SEGMENTATION AND REGISTRATION OF 2D MULTIPHOTON MICROSCOPY IMAGES

A Thesis

Submitted to the Faculty

of

Escola Tècnica Superior d'Enginyers de Telecomunicació de Barcelona

(Universitat Politècnica de Catalunya)

by

Francisco Serrano Cebollero

In Partial Fulfillment of the

Requirements for the Degree

of

Enginyer Tècnic Superior de Telecomunicacions

May 2009

Purdue University

West Lafayette, Indiana

Habrá un día en que todos
al levantar la vista,
veremos una tierra
que ponga libertad...
-José Antonio Labordeta-

To my family...

ACKNOWLEDGMENTS

I would like to express my sincere thanks to my advisor Edward J. Delp for his guidance and patience. During the last 8 months I have encouraged in so many aspects. Thanks Professor Paul Salama for his support every Wednesday at noon. Kevin S. Lorenz, this work wouldn't have been possible if we hadn't worked as a team. Thanks Kevin for your understanding even if I sometimes drove you crazy. I think that the four of us have performed a good team.

A would also like to thank Prof. Luis Torres, Prof. Ferran Marqués and Prof. Phillipe Salembier. They have taught me all the knowledge I needed for developing this project.

The support of everyone at VIPER Lab: Golnaz, Ka ki, Kevin, Ying, Marc, Satyam, Maggie, Aravind, Nitin and Deen has helped me during my stay at Purdue.

I would like to thank the neprhology group at Indiana University who have provided us the data for analysis.

Finally, thanks all the people around the world who have been there when I needed their support. I would specially like to mention my family: Francisco, María Dolores, Amalia, Ricardo, Sandra, Gregorio, Irene, Lorea and Abril. Gracias!

This work was partially supported by Fundación VODAFONE Spain and Obra social Bancaja.

PREFACE

Nuevas técnicas en adquisición de imagen a través de microscopio se están desarrollando en los últimos años [1]. Antes de que aparecieran dichas técnicas, la microscopía era utilizada como una herramienta de simple visualización. Gracias a la obtención de imágenes digitales, se pueden desarrollar técnicas de procesado de imagen que de forma automática o semiautomática, evalúen los diferentes tejidos. Este proyecto trata la segmentación y el alineamiento de pilas de imágenes tomadas a través de microscopios multifotón. Dichos microscopios se basan en un láser de luz cuyos fotones excitan el objeto a visualizar. Figure A.2 muestra el diagrama de este tipo de microscopios.

Las estructuras biológicas que se analizan en este estudio consisten en su mayor parte en vasos sanguíneos de hígados y riñones de ratas vivas. El objetivo del estudio, con las dificultades que conlleva, es el análisis de dichas imágenes teniendo en cuenta que los especímenes están vivos y por lo tanto se están produciendo movimientos debidos a que los animales respiran. Es por ello que el problema del alineamiento de las imágenes se hace todavía más exigente.

En este documento se describen dos técnicas utilizadas para el procesado de las imágenes. En la Section 2 se explica la primera idea desarrollada que más adelante se abandona en favor de la segunda. En el siguiente apartado, se opta por un filtrado paso bajo, binarización adaptativa y filtrado morfológico con un alineamiento de las imágenes como paso intermedio. Todo ello es explicado en la Section 3. Seguidamente, se muestra un amplio capítulo donde se presentan y se evalúan resultados. El siguiente capítulo (Section 5) muestra alternativas que se han estudiado en el proyecto. Finalmente, se exponen las conclusiones y se proponen ideas que pueden ser estudiadas en el futuro. Además, se incluyen tres anexos, el primero de ellos desarrolla el funcionamiento de los microscopios utilizados para la adquisición de las

imágenes, el segundo amplía resultados y el tercero sirve de ayuda para el acceso a las imágenes analizadas vía web.

El método descrito en Section 2 hace, en primer lugar, la descomposición multinivel con filtros de apertura (Ver Figure 2.2) de la imagen obteniendo diferentes objetos en las consecutivas imágenes dependiendo del tamaño de los mismos. La suma de algunas de estas imágenes hace obtener estructuras deseadas despreciando el ruido de fondo. Después de ello, se segmentan las imágenes utilizando el algoritmo de contornos activos, produciendo a su salida una imagen binaria. Dicha imagen es posteriormente filtrada con un filtro morfológico de apertura y luego cerradura con reconstrucción para eliminar pequeños objetos indeseados que han sido segmentados y también rellenar los objetos segmentados ya que no son uniformes. Finalmente, se toma la imagen binaria como máscara para reconstruir la original. El siguiente paso del método consiste en cambiar el algoritmo de contornos activos por una simple binarización. Tras el filtro multiescala, se diferencia claramente entre vasos sanguíneos y fondo de la imagen lo que hace posible que sólo sea necesaria la binarización para obtener los dos niveles (vasos y fondo de imagen). Debido a la diferencia de iluminación entre las primeras imágenes de cada pila y las últimas, es necesario algún algoritmo que se adapte a dicha iluminación. Se decide implementar K-means cuya función es separar los píxeles en dos grupos minimizando la suma de cuadrados. Resultados de este proceso se pueden ver en la Figure 4.1 para las imágenes introducidas en la Figure 1.1. El motivo por el cuál se desprecia este método y se opta por el siguiente es que, dependiendo del tipo de imágenes, la selección de las diferentes salidas de los filtros no es uniforme. Además, el siguiente método trata el problema del alineamiento de las imágenes.

En la Section 3 se desarrolla el algoritmo que en la Section 4 presenta extensos resultados. Figure 3.1 muestra el diagrama de bloques del proceso. En primer lugar, y debido a que las características del ruido introducido por el microscopio se modelan como de Poisson, se aplica un filtro MAP (Maximum a Posteriori). El objetivo de este filtro es disminuir el ruido. Como se puede ver en la Figure 2.1, después de

que las imágenes son filtradas, el ruido ha disminuido aunque todavía es notable. Después se aplica un filtrado espacio-temporal para apaciguar el ruido ya que el ruido está presente en forma de impulsos. En primer lugar, la imagen es filtrada en el dominio espacial usando un filtro triangular de tamaño 5x5 píxeles. Sólo si todas las imágenes están alineadas, se puede aplicar el filtro temporal. Por ello se introduce en este punto el alineamiento de las imágenes explicado en el siguiente párrafo. Para el filtrado temporal de la imagen se opta por un filtro triangular de tamaño 5 imágenes (2 previas y 2 posteriores). Para la binarización de la imagen, se aplica el mismo algoritmo explicado en el párrafo anterior (K-means). El siguiente bloque consiste en un filtro de apertura y cerradura para eliminar pequeños objetos tal como se explica en el primer método. Finalmente, y a partir de la imagen binarizada se construye la imagen original segmentada y alineada.

