

© 2006 by Sang-Chul Lee. All rights reserved.

THREE-DIMENSIONAL VOLUME RECONSTRUCTION FROM FLUORESCENT
CONFOCAL LASER SCANNING MICROSCOPY IMAGERY

BY

SANG-CHUL LEE

B.E., Inha University, 1998

M.S., University of Illinois at Urbana-Champaign, 2003

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Computer Science
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2006

Urbana, Illinois

Abstract

In this dissertation, I present a problem of three-dimensional volume reconstruction from fluorescent confocal laser scanning microscopy (CLSM) imagery. I overview a three-dimensional volume reconstruction framework which consists of (a) volume reconstruction procedures using multiple automation levels, feature types, and feature dimensionalities, (b) a data-driven registration decision support system, (c) an evaluation study of registration accuracy, and (d) a novel intensity enhancement technique for 3D CLSM volumes.

The motivation for developing the framework came from the lack of 3D volume reconstruction techniques for CLSM image modality. The 3D volume reconstruction problem is challenging due to significant variations of intensity and shape of cross sectioned structures, unpredictable and inhomogeneous geometrical warping during medical specimen preparation, and an absence of external fiduciary markers. The framework addresses the problem of automation in the presence of the above challenges as they are frequently encountered during CLSM-based 3D volume reconstructions used for cell biology investigations.

The objectives of the presented three-dimensional volume reconstruction framework are summarized as follows: (1) automate alignment of sub-volumes (physical sections) from multiple cross sections, (2) obtain high resolution image frames by mosaicking (i.e., stitching together), (3) quantify the accuracy of volume reconstruction using multiple techniques, and (4) visualize the reconstructed volumes in three-dimensional environments for visual inspection and quantitative interpretation.

In this dissertation, the three-dimensional sub-volume registration problem is viewed primarily as an alignment problem. It is approached by extracting two- or three-dimensional

features from each sub-volume and registering the sub-volumes based on the analysis of detected features. I present three sets of techniques classified as pre-processing, main-processing, and post-processing techniques for 3D volume reconstruction. First, the pre-processing steps include (a) sub-volume intensity analysis for image frame selection and feature detection, (b) tile mosaicking using different automation levels and user expertise followed by accuracy evaluation, (c) 2D region or 3D volume segmentation using disk/sphere-based region/volume growing technique, and (d) feature detection based on 2D or 3D segmentation for accurate feature matching and registration alignment optimization. Second, the main-processing steps aim at achieving the most accurate sub-volume alignment, and include (a) feature matching (feature correspondence) using different levels of automation and collaborative mechanisms with web services followed by accuracy evaluations, (b) registration refinement based on different registration accuracy evaluation criteria, (c) optimal global transformation estimation, and (d) sub-volume transformation to construct a 3D volume for visualization. Finally, the volume post-processing step enhances visual saliency of the reconstructed 3D volume by minimizing distortions of the local image intensities (e.g., gradients of edges), and provides comparative results for enhancement with the existing methods using several image quality assessment metrics.

The primary contribution of this dissertation is the presentation of a new theoretical model for three-dimensional volume reconstruction that includes reconstruction methodology, a data-driven registration decision support, automation, intensity enhancement for processing volumetric image data from fluorescent confocal laser scanning microscopes (CLSM). Researched methods have been fully implemented in the Image to Knowledge (I2K) software package developed at the National Center for Supercomputing Applications (NCSA).

The broader impact of my work is in providing the algorithms in a form of web-enabled tools to the medical community so that medical researchers can minimize laborious and time intensive 3D volume reconstructions using the tools and computational resources at NCSA.

To my family.

Acknowledgements

First of all, I would like to thank my research advisor, Dr. Peter Bajcsy, who I indebted for his encouragement and advice to accomplish this project. The quality of work during my Ph.D. study was greatly improved by his comments and guidance. I would also like to thank Professor Narendra Ahuja who directed my dissertation and provided very valuable comments for my Ph.D. research.

I have been fortunate to have had many excellent collaborative researchers who helped me to boost the progress of my Ph.D. work. I especially thank Rob Kooper for his friendly help toward my research, technical support, and social interactions in our research group. I also thank Dr. Robert Folberg and Dr. Amy Lin at the department of Pathology, University of Illinois at Chicago, for their responsive collaborative work and discussion of significant research. I also thank Marquita Miller and Bruce Mather for their very friendly support.

My dissertation was greatly improved by the comments from my committee members, Professor Thomas Huang, Professor David Forsyth, Professor John Hart, and Professor Sylvian Ray, who I thank very much.

I am grateful for the many friends at the University of Illinois at Urbana-Champaign who have made my life more enjoyable. I thank Jongwoo Lim, Myeong-wuk Jang, Jin-yeop Chang, Seung Yi, Hyung-Seok Hahm, Hwanjo Yu, and Darrin Trulock.

Finally, I thank my wife, Eun Joo Choi, who has given me endless support at all her sacrifices.

This work was partially supported by the National Institute of Health under Grant No. R01 EY10457 and by the National Center for Supercomputing Applications (NCSA).

Table of Contents

| | |
|--|-----------|
| List of Tables | ix |
| List of Figures | xi |
| Chapter 1 INTRODUCTION | 1 |
| 1.1 Problem Statements | 1 |
| 1.2 Motivation | 4 |
| 1.3 Applications | 5 |
| 1.4 Input Data Description | 7 |
| 1.4.1 Data preparation | 7 |
| 1.4.2 Input data properties | 10 |
| 1.5 Dissertation Overview | 11 |
| Chapter 2 FRAMEWORK FOR THREE-DIMENSIONAL VOLUME RE- CONSTRUCTION | 13 |
| 2.1 Introduction | 13 |
| 2.2 Related Work | 14 |
| 2.2.1 Registration approaches | 16 |
| 2.2.2 Registration for CLSM data | 19 |
| 2.3 Overview of Volume Reconstruction from CLSM Imagery | 20 |
| 2.4 Registration Decision Support System | 24 |
| 2.4.1 Impacts of registration decisions | 26 |
| 2.4.2 Optimization tools for registration decisions | 29 |
| 2.4.3 Results | 32 |
| 2.5 Summary | 42 |
| Chapter 3 SUB-VOLUME PRE-PROCESSING | 44 |
| 3.1 Introduction | 44 |
| 3.2 Related Work | 45 |
| 3.3 Sub-volume Intensity Analysis | 47 |
| 3.3.1 Frame selection | 48 |
| 3.4 Tile Mosaicking | 49 |
| 3.5 Segmentation | 52 |
| 3.5.1 Two-dimensional segmentation | 53 |
| 3.5.2 Three-dimensional segmentation | 55 |

| | | |
|-------------------|--|------------|
| 3.6 | Feature Detection | 56 |
| 3.7 | Results | 56 |
| 3.7.1 | Frame selection | 57 |
| 3.7.2 | Tile mosaicking | 59 |
| 3.7.3 | Segmentation | 66 |
| 3.7.4 | Feature detection | 69 |
| 3.8 | Summary | 70 |
| Chapter 4 | SUB-VOLUME ALIGNMENT | 72 |
| 4.1 | Introduction | 72 |
| 4.2 | Related Work | 74 |
| 4.3 | Feature Matching | 77 |
| 4.3.1 | Manual alignment approach using pixel selection | 78 |
| 4.3.2 | Semi-automated alignment approach using region centroids | 78 |
| 4.3.3 | Fully-automated alignment approach | 80 |
| 4.3.4 | Geographically distributed alignment process | 83 |
| 4.4 | Matching Refinements | 92 |
| 4.4.1 | Optional intensity-based optimization for 2D alignment | 92 |
| 4.4.2 | Trajectory fusion for 3D alignment | 94 |
| 4.5 | Global Transformation Estimation and Sub-volume Transformation | 99 |
| 4.6 | Results | 101 |
| 4.6.1 | Feature matching | 101 |
| 4.6.2 | Matching refinements | 113 |
| 4.6.3 | Global transformation estimation and sub-volume transformation | 117 |
| 4.7 | Summary | 119 |
| Chapter 5 | VISUAL ENHANCEMENT OF 3D CLSM VOLUME | 124 |
| 5.1 | Introduction | 124 |
| 5.2 | Related Work | 126 |
| 5.3 | Visual Enhancement | 127 |
| 5.3.1 | 2D filter model | 128 |
| 5.3.2 | Background separation | 130 |
| 5.3.3 | Kernel size optimization | 133 |
| 5.3.4 | Speckle noise removal and intensity rescaling | 134 |
| 5.4 | Results | 134 |
| 5.4.1 | Simulation | 134 |
| 5.4.2 | Real CLSM images | 141 |
| 5.5 | Summary | 148 |
| Chapter 6 | CONCLUSIONS | 151 |
| References | | 155 |
| Vita | | 167 |

List of Tables

| | | |
|-----|---|-----|
| 2.1 | Summary of registration errors obtained using rigid and affine transformation models for the same set of four matching pairs of points. The residual errors X, Y and (X, Y) are computed as a sum of squared differences between the transformed coordinates of image 1 and the user chosen coordinates in image 2. The original residual error (X, Y) before any transformation was equal to 12.782. | 36 |
| 3.1 | A summary of mosaicking experiments: 3 by 3 tiles with 512 by 512 pixel resolutions. Full automatic methods are performed by the normalized cross-correlation (NC) and by the normalized mutual information (NMI) on Pentium 4, 3.0 GHz. | 65 |
| 3.2 | The paired t-test result for errors of the pixel-based and the feature-based methods in Table reftable-mosaic. | 65 |
| 4.1 | Tradeoffs considered during a prototype system design using web services for 3D volume reconstruction. C refers to a constant, + indicates increase and - is decrease. Multiple + or - symbols indicate the magnitude of the value specified in each column. | 89 |
| 4.2 | A summary of image alignment | 106 |
| 4.3 | The paired t-test result for errors of the pixel-based and the feature-based methods in Table 4.2. | 107 |
| 4.4 | A summary of automatic image alignment experiments without optimization. | 108 |
| 4.5 | Datasets | 109 |
| 4.6 | A summary of automatic image alignment experiments with and without optimization. | 113 |
| 5.1 | Evaluation results for the test image 1 shown in Figure 5.16 using the four metrics defined at the beginning of Section 3.2. The four techniques are represented by abbreviations, such as Histogram equalization (HE), HE with background threshold (HE (threshold)), Contrast Limited Adaptive Histogram Equalization (CLAHE), and Mean Weight Filtering MWF (threshold, filter width). To highlight performance results, the best result is shown as bold and the second best result is shown as italic. | 147 |

| | | |
|-----|---|-----|
| 5.2 | Evaluation results for the test image 2 shown in Figure 5.18 using the four metrics defined at the beginning of Section 3.2. The four techniques are represented by abbreviations, such as Histogram equalization (HE), HE with background threshold (HE (threshold)), Contrast Limited Adaptive Histogram Equalization (CLAHE), and Mean Weight Filtering MWF (threshold, filter width). To highlight performance results, the best result is shown as bold and the second best result is shown as italic. | 147 |
| 5.3 | Comparison using a combined metric formed by summing together normalized C , S , D and N metrics to $[-1, 1]$ | 150 |

List of Figures

| | | |
|-----|---|----|
| 1.1 | An overview of 3D sub-volumes from fluorescent laser scanning confocal microscope images. | 8 |
| 2.1 | The process of 3D volume reconstruction from fluorescent confocal laser scanning microscope 3D sub-volumes. The dashed box indicates an optional step. | 21 |
| 2.2 | Registration decisions and their impact on registration results. | 25 |
| 2.3 | Supporting registration decisions about image sub-area size. A similarity of a pair of sub-areas from the left and right images is evaluated by computing a normalized correlation value as a function of a sub-area location and its size. | 30 |
| 2.4 | The result of processing multiple image sub-areas selected according to Figure 2.3. The resulting graph shows normalized correlation values of a pair of two band images as a function of a sub-area size. | 30 |
| 2.5 | Supporting registration transformation model selection. A choice of rigid or affine transformation model is evaluated by reporting a registration transformation error and by inspecting the level of distortion in the overlaid registered images. The red and blue disks are used for establishing matching pairs in a pair of images. The left panel shows a pair of original images (overlaid and inspected by moving the transparency slider bar below the panel). The right panel illustrates the registered sets of red and blue disks and the overlaid images after transformation. | 31 |
| 2.6 | Two images used for demonstrating the impact of image sub-area versus entire image registration tradeoffs (image spatial size decision). The images came from one stack of CLSM images and hence the ground truth alignment was known. The slides are 13 frames apart (1.7 microns distance along z-axis) , and each pixel represent 0.478 microns (536 microns by 527 microns for the entire image). | 34 |
| 2.7 | Two binary images derived from the images in Figure 10 by intensity thresholding to demonstrate the impact of image sub-area versus entire image registration tradeoffs (image spatial size decision) for the case of zero intensity variation. | 35 |

| | | |
|------|---|----|
| 2.8 | Normalized correlation coefficient as a function of image sub-area size in pixels for two perfectly aligned images that are 13 frames apart (1.7 microns apart), each image is of pixel size 1123 by 1104 corresponding to spatial resolution 536 microns by 527 microns (entire image). The two curves correspond to original image (blue) and thresholded image (pink). The images are shown in Figure 2.6 and Figure 2.7. The normalized correlation coefficient represents a quantitative metric (similar to visual inspection) that ranges from [0, 1]. For our data set, we decided to use the entire image area equal to 561x551 pixels. | 35 |
| 2.9 | Intensity similarity (normalized cross correlation coefficient) using different low-pass filter size for pairs of frames. The x-axis represents the frame number which is compared with the frame 1, and each curve represents a normalized cross correlation coefficient with different filter size. | 38 |
| 2.10 | The similarity of image pair, frame 1 and frame 24, as a function of low-pass filter size. | 38 |
| 2.11 | The similarity of adjacent image frames. | 39 |
| 2.12 | Image frames from 1 to 24 | 39 |
| 3.1 | Adjacent tiles of CLSM images: overlapping regions have few features (closed regions.) | 51 |
| 3.2 | Image mosaicking problem: (left) Image tiles with colored borders and (right) mosaicked image showing where each tile belongs in the final image based on its color. | 52 |
| 3.3 | Illustration of disk-based segmentation. A segment is formed as a connected set of pixels covered by a disk while the disk moves within the fluorescent boundary. Depending on the disk diameter and contour gaps a segment is detected or not. | 53 |
| 3.4 | Illustration of multiple contour interpretations. A partially closed contour could lead to three detection outcomes depending on a disk diameter. | 54 |
| 3.5 | 3D region growing for segmentation of a volumetric region using a sphere. | 55 |
| 3.6 | An example of a pair of sub-volumes after 3D segmentation. | 57 |
| 3.7 | Evaluation of image frame selection from within each 3D sub-volume to be used for alignment of the sub-volumes. The graphs show combined visual saliency score as a function of image frame for two 3D volumes. The frame with maximum visual saliency score within each sub-volume would be used for alignment. | 58 |
| 3.8 | Evaluation of inter-frame similarity for two spatially adjacent sub-volumes with 22 frames after alignment. The similarity is measured by normalized correlation (vertical axis). The horizontal axis refers to the pairs of frames that start with the end frames of sub-volumes (21 – 0 \equiv (Sub-Vol 1, frame 21)–(Sub-Vol 2, frame 0)) and finish with the middle frames (11 – 10 \equiv (Sub-Vol 1, frame 11)– (Sub-Vol 2, frame 10)). | 60 |
| 3.9 | Software interface for (a) manual mosaicking, and (b) semi-automatic mosaicking with highlighted regions. | 64 |

| | | |
|------|---|-----|
| 3.10 | Mosaicking errors for all human subjects performing pixel-based (manual), feature-based (semi-automated), and intensity-based (fully automated) tile mosaicking. | 64 |
| 3.11 | Example images that have to be aligned. Human tonsil tissue stained for laminin. | 67 |
| 3.12 | Segmentation of input images shown in Figure 3.11. Segmentation is performed by thresholding that is followed by connectivity analysis with a disk. The two images illustrate results obtained with different disk parameters [Left image - T.S. (threshold S) = 10, right image - T.T. (threshold T) = 8, M.R. (minimum size of a region) = 80, and D.D. (disk diameter) = 3]. | 68 |
| 3.13 | A pair of CLSM sub-volumes: (left) sub-volume 1 (upper physical section) and (right) sub-volume 2 (lower physical section). | 68 |
| 3.14 | Segmented sub-volumes of Figure 3.13. | 69 |
| 3.15 | The centroids are labeled from the segments shown in Figure 3.12. | 69 |
| 3.16 | A pair of sets of (exaggerated) trajectories from adjacent sub-volumes. Numbers indicate the label of matching trajectories. Red line segments indicate centroids, and green curves indicate a polynomial function with degree 3. | 70 |
| 4.1 | Morphology quantification in a CLSM stack: x-axis represents a frame index (along depth) which is compared with the first frame. | 79 |
| 4.2 | Illustrations of the case when distance-based matching leads to an erroneous match of the segments labeled as 3 (left) and as 4 (right). | 81 |
| 4.3 | Illustration of the correspondence problem for two sets of features S_i and T_j with unequal number of features $N^S = 3$ (left) and $N^T = 5$ (right). Segments are shown as disks characterized by their index i , area a_i^S and centroid location \bar{c}_i^S . Dashed lines represent the Euclidean distances between any two centroid locations. | 82 |
| 4.4 | Illustration of distance-based matching for two pivot segments $S_i = \{\bar{c}_i^S, a_i^S\}$ and $T_j = \{\bar{c}_j^T, a_j^T\}$. In order to find the best match, distance and area ratio of pairs of segments are compared to satisfy $D_{ij}(u, v) < \delta_1$ and $A_{ij}(u, v) < \varepsilon_1$ | 83 |
| 4.5 | An overview of the current application scenario. | 85 |
| 4.6 | The overall architecture of the proposed system. | 86 |
| 4.7 | Image selection menu. | 90 |
| 4.8 | Selection of registration points using the centroid feature approach. Note that sub-areas the large images in Figure 4.7 are shown. ("+" marks represents the region centroids) | 91 |
| 4.9 | Job execution window after pressing the button "View Result". | 91 |
| 4.10 | Overview of the trajectory-based alignment. | 95 |
| 4.11 | Trajectory fusion by connecting end points | 97 |
| 4.12 | Trajectory fusion by residual minimization of a fitted model. | 98 |
| 4.13 | Three pairs of image examples used for alignment evaluation | 105 |
| 4.14 | Illustration of a strong correlation between the compactness measure and the alignment error. | 106 |

| | | |
|------|--|-----|
| 4.15 | Alignment errors for all human trials including pixel-based (manual) and feature-based (semi-automatic) alignment. | 107 |
| 4.16 | The correspondence outcome from two phases for the segments. The left and right images are to be aligned. Overlays illustrate established correspondences between segments that are labeled from 1 through 17. The centroid locations of segments are sorted based on the correspondence error from the smallest error to the largest error. | 109 |
| 4.17 | The result of automated feature selection after processed to establish segment correspondences shown in Figure 4.16. The circles of different colors represent three pairs of centroids selected automatically according to the compactness measure defined in Section 4.5. | 110 |
| 4.18 | The images after mosaicking. The edges of each tile are visualized using green lines. On the left is a frame from dataset 1 and on the right is a frame from dataset 2. | 111 |
| 4.19 | 3D view of the reconstructed 3D volume using web service. | 112 |
| 4.20 | A pair of sets of (exaggerated) trajectories from adjacent sub-volumes. Numbers indicate the label of matching trajectories. Red line segments indicate centroids, and green curves indicate a polynomial function with degree 3. . . | 115 |
| 4.21 | Trajectory discontinuity as function of physical gaps between adjacent sub-volumes: EX() and RM() refer to extrapolation method and residual minimization method respectively. Note that extrapolation method is plotted as dotted lines, and residual minimization method is plotted by solid lines. . . . | 116 |
| 4.22 | Trajectory residual as function of physical gaps between adjacent sub-volumes: EX() and RM() refer to extrapolation method and residual minimization method respectively. Note that extrapolation method is plotted as dotted lines, and residual minimization method is plotted by solid lines. | 117 |
| 4.23 | (left) Final 3D volume without trajectory fusion (gap = 1), and (right) sub-region at the box in the left image. | 118 |
| 4.24 | (left) Final 3D volume after trajectory fusion with extrapolation (gap = 1, linear model), and (right) sub-region at the box in the left image. | 118 |
| 4.25 | (left) Final 3D volume after trajectory fusion with residual minimization (gap = 1, cubic model), and (right) sub-region at the box in the left image. . . . | 119 |
| 4.26 | 3D view of the reconstructed 3D volume from four consecutive sub-volumes of CLSM image stacks (top view). | 120 |
| 4.27 | 3D view of the reconstructed 3D volume from four consecutive sub-volumes of CLSM image stacks (side view). | 121 |
| 5.1 | Intensity Correction Problem: (a) measured intensity profile for CLSM images with intensity bias (solid curve) and (b) corrected intensity profile. Intensity bias has been corrected while preserving local intensity gradients) | 125 |
| 5.2 | Intensity distribution for a CLSM image. $\Phi(v)$ is illustrated as a Gaussian function. | 131 |
| 5.3 | Sum of the squared error of the estimated function for (a) and (b). The intensity value k minimizing R are 26 and 9 respectively. | 132 |

| | | |
|------|--|-----|
| 5.4 | Images of intensity bias fields (intensity distortion maps.) | 135 |
| 5.5 | Simulation images: a bias-free image with known structures. | 135 |
| 5.6 | Simulation images: three images distorted by applying the bias fields shown in (a), (b) and (c) in Figure 5.4. Thickness and the spacing of the bar structures are decreasing toward the lower right corner. | 136 |
| 5.7 | Background thresholding. | 137 |
| 5.8 | (a) Contrast C , (b) high frequency distortion D , and (c) image saliency as a function of the kernel size. The curves are computed for image data shown in Figure 5.6 (a), (b) and (c). | 137 |
| 5.9 | Intensity correction maps with (a) $\sigma = 20; a = b = 33$, (b) $\sigma = 20; a = b = 19$, and (c) $\sigma = 20; a = b = 47$. Note that the weight is pseudo colored, where blue denotes low and red denotes high values. | 138 |
| 5.10 | Intensity corrected images using the Mean-weight filtering from Figure 5.9. For each of the three images, the intensity bias has disappeared. | 138 |
| 5.11 | Result of Histogram Equalization | 139 |
| 5.12 | Result of Histogram Equalization with background thresholding | 139 |
| 5.13 | Result of CLAHE. | 140 |
| 5.14 | Intensity profile along the red line in Figure 5.5. | 140 |
| 5.15 | Image Saliency Measure for automatic kernel selection: the optimal kernel sizes are 15 and 21 for the two test image shown in Figure 5.16 (a) and Figure 5.18 (a). | 142 |
| 5.16 | CLSM test image 1: (a) Original test image, (b) histogram equalization, (c) histogram equalization with background thresholding, (d) CLAHE ("adapthis-teq()") with default setting by Matlab Image Processing Toolbox, Build. R12), (e) mean weight filtering, and (f) intensity correction maps by the mean weight filtering. | 143 |
| 5.17 | Sub-region of Intensity corrected images in Figure 5.16 (From left, original image, Histogram Equalization, Histogram Equalization with background threshold, CLAHE, and Mean-weight filtering.) | 144 |
| 5.18 | CLSM test image 2: (a) Original test image, (b) histogram equalization, (c) histogram equalization with background thresholding, (d) CLAHE ("adapthis-teq()") with default setting by Matlab Image Processing Toolbox, Build. R12), (e) mean weight filtering, and (f) intensity correction maps by the mean weight filtering. | 144 |
| 5.19 | Sub-region of Intensity corrected images in Figure 5.18: (a) original image, (b) Histogram Equalization, (c) Histogram Equalization with background threshold, (d) CLAHE, and (e) Mean-weight filtering. | 145 |
| 5.20 | High frequency distortion D (a) for Figure 5.16 and (b) Figure 5.18. | 146 |
| 5.21 | Low frequency intensity heterogeneity S for the original test image 1 in Figure 5.16 (results are shown in the left graph denoted as (a)) and the original test image 2 in Figure 5.18 (results are shown in the right graph denoted as (b)). | 146 |
| 5.22 | Mean weight filtering with different kernel sizes: From left, the kernel width is equals to 3,7,9,21,51. | 149 |

*Things which we see are not by themselves what we see ...
It remains completely unknown to us what the objects may be by themselves and
apart from the receptivity of our senses. We know nothing but our manner of
perceiving them.*

- Immanuel Kant

Chapter 1

INTRODUCTION

1.1 Problem Statements

Manual three-dimensional volume reconstruction from stacks of confocal laser scanning microscopy (CLSM) images obtained from the study of serial histological cross sections is considered as an extremely time consuming process. Reconstruction is characterized by (a) significant variations in intensity and in shape of corresponding structures in multiple cross-sections, (b) unpredictable and frequently inhomogeneous geometrical warping during specimen preparation, and (c) lack of internal fiduciary markers for alignment. Given this characteristics, one could approach the problem by selecting the most cost-effective methods with achieving maximum reconstruction quality, e.g., alignment accuracy. For example, with automation, it may be possible to save time and to achieve consistency in three-dimensional reconstructions which would not be possible with manual reconstruction methods.

In this dissertation, I present a novel framework for performing three-dimensional volume reconstruction from fluorescent confocal laser scanning microscopy (CLSM) imagery. The presented framework consists of (1) automation procedures using 2D or 3D information from input cross sections, (2) performance evaluation methodology and experimental result from expert study, and (3) a data-driven registration decision support system for optimal registration performance.

The automation procedure addresses the problem occurred in pre-processing, main-processing, and post-processing. First, the goals of the pre-processing step are the high resolution image construction (denoted as mosaicking) and registration feature detection by

segmentation (2D and 3D). Second, the main-process finds the most optimal transformation by matching detected features in multiple automation levels, refining the coordinates of matching features using intensity or structural morphology, and estimating a global transformation. Finally, in post-processing, individual image frames are processed for visual enhancement by removing spatial intensity heterogeneity and preserving fine edge intensity gradient. To assess the registration accuracy, I provide a performance evaluation methodology and evaluation results in different levels of automation and medical expertise. The registration decision support system provides input data analyses on image sub-area size, transformation model selection, registration feature selection, and automation levels that affect registration performance in terms of alignment accuracy, uncertainty of alignment results, repeatability of alignment, and computation requirements. The framework provides medical researchers with guidelines about an optimal volume reconstruction procedure, and addresses related computer science problems. Next, I describe the main claims of the dissertation as follows.

- **Three-dimensional volume reconstruction framework:** A new framework for three-dimensional volume reconstruction from CLSM imagery has been introduced that incorporates automation procedures, performance evaluation, and registration decision support system (Chapter 2).
- **Automation procedure:** An automation procedure has been developed by incorporating (a) different automation levels (e.g., manual, semi-automated, and fully-automated), (b) measurement space (e.g., intensity or morphology), and (c) feature dimensionality (e.g., centroid point or trajectory). It incorporates data-driven registration decisions to achieve the most accurate and cost-efficient three-dimensional volume reconstruction. (Chapter 3 and Chapter 4)
- **Registration decision support system:** A data-driven registration decision support system has been designed. The registration decision support system includes

multiple variables which make impact on registration accuracy, such as image size selection, rigid or affine transformation model, intensity or morphological invariant feature selection, and manual (pixel-based) or semi-automated (centroid-based) automation level. Specifically, it provides multiple paths for performing 3D volume reconstruction based on tradeoff analyses of available computational resources and data characteristics (Section 2.4).

- **Accuracy Evaluation:** Registration accuracy has been evaluated by studying users in multiple domains (e.g., medical doctors, computer experts, and novices). I provide precise evaluation metrics that show objective accuracy evaluation for our simulation and real experimental results (Section 3.7 and Section 4.6).
- **Sub-volume pre-processing techniques:** Sub-volume pre-processing steps have been developed, and they include automated frame selection, tile mosaicking using multiple automation levels, 2D and 3D segmentation, and feature detection (Chapter 3).
- **Automated feature matching technique:** A new robust fully-automated two-dimensional feature matching method is presented. The method automatically establishes feature correspondence using area and qualitative configurations of feature centroids by solving an extension of the Procrustes problem [24] (Section 4.3.3).
- **Web-based collaborative tools:** A web-based sub-volume alignment method is provided. The proposed web service based solution provides a collaborative mechanism for 3D volume reconstruction (Section 4.3.4).
- **Fully automated registration technique:** Novel automated sub-volume registration methods are presented. Based upon the optimality criteria, two-dimensional (e.g., centroid) and three-dimensional (e.g., trajectory of centroids) features have been used

for automated sub-volume registration. I also discuss some factors that affect registration accuracy in experimental results (Section 4.4).

- **Intensity correction technique:** A novel intensity correction method has been developed. The proposed method solves the spatial intensity heterogeneity problem which is known as a major barrier against accurate analysis of fluorescent CLSM data. I presented simulation and real data experiments using objective performance evaluation metrics. Although the technique is used for visual enhancement as a post-processing, it can be further used in image pre-processing step for more robust intensity-based mosaicking and segmentation (Chapter 5).

The developed framework was applied to the problem of 3D volume reconstruction of blood vessels and vasculogenic mimicry patterns in uveal melanoma from serial fluorescent labeled paraffin sections labeled with antibodies to CD34 and laminin and studied by confocal laser scanning microscopy. The medical significance of the 3D volume reconstruction work lies in discovering looping patterns stained in histological cross sections. It is the 3D visual interpretation and quantitative evaluation that would lead to successful uveal melanoma detection.

1.2 Motivation

The problem of 3D volume reconstruction can be found in multiple application domains, such as, medicine, mineralogy, or surface material science. In almost all applications, the overarching goal is to automate a 3D volume reconstruction process while achieving at least the accuracy of a human operator. The benefits of automation include not only the cost of human operators but also the improved consistency of reconstruction and the eliminated training time of operators.

In medical domain, manual 3D volume reconstruction from stacks of laser scanning confocal microscopy images obtained from the study of serial histological sections is time consum-

ing process. Reconstruction is complicated by potentially significant variations in intensity and shape of corresponding structures, unpredictable and frequently inhomogeneous geometrical warping during specimen preparation, and an absence of internal fiduciary markers for alignment.

Given this assumption, 3D volume reconstruction can be performed by either a semi-automated method with improved accuracy in comparison to manual method [85] or by a fully automated 3D volume reconstruction while achieving at least the accuracy of a human operator. Through automation, it may be possible to save time and to achieve consistency in 3D reconstructions which may not be possible with human-assisted reconstruction methods.

In this dissertation, I study a 3D volume reconstruction framework from stack of confocal laser scanning microscopy images with quantitative performance evaluation of fully automatic, semi-automatic and manual methods [85]. In results, I present experiments with fluorescent confocal laser scanning microscope imagery used for mapping the distribution of extracellular matrix proteins in serial histological sections of uveal melanoma [23, 111].

1.3 Applications

Quantitative analysis of tissue components from archival paraffin blocks: The automated three-dimensional volume reconstruction technique is applied to the problem of comparing the surface area of non-endothelial cell lined, laminin-rich looping vasculogenic mimicry (VM) patterns with the surface area of endothelial cell lined vessels in metastatic uveal melanoma to the liver. VM patterns are known to transmit fluid *in vivo* and contribute to the perfusion of many tumor types. After labeling sections with antibodies to CD34 and laminin, the surface area of VM patterns to vessels was calculated by segmenting out structures that labeled with laminin but not with CD34 from those structures labeling with CD34 or CD34 and laminin. In metastatic uveal melanoma tissues featuring co-localization of

high microvascular density [66.4 microvessels adjusted for 0.313 mm² area (range 56.7-72.7)] and vasculogenic mimicry patterning, the surface area of vasculogenic mimicry patterns was 11.6 fold greater (range 10.8-14.1) than the surface provided by CD34-positive vessels. The techniques used in this study revealed tissue characteristics that could not be identified from 2D sections alone. These techniques may be extended to the visualization and quantification of molecular markers in a variety of pathological entities from archival paraffin embedded tissues.

Remote collaboration: The developed 3D volume reconstruction tools are web-enabled, and applied to the problem of remote collaboration. The use of web services is motivated by the fact that the problem of 3D medical volume reconstruction requires significant computer resources and human expertise in medical and computer science areas. In the collaboration between UIC and NCSA, pre-processed input images at NCSA are made accessible to medical collaborators at UIC for registration. Every time UIC medical collaborators inspected images and selected corresponding features for registration, the web service at NCSA is contacted and the registration processing query is executed using the Image to Learn library of registration methods. The co-registered frames are returned for verification by medical collaborators in a new window.

Software services for registration decision: The goal is to provide data-driven mechanisms for evaluating the tradeoffs between accuracy of 3D volume reconstructions and registration variables. The support developed for registration decisions about 3D volume reconstruction is available to the general community with the access to the NCSA HPC resources.

1.4 Input Data Description

In this section, I present the process of input data preparation for 3D volume reconstruction and the assumptions of the acquired data sets.

1.4.1 Data preparation

It is apparent that using artificially inserted fiduciary markers allows automating 3D volume reconstruction while keeping low registration error. However, there still exist medical experiments with CLSM, where fiduciary markers cannot be inserted into a specimen. For example, the placement of fiduciary markers in a paraffin-embedded tissue is problematic. The introduction of markers internally may distort the tissue and areas of interest. On the other hand, markers placed outside the tissue may migrate during sectioning or expanding the paraffin. The composition of the markers also poses challenges. Rigid material, such as suture, may fragment or distort the tissue when sections are cut. In addition to attempting to locate fiduciary markers into tissues using the aforementioned techniques, it is also attempted to insert small cylindrical segments of "donor tissue" from paraffin-embedded tissues according to the techniques used to construct tissue microarrays [106]. It is discovered that the round outlines of the donor tissue cores were inconsistent between tissue sections, making it impossible to use these donor samples as reliable internal fiduciary markers.

Although the 3D volume reconstruction problem without artificially inserted fiduciary markers into paraffin-embedded tissue is addressed, medical experts still need to identify an internal specimen structure for registration that would be visually salient. For this purpose, tonsil tissue was used because it contained structures of interest, e.g., blood vessels. The tonsillar crypts provided a complex edge against which alignment was possible, and the epithelial basement membrane followed its contour. The blood vessels were stained with an antibody to laminin that also stained the epithelial basement membrane. Therefore, by using the epithelial basement membrane - a normal constituent of the tissue - as the visually

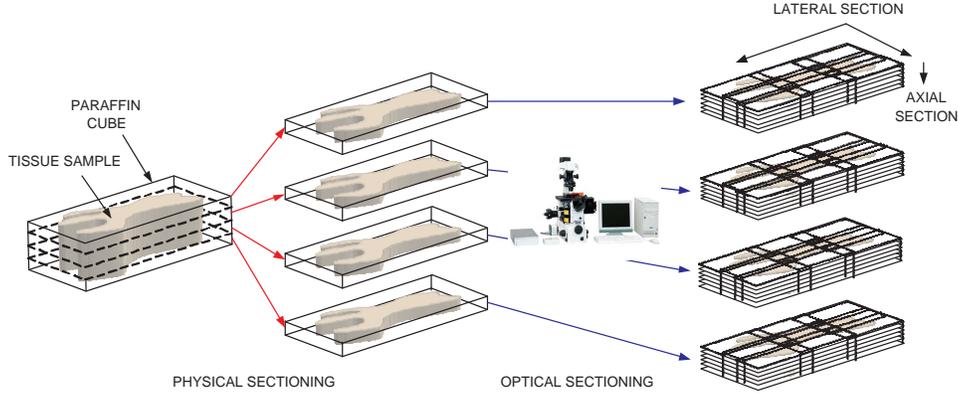


Figure 1.1: An overview of 3D sub-volumes from fluorescent laser scanning confocal microscope images.

salient registration feature in the input CLSM image, medical experts were able to align the tissue sections. Thus, CLSM images of tonsil tissue sections were used for 3D volume reconstruction accuracy evaluations.

In order to demonstrate the issues related to selecting registration variables, we consider 3D volume reconstruction of blood vessels in histological sections of uveal melanoma [21] from paraffin-embedded serial sections labeled with antibodies to CD34 and laminin and studied by confocal laser scanning microscopy (CLSM) imagery [15, 70]. Input data is prepared in following steps as illustrated in Figure 1.1.

1. Formalin-fixed, paraffin-embedded uveal melanoma tissue samples were sectioned at $4\mu m$ thickness. These sections are denoted as *physical sections* or *histologic serial sections*. Note that the slide can be slightly distorted by shear due to slicing procedure.
2. Slides were deparaffinized in xylene and dehydrated through a decreasing ethanol gradient. Slides were rinsed in distilled water followed by antigen unmasking using Target Retrieval Solution 10× Concentrated (DAKO, Carpinteria, CA) according to the manufacturer’s instructions and then rinsed in Phosphate Buffered Saline (PBS) for 5 minutes. Slides were incubated with monoclonal mouse anti-laminin antibody Sigma L8271, clone LAM 89 (Sigma, St. Louis, MO) at a dilution titer of 1:200 for 30 minutes

at room temperature. Slides were rinsed in protein blocking solution (DAKO) for ten minutes followed by detection with Alexa Fluor 488 goat anti-mouse IgG (Molecular Probes, Eugene, OR) for 30 minutes at a dilution of 1:400. Slides were rinsed in buffer then mounted in Faramount Aqueous Mounting Medium (DAKO). For all staining procedures, secondary antibody was omitted in negative controls.

3. All histologic serial sections were examined with a Leica SP2 laser scanning confocal microscope (Leica, Heidelberg, Germany) using the 40 \times objective. Images were stored in tagged information file format (TIFF). Each physical section is imaged by changing optical focal length and spatial field of view as following:
 - (a) In a spatial region (lateral section), acquire a set of depth adjacent images (also denoted as *frames* or *optical sections*) by only changing focal length (axial section) of the CLSM.
 - (b) Mechanically move the specimen stage to change the field of view (e.g., image different physical regions) by increasing/decreasing x and/or y offset of the slide, i.e., translational transformation (two parameters t_x and t_y). An image from one field of view is denoted as a *tile* and a series of tiles are denoted as *stack*. Note that stacks in lateral plane must have overlapping regions with adjacent stacks from different field of view (experimentally, at least 15% of the tile width should be overlapped). For example, in Figure 1.1, each physical section is image into nine spatially adjacent stacks with some overlapping regions.
 - (c) Perform (a)-(b) to cover entire spatial (x-y) area of the slide.
 - (d) Replace the slide to image next depth adjacent physical section. Note that the slide can be rotated and/or translated in this procedure.
4. Label each image with a slide number (i.e., physical section number, e.g., S2), a frame number (i.e., optical section number or frame depth, e.g., z002), a spatial field of

view (i.e., spatial tile location index, e.g., X1Y2), and imaged fluorescent channel (i.e., wavelength, e.g., ch01). For example, a filename can be labeled as 022704-S4X2Y1-z003-ch00.tif with the acquisition date 02-27-04.

1.4.2 Input data properties

The set of spatial tiles is acquired by CLSM and consists of images that came from one cross section (same axial coordinate) in different lateral coordinates or from multiple cross sections of a 3D volume (different axial coordinates). I assume the following properties of acquired image data.

- All physical sections are assumed to be parallel to the same two-dimensional plane.
- During the slide preparation, a tissue slide could be rotated, translated, and slightly sheared. Thus, I use affine transformation for alignment of physical cross sections.
- Image tiles in the same lateral field of view and physical cross section are perfectly aligned in the axial direction.
- The image tiles are acquired by translational motions of the microscope to capture images in different field of view.
- Identical frame indices have the same physical axial depth if they came from the same physical cross section but different image tiles.
- The same frame index among different stacks (field of view) is assumed as the same physical depth of the imaged slide.
- Fluorescent intensity is heterogeneous in both lateral section and axial section. The intensity heterogeneity model is neither known nor controllable (An example is shown in Figure 2.12).

- Structural deformation is negligible in overlapping regions between spatially adjacent tiles when compared at the same depth (e.g., frame index).

1.5 Dissertation Overview

This chapter focused on the problem statement and input data description used for this dissertation. Chapter 2 summarizes the overall methodology of three-dimensional volume reconstruction from CLSM data [10]¹, and outlines a decision support system to optimize 3D volume reconstruction [9]².

Chapter 3 provides a sub-volume pre-processing steps such as sub-volume intensity analysis, tile mosaicking, two-dimensional region or three-dimensional volume segmentation, and feature detection.

The main-processing step is presented in chapter 4 that pursues the most accurate sub-volume alignment using detected features in pre-processing step. The process includes (a) feature matching (e.g., feature correspondence) [86]³ using different levels of automation [85, 88]⁴⁵ and different geographical experimental setups (e.g., web services) [76]⁶ with precise accuracy evaluation, (b) matching refinement, (c) optimal global transformation estimation, and (d) sub-volume transformation to construct an entire three-dimensional volume for visualization.

Finally, in chapter 5, the volume post-processing step provides a novel method to enhance visual saliency of reconstructed volume by minimizing distortions of the local intensity characteristics of features, e.g., gradients of edges, and comparative results with existing methods by providing objective image quality analysis metrics [87]⁷.

¹Related material is published in Journal of Microscopy, 2006

²Related material is published in SPIE international symposium on Medical Imaging, 2006

³Related material is published in SPIE international symposium on Medical Imaging, 2006

⁴Related material is published in SPIE international symposium on Medical Imaging, 2005

⁵Related material is published in EURASIP JASP, 2006

⁶Related material is published in IEEE International Conference on Web Services, 2005

⁷Related material is published in Journal of Microscopy, 2006

Chapter 5 summarizes the contribution made in this dissertation.

Chapter 2

FRAMEWORK FOR THREE-DIMENSIONAL VOLUME RECONSTRUCTION

2.1 Introduction

In general, a feature-based 3D volume reconstruction without fiducial markers requires performing the following steps. First, define a reference coordinate system for a reference image. Second, determine locations of salient features in multiple input image sets, which we call feature detection. Third, establish correspondences of detected feature sets between adjacent sections (images). This step is also denoted as finding spatial correspondences. Fourth, estimate a registration transformation that will compensate for geometric distortions. Fifth, transform sets of input images with respect to the reference coordinate system using estimated transformation in the previous step. Sixth, evaluate registration accuracy with a selected metric.

In this dissertation, the main objectives of 3D volume reconstruction from CLSM data are (1) to obtain high resolution image frames by mosaicking (stitch together) spatial tiles that came from the same depth of a sub-volume, (2) to align sub-volumes (physical sections) from multiple cross sections, (3) to evaluate the accuracy of 3D volume reconstruction using multiple techniques, and (4) to visualize the reconstructed volumes in 3D environment for user-driven or quantitative analysis. Our 3D volume reconstruction work is based on the assumption that there is no prior information about (a) tile locations and their spatial overlap, (b) locations and types of cross section features, and (c) evaluation methodology, ground truth data, and accuracy metrics.

Section 2.4 addresses the problem of optimal registration decisions during 3D medical

volume reconstruction and their impact on anticipated accuracy of aligned images, uncertainty of obtained results, repeatability of alignment, and computational requirements. The registration decisions include (1) image size used for registration, (2) transformation model, (3) invariant registration feature (intensity or morphology), (4) automation level, (5) evaluations of registration results (multiple metrics and methods for establishing ground truth), and (6) assessment of resources (geographically local or distributed computational resources and human expertise). The goal is to provide data-driven mechanisms for evaluating the tradeoffs between accuracy of 3D volume reconstructions and registration variables. First, I present links between registration decisions and 3D reconstruction results in terms of accuracy, uncertainty, consistency and computational complexity characteristics. Second, I have built software tools that enable geographically distributed researchers to optimize their data-driven registration decisions by using web services and high performance computing (HPC) resources. The support developed for registration decisions about 3D volume reconstruction is available to the general community with the access to the NCSA HPC resources. Next, I illustrate the performance of our registration decision support system by considering 3D volume reconstruction of blood vessels in histological sections of uveal melanoma from serial fluorescent labeled paraffin sections. Finally, I discuss the complexity of building a web-enabled, web services based, data-driven, registration decision support system for 3D volume reconstruction.

2.2 Related Work

In this dissertation, 3D volume reconstruction of CLSM data is considered as a registration problem of spatial- and depth-adjacent image stacks. Generally, medical image registration refers to one-to-one mapping between two coordinate spaces, such that pairs of coordinates correspond to the same anatomical structure [103]. This correspondence can be measured

by hardware, e.g., stereotactic frame [79], or estimated by software analysis of measured images. Without hardware support, medical image registration is typically performed using either extrinsic or intrinsic methods [97].

Extrinsic methods introduce foreign objects into a test specimen, so that the object can be detected easily and relatively accurately after specimen imaging [43, 89, 103, 91, 32, 99, 100, 147]. The foreign objects are also denoted as fiduciary markers. The main advantages of extrinsic methods are in fast and easy detection of corresponding coordinates, relatively accurate registration ($0.5mm$ minimum error [151, 37, 152]), and feasibility of full automation of the registration procedure. However, in many cases of medical CLSM data, for example, uveal melanoma tissues, it is difficult to use fiduciary markers [133]. For example, the introduction of markers internally may distort paraffin embedded tissue and its areas of interest. On the other hand, markers placed outside the tissue may migrate during sectioning or expansion of the paraffin.

Intrinsic methods are performed by using measured image contents, e.g., intensity, edge structures, etc., without introducing artificially inserted objects. Registration is typically performed by intensity-based, feature-based, or moment-based approach [103]. Since intrinsic methods do not require artificially inserted markers, they do not distort tissue structure, and make the preparation process much easier. However, since medical structures are highly deformable and contain high intensity variation, accurate registration without fiduciary markers is considered as a very challenging problem. In our problem, based on the material preparation and laminin-positive structures function (i.e., lit up pixels on the stained tissue) imaged by fluorescent CLSM, the automated 3D reconstruction process without fiduciary markers is based on the following two assumptions. First, at least one frame from each 3D sub-volume contains a set of closed or partially opened visually salient contours representing laminin-positive structures (presence of registration features). Second, certain shape characteristics of these salient contours, e.g., centroid or area, remain invariant under translation and rotation transformations (shape invariance of registration features across

two 4μ sections). In the remainders of section 2.2, I survey different types of registration methods for medical specimen without external fiduciary markers.

2.2.1 Registration approaches

Medical image registration approaches have been categorized in different ways by many surveys [97, 103, 64, 16, 26]. In this dissertation, I used the four main categories according to [26]: (1) point based, (2) feature based, and (3) moment based, and (4) correlation based approaches.

Point based approaches

A registration using a point based approach is performed by establishing coordinates of corresponding points in different images usually by manual or semi-automated methods. Relying on intrinsic features only, the point selection can be performed by user interactive pixel selection of some identifiable points, e.g., blood vessel bifurcation points, intersections of structures, or particularly identifiable topographic points [61, 62, 63].

Given the coordinates of corresponding points and a transformation model, the model parameters are estimated [103] using a singular value decomposition [4], an eigenvalue decomposition [66], or unit quaternions [35, 65].

Point-based approaches significantly reduce erroneous matches, e.g., outliers, if (1) point selection is performed by knowledgeable users and (2) there are enough identifiable features, e.g., intrinsic features with anatomical meaning. However, since the task generally requires user interactions, the manual process requires very intensive labor, a good user interface, and anatomical knowledge [30, 31, 61, 62, 63]. More importantly, the registration quality is heavily dependent upon a software environment, e.g., user interfaces, and the expertise of end users.

Feature based approaches

A registration using feature based approach is achieved by identifying a large enough distinctive anatomical structure, e.g., curve or line, or by detecting multiple sufficiently salient point features. A curve matching method has been mainly used for registering cardiac images with a rigid transformation (also known as the “head-and-hat” method [112]). The method searches for the best matches of a pair of contours iteratively using a rigid transformation after manually extracting the boundary (surface) of structures from images. Advancements of this method using a multi-resolution based approach were proposed in [140] to reduce the computational cost. For smooth surfaces, other researchers presented a good registration quality using crest lines [54, 113], hash tables of geometrical invariants together with the Hough transform [55], or the iterative closest point algorithm [12]. Although curve based registration methods are used in many CT or MR images, the major drawbacks of these methods are as follows: (1) there must exist confidently identifiable features, i.e., limited to some application domains, (2) image segmentation must be performed manually by knowledgeable users, (3) a good initial position needs to be manually defined, (4) local deformations of a template (i.e., an image to be searched) can cause erratic registration if the target structure differs from the template, and (5) the registration accuracy heavily depends on segmentation quality.

