.

2. The moving flat box reaches the end of the vessel and there is no possible bifurcation points, vessel curves and volume gaps left to be traced for the vessel.

In the first condition, if part of the vessel lies on the boundary surface of the volume, the tracing would not terminate when the predicted centerline point reaches the boundary of the volume. The moving flat box continues tracing the vessel on the boundary if the vessel does not extend outside of the volume as shown in Figure 3.15a. However, if the previous trace trends towards the outside of the volume boundary, then the flat box would terminate tracing as shown in Figure 3.15b.

In the second condition, bifurcation points, vessel curves and volume gaps discussed in sections 3.3.4-3.3.6 are taken into account. When the second condition is satisfied, an integrated structure of the vessel has been traced.
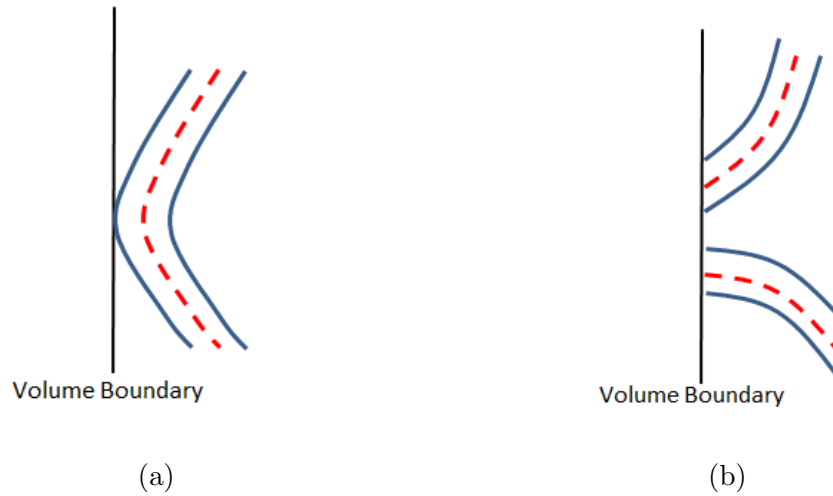
Figure 3.15: Vessel tracing termination criteria. (a) The process continues tracing the vessel when it reaches the boundary in this case. (b) If the vessel trends towards the outside of the volume, the tracing of this vessel will terminate. The other vessel branch in the image on the boundary surface is traced from another seed point.

3.3.8 Vascular network exploration

Vascular network exploration utilizes the tracing steps described in sections 3.3.1-3.3.7. This section illustrates how the moving flat box explores the vascular network within the volume.

The process of exploration consists of the following steps:

1. Starting from each seed point, the moving flat box traces a vessel by keeping a list of predicted centerline points. The key structure of a vessel can be represented by the list of centerline points.

2. In tracing the main vessel, the possible child vessels are explored by breadth first search. Possible bifurcation points of a vessel are inserted into a queue for further exploration.

3. At the end of tracing for a vessel, curves and gaps are checked to trace an integrated vessel. Since the curves and gaps in the vessel are the extensions of the main vessel, the predicted points are concatenated to the end of the main vessel centerline.

4. The moving flat box traces from each point in the queue to trace the bifurcations. The coordinates of the validated child vessels are documented in a new list and the bifurcation relation is documented as well.

5. Remove the point from the queue once it is checked. The exploration terminates when the queue is empty.

Vascular network exploration can also be performed by depth first search with a stack rather than a queue.

The centerline points of the vasculature are recorded while exploring the network. The tracing result of an integrated vessel is recorded as a list of points in the volume. The vascular network contains multiple lists of points where each list represents a vessel. The tracing result is rendered by connecting the two neighboring points in a list using vtk format and rendered in the software Paraview, a visualization program.

3.4   Data analysis

This section presents the analysis of the traced vasculature. Three properties, radius, length, and volume of vessel, are analyzed to demonstrate how much vascular structure is extracted in the tracing process.

First, the radius of the vessel is estimated by the area of the cross section. In each iteration i of the tracing process, the area of vessel cross section is the total number of pixels within the region. Based on the equation $S_i = \pi \times r_i^2$, the radius r_i of the cross section can be estimated.

Second, the length of the vessel is the sum of the distance between the two neighboring centerline points. Since the step length of moving flat box is relatively small compared to the length of the vessel, the length of the two neighboring points (x_1, y_1, z_1) and (x_2, y_2, z_2) on the centerline is computed as straight-line distance between the two points. $l_i = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}$

Finally, the volume of the vessel is approximately computed as the sum of vessel volume between the consecutive vessel cross sections in the tracing process. The volume of each vessel segment is estimated as the multiplication of the vessel cross section area and the step length, i.e. $V_{v_i} = S_i \times l_{step}$.

4.   RESULTS AND ANALYSIS

In the previous chapter, the method of the automated vasculature tracing was explained. In this chapter, the results and analysis of the method will be presented. In the first part of this chapter, the results of seed point selection are presented. First, the accuracy of the automated seed point selection method is compared to manual method of seed point selection. Next, the results of vascular tracing performed on the sub-networks of the KESM India ink data set are presented. Finally, the accuracy of the method is evaluated using synthetic data and the efficiency of the method evaluated based on the processing time of the tracing method in the sub-networks of the KESM India ink data.