El alineamiento de la imagen se lleva a cabo solamente teniendo en cuenta movimientos horizontales y verticales entre consecutivas imágenes. La métrica utilizada minimiza en el error cuadrático medio. Además, el algoritmo de descenso de gradiente es utilizado en lugar de una búsqueda exhaustiva de forma que se minimiza el tiempo de computación. También se utiliza interpolación bilineal de forma que en lugar de utilizar los píxeles de la imagen como rejilla, se utilizan fracciones de los mismos. Finalmente y como condición de borde se utiliza la condición de Neumann de forma que se aplica el valor más cercano dentro de la imagen para el computo de la métrica.

Gran cantidad de pilas de imágenes han sido procesadas y comparadas en este trabajo. Section 4 y Section B muestran, comparan y comentan gran parte de ellas. Debido a que la mejor forma de visualizar el alineamiento de las imágenes es mediante la reproducción de toda la pila en forma de video, y ya que eso no es posible en el documento escrito, Section C explica como acceder vía web a estos archivos. Los resultados incluidos en este documento consisten en la representación de 2 consecutivas imágenes y seguidamente los resultados del procesado de las mismas (Véase Figure 4.2, Figure 4.3, Figure 4.4, Figure B.3, Figure B.5, Figure B.7 y Figure B.9). Además de estos resultados, y en las siguientes respectivas figuras, se pueden ver las

gráficas de los movimientos verticales y horizontales además de la gráfica del valor del umbral obtenido después de la binarización.

Debido a la naturaleza del movimiento entre imágenes, en algunas de las pilas no se consigue el perfecto alineamiento, pues es necesario tener en cuenta otro tipo de movimientos (rotación, escalado). También se requiere el alineamiento de los diferentes objetos en la imagen debido a la respiración de los animales. Así, en la pila *ruben34* (Véase Section C), se pueden apreciar movimientos ondulatorios que con los métodos arriba expuestos, pueden ser corregidos. La segmentación también puede ser perfeccionada con métodos más avanzados como se explica en la Section 6

Se pretende que con el conocimiento de la forma, tamaño, movimiento de estructuras biológicas, se llegue a una interpretación y futuro desarrollo de soluciones para problemas médicos. Por ello, el trabajo realizado espera aportar un grano de arena al complejo mundo de la bioingeniería.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

Figure Page

ABSTRACT

Serrano Cebollero, Francisco , Purdue University, May 2009. Segmentation and Registration of 2D Multiphoton Microscopy Images. Major Professor: Edward J. Delp.

Optical microscopy exhibits many challenges for digital image analysis. In general, microscopy volumes are inherently anisotropic, suffer from decreasing contrast with tissue depth, and characteristically have low signal levels. This thesis describes the initial work in segmenting and registering multiphoton fluorescent microscopy images via a combination of methods. In particular, it describes a method that utilizes image enhancement and spatial filtering along with registration (to correct translational motion) and temporal filtering. Experimental results indicate the methods are promising.

1. INTRODUCTION

Although traditionally used as an observational tool, optical microscopy has developed into a quantitative approach with the development of digital imaging and image analysis methods. For example, various light microscopy imaging techniques have been developed in nephrology. A review of many of these techniques is presented in [1] where their advantages, imaging properties, and limitiations are discussed. When combined with digital image analysis, microscopy is capable of quantifying physiology and structure at a cellular and subcellular level [2]. However, automated methods of quantitative 3D microscopy are still in their infancy. This primarily reflects the lack of image analysis tools suited to the unique properties of microscope image volumes. First, microscopy volumes are inherently anisotropic, with aberrations and distortions that vary in different axes. At a larger scale, contrast decreases with depth in biological tissues. This contrast decrease aggravates a general problem of fluorescence, which characteristically have low signal levels [3], consisting of as little as a single photon. Signal levels are further decreased by the need for high image capture rates necessary to image dynamic biological structures. Finally, biological structures often consist of many different kinds of irregular and complicated structures that are frequently incompletely delineated with fluorescent probes. Each of these characteristics contribute to low contrast and small intensity gradients in multiphoton image volumes, making segmentation and rendering results very sensitive to small changes in parameters, causing the failure of typical image analysis methods [4]. Nonetheless, automatic image segmentation is absolutely necessary for quantification of multiphoton image volumes, whose size and complexity makes manual image segmentation impractical.

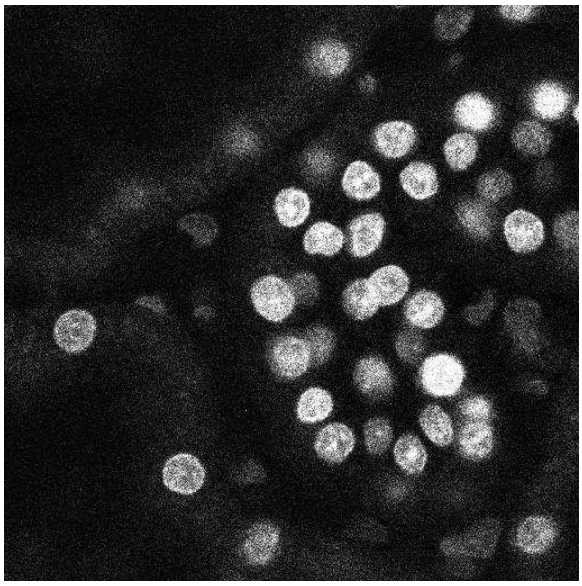
Knowledge of the shape, size, and motion of biological structures may aid in understanding and developing solutions for human and non-human diseases. Gaining

this knowledge often requires the analysis of microscopy image sets and using segmentation methods to identify and separate relevant biological structures. In attempts to perform automated segmentation of cellular structures in microscopy images, several segmentation methods have been proposed for microscopy data, including methods that rely on *a priori* shape information and methods that do not. For example, methods proposed in [5] identify cells with a circular structure based on *a priori* knowledge, while in [3] active contours are used to segment and track cells by minimizing a piece-wise constant energy function that includes terms that penalize the variance of the intensity inside and outside the boundary as well as a regularization term that penalizes total boundary length. In [6] a combination of thresholding and watershed segmentation is used to segment nuclei in time-lapse fluorescence microscopy images. This was followed by a scheme in which segmented nuclei less than some experimentally determined size threshold are merged together and those greater than this size threshold are split. In [4] chromosome feature extraction by using a Kohonen self organizing feature map is described. Neuron segmentation is described in [7] by identifying edges via a multiscale wavelet edge detection technique [8].