Instead of relying on a single detected feature, one can select multiple features, e.g., a set of closed contours, with lower confidence in comparison to contour based methods. Correspondences of detected features can be established by semi-automatic [85] or fully automatic method [86]. Semi-automatic methods display segmented regions for shape matching by users. Next, transformation model parameters are estimated by the least-squares fitting based on the corresponding feature sets. Given a transformation model, (erroneous) outliers can be efficiently removed by examining residuals after fitting. Fully automated methods can be utilized by solving Procrustes problem when a sufficient number of features (e.g.,

closed regions) is detected [52, 68, 78, 131, 121, 24]. The main advantage of using multiple features is that the computation of centroid points after region selection greatly improves the registration accuracy and consistency. This improvement has been documented with semi-automated and fully automated method approaches regardless of user expertise [9].

Moment based approaches

A registration using moment based approaches has been typically used for computing a rigid transformation from a centroid and a principal axis of a set of features [27, 2, 80, 141, 142]. After extracting a set of features, e.g., volumes, surfaces, scattered points, the translation is computed from the centroid of the set of feature coordinates, and the rotation is derived by matching principal axis of the features in the compared images.

Since the accuracy of moment based methods heavily depends on the accuracy of extracted features, semi-automatic method is generally used for feature extraction to remove inconsistently placed features [141, 142]. In addition, it is often the case that the principal axis is not uniquely defined, i.e., features are evenly distributed. For the CLSM data of interest, such as cellular level tissue, the accurate registration might be difficult due to high morphological deformation of structures, intensity heterogeneity, and unpredictable distribution of features, e.g., blood vessels.

Correlation based approaches

A registration using correlation based approaches is usually performed by searching for the optimal transformation parameters that maximize the similarity of compared images. The method usually requires a large amount of computation since it considers full image contents and all possible transformation parameters to compare similarity. For this reason, correlation based methods are considered to be able to be automated if high performance computing resources are available. For more efficient computation, a multi-resolution based method has

been used [6].

The most critical part of accurate registration is the definition of the similarity metric [64]. Some widely used similarity metrics include normalized cross-correlation, normalized mutual information, phase correlation in Fourier domain, and correlation of image variance, optical flow, zero-crossing and so on. [97]. A correlation based approach usually performs well for mono-modal image registration with relatively homogeneous intensity distribution [103]. One should avoid using the correlation based method on images with high noise, morphological distortion, and/or intensity heterogeneity, since the search for similarity maxima could result in erroneous registration by being trapped in its local maxima. The problem of local maxima could be solved by genetic algorithms or simulated annealing [101]. Nevertheless, in our case, high intensity heterogeneity and morphological deformation of CLSM images would still lead to high probability of inaccurate alignments along z-axis (sub-volume alignment).

2.2.2 Registration for CLSM data

The 3D reconstruction of vasculogenic mimicry patterns from CLSM data is considered as a registration problem [97] requiring image pre- and post-processing steps to permit 3D visualization and quantification of these geometrical structures [157]. There is an abundance of registration techniques and publications in the literature [64, 16, 162]. In the medical domain, several 3D volume reconstruction techniques have been developed based on specialized image acquisition procedures, e.g., using a linear differential transformer [1], or truncated pyramid representation [109]. There also exist many commercial tools from multiple vendors that could be used for manual registration. For instance, an overview of 3D registration tools for MRI, CT, confocal, and serial-section data for medical or life-sciences imaging is provided at the Stanford ¹ or at the NIH web sites ². Most of these tools use manual registration

¹Stanford web page with references to 3D volume reconstruction software packages.
<http://biocomp.stanford.edu/3dreconstruction/refs/index.html>
and <http://biocomp.stanford.edu/3dreconstruction/software/>

²NIH website.
<http://www.mwrn.com/guide/image/analysis.htm>

methods because automation of 3D volume reconstruction is very difficult. Some software packages include semi-automated or fully automated 3D volume reconstruction for specific imaging modalities under the assumption that visually salient markers have been inserted artificially in imaged specimens.

Based on my knowledge, there has been limited work on understanding accuracy, uncertainty, consistency and computational complexity characteristics of 3D volume reconstruction and their relationships to registration decisions. The past work usually addressed only certain aspects of registration decisions, for example the choice of transformation models [97], the combination of invariant registration features [18], the image data quality evaluation metrics [155], the choice of shape metrics [118], or the process of geometric (spatial registration related) and radiometric (intensity related) adjustments [122]. The past work has originated primarily from the computer vision community when tackling the problem of matching and alignment from points and frames while modeling rigid motion of objects [41]. For example researchers, such as Pennec and Thirion [113, 114], have developed a theoretical model defining the relationship between uncertainty of a rigid transformation applied to a set of 3D points or 2D frames and the registration accuracy. However, the model is defined for only a very small subset of typical registration decisions.

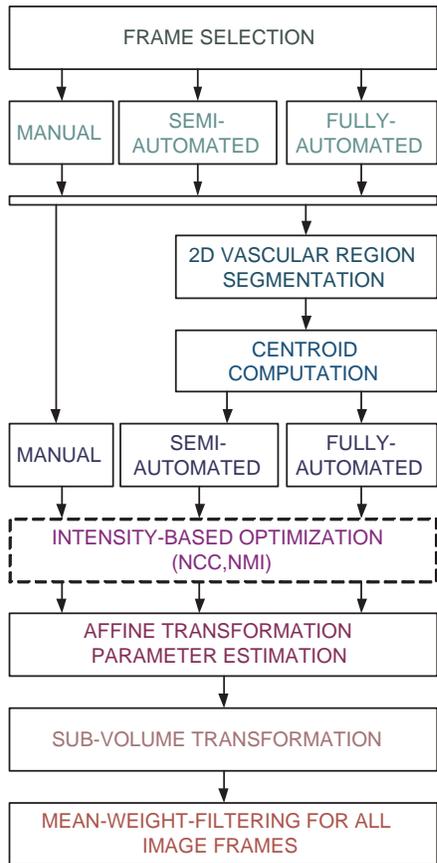
2.3 Overview of Volume Reconstruction from CLSM Imagery

I provide a multi-step 3D volume reconstruction framework. The framework does not assume the presence of fiducial markers since they could not be used according to our medical collaborators. Depending on the dimensionality of registration features, the process of 3D volume reconstruction is divided into two sub-processes as shown in Figure 2.1.

The overall volume reconstruction process is performed in the following steps:

3D VOLUME RECONSTRUCTION PROCESS FROM CLSM IMAGERY

REGISTRATION USING 2D INFORMATION



REGISTRATION USING 3D INFORMATION

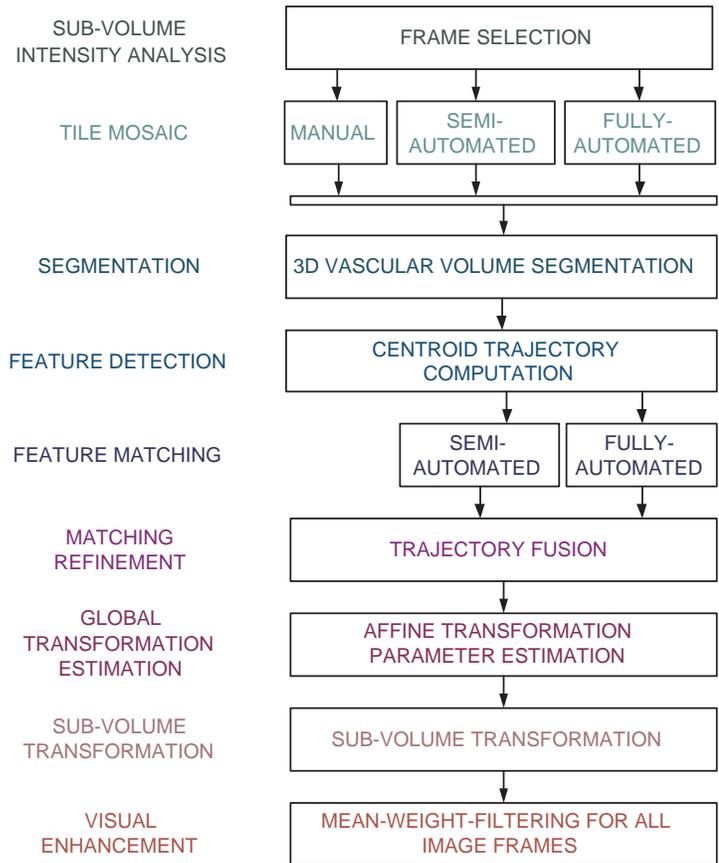


Figure 2.1: The process of 3D volume reconstruction from fluorescent confocal laser scanning microscope 3D sub-volumes. The dashed box indicates an optional step.

1. **Sub-volume intensity analysis:** Analyze image intensity quality from input sub-volumes to select the most salient image frame. The selected frame is used for image mosaicking and alignment process (segmentation, feature detection, feature matching, and matching refinement) when using 2D information. For registration using 3D information, all frames are used to extract 3D features, and therefore, this step is only used for tile mosaicking (Section 3.3).
2. **Tile mosaic:** Mosaic image tiles in lateral coordinate (e.g., x-y coordinate) using manual, semi-automated, or fully automated method. First, manual method refers to manual selections of corresponding pixel coordinates by users. Second, semi-automated method uses 2D centroid features (as the same way in 2D steps in Figure 2.1 "segmentation" and "feature detection") for user region selection. Third, fully automated method is performed by intensity-based correlation method (e.g., normalized cross correlation). Note that the intensity-based method performs relatively well for mosaicking since the selected frames typically present relatively high contrast, entropy, and homogeneous intensity distribution. In addition, based on the assumption from instrumental measurements, mosaicking only requires translational transformation which is feasible to implement in terms of computational complexity. In result, I show a performance evaluation by a user study (Section 3.4).
3. **Segmentation:** Segment out closed or partially opened 2D regions or 3D volumes (Section 3.5).
 - 2D segmentation: 2D (partially) closed regions are segmented out using a region growing method based on a disc. Instrument noise is effectively removed by setting relatively high background threshold with large disk size where background threshold and disk size is selected from data-driven analysis (Section 3.5.1).
 - 3D segmentation: 3D (partially) closed volumes are segmented out using volume growing method based on a 3D sphere. 3D sphere generates a smoother feature

trajectory of a volumetric feature in comparison to a series of 2D centroid regions along axial (z-axis) direction (Section 3.5.2).

4. **Feature detection:** Compute features of segmented regions (e.g., centroids and areas) or volumes (e.g., centroid trajectory). Due to high deformation in cellular structures, precise shape descriptors (such as contour shape) may result in poor matching results. Instead, I approached the problem of feature selection with rather simple region shape descriptors (such as area and centroid) that are relatively more invariant to morphological deformation (Section 3.6).
5. **Feature matching:** Find a set of pairs of matching features. As described in the "feature detection" step, robust features matching is difficult when using individual region descriptors. Therefore, fully automated method uses qualitative configurations of centroid locations to establishing feature correspondence, considered as an extension of the Procrustes problem [24]. Experiments show that this method is more robust to segmentation error and image noise. In this step, I developed manual, semi-automated, and fully automated methods (Section 4.4).
6. **Matching refinement:** Refine matching coordinates based on multiple evaluation categories.
 - 2D refinement: intensity similarity is one of the evaluation categories for finding an optimal (pixel-level) matching of the corresponding features. Since this process is computationally extremely expensive, it was only performed in a relatively small pixel neighborhood of selected features. I discuss the computation cost and accuracy of the results (Section 4.5.1).
 - 3D refinement: the main purpose of 3D refinement is in smoothing sub-volume boundaries across physical cross sections. I used two evaluation metrics, such as structural connectivity and model fitting (Section 4.5.2).

7. **Global transformation estimation:** Transform sub-volumes to the reference coordinate system based on the computed global transformation. Based on the assumptions from data acquisition, I used an affine transformation model for global transformation (Section 4.6).
8. **Visual enhancement:** Enhance voxel intensities of a reconstructed volume for visual inspection purposes. I developed an intensity correction method and evaluation metrics for multiple visual enhancement techniques (Section 5).

2.4 Registration Decision Support System

This section describes data-driven optimization approaches to four registration decisions including image size selection, rigid or affine transformation model, intensity or morphological invariant feature selection, and manual (pixel-based) or automated (centroid-based) automation level. The reason for using data-driven approaches lies in the large variability of objects of interest, specimen preparation, imaging modality, specific instrumentation characteristics, and so on, that is extremely difficult to model analytically with any generality whatsoever. I frame the problem first by presenting the links between 3D reconstruction accuracy and registration variables. Then, I lead to the data-driven methods for optimal choice of registration variables. The experimental results illustrate how the developed web-enabled software would be used by researchers to make the most optimal data-specific registration decisions.

The overview of most common registration decisions is provided in Figure 2.2. In Figure 2.2, one should view vertical black arrows as possible decision outcomes during the registration process and hence any possible combination of user decisions would characterize the obtained registration result. Some of the decisions could be easily expanded, e.g., other transformation models. Other decisions could be elaborated, such as methods for establishing ground truth could be classified into visual inspection, comparison with ground truth data, or measuring the degree of deviations from assumed data model. The purpose of Fig-

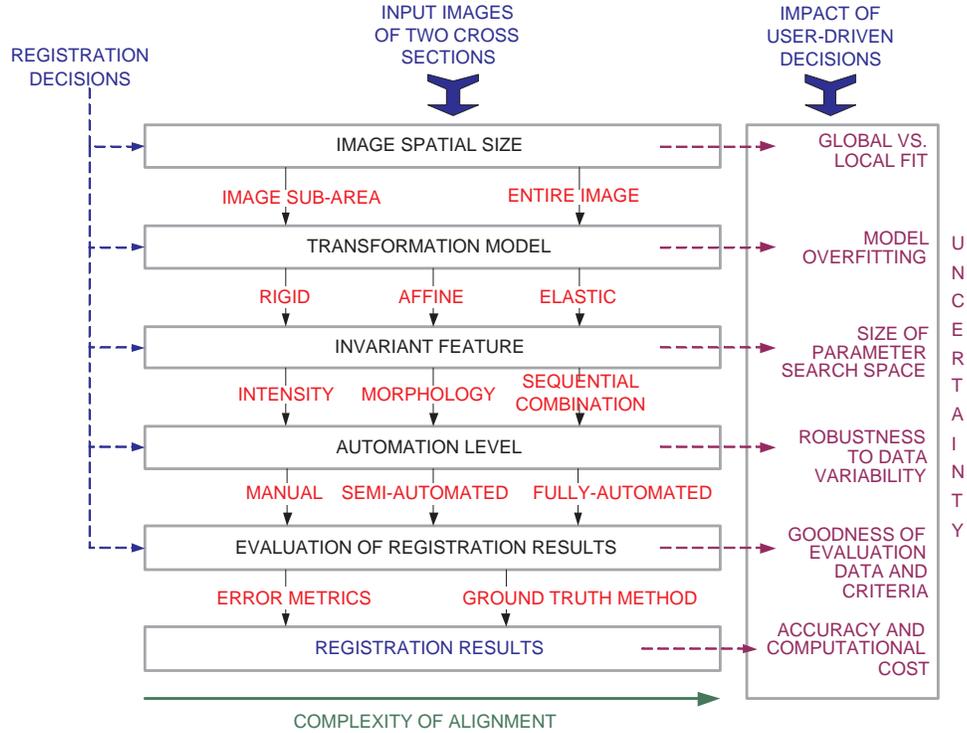


Figure 2.2: Registration decisions and their impact on registration results.

Figure 2.2 is to present basic registration decisions rather than an exhaustive list of possible decision selections.

The user-driven registration decisions define the complexity of (a) registration model, (b) model parameter estimation, (c) registration computations to be performed, and (d) evaluation strategy. For example, the case of a manual registration (alignment) of two image sub-areas containing a few features (visually salient pixel arrangements) using rigid transformation (rotation and translation) by overlaying two sub-areas and visually assessing the quality of alignment would be considered as a low complexity registration. It would use the simplest transformation model (rigid), subjective parameter estimation (visual), no computation (manual), and visual method for evaluating registration quality. In contrast, the case of a fully-automated alignment of two large images containing several millions of features using affine transformation (rotation, translation, scale and shear) by exhaustively evaluating the range of affine transformation parameters based on invariance of intensity (e.g., using

normalized cross correlation or normalized mutual information) would be considered as a high complexity registration. In this case, the registration uses consistent parameter estimation by evaluating invariance of intensity, consuming significant computational resources and performing registration quality evaluations using mathematically defined metrics and based on a set of assumptions about data.

2.4.1 Impacts of registration decisions

Registration decisions during medical cross section alignment have a great impact on (1) anticipated accuracy of aligned images, (2) uncertainty of obtained results, (3) repeatability of alignment, and (4) computational requirements. As illustrated in Figure 2.2, the registration decisions affect (a) spatial distribution of registration error (global vs. local registration), (b) registration error composition (model overfitting, mosaicking x-y error, and alignment z error), (c) computational requirements on registration (invariance assumption and its degree of freedom, model complexity, search space to optimize parameters), and (d) validity of the obtained results (robustness of a registration model with respect to data deviations, quality of evaluation data and criteria).

Alignment Accuracy: The accuracy of aligned images can be measured either by visually inspecting anticipated structures or by defining quantitative metrics to evaluate accuracy with respect to ground truth data (or a data model defined a priori). The analyses consist of simulations with rigid and affine registration transformation and development of registration error models for better understanding of alignment results.

Uncertainty of Alignment Results: As illustrated in Figure 2.2, multiple registration decisions impact alignment uncertainty and include the decisions about (1) the tradeoffs

between global or local registration fit (image spatial size decision), (2) the issues of transformation model overfitting (transformation model decision), (3) the degree of assumed intensity and morphological invariance (feature invariance decision), (4) the size of parameter search space and the algorithmic robustness to model deviations (automation level decision), and (5) the goodness of evaluation criteria for a registration problem. These uncertainties are very difficult to evaluate analytically and are very much data specific.

In order to understand registration uncertainties, data-driven evaluation software tools might provide insight about the anticipated result quality. Since some level of uncertainty is introduced to alignment results from every registration decision, I developed data-driven analysis which calculates (1) normalized correlation based error as a function of image sub-area size, (2) error residuals as a function of rigid or affine transformation models (model complexity), (3) the degree of intensity and morphological invariance, (4) the total search space size for semi-automated and fully-automated registration, and (5) error residuals as a function of evaluation metrics (normalized cross correlation or normalized mutual information).

Repeatability of Alignment: Repeatability can be viewed as the consistency of alignment results obtained using multiple methods and processes [132]. The alignment processes usually include humans and computer algorithms. In general, alignment repeatability varies depending on the level of automation, registration methods, and the complexity of human input. The higher automation level leads to better convergence of registration methods to global extreme [18]. In the semi-automated case, the less complex human input will lead to higher alignment repeatability.

The repeatability issue is usually addressed by performing studies using human subjects [85]. The studies are based on either a class of images acquired by the same imaging techniques and from similar specimens or a class of synthetic images that simulate different degrees of deviations from a registration model. Algorithmic repeatability is often evaluated

together with its robustness to test that the algorithm avoids getting trapped in local minima, and can reliably find the best global minimum in complex landscapes defined by objective functions. Similarly, any measured repeatability due to automation is based on (a) making assumptions about acquired data, for instance, assuming feature invariance, (b) allowing only a subset of possible registration transformations (model constraints), or (c) searching only a subspace of possible transformation parameters. Thus, usually the robustness and accuracy of cross section alignment is decreasing with an increasing level of automation for data sets deviating from the automation model and violating the registration assumptions.

In order to understand repeatability issues, we developed web accessible registration tools that allow researchers to upload their data and collect measurements about repeatability as a function of (a) the complexity of human input (pixel or segment selection), (b) the level of automation (manual or semi-automated registration), (c) human expertise (experts and novices), and (d) data sets deviating from automation models.

Computational requirements: The computational requirements of alignment are directly proportional to the level of automation, the complexity of transformation model, and to the search space of transformation parameters. Furthermore, computational requirements for accommodating all researchers interested in using the same software with multiple registration tasks also have to be addressed.

In the prototype software, we utilize high performance computing resources at the National Center for Supercomputing Applications (NCSA) to guarantee sufficient computing power. The registration decision support software is also designed using web services technology that enables connecting multiple researchers at several geographical locations with computational resources at NCSA [76].

2.4.2 Optimization tools for registration decisions

Given a set of registration decisions and their impacts on accuracy, uncertainty, repeatability and computational requirements, there is a need for a set of methods and software tools to support optimal registration decisions. This work aims at providing such a solution that can be described as a web-enabled, web services-based and data-driven registration decision support system. The system includes methods and software tools that provide data specific understanding of registration decision tradeoffs, such as (1) image sub-area size and location selections, (2) transformation model selection as a function of registration accuracy and possible image distortions (e.g., rigid or affine model), (3) registration feature selection (invariance of intensity or morphological features), and (4) choice of automation level (manual, semi-automated, or fully-automated). All of the above tradeoffs are evaluated as a function of registration accuracy. In addition, the system provides guidelines for an assessment of needed resources (geographically local or distributed computational resources and human expertise).

Image sub-area size and location selection: Specifically, the first method supporting registration decisions enables uploading two images from a 3D volume, selecting a sub-area of interest, and then it reports a visualization of a correlation coefficient as a function of incrementally increasing sub-area size, as shown in Figure 2.3. The correlation coefficient is used as registration accuracy metric in this case (see Figure 2.4). If two identical frames would be compared then the correlation metric value equals to one. The frames that would be completely dissimilar would lead to the value of zero.

Transformation model selection: The second method allows uploading two images to be registered, manually placing disks of variable size over visually matching features, and selecting rigid or affine transformation to compute the registered image pair. An example

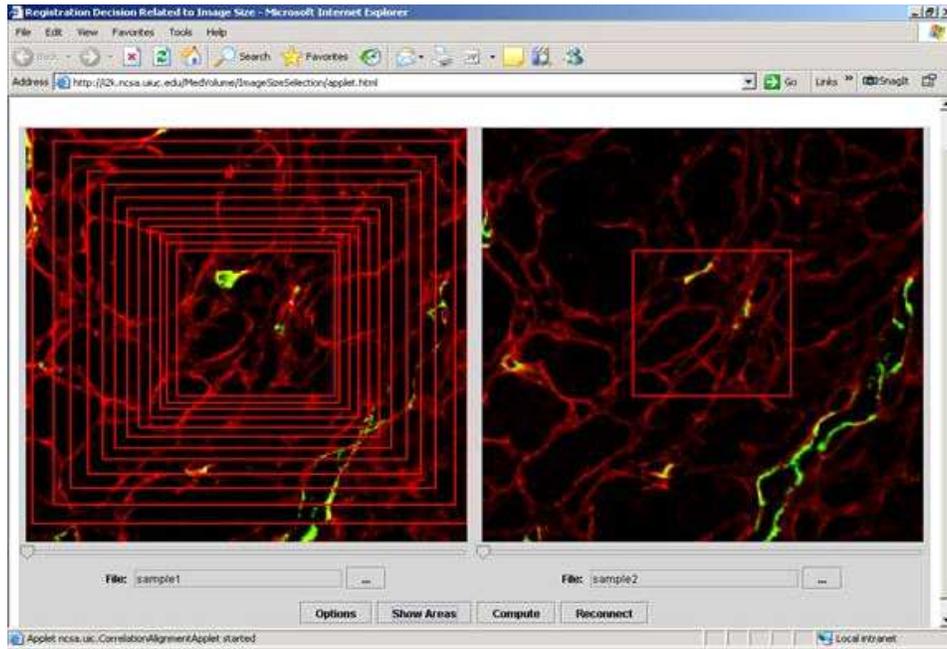


Figure 2.3: Supporting registration decisions about image sub-area size. A similarity of a pair of sub-areas from the left and right images is evaluated by computing a normalized correlation value as a function of a sub-area location and its size.

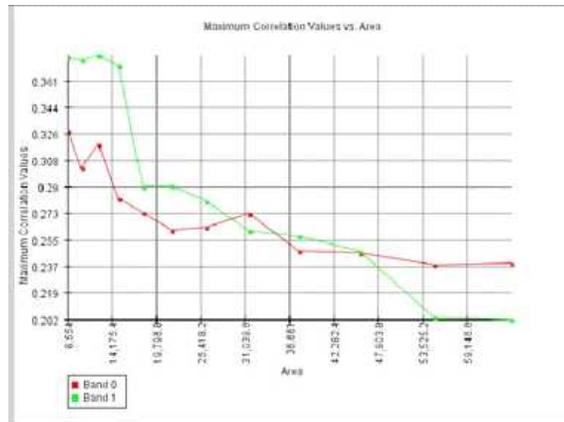


Figure 2.4: The result of processing multiple image sub-areas selected according to Figure 2.3. The resulting graph shows normalized correlation values of a pair of two band images as a function of a sub-area size.

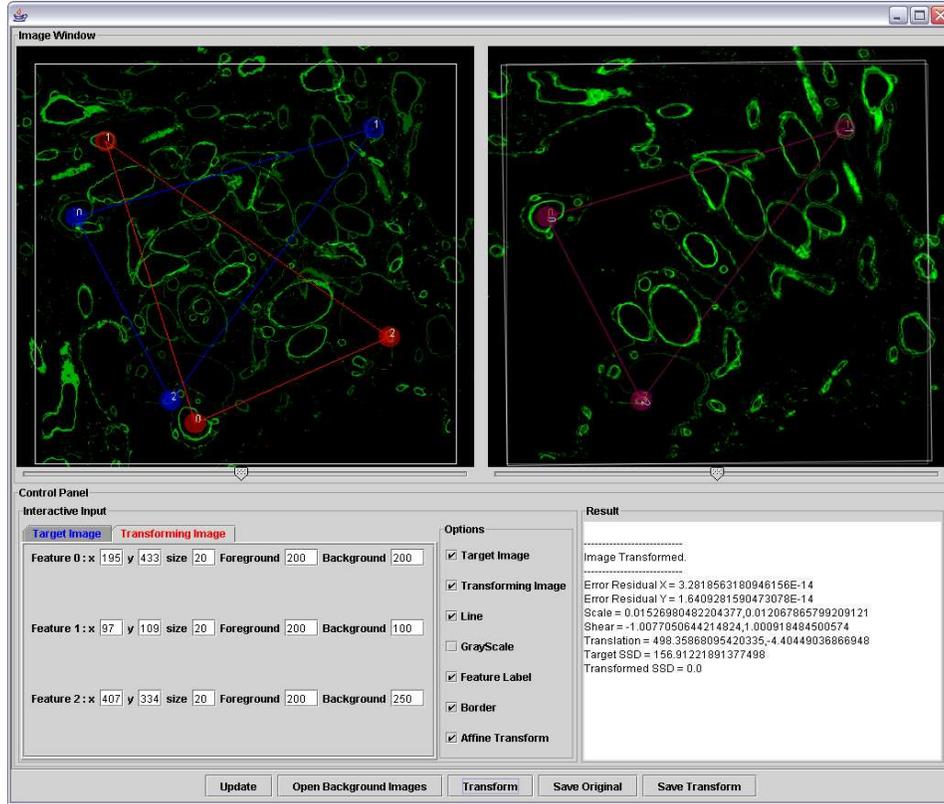


Figure 2.5: Supporting registration transformation model selection. A choice of rigid or affine transformation model is evaluated by reporting a registration transformation error and by inspecting the level of distortion in the overlaid registered images. The red and blue disks are used for establishing matching pairs in a pair of images. The left panel shows a pair of original images (overlaid and inspected by moving the transparency slider bar below the panel). The right panel illustrates the registered sets of red and blue disks and the overlaid images after transformation.

is shown in Figure 2.5. The reported registration transformation error and the level of distortion in the overlaid registered images with registration disks are used for aiding the choice of registration model.

Registration feature selection: The third method supporting registration decisions provides degree of intensity variation for intensity-based approach. If morphological change in a sub-volume is negligible, the spatial intensity-variation along depth (z-axis) can be measured by computing intensity similarity (e.g., normalized cross correlation coefficient) between an

image pair. To deal with the presence of morphological changes, I used low-pass filtered images for intensity similarity comparison. Since the low frequency component of an image suppresses sharp morphology changes and noise such as internal structures (e.g., lines, contours), the low-pass filtered image can be used as an estimate of image with intensity variation only within a sub-volume. This method supports decision of usability of intensity-based method based on a user-defined threshold of intensity variation. The decision rule can be defined as follows: if $degree(\text{intensity variation}) > \delta_{\text{threshold}}$ then use feature-based method (manual, semi-automated, or fully automated), else use intensity-based method (normalized-cross correlation or normalized mutual information), where $\delta_{\text{threshold}}$ are user-defined threshold.

Automation levels: The fourth method provides an interface to pixel-based (manual) and segment centroid-based (semi-automated) registration methods. A user can choose any two images from a 3D stack of pre-segmented images, perform registration using manual or semi-automated registration and compare the registration accuracy. If pre-segmented results are not available then segmentation and segment centroid extraction will be conducted.

2.4.3 Results

This section describes how the set of developed methods was used for making optimal registration decisions during 3D volume reconstruction of the specimens imaged by CLSM. Other researchers processing CLSM images may follow similar data-driven procedures to choose optimal 3D volume reconstruction parameters, such as registration area size, rigid or affine transformation model, intensity or morphology as invariant feature across cross sections, and pixel-based (manual) or centroid-based (semi-automated) alignment method.

Data-driven analysis

Image Spatial Size: From a medical user view point, one would like to obtain 3D volume reconstructions over a large spatial area at a high spatial resolution, and with the best possible visual alignment of all salient image features. It is assumed that the visual alignment of salient image features is measured by a normalized correlation coefficient. Then, the goal is to provide data-driven analysis of input data to understand the tradeoffs between these two conflicting requirements, such as to maximize aligned image area and its measure of alignment goodness (e.g., normalized correlation coefficient) with other images.

In order to compute a data-driven dependency between normalized correlation metric and image sub-area size, at least two frames have to be aligned. Otherwise, other registration unknowns would be inseparable from the variables under our scrutiny. In the case of CLSM, we can establish the data driven dependency by using any two frames from one stack of images. The assumptions are that these two selected frames are representatives of the two frames from two adjacent cross sections without any distortion due to specimen slicing, and the frame-to-frame morphology and intensity changes in the selected frames from one sub-volume are similar to those in the adjacent sub-volumes.

For the input data set with four sub-volumes, I used varying sub-area sizes of images shown in Figure 2.6 from 24648 pixels (79×78) to 1236444 pixels (561×551) by increments of 2% of the entire number of pixels (approximately 25000 pixels). Figure 2.8 illustrates the dependency of the level of similarity measured by normalized correlation as a function of pixel count (sub-area size in pixels). The normalized correlation coefficients represent the maximum value obtained by optimizing spatial overlays using the affine transformation for every pair of image sub-areas. We can observe that the maximum similarity is obtained for the smallest experimental sub-area, and the minimum similarity is reported for the full image. Thus, alignment results for a sub-area that is approximately 37 times smaller than the full image size ($915,788/24,648$) will lead to 13% absolute improvement ($0.85 - 0.72$),

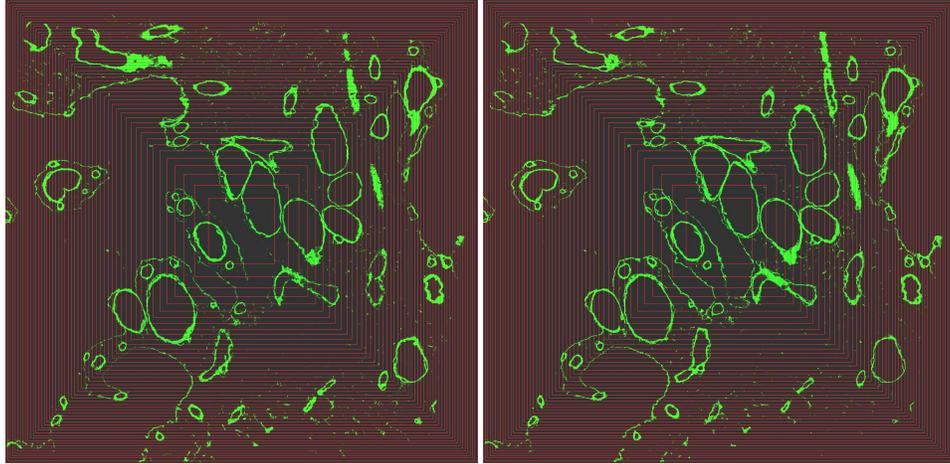


Figure 2.6: Two images used for demonstrating the impact of image sub-area versus entire image registration tradeoffs (image spatial size decision). The images came from one stack of CLSM images and hence the ground truth alignment was known. The slides are 13 frames apart (1.7 microns distance along z-axis) , and each pixel represent 0.478 microns (536 microns by 527 microns for the entire image).

and 46% relative improvement with respect to the best (1.0) and worst (0.72) alignment of the two frames $(0.85 - 0.72)/(1 - 0.72)$. If one would like to explore the relationship without intensity variations then thresholded images can be evaluated the same way as. The thresholded images are shown in Figure 2.7 and the dependency is shown in Figure 2.8.

Transformation Model: Let us assume that a class of CLSM images acquired from similar tissues can be perfectly registered by an unknown transformation with N parameters. The N parameters define the order of transformation model. It is well known that a higher order transformation model than N would lead to data overfitting while a lower order transformation model than N would lead to large registration errors. In our case, overfitting would lead to distortions in the transformed image that could never happen in the original cross sections although the registration error would be small and indicate a good alignment. Using lower order transformation model might likely never satisfy the transformation error defined as user requirements.

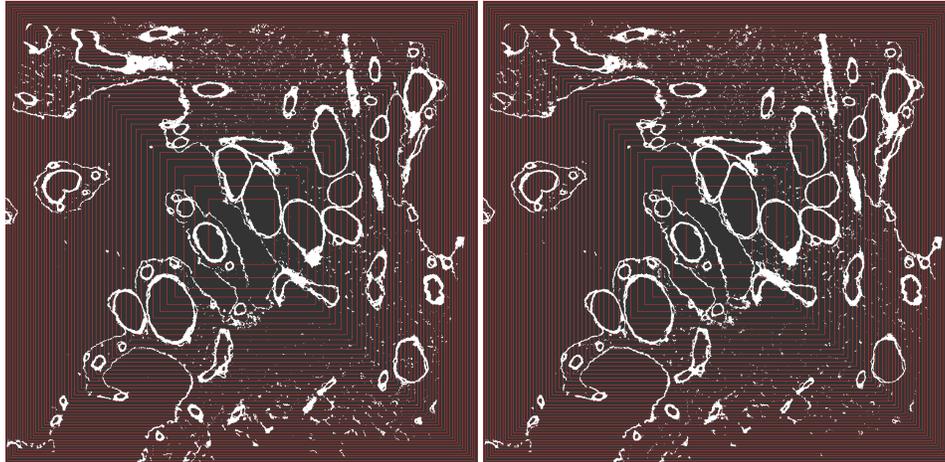


Figure 2.7: Two binary images derived from the images in Figure 10 by intensity thresholding to demonstrate the impact of image sub-area versus entire image registration tradeoffs (image spatial size decision) for the case of zero intensity variation.

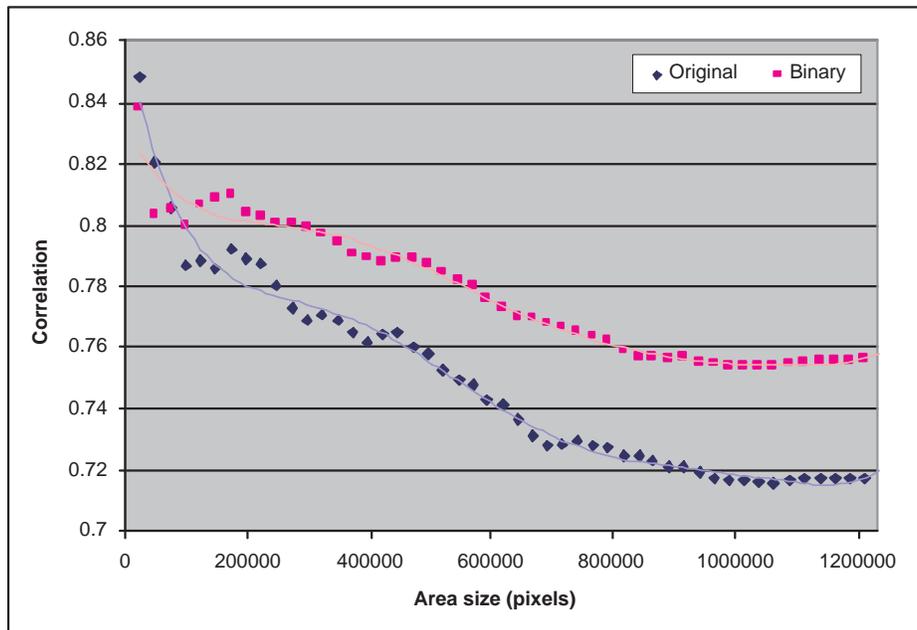


Figure 2.8: Normalized correlation coefficient as a function of image sub-area size in pixels for two perfectly aligned images that are 13 frames apart (1.7 microns apart), each image is of pixel size 1123 by 1104 corresponding to spatial resolution 536 microns by 527 microns (entire image). The two curves correspond to original image (blue) and thresholded image (pink). The images are shown in Figure 2.6 and Figure 2.7. The normalized correlation coefficient represents a quantitative metric (similar to visual inspection) that ranges from $[0, 1]$. For our data set, we decided to use the entire image area equal to 561x551 pixels.

Table 2.1: Summary of registration errors obtained using rigid and affine transformation models for the same set of four matching pairs of points. The residual errors X, Y and (X, Y) are computed as a sum of squared differences between the transformed coordinates of image 1 and the user chosen coordinates in image 2. The original residual error (X, Y) before any transformation was equal to 12.782.

| Transformation Model | Residual X | Residual Y | Residual (X,Y) |
|----------------------|------------|------------|----------------|
| Rigid | 1.528 | 1.802 | 1.225 |
| Affine | 0.927 | 1.581 | 0.935 |

The tradeoffs between transformation model complexity and registration error are usually resolved by practitioners based on a good understanding of medical specimen preparation. For example, for imaging cross sections of solid and hard specimens, rigid transformation might be appropriate (rotation and translation). If cutting the specimen introduces additional shear and scale changes, then affine transformation would be appropriate (rotation, translation, scale and shear). Our goal is to provide data-driven analysis of input data to understand the tradeoffs between these two conflicting requirements, such as to minimize transformation model complexity and registration error.

The task can be achieved by uploading two frames from two adjacent sub-volume images, manually placing disks of variable size over visually matching features, and selecting rigid or affine transformation to compute the registered image pair. The registration error for these two transformation models is summarized in Table 2.1. It provides a quantitative comparison of complexity versus error tradeoffs. For our data set, I used affine transformation model.

Invariant Registration Feature: To quantify the degree of intensity variation, I followed the steps: (1) perform low-pass filtering to all the frames in a sub-volume with different filter sizes, (2) compute a set of normalized cross correlation coefficients by comparing the first frame (e.g., frame 1) and other frames (e.g., frame 2, frame 3, etc), (3) decide a low-pass filter size which mostly suppresses morphological deformation (i.e., images that estimate

intensity variation only), (4) find the maximum intensity variation (normalized cross correlation coefficient) using low-pass filtered images with intensity variation only (computed filter size in step 3), (5) decide whether the intensity-based method can be used based on the user defined threshold $\delta_{threshold}$.

Figure 2.9 shows the intensity similarity (normalized cross correlation coefficient) using different low-pass filter size for multiple pairs of frames. The x-axis represents a frame number of an image which is compared with the frame 1, and each curve represents a normalized cross correlation coefficient with different filter sizes (e.g., 1 100). One can observe that the similarity decreases along large gaps (i.e., frame depth) of the selected frames. Therefore, we could select the frame pair that represents the largest morphological deformation (frame 1 and frame 24) to decide the low-pass filter size. Figure 2.10 shows the similarity of the image pair, frame 1 and frame 24, as a function of the low-pass filter sizes. One can notice that the similarity increases relatively rapidly up to the filter size equal to 40, and then stabilizes after. It implies that the morphology hardly contributes to the similarity relatively. As a result, we chose the low-pass filter size as 40 pixels neighborhood to compute the intensity variation. Finally, the lowest similarity due to intensity variation can be found by examining the curve with the filter size of 40 pixels in Figure 2.9, which is 0.702. Note that this value is relatively low since the result is based on the assumption that there is only intensity variation (e.g., no morphological change).

In Figure 2.9, one could argue that the morphological deformation is not removed well based on the curve which keeps falling, but in our data set, it is observed this is caused not by morphology but by intensity variation. Figure 2.11 supports the argument by showing the similarity of adjacent image frames. Assuming the rate of morphological deformation is constant along depth, Figure 2.11 shows low similarity at frame 15 through 23. It is interpreted that the intensity variation is more severe in this range. One could also visually verify the fact from Figure 2.12.

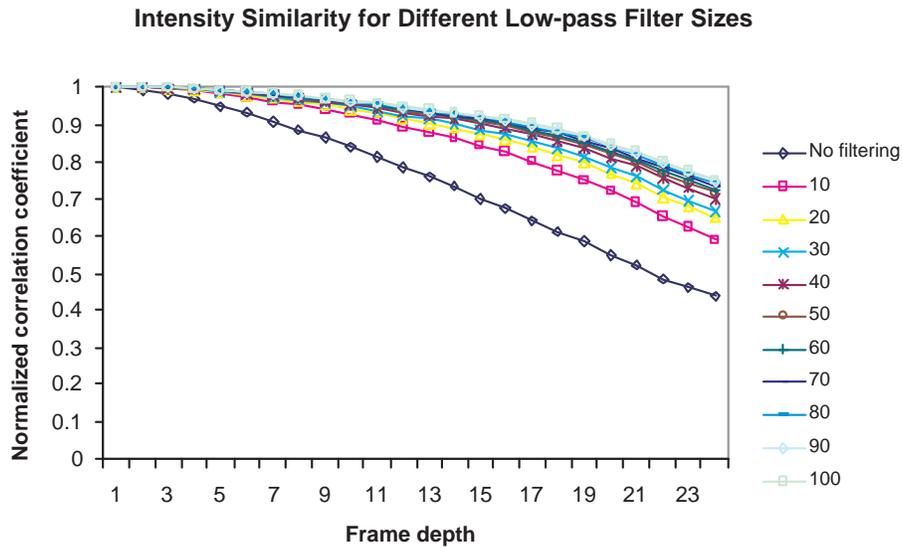


Figure 2.9: Intensity similarity (normalized cross correlation coefficient) using different low-pass filter size for pairs of frames. The x-axis represents the frame number which is compared with the frame 1, and each curve represents a normalized cross correlation coefficient with different filter size.

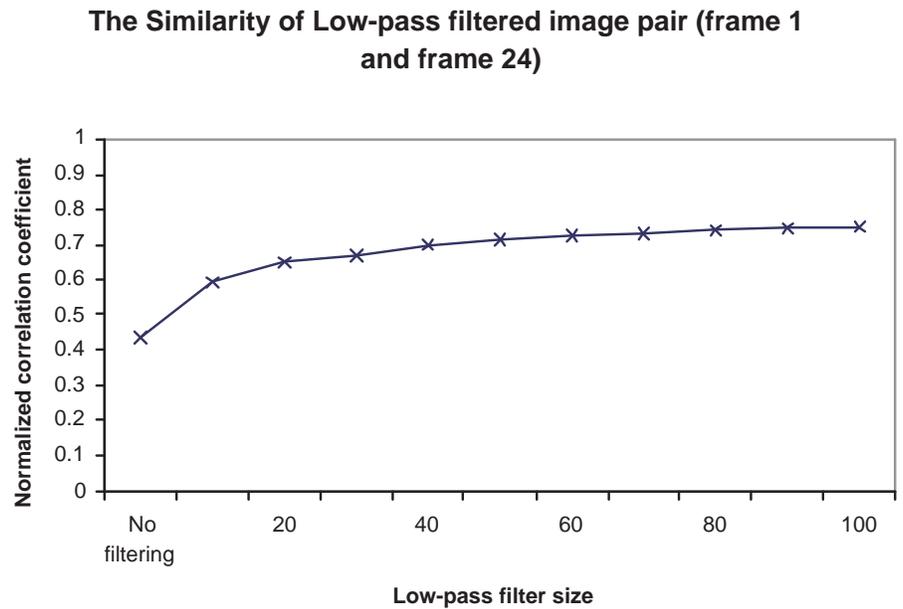


Figure 2.10: The similarity of image pair, frame 1 and frame 24, as a function of low-pass filter size.

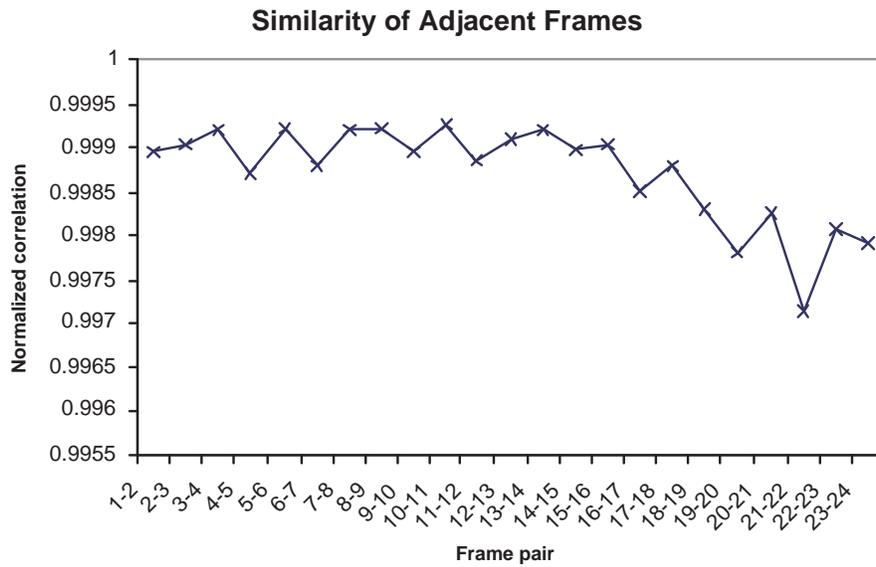


Figure 2.11: The similarity of adjacent image frames.

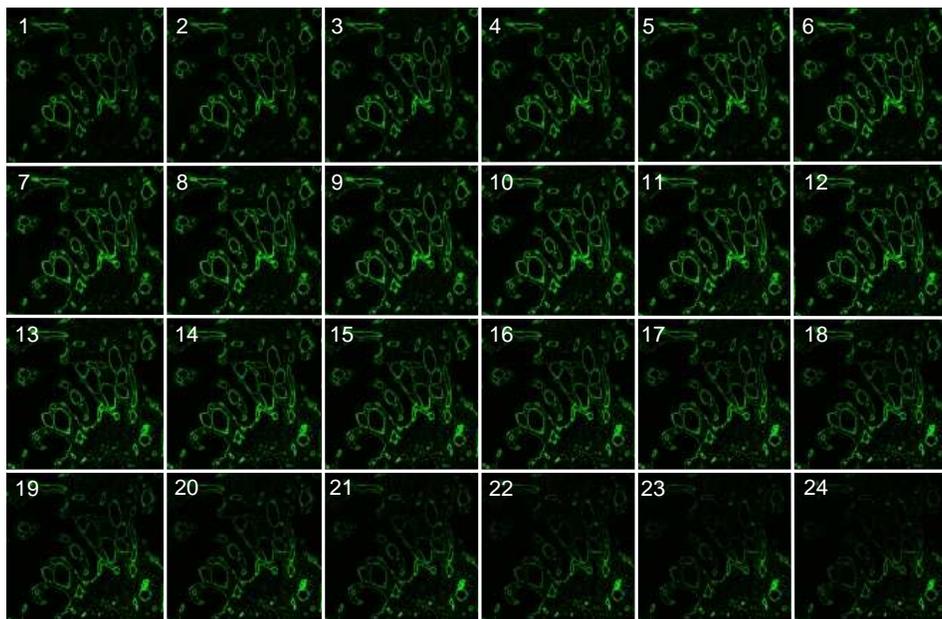


Figure 2.12: Image frames from 1 to 24

Automation Level: The goal of this step is to decide the level of automation based on the registration accuracy, computational resources, and geographic locations of needed 3D volume reconstruction expertise. We approached this problem as follows: The evaluations of registration accuracy at multiple levels of automation require involvement of human subjects and careful preparation of ground truth data. We have conducted a user study in the past using manual or semi-automated registration techniques [85] and decided to use the semi-automated (segment centroid based) automation level. By web-enabling the developed software, other researchers can perform similar studies to the study published in [76] and support their data-specific decision about the level of automation.