4.1   Results of seed point selection

As discussed earlier, the seed points should be selected on the six surfaces of the vascular volume in order to trace vessels within the volume. The seed point selection method was tested on a sample sub-network of $100 \times 100 \times 200$ voxel volume and a sample sub-network of $200 \times 200 \times 200$ voxel volume from the KESM India ink data set. The seed points which are the centroids of vascular cross sectional regions on the six surfaces of the volume are computed and the initial tracing directions are assigned for each seed point based on the orientation of the surface where the seed points are located. The results of the $100 \times 100 \times 200$ voxel volume and the $200 \times 200 \times 200$ voxel volume are shown in Figure 4.1 and Figure 4.2 .The selected seed point coordinates were marked in 3D and superimposed on the original KESM volume data using the software Paraview. In the figures, the gray tube-like structures are the vessels and the pink dots represent the selected seed points.
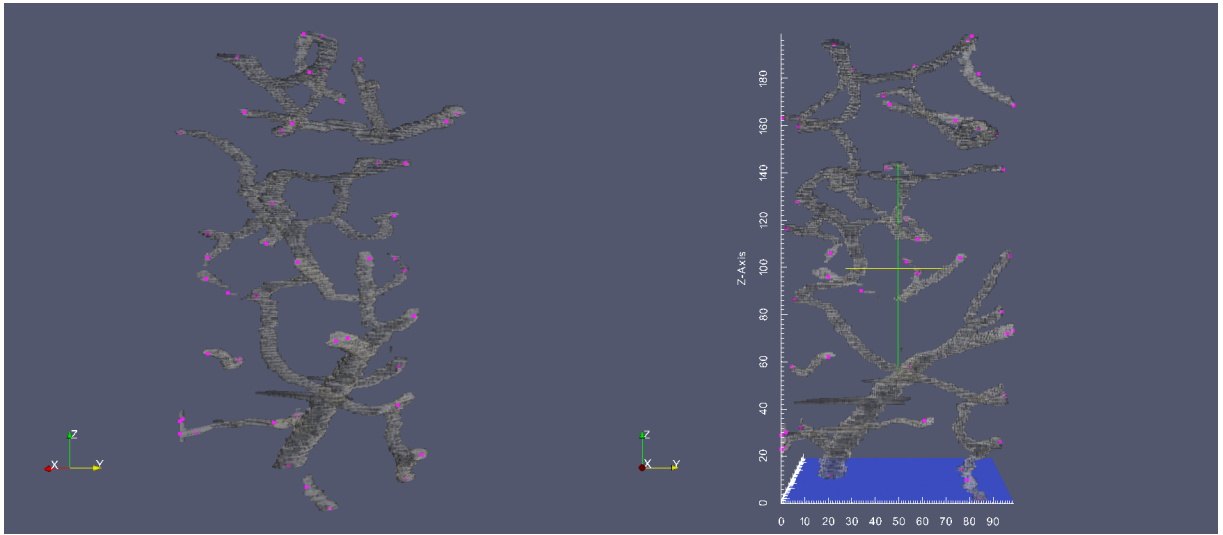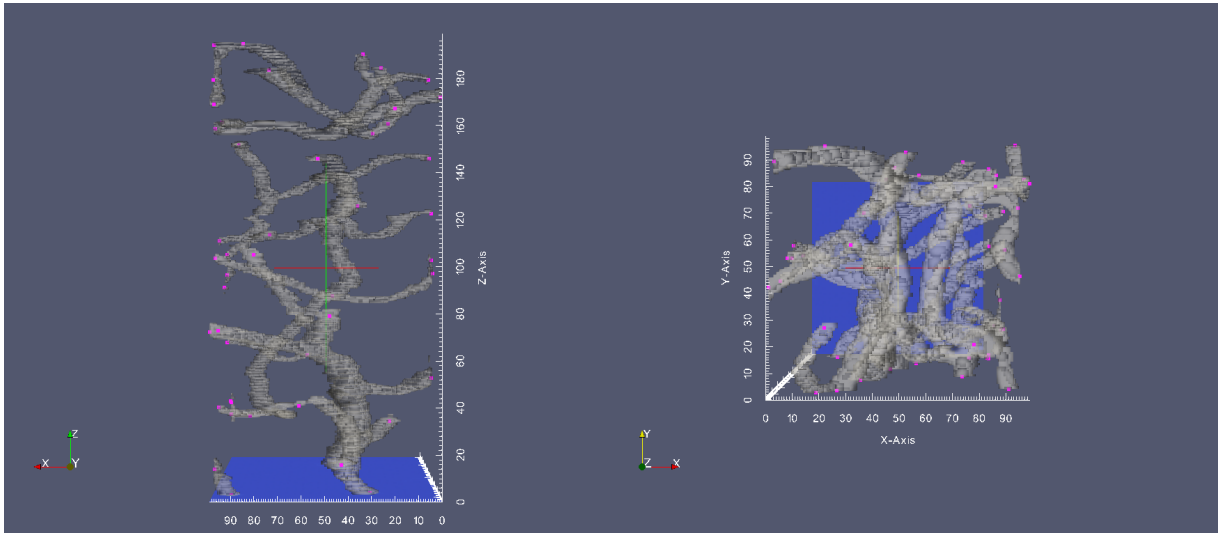
Figure 4.1: Seed point selection result of a $100 \times 100 \times 200$ voxel volume. (a) The 3D view of seed points on the six surfaces of the volume. (b) The view of (a) along the x- axis. (c) The view of (a) along the y- axis. (d) The view of (a) along the z-axis.
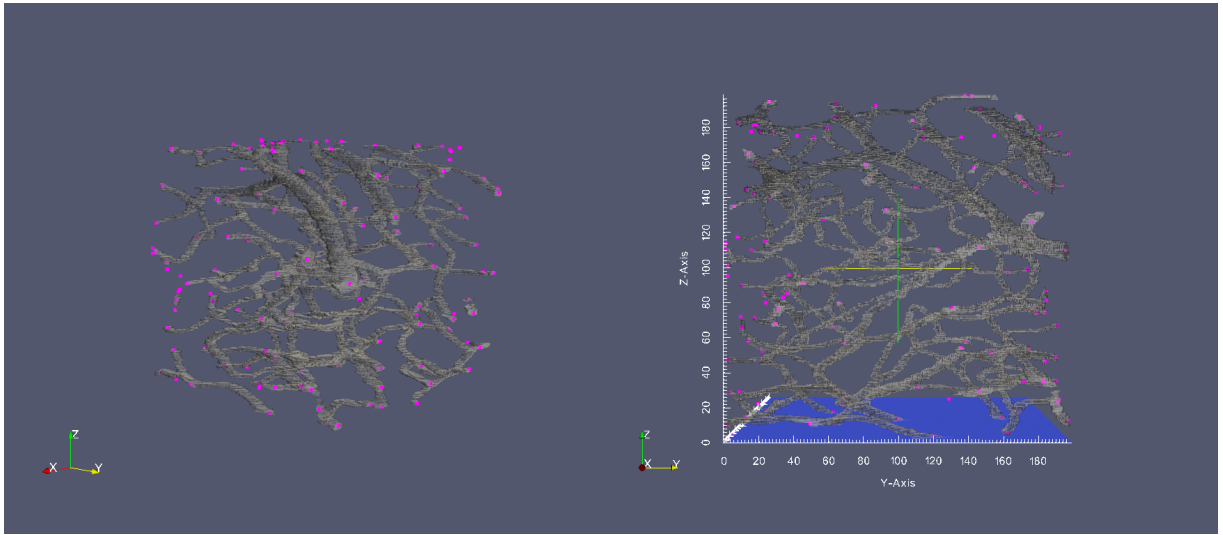
(a)   (b)
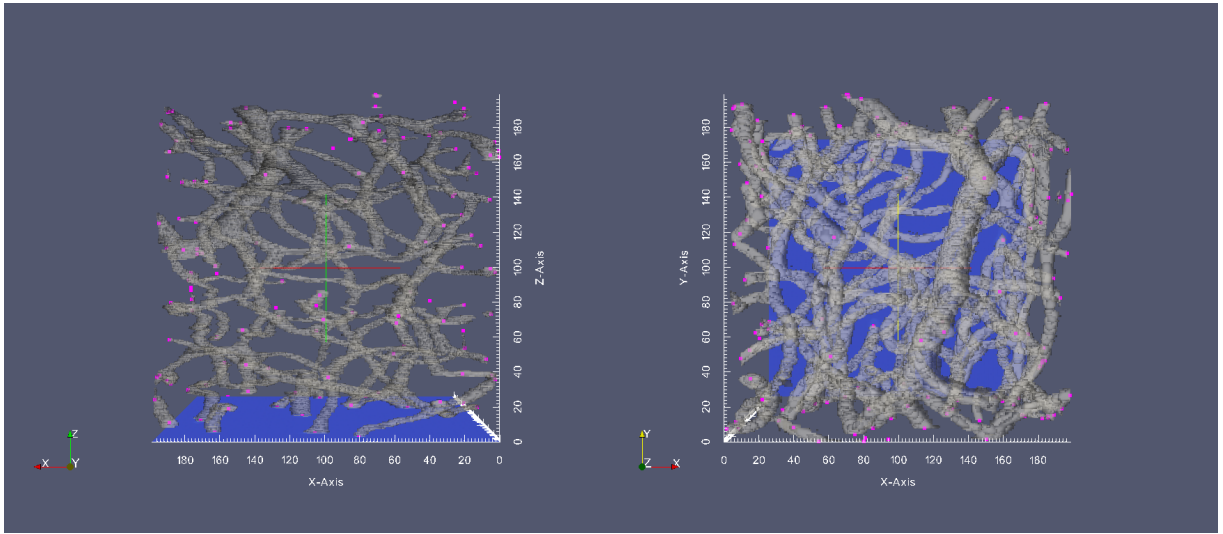
(c)   (d)

Figure 4.2: Seed point selection result of a $200 \times 200 \times 200$ voxel volume. (a) The 3D view of seed points on the six surfaces of the volume. (b) The view of (a) along the x- axis. (c) The view of (a) along the y- axis. (d) The view of (a) along the z-axis.