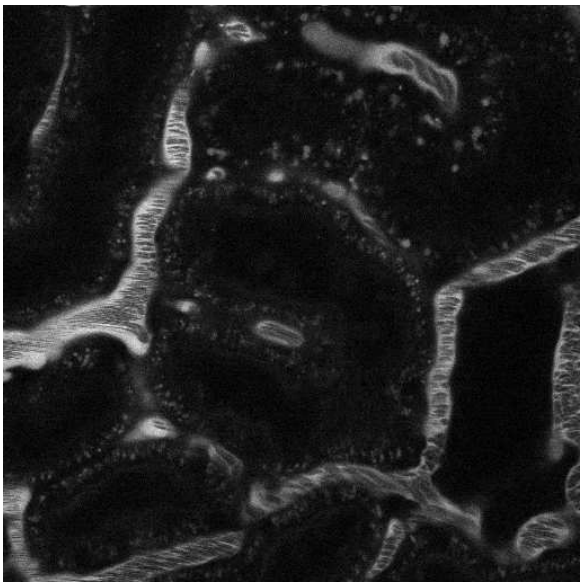
Microscopy image sets are frequently obtained from live specimens. Therefore, motion artifacts are introduced into the scene as a result of factors such as respiration and voluntary muscle contractions. It has also been demonstrated that using registration schemes can aid with segmentation, especially with volumetric data. Image registration involves matching two or more images taken at different times or from different viewpoints. Registration aligns the images so that they all share a common coordinate system, a condition aiding in future image analysis. The process is often described as finding an explicit function that performs a backward mapping of a target image onto a source image [9, 10]. Methods proposed in [11] use motion patterns of cells to register consecutive images, while methods in [6] use a matching process between consecutive images to determine displacement distances. Nonetheless, there exists much controversy over approaches to image segmentation with confocal and multiphoton microscopy. Some argue [12] that user interaction is

best suited since no automated procedure can truly reproduce the accuracy of the human visual system in correctly identifying individual cellular objects. While the entire segmentation process is not completely manual, it requires at least a manual input of an initial seed value or starting point. Additionally, several authors have based their registration schemes on manually selecting landmarks [13–15]. Finding an optimal relation between images given specified landmarks is described as NP-hard, and therefore heuristics is often used to make the matching process practical [16]. Introducing a smoothness/regularity constraint to reduce the effects of noise on elastic materials have been promising [17]. Other authors have mixed the use of landmarks with contribution from the image [18]. However, others argue [2] that a completely automated process is necessary to obtain quantitative, objective, and repeatable results. User interaction causes varying results for identical input data sets which may be unacceptable in certain applications.

This document describes a method to segment multiphoton fluorescent microscopy images via a combination of segmentation and registration methods. In particular, the proposed method utilizes image enhancement and spatial filtering along with registration (to correct for translational motion) and temporal filtering. Motivation for the development of our proposed technique is provided by the desire to analyze two types of data sets of multiphoton fluorescent microscopy images. One data set type consists of a progression of images corresponding to focal planes looking deeper in the tissue. Outside literature may refer to this type of data set as "z-series" data, as the z-axis typically corresponds with depth. A specific example of this data set type used in this work is referred to as the *nuclei* data set, which contains approximately one-hundred images of rat kidney cells. The other data set type consists of a progression of images corresponding to a series of time instances. Outside literature may refer to this type of data set as "time-series" data. A specific example of this data set type used in this work is referred to as the *vascular flow* data set, which contains approximately two-hundred images comprising vascular flow in a rat liver. Example images from each data set are shown in Figure 1.1. Both data sets have been corrupted with noise



(a) Image 29 of *nuclei* data set



(b) Image 70 of *vascular flow* data set

Fig. 1.1. Example of Images Used in Our Study

from a variety of sources, including photon shot noise and detector noise due to the small number of photons available to illuminate each pixel or voxel, creating low SNR values.

The goals of this project include:

Image segmentation The goal is to identify and segment individual nuclei from the *nuclei* image set for the purpose of quantifying the number and volume of nuclei in the scene. In this data set, given any particular image, some nuclei appear with significantly increased intensity. This is expected, as cells that are currently in the middle of the focal plane will appear with more intensity, while cells that are on the edge of the focal plane will appear with less intensity. This is a common occurrence with multiphoton microscopy as opposed to confocal microscopy. This effect is caused by the dye concentrating within the middle of the nuclei creating a non-homogeneous distribution. Naturally, selecting which cells are relevant in any given image is subjective. Additionally, the objective is to identify the path of blood flow in the *vascular flow* image set for the purpose of quantifying the area occupied by vascular vessels in the scene.

Image registration In the *nuclei* data set, since there does not appear to be significant movement between images, registration techniques may not need to be utilized. This is expected as the tissue was fixed and not live. Regarding the *vascular flow* data set, movement between images is much more significant due to specimen motion during acquisition. Therefore, registration is needed to correct these movements between images. Registration is expected to improve segmentation performance. However, registration is also a goal in itself of the project.

The general goal of this work is to segment a sequence or "stack" of confocal fluorescent microscopy images. Concerning data set types consisting of progressive tissue depths such as the *nuclei* data set, the images are acquired at various (known) depths across a 3D sample of renal structures in a rat. Regarding time series data set types such as the *vascular flow* data set, the images are acquired at evenly spaced time intervals across a single focal plane of liver structures in a rat. Two fundamental approaches have been developed to address the aforementioned issues, with the second approach replacing the initial approach. This thesis is organized as fol-

lows: Section 2 describes the initial approach for image segmentation despite later being abandoned, Section 3 describes the second approach for image segmentation and registration, Section 4 shows experimental results from the second approach, Section 5 discusses alternatives and modifications available to the second approach, and Section 6 discusses future work and concludes the document.

2. INITIAL APPROACH

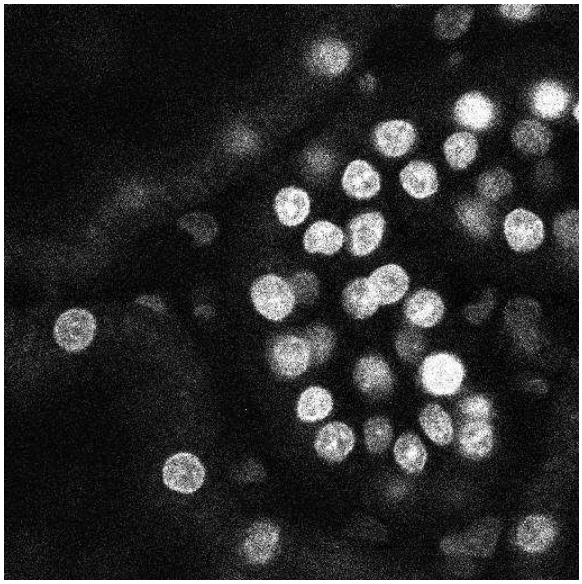
This section describes the initial approach for image segmentation. Despite sufficient performance and results as shown in Figure 4.1, this approach is later abandoned in favor of the approach described in Section 3.