The problem of available computational resources and geographically distributed expertise was solved by developing a web service-based mechanism for registration that provides access to image data at the location with 3D volume reconstruction expertise and performs computation at the location with computational resources. Other researchers would be able to use the same prototype developed for our 3D volume reconstruction and evaluate their data-driven decision about an appropriate automation level.

Registration decisions

Based on the experimental data-driven evaluations, registration decisions are supported to choose (1) an entire image size for registration, (2) affine transformation model to compensate for observed warping, (3) morphology as the more invariant feature than intensity and (4) semi-automated (or fully automated) centroid-based technique over pixel-based registration method. These registration decisions are optimal given the choice of accuracy quality metrics and algorithmic approaches. We provide next the discussion about (a) the need to have 3D volume reconstruction quality specifications, (b) the pros and cons of analytical and data-driven modeling approaches to registration decisions, (c) the challenges associated with computing and comparing the degrees of intensity variation and morphological distortions

and (d) the difficulties with evaluations that support registration decisions.

3D volume reconstruction quality specifications: The outlined procedure for optimizing registration decisions brings up several issues related to well-defined 3D volume reconstruction requirements. In a real world scenario, input data are not ideal, and 3D volume reconstruction accuracy requirements are limited by the nature of specimen preparation, data acquisition, and processing. It was demonstrated that the requirements for the 3D volume reconstruction presented in this section have to include several user-defined parameters. First, a user has to define the minimum area of the interesting region and the minimum intensity similarity (e.g., normalized cross correlation coefficient) for accepting the alignment results (image sub-area size decision). Second, there is a need to specify the minimum registration error improvement due to increased transformation model complexity for adopting more complex transformation model (registration model decision). Third, one would specify the instrument noise threshold value to separate intensity variations from morphological deviations, as well as the 3D model for segmented structures and the maximum estimated deviation from the model for accepting registration feature invariance (invariant feature decision). As for the automation level decision, a user has to decide what accuracy and repeatability of reconstructions would be acceptable as a function of automation.

Analytical and data-driven modeling approaches: To consider the impacts of registration decisions on (1) anticipated accuracy of aligned images, (2) uncertainty of obtained results, (3) repeatability of alignment, and (4) computational requirements, we have offered data-driven models that facilitate the understanding of real data acquired from instrumentation. The reason for choosing data-driven models instead of analytical models lies in the large variability of objects of interest, specimen preparation, imaging modality, specific in-

strumentation characteristics, and so on. The use of analytical models might not be feasible (lack of knowledge) or introduce too many assumptions about data so that the generality of the models is very small. As illustrated in the previous sections, the data-driven approach to support 3D volume reconstructions is more feasible and general. For example, it is addressing several basic registration decision problems related to (a) what variables play major roles during 3D volume reconstruction, (b) how relevant 3D reconstruction variables (parameters) depend on each other and how they impact resulting 3D reconstruction, (c) how to optimize the choice of 3D reconstruction parameters, and (d) what solutions could accommodate computational and expertise requirements during 3D volume reconstruction.

Evaluations supporting registration decisions: Ideally, one would like to perform data-driven evaluations with all possible registration accuracy metrics, all registration methods and approaches, and with unlimited computational resources. Each researcher would assess his/her data variability and quality with respect to the 3D reconstruction task under multiple modeling assumptions and without computational constraints. While this is currently unobtainable, the work presents a methodology how to obtain better 3D volume reconstruction understanding given a limited subset of metrics, methods, and resources. In addition, the evaluation problems related to human subject selection, ground truth data preparation, and 3D volume validation (just to name a few), are beyond the scope of this dissertation but might be of interest to a user evaluating automation levels.

2.5 Summary

I presented a framework of 3D volume reconstruction using blood vessels and vasculogenic mimicry patterns in histological sections by CLSM. To my knowledge, there have not been

developed robust and accurate fully automated 3D volume reconstruction methods from a stack of CLSM images without artificially inserted fiducial markers. The proposed framework is also unique in terms of forming a complete system for medical inspection of 3D volumes reconstructed from CLSM images. The approach to automate the reconstruction process incorporates the tradeoffs between computational requirements and uncertainty of resulting reconstructions introduced by each processing step.

In addition, I addressed the problem of optimal registration decisions during 3D medical volume reconstruction. The registration decisions of interest included (1) image spatial size (image sub-area size), (2) transformation model (rigid or affine), (3) invariant registration feature (intensity or morphology), and (4) automation level (manual or semi-automated). I provided mechanisms for evaluating the tradeoffs of each registration decision in terms of 3D volume reconstruction accuracy, repeatability and computational requirements. Furthermore, I presented software tools for geographically distributed researchers to optimize their data-driven registration decisions by using web services and supercomputing resources. Finally, I presented an example of optimal registration decisions supported by the developed software and analyses. The example showed the case of 3D volume reconstruction of blood vessels in histological sections of uveal melanoma from serial fluorescent labeled paraffin sections imaged by fluorescence CLSM imaging.

Chapter 3

SUB-VOLUME PRE-PROCESSING

3.1 Introduction

In this chapter, I describe sub-volume pre-processing steps such as (a) sub-volume intensity analysis, (b) tile mosaicking, (c) segmentation, and (e) feature detection as shown in the 3D volume reconstruction process in Figure 2.1. The input to the pre-processing step is a set of images measured by CLSM and labeled with a slide number (physical section), a spatial tile index, a frame number (optical section), and a measured channel number.

In section 3.3, I describe a frame selection method to select the most salient frame within each image stack based on contrast and entropy analysis of every image frame. Next, in the mosaicking, based on our data assumptions (see Section 1.4), I chose laterally adjacent images from the same frame depth. The selected set of frames is mosaicked by manual (pixel-based), semi-automated (feature-based), and fully automated (intensity-based) method described in section 3.4. I also show the result of a user study to demonstrate the mosaicking accuracy using different automation levels.

In section 3.5, I show 2D/3D segmentation methods applied to the mosaicked image stacks (also denoted as sub-volume). In section 3.6, I describe the registration feature extraction from the segmented regions. The goal of the 2D segmentation is to extract a set of features for matching described in chapter 4. 3D segmentation provides high dimensional features (e.g., centroid trajectory) for 3D matching refinement outlined in section 4.4.

The outputs of the pre-processing step are a set of features including centroid coordinates associated with region areas and 3D centroid trajectories of 3D segments (volumes).

3.2 Related Work

There are several known factors that cause spatial intensity heterogeneity, such as photo-bleaching, fluorescent attenuation along confocal (depth) axis, image acquisition factors [11, 124, 42], variations of illumination exposure rate, spatially uneven distribution of dye and the spatial characteristics of illumination beams [107], and fluorochrome micro-environment, e.g., pH, temperature, embedding medium, etc [138]. In particular, z-axis intensity heterogeneity can have a serious effect on CLSM images. It continues to bleach entire illuminated three-dimensional volume equally, all of the time [135, 28].

Image restoration approaches toward z-axis intensity correction have been suggested based on empirical correction methods for intensity loss [127], constant thresholding [69], iterative correction methods [146, 126], 2D histogram [93], or estimations of intensity decay function [75]. Beam attenuation can be corrected mathematically [124], but the bleaching problem is non-linear and far more difficult to correct mathematically. Since the intensity heterogeneity in our data set is caused by multiple factors described above, image restoration techniques cannot be simply used since the method could amplify noise or distort structural features, e.g., intensity enhancement of outliers.

Based on the assumption of intensity heterogeneity along z-axis and data acquisition, for registration purpose, I select an image frame with the most salient structural information based on a combination of a contrast measure and an entropy measure. After the frame selection, the image tiles are mosaicked to generate a high resolution image frame (i.e., wide field of view). In various scientific fields, mosaicking has been extensively used such as in geo-referencing, mapping/localization for navigation, remote sensing, underwater imaging, astronomy, medical imaging, etc [82, 29, 44, 83, 144, 105, 57, 123, 17, 34]. Generally, image mosaicking is performed in the following order: First, establish correspondence between ad-

jacent images. Second, estimate optimal transformation parameters with a known model. Third, blend overlapping region of the images. Without considering blending operation, the mosaicking can be considered as a sub-problem of alignment problem. The correspondence problem can be approached by intensity-based or feature-based method (as described in next section), and the parameter estimation can be performed using the RANSAC algorithm [36] that evaluates multiple corresponding pixels randomly. However, in the case of CLSM data, this problem can be significantly simplified due to the data acquisition assumptions such as (1) tiles can be merged by using a translational transformation (two parameters), (2) morphological deformation at the same depth (the same frame number) is minimal, and (3) spatial intensity heterogeneity is minimal in the selected frames after the frame selection step. Therefore, in this work, the mosaicking problem is reduced to matching problem (intensity or feature) in 2D images. One of the main requirements in our medical image analysis is that the image morphology (structure) must not be distorted during image processing. Thus, I consider neither local image warping for smooth transition nor blending of imaged structures.

Segmentation of the mosaicked image stacks (also denoted as sub-volumes) follows the mosaicking step. Segmentation is usually referred to as a partitioning process of an image/volume of interest into a set of regions with homogeneous intensity, texture, or structural characteristics [58, 108]. The segmented 2D regions can be used for quantitative analysis [84], structural diagnosis [156], pathological localization [161], or feature extraction for image registration [85, 88].

Generally, segmentation methods are divided into eight categories based on their approaches; thresholding, region growing, classification, clustering, Markov random field models, artificial neural network, deformable model, and atlas guided method [161]. Although there is an abundance of literature about segmentation of medical images, it is known that the performance of each segmentation method heavily depends on specific imaging modality and application domains [115]. For this reason, a segmentation approach needs to be selected

based on appropriate assumptions about input data. In the case of automated vascular region detection in CLSM data, model-based approaches (classification, clustering, Markov random field model, artificial neural network, deformable models, and atlas guided method) cannot be used due to no prior knowledge about the structural shape, a high variety of vascular region shapes, and high intensity heterogeneity. In segmentation, I used a method that combines thresholding and region growing methods [9].

3.3 Sub-volume Intensity Analysis

As introduced in the framework in Chapter 2, sub-volume intensity analysis is performed to find the most salient image sets that are to be used for mosaicking and 2D regions-based alignment process.

First, based on the assumption of the material preparation and the image acquisition process, a pair of sub-volumes can be mosaicked or aligned using one representative global transformation since optical sectioning process does not change lateral configurations of frames geometrically. For example, frames are perfectly aligned in a single sub-volume. For mosaicking purpose, I used a set of representative image tiles (frames) in the same axial depth from each spatially adjacent (in lateral plane) image stacks. For example, choose k -th frame from all spatially adjacent image stacks. Therefore, it is preferred to select the most salient frames from the same axial coordinate of sub-volumes, which provides high intensity similarity and segmentation performance.

Second, in case of axial sub-volume alignment (z -axis) using 2D image information, it is preferred to align a pair of sub-volumes based on a pair of depth adjacent frames, e.g., the last frame of the first sub-volume and the first frame of the second sub-volume, because a structural discontinuity would be expected to be minimal. Unfortunately, these end frames of a sub-volume acquired by CLSM are usually characterized with much lower fluorescent

intensities (i.e., signal-to-noise ratio) than other frames inside of a sub-volume due to the physical constraints of fluorescent confocal imaging. For these end frames with low signal-to-noise ratios, the alignment is difficult. Therefore, for sub-volume alignment based on 2D information, I also assume that the typical depth of each sub-volume is small in comparison with the rate of structural deformation, and therefore the frames other than the end frames could be used for image comparison of depth adjacent sub-volumes.

3.3.1 Frame selection

This selection is performed in such a way that the pair of frames would provide high confidence in any found features. In general, high confidence in image features is related to their level of visual saliency (image intensity amplitude, contrast, spatial variation and distribution). From this viewpoint, I select an optimal pair of image frames that provides maximum saliency of image features and is defined as:

$$I_{opt} = \arg \max_{I_k \in V_i} \{ENTROPY(I_k) \times CONTRAST(I_k)\} \quad (3.1)$$

where I_k is an image frame ($I_k \in V_i$), V_i is a sub-volume, I_{opt} is the most salient image frame. Note that *ENTROPY* is the information entropy based score, and *CONTRAST* is the contrast based score as described below.

Information entropy: This saliency measure is based on evaluating each image frame separately using the information entropy measure [129] defined below.

$$ENTROPY(I_k) = - \sum_{i=1}^m p_i(I_k) \ln p_i(I_k) \quad (3.2)$$

where $ENTROPY(\cdot)$ is the entropy measure, $p_i(I_k)$ is the probability density of a fluorescent intensity value i of the image frame I_k , and m is the number of distinct intensity. The probabilities are estimated by computing and normalizing a histogram of intensity values. Generally, if the entropy value is high then the amount of information in the data is large and a frame is suitable for further processing.

Contrast measure: This measure is based on the assumption that a suitable frame for registration would demonstrate large intensity value discrimination (image contrast) from the background. Thus, I evaluate contrasts at all spatial locations and compute a contrast score according to the mathematical formula shown below.

$$CONTRAST(I_k) = \sum_{i=1}^m \|h_i(I_k) - E(h(I_k))\| \times h_i(I_k) \quad (3.3)$$

where $h(\cdot)$ is the histogram (estimated probability density function) of all contrast values computed across one band by using Sobel edge detector [129](Chapter 4) and $E(h(I_k))$ is the sample mean of the histogram $h(I_k)$. The equation includes the contrast magnitude term and the term with the likelihood of contrast occurrence. In general, image frames characterized by a large value of $CONTRAST(\cdot)$ are more suitable for further processing than the frames with a small value of $CONTRAST(\cdot)$.

Section 3.7.1 shows the frame selection results from our CLSM input data.

3.4 Tile Mosaicking

Image mosaicking can be performed by visually inspecting spatially adjacent image frames in x-y plane, selecting one or more pairs of corresponding points in the overlapping image area, and computing translational transformation parameters for stitching image tiles together. This approach is denoted as manual mosaicking and is supported by a software that enables

pixel selection of matching pairs of points and computation of transformation parameters from a set of control points. If a computer pre-computes salient feature candidates and a user interaction specifies correspondences between any two features, then the method is referred as semi-automated mosaicking. If images are stitched together without any human intervention then it is referred to fully automated mosaicking. Based on the underlying registration mechanism, I also denote manual mosaicking as the pixel-based method, semi-automated mosaicking as the feature-based method, and fully-automated mosaicking as the intensity-based method.

First, I present a manual mosaicking method that displays two spatially overlapping image tiles to a user. A user selects a pair of matching pixels, and then image tiles are stitched. In the next step, a user is presented with the already stitched image and a new tile to select matching pixels. Manual mosaicking is performed in this way till all images are stitched together and the final mosaicked image can be viewed for verification purposes. Second, I present a semi-automated method that (1) highlights segmented vascular regions (closed contours) as salient feature candidates and (2) computes a pair of region centroids, as control points for registration, after a user defined two region correspondence. This semi-automatic method is designed based on the observations that (a) an accurate point selection is much harder for a human than an accurate region (segment) selection, (b) a centroid selection of any region is less accurate by a human than by a computer, and (c) registration based on structural shape of a region rather than on intensity-defined point is more robust to noise. Third, I present a fully automated mosaicking method. Full automation can be achieved by either automating feature-based registration process [49, 38, 92], or maximizing pixel intensity correlation using computationally feasible search techniques (translational transformation only) with normalized cross-correlation or mutual information metrics [119, 145].

To compare the mosaic accuracy it would be more natural to achieve full automation by automating feature registration process. However, in CLSM imaging, it is not always feasible to detect good features. For example, there can be lack of detected features in the

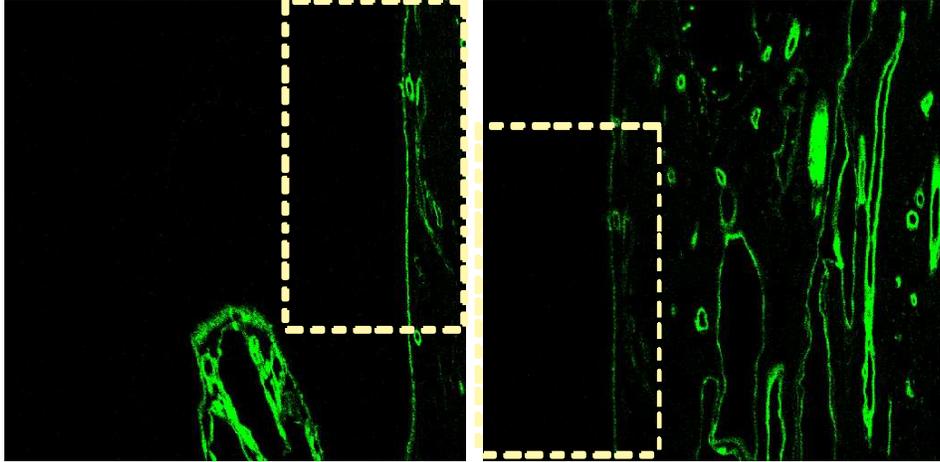


Figure 3.1: Adjacent tiles of CLSM images: overlapping regions have few features (closed regions.)

overlapping region as it is illustrated in Figure 3.1. In this case, the manual method or intensity-based fully automated method needs to be used. As another case, if the images suffer from high intensity heterogeneity (usually not as the case after frame selection), the manual method is the only choice for tile mosaicking. In this work, performing accuracy evaluations by comparing multiple techniques is independent of their underlying principles and I focus only on their resulting mosaicking accuracy. One could also inspect the experimental result to realize that the intensity based method (fully automatic) can lead to better or worse results than the feature based method (manual and semi-automatic). In this work, I demonstrate that the region centroid based registration method significantly improves performance for 3D volume reconstruction of CLSM images in terms of achieved registration accuracy, consistency of the results, and performance time.

For fully automated mosaicking, I used normalized mutual information and normalized cross-correlation metrics to find the best match of spatially adjacent tiles and to provide the sought translational offset for tile stitching. The main mosaicking advantages of these intensity correlation based methods are in (a) their feasible computational cost for translation only, (b) robust performance for image tiles acquired with the same instrumentation setup, and (c) no user interaction (full automation). For example, Figure 3.2 shows how a

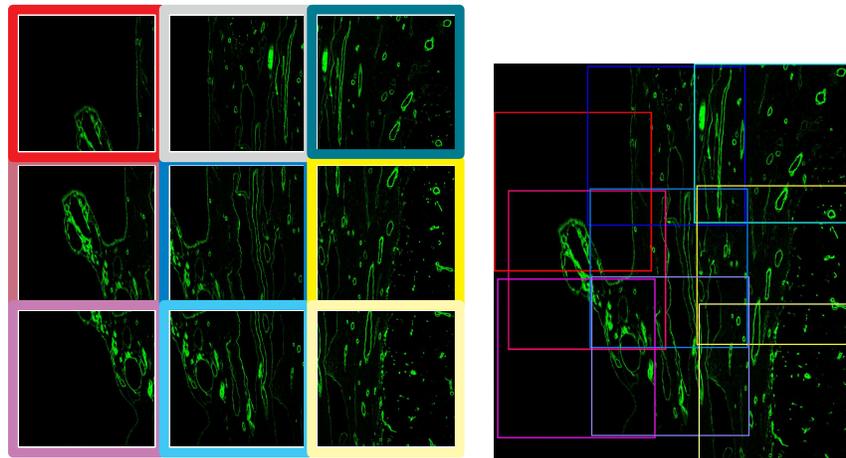


Figure 3.2: Image mosaicking problem: (left) Image tiles with colored borders and (right) mosaicked image showing where each tile belongs in the final image based on its color.

high-resolution mosaicked image is constructed from nine image tiles.

3.5 Segmentation

The goal of this step is to segment out structural features that could be matched in image frames or volumes from depth adjacent sub-volumes. A simple intensity thresholding followed by 2D or 3D connectivity analysis [25] would lead to segments that are enclosed by fluorescent pixels above a threshold value. Nonetheless, it is not always the case that all pixels along a segment circumference are lit up above the threshold that separates background (no fluorescence) and foreground (fluorescent signal) because of specimen preparation imperfections and limitations of fluorescent imaging. Disregarding partially open/closed contours or volumes of lit up pixels could lead to an insufficient number of segments necessary for computing registration transformation parameters.

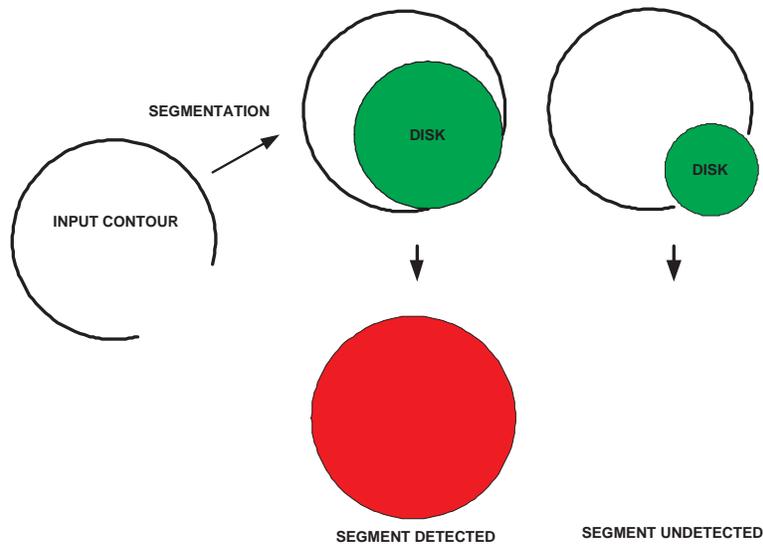


Figure 3.3: Illustration of disk-based segmentation. A segment is formed as a connected set of pixels covered by a disk while the disk moves within the fluorescent boundary. Depending on the disk diameter and contour gaps a segment is detected or not.

3.5.1 Two-dimensional segmentation

I present a 2D region segmentation method that labels segments (regions) with partially closed high intensity contours. The algorithm is based on connectivity analysis of a thresholded image with a disk of a finite diameter, as it is illustrated in Figure 3.3. A disk is placed at every background pixel location that has not yet been labeled. A segment (region) is formed as a connected set of pixels covered by a disk while the disk moves within the fluorescent boundary. The diameter of a disk determines what contours would lead to a detected segment that represents closed contours with a few permissible gaps.

Selection of a disk diameter might affect interpretations of a contour depending on a gap size. For example, Figure 3.4 shows multiple contour interpretations as a function of disk diameter resulting to a line or one closed contour or two touching contours. These interpretations are consistent with methods used by human investigators to select regions of interest from a partially closed contour.

In order to automate the segmentation process, the threshold value and disk diameter parameters must be chosen. The threshold value is usually based on the signal-to-noise ratio

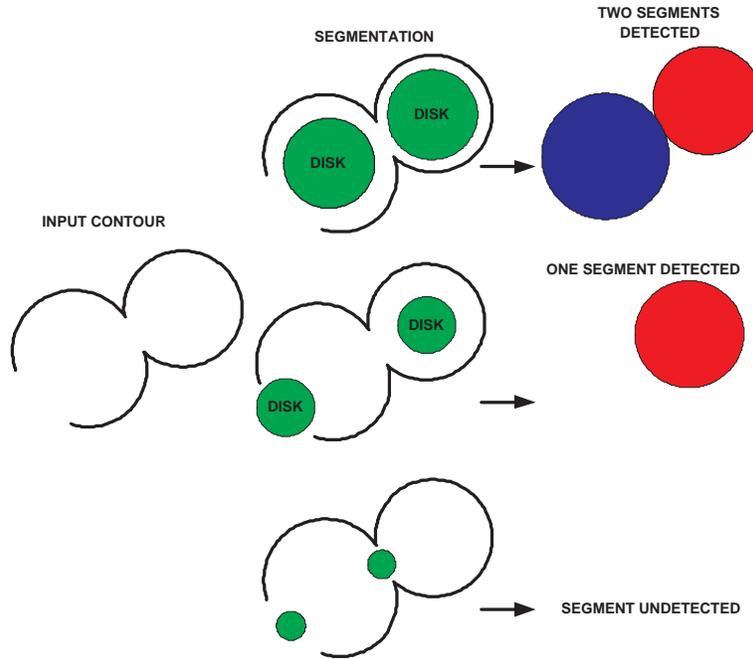


Figure 3.4: Illustration of multiple contour interpretations. A partially closed contour could lead to three detection outcomes depending on a disk diameter.

(SNR) of a specific instrument. It is possible to optimize the threshold value by analyzing the histogram of labeled region areas as a function of the threshold value because a large number of small areas (occurring due to speckle noise) disappear at certain range of threshold values. Discontinuity of closed contours caused by the thresholding is later recovered by disk-based region growing. The choice of disk diameter for connectivity analysis is much harder to automate because it ties to the medical meaning of each closed or partially closed contour that would be selected by an expert. I automated the choice of a disk diameter by imposing lower and upper bounds on the number of segmented regions and evaluating multiple segmentation outcomes.

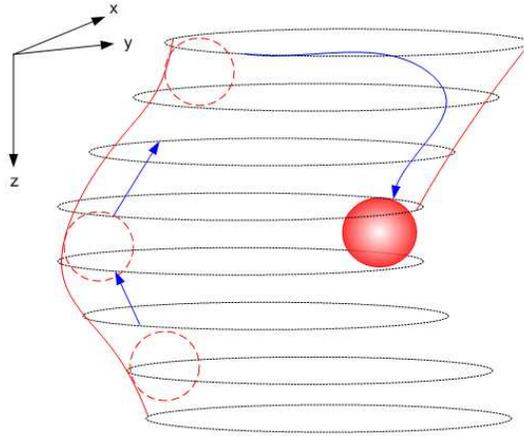


Figure 3.5: 3D region growing for segmentation of a volumetric region using a sphere.

3.5.2 Three-dimensional segmentation

I used a volume growing method that can segment out partially closed volumetric structures that correspond to vascular volume. These partially closed structures (surface) occur in imaged specimens due to photo bleaching or loss (possibly lack) of a fluorescent dye. The motivation for detecting such vascular volume is that they can be used in the absence of artificially inserted fiducial markers for computer-assisted 3D volume reconstruction including mosaicking of image tiles and alignment of multiple cross sections. This volume growing method can recover vascular volume in such cases when (a) a surface region of a color homogeneous closed volume is discontinued due to photo-bleaching, (b) a surface region is partially destroyed during noise thresholding (especially if a high threshold value has to be applied to remove background clutter), or (c) a surface regions is missing since it is partially outside of the imaged area of a microscope. Figure 3.5 shows an illustration of 3D segmentation by growing interior volume of a vascular volume using a virtual sphere.

3.6 Feature Detection

In this step, I identify the features of segments that are most descriptive but insensitive to structural morphology changes. The detected features extracted from images/volumes will be used for computing image registration parameters. Ideally, these descriptive features include parameters describing each segment shape so that homologous segments can be identified in a pair of images.

Registration using 2D information: Based on the choice of the image transformation model for finding segment correspondences that consist of rotation and translation with small shear, I selected segment centroids and areas as the primary shape features. It is known that the segment areas, as well as, the mutual distances between any two centroids of segments, are invariant under rotation and translation. Thus, I utilized the invariance of these two shape features during feature matching and registration parameter estimation. Both of the selected segment features were extracted after performing a connectivity analysis by simple pixel count (areas) and average (centroid) operations.

Registration using 3D information: I computed trajectory of centroids from a series of optical sections (lateral planes) for each 3D segments. Figure 3.6 illustrates an example of a pair of sub-volumes after 3D segmentation with centroid trajectory within each volume.

3.7 Results

In this section, I show (a) frame selection results from two sets of CLSM data, (b) performance evaluation of image mosaicking in three different automation levels by user study, (c)

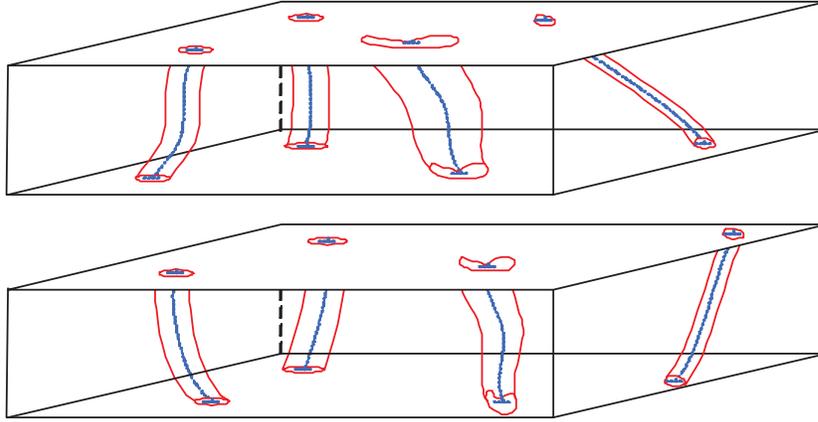


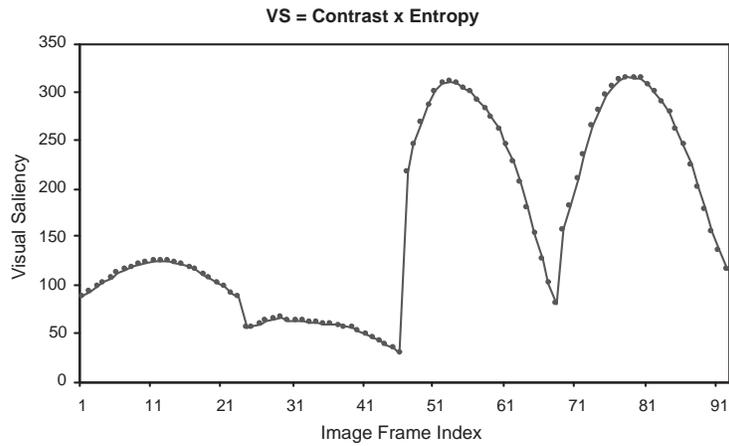
Figure 3.6: An example of a pair of sub-volumes after 3D segmentation.

2D/3D segmentation results, and (d) feature detection results for 2D and 3D cases.

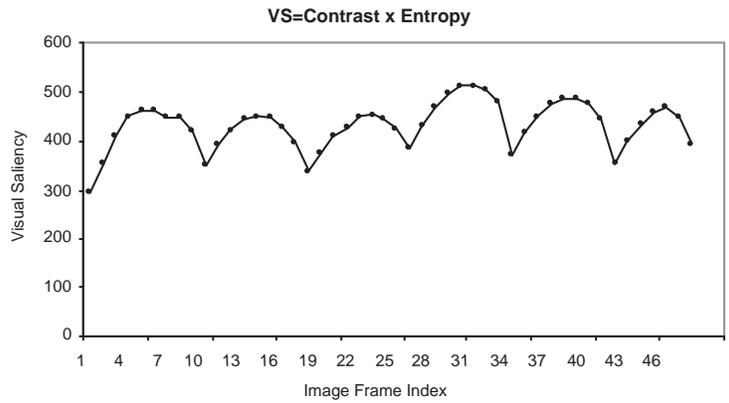
3.7.1 Frame selection

I evaluated two 3D volumes experimentally. One 3D volume was formed from four consecutive sub-volumes consisting of 96 image frames, and another one was formed from six sub-volumes consisting of 48 image frames. The results are shown in Figure 3.7 as a function of image frames index. These graphs demonstrate the combined entropy and contrast variations as a function of image frames from all sub-volumes. The frame with maximum visual saliency score is selected automatically within each sub-volume for alignment. For example, the frame indices 11 and 28 would be used for alignment of the first two depth adjacent sub-volumes because they correspond to the maxima in Figure 3.7 (a).

Based on Figure 3.7, one could always choose the middle frame of each physical section (sub-volume) for image alignment and skip this processing step. From an accuracy viewpoint, this approach would lead to a sub-optimal solution with some computational savings. I discovered that the computational savings are negligible in comparison with all registration related processing (less than a second per frame in comparison with several hours per frame for other processing steps). Experimental evaluations of the sub-optimal accuracy loss led to



(a)



(b)

Figure 3.7: Evaluation of image frame selection from within each 3D sub-volume to be used for alignment of the sub-volumes. The graphs show combined visual saliency score as a function of image frame for two 3D volumes. The frame with maximum visual saliency score within each sub-volume would be used for alignment.

the range of normalized correlation coefficients in $[0.59, 0.635]$ by computing the correlation of the middle frame (sub-optimal solution) with all other 21 image frames inside of one sub-volume (potentially optimal solutions). Additional evaluations of normalized correlations for pairs of middle frames led to the range of values in $[0.21, 0.27]$. Thus, I concluded that based on the experimental data one would expect relatively considerable inaccuracy due to a sub-optimal solution. Statistically, the distribution of the introduced inaccuracy would be skewed toward smaller values and would be data dependent. In my experimental data shown in Figure 3.7 (b) (altogether 10 sub-volumes), 10% of middle frames coincided with the optimal solution and the middle frame was on average 2.5 frames displaced from the optimal frame. In order to avoid these uncertainties, frame selection step should be performed.

When I investigated tradeoffs between choosing the end frames of two spatially adjacent sub-volumes (as opposed to selecting the frames closer to the middle frames within each sub-volume), I discovered that this approach would still demonstrate the tradeoffs between intensity variations and morphological distortions along z-axis (frame index), and hence guide me in an optimal frame selection. Figure 3.8 illustrates data that support the frame selection according to our proposed criteria rather than choosing the end frames. Note that the index "21-0" represents a pair of frames with minimal morphological change and maximal intensity variation. To the right of the x-axis of the graph in Figure 3.8, morphological change increases, and intensity variation decreased. Based on the result, the similarity (normalized correlation coefficient) loss due to intensity variation (see "21-0") is much more significant than the similarity loss due to morphological distortion (see peak at "14-7").

3.7.2 Tile mosaicking

Accuracy evaluation metric for image mosaicking

The proposed accuracy evaluation aims at estimating upper error bounds for automatic, semi-automatic and manual tile mosaicking techniques. To achieve the aim, I introduced

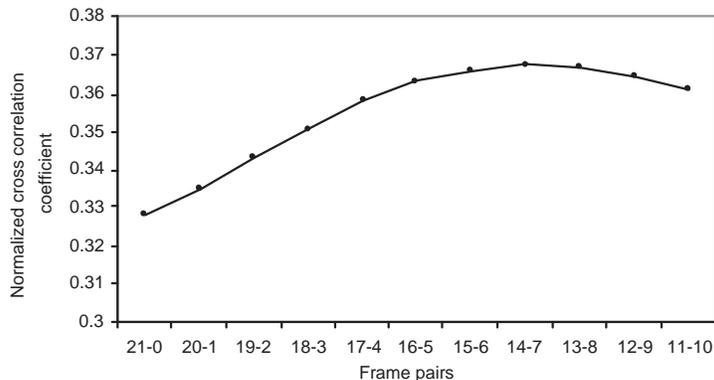


Figure 3.8: Evaluation of inter-frame similarity for two spatially adjacent sub-volumes with 22 frames after alignment. The similarity is measured by normalized correlation (vertical axis). The horizontal axis refers to the pairs of frames that start with the end frames of sub-volumes ($21 - 0 \equiv (\text{Sub-Vol 1, frame 21}) - (\text{Sub-Vol 2, frame 0})$) and finish with the middle frames ($11 - 10 \equiv (\text{Sub-Vol 1, frame 11}) - (\text{Sub-Vol 2, frame 10})$) .

three mosaicking methods (registration of x-y image tiles in a single frame of a physical section). Next, I designed an experimental evaluation methodology that addresses the issues of (a) defining optimality criteria for assessing registration accuracy, and (b) obtaining the ground truth (or reference) images, as encountered in real medical registration scenarios. After conducting experiments with human subjects consisting of experts and novices, I drew conclusions about the mosaicking methods, and analyzed the driving factors behind the results.

For an experimental setup, one could carve out several spatially overlapping tiles from one large image and use the original image as the reference (ground truth) image. However, this evaluation setup would not simulate the real problem of mosaicking multiple tiles acquired at different time instances. For example, it would not represent unpredictable intensity variations due to fluorescent imaging physics. Thus, I chose to establish the ground truth image and the locations of all n tiles in this image (denoted as T^{GT} in the following way.

First, I took an overview image of a specimen at 20x optical magnification and 3x3 high resolution image tiles at 63x optical magnification ($n = 9$). The overview image became the ground truth image. Second, tile images (63x magnification) were digitally down-sampled

to match the resolution of the overview image (20x magnification). Third, I found the best match between a down-sampled tile and the overview image with a template-based search using a normalized cross correlation metric. Fourth, the location of the best tile match was re-scaled to the original tile resolution. Fifth, steps one through four were repeated for all nine tiles to obtain a matrix of tile locations T^* . Sixth, the matrix T^* was normalized with respect to the tile location in the upper left corner (t_{1x}, t_{1y}) of the final mosaic image. Note that I used a bi-linear interpolation method for down- and up-sampling processes.

The uncertainty (pixel error distance) caused by the re-sampling (e.g., interpolation) procedure can be easily computed from the magnification factors. For example, for the re-sampling factor equal to $63/20$ ($= 3.15$), a down-sampled pixel will have contributions from an 3.15 by 3.15 pixel neighborhood. Thus, the uncertainty of the down-sampled and re-scaled pixel is equal to the maximum pixel distance in a 3.15 by 3.15 pixel region ($3.04(= 2.15\sqrt{2})$ pixels. Note that, geometrically, the maximum distance is a Euclidian distance between the *centers* of pixels in a region).

Finally, I denote the normalized matrix as the ground truth matrix T^{GT} of tile locations.

$$\vec{T}^{GT} = \begin{pmatrix} t_{1x}^{GT} & t_{1y}^{GT} \\ t_{2x}^{GT} & t_{2y}^{GT} \\ \cdot & \cdot \\ t_{nx}^{GT} & t_{ny}^{GT} \end{pmatrix} = T^* - \begin{pmatrix} t_{1x}^* & t_{1y}^* \\ t_{1x}^* & t_{1y}^* \\ \cdot & \cdot \\ t_{1x}^* & t_{1y}^* \end{pmatrix}, \text{ where } T^* = \begin{pmatrix} t_{1x}^* & t_{1y}^* \\ t_{2x}^* & t_{2y}^* \\ \cdot & \cdot \\ t_{nx}^* & t_{ny}^* \end{pmatrix} \quad (3.4)$$

Any other result of mosaicking is represented by a matrix of tile locations $T((t_{ix}, t_{iy}) \in T)$, and compared with T^{GT} . The mosaicking registration error $E_{translation}$ is computed as an average error distance according to the following.

$$E_{translation} = \frac{1}{n} \sum_{i=1}^n \sqrt{(t_{ix}^{GT} - t_{ix})^2 + (t_{iy}^{GT} - t_{iy})^2} \quad (3.5)$$

Note that the smaller the error implies the better mosaicking accuracy.

The proposed mosaicking evaluation methodology, using (1) the overview image acquired at low optical magnification as the true reference image and (2) the normalized correlation based estimation of tile locations T^{GT} , more closely simulates real image tile data than a set of carved out tiles from one image. Furthermore, the bias of tile locations T^{GT} coming from normalized correlation based matching can be quantitatively expressed by the correlation values in the vicinity of the best tile match with the overview image. The final remark is related to the selection of the error metric $E_{translation}$. Due to the intensity variations of CLSM images, it is preferred to use a registration accuracy metric based on spatial matches of salient structures rather than based on pixel intensity matches. The appropriateness of this metric selection could be demonstrated by taking images of the same specimen multiple times without moving it. If the metric would be based on pixel intensity matches then the metric would indicate falsely mis-registration in contrary to the metric based on spatial matches.

Statistical performance evaluation I describe a statistical test method to evaluate accuracy improvement of the feature-based approach against pixel-based approach. Let $\{E_i^P\}$ and $\{E_i^F\}$ be two paired sets of N measured error values for the pixel-based method and the feature-based method respectively obtained with the same data. In our experiments, the size of the set is relatively large ($N = 50$ for mosaicking and $N = 78$ for alignment). I assume that the paired error values are independent and follow a Gaussian distribution. The null hypothesis in our tests states that there is no improvement of the feature-based registration approach in comparison with the pixel-based registration approach. I performed the Student's t-test to prove or reject the null hypothesis [51]. I compute $\hat{E}_i^P = (E_i^P - \bar{E}_i^P)$ and $\hat{E}_i^F = (E_i^F - \bar{E}_i^F)$, where \bar{E}_i^P and \bar{E}_i^F are the average errors of each set. Then, I calculated

the t value for the paired t-test according to the equation below.

$$t = (\bar{E}^P - \bar{E}^F) \sqrt{\frac{N(N-1)}{\sum_{i=1}^N (\hat{E}_i^P - \hat{E}_i^F)^2}} \quad (3.6)$$

Given t , I obtained the confidence interval (p-value [51]) to prove or reject the null hypothesis (no improvement) using one-tailed cumulative probability distribution function $P(X \leq t)$ with $N - 1$ degrees of freedom. The results of statistical comparisons are shown in the next section.

Experimental result for image mosaicking

The overall experiments consist of mosaicking 3x3 image tiles. I report the results obtained from twenty human subjects (fifteen experts and five novices) who participated in the user study, and performed manual and semi-automatic image mosaicking. To assess registration consistency, novices performed registration three times with any given data set. Although the results from novices may be biased by “a learning effect”, I did not observe it in our experiments due to the small number of trial repetitions.

Figure 3.9 (a) shows the user interface for selecting matching points in two image tiles. Users selected one pair of feature points, one from each tile. Figure 3.9 (b) illustrates the interface for selecting regions that would be used for centroid calculation. In order to construct a mosaicked image, eight pairs of points or regions had to be selected. I used a set of nine images from a single physical section for mosaicking, and the experimental results are summarized in Figure 3.10 and Table 3.1, and the t-test result comparing the pixel-based and feature-based mosaicking is shown in Table 3.2.

The results lead to the following conclusions. First, fully automatic mosaicking using normalized cross correlation similarity is the fastest method, followed by semi-automatic (feature-based) and manual mosaicking. Second, manual pixel based image mosaicking is

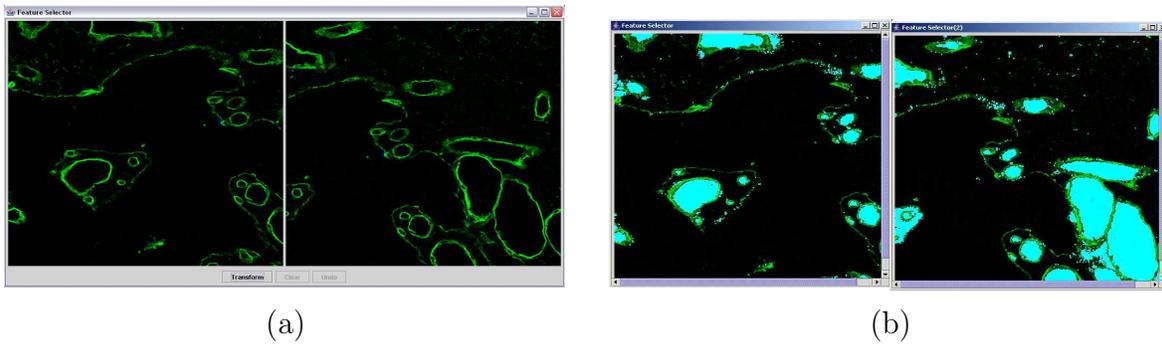


Figure 3.9: Software interface for (a) manual mosaicking, and (b) semi-automatic mosaicking with highlighted regions.

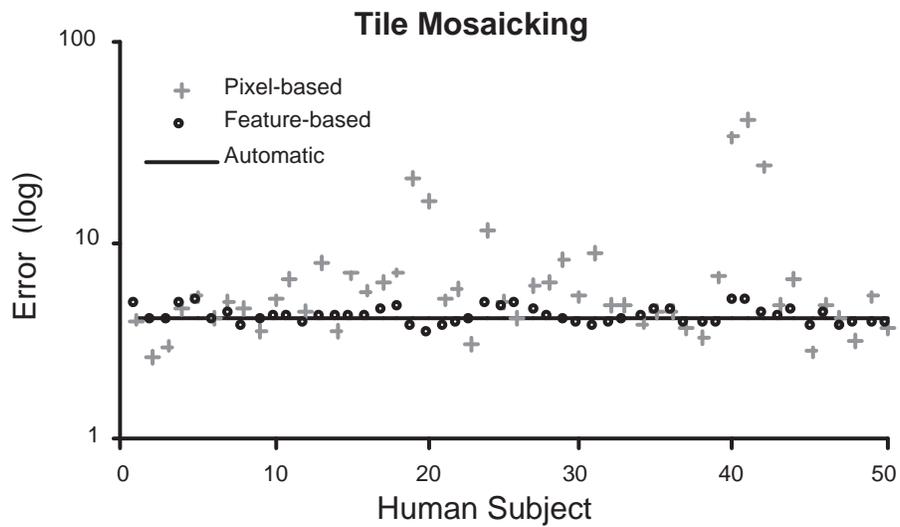


Figure 3.10: Mosaicking errors for all human subjects performing pixel-based (manual), feature-based (semi-automated), and intensity-based (fully automated) tile mosaicking.

Table 3.1: A summary of mosaicking experiments: 3 by 3 tiles with 512 by 512 pixel resolutions. Full automatic methods are performed by the normalized cross-correlation (NC) and by the normalized mutual information (NMI) on Pentium 4, 3.0 GHz.

| | Error (pixels) | | | | | |
|----------------------|----------------|--------|---------------|--------|------|------|
| | Pixel-based | | Feature-based | | Auto | |
| | expert | novice | expert | novice | NC | NMI |
| Average | 5.72 | 10.65 | 4.04 | 4.22 | 4.12 | 4.12 |
| Standard deviation | 3.42 | 11.83 | 0.32 | 0.47 | 0 | 0 |
| Total average | 6.96 | | 4.07 | | 4.12 | |
| Total std. deviation | 6.82 | | 0.35 | | 0 | |
| Upper bound (99.73%) | 27.42 | | 5.12 | | 4.12 | |
| | Time (seconds) | | | | | |
| | Pixel-based | | Feature-based | | Auto | |
| | expert | novice | expert | novice | NC | NMI |
| Average | 211.56 | 153.47 | 125.27 | 101 | 68 | 480 |
| Standard deviation | 132.32 | 95.06 | 56.96 | 45.66 | 0 | 0 |
| Total average | 197.03 | | 119.2 | | 274 | |
| Total std. deviation | 125.88 | | 55.01 | | 0 | |
| Upper bound (99.73%) | 574.67 | | 284.23 | | 274 | |

Table 3.2: The paired t-test result for errors of the pixel-based and the feature-based methods in Table reftable-mosaic.

| | Pixel-based vs. Feature-based |
|--------------------|-------------------------------|
| Degrees of freedom | 49 |
| t value | 3.019 |
| p value | 0.998 |

the least accurate with the highest standard deviation among all methods. Third, semi-automatic and fully automatic mosaicking methods are approximately equally accurate. Fourth, experts using the manual (pixel-based) mosaicking method selected one pair of points/regions more accurately (small average error) and consistently (small standard deviation) than novices although it took them more time. Fifth, the difference in mosaicking average errors and their standard deviations between experts and novices using the pixel-based method disappears when human subjects start using the feature-based mosaicking method. Sixth, the upper error bound of each mosaicking method can be estimated in pixels as the average plus three times standard deviation (99.73% confidence interval), which leads to about 4.12, 5.12 and 27.42 pixel error for the fully automatic, semi-automatic and manual methods respectively. Seventh, the t-test result in Table 3.2 shows that the null hypothesis (no improvement) is rejected with 99.8% confidence. Finally, the time-saving for experts and novices using semi-automatic method with respect to manual method is 41% and 36% respectively.