To measure the seed point selection performance, the precision(p), recall(r), and F-measure(F) were calculated. Let n_1 be the true number of seed point, n_2 be the

number of all selected seed points, and n_3 be the number of correctly selected seed points. The precision, recall and F-measure of seed point selection are derived as:

$$p = n_3/n_2$$

$$r = n_3/n_1$$

$$F = \frac{2pr}{p + r}$$

In the experiment on the $100 \times 100 \times 200$ voxel volume, precision, recall, and F-measure of seed point selection were 0.889, 1, and 0.941, respectively. In the experiment on the $200 \times 200 \times 200$ voxel volume, precision, recall, and F-measure of seed point selection were 0.978, 0.875, and 0.923, respectively.

4.2   Results of vascular tracing

In this section, both the sample volume of the KESM India ink data and the synthetic data are tested to evaluate the vascular tracing method. In the first part, the experiment on the KESM data is presented. In the following part, the results of two sample synthetic data as well as the comparison between the ground truth and tracing result are presented.

4.2.1   Experiment on KESM data

The KESM sample sub-networks of the $100 \times 100 \times 200$ voxel volume and the $200 \times 200 \times 200$ voxel volume discussed in section 4.1 were tested in this experiment. The seed points on the six surfaces of the volume were the seed point selection results described in the previous section. Figure 4.3 and Figure 4.4 present the 3-dimensional view of tracing results overlaid with the original vascular volume data. In the figure, the white curves within the vasculature represent the tracing results.
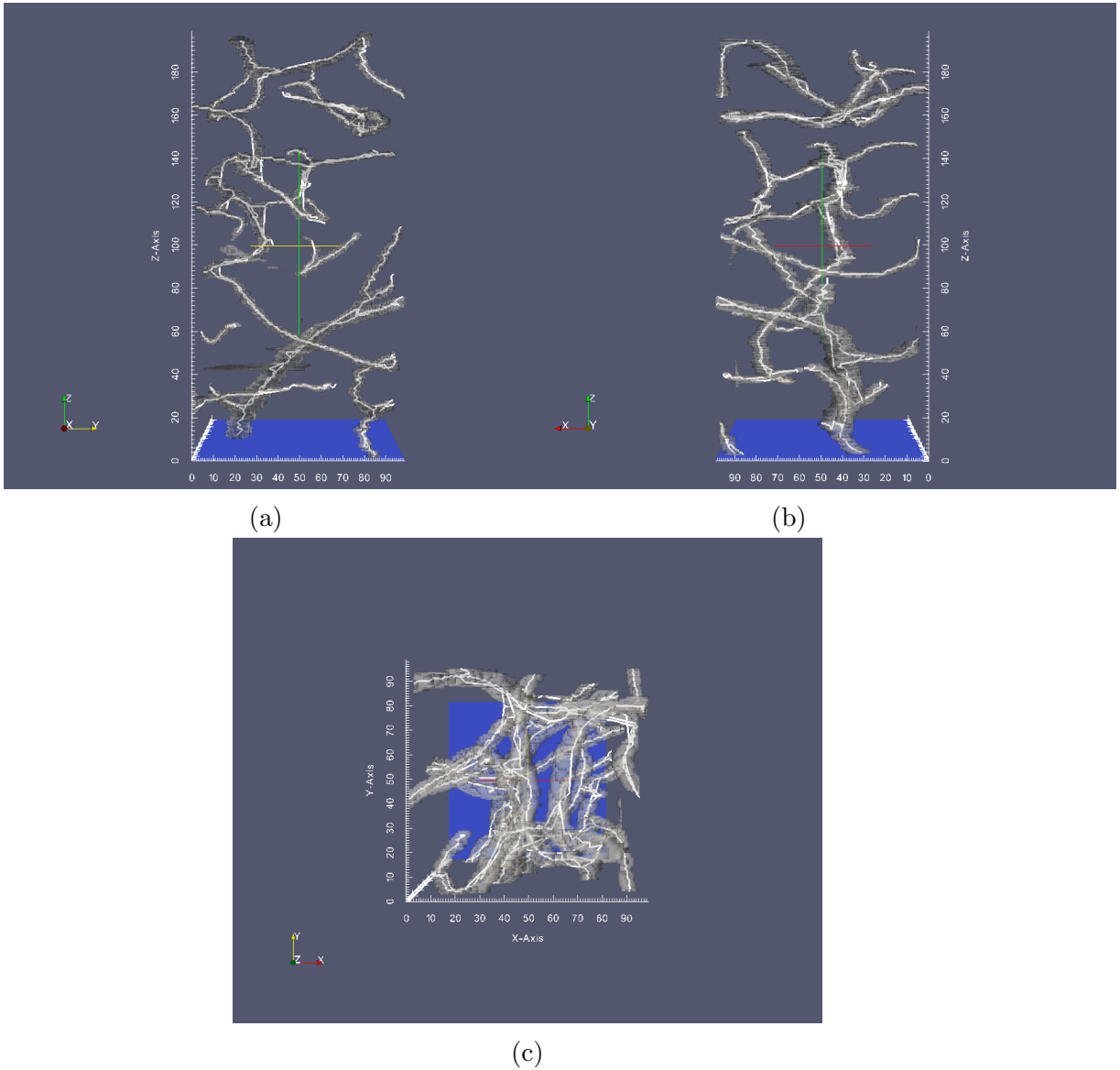
Figure 4.3: The vascular tracing result of KESM $100 \times 100 \times 200$ voxel volume. (a) View along the x- axis. (b) View along the y- axis. (c) View along the z- axis.

The tracing result shows the vessels in the volume are fully explored. The extracted vascular structures are also within the boundary of the vessels. In addition, bifurcations and inner vessel curve structures in the volume are traced.