2.1 Maximum a Posteriori Filtering

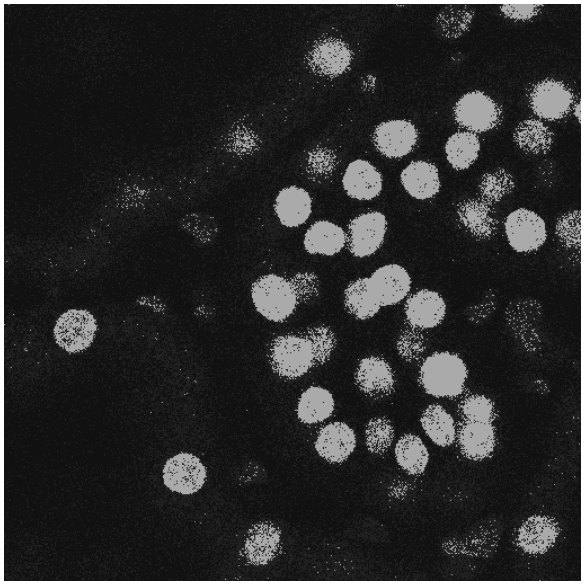
The first step in this initial approach is a denoising filter. Specifically, this filter is a maximum a posteriori (MAP) filter [19] which takes advantage of the known characteristics of the noise, as photon noise is known to predominantly have a Poisson distribution. By utilizing this additional information about the characteristics of photon noise, particularly in the *nuclei* data set, cellular objects become more well-defined to enhance the performance of the segmentation process to follow. Figure 2.1 shows the effects of the filter on an example image.

2.2 Morphological Filtering and Segmentation

With respect to image segmentation, the initial approach was to consider a grayscale morphological multi-scale analysis of both data sets [20], as shown in Figure 2.2. Raw data images are provided as the input to the system, shown at the bottom of the figure as x_i, indicating a single image. Each operator F_k represents a morphological opening followed by a morphological reconstruction, both using the same structuring element. Morphological reconstruction is a series of repeated dilations of an image, known as the marker image, until the profile of the marker image is entirely less than the profile of a second image, known the mask image. This process causes the peaks in the marker image to dilate [21]. The marker image is given as the morphologically



(a) Original Image



(b) MAP Filtered Image

Fig. 2.1. Maximum a Posteriori Filtering of Image 29 from *Nuclei* Data Set

opened image, while the mask image is given as the original input image. Each operator F_k uses a flat, square structuring element of size $2k+1 \times 2k+1$. Due to the multi-scale structure, subsequent images, l_k, extract increasingly larger objects from the previous g_{k-1} images. The g_k images contain the remaining objects after objects segmented in l_k have been removed. The results of the morphological multi-scale analysis for an example image in the *nuclei* data set are shown in Figure 2.3, and results for an example image in the *vascular flow* data set are shown in Figure 2.4.

In selecting which filter outputs should appear in the final output, the analysis was subjective. Compositions such as $l_4 + l_5 + l_6 + g_6$ and g_4 appeared to produce acceptable results. However, the segmented objects in these composition images often contained small holes. To eliminate these holes, a single morphological closing operation with a flat, square structuring element of size 5×5 was used. The size of this structuring element was selected empirically. The resulting image now contained large, undesired dark artifacts. Only the cells with the greatest intensity are desired. Therefore, a binary image was created by using a simple thresholding technique, where the threshold was chosen subjectively to eliminate the dark artifacts and retain the desired cell objects. From this binary image, the final output image was created. For black pixels in the binary image, the corresponding pixels in the output image remain black. For white pixels in the binary image, the corresponding pixels in the output image are restored to the corresponding pixels in the original image. The binary image acts as a segmentation mask for the original image:

$$F(m,n) = B(m,n)O(m,n) \tag{2.1}$$

where F is the final image, B is the binary image, and O is the original image. It should be noted that this approach is a causal system. Furthermore, the current output image only depends on the current input image. No dependencies on previous or future images are created. However, a drawback with this approach became apparent due to the nature of the data set. In the majority of the images, individual cells are composed of individual pixels that are not very contiguous. Images of the cells

do not appear as solid objects and often contain many black pixels interspersed between white pixels. This is due to the cell itself not having a rigid boundary, or more likely due to characteristics of the image acquisition system. For extreme cases, the morphological opening operations eliminated all of these cells, and the corresponding images turned out to be essentially completely black, particularly for beginning and ending images in the sequence.

In an attempt to improve on this approach, active contours were introduced as a replacement for the subjective threshold.[1] However, the results showed no noticeable improvement, but greatly increased computational complexity. Therefore, active contours were abandoned and the simple subjective threshold was used.

[1]The active contour based segmentation was provided by: F. Marqués and P. Salembier, Signal Theory and Communications Department, Technical University of Catalonia