Although the intensity-based automatic methods look pretty attractive, note that there are mosaicking cases when the overlapping area of two adjacent tiles is characterized by either a lack of detected vascular features (feature based techniques fail) or significant spatial intensity heterogeneity (intensity based techniques fail). Figure 3.1 illustrates the former case. Thus, there is a need to evaluate manual and semi-automated mosaicking techniques for those cases when the intensity based techniques fail. In addition, it is not always the case that the fully automatic method will outperform the manual and semi-automatic methods (see Figure 3.10 and Table 3.1).

3.7.3 Segmentation

Figure 3.11 shows a pair selected frames from adjacent sub-volumes. To demonstrate the registration performance in Section 4, one image (right) is rotated about 90-degrees in addition

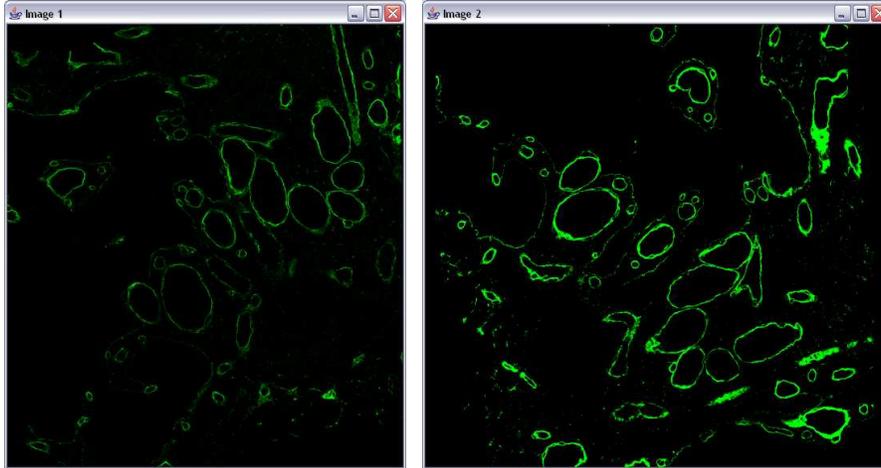


Figure 3.11: Example images that have to be aligned. Human tonsil tissue stained for laminin.

to measurement deformation (e.g., translation, rotation, and shear). The images are taken by CLSM from human tonsil tissue stained by laminin. Green closed contours represent blood vessels and some of them are partially opened. Figure 3.12 shows the corresponding 2D segmented images using a disk-based region growing method with threshold of 10 (left image) and 8 (right image), minimum area size of 80, and disk diameter of 3. It detected 28 (left image) and 32 (right image) closed contour regions for feature matching.

Figure 3.13 shows a 3D visualization of a pair of adjacent input sub-volumes that were used in Figure 3.11. Cylindrical structures are segmented out using sphere-based 3D segmentation method with sphere diameter of 3, threshold of 30, maximum number of voxels in a volume feature of 50000 (for both left and right sub-volumes) as shown in Figure 3.14. The visualized sub-volume dimension is $562 \times 552 \times 23$ (width \times height \times depth) (the sub-volumes are cropped by negligible amount in lateral coordinates to match with the image dimension of the reference sub-volume).

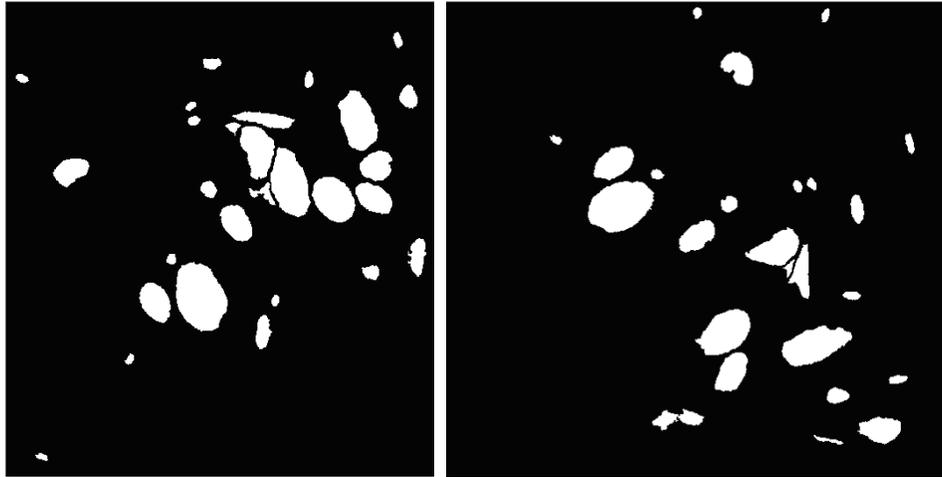


Figure 3.12: Segmentation of input images shown in Figure 3.11. Segmentation is performed by thresholding that is followed by connectivity analysis with a disk. The two images illustrate results obtained with different disk parameters [Left image - T.S. (threshold S) = 10, right image - T.T. (threshold T) = 8, M.R. (minimum size of a region) = 80, and D.D. (disk diameter) = 3].

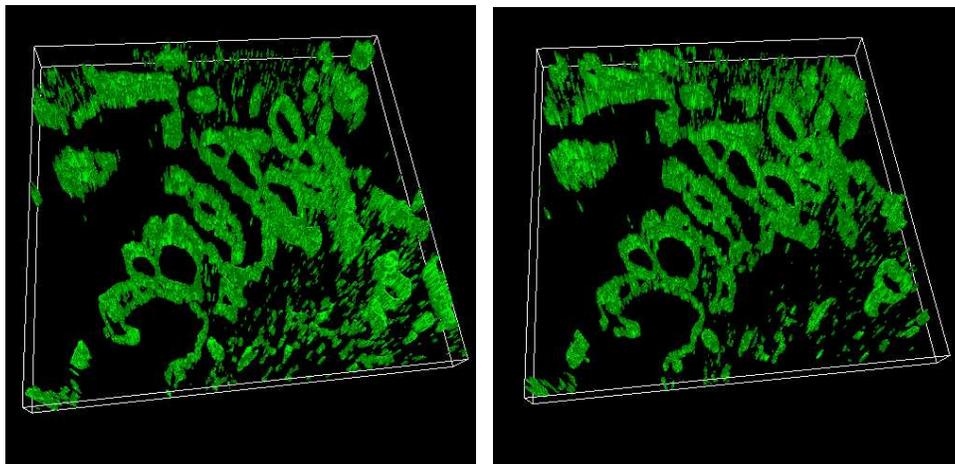


Figure 3.13: A pair of CLSM sub-volumes: (left) sub-volume 1 (upper physical section) and (right) sub-volume 2 (lower physical section).

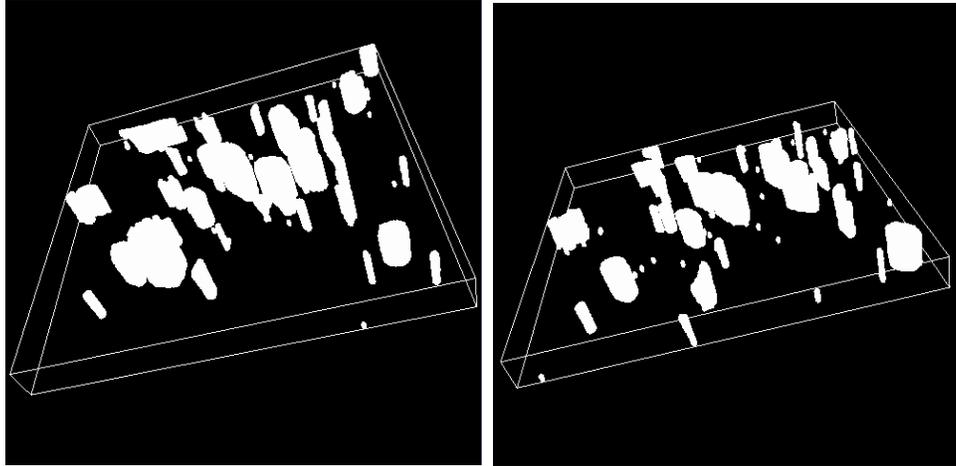


Figure 3.14: Segmented sub-volumes of Figure 3.13.

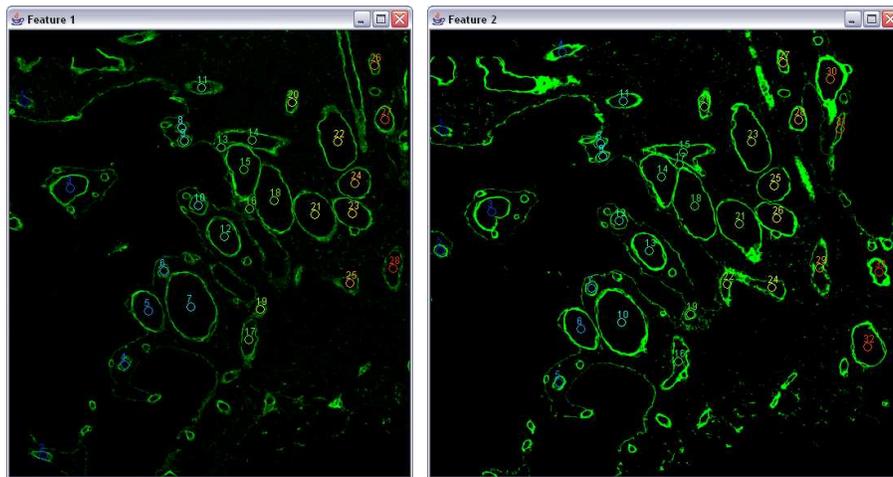


Figure 3.15: The centroids are labeled from the segments shown in Figure 3.12.

3.7.4 Feature detection

Based on the segmentation shown in Figure 3.12, I computed a set of centroids as shown in Figure 3.15. The centroid labels does not necessarily match in the shown pair of images.

Figure 3.16 shows a pair of *exaggerated* trajectories from the volume segmentation of the sub-volumes in Figure 3.14. The label does not necessarily matches with matching pairs. Red line-segments represent centroid locations, and green curve segments represents third degree polynomial curves by least squares fitting to the centroid trajectories.

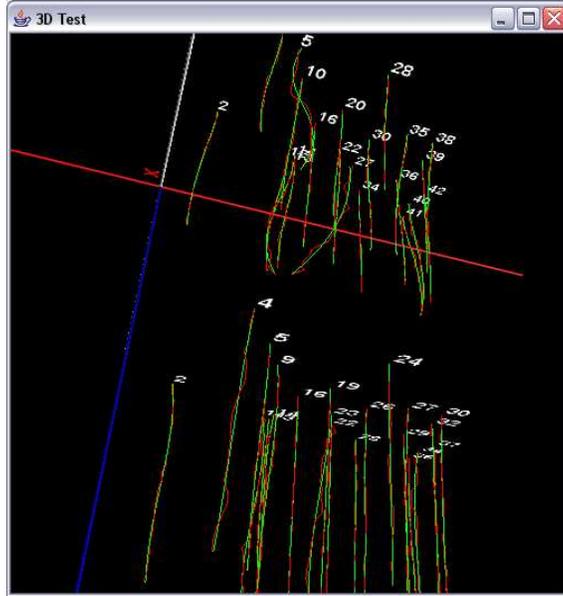


Figure 3.16: A pair of sets of (exaggerated) trajectories from adjacent sub-volumes. Numbers indicate the label of matching trajectories. Red line segments indicate centroids, and green curves indicate a polynomial function with degree 3.

3.8 Summary

In this section, I presented methods of sub-volume preprocessing required for automated 3D volume reconstruction process. The pre-processing step consists of sub-volume intensity analysis, tile mosaicking, segmentation, and feature detection. Frame selection and mosaicking (manual, semi-automated, and fully automated) steps are performed regardless which registration methods the volume reconstruction uses, e.g., 2D or 3D information. In case of using 2D information for registration, the selected frame from the sub-volume analysis process is further needed through entire registration process, such as 2D segmentation, feature matching, and intensity-based optimization. Segmentation and feature detection processes are performed based on the dimensionality of registration method, such as region segmentation followed by centroid computation in 2D case, and the volume segmentation followed by centroid trajectory computation in 3D case.

In addition, I would like to add a few comments about the performance robustness of fully automatic and semi-automatic methods. Fully automatic mosaicking method based on

normalized correlation or normalized mutual information might not achieve the best performance when corresponding salient features have spatially mismatched intensity variations. Semi-automatic method based on region centroids might not be used when closed regions cannot be detected due to the spatial structure of an imaged specimen or a very low image quality, for instance, a low signal-to-noise (SNR) ratio and a large amount of intra-region noise.

In next chapter, I deal with feature matching used for 2D and 3D case in different automation levels, and matching refinement problem for optimal global transformation parameter estimation.

Chapter 4

SUB-VOLUME ALIGNMENT

4.1 Introduction

In this chapter, I describe multiple approaches to sub-volume alignment (registration along axial direction). The problem consists of feature matching using multiple automation levels, matching refinement based on intensity or morphology (e.g., centroid trajectory), global transformation estimation, and sub-volume transformation. The input to this process is the set of frames selected from the sub-volume intensity analysis step and computed features in the pre-processing step (e.g., centroid coordinates, area, and centroid trajectories). The output is a set of transformed sub-volumes by an affine transformation. The transformed sub-volumes are adjoined to reconstruct an entire 3D volume for visualization.

I define the 3D reconstruction problem as a registration problem [97]. The goal of 3D reconstruction is to form a high-resolution 3D volume with large spatial coverage from a set of spatial tiles (small spatial coverage and high-resolution 2D images or 3D cross section volumes). The main challenges of image alignment include the transformation technique and model selection problems. In the past, the transformation technique based on intensity correlation has been applied to many medical image modalities [60] other than the fluorescent CLSM modality. Nonetheless, applying the same techniques to the image alignment problem of CLSM images is more difficult due to (1) high computational cost, (2) spatial intensity heterogeneity, and (3) noise issues as explained below.

First, the computational difficulty arises from a large image size and high degrees of freedom for complex transformation models. In our case, the computational complexity due

to a large amount of data (3D stacks with many physical sections to obtain sufficient depth information) with high spatial resolution (around 2500 by 2500 pixels) should be considered when applying an affine transformation (6 degree-of-freedom). One could find methods in the literature that process large data of other imaging modalities by using multi-resolution (or pyramid) based techniques [60]. However, in the case of CLSM images, the local minima problem is much more severe due to high spatial and depth intensity heterogeneity (attenuation) [73]. Second, varying signal to noise ratio due to aforementioned intensity heterogeneity should be considered. Third, noisy (spurious) features with high intensity values due to unbound fluorescence have to be handled.

In Section 4.3, I propose feature matching methods in three different automation levels, such as manual (denoted as pixel-based), semi-automated (feature-based), and fully automated (feature-based) method. One could notice that one major difference between mosaicking and alignment is the approach to the fully automated method. As discussed earlier, for affine transformation model, intensity-based method is infeasible due to intensity heterogeneity and high computational cost. For sub-volume alignment, I used fully automated feature-based method by solving a variation of the Procrustes problem [24] using centroid and area of detected segmentations. Suggested manual and semi-automated methods require human intervention such as pixel or region selection. As well as on-site user intervention, I present manual and semi-automated method using web services.

The use of proposed web services is motivated by the fact that the problem of 3D medical volume reconstruction requires significant computer resources and human expertise in medical and computer science areas. Specifically, web services are implemented as an additional layer to a data flow framework called Data to Knowledge (D2K). In the collaboration between UIC and NCSA, pre-processed input images at NCSA are made accessible to medical collaborators for registration. Every time UIC medical collaborators inspected images and selected corresponding features for registration, the web service at NCSA is contacted and

the registration processing query is executed using the Image to Knowledge (I2k) library of registration methods. Co-registered frames are returned for verification by medical collaborators in a new window. I present 3D volume reconstruction problem requirements and the architecture of the developed prototype system at <http://i2k.ncsa.uiuc.edu/MedVolume>. I also explain the tradeoffs of our system design and provide experimental data to support our system implementation. The prototype system has been used for multiple 3D volume reconstructions of blood vessels and vasculogenic mimicry patterns in histological sections of uveal melanoma studied by fluorescent confocal laser scanning microscope.

In Section 4.4, I propose matching refinement processes of optional intensity-based approach in 2D registration and trajectory fusion approach in 3D registration. The common goal of two approaches is to compute an optimal global affine transformation using different registration quality evaluations. First, the intensity-based approach is aimed to achieve higher image intensity similarity (e.g., normalized cross correlation), and the problem is viewed as a search problem for an optimal transformation. Second, the trajectory fusion approach is to optimize global transformation which preserves morphological smoothness of medical structures (called features, e.g., blood vessels) inside of the reconstructed 3D volume.

4.2 Related Work

There exist many techniques for 3D volume reconstruction and many commercial tools from multiple vendors that could be used for image registration [162, 64, 16, 22, 50, 143]. An overview of 3D registration tools for MRI, CT, confocal, and serial-section data for medical/life-sciences imaging is provided at the Stanford ¹ or at the NIH web sites ². One

¹Stanford web page with references to 3D volume reconstruction software packages.
<http://biocomp.stanford.edu/3dreconstruction/refs/index.html>
and <http://biocomp.stanford.edu/3dreconstruction/software/>

²NIH website.
<http://www.mwrn.com/guide/image/analysis.htm>

could list a few software tools that have been developed specifically for CLSM, for example, 3D-Doctor, Science GL, MicroVoxel, 3DVIEWNIX or Analyze. Most of these tools use manual registration methods, and users have to make manual selections before any particular software reports registration error associated with registered images. Some software packages include semi-automatic or fully automatic 3D volume reconstruction for specific imaging modalities under the assumption that visually salient markers have been inserted artificially in imaged specimens. For instance, 3D-Doctor provides a maximum likelihood algorithm for aligning slices under such assumption.

The problems of 3D volume reconstruction and medical cross section registration have been approached by an overwhelming number of researchers over the past several decades [60, 64, 104, 109, 118, 145, 157] and remain still an open problem. There have been several survey papers about registration approaches that include selection of registration variables based on user decisions [16, 97, 162]. In section 2.4, I specifically focused on optimal selections of registration variables that are inherent parts of 3D volume reconstruction process. The 3D volume reconstruction problem is defined as a registration problem without fiducial markers [97]. The goal of 3D reconstruction is to form a high-resolution 3D volume with large spatial coverage from a set of spatial tiles (small spatial coverage and high-resolution 2D images or 3D cross section volumes).

Regarding geographically distributed alignment process, there exist commercial software packages that address the problems of image access and navigation with other than web service approaches. In the medical domain, these solutions are known as “Virtual Microscopes” and primarily owned by companies (like Bacus Laboratories, Inc. and Aperio Technologies). In the GIS domain, the solution for accessing all IKONOS aerial photos before 9/11 was developed by Microsoft (navigation capability without annotation or computation capability).

Among the most recent solutions using web services, I should mention a new suite of web service tools to facilitate multi-sensor investigations in Earth System Science that is sponsored by NASA [159], and the ArcWeb tools which are web services implemented for

GIS data operations. The tools for NASA are developed based on a framework using grid workflows (known as SciFlo) [154]. Other workflow frameworks, like Kepler [3, 96], have not been used for applications using web services. The ArcWeb services are proprietary, and focus primarily on (a) accessing images with the size in the terabytes, and (b) reducing data storage and maintenance costs.

In contrast to the previous work, the presented work for geographically distributed alignment process is based on the data flow framework called D2K [149]. D2K is a visual programming environment and data flow execution engine developed at NCSA for data mining applications (prediction, discovery, and anomaly detection with data management and information visualization). The underlying 3D volume reconstruction algorithms came from a library of image analysis tools called I2K [8], also developed at NCSA. Our prototype system has been used in practice for 3D reconstruction of uveal melanoma tissues based on the NIH-funded collaboration between UIC and NCSA [76].

The previous methods developed for the problem of pair-wise sub-volume alignment can be classified as intensity based methods (e.g., normalized cross correlation or normalized mutual information) or morphological feature-based methods (e.g., shape matching, semi-automated method) [9]. Intensity-based methods are typically performed by selecting or generating a pair of representative images from adjacent sub-volumes [85], determining a global transformation parameters by minimizing a similarity metric for all possible transformations, and applying the computed global transformation to the sub-volume. Intuitively, using the intensity-based approach, one could select a pair of end frames that are near the boundary of adjacent physical sections to minimize morphological distortion of the structural changes. However, it is well known that due to spatial intensity heterogeneity of end frames in CLSM sub-volumes, correlation techniques would result in very low similarity measures, which leads to undesirable sensitivity to noise and inaccurate alignments [11],[73]. To overcome the problems with end frames, alternative approaches for frame selection have been suggested that would select the highest contrast images [85, 88] or generate a representative

image based on sub-volume analysis, e.g., a projection [20]. Although the intensity based methods are feasible in some cases, it should be noted that they are computationally expensive due to a large search space of transformation parameters. If the search is limited then the alignment accuracy might be compromised and could result in discontinuities along z-axis (depth).

Since partial 3D information about sub-volumes are available in CLSM data, e.g., physical section with limited thickness, we approach the registration problem by utilizing the 3D information from each sub-volumes such as trajectories of volumetric segment centroids. Trajectory descriptors are mainly used in a motion tracking problem in computer vision [41, 150]. It is typically modeled as 2+t dimension, i.e., 3-dimension with replacing z with t [74] in video frame sequences. The trajectory fusion is performed by estimating spatial alignment and temporal synchronization (temporal alignment) between the two sequences, using all available spatio-temporal information [158]. This work is related to trajectory tracking problem in terms of fusing trajectory segments into one trajectory [139]. However, the goal of the trajectory fusion in 3D medical volume reconstruction is to optimize visual registration accuracy by estimating a global transformation as opposed to individually correlating each pairs of trajectory segments.

4.3 Feature Matching

The feature matching problem in this section is reduced to a problem of establishing correspondences of pixel locations of matching pairs in a pair of images from depth adjacent sub-volumes. In this section, I provide feature matching methods in three different levels of automation such as manual, semi-automated, and fully-automated method. Manual and semi-automated methods are implemented and studied by (1) on-site experiment, e.g., software packaging for users and (2) off-site experiment, e.g., geographically distributed alignment process (web service). The fully-automated method is performed by solving a

variation of the Procrustes problem to establish the feature correspondences [24].

4.3.1 Manual alignment approach using pixel selection

Manual alignment method is one of the most popular method and still being used in many clinical applications. The typical process of the manual alignment consists of (1) displaying pairs of images to be registered in GUI (Graphic User Interface), (2) having a user select sets of matching pairs of points, and (3) computing a global transformation parameters based on the user pixel selection. Another possible scenario may be overlaying a pair of images and adjusting transformation manually while visually inspecting the images. The main advantages of the manual alignment method include (1) low software implementation cost (simple software design), (2) flexibility of input data set (e.g., multi-modal alignment, independence of image contents, and (3) possible integration (exploitation) of user expertise. However, the alignment accuracy (quality) may significantly depend on user expertise, experiment condition, and user interface (e.g., HCI). In this work, I present a manual alignment method by pixel selection using a same user-interface in image mosaicking in Section 3.4, and use the manual alignment method as a benchmark for performance evaluation when comparing with semi-automated and fully automated approach.

4.3.2 Semi-automated alignment approach using region centroids

The registration was conducted by semi-automated, i.e., region centroid-based, alignment. To support the claims about the spatial intensity heterogeneity and the presence of noise, I evaluated an intensity-based similarity metric (normalized correlation) for pairs of images from two *registered* CLSM sub-volumes. The low magnitudes of these similarity values (approximately in the interval [0.325, 0.365]) proved that the intensity-based automatic alignment would not be robust and would very frequently fail. As a consequence, I did not apply the correlation based technique developed for image mosaicking to the image align-

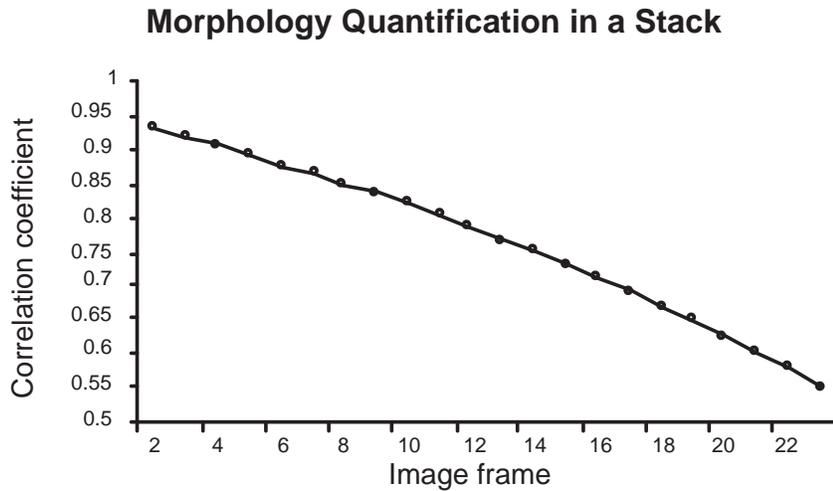


Figure 4.1: Morphology quantification in a CLSM stack: x-axis represents a frame index (along depth) which is compared with the first frame.

ment problem.

Furthermore, the problem of image alignment (or registration along z-axis) is much harder to automate than the problem of mosaicking because images of cross sections are much less similar than images of spatial tiles due to the process of cross section specimen preparation (sample warping due to slicing), intensity variation (confocal imaging), and structural changes (bifurcating structures). In Figure 4.1, I quantified the morphological changes along depth in a single physical cross section (sub-volume) by computing normalized cross-correlation coefficient between the first and other image frames.

The manual and semi-automatic methods for image alignment differ from the methods described for image mosaicking by the need to select at least three pairs of corresponding registration points as opposed to one pair of points sufficient in the case of image mosaicking. Feature matching using semi-automated method is performed by (1) displaying a pair of images that are selected from depth adjacent sub-volumes, (2) computing centroids of closed pre-segmented regions on user selection, and (3) establishing coordinate correspondence of selected centroids.

4.3.3 Fully-automated alignment approach

The homology of a pair of features is closely related to an image transformation model usually defined a priori. It is selected based on expected deformations during specimen preparation and image acquisition. Depending on a selected transformation, one finds representative invariant features under the selected transformation (e.g., area remains invariant under the rigid transformation). I used a combination of rigid transformation model (translation and rotation; 3 parameters) and an affine transformation model (translation, rotation, scale and shear; 6 parameters) for the automated feature matching problem. I investigated methods for establishing feature correspondences and their assumptions about registration transformations (as known as the Procrustes problem).

I introduce a method to establish a correspondence between two sets of segment features. The problem is a variation of the Procrustes problem [24], where one estimates transformation parameters based on centroid and area characteristics of segments (segment feature). The developed solution to the correspondence problem consisted of two phases. First, a coarse rigid transformation is estimated by (a) matching Euclidian distances between pairs of centroids (denoted as distance-based matching) and (b) comparing segment areas. Although this type of correspondence estimation is robust to partial mismatches, it is insensitive to angular differences (see Figure 4.2). Second, the set T (see Figure 4.3) is rotated and translated to the coordinate system of the set S according to the parameters computed in the first phase. Finally, I found correspondences by matching vector distances, as opposed to Euclidean distances used in the first phase (denoted as vector-based matching). This computation is appropriate for correcting wrong correspondences from the first phase, but would not be robust on its own for highly rotated feature sets.

The developed solution to the correspondence is presented here. Let $S_a(a = 1, \dots, N^S)$ and $T_b(b = 1, \dots, N^T)$ be a pair of extracted sets of shape features from selected image frames I_a^S and I_b^T such as $S_i = \{\vec{c}_i^S, a_i^S\}$ and $T_j = \{\vec{c}_j^T, a_j^T\}$, where \vec{c}_k^Ω are the 2D centroid locations

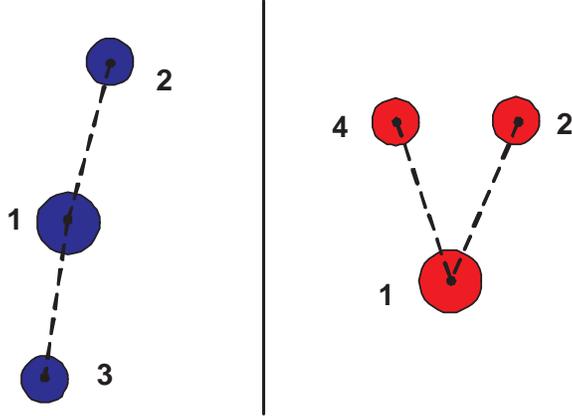


Figure 4.2: Illustrations of the case when distance-based matching leads to an erroneous match of the segments labeled as 3 (left) and as 4 (right).

and a_k^Ω are the area values. The number of features in each image is N^Ω , $\Omega = \{S, T\}$, and N^S is not necessarily the equals to N^T . An illustration of the correspondence problem is presented in Figure 4.3 for $N^S = 3$ and $N^T = 5$.

In the first phase, the computation consists of (1) calculating a matrix of mutual distances d_{iu}^S and d_{jv}^T ($i, u = 1, \dots, N^S$ and $j, v = 1, \dots, N^T$) for each set of feature centroids (e.g., dotted lines in Figure 4.3), (2) finding a set of matching pairs of segments M_k , and (3) sorting the matched segments based on the residuals of all matched centroids with respect to the estimated rigid transformation. To find a matching pair of segments, I first select a pair of segments S_i and T_j denoted by *pivot segments*, and introduce three similarity measures, such as (a) the area ratio $A_{ij}(u, v)$, (b) the difference of Euclidian distances (residual) $D_{ij}(u, v)$, and (c) the number of matching segments Q_{ij} .

Pivoted by segments (S_i, T_j) , I first compute the area ratio as for matching segments $A_{ij}(u, v) = |a_u^S/a_v^T - 1|$. Next, the difference of Euclidian distance is computed as $D_{ij}(u, v) = |d_{iu}^S - d_{jv}^T|$. Finally, the number of matching pairs Q_{ij} is calculated by counting the number of pairs that satisfies both $D_{ij}(u, v) < \delta_1$ and $A_{ij}(u, v) < \varepsilon_1$, where the value of δ_1 is the dissimilarity upper bound of a pair of distances, and ε_1 is the divergence upper bound of a ratio of two segment areas from one. I not only maximize Q_{ij} but also remove the matches that do not satisfy the inequality $\frac{Q_{ij}}{\text{Min}(N^S, N^T)} \geq \lambda_1$, where $\lambda_1 \in [0, 1]$ is the lower bound of

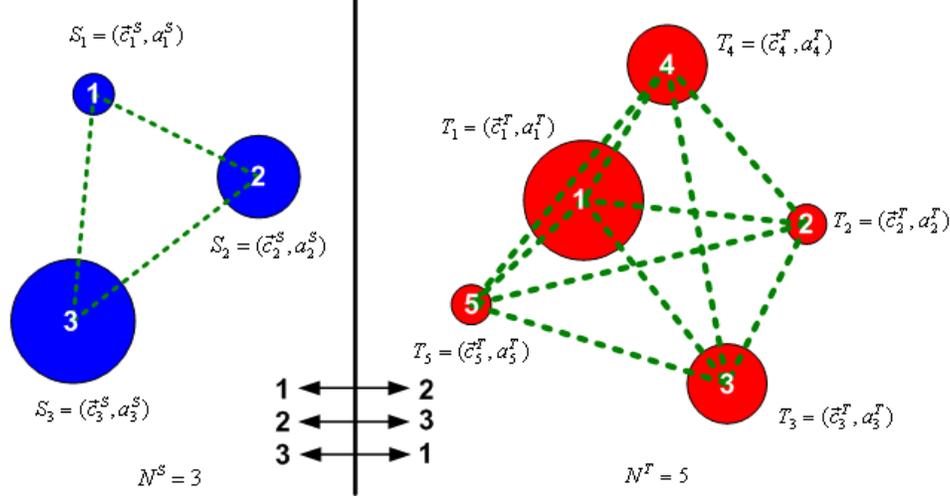


Figure 4.3: Illustration of the correspondence problem for two sets of features S_i and T_j with unequal number of features $N^S = 3$ (left) and $N^T = 5$ (right). Segments are shown as disks characterized by their index i , area a_i^S and centroid location \bar{c}_i^S . Dashed lines represent the Euclidean distances between any two centroid locations.

normalized number of matches for a single pivot segment. Figure 4.4 shows the description of the defined similarity metrics.

To find a final match M_k , we maximize a score function $f(\cdot)$ by incorporating three similarity metrics as follows:

$$M_k = \left\{ (i, j) \mid (i, j) = \arg \max_{i, j} \left\{ f \left(Q_{ij}, \frac{1}{\bar{D}_{ij}}, \bar{A}_{ij} \right) \right\} \right\} \quad (4.1)$$

where \bar{D}_{ij} and \bar{A}_{ij} are average error distance and area ratio pivoted by (S_i, T_j) for all matching pairs of (u, v) , $0 \leq k \leq N^{ST}$, and N^{ST} is the number of matching pairs of segments (pivots). The function $f(\cdot)$ may be defined as an energy function or, more simply, a weighted product of all components. In implementation, I used a weighted product of each component with normalization.

In the second phase, I first rotated the features in the set $T_b (b = 1, \dots, N^T)$ to match with the coordinate system of the set $S_a (a = 1, \dots, N^S)$ according to the parameters which is computed by selecting two best matches in the first phase. Next, vector-based matching is performed in a similar way as distance-based matching in the first phase. The major differ-

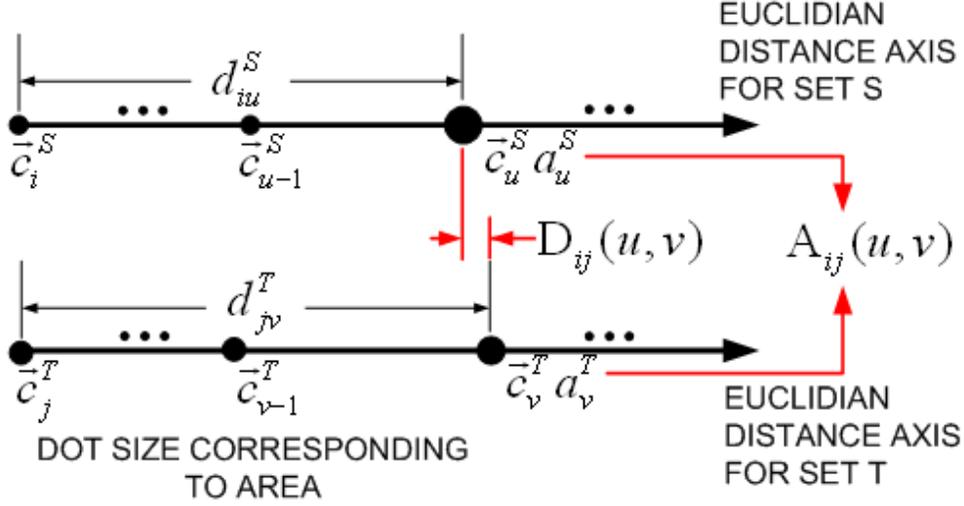


Figure 4.4: Illustration of distance-based matching for two pivot segments $S_i = \{\bar{c}_i^S, a_i^S\}$ and $T_j = \{\bar{c}_j^T, a_j^T\}$. In order to find the best match, distance and area ratio of pairs of segments are compared to satisfy $D_{ij}(u, v) < \delta_1$ and $A_{ij}(u, v) < \varepsilon_1$.

ence is in replacing the Euclidian distance metric with the vector distance metric $D_{ij}^*(u, v)$ defined in following equation:

$$D_{ij}^*(u, v) = \|(\bar{c}_i^S - \bar{c}_u^S) - (\bar{c}_j^T - \bar{c}_v^T)\| \quad (4.2)$$

The vector distance incorporates both Euclidian distance and angular constrains about a pivot segment. Therefore, the matching performance is greatly increased comparing with the first phase in terms of accuracy (mis-match) and the number of matching segments. The second phase contains the same three parameters as the first phase, and we denoted them as ε_2 , δ_2 and λ_2 . In almost all experimental cases, we have set $\varepsilon_1 = \varepsilon_2$, $\delta_1 = \delta_2$ and $\lambda_1 = \lambda_2$.

4.3.4 Geographically distributed alignment process

The problem of 3D medical volume reconstruction requires significant computer resources and human expertise in medical and computer science areas. Web services can be viewed as the mechanism for establishing a collaborative environment between medical and computer science collaborators and combining their geographically distributed expertise. Web services

also allow us to quickly prototype an application and send a URL to UIC to obtain fast user feedback. In our collaborative environment, the data of 3D cross sections are large in size, and the registration computation requirements are significant. Therefore, the above benefits of web services outweigh the timing overhead of web services.

In a collaborative environment with medical and computer science collaborators, the goal is to reconstruct 3D medical volume from high resolution microscopy images of several cross sections. High resolution mosaic images of cross sections are formed from a large set of tiles, and then the mosaic images are aligned to construct a 3D volume. From a medical collaborator viewpoint, 3D volume reconstruction requires (a) sophisticated 3D volume reconstruction algorithms and (b) computation and storage beyond the capability of a desktop computer, i.e., computer science expertise and resources. From a computer science collaborator viewpoint, 3D volume reconstruction requires selecting pairs of matching features for cross section alignment, i.e., medical expertise. Thus, there is need to develop a cyber-infrastructure environment where the computational resources and the expertise of remotely located medical and computer science collaborators can be integrated.

Application scenario

I address the problem of 3D volume reconstruction in a collaborative environment by using web services. Our prototype system, shown in Figure 4.5, enables medical and computer science researchers to solve 3D volume reconstruction problems using web services.

The developed solution consists of the following workflow. First, a medical collaborator, e.g., from UIC, acquires images and sends data to his or her computer science collaborator. It is also possible that the medical collaborator uploads the data assuming a high bandwidth connection. Second, a computer program automatically mosaics image tiles, selects the most salient frame from each sub-volume, segments the selected frames and pre-computes centroids of all segments. The pre-processed images and centroid information are packaged

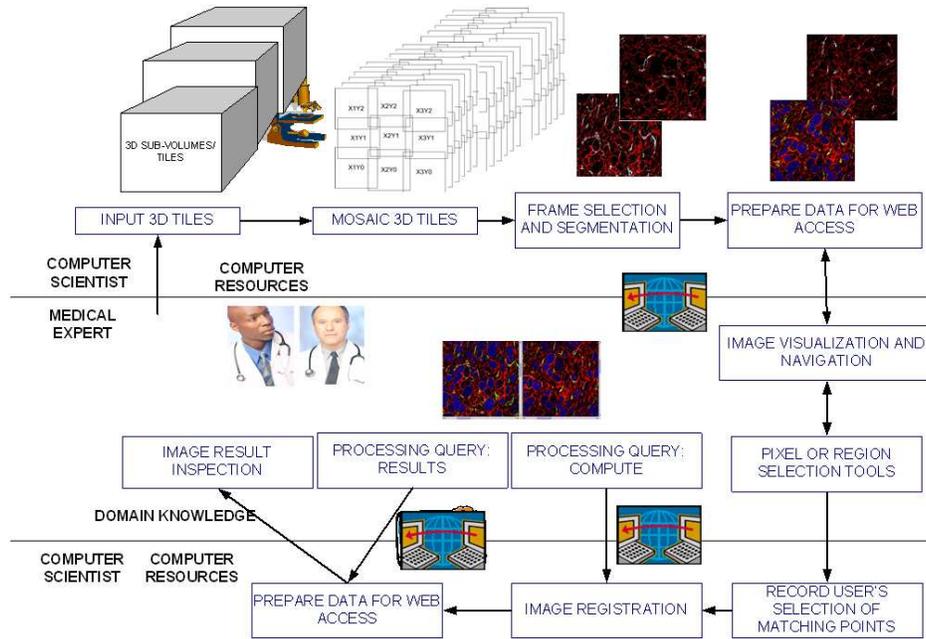


Figure 4.5: An overview of the current application scenario.

for web access. Third, the medical collaborator will be notified about the URL designed for accessing and navigating the image data, as well as for selecting registration points and visualizing registration results.

The medical collaborator selects matching features for cross section alignment by using standard human computer interfaces (HCI), and our developed image navigation tools [85, 88]. The points are saved at NCSA for use later when the final 3D volume is reconstructed. Once the medical collaborator has selected their points a query is sent from UIC to NCSA to request registration computation. After the computation is completed, the results can be previewed by the medical collaborator. The preview shows how well the image alignment was done. If the medical collaborator is not satisfied with the image, registration they can select more points or modify the already selected points. In the aforementioned workflow, all operations that require intensive computation are performed at NCSA (computer resource location), while all operations that require medical domain knowledge (image acquisition, registration point selection, and volume inspection) are performed at UIC (do-

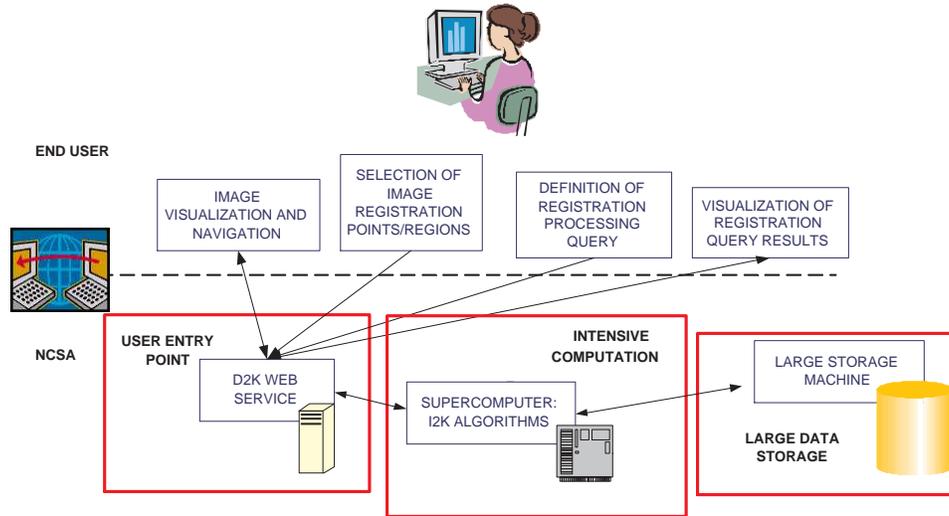


Figure 4.6: The overall architecture of the proposed system.

main expertise location).

System design using web services

The overall architecture of the proposed system is illustrated in Figure 4.6. It is assumed that (a) all connections inside of NCSA are high speed, and (b) there might be additional high speed connections between the storage machines and the NCSA supercomputer. A user (a medical expert) will be able to interact with the system if they have a low bandwidth connection to the Internet. If the image size is larger than user's memory or screen size then the image access would be enabled by (a) creating a multi-resolution image pyramid, (b) tiling images at multiple resolutions, and (c) storing pyramid and tile information in a database. The current implementation assumes that the two image frames to be registered do not exceed available memory of the medical expert's desktop.

The problem of image navigation is solved by (a) selecting and displaying image sub-areas at a chosen resolution, (b) panning through spatially large images using vertical and horizontal slider bars, and (c) selecting sub-set of bands to display to original data, pre-processed data (segmentation) and registered data. I should note that in our architecture,

the user interface was implemented using Java applets. The reason for choosing Java applets comes from the fact that web services have been designed for web-based software interoperability but not for image visualization and data interaction purposes. The additional problem of registration feature selection is approached by providing tools for either pixel selection or region selection that is converted to a region centroid. The problem of intensive computation and extensive storage is tackled by (a) preparing a set of processing algorithms accessible by web services, and (b) using storage and computational resources at NCSA.

The proposed approach is based on the storage-computation paradigm, where a user is running only a "thin" client applet with small storage-computation resources and all storage demanding and computationally intensive operations are performed at NCSA. If a user would like to perform all operations at his/her desktop then he/she could install the entire system locally on his/her machine. We have not pursued the paradigm where data sets would be transferred to a local machine and computationally intensive processing of large size image data would be performed on a local machine.

Tradeoffs of System Design: When designing the prototype we tried to optimize the system so that an end user (the medical expert) would have a responsive system that would show the most up-to-date information with a limited amount of resources used. We considered the following four tradeoffs: image data transfer, image segmentation computation, image compression, and image transformation.

- Image data transfer: Image transfer can be executed as a transmission of the full size image vs. the use of image pyramids [128] [19], in which the system would load the image as needed, for example, when a user zoomed in/out of the image and panned around the image. In general, medical images can be large in size. In our test case, we worked with images that are approximately 1300 x 2700 pixels, but could poten-

tially be larger. Using image pyramids will reduce the initial transfer of data but will continuously download data of the image tiles.

- Image segmentation computation: Segmentation of the input images can be performed on a server side or on a client side. I segment the input images in order to improve registration accuracy by replacing a pixel location with a more reliable region centroid location. The segmentation could be performed before a user selects a region (segmentation results have to be transferred) or after a user chooses a location to segment locally (computation has to occur on a client side). Segmentation on the client side is a CPU intensive task. Segmentation on the server will reduce the client CPU time but will increase the image size that needs to be transferred.
- Image compression: Medical image data can be transferred compressed or uncompressed. Transferring compressed input images will result in smaller data transfer, but will require some client CPU time to decompress the images. Transferring uncompressed images will increase the data transfer requirements. In our prototype, compressed images led to a size reduction from 15.3Mb to 2.27Mb.
- Image transformation: Computation of image transformation parameters and the transformed images can occur on a client side or on a server side. Once the expert has selected at least three points or regions in each image, we can calculate the image transformation parameters for an affine transformation and transform one image into the coordinate system of the other image. If the computation is performed on a client side, then there would be less network traffic but higher demands on the CPU usage on the client side. If it is performed on a server side, then there would be less CPU usage on a client side but more network traffic to transfer transformed image data.

In our application scenario, the objective was to create a prototype that would be very responsive to a medical expert, and it would not require significant computational resources on a client side. Thus, we tried to minimize the computer requirements for a medical expert

Table 4.1: Tradeoffs considered during a prototype system design using web services for 3D volume reconstruction. C refers to a constant, + indicates increase and - is decrease. Multiple + or - symbols indicate the magnitude of the value specified in each column.

| | Client CPU | Client RAM | Bandwidth | Number of Queries |
|-------------------------------|------------|------------|-----------|-------------------|
| Image pyramid | C | - | ++ | + |
| Client does segmentation | +++ | - | - | C |
| Compression | + | + | — | C |
| Transformation on client side | +++ | + | - | - |

and leverage the virtually unlimited resources of NCSA. Table 4.1 lists the tradeoffs we have considered during the system design. The tradeoff considerations revolved around the following metrics: (1) CPU time on a client side, (2) RAM on a client side, (3) bandwidth for image transmission, and (4) an overhead associated with number of queries. According to Table 4.1, we chose the options that would limit the resources at the client (UIC) side. Our final choices were (1) to transfer full images, (2) to perform segmentation on a server side, (3) to use image compression before sending data, and (4) to compute image transformation parameters and transformed images on a server side. The images used in our prototype were small enough in size to fit into RAM memory of a standard desktop computer, and hence we decided not to use the image pyramids.

Prototype Solution The prototype solution is available in our web site³. I have tested the prototype with the Microsoft Internet Explorer 6.0 using the Sun Java Runtime Environment (JRE) 1.4.2 browser plugin. The left lower corner of the browser conveys messages about the execution steps, e.g., Applet ncsa.uic.RegistrationApplet started. A user can select image frames to register using the drop down menu with image names. Selected images are compressed at NCSA site, transferred to a client site, decompressed and displayed in the image panels (see Figure 4.7).