The geometric statistics were collected in the tracing process. The ground truth
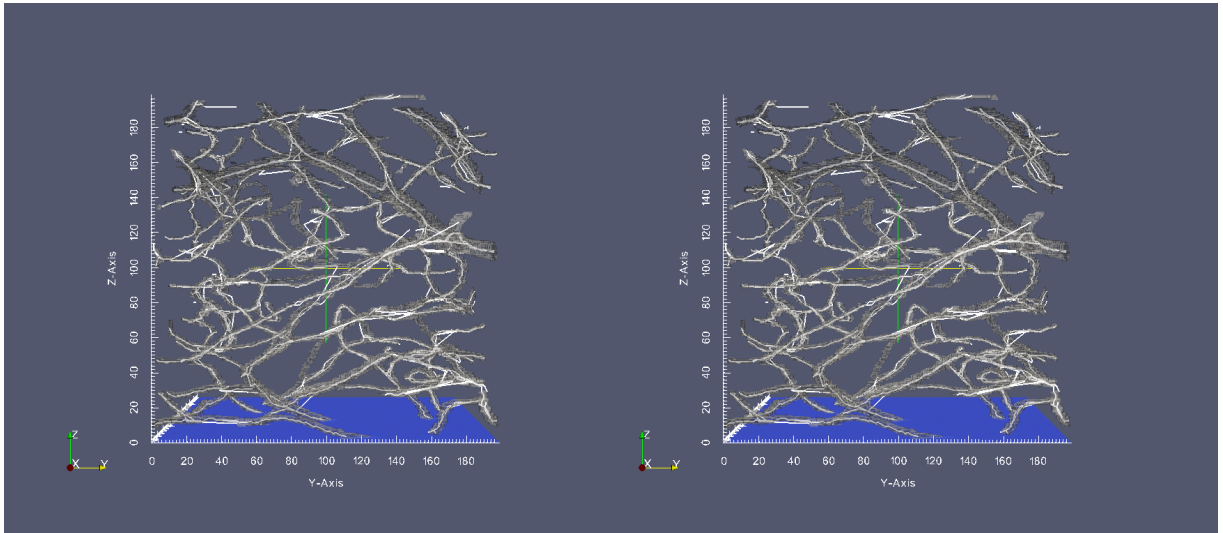
(a)

(b)
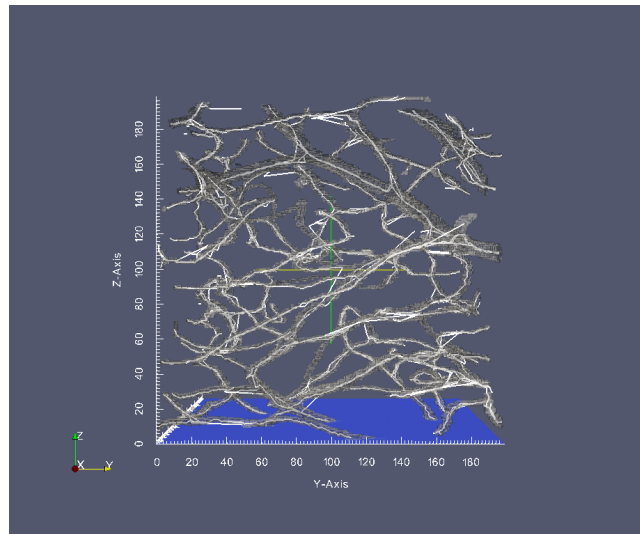
(c)

Figure 4.4: The vascular tracing result of KESM $200 \times 200 \times 200$ voxel volume. (a) View along the x- axis. (b) View along the y- axis. (c) View along the z- axis.

of total vasculature volume was computed as the number of voxels within the vasculature. In the $100 \times 100 \times 200$ volume, the total volume of traced vasculature was 62590 voxels and the ground truth of total vasculature volume was 69653 voxels. The total length of traced vasculature was 1565 voxels . In the $200 \times 200 \times 200$ volume,

41

the volume of traced vasculature was 252291 voxels and the ground truth of total vasculature volume was 281292 voxels. The total length of traced vasculature was 5922 voxels.

Tracing result also shows that there are some redundant traces along some vessel segments. This is caused by tracing from seed points on both ends of the vessel. To reduce duplicated tracing results, redundant seed points need to be eliminated. However, removing some seed points may cause missed vessel to be traced, thus there is a trade-off .

4.2.2 Experiment on synthetic data

To quantitively evaluate the tracing method, synthetic vascular data was generated from a known ground truth. The process of generating synthetic data is a simple simulation of generating tube-like structure. The tube-like vessel structure is simplified into a series of truncated generalized cones. In order to make sure there is no gap at the connections of the truncated generalized cones, each point extends at a certain degree to overlap with neighboring cones as shown in Figure 4.5.
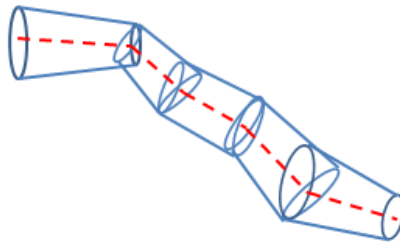
Figure 4.5: Synthetic data generation is simplified into a series of truncated generalized cones.

The process of generating the synthetic data is as follows. First, given some control points a linear bezier curve passing the control points can be obtained. Let the curve be the centerline of the vessel. At each point of the curve, a randomized radius r_i within a small range is assigned to the point p_i. The radius at each centerline point within the range of the two caps of the cone is computed by linear interpolation.

The tube-like vessel is plotted in in 3-dimensional space by marking all the points within the vessel boundary. The vessel boundary is determined by the radius relative to the centerline point. Therefore we need to find all the points whose distances to the centerline are less than the radius. The criteria of determining a point within the boundary of a vessel is illustrated in Figure 4.6. If point v is between the caps of the truncated generalized cone, the equation can be derived as:

$$0 < (\vec{v} - \vec{p}) \cdot (\vec{q} - \vec{p}) < ||\vec{q} - \vec{p}|| \tag{4.1}$$

where the points $\vec{p}$ and $\vec{q}$ are the center of the caps. The distance of the point v to the centerline of the truncated generalized cones is given by:
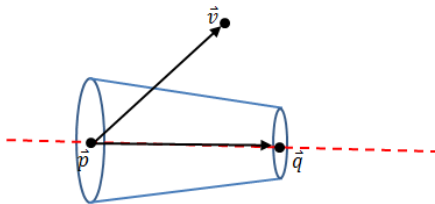
Figure 4.6: Deciding if a point $\vec{v}$ is within the truncated generalized cones. $\vec{p}$ and $\vec{q}$ are the centeroids of the caps.

43

$$r_{\vec{v}} = ||(\vec{v} - \vec{p}) - \vec{d} \cdot (\vec{q} - \vec{p})/||\vec{q} - \vec{p}||||$$ (4.2)

where $\vec{d} = (\vec{v} - \vec{p}) \cdot (\vec{q} - \vec{p})$. $r_{\vec{v}}$ should be less than or equal to the interpolated radius at $\vec{v}$ if it is located within the boundary of the vessel.

The resulting structures are then converted to volume data and the centerlines are regarded as the ground truth to be validated against. Figure 4.8 shows the two sample synthetic data generated by this method. The ground truth of centerline is marked in pink lines in the volume.

4.2.3   Synthetic test 1

In synthetic test 1, the tracing algorithm was tested on a synthetic vascular volume of $100 \times 100 \times 100$. The seed point of the volume $(60, 60, 0)$ was given at the beginning of tracing. The volume contains 11 distinct vessels and 5 bifurcation points. The topology of the synthetic vasculature is illustrated in Figure 4.7. The tracing result is plotted in 3-dimensional space as shown in Figure 4.9.