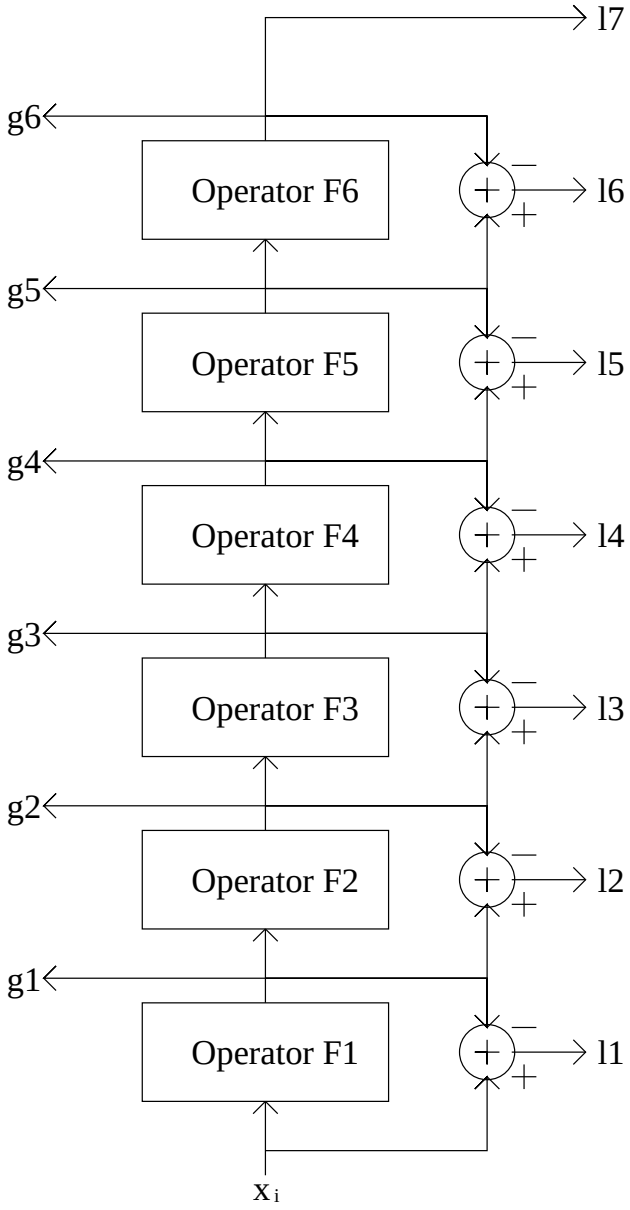
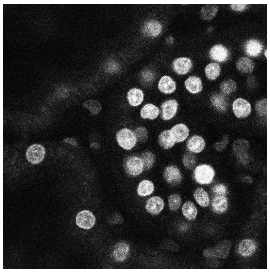
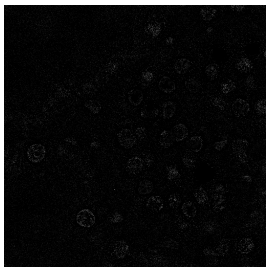


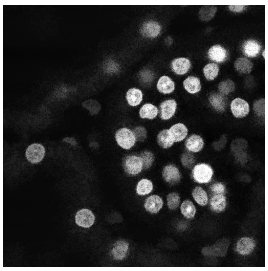
Fig. 2.2. Morphological Multi-Scale Analysis



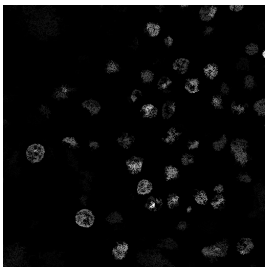
(a) Original Image



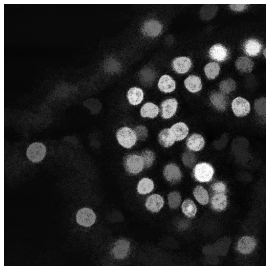
(b) Image l1



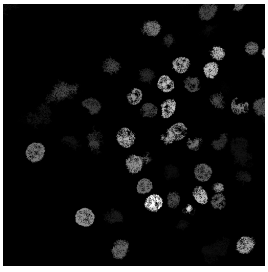
(c) Image g1



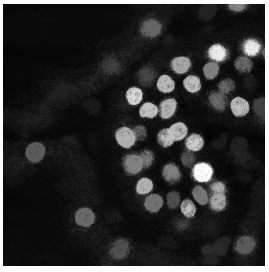
(d) Image l2



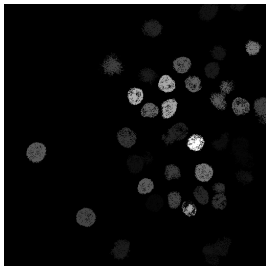
(e) Image g2



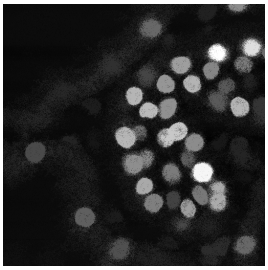
(f) Image l3



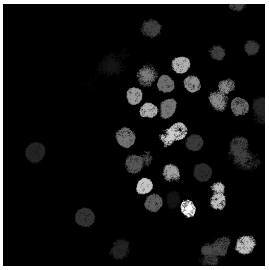
(g) Image g3



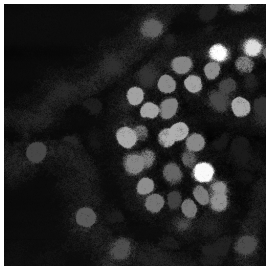
(h) Image l4



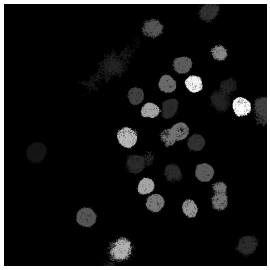
(i) Image g4



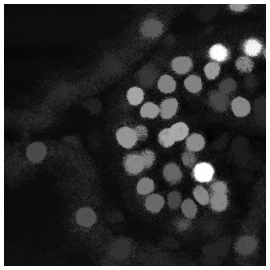
(j) Image l5



(k) Image g5

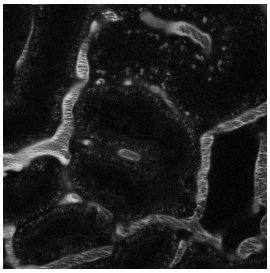


(l) Image l6

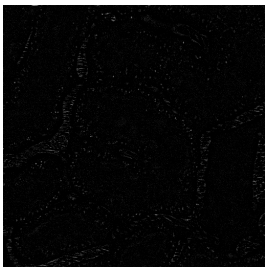


(m) Image g6

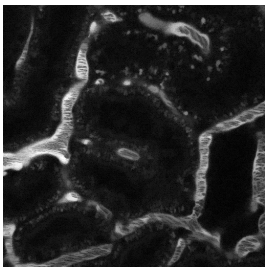
Fig. 2.3. Multi-Scale Decomposition of Image 29 from *Nuclei* Data Set