After selecting at least three pairs of matching points/segments, the compute button is

³<http://i2k.ncsa.uiuc.edu/MedVolume/>

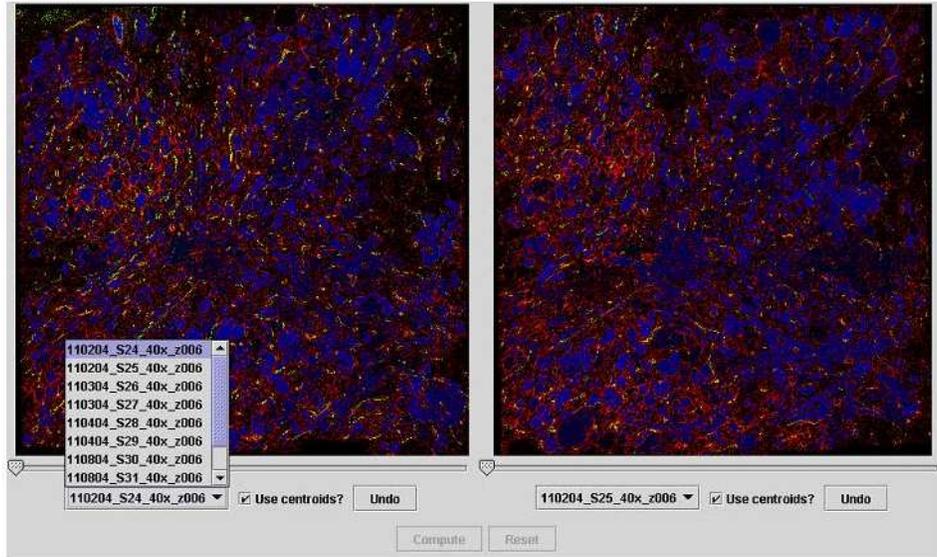


Figure 4.7: Image selection menu.

enabled and the D2K Web Service can be contacted. If an image panel is in the point selection mode then the left mouse click will define the pixel location to be used for registration. If an image panel is in the segment selection mode then the left mouse click will be replaced with the centroid location of the segment that contains the mouse click location. Selection of registration points using the centroid feature approach is illustrated in Figure 4.8.

After launching web services by pressing the button "Compute", a new Java window labeled "Executing Job" will appear containing information about the compute job while the affine transformation parameters are computed. When the job is completed, the button "View Results" becomes active and the results can be visualized on the client side as it shown in Figure 4.9. The visualization will contain a seven band image that was sub-sampled in order to decrease the file size. The seven band image contains the original left image bands, the transformed right image bands, and one black band for visualization purposes.

If the registration accuracy is satisfactory to the medical collaborators then they record the job ID number displayed in the top section of the screen. This ID number allows computer scientists at NCSA to retrieve the points associated with the correct registration session and complete the 3D volume reconstruction. Based on the set of points selected by medical

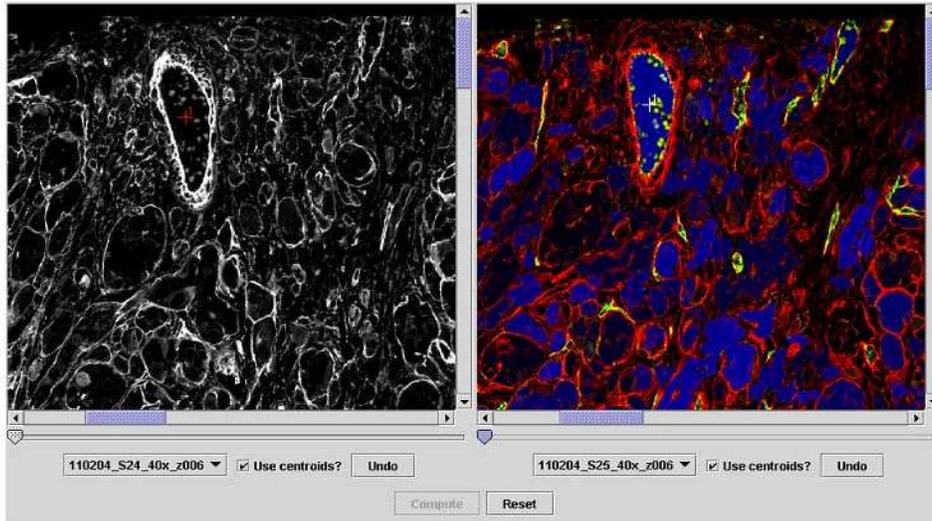


Figure 4.8: Selection of registration points using the centroid feature approach. Note that sub-areas the large images in Figure 4.7 are shown. (“+” marks represents the region centroids)

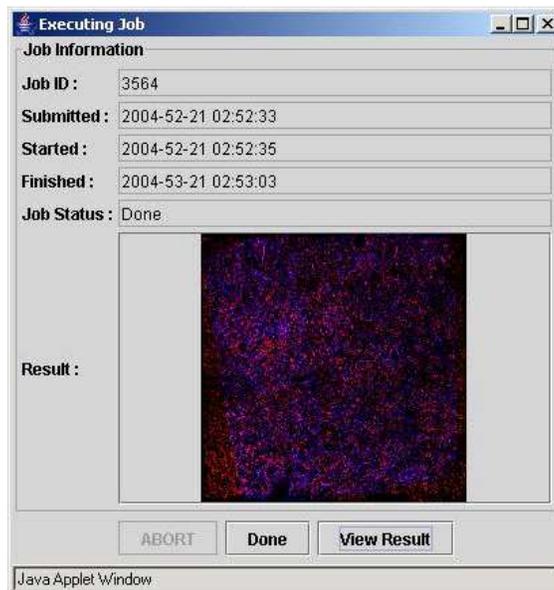


Figure 4.9: Job execution window after pressing the button “View Result”.

experts and saved at NCSA, multiple confocal laser scanning microscope sub-volume are transformed into a reference coordinate system at NCSA and the 3D volume reconstruction results are posted for downloading at an ftp site or a web site.

4.4 Matching Refinements

In this section, I introduce the processes to enhance the global transformation accuracy using intensity of global images and morphology of detected features. First, an optional intensity-based optimization can be performed to the results from alignment method using 2D information (e.g., frame selection). As discussed earlier, the intensity-based method can only be performed as an optimization step with, at least coarsely, known transformation due to (a) its registration abnormality from local minima caused by intensity heterogeneity, and (b) extremely high computation cost for affine transformation model. I propose an intensity-based optimization that adjusts pixel correspondence of selected pairs of locations in limited local neighborhood area. Although the matching refinement process is defined as an optional optimization process in 2D registration, it is a necessary step when performing 3D registration. After establishing correspondence, the alignment parameters (e.g., pixel coordinate correlation) are modified based on structural changes of cylindrical features (e.g., 3D trajectories). The main goal of the trajectory-based refinement is to preserve the smooth connectivity of morphological structures.

4.4.1 Optional intensity-based optimization for 2D alignment

In this step, there is an option to accommodate a typical visual alignment by using the normalized correlation approach [25]. The feature-based alignment performed so far could be refined based on intensity information the same way as a human would incorporate morphology and intensity cues during alignment. The intensity-based optional refinement

can be performed by following:

1. Select three pairs of matching coordinates (x_1, y_1) , (x_2, y_2) , and (x_3, y_3) from the *transforming image* (e.g., right image in Figure 4.16) based on the compactness measure in Section 4.5.
2. Define neighborhood adjustment vectors (u_1, v_1) , (u_2, v_2) , and (u_3, v_3) . For example, $(u_k, v_k) = ([-M/2, +M/2], [-M/2, +M/2])$, where M is the neighborhood window size, and $k = [1, 3]$.
3. Adjust three selected coordinates (x_1, y_1) , (x_2, y_2) , and (x_3, y_3) by adding (u_1, v_1) , (u_2, v_2) , and (u_3, v_3) .
4. Compute affine transformation parameters based on adjusted coordinates $(x_1 + u_1, y_1 + v_1)$, $(x_2 + u_2, y_2 + v_2)$, and $(x_3 + u_3, y_3 + v_3)$.
5. Transform the image chosen in step 1 based on the affine transformation from step 4.
6. Compute the normalized cross correlation coefficient between target image and transformed image in step 5.
7. Perform step 3-6 to find the optimal affine transformation by adjustment vectors (u_1, v_1) , (u_2, v_2) , and (u_3, v_3) that maximizes the normalized cross correlation coefficient.
8. Output the optimal affine transformation.

This process is very computationally expensive with the computational complexity of $O(\text{image size} \times \text{complexity of normalized correlation} \times M^6)$, where $M \times M$ is the size of neighborhood. One would select only a small spatial neighborhood of the three points to refine their locations based on the highest correlation value among all possible perturbations of point locations. In practice, the size of the point neighborhood determines the importance assigned to morphology and intensity information during the alignment.

4.4.2 Trajectory fusion for 3D alignment

In this section, I propose a trajectory-based sub-volume refinement method using trajectory fusion objectives, such as minimum discontinuity across adjacent sub-volumes and minimum residual of a polynomial fit to the corresponding trajectory points from adjacent sub-volumes. I approach the refinement problem by learning morphological characteristics (centroid trajectories) of structures inside of each CLSM sub-volume first, and then aligning sub-volumes by fusing the structures using their characteristics. The entire alignment process using 3D information is described in the following steps (shown in Figure 4.10):

1. Segment out salient medical structures in each sub-volume, e.g., cylindrical structures (3D segmentation).
2. Establish correspondences between structures from adjacent sub-volumes (feature matching).
3. Compute 2D centroids of the segmented structures in each frame (lateral plane) to obtain a discrete set of points from a 3D trajectory for each medical structure within a sub-volume, and use a polynomial model to estimate 3D trajectories from sets of points (trajectory computation).
4. Fuse a pair of corresponding medical structures by combining their 3D trajectories (trajectory fusion).
5. Compute a global 3D volume by transforming all sub-volumes based on the set of affine transformations derived from fused structures.

Fusing corresponding sets of trajectory points

The goal of this computation is (a) to compute volumetric segment trajectory points by estimating 2D centroids in each 2D frame of a sub-volume, and (b) to determine pairs of

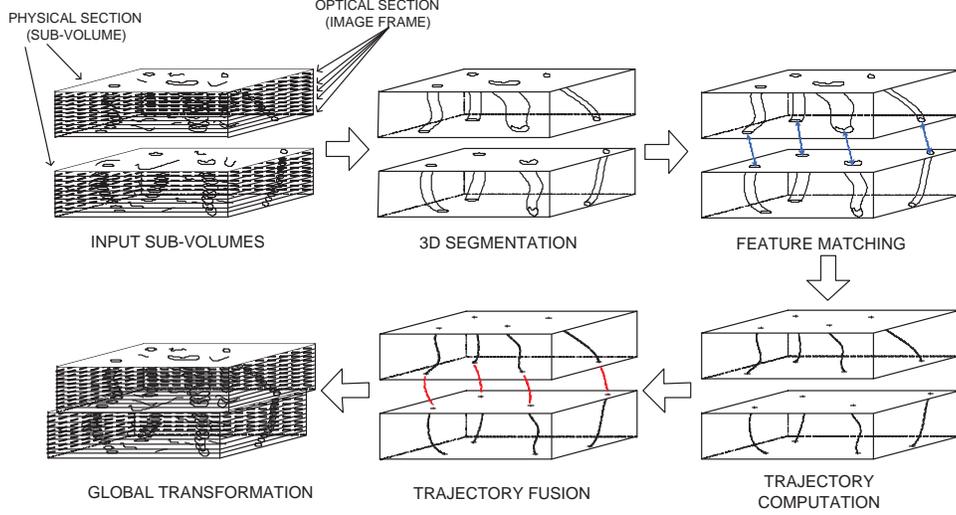


Figure 4.10: Overview of the trajectory-based alignment.

alignment control points from adjacent sub-volumes by fusing corresponding sets of trajectory points.

I define a set of 3D trajectory points derived for the i -th volumetric segment in the k -th sub-volume as $t_i^k = \{(x_1, y_1, z_1), (x_2, y_2, z_2), \dots, (x_m, y_m, z_m)\}$, where m is the number of frames (depth) in the sub-volume k . The lateral coordinate (x, y) is a centroid location in a frame, and the z coordinate represents the depth (frame index normalized with respect to lateral pixel coordinate system) in a volumetric segment.

This task is achieved by performing trajectory fusion under a set of optimization objectives. The goal of trajectory fusion is to determine pairs of *adjusted* control points to estimate the most accurate global transformation $\alpha : \mathcal{R}^2 \rightarrow \mathcal{R}^2$ applied to lateral planes (frames) of sub-volumes. First, given a set of matching pairs of depth-adjacent trajectories t_i^k and t_i^{k+1} ($i \in$ detected features), I compute $t_i^{*k+1} = \{(x_1 + u_i, y_1 + v_i, z_1), (x_2 + u_i, y_2 + v_i, z_2), \dots, (x_m + u_i, y_m + v_i, z_m)\}$ by trajectory fusion methods. Then, t_i^{*k+1} is considered as a set of modified (translated in lateral plane) trajectories that best fuse (connect) the corresponding features in t_i^k . Next, I compute the global affine transformation α by using

all features based on a least-squares fit defined as follows:

$$\alpha(\cdot) = \arg \min_{\alpha(\cdot)} \left(\sum_{i=1}^n [t_i^{*k+1} - \alpha(t_i^{k+1})]^2 \right) \quad (4.3)$$

where n is the number of matching volume features. Finally, I transform a set of images (frame) in the sub-volume $k + 1$ by the estimated α to create the final aligned sub-volume.

I propose two approaches for trajectory fusion denoted as fusion by extrapolation and by polynomial model fitting in followings:

(a) Fusion by extrapolation: The trajectory fusion by extrapolation is inspired by maximizing connectivity of matching trajectories. Assuming that there are some gaps between adjacent sub-volumes, I extrapolate a pair of points $(x_{m+1}^k, y_{m+1}^k, z_{m+1}^k) \in t_i^k$ and $(x_0^{k+1}, y_0^{k+1}, z_0^{k+1}) \in t_i^{k+1}$ where $z_m^k < z_{m+1}^k = z_0^{k+1} < z_1^{k+1}$. Then, I compute the translational offset (u_i, v_i) of feature i as following:

$$(u_i, v_i) = (x_0^{k+1} - x_{m+1}^k, y_0^{k+1} - y_{m+1}^k) \quad (4.4)$$

To generate the pair of extrapolated points $(x_{m+1}^k, y_{m+1}^k, z_{m+1}^k)$ and $(x_0^{k+1}, y_0^{k+1}, z_0^{k+1})$, I used different methods such as an end-point duplication method and a polynomial fitting method with three different degrees. First, the end-point duplication is performed by replicating (x, y) coordinate of the adjacent end trajectory points such as followings:

$$\begin{aligned} (x_{m+1}^k, y_{m+1}^k, z_{m+1}^k) &= (x_m^k, y_m^k, z_m^k + \Delta_{gap}) \\ (x_0^{k+1}, y_0^{k+1}, z_0^{k+1}) &= (x_1^{k+1}, y_1^{k+1}, z_1^{k+1} - \Delta_{gap}) \end{aligned} \quad (4.5)$$

An example of trajectory fusion by end-point duplication is illustrated in Figure 4.11.

Second, the polynomial-based extrapolation is done by estimating degree γ polynomial functions $f_i^{k(\gamma)}$ and $f_i^{k+1(\gamma)}$ based on t_i^k and t_i^{k+1} . To compute a 3D polynomial curve, I

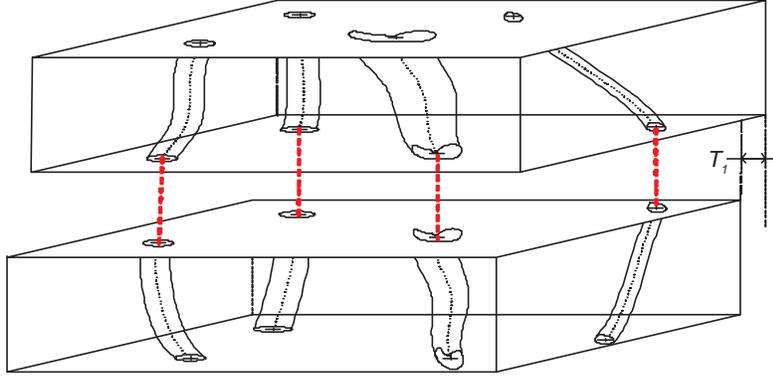


Figure 4.11: Trajectory fusion by connecting end points

estimate a function $f_i^{k(\gamma)}$ as followings:

$$\begin{aligned} f_{ix}^{k(\gamma)}(z) &= a_0 + a_1 z + \dots + a_\gamma z^\gamma \\ f_{iy}^{k(\gamma)}(z) &= b_0 + b_1 z + \dots + b_\gamma z^\gamma \end{aligned} \quad (4.6)$$

where

$$\begin{bmatrix} m & \sum_{j=1}^m z_j & \cdots & \sum_{j=1}^m z_j^\gamma \\ \sum_{j=1}^m z_j & \sum_{j=1}^m z_j^2 & \cdots & \sum_{j=1}^m z_j^{\gamma+1} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{j=1}^m z_j^\gamma & \sum_{j=1}^m z_j^{\gamma+1} & \cdots & \sum_{j=1}^m z_j^{2\gamma} \end{bmatrix} \begin{bmatrix} a_0 & b_0 \\ a_1 & b_1 \\ \vdots & \vdots \\ a_\gamma & b_\gamma \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^m x_j & \sum_{j=1}^m y_j \\ \sum_{j=1}^m z_j x_j & \sum_{j=1}^m z_j y_j \\ \vdots & \vdots \\ \sum_{j=1}^m z_j^\gamma x_j & \sum_{j=1}^m z_j^\gamma y_j \end{bmatrix}. \text{ Next,}$$

a pair of extrapolated points is generated as followings:

$$\begin{aligned} (x_{m+1}^k, y_{m+1}^k, z_{m+1}^k) &= (f_{ix}^{k(\gamma)}(z_m^k + \Delta_{gap}), f_{iy}^{k(\gamma)}(z_m^k + \Delta_{gap}), z_m^k + \Delta_{gap}) \\ (x_0^{k+1}, y_0^{k+1}, z_0^{k+1}) &= (f_{ix}^{k+1(\gamma)}(z_1^{k+1} - \Delta_{gap}), f_{iy}^{k+1(\gamma)}(z_1^{k+1} - \Delta_{gap}), z_1^{k+1} - \Delta_{gap}) \end{aligned} \quad (4.7)$$

(b) Fusion by model fitting: The assumption behind the fusion by model fitting is that relatively short trajectory of a blood vessel follows a polynomial curve with degree of less than three (cubic). Although the exact model of the trajectories cannot be defined, I show that maximum three degrees of polynomial fit to the real data quite well based on the experimental results.

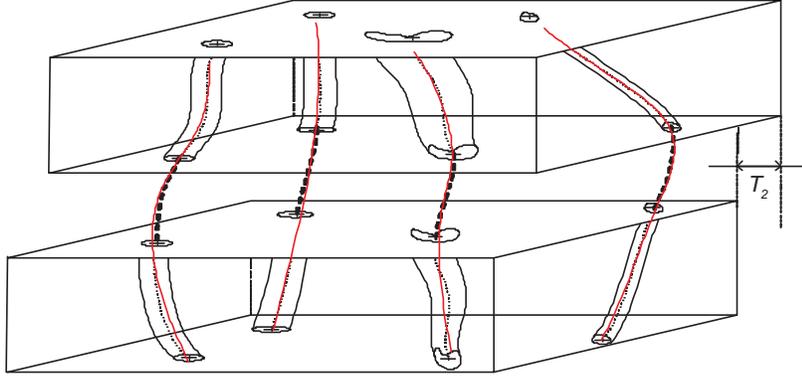


Figure 4.12: Trajectory fusion by residual minimization of a fitted model.

First, I define $r^{(\gamma)}(t_i^k)$ as a residual after fitting γ -degree polynomial function $f_i^{k(\gamma)}$ to t_i^k as following:

$$r^{(\gamma)}(t_i^k) = (r_x^{(\gamma)}(t_i^k) + r_y^{(\gamma)}(t_i^k))/2 \quad (4.8)$$

where $r_x^{(\gamma)}(t_i^k) = \sqrt{\frac{1}{m} \sum_{j=0}^m [x_{ij} - f_{ix}^{k(\gamma)}(z_i)]^2}$ and $r_y^{(\gamma)}(t_i^k) = \sqrt{\frac{1}{m} \sum_{j=0}^m [y_{ij} - f_{iy}^{k(\gamma)}(z_i)]^2}$.

To compute the residual of fused trajectory from a pair of trajectories, I introduce a binary operator \oplus_{Δ_k} which merges a pair of trajectories between the sub-volume k and $k+1$ with a physical gap Δ_k , e.g., missing frames. Δ_k is assumed to be a constant for all sub-volumes (typically no more than 3 frame length). Then residual of the fused trajectory can be computed by $r^{(\gamma)}(t_i^k \oplus_{\Delta_k} t_i^{k+1})$.

Next, for each pair of matching trajectories i , I search the translational offset (u_i, v_i) for t_i^{*k+1} by minimizing the residual $r^{(\gamma)}(t_i^k \oplus_{\Delta_k} t_i^{*k+1})$ as following:

$$(u_i, v_i) = \arg \min_{(u_i, v_i) \in \text{neighborhood}} (r^{(\gamma)}(t_i^k \oplus_{\Delta_k} t_i^{*k+1})) \quad (4.9)$$

An example of trajectory fusion by residual minimization is shown in Figure 4.12.

4.5 Global Transformation Estimation and Sub-volume Transformation

Rigid transformation with only translation and rotation has been one of the most popular lower order transformation models designed for rigid structures such as bones. However, in cellular imaging, higher order (elastic) model or local transformation are preferred to achieve smooth transition of image structures across slides. However, the difficulty with higher order models is (a) in robust parameter estimation due to intensity variation (noise) and deformation exceeding the order of the chosen model, or (b) in bifurcation (appearing and disappearing structures). Although non-rigid optimization can be applied for only local features after a global alignment, I limited our transformation model as an affine because the transformations using higher order models could lead to erroneous alignment due to the well-known leaning tower problem [104]. As a result, it could ultimately distort the 3D anatomical structures (features) by matching accurately small regions while significantly distorting other regions.

Considering the medical specimens of my interest, I chose an affine transformation for modeling cross section distortions and expected to detect only a small amount of scale and shear deformations. In this case, the paraffin-embedded tonsil tissue represents a non-rigid structure and has to include deformation like shear due to tissue slicing. The use of affine model for image transformation accommodates small amounts of scale and shear that are inevitable during material preparation.

Given the affine transformation model $\alpha : \mathcal{R}^2 \rightarrow \mathcal{R}^2$, the image alignment can be performed by selecting at least three pairs of corresponding points and computing six affine transformation parameters shown below.

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \quad (4.10)$$

The values $(x', y') = \alpha(x, y)$ are the transformed coordinates of the original image coordinates (x, y) by affine transformation $\alpha(\cdot)$. The four parameters, a_{00} , a_{10} , a_{01} , and a_{11} , represent a 2 by 2 matrix compensating for scale, rotation and shear distortions in the final image. The two parameters, t_x and t_y , represent a 2D vector of translation.

Based on the model assumption from CLSM imaging, the limited range of shear and scale values can be verified by scrutinizing entries in the affine transformation matrix (a_{01} and a_{10} entries for shear, and a_{00} and a_{11} matrix entries for scale). This approach enabled me to apply the affine transformation carefully in order to avoid any unreasonable registration artifacts. As in any experimental setting, the registration decision about transformation models is critical to automating 3D volume reconstruction. If other data sets have to be modeled with more complex deformations (transformation models), then the steps of feature selection, matching and registration parameter estimation should be re-visited.

From the image alignment accuracy viewpoint, corresponding pairs of segment centroids should be well spatially distributed in each image, and should not be collinear. If points are close to be collinear then the affine transformation parameters cannot be uniquely derived from a set of linear equations (more unknowns than the number of equations), which leads to large alignment errors. If points are locally clustered and do not cover an entire image spatially then the affine transformation is accurate only in the proximity of the selected points. However, the affine transformation inaccuracy increases with the distance from the selected points, which leads to large alignment errors because a registration error metric takes into account errors across the entire image area. In order to assess the pairs of matched centroid points in terms of their distribution and collinear arrangement, I have designed a compactness measure. It is defined as a ratio of the entire image area ($Area_{IMAGE}$) divided by the largest triangular area formed from the points ($Area_{TRIANGLE}$). The measure is defined mathematically in following Equation.

$$Compactness\ Measure = \frac{Area_{IMAGE}}{Area_{TRIANGLE}} \quad (4.11)$$

4.6 Results

This section provides the sub-volume alignment results in feature matching, matching refinement, and global transformation estimation and sub-volume transformation.

4.6.1 Feature matching

First, I outline the methodology for assessing upper error bounds of automatic, semi-automatic, and manual 3D volume reconstruction techniques. The experimental variables include the type of registration methods (automatic, semi-automatic and manual) and the type of human subjects (experts and novices) performing registration. Human subjects were labeled as experts if they had the knowledge about CLSM imaging, imaged specimen and its anatomical/structural properties, and/or principles of the affine transformation based registration algorithm. This type of knowledge was critical for establishing feature correspondences and obtaining accurate registration results.

The result demonstrates benefits of automation for 3D volume reconstruction in terms of achieved accuracy, consistency of results, and performance time. I also outline the limitations of fully automated and manual 3D volume reconstruction systems, and describe related automation challenges. The result shows that given computational resources and repetitive experimental data, the automated alignment provides more accurate and consistent results than a manual alignment approach. With the proposed approach, the automation will reduce the alignment execution time and cost in the future, as the cost of more powerful computers goes continuously down.

Alignment accuracy evaluation methodology

Similarly to the case of image mosaicking, it can be simulated by creating a pair of misaligned images by applying a known affine transformation to a duplicate of an image and presenting

the original and transformed images to a user for accuracy evaluation purposes. However, this evaluation setup would not simulate the real problem of image alignment where two real cross sections might have missing or new or warped structures with a priori unknown intensity variations. Thus, I chose to establish the reference image and its corresponding affine transformation parameters in the following way.

1. A stack of CLSM images (optical sections co-registered along z-axis) is acquired from a physical section.
2. Multiple stacks of CLSM images are aligned by a manual alignment method, and the representative of all resulting affine transformations is recorded, e.g., maximum translation, rotation and shear.
3. A pair of misaligned images is constructed for accuracy evaluations by taking a pair of images (apart from relatively large number of frames to simulate morphological deformation) along the z-axis of *one* CLSM physical section and applying the representative affine transformation (recorded in the second step) to the second image.

The first and the transformed images become the evaluation images with the known ground truth affine transformation $\alpha^{GT}(\cdot)$. All pixel coordinates of the transformed (ground truth) image $\mathbf{P}^{GT} = \{p_1^{gt}, p_2^{gt}, \dots, p_n^{gt}\}$ are then defined by the affine transformation $\alpha^{GT} : p_i \rightarrow p_i^{gt}$. Based on user's registration input, an affine transformation $\alpha^{USR}(\cdot)$ is estimated. I denote the corresponding set of transformed pixel coordinates as $\mathbf{P}^{USR} = \{p_1^{usr}, p_2^{usr}, \dots, p_n^{usr}\}$, where $\alpha^{USR} : p_i \rightarrow p_i^{usr}$. The final image alignment registration error E_{affine} is then calculated as an average Euclidian error distance over all pixels coordinates according to the following equation, where m is the number of transformed pixels.

$$E_{affine} = \frac{1}{m} \sum_{i=1}^m \sqrt{(p_{ix}^{gt} - p_{ix}^{usr})^2 + (p_{iy}^{gt} - p_{iy}^{usr})^2} \quad (4.12)$$

Once again, with the smaller the error E_{affine} , the better image alignment accuracy is achieved.

The proposed image alignment evaluation methodology utilizes confocal imaging to obtain required image frames, and empirically observed affine distortions to prepare test alignment data as close to real data as possible. The justification for choosing the alignment error metric E_{affine} is twofold. First, similar to the explanation provided for the choice of the mosaicking error metric, an error metric based on pixel locations seems more appropriate than a metric based on intensity comparisons due to CLSM intensity variations. Second, it would not be fair to compute differences of affine transformation parameters since they represent a mix of distortions (translation, rotation, scale and shear). Euclidean distances over the registered area reflect the degree of misalignment. It would be possible to consider a metric that would include the spatial mismatch only over the set pixels that are above a certain intensity threshold. However, I decided to avoid introducing a threshold parameter into our evaluation metric due to different unknown intensity ranges and distributions of a pair of compared images.

Performance evaluation in different automation levels

I present an accuracy evaluation of manual, semi-automated, and fully automated registration technique for 3D volume reconstruction from fluorescent confocal laser scanning microscope (CLSM) imagery. The presented region-based (semi-automated and fully automated) method is designed based on the observations that (a) an accurate point selection is much harder than an accurate region (segment) selection for a human, (b) a centroid selection of any region is less accurate by a human than by a computer, and (c) registration based on structural shape of a region rather than based on intensity-defined point is more robust to noise and to morphological deformation of features across stacks. I applied the method to image alignment registration steps and evaluated its performance with 20 human subjects

on CLSM images with stained blood vessels. The experimental evaluation showed significant benefits of automation for 3D volume reconstruction in terms of achieved accuracy, consistency of results and performance time. In addition, the results indicate that the differences between registration accuracy obtained by experts and by novices disappear with the proposed semi-automatic registration technique while the absolute registration accuracy increases.

Manual vs. semi-automated alignment method: For the image alignment experiments, I used the same user interfaces for selecting multiple points and regions as shown in Figure 3.9 for test images in Figure 4.13. It was recommended that human subjects select at least three points or regions, in such a way that they would be well spatially distributed in each image but would not be collinear. If points are close to be collinear then the affine transformation parameters cannot be uniquely derived from a set of linear equations (more unknowns than the number of equations), which leads to large alignment errors. If points are locally clustered and do not cover an entire image spatially then the affine transformation is very accurate only in the proximity of the selected points. However, the affine transformation inaccuracy increases with the distance from the selected points, which leads to large alignment error since the error metric takes into account errors across the entire image area. In order to assess the points selected by a user in terms of their distribution and collinear arrangement, I showed a compactness measure defined in Section 4.6.

I observed large alignment error when human subjects selected almost collinear points or locally clustered points regardless of our recommendations. Figure 4.14 shows the relationship between compactness and alignment error modeled with a linear fit. I used three different pairs of adjacent physical sections for alignment study, and the error results of all experiments as a function of human subject trials are shown in Figure 4.15 and summarized in Table 4.2. The t-test values for comparing the pixel-based and feature-based mosaicking

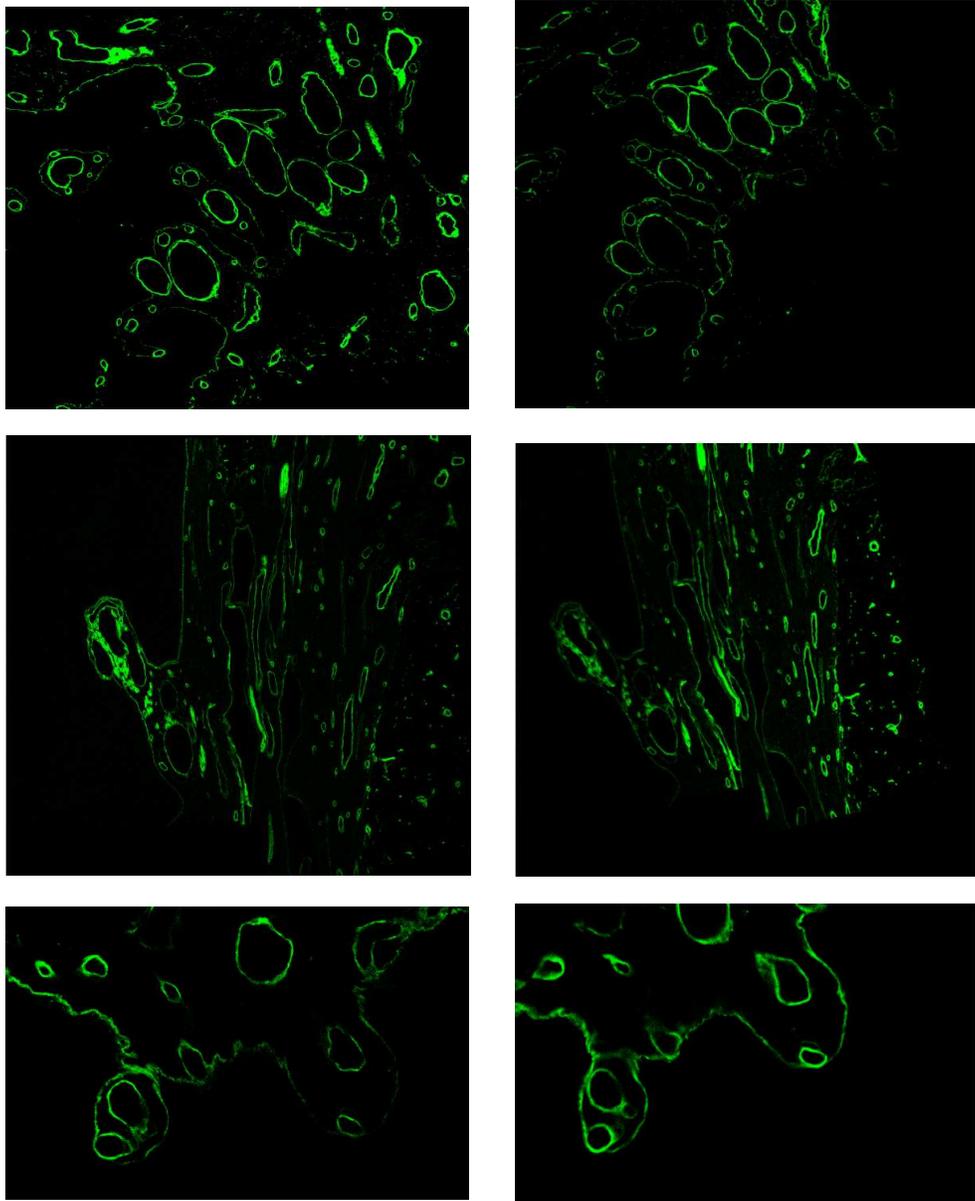


Figure 4.13: Three pairs of image examples used for alignment evaluation

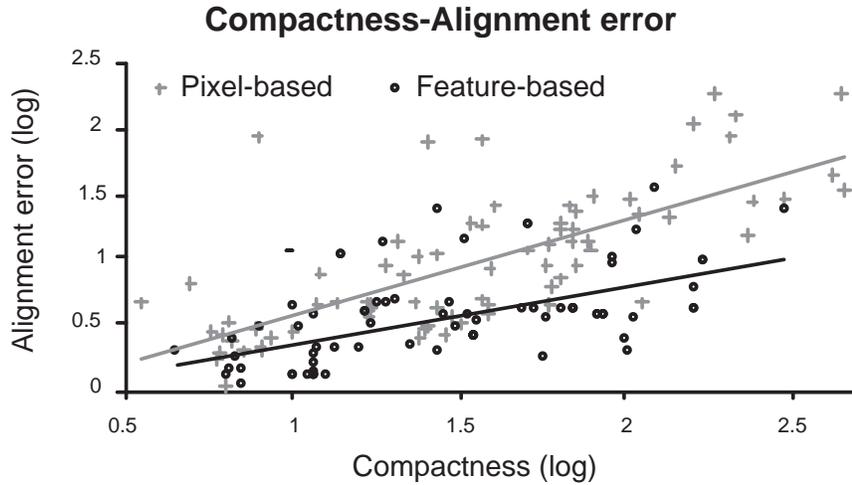


Figure 4.14: Illustration of a strong correlation between the compactness measure and the alignment error.

Table 4.2: A summary of image alignment

| Error (pixels) | manual (pixel-based) | | semi-automated (feature-based) | |
|---------------------------------|----------------------|--------|--------------------------------|--------|
| Expertise | expert | novice | expert | novice |
| Average | 17.32 | 27.98 | 4.85 | 5.83 |
| Standard deviation | 27.12 | 43.28 | 5.63 | 6.71 |
| Total average | 22.28 | | 5.28 | |
| Total std. deviation | 35.74 | | 6.11 | |
| Upper bound (99.73% confidence) | 129.5 | | 23.61 | |

are shown in Table 4.3.

The image alignment results lead me to the following conclusions. First, manual (pixel-based) image alignment is less accurate and less consistent (large standard deviation) than the semi-automatic (feature-based) alignment. Based on the t-test result in Table 4.3, the null hypothesis (no improvement) can be rejected with 99.9% confidence. Second, selection of (a) collinear features, or (b) spatially dense points or regions, can have a detrimental effect on alignment accuracy. Third, experts achieved higher average alignment accuracy than novices with both methods. Finally, the difference in alignment errors between experts and novices using the pixel-based method is significantly reduced when human subjects start using the feature-based alignment method. I should also mention that the majority of human

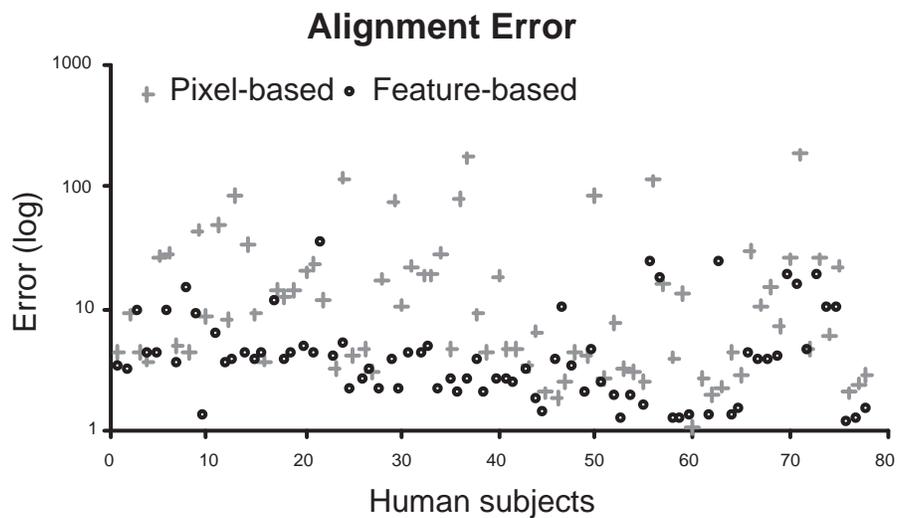


Figure 4.15: Alignment errors for all human trials including pixel-based (manual) and feature-based (semi-automatic) alignment.

Table 4.3: The paired t-test result for errors of the pixel-based and the feature-based methods in Table 4.2.

| | manual vs. semi-automated |
|--------------------|---------------------------|
| Degrees of freedom | 77 |
| t value | 4.109 |
| p value | 0.999 |

Table 4.4: A summary of automatic image alignment experiments without optimization.

| | Automatic without optimization |
|------------------------|--------------------------------|
| Average error [pixels] | 3.0949 |
| Standard Deviation | 2.0691 |

subjects selected only three points or regions for aligning two images. To demonstrate the effect of the number of selected points on the registration accuracy, I computed the accuracy by using all matching pairs of features detected by segmentation (27, 21, and 4 pairs for each test in Figure 4.13). The estimated affine transformation results in 1.21, 1.12, and 2.54 pixel error distances for each test data, respectively. The average pixel error distance is equal to 1.62 pixels and the standard deviation is 0.79. This result indicates that (a) more well-matched points lead to more accurate alignment, and (b) instructing human subjects to choose the maximum number of the features detected by segmentation would lead to higher alignment accuracy.

Fully automated alignment method Unlike manual and semi-automated alignment method, fully-automated method does not required a human intervention during the alignment process, and hence the number of selected features is not limited to three or four (that is, usually selected by human subjects) but all detected features are used to search for the best alignment by best feature selection using the compactness measure. Table 4.4 contains results obtained using automated alignment from the test images in Figure 4.13. The automated alignment leads to the same result every time the algorithm is executed with the same data therefore the standard deviation is equal to zero.

Figure 4.16 illustrates the correspondences found for the features by considering centroid and area using the fully automated method described in Section 4.3.3. In Figure 4.17, the circles denote three pairs of centroid points that were selected from the set of all pairs shown in Figure 4.16 according to the proposed compactness measure described in Section 4.6.

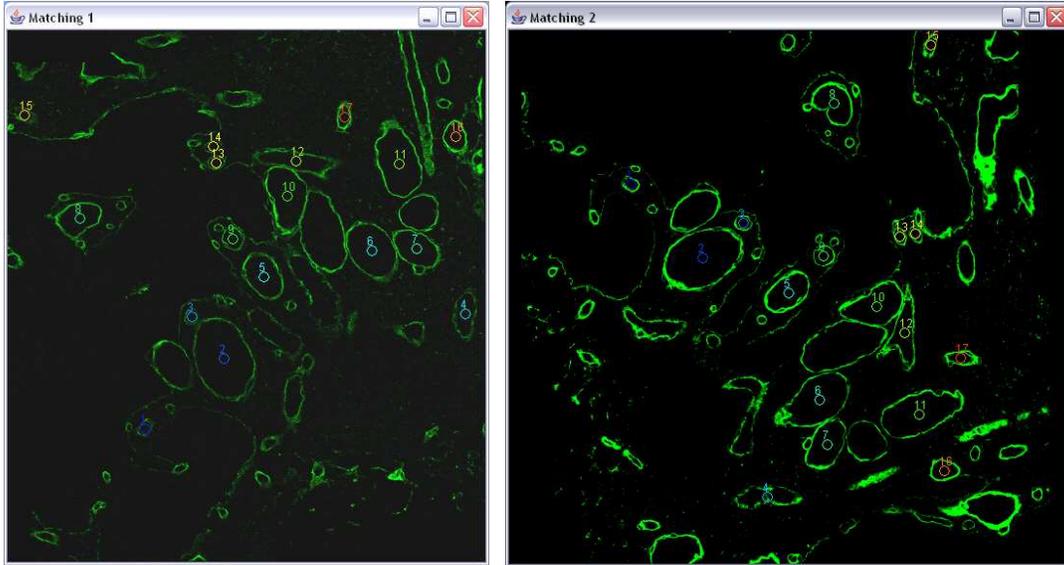


Figure 4.16: The correspondence outcome from two phases for the segments. The left and right images are to be aligned. Overlays illustrate established correspondences between segments that are labeled from 1 through 17. The centroid locations of segments are sorted based on the correspondence error from the smallest error to the largest error.

Table 4.5: Datasets

| | Dataset 1 | Dataset 2 |
|--------------------------------------|---------------|---------------|
| Tile Size (pixels) | 512 x 512 | 512 x 512 |
| Tiles per frame (row x column) | 7 x 3 | 5 x 5 |
| Frame size after mosaicking (pixels) | 2746 x 1170 | 2300 x 2300 |
| Frames in a subvolume | 13 | 13 |
| Subvolumes | 16 | 10 |
| Channels | 2 | 2 |
| Memory Size | 1336.5 Mbytes | 1375.4 Mbytes |

Performance evaluation for web-service based registration: I performed two user studies to see how well the web-service based registration worked. NCSA received from UIC two datasets. Table 4.5 gives some information about the size of the datasets. The data delivered to NCSA was just the tiles for each dataset. Some preprocessing was needed before the datasets could be used in the application described previously.

The first preprocessing step done was to mosaic the tiles into frames. Figure 4.18 shows

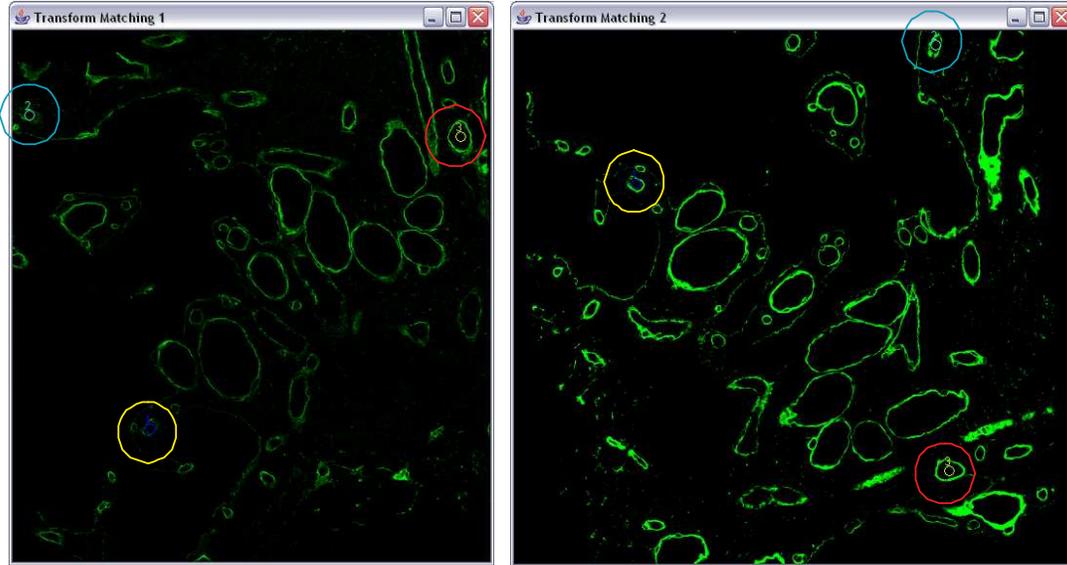


Figure 4.17: The result of automated feature selection after processed to establish segment correspondences shown in Figure 4.16. The circles of different colors represent three pairs of centroids selected automatically according to the compactness measure defined in Section 4.5.

a single frame selected from the each dataset after the mosaicking. The fully automated mosaicking preprocessing step took one hour for dataset 1 and four hours for dataset 2. When scanning the sample, each tile was scanned at multiple depths resulting in a subvolume. The next step is to combine the frames into the subvolumes again. Next the best frame of each subvolume is selected and segmented [85, 88] (4 hours for each dataset). The information about the segments and the centroid for each segment is packaged with the frame for use in the web service application (1 hour for each dataset).

The URL of the web service application is send back to UIC for the medical expert to do the alignment of frames. The medical expert will look at the frames and find corresponding points in each successive frame and view the results. This process took 140 minutes for dataset 2 and 129 minutes for dataset 1. Even though dataset 1 contains more subvolumes, and hence more frames that need to be aligned, this dataset was finished faster than dataset 2. This could be explained by the fact that (a) the dataset 1 contains more visually salient features than the dataset 2 and (b) dataset 2 was the first dataset processed by the medical

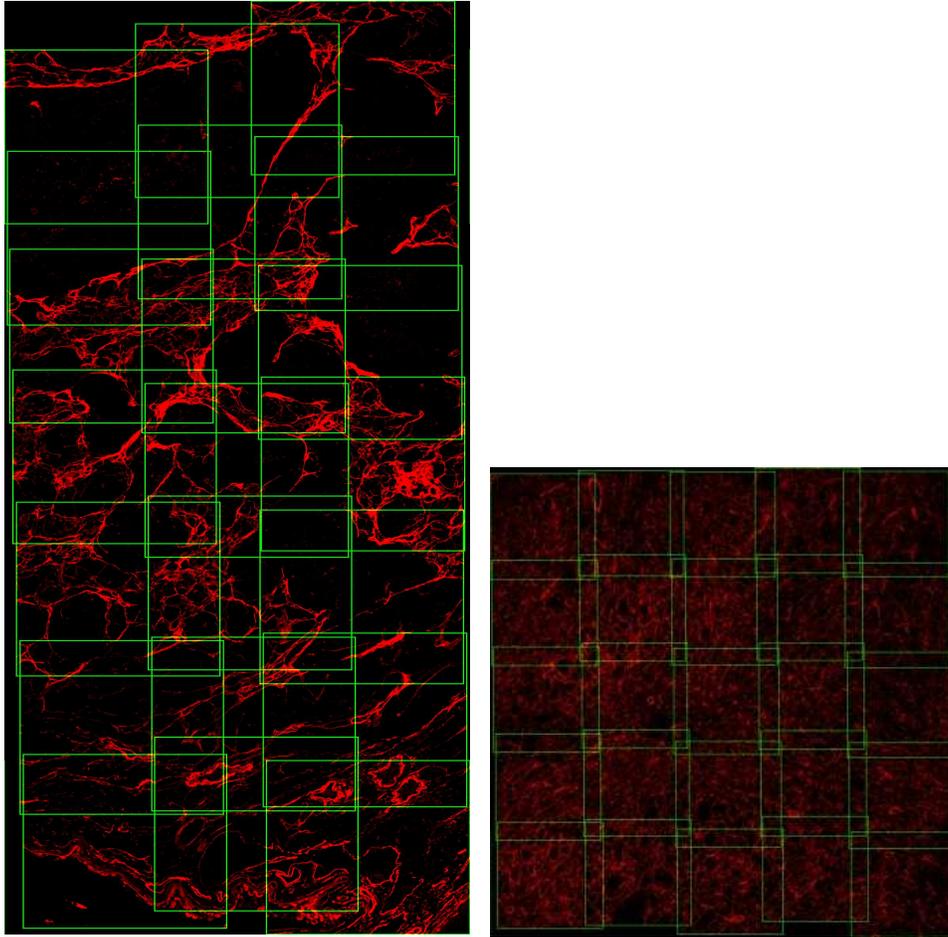


Figure 4.18: The images after mosaicking. The edges of each tile are visualized using green lines. On the left is a frame from dataset 1 and on the right is a frame from dataset 2.

expert and part of the extra time could be attributed to the learning process of the new tool.