Figure 4.7: The topology of the vascular network in synthetic test 1.
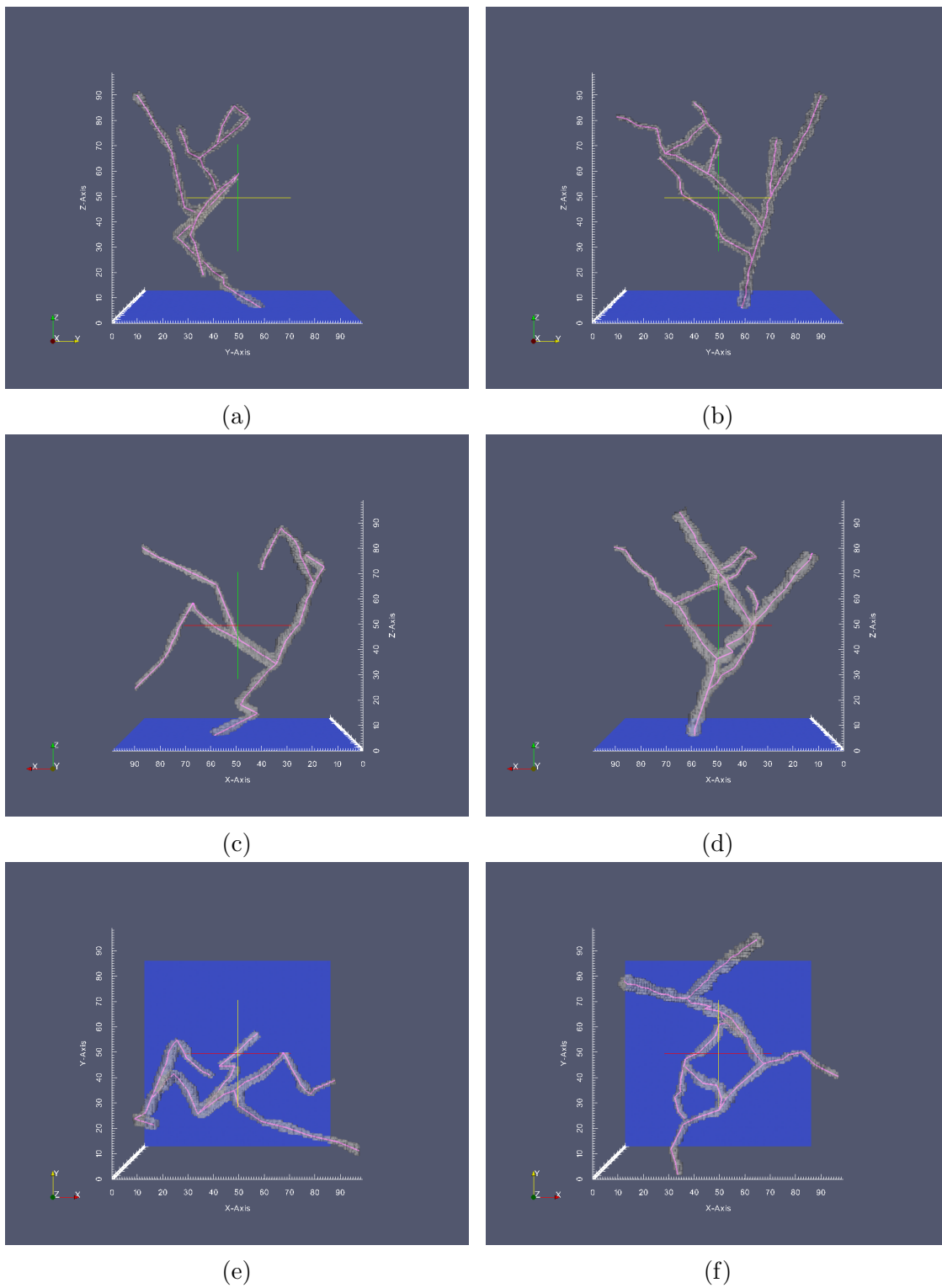
Figure 4.8: Synthetic data with ground truth. (a) (c) (e) are the x,y,z axis view of synthetic data 1. (b) (d) (f) are the x,y,z axis view of synthetic data 2. The pink curves represent the ground truth.
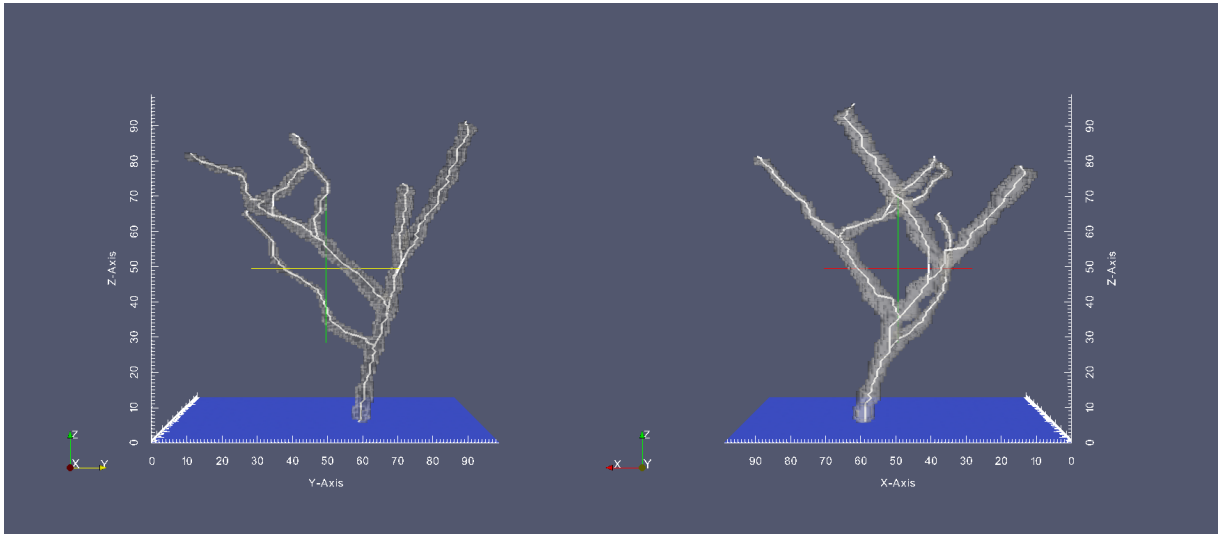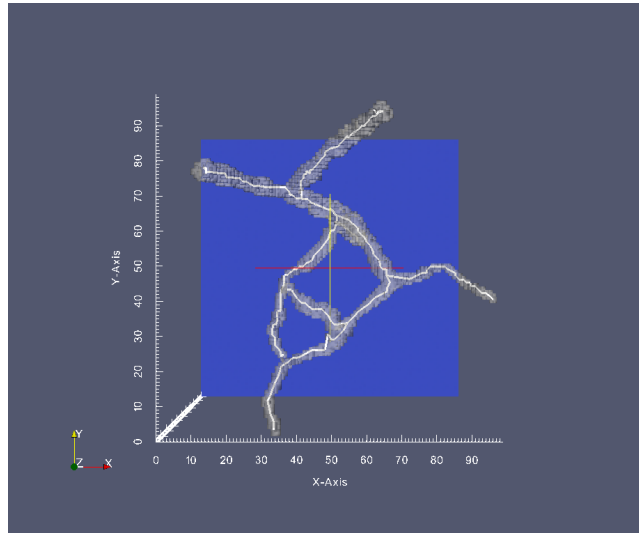
Figure 4.9: Vascular tracing result of the synthetic data 1. (a) View along the x-axis. (b) View along the y- axis. (c) View along the z- axis.