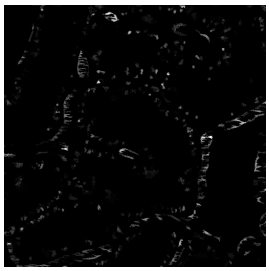
(a) Original Image



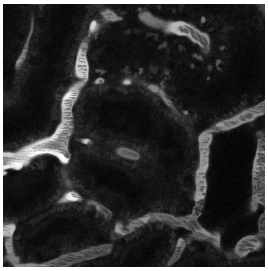
(b) Image l1



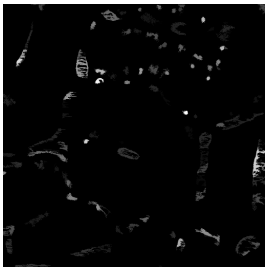
(c) Image g1



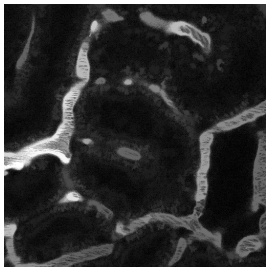
(d) Image l2



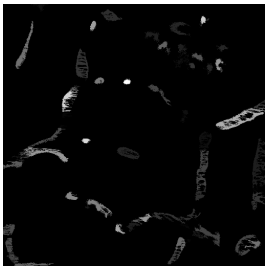
(e) Image g2



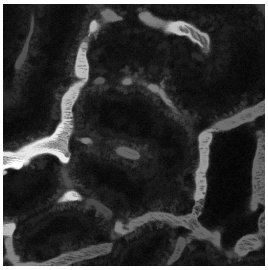
(f) Image l3



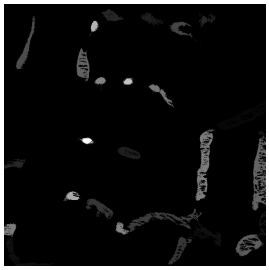
(g) Image g3



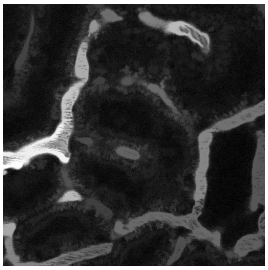
(h) Image l4



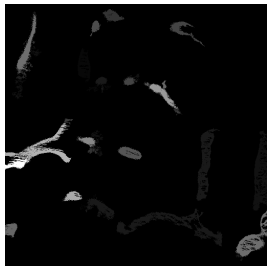
(i) Image g4



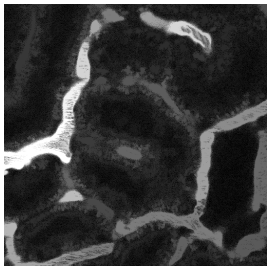
(j) Image l5



(k) Image g5



(l) Image l6



(m) Image g6

Fig. 2.4. Multi-Scale Decomposition of Image 70 from *Vascular Flow* Data Set

3. PROPOSED APPROACH

The initial approach discussed in Section 2 only addressed image segmentation. This implementation now considers both image segmentation and registration. By introducing registration along with segmentation, motion artifacts within the scene may be reduced. Changes in shape, size, and motion of cellular structures may be better visualized. These artifacts were not addressed using the initial approach described in Section 2. For example, the motion of particular cellular objects may be exaggerated by specimen respiration. Therefore, the results would not reflect the true motion of just the internal biological structures.

The proposed analysis technique is shown in Figure 3.1. First, each image in the image stack is denoised via a Bayesian estimation technique [19] as used in the initial approach described in Section 2.1. Each image is then low-pass filtered to provide some local blurring and is subsequently spatially registered. One dimensional smoothing along the vertical (space/time) axis after the images have been aligned is then done followed by adaptive thresholding to produce a binary image using K-means. Finally, the segmentation mask constructed using K-means is filtered using morphological techniques. This new mask is then used to segment each image.

3.1 2-D Spatial Low-Pass Filtering

After denoising, each image is then low-pass filtered to provide some local blurring using a 3×3 low-pass filter with impulse response:

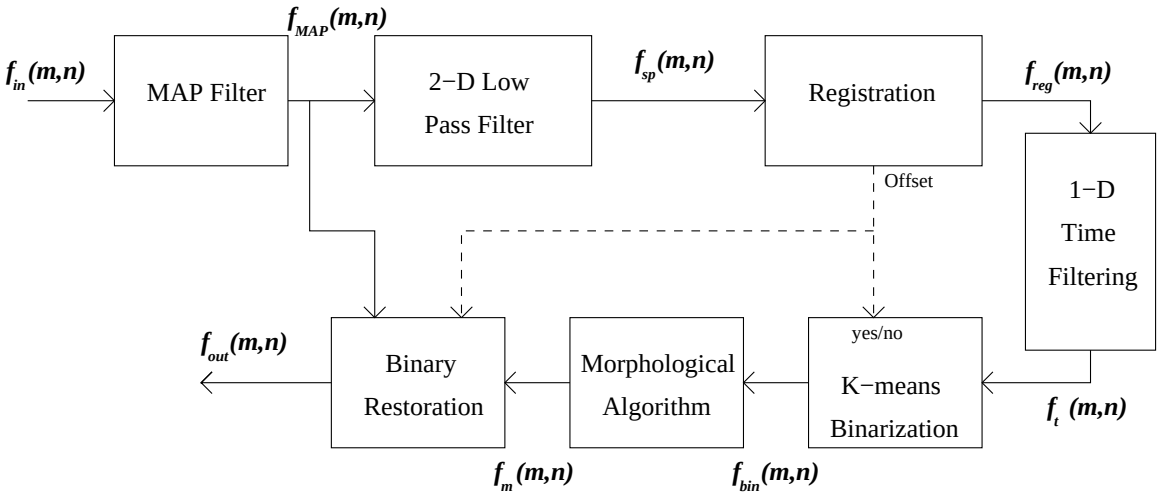


Fig. 3.1. Block Diagram of Our Proposed Approach

$$h(m,n) = \begin{cases} 0.2 & \text{if } m,n = 0 \\ 0.1 & \text{if } -1 \leq m,n \leq 1 \\ 0 & \text{else} \end{cases} \tag{3.1}$$

Our experiments have shown that this blurs the image so that segmentation is more robust. The blurring process mitigates shot noise remaining in the image after the Bayesian denoising filter is used. A similar approach was adopted in [11]. It is a commonly held belief that blurring traditionally reduces segmentation performance, as object boundaries become less well-defined. However, object boundaries for our images are poorly defined in the sense that they are not composed of rigid continuous edges. In this case, blurring the objects improves the segmentation by creating more well-defined continuous boundaries in contrast to ill-defined sparse boundaries. The coefficients of this filter were selected empirically.

The comparison between the input and output of the filter can be seen on figure Figure 4.7(a) and Figure 4.7 respectively. It can be seen that the filtered image has a more uniform light intensity values than the original image.