After selecting the registration points an email message was send to NCSA with the job numbers of the best alignments. These jobs are then used to extract the points selected by the medical expert and used to create the final 3D volume. Dataset 1 resulted in a final 3D volume of $2746 \times 1170 \times 208$ with 2 colors for each pixel in this 3D volume. Dataset 2 resulted in a 3D volume of $2300 \times 2300 \times 130$ with again 2 colors per channel. Figure 4.19 show what the final 3D volume would look like for Dataset 1.

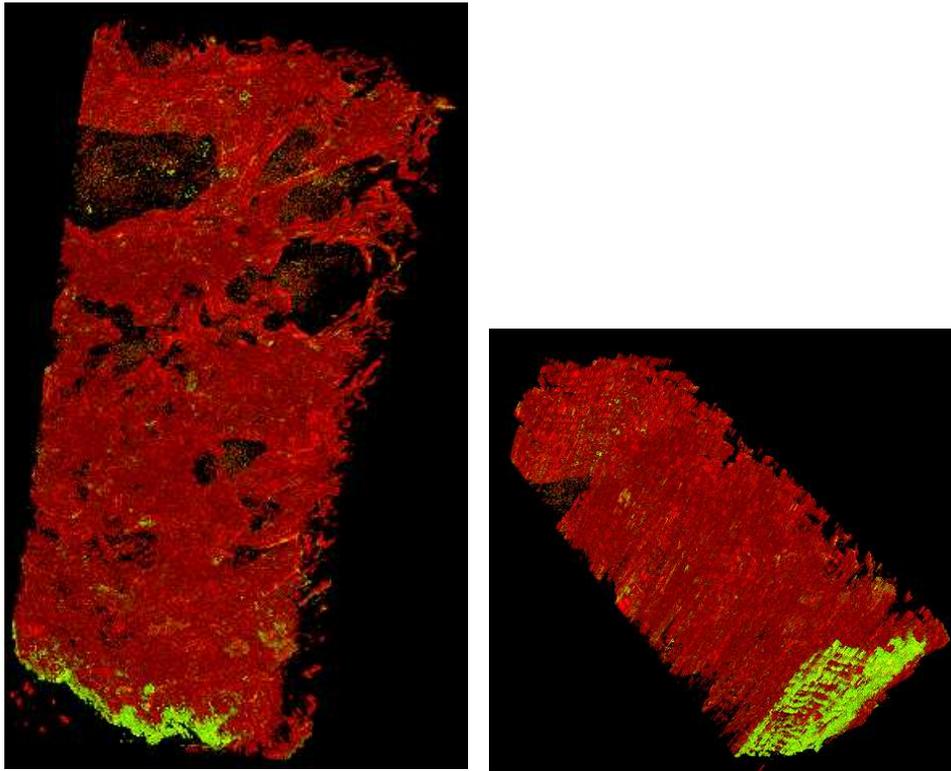


Figure 4.19: 3D view of the reconstructed 3D volume using web service.

Table 4.6: A summary of automatic image alignment experiments with and without optimization.

| Error [pixels] | Automatic with and without optimization | | | | | | | |
|--------------------|---|--------|---------|--------|---------|--------|---------|--------|
| Test images | Image 1 | | Image 2 | | Image 3 | | Average | |
| Optimization | no | yes | no | yes | no | yes | no | yes |
| Average | 1.2839 | 1.4009 | 2.6508 | 2.1993 | 5.3500 | 2.2280 | 3.0949 | 1.9427 |
| Standard Deviation | 0 | 0 | 0 | 0 | 0 | 0 | 2.0691 | 0.4694 |

4.6.2 Matching refinements

Intensity-based optimization

I performed the optional intensity-based optimization step for comparison using the fully-automated results in section 4.7.1. Table 4.6 contains results obtained using automated alignment without and with the optional optimization step.

According to Table 4.6, the optimization step improved alignment accuracy for two out of the three test image pairs (test image pairs 2 and 3). The optimization step was conducted by choosing the correlation spatial neighborhood size $M = 5$ to control (a) the amount of needed computation and (b) the maximum range of shear and scale. In our experiments, all processing steps except the optimization took several seconds on a regular desktop computer. The optimization step on its own took on average 19.66 minutes (16, 21 and 22 minutes) on a single processor machine (3.0GHz), for the three pairs of test images of the pixel dimensions 579×549 , 584×649 , and 789×512 . The offsets of the three centroids due to optimization were $(+1, +1)$; $(0, 0)$; $(+1, 0)$ for image 1, $(-2, 2)$; $(1, -2)$; $(1, 0)$ for image 2, and $(2, -1)$; $(2, 2)$; $(2, 2)$ for image 3. The normalized correlation coefficient increased with optimization by 7.8%, 21.3% and 12.4% for the test image pairs 1, 2 and 3 respectively.

Trajectory fusion

Evaluation metrics for alignment accuracy : In this section I define two evaluation metrics to be used after applying the final global transformation such as overall residual and discontinuity.

Global residual: The global residual can be calculated by following:

$$R_k^{(\gamma)} = \frac{1}{n} \sum_{i=1}^n r_i^{(\gamma)}(t_i^k \oplus_{\Delta_k} \alpha(t_i^{k+1})) \quad (4.13)$$

where $r_i^{(\gamma)}(t_i^k \oplus_{\Delta_k} \alpha(t_i^{k+1}))$ is a residual of i -th feature in sub-volume k and $k + 1$, and n is the number of matching trajectories (features). I assume that sub-volume correspondence is known, for example, t_i^k and t_i^{k+1} are assumed to be a pair of matching volume features.

Discontinuity: I define a trajectory discontinuity $d^{(\gamma)}(t_i^k, t_i^{k+1})$ as a 2D Euclidian distance (in lateral plane) of the pair points that connects the adjacent trajectories.

$$d^{(\gamma)}(t_i^k, t_i^{k+1}) = \left\| e_{last}^{(\gamma)}(t_i^k) - e_{first}^{(\gamma)}(t_i^{k+1}) \right\| \quad (4.14)$$

where $e_{last}^{(\gamma)}(t_i^k)$ and $e_{first}^{(\gamma)}(t_i^{k+1})$ are the adjacent (the last of t_i^k and the first of t_i^{k+1}) 2D lateral coordinates on the polynomial after fitting with degree γ . These coordinates are acquired by evaluating the same axial z -coordinate with two different polynomial functions $f_i^{k(\gamma)}$ and $f_i^{k+1(\gamma)}$ estimated from t_i^k and t_i^{k+1} . To incorporate the sub-volume gap Δ_k for end point evaluation, the axial z -coordinate need to be adjusted by $+(\Delta_k + 1) \times 0.5$ and $-(\Delta_k + 1) \times 0.5$ for t^k and t^{k+1} respectively. For our notational convenience, note that $\gamma = 0$ implies that I compute the 2D lateral Euclidian distance of physical end points (no polynomial evaluation).

Finally, the overall discontinuity $D_k^{(\gamma)}$ of sub-volume k and $k + 1$ is defined as follows:

$$D_k^{(\gamma)} = \frac{1}{n} \sum_{i=1}^n d^{(\gamma)}(t_i^k, t_i^{k+1}) \quad (4.15)$$

Trajectory fusion results: I evaluated the global residual and discontinuity by considering three variables for trajectory based 3D volume reconstruction from CLSM: (1) polynomial degree γ of centroid trajectories along axial direction (z -axis of the sub-volume), (2) assumed gaps between adjacent sub-volumes Δ_k , (3) trajectory fusion approach (extrapolation versus

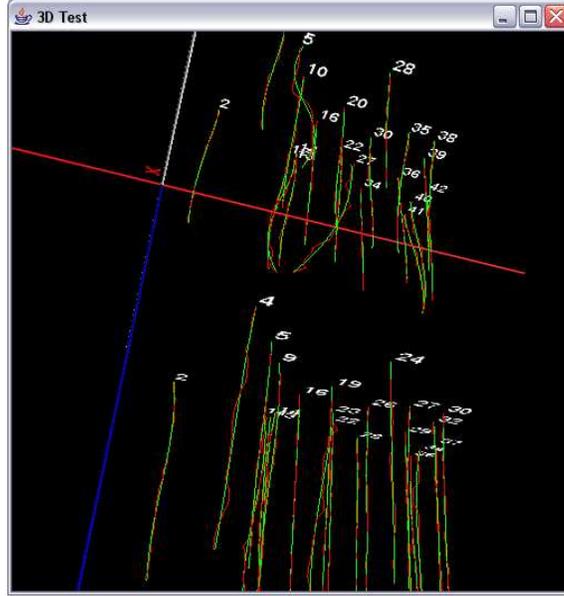


Figure 4.20: A pair of sets of (exaggerated) trajectories from adjacent sub-volumes. Numbers indicate the label of matching trajectories. Red line segments indicate centroids, and green curves indicate a polynomial function with degree 3.

residual minimization).

After acquiring the set of volume features after feature detection on volume segments as shown in Section 3.7.4, feature correspondence are established based on feature matching described in Section 4.4. Next, I compute trajectories $t_i^{sub-volume1}$ and $t_i^{sub-volume2}$ ($i \in$ matching features) based on the series of centroids from 2D optical sections in each volume feature. An example of a pair of 3D trajectories is shown in Figure 4.20.

Figure 4.21 shows the result of discontinuities from trajectory fusion using extrapolation method and residual minimization method in linear, quadratic, and cubic models with respect to different sub-volume gaps. Based on the result, I conclude the followings; (1) the extrapolation method achieves lower (better) discontinuity than residual minimization method regardless of the polynomial model complexity; (2) discontinuity is not necessarily related to the size of the physical gaps; and (3) using same method, discontinuity is relatively insensitive to the degree of the polynomial model complexity (e.g., curves in the graph using the same method remain close to each other).

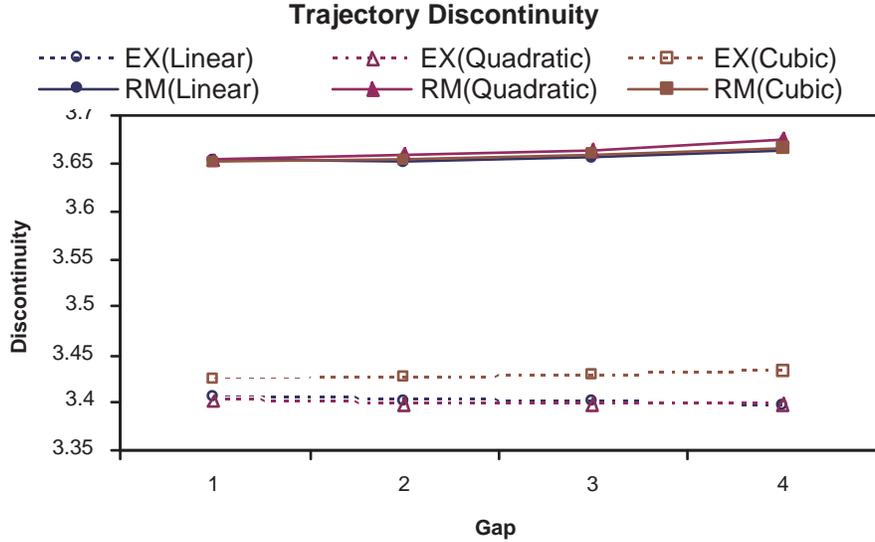


Figure 4.21: Trajectory discontinuity as function of physical gaps between adjacent sub-volumes: EX() and RM() refer to extrapolation method and residual minimization method respectively. Note that extrapolation method is plotted as dotted lines, and residual minimization method is plotted by solid lines.

Figure 4.22 shows the result of trajectory residuals after fusing depth-adjacent trajectory segments using extrapolation method and residual minimization method. From the result, I conclude that; (1) using same degree of polynomial model complexity, residual minimization method always achieves lower (better) residual than extrapolation method; (2) all residuals (all method and degrees of polynomial model complexity) tend to decrease about the size of physical gaps of the adjacent sub-volumes; and (3) the degree of polynomial model complexity is relatively more sensitive than the method.

Figure 4.23, Figure 4.24, and Figure 4.25 show a 3D visualization without optimization, with optimization using extrapolation method, and with optimization using residual minimization method respectively. The figures on the right shows the side views of regions that demonstrate structural discontinuity of volume segments in a sub-region marked in the images on the left. It shows that feature segments has high discontinuity when trajectory-based refinement has not been performed (see Figure 4.23 (b)). One could notice that refinement us-

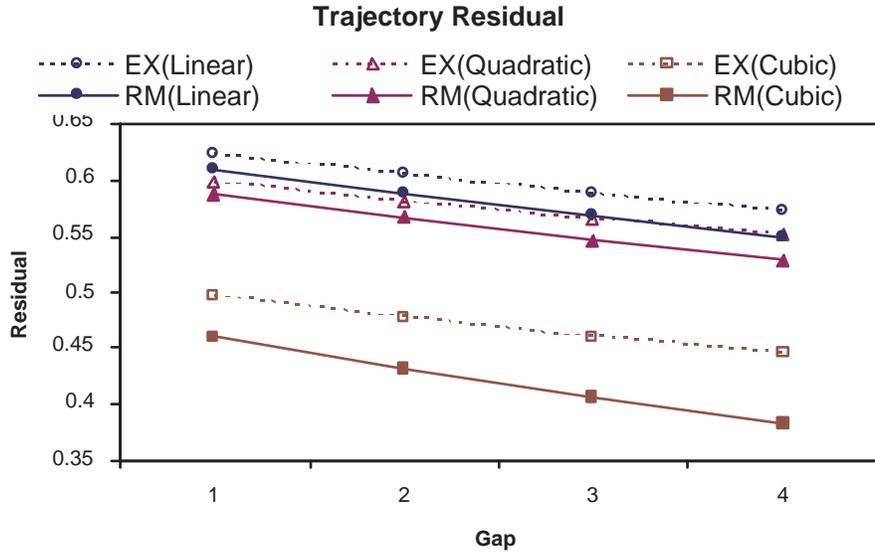


Figure 4.22: Trajectory residual as function of physical gaps between adjacent sub-volumes: EX() and RM() refer to extrapolation method and residual minimization method respectively. Note that extrapolation method is plotted as dotted lines, and residual minimization method is plotted by solid lines.

ing extrapolation or residual minimization significantly improves the discontinuity of feature segments (see Figure 4.24 (b) and Figure 4.25 (b)).

4.6.3 Global transformation estimation and sub-volume transformation

Figure 4.26 shows a final 3D volume which consists of four consecutive sub-volumes from CLSM image stacks. Three sub-volumes (sub-volume 2, sub-volume 3, and sub-volume 4) are transformed to the target coordinate (the original coordinate system of sub-volume 1), and merged along axial direction (depth). One might detect some cylindrical structures (e.g., blood vessels) deforming along depth. Figure 4.27 shows a side view of the volume in Figure 4.26. It shows some intensity discontinuities along axial direction at physical boundaries of the slides. However, note that the visual discontinuity arises mostly from intensity heterogeneity (which can be corrected in visual enhancement), and structural morphology

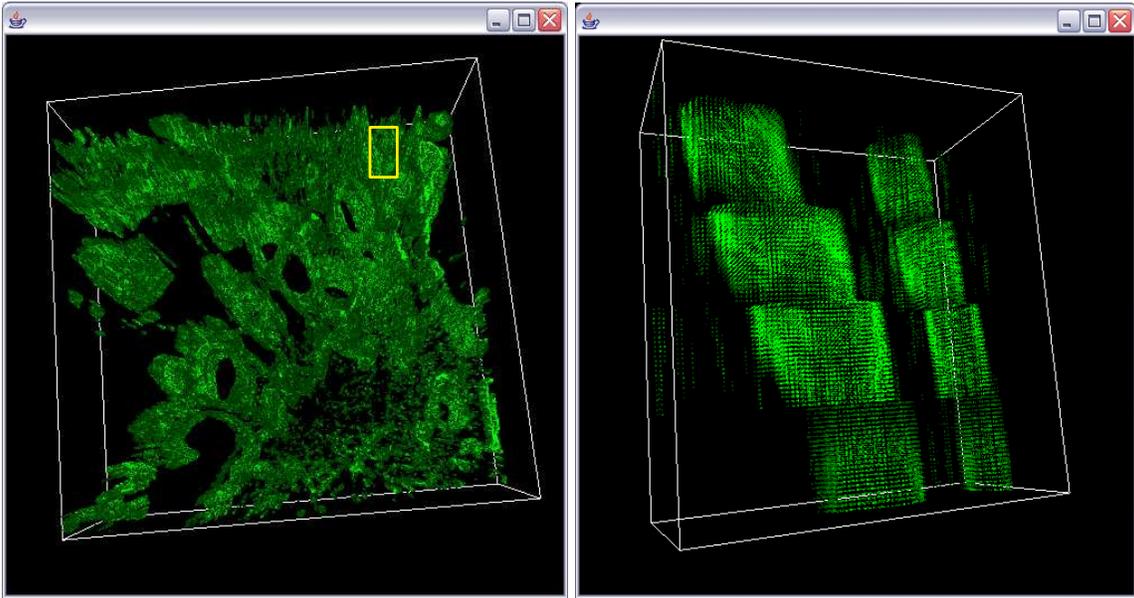


Figure 4.23: (left) Final 3D volume without trajectory fusion (gap = 1), and (right) sub-region at the box in the left image.

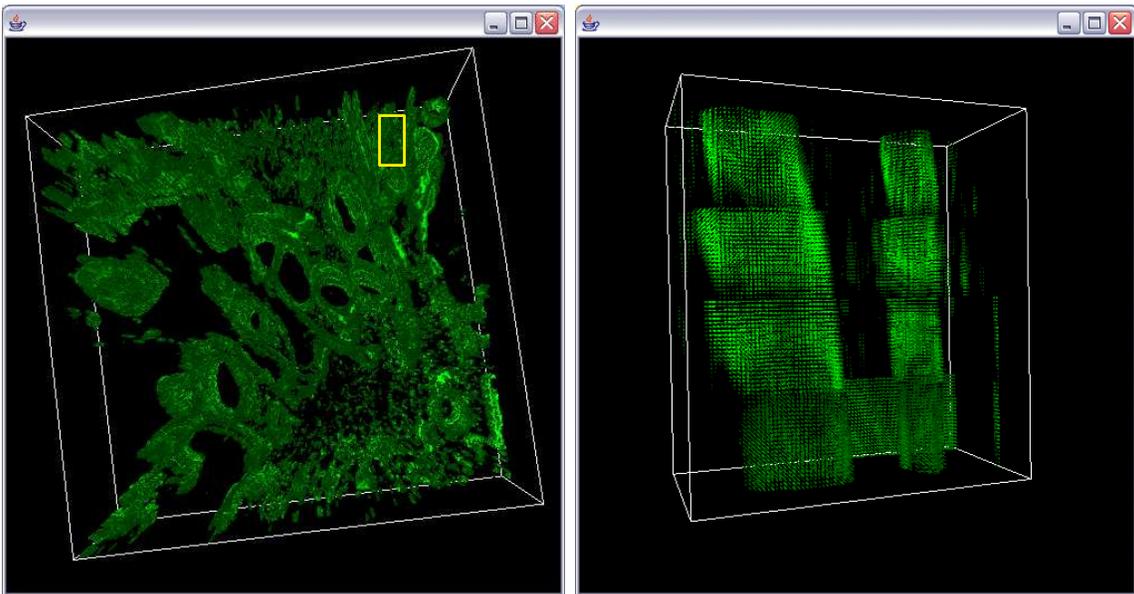


Figure 4.24: (left) Final 3D volume after trajectory fusion with extrapolation (gap = 1, linear model), and (right) sub-region at the box in the left image.

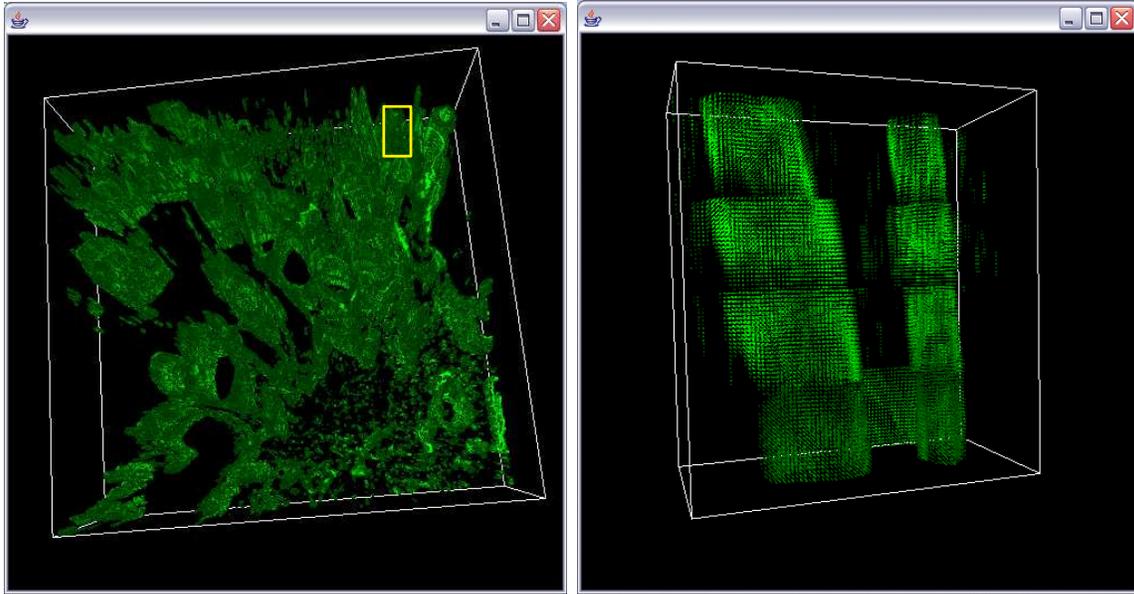


Figure 4.25: (left) Final 3D volume after trajectory fusion with residual minimization (gap = 1, cubic model), and (right) sub-region at the box in the left image.

is quite well preserved.

4.7 Summary

I report accuracy evaluations for (1) three registration methods including manual (pixel-based), semi-automatic (region centroid feature-based) and fully automatic (correlation-based) registration techniques and (2) two groups of human subjects (experts and novices). The study demonstrates significant benefits of automation for 3D volume reconstruction in terms of achieved accuracy, consistency of results, and performance time. In addition, the results indicate that the differences between registration accuracy obtained by experts and by novices disappear with an advanced automation while the absolute registration accuracy increases. If one is interested in performing data-specific evaluations then we prepared web based tools [76] for better data understanding and analysis at the URL ⁴.

⁴<http://i2k.ncsa.uiuc.edu/MedVolume/>

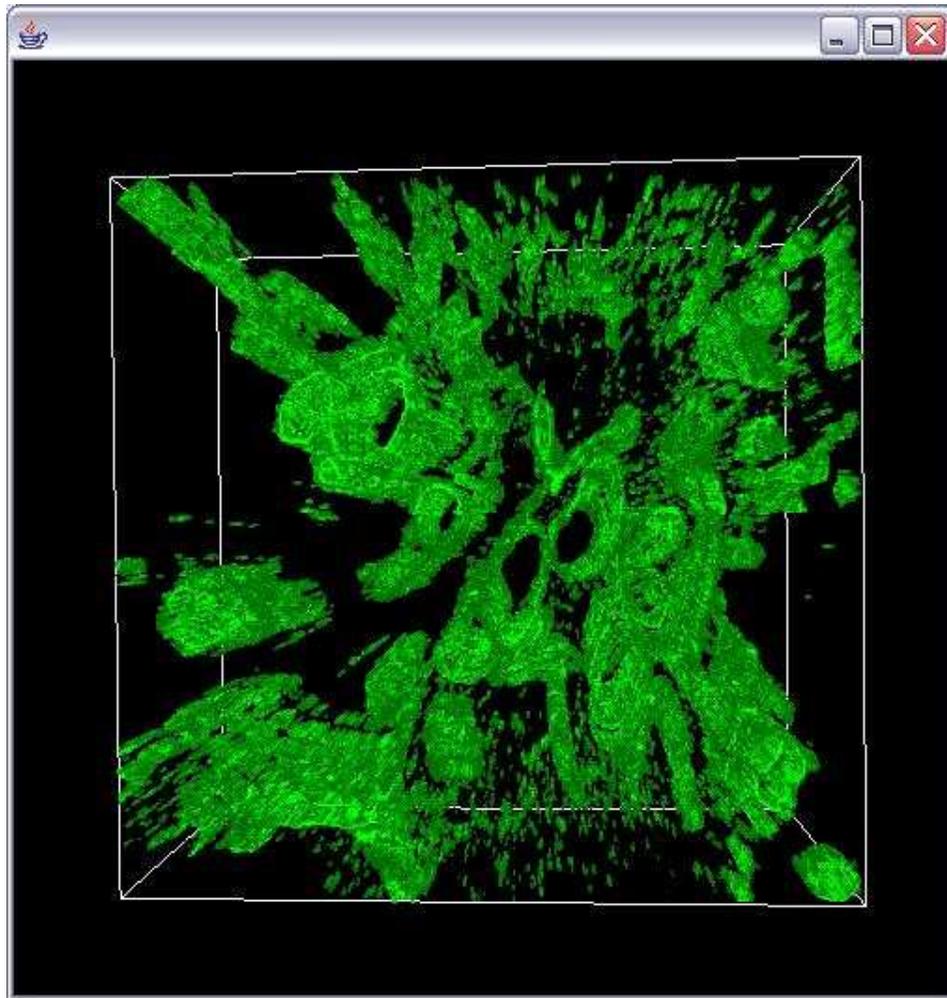


Figure 4.26: 3D view of the reconstructed 3D volume from four consecutive sub-volumes of CLSM image stacks (top view).

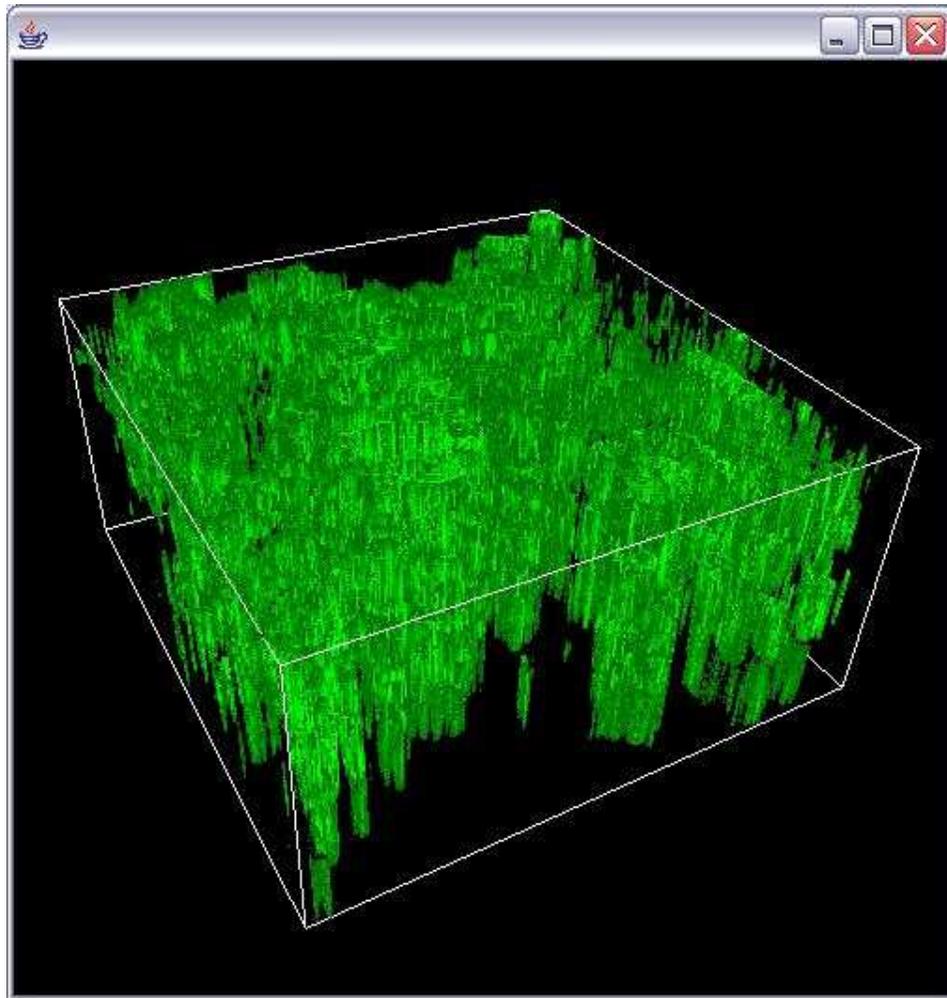


Figure 4.27: 3D view of the reconstructed 3D volume from four consecutive sub-volumes of CLSM image stacks (side view).

I investigate the main factors behind the summarized experimental results and present them in this section. First, the feature-based registration is faster and more accurate than pixel-based registration for alignment problems. The claims in accuracy improvement are supported by the paired t-test result. Second, the image alignment upper bound errors (23.61 for semi-auto and 129.5 for manual) are much higher than the mosaicking upper bound errors (4.12 for auto, 5.12 for semi-auto and 27.42 for manual). I believe that the main factors behind these differences are (1) one order higher complexity of the alignment problem (intensity and spatial structure variations across slides) in comparison with the mosaicking problem (intensity variations across tiles), (2) a larger degree of freedom in occurring image alignment transformations (rotation, scale, shear and translation) than in mosaicking transformations (translation), and (3) significantly larger sensitivity to human inconsistency in selecting points (attention level, skills, fatigue, display quality). Human inconsistency is expressed by a much larger standard deviation of errors in the case of alignment (35.74 for manual and 6.11 for semi-auto) than in the case of mosaicking (6.82 for manual and 0.35 for semi-auto).

Suggested fully automated method demonstrates significant benefits of automation for 3D volume reconstruction in terms of achieved accuracy (on average, 9.4 or 15 times smaller alignment error without or with optimization), consistency (25.8 or 113.7 times smaller standard deviation without or with optimization) of results and performance time. I also outlined the limitations of fully automated and manual 3D volume reconstruction systems, and described related automation challenges. I showed that given computational resources and repetitive experimental data, the automated alignment provides more accurate and consistent results than a manual alignment approach. With the proposed approach, the automation will reduce the alignment execution time and cost in the future, as the cost of more powerful computers goes continuously down.

Next, I presented a prototype solution for 3D medical volume reconstruction that was used in practice by UIC and NCSA collaborators. I overviewed the 3D volume reconstruc-

tion problem requirements, the architecture of the developed prototype system using web services and the tradeoffs of our system design. In a summary, the web services based approach provides two major advantages. First, a user will be able to perform computationally intensive image operations (a) with large size image data and (b) with sophisticated 3D volume reconstruction analysis methods. Furthermore, a user will not have to invest into (a) computational and storage resources and (b) development of complicated analysis software. Second, the currently advertised interoperability feature of web services will enable us in the future (a) to customize system front end (graphical user interfaces (GUI) for the user entry point) without changing system back end (complex algorithms that perform desired computations), (b) to upgrade algorithms and fix software bugs without any involvement of a user, and (c) to integrate distributed web services that will be available on the Internet.

Matching refinements are performed in two different senses: (1) optional intensity-based optimization for 2D alignment and (2) registration optimization based on trajectory fusion using 3D information. The intensity-based optimization showed improved alignment accuracy for two out of the three test image pairs (test image pairs 2 and 3). By using intensity-based alignment as an optimization step (rather than using alignment method), I could achieve low erroneous alignment with reasonably low computational cost. The registration optimization using trajectory fusion is aimed for preserving morphological smoothness of medical structures inside of the reconstructed 3D volume. To preserve morphological smoothness of the reconstructed 3D medical structures, I defined two metrics for morphological continuity labeled as end point discontinuity and polynomial curve fitting residual, and then minimize the metrics across adjacent sub-volumes and all salient structures.

Chapter 5

VISUAL ENHANCEMENT OF 3D CLSM VOLUME

5.1 Introduction

Visual inspection of medical specimens is one of the most common techniques in a medical domain used for learning and diagnosis. Based on the need to investigate specimen characteristics at high spatial resolution, fluorescent Confocal Laser Scanning Microscopy (CLSM) imaging is frequently used for obtaining 2D images of cross sections. The fluorescent labeling of medical specimens creates the basis for researchers to inspect visually spatial locations and morphology of structures of interest. Ideally, with confocal imaging and 3D volume reconstruction in place, CLSM image intensities provide 3D information about not only presence or absence but also abundance. Unfortunately, the intensity information is distorted due to multiple specimen preparation and image acquisition limitations. Thus, there is a need to investigate techniques for fluorescent CLSM image intensity adjustment to support visual inspection tasks.

The work primarily supports visual inspection of 3D volumes in virtual reality environments [21], where the 3D volumes are obtained by tile mosaicking and cross section alignment, then quantitatively analyzed (e.g., segmented) and correlated with information from other sensors (data integration). It is well recognized that spatial intensity heterogeneity is a major barrier in acquiring reliable results during the above analyses [11]. For example, major difficulties of 3D volume segmentation arise from spatial intensity heterogeneity rather than from intensity noise [73].

In this chapter, I introduce an intensity heterogeneity correction technique with auto-

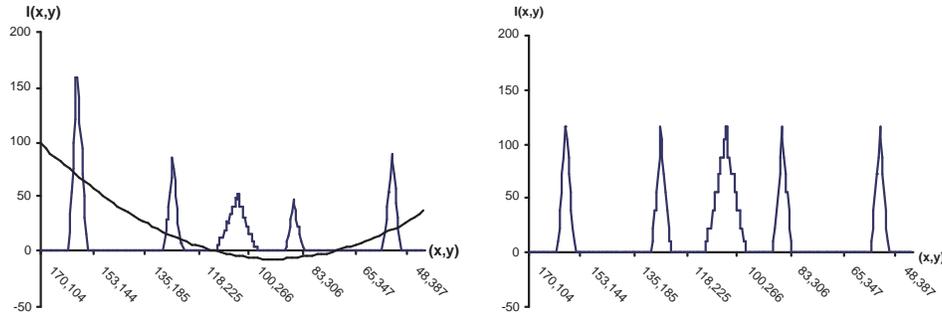


Figure 5.1: Intensity Correction Problem: (a) measured intensity profile for CLSM images with intensity bias (solid curve) and (b) corrected intensity profile. Intensity bias has been corrected while preserving local intensity gradients)

matic parameter selection based on image analysis [87]. The technique (a) adjusts intensity heterogeneity in x - y plane, (b) preserves fine structural details, and (c) enhances image contrast by performing spatially adaptive filtering, denoted as *mean-weight filtering*. I also provide novel numerical comparison metrics to objectively evaluate the quality of multiple intensity correction methods, such as low frequency and high frequency based metrics.

Although the intensity heterogeneity correction problem may be viewed as a restoration problem, I formulate the problem as an optimization problem since it is impossible to obtain the true uncorrupted intensity values for comparative purposes. Thus, I formulate the intensity heterogeneity correction problem as a search for an optimal, spatially adaptive, intensity transformation that (1) maximizes intensity contrast with respect to background, (2) minimizes overall spatial intensity variation for large area, e.g., low frequency domain, and (3) minimizes distortion of intensity gradient for local features, e.g., high frequency domain, as shown in Figure 5.1. I assume that the input image contains a single band (or a grayscale image) with spatially varying intensities.

5.2 Related Work

There are several known factors that cause spatial intensity heterogeneity, such as (a) photo-bleaching, (b) fluorescent attenuation along confocal (depth) axis, (c) image acquisition factors [11, 124, 42], (d) variations of illumination exposure rate, (e) spatially uneven distribution of dye and the spatial characteristics of illumination beams [107], and (f) fluorochrome micro-environment, e.g., pH, temperature, embedding medium, etc [138]. However, it may not be trivial to monitor all exact states of a fluorescent dye in any imaged specimen at a pixel resolution, as well as to know many of the image acquisition factors. Thus, it is very challenging to develop an intensity heterogeneity compensation method that would be directly linked to the sources of intensity heterogeneity coming from specimen preparation and image acquisition steps.

Given the complexity of factors introducing intensity heterogeneity, there are two approaches to improve CLSM images in a lateral plane. First, one could focus on improving specimen preparation and imaging conditions, e.g., improved fluorescent dyes [13]. Second, one could attempt correcting pixel intensities after image acquisition as it is the case of image restoration [129]. Image restoration can be performed based on empirical correction methods for intensity loss [127], constant thresholding [69], iterative correction methods [146, 126], 2D histogram [93], or estimations of intensity decay function [75]. Most of these methods assume that the rate of photo-bleaching (a) is spatially homogeneous in a lateral plane, (b) can be characterized by an exponential function in depth [124], and (c) mainly contributes to intensity loss along the specimen depth axis (z-axis) [107]. However, these assumptions might not always be satisfied as it was the case in the experiments.

In the past, Histogram Equalization (HE) has been used in early applications [48] to correct spatial intensity heterogeneity in lateral (x-y) plane. Histogram equalization leads to a uniform global intensity distribution in output image. However, it cannot effectively enhance local intensity variation due to its global property. To address this problem, Adap-

tive Histogram Equalization (AHE) has been used to adjust intensity variation locally by computing local histograms within spatially different windows [117]. A major problem of Adaptive Histogram Equalization is high sensitivity to noise, which results in amplification of undesired noise values. An improved approach to adjust local intensity variation is the Contrast Limiting Adaptive Histogram Equalization (CLAHE) [116]. It reduces noise amplification due to Adaptive Histogram Equalization by setting clipping limits and so removes boundary artifacts by background subtraction. Nevertheless, the main drawbacks of CLAHE are (1) the parameters need to be manually selected and (2) there could be loss of fine details caused by intensity saturation. Some researchers proposed an intensity correction method based on the bias field, e.g., intensity distortion map, to a polynomial function which is defined in prior [137]. However, this method requires prior knowledge about images, e.g., the degree of a polynomial function, and the computational complexity tends to increase exponentially upon the degree of a polynomial function.

5.3 Visual Enhancement

To solve the problem as formulated above, I designed a spatial filter and several image quality measures according to the optimization criteria. The approach combines a kernel-based spatial filtering, and incorporates local and global image intensity analysis.

Intuitively, 3D voxel intensity adjustment should be performed by examining neighboring voxels and adjusting (filtering) intensities to meet visual inspection criteria without distorting true intensities too much. However, the 3D voxel intensity adjustment approach ignores the fact that intensity heterogeneity observed in a lateral plane (x-y) and along an axial coordinate (z or depth) are fundamentally different in their nature. In a lateral (x-y) plane, intensity heterogeneity can be characterized by relatively low frequency variations while intensities along an axial coordinate change more rapidly over a small physical distance

(thickness of a cross section). I have primarily focused on designing a 2D filtering technique that preserves edge intensity gradient in a lateral (x-y) plane and minimizes rapid intensity variations along an axial coordinate. The filtering along axial direction is embedded into the filtering of any lateral (x-y) plane by performing frame mean intensity based background separation and intensity rescaling after speckle noise removal to cover the full dynamic range, e.g., 0 to 255 for a byte image.

The proposed 3D intensity correction process consists of (1) determining the background threshold intensities for each frame, (2) optimizing the kernel size, (3) constructing a set of 2D intensity correction maps for the volume, (4) multiplying the intensity correction maps to the frames, (5) remove outlier pixels (speckle noise) from the corrected images based on statistical value ranges, and (6) rescaling the intensity range of each frame to cover the full dynamic range of the data. The method is denoted as *Mean-weight filtering*. In this section, I describe the filter model and the automatic kernel parameter selection based on the image analysis.

5.3.1 2D filter model

An image filtering approach to the presented intensity correction problem can be described as:

$$g_{[a,b]}(x, y) = I(x, y) \cdot w_{[a,b]}(x, y) \quad (5.1)$$

where $g_{[a,b]}(x, y)$ and $I(x, y)$ are the output and input pixel values at (x, y) respectively and $w_{[a,b]}(x, y)$ is the weighting coefficient computed over a pixel neighborhood $([x - a, x + a], [y - b, y + b])$. The spatial neighborhood, also denoted as a filtering kernel of size $(2a + 1) \times (2b + 1)$ is introduced to meet the requirement on local intensity gradient. Other requirements on intensity contrast and global spatial intensity variations are incorporated into the filter design

by (a) separating background according to Equation (5.2), and (b) computing the weighting coefficient $w(x, y)_{[a,b]}$ as a ratio of global and local sample means (see Equation (5.3)).

$$\hat{I}(x, y) = \begin{cases} 0 & \text{if } I(x, y) \leq \sigma \\ I(x, y) & \text{if } I(x, y) > \sigma \end{cases} \quad (5.2)$$

$$w_{[a,b]}(x, y) = \begin{cases} 1 & \text{if } I(x, y) \leq \sigma \\ 0 & \text{if } \mu_{L[a,b]}(x, y) = 0 \\ \frac{\mu_G}{\mu_{L[a,b]}(x, y)} & \text{if } \mu_{L[a,b]}(x, y) \neq 0 \end{cases} \quad (5.3)$$

where μ_G and $\mu_{L[a,b]}(x, y)$ are the global and local estimated sample means computed according to Equations (5.4) and (5.5), and σ is a background threshold intensity value.

$$\mu_G = \frac{1}{NM - |\phi_G|} \cdot \sum_{i=1}^N \sum_{j=1}^M \hat{I}(i, j) \quad (5.4)$$

$$\mu_{L[a,b]}(x, y) = \frac{1}{(2a+1)(2b+1) - |\phi_{L[a,b]}(x, y)|} \cdot \sum_{s=-a}^a \sum_{t=-b}^b \hat{I}(x+s, y+t) \quad (5.5)$$

where $\phi_G = \{(x, y) | \hat{I}(x, y) = 0; 1 \leq x \leq N, 1 \leq y \leq M\}$, $\phi_{L[a,b]}(x, y) = \{(s, t) | \hat{I}(x+s, y+t) = 0; -a \leq s \leq a, -b \leq t \leq b\}$, and N and M are the width and height of the input image. I denoted the threshold image as $\hat{I}(x, y)$ and the estimated weighting coefficients as $w_{[a,b]}(x, y)$.

It is apparent that the weighting coefficients could be ill-defined when the local sample mean takes very small values (the spatial kernel belongs to background with some noise, $\mu_{L[a,b]}(x, y) \approx 0$.) To avoid this problem, input image is thresholded first, and then the values of global and local sample means are computed only over foreground pixels. In the filter design, I introduced two parameters, such as a background threshold σ and a kernel size

$(2a + 1) \times (2b + 1)$. In order to automate the filtering method and achieve a parameter-bias free performance, I describe novel data driven parameter selection methods in the next two sections.

5.3.2 Background separation

To separate background from foreground in image, one would like to decide on an appropriate threshold value σ for noise removal due to variable camera exposures or sensitivity of fluorescent photo multiplier tube. In general, the background threshold value could be determined by purely depending on images, such as variogram [130]. However, it is preferred that the noise model of different imaging techniques should be modeled differently based on known imaging physics such as Exponential, Rayleigh, Gaussian, Gamma, Poisson, or Weibull function [67]. Although those functional models fit well for specific imaging modalities, it is well known that the noise in laser imaging can be modeled by an exponential function [117]. In background separation, I assume an exponential noise model for the CLSM background noise.

To select a set of threshold values for a stack of images, I first select the brightest image frame based on the foreground sample means and select the threshold value. Next, each threshold value for individual frame is selected by adjusting the selected threshold values with respect to its sample mean. This process enables us to separate out the background with high confidence since the brightest image offers (1) low intensity frequency in low intensity range, (2) concentrated intensity frequency distribution in high intensity value, and (3) low mixture of foreground intensity frequency with background noise.

I derived σ by modeling a frequency function of the pixel intensities, i.e., a histogram, and by comparing with underlying physical models of CLSM imaging. I model the frequency function of intensity values as a conjunction of (a) background noise $\Lambda(v)$ in a photo-multiplier tube and (b) foreground fluorescent pixel intensity distribution $\Phi(v)$, where v is an intensity value, e.g., bin number in histogram. Assuming that $\Lambda(v)$ and $\Phi(v)$ follow

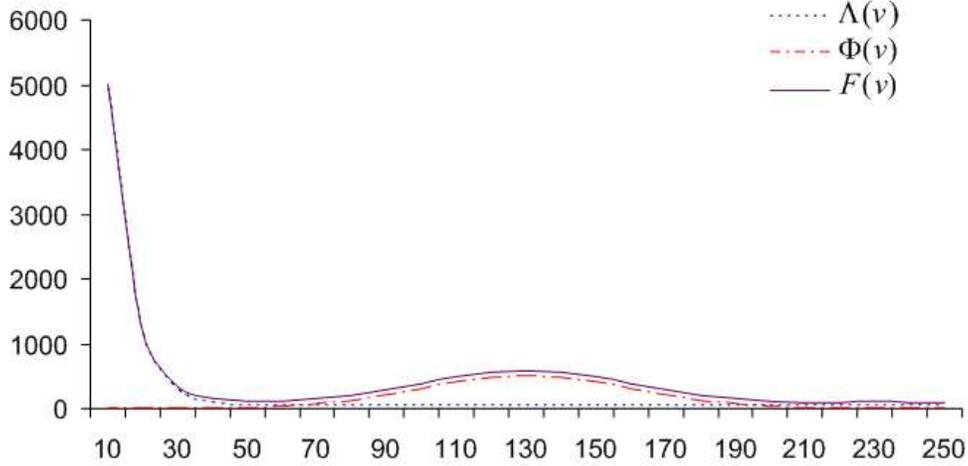


Figure 5.2: Intensity distribution for a CLSM image. $\Phi(v)$ is illustrated as a Gaussian function.

an exponential and unknown exponential family of functional model, I define the frequency functions as $\Lambda(v) = \alpha e^{\beta v}$ and $\Phi(v) = \gamma(v)e^{\kappa(v)}$ where α, β are constants and $\gamma(\cdot), \kappa(\cdot)$ are some functions for foreground intensity values. Assuming the defined components are independent and noise are additive, the frequency function $F(v)$ of output intensity can be modeled as a sum of $\Lambda(v)$ and $\Phi(v)$, which is $F(v) = \Phi(v) + \Lambda(v)$ as illustrated in Figure 5.2.

In particular, since the background pixels usually appear in low-intensity ranges and I assume no prior knowledge about the foreground intensity frequency model, for thresholding purpose, I only consider the left part of the $F(v)$ where foreground intensity starts to contribute to $F(v)$ by fitting the noise model $\Lambda(v)$ to the low intensity range of $F(v)$. To compute the background frequency function, I estimated the parameters α, β using the least-squared methods [148] as following:

$$\begin{bmatrix} \sum_{v=1}^k F(v) & \sum_{v=1}^k vF(v) \\ \sum_{v=1}^k vF(v) & \sum_{v=1}^k v^2F(v) \end{bmatrix} \begin{bmatrix} \alpha \\ \ln(\beta) \end{bmatrix} = \begin{bmatrix} \sum_{v=1}^k F(v) \ln(F(v)) \\ \sum_{v=1}^k vF(v) \ln(F(v)) \end{bmatrix} \quad (5.6)$$

where $F(v)$ is the frequency (e.g., pixel count) of the intensity value v in the input image, k is the bin number (intensity value) in $[1, n]$, and n is the total number of bins.