The overlay of tracing results and the original volume shows that the structure of the vascular network was extracted correctly. Also, given one seed point the branches of the main vessels have been traced completely.

The geometric statistics collected in the tracing process shows that the total vol-

ume of traced vasculature was 10826 voxels and the total length of traced vasculature was 261 voxels. In the synthetic data 1, the total volume of ground truth was 11812 voxels and the total length of the ground truth was 271 voxels. The statistics also show that most of the vascular network was traced.

4.2.4 Synthetic test 2

The volume in synthetic test 2 was a $100 \times 100 \times 100$ cube. Tracing starts from the seed point $(60, 60, 0)$ using my tracing method. The topology of the vascular network in the synthetic data set is shown in Figure 4.10. In the network, there are some vessel curves as described in the methods chapter. Also, the volume contains bifurcation points. Vessel 5 and vessel 6, vessel 8 and vessel 9 described in topology are vessel curves, while the intersections of vessel 2 and vessel 3, vessel 4 and vessel 5, vessel 7 and vessel 8 are the bifurcation points that need to be detected.
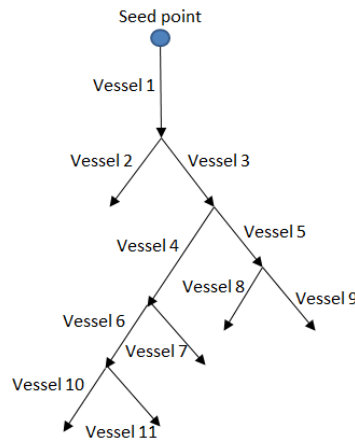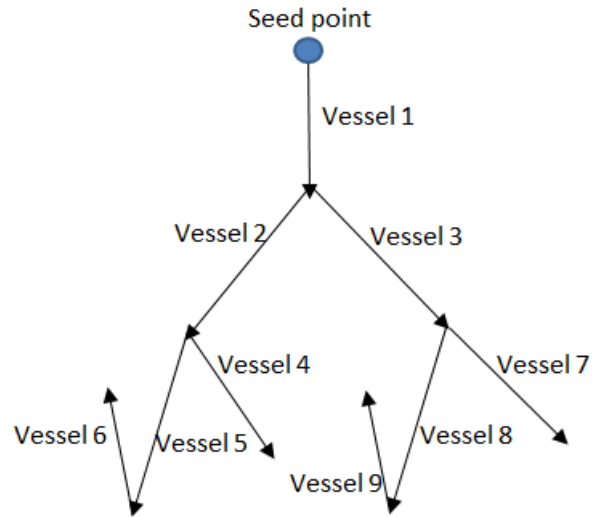

Figure 4.10: The topology of the vascular network in synthetic test 2.
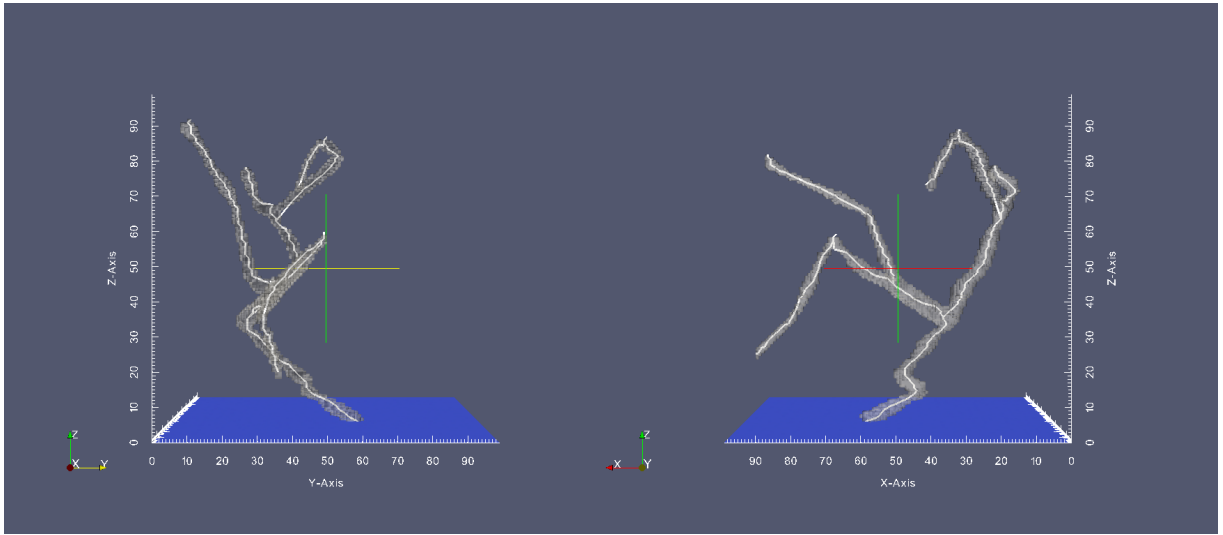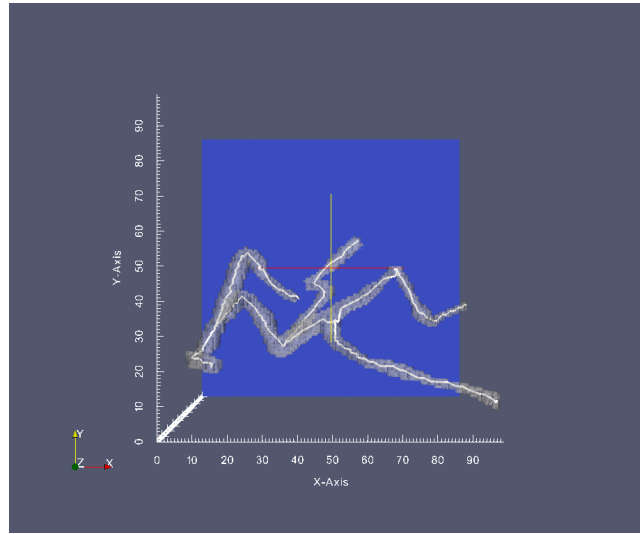
Figure 4.11: Vascular tracing result of the synthetic data 2. (a) View along the x-axis. (b) View along the y- axis. (c) View along z- axis.

The extracted centerlines are shown in Figure 4.11. Both the vessel curves and the bifurcations are traced correctly in the volume.

The geometric statistics collected in the tracing process shows that the total volume of traced vasculature was 9156 voxels and the total length of traced vasculature

was 220 voxels. In the synthetic data 2, the total volume of ground truth was 9239 voxels and the total length of the ground truth was 248 voxels. The statistics show that most of the vascular network was traced in this experiment. The deviation between the tracing result and the ground truth may be due to the fact that the estimates assume a circular cross sectional area while the actual shape could be oval.