3.2 Registration

After each image has been blurred, it is then registered to remove translations due to respiration and other motion artifacts. The registration technique used was obtained from ITK[1] [22]. Registration within ITK requires several parameters to be specified. An extended description of these parameters along with an overview of all components of the registration process can be found in [23]. All user adjustable parameters and their corresponding options for registration are presented in Table 3.1, with the selected options used in this approach shown in bold. The particular registration parameters selected for this approach include a Neumann boundary condition, where pixel values outside of the image boundary are given the values of the nearest

[1]National Library of Medicine Insight Segmentation and Registration Toolkit (http://www.itk.org)

pixel within the image boundary. Additionally, registration is performed in a simulated continuous space, not in the discrete space of the image grid. Equivalently, a fixed sub-pixel resolution is not specified. The value at any arbitrary fraction of a pixel location may be requested, and a bilinear interpolation is used to obtain pixel values at these non-integer sub-pixel locations. Some authors have compared the various interpolation methods [24, 25]. Furthermore, an exhaustive search is not performed to compute the registration solution. Instead, several optimization methods (see Table 3.1) have been studied [26, 27]. Our registration implementation employs a step gradient descent optimizer. This optimization method greatly reduces computational complexity by computing the error metric only for locations along the direction with the steepest gradient and using a bipartition scheme to compute the step size [23].

The metric employed was the mean squared error metric:

$$(u, v) = \underset{(u,v)}{\operatorname{argmin}} \left\{ \sum_m \sum_n (x_{i-1}(m, n) - x_i(m - u, n - v))^2 \right\} \tag{3.2}$$

where $x_{i-1}(m, n)$ and $x_i(m, n)$ denote the pixels at location (m, n) within the reference image x_{i-1} and target image x_i, respectively. A full comparison of the use of metrics in Medical image registration is done in [28]. The reference image used is always the current image's prior neighbor. If the registration of the current image in reference to the previous image is denoted as $\mathbf{r}(x_i, x_{i-1}) = x_i \circ x_{i-1}$, the output of the registration process in terms of the first image can be represented as a concatenation of all previous registrations:

$$\mathbf{r}(x_i, x_1) = \mathbf{r}(x_i, x_{i-1}) \circ \mathbf{r}(x_{i-1}, x_{i-2}) \circ \cdots \circ \mathbf{r}(x_2, x_1) \tag{3.3}$$

as was similarly performed in [29]. This was found to greatly improve registration performance instead of using the first image in the stack as the reference image, as well as reduce the computational complexity and consequently the time required to analyze the entire series. This result was especially true for z-series data, or image sequences corresponding with increasing tissue depth. Using the first image in the stack as the reference image may be suitable for time-series data, or image sequences

corresponding with increasing time instances, as the same cellular objects are expected to be in view for every image in the sequence.

Future enhancements to this registration process will involve considering other linear and non-linear transformations in addition to translation. This will further reduce artifacts caused by specimen movements and data acquisition limitations.

It can be seen on Figure 4.7(c) that the input image has been placed, after registering, on a bigger image. The first image of the stack is located at the center of the output image and the consecutive images are placed in their respective locations taking into account the movement values of the registration process.

3.3 1-D Low-Pass Filtering

Thus far, any information about the current image that is contained in previous and/or future images has not been utilized. If a particular object was segmented in the previous image, it is highly likely that this same object should also appear and be segmented in the current image. It has been observed in our images, particularly for the *nuclei* data set, that the intensity of a specific pixel related to an object of interest, as a function of depth (image number), starts out low, increases as the dye concentration in the object of interest increases in the current focal plane, peaks, and then decreases as the object fades away. However, noise interferes with this expected behavior, as seen in Figure 3.2(a), which displays an example of the pixel intensity as a function of image number. With the noise inducing sharp discontinuities in pixel intensity, objects segmented in the current image may disappear in the next image, then reappear in the following image. Therefore, to address this effect and restore the expected behavior of a pixel's intensity as a function of image number, a 1-D low-pass filter is used across the images. The filter is given by:

$$T_n = \frac{1}{10}I_n + \frac{2}{10}I_{n-1} + \frac{4}{10}I_{n-2} + \frac{2}{10}I_{n-3} + \frac{1}{10}I_{n-4} \tag{3.4}$$

Parameter	*Options Available* ***Option Selected***
Boundary condition	**Neumann** Constant padding Circular
Error metric	**Mean squared error** Normalized correlation Mean reciprocal squared difference Mutual information Kullback Liebler distance metric Normalized mutual information Mean squares historgram Correlation coefficient histogram Cardinality match Kappa statistics Gradient difference
Optimizer	Amoeba Conjugate gradient Gradient descent Quaternion rigid transform gradient descent Limited memory minimization Bounded limited memory minimization One plus one evolutionary **Regular step gradient descent** Powell Simultaneous perturbation stochastic approximation Versor transform Versor rigid 3D transform
Interpolator	Nearest neighbor **Bilinear** B-Spline Windowed sinc

Table 3.1

where T_n denotes the n-th filtered image corresponding to the n-th input image I_n and its four preceding images I_{n-1}, I_{n-2}, I_{n-3}, and I_{n-4}, respectively. The coefficients of this filter were selected empirically. The filter is chosen so that the system remains causal, with the idea of establishing a real-time system. The largest coefficient is chosen to correspond with the second previous image. Therefore, the system will have a time delay of two images. For images at the beginning of the sequence, where not all of the four previous images are available, the n-th filtered image is obtained by:

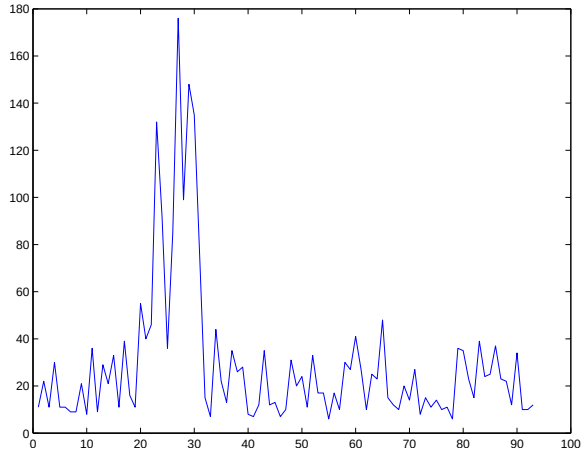
$$\begin{aligned} T_n &= \frac{1}{4}I_n + \frac{1}{4}I_{n-1} + \frac{1}{4}I_{n-2} + \frac{1}{4}I_{n-3} \\ T_n &= \frac{1}{3}I_n + \frac{1}{3}I_{n-1} + \frac{1}{3}I_{n-2} \\ T_n &= \frac{1}{2}I_n + \frac{1}{2}I_{n-1} \\ T_n &= I_n \end{aligned} \tag{3.5}$$

The filtered intensity profile corresponding to Figure 3.2(a) is shown in Figure 3.2(b). It has been observed that if consecutive images have significant movement or displacement, then any analysis following the 1-D filtering fails. This is because the relationship between image number and pixel intensity shown in Figure 3.2 is no longer valid. Therefore, this movement must be corrected using registration before the 1-D filter.