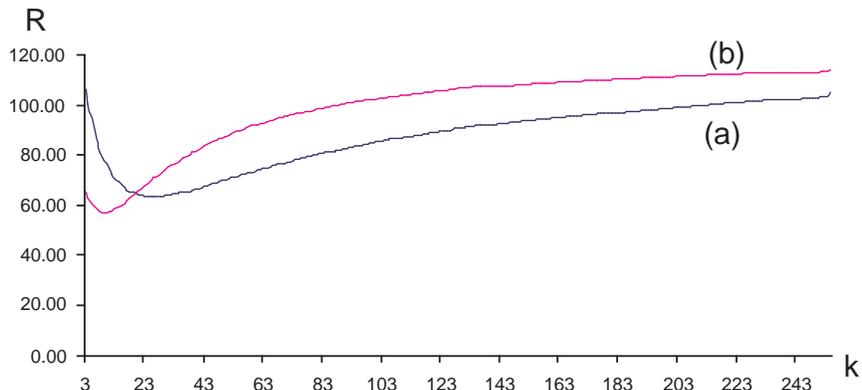


Figure 5.3: Sum of the squared error of the estimated function for (a) and (b). The intensity value k minimizing R are 26 and 9 respectively.

In order to find the threshold value which is dominantly determined by the background noise by photo-multiplier tube, I computed the sum of squared error

$R(k) = \left(1/0.1n \sqrt{\sum_{v=1}^{0.1n} (F(v) - \Lambda_k(v))^2} \right)$ by increasing k from 3 to n , where $\Lambda_k(v)$ is an exponential function fit to first k bins. Since I approximated the foreground pixel intensity distribution for the left tail only, I calculated the sum of squared error for the pixel intensity values in 10% of low intensity value range, i.e., $[1, 25]$ in byte (8-bit) image. It approximates $\Lambda_k(v)$ in the input image well enough since R is not affected much by large intensity values (see Figure 5.2). Figure 5.3 shows the sum of squared error R as a function of k for the real CLSM images used in the experiments. The graph shows that (a) the estimate of the intensity multiplication factor only for low intensity values results in high error, (b) there exists an intensity value k which minimize the error, and (c) the sum of squared error increases when the $\Lambda(v)$ starts mixing with $\Phi(v)$, i.e., fitting the entire curve.

Finally, the background threshold value σ is determined by the value k where the estimated function $\Lambda_k(v)$ best fits with the background noise in the input image such as $\sigma = \arg \min_{k \in [1, n]} (R(k))$. In the example shown in Figure 5.3, I acquired the threshold values 26 and 9 for each image.

5.3.3 Kernel size optimization

In order to compute local sample means, one has to choose the size and shape of a filtering kernel for an input image. I first constrained the kernel shape to a square. Then I selected a kernel size based on maximizing the global contrast and minimizing the gradient distortion, e.g., high frequency distortion. The global contrast metric [7] incorporates one of the requirements of the intensity heterogeneity correction problem and is mathematically described below.

$$C = \sum_{i=1}^m \|f_i(I) - E(f(I))\| \times f(I)_i \quad (5.7)$$

where $f(I)$ is the histogram (estimated probability density function) of all contrast values by using Sobel edge detector in an image I , $f_i(I)$ is the density of i -th bin, $E(f(I))$ is the sample mean of the histogram $f(I)$, and m is the number of distinct contrast values in a discrete case. The equation includes the contrast magnitude term and the term with the likelihood of contrast occurrence. In general, image frames characterized by a large value of C are more suitable for further processing than the frames with a small value of C .

To demonstrate the high frequency image difference between the original image and the processed image, I define a metric D as follows:

$$D = \sum_v^M \sum_u^N (I_{hf}^{org}(u, v) - I_{hf}^{adj}(u, v))^2 \quad (5.8)$$

where I_{hf}^{org} and I_{hf}^{adj} are the high-pass filtered images of the original and the intensity adjusted image respectively. Finally, the filter size is selected by evaluating the maximum value of the ratio C/D denoted as a measure of image saliency.

5.3.4 Speckle noise removal and intensity rescaling

One of the side products of the mean-weight filtering is an easy detection of speckle noise in the background. Speckle noise is characterized by a very few or no neighboring pixels and the mean-weight filtering generates very high intensity correction value for the speckle pixels. I have eliminated speckle noise by removing the pixels with high intensity correction values, and accepted the pixels with values within the range $[0, \mu + 4\sigma]$ (99.99% of pixels are included), where μ and σ are the sample mean and standard deviation of the intensity corrected image. Finally, the values within the range $[0, \mu + 4\sigma]$ are normalized (rescaled) to $[0, 255]$ to meet the dynamic range of output images (8 bits per pixel).

5.4 Results

In this section, first, I show simulation results of three synthetic images with pre-defined bias fields (intensity distortion map). I apply the bias fields to images with known structures in order to demonstrate performance of the developed method in multiple well-understood intensity distortion scenarios. Next, I report comparisons of the developed mean-weight filtering method and other techniques using real CLSM images. In these comparisons, it is assumed that the intensity bias fields are unknown.

5.4.1 Simulation

Synthetic data preparation: First, I created three pre-defined intensity variations (bias fields) as shown in Figure 5.4 (a), (b) and (c). The intensity variations demonstrate spatially localized bias (Figure 5.4 (a) and (c)) with or without significant saturation, and spatially distributed bias (Figure 5.4 (b)). Second, an intensity bias free image of 512 x 512 pixel resolution with known structures (shown as horizontal and vertical bars with different thickness and spacing) was formed to simulate fluorescent boundaries of medical specimen (e.g., blood vessel boundaries) as shown in Figure 5.5. Third, the bias free image is distorted by

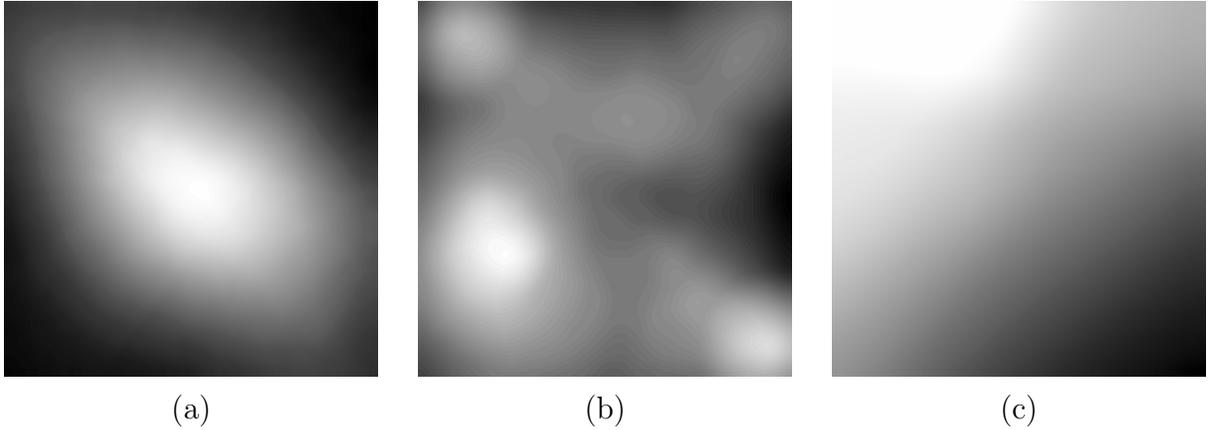


Figure 5.4: Images of intensity bias fields (intensity distortion maps.)

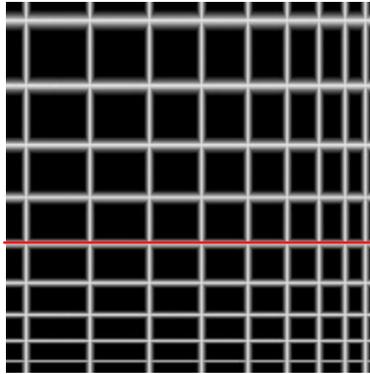


Figure 5.5: Simulation images: a bias-free image with known structures.

applying the intensity bias fields in Figure 5.4. Finally, random exponential synthetic noise is added into each image (with density function $\Lambda(v) = 0.3e^{0.3v}$). The final synthetic images for the simulation are shown in Figure 5.6, (a), (b) and (c).

Image processing: I performed intensity correction of the three simulated images by finding the threshold value for background separation and optimizing the kernel size described in the previous section. First, background threshold value was selected by finding the lowest residual after fitting exponential function to the low intensity range. Figure 5.7 shows the three curves of residuals for the simulation images. Based on the residual, I selected 20, 20, and 20 for the background thresholding of the simulation images.

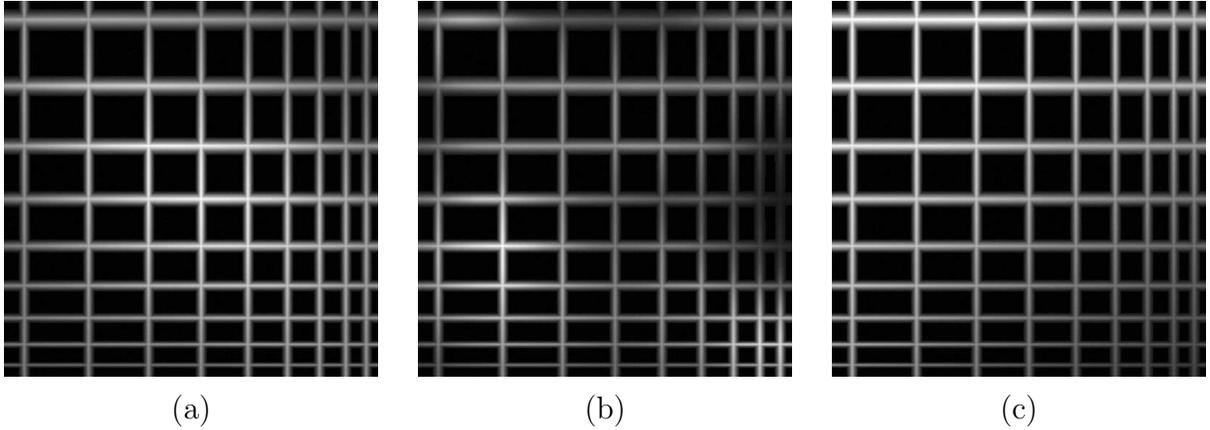


Figure 5.6: Simulation images: three images distorted by applying the bias fields shown in (a), (b) and (c) in Figure 5.4. Thickness and the spacing of the bar structures are decreasing toward the lower right corner.

Next, to optimize the kernel size for each image, I calculated the image saliency measure (C/D) using different kernel sizes. The results shown in Figure 5.8 (c) illustrate that the image saliency measure is low for small kernel sizes (affected by large D), having a peak, and decreasing for large kernels (due to decreasing C). One can conclude that (1) for small kernels, the contrast C is maximized but the high frequency distortion is large, (2) for large kernels, the high frequency distortion is minimized but the contrast C is compromised. Therefore, the simulation results verified that the optimal kernel size is achieved by maximizing the contrast divided by the high frequency difference (C/D). In the simulation experiments, I calculated C/D using kernel sizes from 3 to 51 pixels wide, and obtained the optimum values equal to 33, 19 and 47 pixel widths respectively (see the peaks in Figure 5.8 (c)).

To generate intensity corrected images, first, I computed the weight for each pixel by analyzing distorted images (unknown bias field). Figure 5.9 shows the distribution of weights (also denoted as the intensity correction map) for three distorted synthetic images. Next, each intensity correction map is applied to the distorted image to generate the intensity corrected images.

Figure 5.10 shows the intensity corrected images using the intensity correction maps

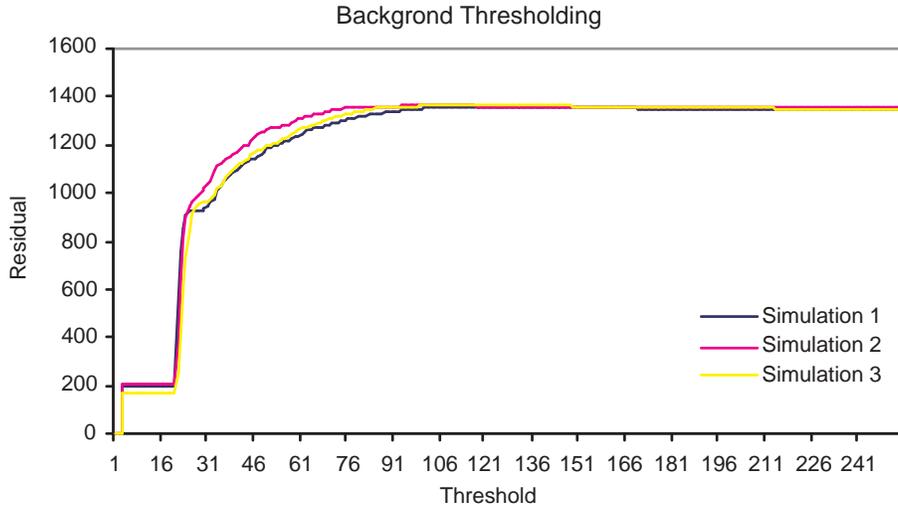


Figure 5.7: Background thresholding.

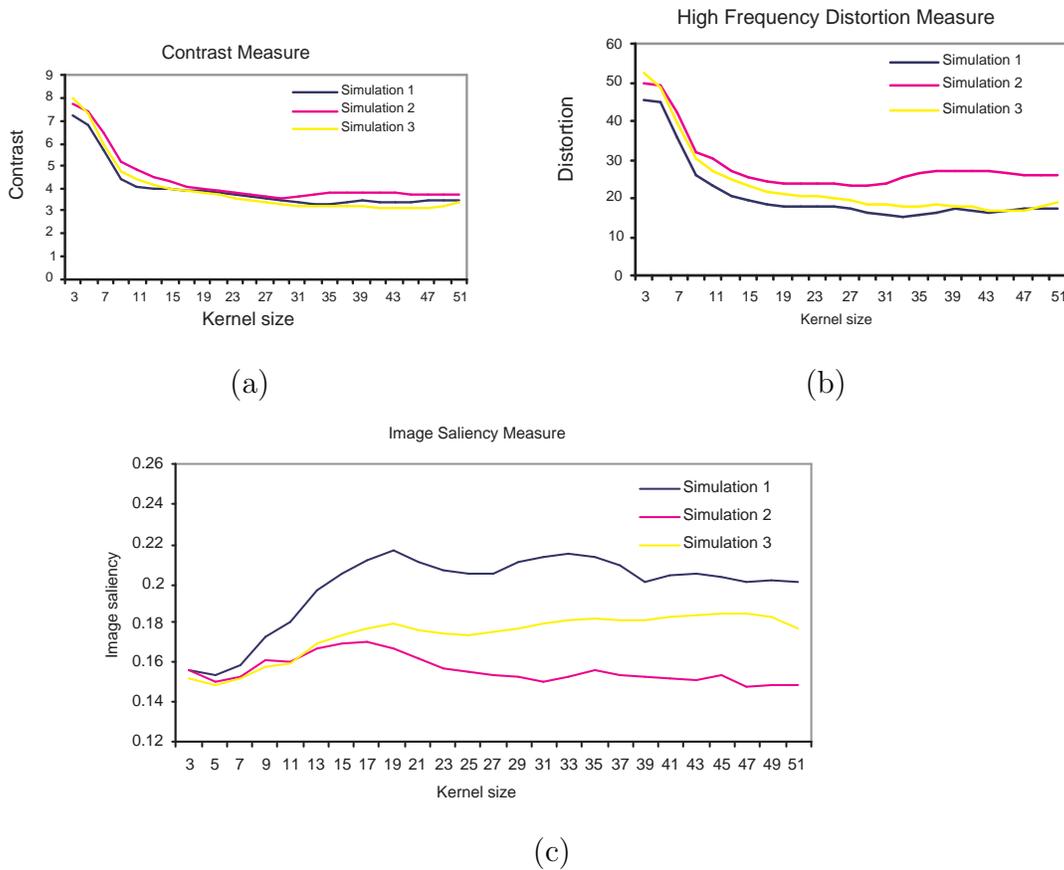


Figure 5.8: (a) Contrast C , (b) high frequency distortion D , and (c) image saliency as a function of the kernel size. The curves are computed for image data shown in Figure 5.6 (a), (b) and (c).

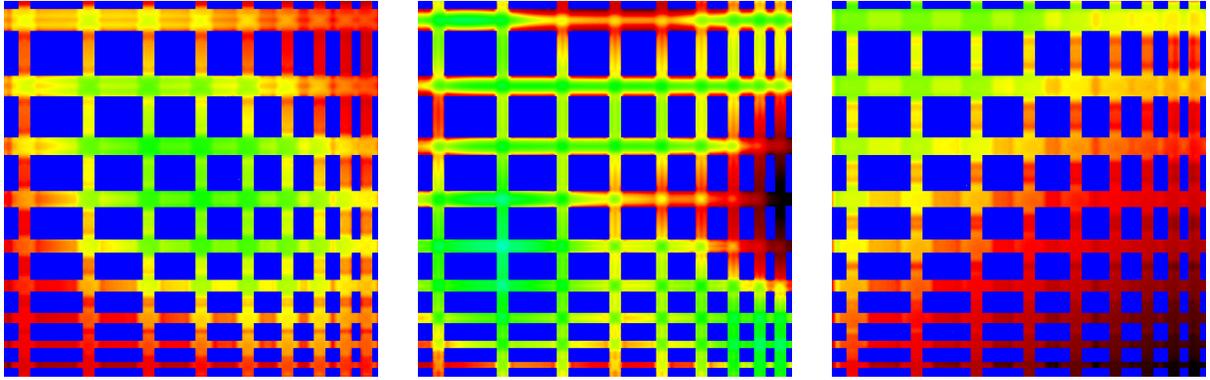


Figure 5.9: Intensity correction maps with (a) $\sigma = 20; a = b = 33$, (b) $\sigma = 20; a = b = 19$, and (c) $\sigma = 20; a = b = 47$. Note that the weight is pseudo colored, where blue denotes low and red denotes high values.

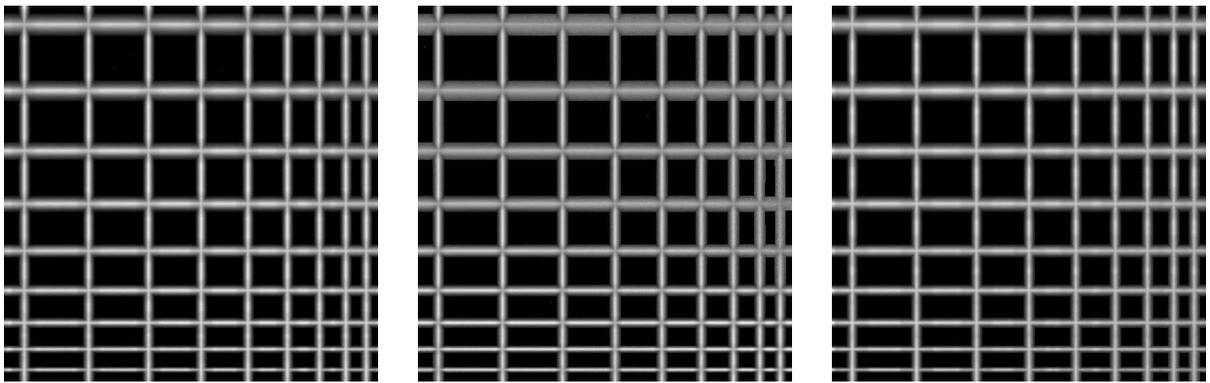


Figure 5.10: Intensity corrected images using the Mean-weight filtering from Figure 5.9. For each of the three images, the intensity bias has disappeared.

shown in Figure 5.9. Regardless of the type of a bias field, all three intensity corrected images show noticeable results that the mean weight filtering corrects intensity heterogeneity in spatially large area of the images while preserving the edge gradients (minimum high frequency distortion).

Figure 5.11, Figure 5.12 and Figure 5.13 show the intensity correction results for the three simulation images, using Histogram equalization, Histogram equalization with background separation, and CLAHE. One could notice that (a) Histogram equalization saturated most of pixels as well as background noise, and removed edge gradient significantly (b) Histogram equalization with background separation separated out background noise, but the

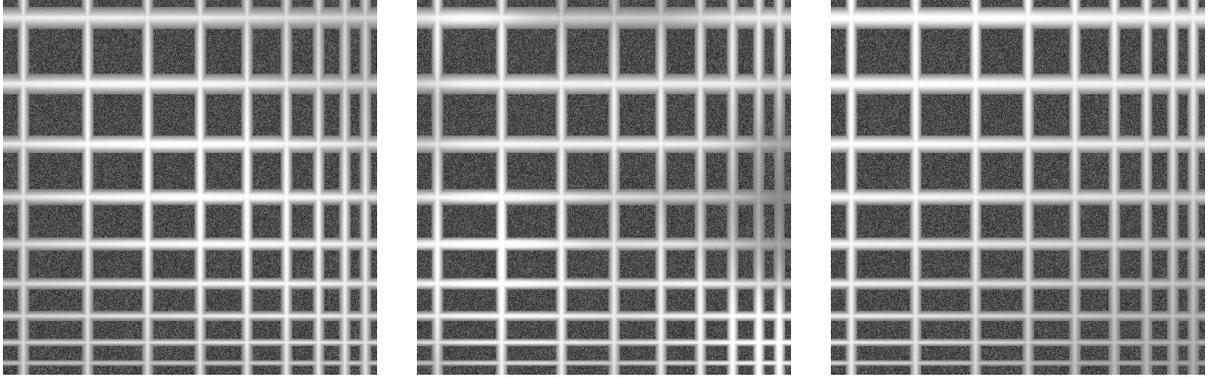


Figure 5.11: Result of Histogram Equalization

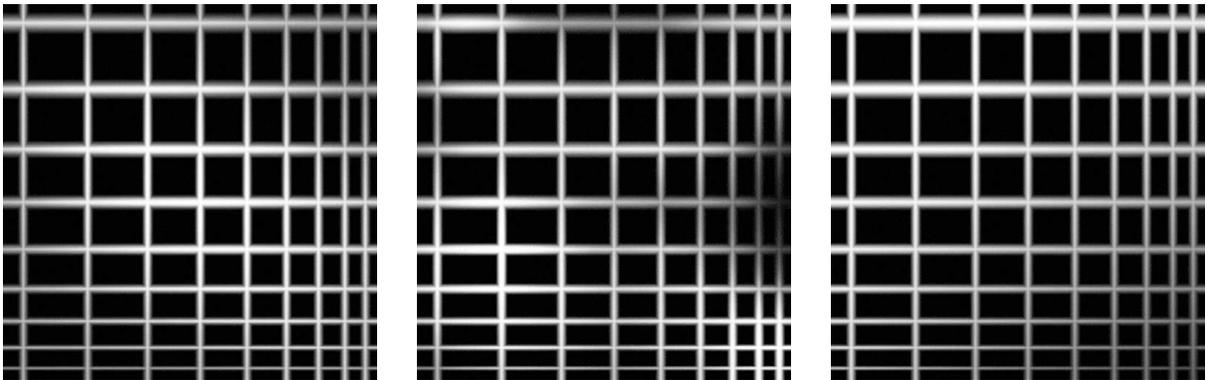


Figure 5.12: Result of Histogram Equalization with background thresholding

edge gradient is not preserved well similarly to the Histogram equalization, and (c) CLAHE relatively well preserved the edge gradient, but the spatial (x-y) intensity heterogeneity has not been corrected in comparison to the Mean weight filtering.

To demonstrate the spatial intensity correction result, Figure 5.14 shows the intensity profiles (along the red line Figure 5.5) of the test image (c) in Figure 5.6 for the original (uncorrected) image, existing methods, and the mean weight filtering. The intensity profile clearly demonstrates that the intensity along x-y plane is best corrected by the mean weight filtering, i.e., intensity value remains between 140 and 200.

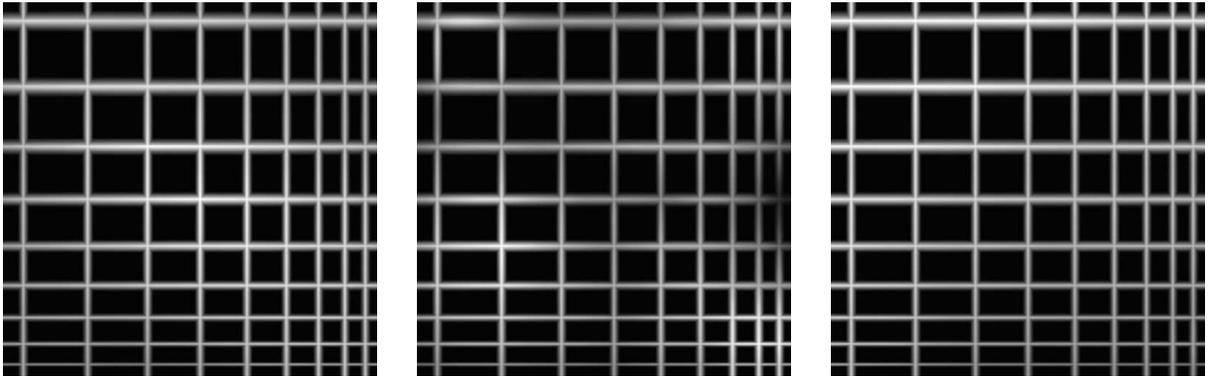


Figure 5.13: Result of CLAHE.

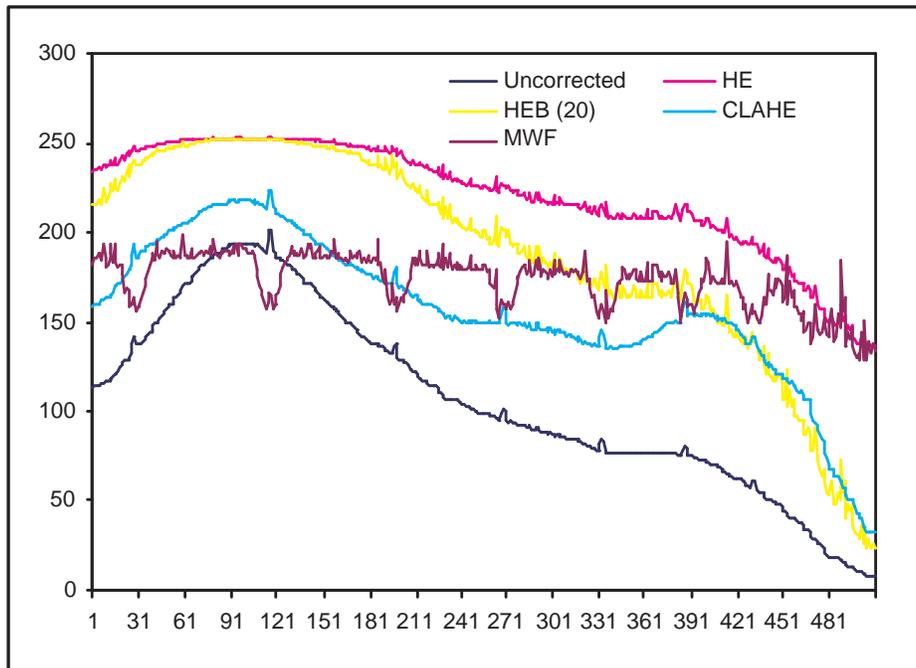


Figure 5.14: Intensity profile along the red line in Figure 5.5.

5.4.2 Real CLSM images

Material preparation: Formalin-fixed, paraffin-embedded uveal melanoma tissue samples were sectioned at 4 μ m thickness. The use of archival human tissue in this study was approved by the Institutional Review Board of the University of Illinois at Chicago. Slides were deparaffinized in xylene and rehydrated through a decreasing ethanol gradient. Slides were rinsed in distilled water followed by antigen unmasking using Target Retrieval Solution 10X Concentrated (DAKO, Carpinteria, CA) according to the manufacturer's instructions and then rinsed in Phosphate Buffered Saline (PBS) for 5 minutes. Slides were incubated with monoclonal mouse anti-laminin antibody Sigma L8271, clone LAM 89 (Sigma, St. Louis, MO) at a dilution titer of 1:200 for 30 minutes at room temperature. Slides were rinsed in protein blocking solution (DAKO) for ten minutes followed by detection with Alexa Fluor 488 goat anti-mouse IgG (Molecular Probes, Eugene, OR) for 30 minutes at a dilution of 1:400. Slides were rinsed in buffer then mounted in Faramount Aqueous Mounting Medium (DAKO). For all staining procedures, secondary antibody was omitted in negative controls.

All histological serial sections were examined with a Leica SP2 laser scanning confocal microscope (Leica, Heidelberg, Germany) using the 40X objective with 605–700 nm excitation wavelength range for the test specimens. Images were stored in tagged information file format (TIFF) with 512 by 512 pixel resolution. One 3D volume was formed from 13 image frames along axial coordinates (z-coordinate or depth) in a same lateral area. Therefore, a sub-volume consists of 512 by 512 by 13 voxels, which is equivalent to 375 x 375 x 3.91 microns in physical dimension. The red structures represent extravascular matrix patterns (loops) or blood vessels. Two unprocessed images with spatial intensity heterogeneity are shown in Figure 5.16 (a) and Figure 5.18 (a).

Compared intensity correction techniques: The intensity of each image is corrected by (1) Histogram Equalization, (2) Histogram Equalization with background threshold, (3)

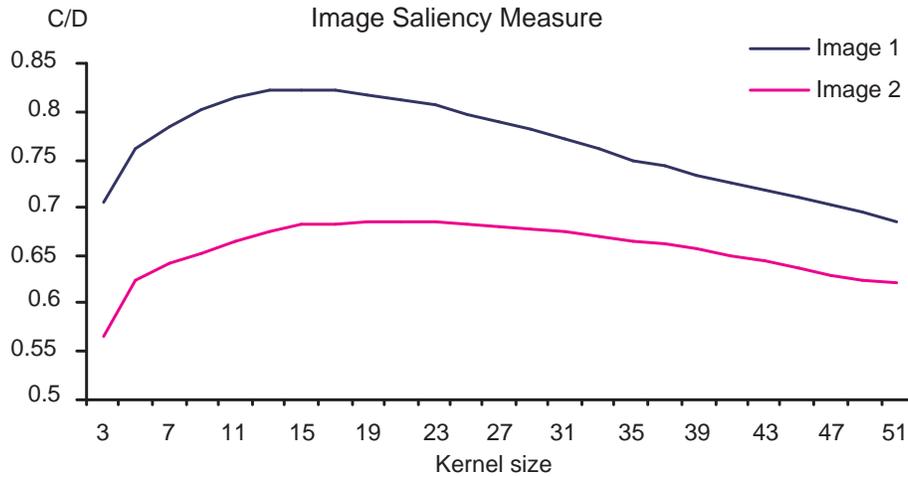


Figure 5.15: Image Saliency Measure for automatic kernel selection: the optimal kernel sizes are 15 and 21 for the two test image shown in Figure 5.16 (a) and Figure 5.18 (a).

Contrast Limited Adaptive Histogram Equalization, and (4) Mean Weight Filtering with automatic parameter selection. To apply the mean weight filtering, the background threshold and kernel size are automatically selected as illustrated in Figure 5.3 and Figure 5.15. According to Figure 5.3, the optimal threshold parameters are 26 and 9 for the test images in Figure 5.16 (a) and Figure 5.18 (a). Similarly, according to Figure 5.15, the image saliency measure the optimal kernel sizes are 15 and 21 (selected as the peak of each curve).

Using the optimal parameters, corresponding intensity correction maps (weight coefficients) are generated and presented in Figure 5.16 (f) and Figure 5.18 (f). The dark edge pixels in each original image are weighted by larger value coefficients, and the bright edge pixels are weighted to remain approximately the same.

Figure 5.16 and Figure 5.18 show the original CLSM image and processed results by histogram equalization (shown as the image (b)), histogram equalization with background thresholding (shown as the image (c)), CLAHE (shown as the image(d)), and mean weight filtering (shown as the image (e)). One could visually notice that the histogram equalization method removes best intensity heterogeneity in large spatial regions but maintains low visual saliency (low contrast) and removes edge details (high frequency distortion). The

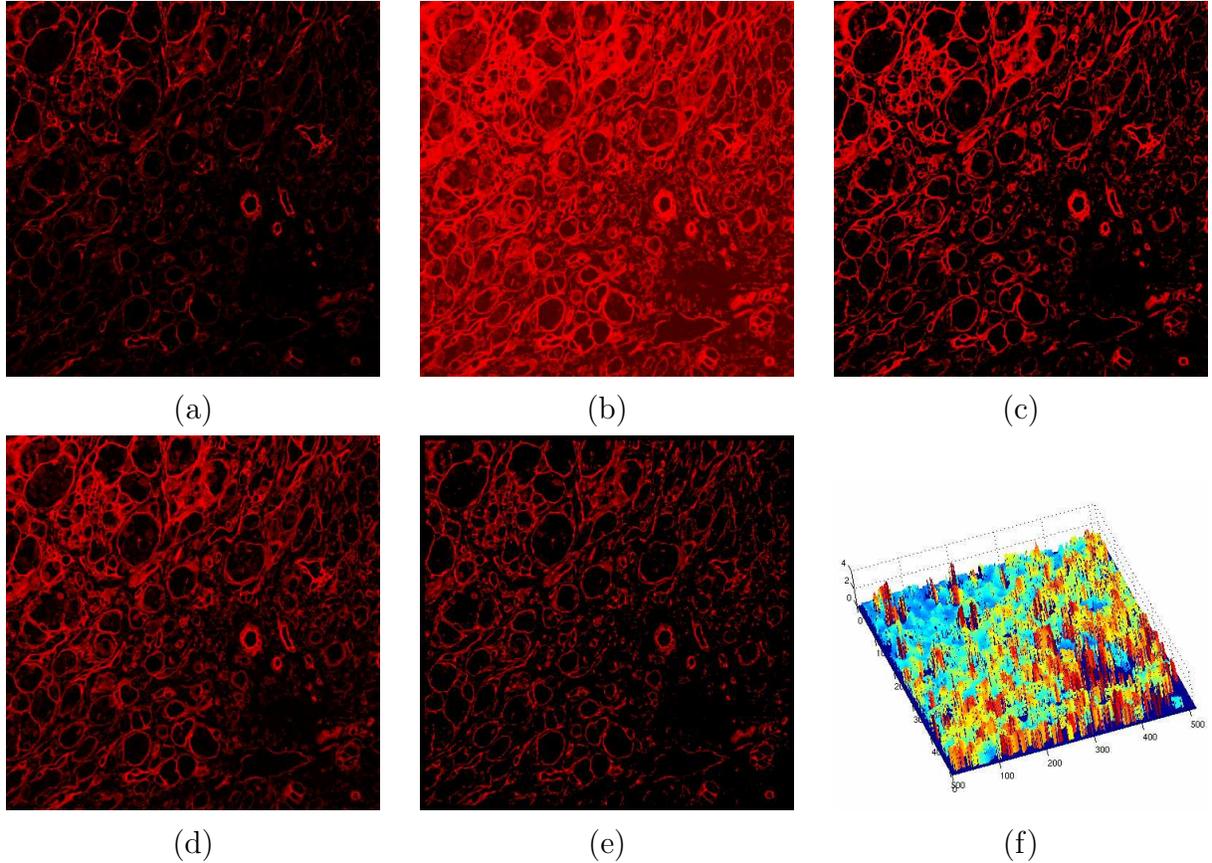


Figure 5.16: CLSM test image 1: (a) Original test image, (b) histogram equalization, (c) histogram equalization with background thresholding, (d) CLAHE ("adaphisteq()") with default setting by Matlab Image Processing Toolbox, Build. R12), (e) mean weight filtering, and (f) intensity correction maps by the mean weight filtering.

histogram equalization with background thresholding and Contrast Limited Adaptive Histogram Equalization demonstrate similar removal of edge gradient and edge details. To demonstrate the effect of edge gradient, I show a sub-region of the intensity corrected images in Figure 5.17 and Figure 5.19. One could see that most of pixel values around the edge was saturated using HE, HEB, or CLAHE. Next, I show the quantitative evaluation of the compared techniques.

Evaluation Metrics: In order to quantitatively asses the quality of multiple filtering techniques for real CLSM images, multiple intensity correction results are compared by (a) image

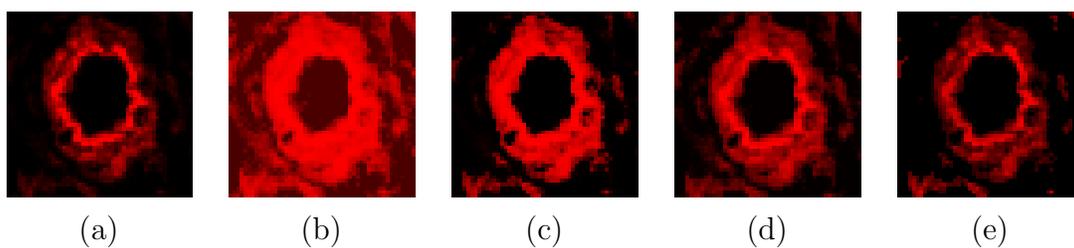


Figure 5.17: Sub-region of Intensity corrected images in Figure 5.16 (From left, original image, Histogram Equalization, Histogram Equalization with background threshold, CLAHE, and Mean-weight filtering.)

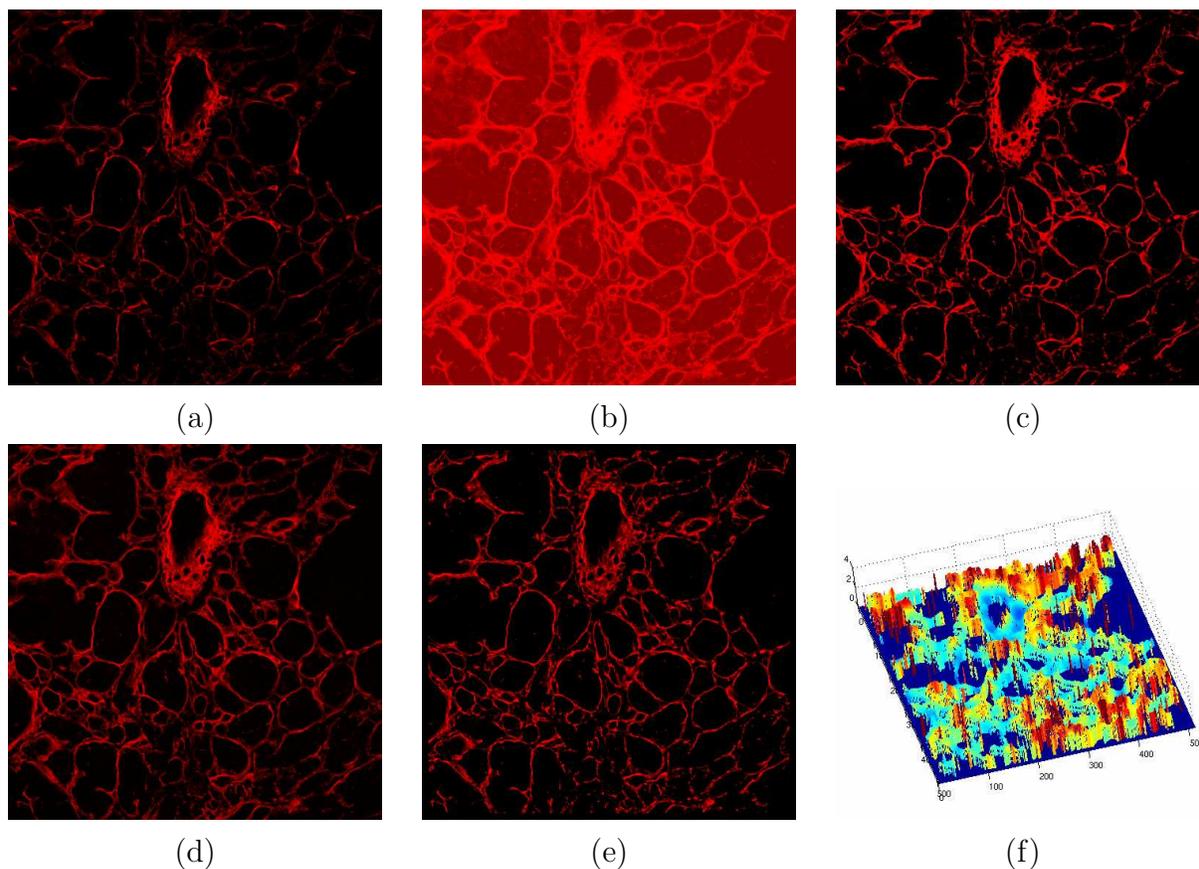


Figure 5.18: CLSM test image 2: (a) Original test image, (b) histogram equalization, (c) histogram equalization with background thresholding, (d) CLAHE ("adaphstheq()") with default setting by Matlab Image Processing Toolbox, Build. R12), (e) mean weight filtering, and (f) intensity correction maps by the mean weight filtering.

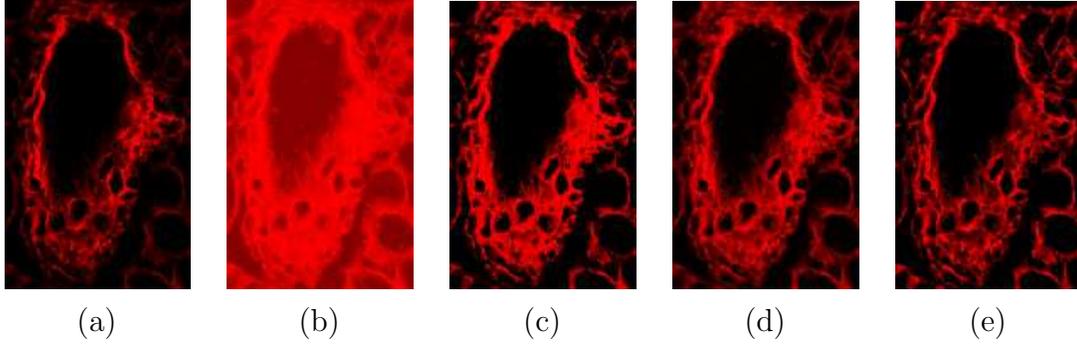


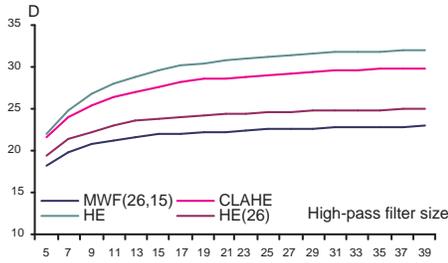
Figure 5.19: Sub-region of Intensity corrected images in Figure 5.18: (a) original image, (b) Histogram Equalization, (c) Histogram Equalization with background threshold, (d) CLAHE, and (e) Mean-weight filtering.

contrast C , (b) high frequency distortion D , (c) the low frequency intensity heterogeneity S (evaluated by the image entropy in low frequency domain) [98], and (d) the number of saturated pixels N . The metric N is defined in following equation.

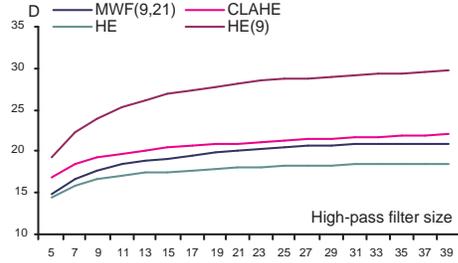
$$N = \sum_v^M \sum_u^N sat(I^{org}(u, v), I^{adj}(u, v)) \quad (5.9)$$

where $sat(I^{org}(u, v), I^{adj}(u, v)) = \begin{cases} 1 & \text{if } I^{org}(u, v) \neq l, I^{adj}(u, v) = l \\ 0 & \text{otherwise} \end{cases}$, I^{org} and I^{adj} are the original and intensity adjusted images, and l is the maximum intensity value of the original image, e.g. 255 in a byte image. For best image quality, it is desirable to achieve large C and small S , D , and N .

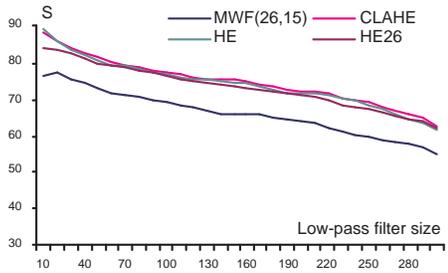
Comparative results: Figure 5.20 and Figure 5.21 present the results obtained by multiple intensity correction techniques, and evaluated with high frequency difference and low frequency intensity heterogeneity metrics D and S as a function of filter size. Based on these results, the mean-weight filtering method achieves consistently smaller (better) values (except Histogram Equalization) of (a) the high frequency difference D for the evaluated



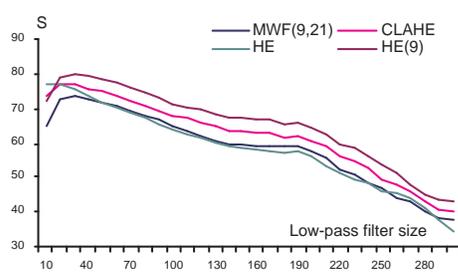
(a)



(b)

Figure 5.20: High frequency distortion D (a) for Figure 5.16 and (b) Figure 5.18.

(a)



(b)

Figure 5.21: Low frequency intensity heterogeneity S for the original test image 1 in Figure 5.16 (results are shown in the left graph denoted as (a)) and the original test image 2 in Figure 5.18 (results are shown in the right graph denoted as (b)).

range of filter sizes [3,39] (see Figure 5.20), and (b) low frequency intensity heterogeneity S for the evaluated filter sizes [10, 200] (see Figure 5.21). Although I included Histogram equalization method for comparison, the result shows that it is highly sensitive to image data and it demonstrates poor visual perceptual quality.

Table 5.1 and Table 5.2 numerically summarize the results with four quality metrics with two different images. In general, an optimal intensity correction technique has to meet multiple optimization criteria, for example, minimize S , D and N , and maximize C , as defined in the work. For both measured test images, Table 5.1 and Table 5.2 show that the mean-weight filtering achieves a normalized metric that is about 1.5 to 1.7 times larger than the second best performing technique.

Table 5.1: Evaluation results for the test image 1 shown in Figure 5.16 using the four metrics defined at the beginning of Section 3.2. The four techniques are represented by abbreviations, such as Histogram equalization (HE), HE with background threshold (HE (threshold)), Contrast Limited Adaptive Histogram Equalization (CLAHE), and Mean Weight Filtering MWF (threshold, filter width). To highlight performance results, the best result is shown as bold and the second best result is shown as italic.

| CLSM Image 1 | HE | HE (26) | CLAHE | MWF (26,15) |
|--------------|-------|--------------|--------------|--------------|
| C | 17.15 | 17.34 | <i>17.20</i> | 16.26 |
| S | 74.27 | <i>73.40</i> | 74.86 | 66.24 |
| D | 29.88 | <i>23.93</i> | 28.02 | 21.98 |
| N | 330 | <i>330</i> | 334 | 0 |

Table 5.2: Evaluation results for the test image 2 shown in Figure 5.18 using the four metrics defined at the beginning of Section 3.2. The four techniques are represented by abbreviations, such as Histogram equalization (HE), HE with background threshold (HE (threshold)), Contrast Limited Adaptive Histogram Equalization (CLAHE), and Mean Weight Filtering MWF (threshold, filter width). To highlight performance results, the best result is shown as bold and the second best result is shown as italic.

| CLSM Image 2 | HE | HE (9) | CLAHE | MWF (9,21) |
|--------------|--------------|--------------|------------|--------------|
| C | 8.10 | 13.39 | 9.49 | <i>11.43</i> |
| S | 58.36 | 65.24 | 61.88 | <i>58.38</i> |
| D | 17.63 | 27.27 | 20.60 | <i>19.47</i> |
| N | 181 | 181 | <i>105</i> | 0 |

5.5 Summary

This chapter addressed the problem of intensity correction of fluorescent confocal laser scanning microscope (CLSM) images. CLSM images are frequently used in medical domain for obtaining 3D information about specimen structures by imaging a set of 2D cross sections and performing 3D volume reconstruction afterwards. However, 2D images acquired from fluorescent CLSM images demonstrate significant intensity heterogeneity, for example, due to photo-bleaching and fluorescent attenuation in depth. I developed an intensity heterogeneity correction technique that (a) adjusts intensity heterogeneity of 2D images, (b) preserves fine structural details, and (c) enhances image contrast, by performing spatially adaptive mean-weight filtering. The solution is obtained by formulating an optimization problem, followed by filter design and automated selection of filtering parameters. The proposed filtering method is experimentally compared with several existing techniques by using four quality metrics, such as contrast, intensity heterogeneity (entropy) in low frequency domain, intensity distortion in high frequency domain, and saturation. Based on the experiments and the four quality metrics, the developed mean-weight filtering outperforms other intensity correction methods by at least a factor of 1.5 when applied to fluorescent CLSM images.