4.2.5 Validation

From the results above, we can see that the vascular network is fully traced. To evaluate the detailed difference between the ground truth and my tracing results, quantitive measurements at the voxel level were performed on the synthetic data sets.

In the two synthetic test cases, the data were divided into 10 pieces of vascular segments to validate my tracing method. Two measurements were used to evaluate the difference between ground truth and the results of the tracing method. The quantitive measurements were based on Han's validation method [7]. The first measurement is the length difference ϕ. Suppose R represents the ground truth and A represents my tracing method's results. The length difference is computed as:

$$\phi = |1 - L_R/L_A| \tag{4.3}$$

where L_R and L_A are the length in the ground truth and the vascular segments extracted by my tracing method, respectively. The second measurement is the centerline deviation φ which is derived as:

$$\varphi = V_{ox}(l_R, l_A)/L_A \tag{4.4}$$

where $V_{ox}(l_R, l_A)$ is the voxel-wise distance between l_R and l_A, where l_R is the

ground truth centerline and l_A is the extracted centerline by my tracing method. These two measurements are collected from the two synthetic test cases.

Table 4.1 shows the difference between my tracing result and the ground truth in synthetic test 1 and 2. The results show high similarity between the centerline ground truth and my tracing results.

Table 4.1: The mean (μ) and standard deviation (σ) of length difference (ϕ) and centerline deviation (φ) relative to the ground truth of the synthetic data sets.

	ϕ		φ	
	μ	σ	μ	σ
Synthetic test 1	0.0494	0.0416	0.6753	0.1863
Synthetic test 2	0.0573	0.0308	0.7632	0.0837

4.3   Tracing speed

In order to perform tracing in very large vascular data sets, the tracing method has to be fast. The implementation of my tracing method was optimized to achieve high efficiency. The speed of the tracing method was tested on sub-networks of the KESM India ink data set with different sizes. The processing times of vascular tracing on the three volumes were 24.164, 49.291, and 112.599 seconds for the total traced vessel length of 2936, 6756, and 16702 voxels, respectively. The experiments were performed on a PC with Intel Core 2 (2.13 GHz) processor and 4 GB of memory under Windows 7 operating system in Debug mode of Microsoft Visual Studio 2010.

In Han's method [7], the processing times of vascular tracing were 177, 194, and 1665 seconds for the length of 1214, 1334, and 13700 voxels, respectively. Since

Han's method was tested on a different machine (a PC with Intel Pentium 4 (2.4 GHz) processor, 512MB of memory under Windows XP operating system in Debug mode of Microsoft Visual Studio 2005).

According to publicly available performance benchmark data, the performance of the machine that tested my tracing method is three times better than the performance of Han's machine. The performance benchmark used here is wPrime v1.55 (32M) which measures time taken to calculate square roots of numbers from 1 to 33554431. In the performance benchmark, the time to calculate the square roots in my machine is approximately three times faster than in Han's machine. Vascular tracing is mainly mathematical calculation, therefore, to be fair, the processing times of vascular tracing in Han's results are scaled (i.e. divide by three) in comparison with the processing times of my tracing method(Figure 4.12).
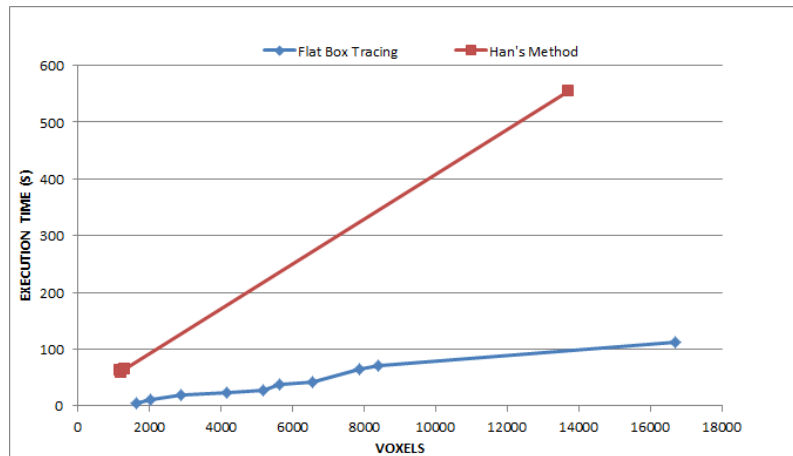
Figure 4.12: Speed comparison between my tracing method and Han's method.

Figure 4.12 shows that the processing time of my flat box tracing method scales linearly in general. In large volume the vasculature structure could be complicated

which leads to longer processing time while tracing. However, the flat box tracing method improved the speed of processing relatively small vascular volume significantly. Figure 4.12 also shows that shorter execution time was needed in processing the same number of voxels in the flat box tracing method compared to Han's method. In Han's method, six surfaces of the MIP cubes were processed in order to predict the vessel direction in 3D. However, in the flat box method, only the top and the bottom surfaces of the flat box are processed to trace the centerline of vessel, and furthermore it does not require expensive computation of the Hessian. All these factors lead to shorter execution time. The comparison shows that the execution time of the flat box method increases slower than the execution time of Han's method as the number of traced voxels grow. Thus, in conclusion, the tracing method proposed in this thesis shows a significant improvement in processing speed over existing methods.

5. DISCUSSION

This thesis presented a novel vascular tracing algorithm for the KESM India ink data set. In this chapter, contribution, open issues, and future works will be discussed.

5.1 Contribution

The vascular tracing method combines automated seed point selection with vector-based vascular tracing which reduces the processing time than the related work. The seed point selection method extracted the vessel cross sectional centroid efficiently on every surface of volume without human assistance. Also, the tracing method explores the entire network of vascular volume with high accuracy in the KESM India ink data set. Since the tracing conducted in each volume is fast without time consuming template matching and computation, the method can be applied to large and high-resolution vascular data sets like those of the KESM. In addition, the vector tracing method deals with image gaps in the data set which can cause some tracing methods like the vessel thinning method to fail to completely trace the network. Finally, the data analysis step collects geometric statistics such as vessel length and volume while tracing. In summary, the proposed tracing method performs well on the KESM India ink data, and possibly on other similar data sets.

5.2 Open issues and future work

In this thesis, an automated vascular tracing and data analysis methods were demonstrated. There are several open issues: (1) Duplicated tracing results due to seed points at the two ends of the vessel, (2) limitation on the volume size in the tracing method, (3) alternative bifurcation cases and parameters that need to

be manually selected for bifurcation detection, and (4) irregular smoothness of the tracing results.

5.2.1 Duplication of the tracing result

Before tracing the vasculature, seed points on the six surfaces of the volume were generated. However, while tracing the vascular volume, the same part of vessel might be traced multiple times if multiple seed points on different surfaces lead to the same vessel. As a result, duplicated tracing results with different tracing directions can exist for a single vessel and the processing time increases as a result. Tracing from all the seed points on every surfaces of volume would reduce missed vessels, but the redundant seed points can cause the duplicated tracing issue at the same time.

To solve this issue, future work could include checking whether the same voxel is visited while tracing the vessel. If the same part of volume is already traced, the current tracing process should be terminated. This way, redundant tracing results would be eliminated. The tracing result will be more clear for studying the vascular architecture.

5.2.2 Limitation on the volume size

The size of the volume tested by the tracing method is limited. If the volume size is too large, there might be vessels within the volume that are not connected to the surface. Since the vascular tracing method starts from the seed points which are selected only on the surfaces of the volume, it is possible that the method misses vessels in the center of the volume. Thus, the size of volume tested by the tracing method should be limited to ensure all the vessels in the volume are traced.

Thus, a large vascular volume needs to be divided into multiple relatively small volumes to perform the tracing. Then, the traced network of each small volume can be merged to form the complete network of the large volume. Vascular tracing in

the sub-volumes can be conducted in parallel to reduce the processing time when tracing large vascular volumes like the entire mouse brain data from the KESM.

5.2.3   Bifurcation dection

Bifurcation of vasculature described in the methods chapter is the most common bifurcation case observed in the KESM data. There might be other bifurcation cases in the vascular volume. To make sure all the bifurcations are detected, a study of alternative bifurcation cases is necessary.

In addition, bifurcation detection is based on the difference of radius of vessel cross sections in each step in the tracing process. The threshold of difference is chosen empirically. In the future, I will develop a more precise method with adaptive threshold to detect the bifurcations. An adaptive threshold to detect bifurcation might increase the accuracy and decrease the number of missed vessels during tracing. However, extra computation time may be needed.

5.2.4   Smoothing of the tracing result

The vector-based vascular tracing described in this thesis does not guarantee smoothness of the tracing result. It connects the extracted centerline points by linear interpolation, however, the centerline of the vessel is a curve. To smooth the tracing result, post-processing of the tracing result is needed. In the future, developing a smoothing algorithm with correct control points on the centerline would guarantee the smoothness of the tracing result and improve the accuracy of tracing.

6. CONCLUSION

The main goal of this research was to develop a fully automated vascular tracing method with shorter processing time to trace vascular networks in the mouse brain KESM India ink data set. The new vascular tracing algorithm was combined with automated seed point selection method to explore completely the vascular network in the volume data. After tracing, I also analyzed the detailed properties such as length and volume of extracted vessels. This thesis presented a robust, highly accurate, and fast vascular tracing method that can be applied to the KESM system. The accuracy of the tracing method was validated against synthetic vascular data. The processing time was tested on different sub-networks of the KESM data set which demonstrated that the algorithm is more efficient than previous approaches. The vascular tracing method proposed in this thesis is expected to help trace and analyze large vascular data sets like the KESM data set.

REFERENCES

[1] Y. Choe, L. C. Abbott, D. Han, P. S. Huang, J. Keyser, J. Kwon, D. May-
erich, Z. Melek, and B. H. McCormick. "Knife-edge scanning microscopy: High-
throughput imaging and analysis of massive volumes of biological microstruc-
tures." *High-Throughput Image Reconstruction and Analysis: Intelligent Mi-
croscopy Applications* (2008): 11-37.

[2] Y. Choe, D. Han, P. S. Huang, J. Keyser, J. Kwon, D. Mayerich, and L. C.
Abbott. "Complete submicrometer scans of mouse brain microstructure: Neurons
and vasculatures." *2009 Neuroscience Meeting Planner, (Chicago, IL: Society for
neuroscience, 2009). Program.* No. 389.10. 2009.

[3] Y. Choe, L. C. Abbott, D. E. Miller, D. Han, H.-F. Yang, J. R. Chung, C. Sung, D.
Mayerich, J. Kwon, K. Micheva, and S. J. Smith. "Multiscale imaging, analysis,
and integration of mouse brain networks." *Neuroscience Meeting Planner. San
Diego, CA: Society for Neuroscience*, 2010.

[4] Y. Choe, D. Mayerich, J. Kwon, D. E. Miller, J. R. Chung, C. Sung, J. Keyser,
and L. C. Abbott. "Knife-edge scanning microscopy for connectomics research."
Neural Networks (IJCNN), The 2011 International Joint Conference on. IEEE,
2011.

[5] J. R. Chung, C. Sung, D. Mayerich, J. Kwon, D. E. Miller, T. Huffman, J.
Keyser, L. C. Abbott, and Y. Choe. "Multiscale exploration of mouse brain
microstructures using the knife-edge scanning microscope brain atlas." *Frontiers
in Neuroinformatics* 5 (2011).

[6] D. Han, J. Keyser, and Y. Choe. "A local maximum intensity projection trac-

ing of vasculature in Knife-Edge Scanning Microscope volume data." *Biomedical Imaging: From Nano to Macro, 2009. ISBI'09. IEEE International Symposium on.* IEEE, 2009.

[7] D. Han. "Rapid 3D tracing of the mouse brain neurovasculature with local maximum intensity projection and moving windows." Ph.D. dissertation, Department of Computer Science and Engineering, Texas A&M University, College Station, Texas, 2009.

[8] J. Kwon, D. Mayerich, Y. Choe, and B. H. McCormick. "Automated lateral sectioning for knife-edge scanning microscopy." *Biomedical Imaging: From Nano to Macro*, 2008. ISBI 2008. 5th IEEE International Symposium on. IEEE, 2008.

[9] D. Mayerich, L. Abbott, and B. McCormick. "Knife-edge scanning microscopy for imaging and reconstruction of three-dimensional anatomical structures of the mouse brain." *Journal of Microscopy,*231.1 (2008): 134-143.

[10] D. M. Mayerich, and J. Keyser. "Filament tracking and encoding for complex biological networks." *Proceedings of the 2008 ACM Symposium on Solid and Physical Modeling.* ACM, 2008.

[11] D. Mayerich, J. Kwon, C. Sung, L. Abbott, J. Keyser, and Y. Choe. "Fast macroscale transmission imaging of microvascular networks using KESM." *Biomedical Optics Express* 2.10 (2011): 2888-2896.

[12] B. H. McCormick. "Development of the brain tissue scanner." *Brain Networks Lab Technical Report* (2002).

[13] N. Otsu. "A threshold selection method from gray-level histograms." *Automatica*, 11.285-296 (1975): 23-27.

[14] K. Palagyi, J. Tschirren, E. Hoffman, and M. Sonka. "Quantitative analysis of pulmonary airway tree structures." *Computers in Biology and Medicine* 36.9 (2006): 974-996.

[15] D. L. Pham, C. Xu, and J. L. Prince. "Current methods in medical image segmentation 1." *Annual Review of Biomedical Engineering* 2.1 (2000): 315-337.

[16] H. Samet, and M. Tamminen. "Efficient component labeling of images of arbitrary dimension represented by linear bintrees." *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 10.4 (1988): 579-586.

[17] J. Shan, H. D. Cheng, and Y. Wang. "A novel automatic seed point selection algorithm for breast ultrasound images." *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on.* IEEE, 2008.

[18] R. Xiao, J. Yang, M. Goyal, Y. Liu, and Y. Wang. "Automatic vasculature identification in coronary angiograms by adaptive geometrical tracking." *Computational and Mathematical Methods in Medicine 2013* (2013).

[19] Y. Zhang, X. Zhou, A. Degterev, M. Lipinski, D. Adjeroh, J. Yuan, and S. T. Wong. "A novel tracing algorithm for high throughput imaging: Screening of neuron-based assays." *Journal of Neuroscience Methods* 160.1 (2007): 149-162.