The experimental results obtained with synthetic images and measured images were discussed in the sections above. Here, I discuss two issues related to kernel size selection and quality metric normalization.

Kernel size selection: I provided a method for automatic selection of a kernel size that optimizes the *global* image saliency. However, there might be image cases with a mixture of sub-regions that require different kernel sizes per one image. If a user chooses to select a kernel size on his own, I would provide the following recommendations: (a) A large kernel tends to preserve the detail of rather large area, e.g., thick edge or spatial intensity heterogeneity in a feature region, and extremely large kernels correct minimally intensities in x-y

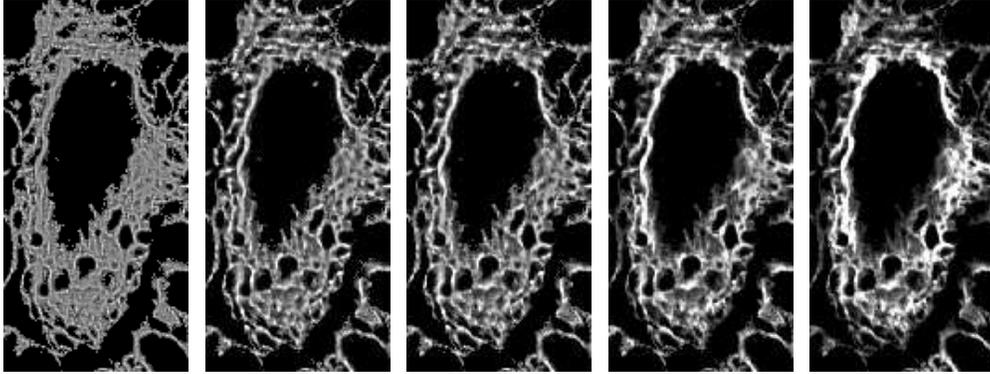


Figure 5.22: Mean weight filtering with different kernel sizes: From left, the kernel width is equals to 3,7,9,21,51.

plane. (b) A small kernel generates visually salient images by highlighting sharp intensity changes, e.g., small intensity discontinuities. However, extremely small kernels correct high frequency intensity change which is typically considered as edge gradient (and need to be preserved). (c) A kernel size could be selected based on edge thickness: for thin edges, a smaller kernel size is preferred since only high frequency component should be corrected. For thick edges, a larger kernel should be used since a low frequency component should be corrected while preserving a high frequency component. Figure 5.22 show the mean weight filtering results with different kernel sizes.

Quality metric normalization: One might desire to compare multiple intensity adjustment methods using a combined quality metric. While considering the metrics presented in Section 3.2, it might not be trivial to decide how to normalize and how to combine the metrics. It is possible to borrow the normalization approach from the data mining field [56] and normalize the values \bar{C} , \bar{S} , \bar{D} , \bar{N} to a fixed range, e.g., $[-1, 1]$ like in the case of artificial neural networks. The problem of combining multiple metrics could be approached by weighted sum of normalized metrics with the assigned weights according to application specific importance. For example, the results presented in Table 5.1 and Table 5.2 would lead to the following overall comparisons assuming equally important metrics in Table 5.3.

Table 5.3: Comparison using a combined metric formed by summing together normalized C , S , D and N metrics to $[-1, 1]$.

| | | | | |
|-----------------|-------|---------|-------|-------------|
| CLSM Image 1 | HE | HE (26) | CLAHE | MWF (26,15) |
| Combined metric | 0.225 | 0.483 | 0.275 | 0.750 |
| CLSM Image 2 | HE | HE (9) | CLAHE | MWF (9,21) |
| Combined metric | 0.500 | 0.250 | 0.465 | 0.859 |

Based on Table 5.3, the mean-weight filtering method would outperform all other intensity adjustment methods for the two test images. Other user-driven choices of techniques for normalization and parameters of metric combination (operators and weights) might be preferred in other application scenarios.

Chapter 6

CONCLUSIONS

This dissertation addressed the problem of three-dimensional volume reconstruction of image data obtained by fluorescent confocal laser scanning microscopes (CLSM). The main purpose of this dissertation is to provide a unified framework of three-dimensional volume reconstruction from CLSM data with performance evaluations and a registration decision support system. The challenges of the proposed problem arise from the characteristics of cellular level fluorescent CLSM data, such as extremely high morphological distortions (deformations) of imaged structures, and unpredictable spatial intensity heterogeneity, such as intensity variations both in lateral and axial plane. The main contribution of this dissertation is described in Section 1.1, such as the presentation of (1) the three-dimensional volume reconstruction framework, (2) automation procedures, (3) the registration decision support system, (4) registration accuracy evaluations, (5) sub-volume pre-processing techniques, (6) the automated feature matching technique, (7) web-based collaborative tools, (8) the fully-automated registration technique, and (9) the intensity correction technique.

In Chapter 2, I proposed a three-dimensional volume reconstruction framework and a registration decision support system. The main goal of the volume reconstruction is to obtain high resolution image frames by mosaicking (stitching together) spatial tiles that came from the same depth of a sub-volume, align sub-volumes (physical sections) from multiple cross sections, evaluate the accuracy of volume reconstruction using multiple techniques, and visualize the reconstructed volumes in a three-dimensional virtual reality environment for user-driven or quantitative analysis. The framework consists of three main steps, such as sub-volume pre-processing, automated alignment, and post-processing techniques. The

proposed framework is unique in terms of forming a complete system for medical inspection of three-dimensional volumes reconstructed from CLSM images. In addition, I proposed the problem of optimal registration decisions during the volume reconstruction. The registration decisions of interest included image spatial size (image sub-area size), transformation model (rigid or affine), invariant registration feature (intensity or morphology), and automation level (manual or semi-automated). I provided mechanisms for evaluating the tradeoffs of each registration decision in terms of 3D volume reconstruction accuracy, repeatability and computational requirements.

The sub-volume preprocessing step in Chapter 3 focus on techniques required for automated 3D volume reconstruction process, which consists of sub-volume intensity analysis, tile mosaicking, segmentation, and feature detection. Frame selection and mosaicking (manual, semi-automated, and fully automated) steps are performed regardless which registration methods the volume reconstruction uses, e.g., 2D or 3D information. In the case of using 2D information for registration, the selected frame from the sub-volume analysis process is further needed through entire registration process, such as 2D segmentation, feature matching, and intensity-based optimization. Segmentation and feature detection processes are performed based on the dimensionality of registration method, such as region segmentation followed by centroid computation in 2D case, and the volume segmentation followed by centroid trajectory computation in 3D case.

The sub-volume alignment techniques are proposed in Chapter 4. The main purpose of this step is to acquire an optimal registration transformation by performing feature matching using multiple automation levels followed by matching refinements in different optimality criteria. I also report accuracy evaluations for proposed registration methods including manual (pixel-based), semi-automatic (region centroid feature-based) and fully automatic (correlation-based) registration techniques by evaluating two groups of human subjects (experts and novices). The study demonstrates significant benefits of automation for 3D volume reconstruction in terms of achieved accuracy, consistency of results and performance time.

In addition I presented a prototype solution for 3D medical volume reconstruction using web services. For optimal sub-volume alignment, matching refinements are performed in two different senses: optional intensity-based optimization for 2D alignment and registration optimization based on trajectory fusion using 3D information. The intensity-based optimization showed improved alignment accuracy for two out of the three test image pairs. The registration optimization approaches using trajectory fusion is aimed at preserving morphological smoothness of medical structures inside of the reconstructed 3D volume by evaluating two metrics: morphological continuity at boundaries (end point discontinuity) and global smoothness (polynomial curve fitting residual).

In Chapter 5, I provided new image enhancement technique which corrects spatial intensity heterogeneity problems for enhanced visualization of fluorescent CLSM imagery. The developed technique, denoted as mean-weight filtering, enhances image stacks by adjusting intensity heterogeneity of image frames, preserving fine structural details, and enhances image contrast. The solution is obtained by formulating an optimization problem, followed by filter design and automated selection of filtering parameters. The proposed filtering method is experimentally compared with several existing techniques by using four quality metrics, such as contrast, intensity heterogeneity (entropy) in low frequency domain, intensity distortion in high frequency domain and saturation.

Provided methods have been researched and fully implemented into the Image to knowledge (I2K) mainframe of the National Center for Supercomputing Applications (NCSA). In particular, the web-enabled software packages provide extremely useful volume reconstruction tools to end-users that minimize laborious and time intensive data preparation, has great impact on the researchers who would make registration decision on CLSM data, and provide a comprehensive educational resources of computerized three-dimensional volume reconstruction process and analysis to a number of medical researchers and experts in clinical domain.

The whole is greater than the sum of its parts - Max Wertheimer

The whole is equal to the sum of its parts - Euclid

References

- [1] J. Alkemper and P.W. Voorhees, “Quantitative serial sectioning analysis.” *J. Microsc.* 201, 388-394, 2001.
- [2] N.M. Alpert, J.F. Bradshaw, D. Kennedy, and J.A. Coreia, “The principal axis transformation - a method for image registration” *J. Nucl. Med.* 31(10):1717-1722, 1990.
- [3] I. Altintas, C. Berkley, E. Jaeger, M. Jones, B. Ludcher, and S. Mock, “system demonstration, Kepler: An Extensible System for Design and Execution of Scientific Workflows, ” *16th Intl. Conf. on Scientific and Statistical Database Management (SSDBM'04)*, 21-23 June 2004, Santorini Island, Greece, 2004.
- [4] L.S. Arun, T. S. Huang, and S. D. Blostein “Least-squares fitting of two 3-D point sets” *IEEE Trans. Pattern Anal. Mach. Intell.*, 9:698-700, 1987.
- [5] D.E. Atkins, K.K. Droegemeier, S.I. Feldman, H. Garcia-Molina, P. Messina, D.G. Messerschmitt, J. Ostriker, and M.H. Wright, “Revolutionizing Science and Engineering through Cyberinfrastructure: ” *Report of the National Science Foundation, Blue ribbon advisory panel on Cyberinfrastructure* , April 19, 2002.
- [6] R. Bajcsy and S. Kovacic, “Multiresolution elastic matching” *Comp Vis Graph Image Proc*, 46:1-21, 1989.
- [7] P. Bajcsy and P. Groves, “Methodology For Hyperspectral Band Selection”, *Photo. Eng. and Remote Sensing J.*, 70, 793-802, 2004.
- [8] P. Bajcsy, R. Kooper, D. Clutter, S-C Lee, and P. Ferak, “Image to Knowledge user manual,” *ALG-2005-005.*, 2005.
- [9] P. Bajcsy, S-C. Lee, and D. Clutter, “Supporting registration decisions during 3D medical volume reconstructions.” *SPIE International Symposium in Medical Imaging*, 2006.
- [10] P. Bajcsy, S-C. Lee, A. Lin, and R. Folberg, “3D Volume Reconstruction of Extracellular Matrix Proteins in Uveal Melanoma from Fluorescent Confocal Laser Scanning Microscope Images” *in Journal of Microscopy*, Blackwell Synergy, 2006.
- [11] D. Benson, J. Bryan, A. Plant, A. Gotto, and L. Smith, “Digital imaging fluorescence microscopy: Spatial heterogeneity of photobleaching rate constants in individual cells.” *J. Cell Biol.* 100, 1309-1323, 1985.

- [12] P.J. Besl and N.D. McKay, "A method for registration of 3-D shapes " *IEEE Trans. Pattern Anal. Mach. Intell.* 14 239-56, 1992
- [13] G. Bock, M. Hilchenbach, K. Schauenstein, and G. Wick, "Photometric analysis of antifading reagents with laser and conventional illumination sources," *J. Histochem. and Cytochem.* 33(7):699-705, 1985.
- [14] F. Bolton, *Pure Corba*, Sams, 2001
- [15] H. Bornfleth, K. Satzler, R. Eils, and C. Cremer, "High-precision distance measurements and volume-conserving segmentation of objects near and below the resolution limit in three-dimensional confocal fluorescence microscopy. *J. Microsc.* 189, 118-136, 1998.
- [16] L.G. Brown, "A Survey of Image Registration Techniques," *ACM Computing Surveys*, 1992.
- [17] P.J. Burt and E.H. Adelson, "A multiresolution spline with application to image mosaics," *ACM Trans. Graphics* 2, 217-236 (1983).
- [18] P. Cachier et al., "Multisubject Non-rigid Registration of Brain MRI Using Intensity and Geometric Features. *LNCS* 2208, 734-742, 2001.
- [19] V. Cantoni and S. Levialdi, "Pyramidal Systems for Computer Vision., " *NATO ASI Series, Series F: Computer and Systems Sciences*, 25., Springer-Verlag, New York, 1986
- [20] H-M. Chan and A.C.S. Chung, "Efficient 3D-3D Vascular Registration Based on Multiple Orthogonal 2D Projections" *Second International Workshop on Biomedical Image Registration, 2003 (WBIR'03)*, pp.301 - 310.
- [21] X. Chen, A.J. Maniotis, D. Majumdar, and R. Folberg, "Uveal Melanoma Cell Staining for CD34 and Assessment of Tumor Vascularity." *Invest. Ophthalm. Vis. Sci.* 43, 2533-2539, 2002.
- [22] L.D. Cohen and I. Cohen, "Deformable models for 3D medical images using finite elements and balloons," *IEEE Computer Vision and Pattern Recognition*, 1992.
- [23] C.L. Collins, J.H. Ideker, and K.E. Kurtis "Laser scanning confocal microscopy for in situ monitoring of alkali-silica reaction," *Journal of Microscopy*, 213 (2): 149, 2004.
- [24] L. Dorst, "First order error propagation of the procrustes method for 3D attitude estimation. " *IEEE Trans. Pattern Anal. Mach. Intell.* 27, 221- 229, 2005.
- [25] R. Duda, P. Hart, and D. Stork, "Pattern classification (Second Edition), " 654, *Wiley-Interscience*, New York, N.Y, 2001.
- [26] P. Elsen, E. Pol, and M. Viergever, "Medical image matching - A review with classification," *IEEE Eng, Med. Biol.* 26-39 (1993)

- [27] L.B. Engelstad, H.J. Meyers, W.A. Hanson, J.W. O'Connell, R.C. Taylor, et al. "Information extraction from multi-modality medical imaging" *In Proc SPIE* Vol. 902-28, SPIE Press, Bellingham, WA, 1988
- [28] A. Entwisle and M. Noble, "The use of Lucifer yellow, bodipy, FITC, TRITC, RITC and Texas red for dual immunofluorescence visualized with a confocal scanning laser microscope" *J. Microscop., Oxford* 168, 219-238
- [29] R. Eustice, O. Pizarro, H. Singh, and J. Howland, "UWIT: Underwater Image Toolbox for optical image processing and mosaicking in MATLAB," *in Proceedings of IEEE International Symposium on Underwater Technology* (IEEE Press, Piscataway, N.J., 2002), pp. 141-145.
- [30] A.C. Evans, S. Marrett, and T.M. Peters, "Anatomical-functional correlative analysis of the human brain using three-dimensional imaging systems" *Proceedings of the International Society of Optical Engineering (SPIE): Medical Imaging III*: 264-274, (1989)
- [31] A.C. Evans, T. Peters, L. Collins, C. Henri, S. Marrett, B. Pike, and W. Dai, "3-D correlative imaging and segmentation of cerebral anatomy, function and vasculature" *IEEE Trans Eng Med Biol* 12(3): 297-98, 1990.
- [32] A.C. Evans, S. Marrett, J. Torrescorzo, S. Ku, and L. Collins, "MRI-PET correlation in three dimensions using a volume of interest (VOI) atlas." *Journal of cerebral blood flow and metabolism*, 11, A69-A78, 1991.
- [33] J. Fang and T.S. Huang, "Some Experiments on Estimating the 3-D Motion Parameters of a Rigid Body from Two Consecutive Image Frames." *IEEE Trans. Pattern Anal. Machine Intell*, 6, 545-554, 1984.
- [34] D. Fedorov, L.M.G. Fonseca, C. Kenney, and B.S. Manjunath, "Automatic Registration and Mosaicking System for Remotely Sensed Imagery" *SPIE 9th International Symposium on Remote Sensing*, Crete, Greece, 2002
- [35] O.D. Faugeras and M. Herbert, "The representation, recognition, and locating of 3-D objects." *Ont. J. Robotics Res.*, 5:27-52, 1986
- [36] M. A. Fischler and R. C. Bolles, "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography." *Comm. of the ACM*, Vol 24, pp 381-395, 1981.
- [37] J.M. Fitzpatrick, D.L.G. Hill, Y. Shyr, J. West, C. Studholme, and C. R. Maurer, "Visual assessment of the accuracy of retrospective registration of MR and CT images of the brain," *IEEE Trans. Med. Imag.*, vol.17, pp.571-585, Aug.1998.
- [38] J. Flusser and T. Suk, "A moment-based approach to registration of images with affine geometric distortion," *IEEE Trans. Geoscience and Remote Sensing* 32: 382-387, 1994.
- [39] R. Folberg and A.J. Maniotis, "Vasculogenic mimicry." *APMIS* 112, 508-525, 2004

- [40] W. Forstner, "Reliability Analysis of Parameter Estimation in Linear Models with Applications to Mensuration Problems in Computer Vision. *CVGIP*, 273-310, 1987.
- [41] D.A. Forsyth and J. Ponce, "Computer Vision: A Modern Approach" Prentice Hall, 2003
- [42] A. Franklin and W. Filion, "A new technique for retarding fading of fluorescence: DPX-BME." *Stain Technol.* 60, 125-153, 1985.
- [43] K.P. Gall and L.J. Verhey, "Computer-assisted positioning of radiotherapy patients using implanted radioopaque fiducials." *Medical physics*, 20(4), 1152-1159, 1993.
- [44] R. Garcia, J. Batlle, X. Cufi, and J. Amat, "Positioning an underwater vehicle through image mosaicking," in *Proceedings of IEEE International Conference on Robotics and Automation* (IEEE Press, Piscataway, N.J., 2001), Part 3, pp. 2779-2784.
- [45] P. Gerlot, Y. Bizais, "Image registration: a review and a strategy for medical applications" in: C.N. de Graaf, M.A. Viergever (Eds.), *Information Processing in Medical Imaging*, Plenum Press, New York, 1988, pp. 81-89.
- [46] P. Gerlot-Chiron and Y. Biza, "Definition and evaluation of a surface overlap criterion for medical image registrationInformation " *Processing in Medical Imaging* ed: D A Ortendahl and J Llacer (New York: Wiley-Liss) pp 429-42, 1991
- [47] P. Gerlot-Chiron and Y. Bizais, "Registration of Multimodality Medical Images Using a Region Overlap Criterion" *GMIP(54)*, No. 5, September 1992, pp. 396-406.
- [48] R.C. Gonzalez and R.E. Woods, "Digital image processing (Second Edition)," *Prentice-Hall Inc.*, 793, New Jersey, 2002.
- [49] A. Goshtasby, G. C. Stockman, and C. V. Page, "A region-based approach to digital image registration with subpixel accuracy," *IEEE Trans. Geoscience and Remote Sensing* 24: 390-399, 1986.
- [50] A. Goshtasby, "Registration of images with geometric distortions," *IEEE Transactions on Geoscience and Remote Sensing* 26: 60-4, 1988.
- [51] C.H. Goulden, "Methods of Statistical Analysis", *2nd ed.* New York: Wiley, pp. 70-55, 1956.
- [52] B.F. Green, "The orthogonal approximation of an oblique structure in factor analysis " *Psychometrika* 17 429-40, 1952
- [53] W.E.L Grimson, "Object Recognition by Computer - The role of Geometric Constraints. *MIT Press*, 1992.
- [54] A. Gueziec and N. Ayache, "Smoothing and matching of 3-D space curves " *Int. J. Comput. Vision* 12 79-104, 1994

- [55] A. Gueziec, X. Pennec, and N. Ayache, "Medical image registration using geometric hashing" *IEEE Comput. Sci. Eng.* 4 29-41,1997
- [56] J. Han and M. Kamber, "Data Mining: Concepts and Techniques", Morgan Kaufmann Publishers, 550, 2001.
- [57] M. Hansen, P. Anandan, K. Dana, G. van der Wal, and P. Burt, "Real-time scene stabilization and mosaic construction," *in Proceedings of IEEE Workshop on Applications of Computer Vision* (IEEE Press, Piscataway, N.J., 1994), pp. 54-62.
- [58] R.M. Haralick and L.G. Shapiro. "Image segmentation techniques." *Comput. Vis. Graph. Im. Proc.*, 29:100-132, 1985.
- [59] T. Hartkens, et al., "Evaluation of 3-D operators for the detection of anatomical point landmarks in MR and CT images. *Comp. Vis. Image Understanding*, 86, 118 - 136, 2002.
- [60] G. Hermosillo and O. Faugeras "Variational methods for multi-modal image matching" , *in International Journal of Computer Vision*, 50 (3): 329 ?343, 2002.
- [61] D.L.G. Hill, D.J. Hawkes, J.E. Crossman, et al. "Registration of MR and CT images for skull base surgery using point-like anatomical features." *Br J Radiology* 64:1030-1035 1991
- [62] D.L. Hill, S.E. Green, J.E. Crossman, D.J. Hawkes, G.P. Robinson, C.F. Ruff, T.C. Cox, A.J. Strong, and M.J. Gleeson, "Visualization of multi-modal images for the planning of skull base surgery" *In Proceedings of the Second Conference on Visualization in Biomedical Computing*, Chapel Hill, NC, 1992, pp 564-573.
- [63] D.L.G. Hill, D.J. Hawkes, Z. Hussain, et al. "Accurate combination of CT and MR data of the head: validation and applications in surgical and therapy planning. " *Comput Med Imaging* 1993;17:357-362
- [64] D. Hill, P. Batchelor, M. Holden, and D. Hawkes, "Medical Image Registration," *Phys. Med. Biol.* 46 R1-R45, 2001.
- [65] B.K.P. Horn, "Closed-form solution of absolute orientation using unit quaternions" *J. Opt. Soc. Amer. A*, 4:629-642, 1987
- [66] B.K.P. Horn, H.M. Hilden, and S. Negahdaripour, "Closed-form solution of absolute orientation using orthonormal matrices" *J. Opt. Soc. Amer. A*, 5:1127-1135, 1988
- [67] J. Hu, A. Razdan, G. Nielson, G. Farin, D. Baluch, and D. Capco, "Volumetric Segmentation Using Weibull E-SD Fields," *IEEE Trans. on Vis. and Comp. Graphics*, 9, 320-328, 2003.
- [68] J.R. Hurley and R.B. Cattell, "The Procrustes program: producing direct rotation to test a hypothesized factor structure " *Behav. Sci.* 7 258-62, 1962

- [69] T. Irinopoulo, J. Vassy, M. Beil, P. Nicolopoulo, D. Encaoua, and J. Rigaut, "3-D DNA image cytometry by confocal scanning laser microscopy in thick tissue blocks of prostatic lesions." *Cytometry*, 27, 99-105, 1997.
- [70] J. Itoh, K. Yasumura, K. Ogawa, K. Kawai, A. Serizawa, Y. Yamamoto, and Y.R. Osamura, "Three-dimensional (3D) imaging of tumor angiogenesis and its inhibition: Evaluation of tumor vascular-targeting agent efficacy in the DMBA-induced rat breast cancer model by confocal laser scanning microscopy (CLSM). *Acta Histochem., Cytochem.*, Technical advancement, 36, 27-36, 2003.
- [71] G.D. Johnson and G.M. De C Nogueira Araujo, "A simple method for reducing the fading of immunofluorescence during microscopy," *J. Immunol. Methods.* 43:349, 1981.
- [72] G.D. Johnson, R.S. Davidson, K.C. McNamee, G. Russel, D. Goodwin, and E.J. Holborow, "Fading of Immunofluorescence during microscopy: a study of the phenomenon and its remedy," *J. Immunol. Methods.* 55:231, 1982.
- [73] M. Jungke, W. von Seelen, G. Bielke, S. Meindl, et al, "A system for the diagnostic use of tissue characterizing parameters in NMR-tomography." *proc. of Information Processing in Medical Imaging, IPMI'87*, 39:471-481, 1987.
- [74] J. Kang, I. Cohen, and G. Medioni, "Continuous multi-view tracking using tensor voting" *IEEE Workshop on Motion and Video Computing*, Orlando, Florida. 5-6 December, 2002.
- [75] C. Kervrann, D. Legland, and L. Pardini, "Robust incremental compensation of the light attenuation with depth in 3D fluorescence microscopy. *J. Microsc.*, 214, 297-314, 2004.
- [76] R. Kooper, A. Shirk, S-C. Lee, A. Lin, R. Folberg, and P. Bajcsy, "3D Medical Volume Reconstruction Using Web Services" in *IEEE International Conference on Web Services*, Orlando, June 2005.
- [77] D. Koppel, C. Carlson, and H. Smilowiz, "Analysis of heterogeneous fluorescence photobleaching by video kinetics imaging: The method of culminants," *J Microsc* 155:199-206, 1989
- [78] M.A. Koschat and D.F. Swayne, "A weighted Procrustes criterion " *Psychometrika* 56 229-39, 1991
- [79] D.Kotz, "Nuclear medicine in the 21st century: Integration with other specialties," *J. Nucl. Med.* ,vol.40,no.7,pp.13N-25N,July1999.
- [80] S. Kovacic, J.C. Gee, W.S.L. Ching, N. Rivich, and R. Bajcsy, "Three-dimensional registration of PET and CT images" In *Ann. Int. Conf. IEEE Comp. Med. Biol. Soc.* Vol 11, Los Alamitos, CA, 548-549,1989

- [81] K.D. Krenik, G.M. Kephart, K.P. Offord, S.L. Dunnette, and G.J. Gleich, "Comparison of antifading reagents used in immunofluorescence," *J. Immunol. Methods.* 117:91-7, 1989.
- [82] R. Kwok, J. C. Curlander, and S. Pang, "An automated system for mosaicking spaceborne SAR imagery," *Int. J. Remote Sens.* 11, 209-223 (1990).
- [83] C. J. Lada, D. L. DePoy, K. M. Merrill, and I. Gatley, "Infrared images of M17," *Astron. J.* 374, 533-539 (1991).
- [84] S.M. Larie and S.S. Abukmeil. "Brain abnormality in schizophrenia: a systematic and quantitative review of volumetric magnetic resonance imaging studies. *J. Psych.*, 172:110-120, 1998.
- [85] S-C. Lee and P. Bajcsy, "Feature based registration of fluorescent LSCM imagery using region centroids," *SPIE International Symposium in Medical Imaging*, 5747, pp. 170-181, 2005
- [86] S-C. Lee and P. Bajcsy, "Automated feature-based alignment for 3D volume reconstruction of CLSM imagery." *SPIE International Symposium in Medical Imaging*, 2006
- [87] S-C. Lee and P. Bajcsy, "Intensity Correction of Fluorescent Confocal Laser Scanning Microscope Images by Mean-Weight Filtering." in *Journal of Microscopy*, Blackwell Synergy, 2006.
- [88] S-C. Lee, P. Bajcsy, A. Lin, and R. Folberg, "Accuracy Evaluation of 3D Volume Reconstruction Techniques at Different Automation Level" in *EURASIP JASP, Performance Evaluation in Image Processing*, Hindawi publishing, 2006.
- [89] K. Leung Lam, R.K. ten Haken, D.L. McShan, and A.F. Thornton, "Automated determination of patient setup errors in radiation therapy using spherical radio-opaque markers." *Medical physics*, 20(4), 1145-115, 1993.
- [90] N. Leavitt, "Are Web Services Finally ready to Deliver?," *IEEE Computer, Computer Society*, pp. 14-18, 2004
- [91] S. Li, C.A. Pelizzari, and G.T.Y. Chen, "Unfolding patient motion with biplane radiographs." *Medical physics*, 21(9), 1427-1433, 1994.
- [92] H. Li, B.S. Manjunath, and S.K. Mitra, "A contour-based approach to multisensor image registration," *IEEE Trans. Image Processing* 4:320-334, 1995.
- [93] A. Liljeborg, M. Czader, and A. Porwit, "A method to compensate for light attenuation with depth in 3D DNA image cytometry using a confocal scanning laser microscope." *J. Microsc.*, 177, 108-114, 1995.
- [94] H. Liu, X. Lin, and M. Li, "Modeling Response Time of SOAP over Http," *IEEE International Conference on Web Services (ICWS 2005)*, 2005

- [95] A. Longin, C. Souchier, M. French, and P.A. Bryon, "Comparison of anti-fading agents used in fluorescence microscopy: image analysis and laser confocal microscopy study," *J. Histochem. and Cytochem.* 41(12):1833-40, 1993 Dec.
- [96] B. Ludscher, I. Altintas, C. Berkley, D. Higgins, E. Jaeger-Frank, M. Jones, E. Lee, J., Tao, and Y. Zhao, "Scientific Workflow Management and the Kepler System, Concurrency and Computation: Practice and Experience," *Special Issue on Scientific Workflows*, 2005
- [97] J.B.A. Maintz and M.A. Viergever, "A survey of medical image registration," *Medical Image Analysis*, 2(1):1-36, 1998.
- [98] J. Mangin, "Entropy minimization for automatic correction of intensity nonuniformity" *Math. Method in Biomed. Image Analysis (MMBIA)*, 162-169, 2000.
- [99] G.Q. Maguire, M. Noz, H. Rusinek, J. Jaeger, E.L. Kramer, J.J. Sanger, and G. Smith, "Graphics applied to medical image registration." *IEEE Computer graphics and applications*, 11(2), 20-28, 1991.
- [100] R.T. Malison, E.G. Miller, R. Greene, G. McCarthy, D.S. Charney, and R.B. Innis, "Computer-assisted coregistration of multislice SPECT and MR brain images by fixed external fiducials." *Journal of computer assisted tomography*, 17(6), 952-960, 1993.
- [101] V.R. Mandava, J.M. Fitzpatrick, and D.R. Pickens III, "Adaptive search space scaling in digital image registration" *IEEE Trans. Med. Imaging*, 8:251-262, 1989
- [102] A. Maniotis, X. Chen, C. Garcia, P. DeChristopher, D. Wu, J. Pe'er, and R. Folberg, "Control of melanoma morphogenesis, endothelial survival, and perfusion by extracellular matrix." *Lab. Investigation* 82, 1031-1043, 2002.
- [103] C.R. Maurer and J.M. Fitzpatrick, "A review of medical image registration," *In Interactive Image-Guided Neurosurgery*, R.J.Maciunas, Ed.Parkridge, IL:Amer.Assoc. NeurologicalSurgeons,1993, pp.7-44.
- [104] K. Montgomery and M.D. Ross, "Non-fiducial, shape-based registration of biological tissue," *3D Microscopy: Image Acquisition and Processing III. SPIE Electronic Imaging*, San Jose, CA, 1996
- [105] S. Negahdaripour, X. Xu, A. Khemene, and Z. Awan, "3-D motion and depth estimation from sea-floor images for mosaic-based station-keeping and navigation of ROV's/AUV's and high-resolution sea-floor mapping," *in Proceedings of IEEE Workshop on Autonomous Underwater Vehicles* (IEEE Press, Piscataway, N.J., 1998), pp. 191-200.
- [106] A. Nocito, J. Kononen, O. P. Kallioniemi, and G. Sauter, "Tissue microarrays (TMAs) for high-throughput molecular pathology research," *Int.J.Cancer* 94 (1):1-5, 2001.
- [107] P. V. Oostveld, F. Verhaegen, and K. Messen, "Heterogenous photobleaching in confocal microscopy caused by differences in refractive index and excitation mode," *Cytometry* 32, 137-146, 1998

- [108] N.R. Pal and S.K. Pal. "A review on image segmentation techniques." *Patt. Rec.*, 26:1277-1294, 1993.
- [109] C. Papadimitriou, C. Yapijakis, and P. Davaki, "Use of truncated pyramid representation methodology in three-dimensional reconstruction: an example." *J. Microsc.* 214, 70-75, 2004.
- [110] S. Parastatidis and J. Webber, "Assessing the Risk and Value of Adopting Emerging and Unstable Web Services Specifications," *Services Computing, 2004 IEEE International Conference on (SCC'04).*, 2004
- [111] J.B. Pawley, Handbook of Biological Confocal Microscopy, *Plenum*, New York, 1990.
- [112] C.A. Pelizzari, G.T.Y. Chen, D.R. Spelbring, R.R. Weichselbaum, and C-T. Chen, "Accurate three-dimensional registration of CT, PET, and/or MR images of the brain" *J. Comput. Assist. Tomogr.* 13 20-6, 1989
- [113] X. Pennec and J-P Thirion, "A framework for uncertainty and validation of 3-D registration methods based on points and frames" *Int. J. Comput. Vision* 25 203-29, 1997
- [114] X. Pennec, "Evaluation of the Uncertainty in Various Registration problems." *Meth. of Eval. Comp. Med. Imaging*, 15-16, 2001.
- [115] L.D. Pham, C. Xu, and J.L. Prince, "a survey of current methods in medical image segmentation." *Annual Review of Biomedical Engineering* Vol. 2: 315-337
- [116] E. Pisano, S. Zong, M. Hemminger, M. De Luca, R. Johnsoton, K. Muller, M. Braeuning, and S. Pizer, "Contrast Limited Adaptive Histogram Equalization Image Processing to Improve the Detection of Simulated Spiculations in Dense Mammograms." *J. Digital Imaging*, 11, 193-200, 1998.
- [117] S.M. Pizer, J.B. Zimmerman, and E. Stabb, "Adaptive grey level assignment in CT scan display." *J. Comp. Assist. Tomography*, 8, 300-305, 1984.
- [118] S.M. Pizer, D.S. Fritsch, P.A. Yushkevitch, V.E. Johnson, and E.L. Chaney, "Segmentation, Registration, and Measurement of Shape Variation via Image Object Shape." *IEEE Trans. Med. Imaging*, 18, 851-865, 1999.
- [119] W. K. Pratt, "Correlation techniques for image registration," *IEEE Trans. on Aerospace Engineering Systems*, 10:353-358, 1974.
- [120] J. Pritchard, "COM and CORBA(R) Side by Side: Architectures, Strategies, and Implementations," *Addison-Wesley Professional.*, 1999
- [121] C. Rao, "Matrix approximations and reduction of dimensionality in multivariate statistical analysis" *Multivariate Analysis* (Amsterdam: North Holland), 1980

- [122] R.J. Radke, S. Andra, O. Al-Kofahi, and B. Roysam, "Image Change Detection Algorithms: A Systematic Survey. *IEEE Trans. Image Processing*, 14, 294-307, 2005.
- [123] E. M. Reynoso, G. M. Dubner, W. M. Goss, and E. M. Arnal, "VLA observations of neutral hydrogen in the direction of Puppis A," *Astron. J.* 110, 318-324 (1995).
- [124] J. Rigaut, J. Vassy, P. Hetlin, F. Duigou, E. Masson, D. Briame, J. Foucrier, S. Carvajal-Gonzales, A. Downs, and A. Mandard, "Three-dimensional DNA image cytometry by confocal scanning laser microscopy in thick tissue blocks," *Cytometry* 12:511-524, 1991
- [125] I. Rigoutsos and R. Hummel, "Distributed Bayesian Object Recognition. *Int. Conf. Comput. Vis. and Pat. Recog.*, 180-186, 1993.
- [126] J. Roerdink and M. Bakker, "An FFT-based method for attenuation correction in fluorescence confocal microscopy." *J. Microsc.*, 169, 3-14, 1993.
- [127] K. Rodenacker, P. Aubele, M. Hutzler, and P. Adiga, "Groping for quantitative digital 3-D image analysis: an approach to quantitative fluorescence in situ hybridization in thick tissue sections of prostate carcinoma." *Anal. Cell. Pathol.*, 15, 19-29, 1997.
- [128] A. Rosenfeld, "Multiresolution Image Processing and Analysis., " *Springer-Verlag*, New York, 1984
- [129] J.C. Russ, "The image processing handbook (Third Edition), " *CRC Press LLC*, 1999.
- [130] L.M. Sanchez-Brea and E. Bernabeu, "On the standard deviation in CCD cameras: a variogram-based technique for non-uniform images. " *J. Electronic Imaging*, 11, 121-126, 2002.
- [131] P.H. Schonemann, "A generalized solution of the orthogonal Procrustes problem " *Psychometrika* 31 1-10, 1966
- [132] C. Schmid, R. Mohr, and C. Bauckhage, "Evaluation of Interest Point Detectors. *Int. J. Computer Vision*, 37, 151 - 172, 2000.
- [133] J. SpaceK, "Three-dimensional reconstruction of astroglia and oligodendroglia cells. " *Anat.* 112, 430-442, 1971.
- [134] J.A. Stark and W.J. Fitzgerald, "An alternative algorithm for adaptive histogram equalization. " *Compu. Vis. Graph. Image Process.* 56, 180-185, 1996.
- [135] J.K. Stevens and J. Trogadis, "Three-dimensional volume investigation of serial confocal data sets" *Proc. Int. Congr. Electron Microsc.*, 12th, 1990, 375-380
- [136] J. Stevens, "Introduction to Confocal Three-Dimensional Volume Investigation", in three-dimensional confocal microscopy: volume investigation of biological systems " *Academic press*, pp. 3-27, 1994

- [137] M. Styner, C. Brechbuhler, G. Szekely, and G. Gerig, "Parametric estimate of intensity inhomogeneities applied to MRI." *IEEE Trans Med Imaging*, 19, 153-65, 2000.
- [138] U. Tauer and O. Hills, "Confocal Spectrophotometry," in *Sci. and Tech. Info., Sp. issue :Confocal Microscopy*, CDR 4, 15-27, 2000.
- [139] S. Thirumalai and N. Ahuja, "Detection and Segmentation of Feature Trajectories in Multiple, Discontinuous Motion Image Sequences" *Proc. DARPA Image Understanding Workshop*, Washington, April 19-21, 1993, 621-628.
- [140] M.A. Oghabian and A. Todd-Pokropek, "Registration of brain images by a multi-resolution sequential method" *In Information Processing in Medical Imaging*, Springer-Verlag, Berlin, 165-174, 1991
- [141] K.D. Toennies, G.T. Herman, and J.K. Udupa, "Surface registration for the segmentation of implanted bone grafts" *In Computer Assisted Radiology*, Springer-Verlag, Berlin, 381-385, 1989
- [142] K.D. Toennies, J.K. Udupa, G.T. Herman, I.L. Wornom III, and S.R. Buchman, "Registration of 3D objects and surfaces" *IEEE Comp Graph Appl* 10(3): 52-62, 1990
- [143] M. Touhy, et al., "Computer-assisted three-dimensional reconstruction technology in plant cell image analysis: applications of interactive computer graphics," *Journal of Microscopy* 147: 83-8, 1987.
- [144] J.M. Uson, S.P. Boughn, and J.R. Kuhn, "The central galaxy in Abell 2029: an old supergiant," *Science* 250, 539-540 (1990).
- [145] P. Viola and W.M. Wells, "Alignment by maximization of mutual information," in *proc. of the Fifth International Conference on Computer Vision*, pp 16, 1995
- [146] T. Visser, F. Groen, and G. Brakenhoff, "Absorption and scattering correction in fluorescence confocal microscopy." *J. Microsc.*, 163, 189-200, 1991.
- [147] B. Wang, C. Toro, T.A. Zeffiro, and M. Hallett, "Head surface digitization and registration: a method for mapping positions on the head onto magnetic resonance images. " *Brain topography*, 6(3), 185-192, 1994.
- [148] E. Weisstein, "Least Squares Fitting–Exponential." *MathWorld* -A Wolfram Web Resource. <http://mathworld.wolfram.com/LeastSquaresFittingExponential.html>.
- [149] M. Welge, W.H. Hsu, L.S. Auvil, C. Bushell, J. Martirano, T.M. Redman, and D. Tchong, "Data to Knowledge (D2K): A Rapid Application Development Environment for Knowledge Discovery in Database, " *Technical Report*, National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign, Champaign, 1999
- [150] J. Weng, T. S. Huang, and N. Ahuja, "Motion and Structure from Image Sequences", Springer-Verlag, 1993.

- [151] J. West, et. al. "Comparison and evaluation of retrospective intermodality brain image registration techniques," *J. Comput. Assist. Tomogr.*, vol.21, pp.554-566, 1997.
- [152] J. West, et. al., "Retrospective intermodality registration techniques for images of the head: Surface-based versus volume-based," *IEEE Trans. Med. Imag.*, vol.18, pp.144-150, Feb. 1999.
- [153] W.M. Trochim, "The Research Methods Knowledge Base, 2e", Cornell University, 2001.
- [154] B.D. Wilson, "GENESIS SciFlo: Enabling Multi-Instrument Atmospheric Science Using Grid Workflows, " *poster SF31A-0716 0800h at American Geophysical Union (AGU) Fall Meeting Special Focus: Advances in Data Acquisition, Management, Analysis and Display*, pp. 13-17, 2003
- [155] D. Wolin, K. Johnson, and Y. Kipma, "The Importance of Objective Analysis in Image Quality Evaluation. *Imaging Sic. Tech.* (IS and T) 603-606, 1998 .
- [156] A.J. Worth, N. Makris, V.S. Caviness, and D.N. Kennedy, "Neuroanatomical segmentation in MRI: technological objectives. " *Int. J. Patt. Rec. Art. Intel.*, 11:1161-1187, 1997.
- [157] J. Wu, B. Rajwa, D.L. Filmer, C.M. Hoffmann, B. Yuan, C. Chiang, J. Sturgis, and J.P. Robinson, "Automated quantification and reconstruction of collagen matrix from 3D confocal datasets. " *J. Microsc.* 210, 158-165, 2003.
- [158] Y. Caspi and M. Irani, "A Step towards Sequence-to-Sequence Alignment" *2000 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'00)* - Volume 2, p. 2682, 2000.
- [159] T. Yunck, B. Wilson, A. Braverman, E. Dobinson, and E. Fetzer, " GENESIS: The General Earth Science Investigation Suite," *The fourth annual NASA's Earth Technology Conference.*, 2004
- [160] Z. Zhang, "Iterative Point Matching for Registration of Free-Form Curves and Surfaces. *Int. J. Comp. Vision*, 13, 119-152, 1994.
- [161] A.P. Zijdenbos and B.M. Dawant, "Brain segmentation and white matter lesion detection in MR images." *Critical Reviews in Biomedical Engineering*, 22:401-465, 1994.
- [162] B. Zitova and J. Flusser, "Image Registration Methods: A Survey" *Image and Vision Computing*, 21:977-1000, 2003.

Vita

SANG-CHUL LEE

Data, Room 1008, NCSA, MC-257
1205 W. Clark St
Urbana, IL 61801
Phone: (217)244-2915, Fax: (217)244-7396
slee40@uiuc.edu

EDUCATION

Ph.D. University of Illinois at Urbana-Champaign, Urbana, Illinois, May 2006
Department of Computer Science
Dissertation: 3D Volume Reconstruction from Fluorescent Confocal Laser Scanning Microscopy Imagery
Advisors: Dr. Peter Bajcsy and Prof. Narendra Ahuja

M.S. University of Illinois at Urbana-Champaign, Urbana, Illinois, August 2003
Department of Computer Science
Dissertation: Landmark-based Exploration, Mapping, and Navigation by Free-region Sweeping
Advisor: Prof. David J. Kriegman

B.E. Inha University, In-cheon, Korea, January 1998
Department of Computer Science and Engineering

RESEARCH EXPERIENCE

National Institutes of Health (NIH) project: 3D Volume Reconstruction from Confocal Laser Scanning Microscopy Imagery (joint work with department of pathology, University of Illinois at Chicago), 08/03-present.

- Automated 3D volume reconstruction for highly deformable stacks of images.
- Accuracy evaluation for 3D volume reconstruction from fluorescent confocal scanning laser microscope (CSLM) imagery for image mosaicking and image alignment.

- Automation of 3D volume reconstruction by segmentation, feature detection, moment analysis, feature correspondence estimation, and registration optimization.
- Intensity correction of CSLM images for robust feature detection and enhanced 2D and 3D visualization.

Network for Earthquake Engineering Simulation (NEES) project: Multi-instrument data analysis methods for autonomous structural health monitoring systems for NEES projects (joint work with Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign), 08/02-08/03.

- Data integration of vector and raster measurements.
- Visualization of advanced sensors for structural analysis.

Other projects: Autonomous mobile robot navigation, 08/99-08/02.

- Landmark-based mobile robot navigation method using omni-directional vision.

PUBLICATIONS

Journal Publications

1. Sang-Chul Lee and Peter Bajcsy, "Intensity Correction of Fluorescent Confocal Laser Scanning Microscope Images by Mean-Weight Filtering", in *Journal of Microscopy*, Blackwell Synergy, 2006.
2. Peter Bajcsy, Sang-Chul Lee, Amy Lin and Robert Folberg, "3D Volume Reconstruction of Extracellular Matrix Proteins in Uveal Melanoma from Fluorescent Confocal Laser Scanning Microscope Images", in *Journal of Microscopy*, Blackwell Synergy, vol. 221 (1), pp. 30-45, 2006.
3. Sang-Chul Lee, Peter Bajcsy, Amy Lin and Robert Folberg, "Accuracy Evaluation of 3D Volume Reconstruction Techniques at Different Automation Level", in *EURASIP JASP, Performance Evaluation in Image Processing*, Hindawi publishing, 2006.
4. Amy Y. Lin, Zhuming Ai, Sang-Chul Lee, Peter Bajcsy, Jacob Pe'er, Lu Leach, Andrew J. Maniotis, and Robert Folberg, "Comparing the surface area of looping vasculogenic mimicry patterns with the surface area of blood vessels", in *Applied Immunohistochemistry and Molecular Morphology*, Lippincott Williams and Wilkins, 2006.

Conference Publications

1. Sang-Chul Lee and Peter Bajcsy, "Automated Feature-based Alignment for 3D Volume Reconstruction of CLSM Imagery", *SPIE International Symposium in Medical Imaging*, 2006

2. Peter Bajcsy, Sang-Chul Lee and David Clutter, "Supporting Registration Decision during 3D Medical Volume Reconstruction", SPIE International Symposium in Medical Imaging, 2006
3. Rob Kooper, Andrew Shirk, Sang-Chul Lee, Amy Lin, Robert Folberg and Peter Bajcsy, "3D Medical Volume Reconstruction Using Web Services", in Proc. of IEEE International Conference on Web Services, 2005
4. Sang-Chul Lee and Peter Bajcsy, "Feature based Registration of Fluorescent LSCM Imagery", in Proc. of SPIE International Symposium in Medical Imaging, 2005
5. Sang-Chul Lee and David Kriegman, "Omnidirectional Vision based Mapping by Free Region Sweeping", in Proc. of IEEE Conference on Robotics, Automation and Mechatronics, 2004
6. Sang-Chul Lee and Peter Bajcsy, "Multi-sensor Raster and Vector Data Fusion Based on Uncertainty Modeling", in Proc. of IEEE International Conference on Image Processing, 2004

Other publications

1. Sang-Chul Lee and Peter Bajcsy, "Multi-instrument Analysis from Point and Raster Data", Technical Report NCSA-ALG-04-0002, Feb 2004
2. Sang-Chul Lee, "Landmark-Based Exploration, Mapping and Navigation by Free-Region Sweeping", M.S. thesis, Computer Science, University of Illinois at Urbana-Champaign, 2003

PROFESSIONAL MEMBERSHIPS

IEEE Computer Society, International Society for Optical Engineering (SPIE) , Korean-American Scientists and Engineer Association (KSEA)

HONORS AND ACTIVITIES

- Reviewer: International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), 02/05.
- Honor Student, Inha University, 02/95-02/98.
- Korean Airline Student Award, Korean Airline, 08/95-02/98.
- Student Representative, Department of Computer Science and Engineering, Inha University, 02/95-02/96.