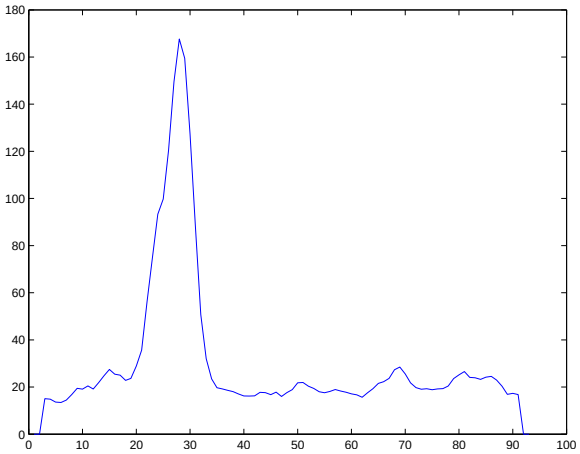
Image Figure 4.7(d) shows the output of that filter compared with the input on previous figure. Note here that the light intensity inside the blood vessels is now more constant which was the objective of the filter. Also, note that the boundaries are less defined.

3.4 K-Means Threshold

After filtering and registration, K-means ($K = 2$) is used to generate binary images to serve as segmentation masks [30]. The objective here is to generate image



(a) Pixel values vs. depth



(b) Pixel values after 1-D filtering

Fig. 3.2. This shows the effects of image number or depth (x-axis) on the pixel values (y-axis)

dependent thresholds, whereby binary segmentation masks can be constructed. The threshold obtained clusters each pixel from every image into one of two classes, segmented and unsegmented objects.

K-means is a method that achieves thresholding based on the shape properties of the histogram. It is an iterative scheme which uses the average of the foreground and background class means. The computation of the mean is computed as follows:

$$\begin{aligned} m_b(T_n) &= \sum_{g=0}^{T_n} gp(g) \\ m_f(T_n) &= \sum_{g=T_n+1}^{G} gp(g) \end{aligned} \tag{3.6}$$

where m_f and m_b are the computed means for the foreground and the background respectively, T_n is the value of the threshold (inizialiced to 256/2) and $p(g)$ is the value of the histogram for the intensity g. After the means are computed, a new threshold is found as:

$$T_{n+1} = \frac{m_f(T_n) + m_b(T_n)}{2} \tag{3.7}$$

Finally, the algorithm stops once $|T_n - T_{n-1}| < \alpha$ with $\alpha = 1$ in our case. It can be seen that for all the images in the stacks we tried, the number of iteration is smaller than 10 iterations.

Each binary mask was initially obtained by setting all pixels whose values fall below the image's corresponding threshold to zero, and those whose values are greater than or equal to the threshold to 255. It was observed, however, that the threshold does not change drastically between consecutive images. Therefore, K-means does not have to be used for every image, rather only when there is significant movement or a scene change. K-means is necessary because a static threshold obtained subjectively may work for a given specific data set, but would not be adaptive to other data sets.

The output of the binarization process can be seen on Figure 4.7(e). We can appreciate that the filter has split the image on two groups but there still are undesired objects that have been segmented.

3.5 Morphological Filtering

Upon examining the resulting binary masks, we observed that they often contained small black holes inside objects and small white objects in the background. To remove these isolated objects, several morphological filters were used [20]. First, a binary morphological opening with a flat, square 7×7 structuring element was used to remove small objects. Next, a morphological closing with a flat, square 4×4 structuring element was used to remove small holes. The sizes of these structuring elements were selected empirically. This resulted in the final binary masks used to segment the original images. Denoting an original image by O, its corresponding mask by B, then the segmented image F is given by $F(m,n) = B(m,n)O(m,n)$, just as was performed in the initial approach.

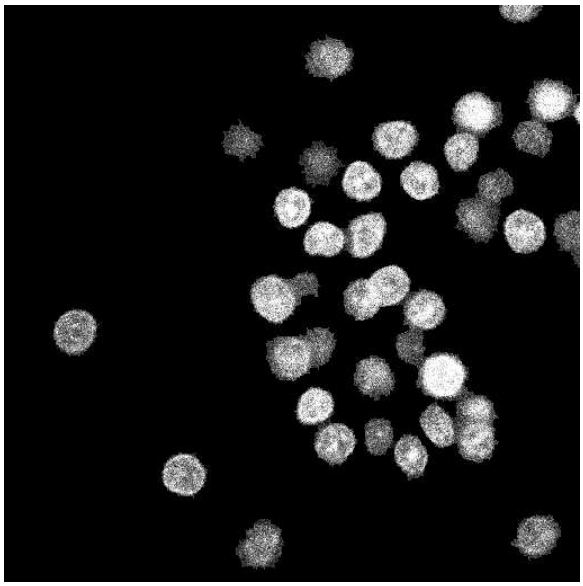
The output of the morphological filter and the reconstruction image can be seen on Figure 4.7(f) and Figure 4.7(g) respectively.

4. EXPERIMENTAL RESULTS

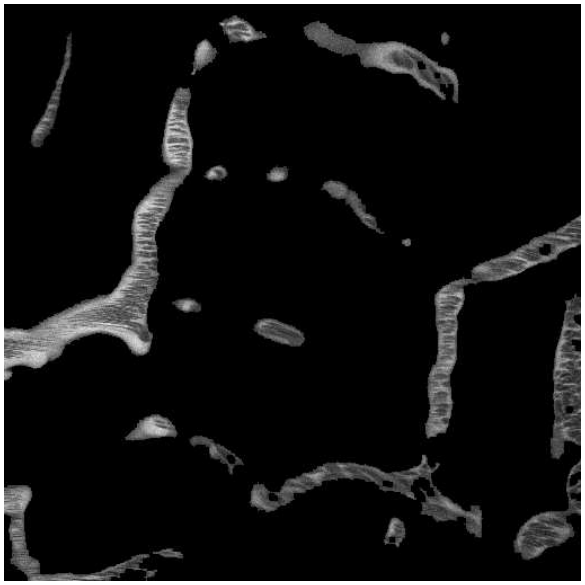
Example images displaying the segmentation results for both the *nuclei* and *vascular flow* data sets for the initial approach described in Section 2 are shown in Figure 4.1. Despite satisfactory segmentation results, this approach did not include any registration techniques and was therefore abandoned.

All of the results henceforth in this section now pertain to the second approach described in Section 3. A table summarizing all data sets processed with the second approach is shown in Table 4.1. Due to the similarity among the *ruben34min*, *ruben62min*, *ruben70min*, and *ruben120min* data sets, only the results from the *ruben34min* data set have been shown in this document for brevity. Likewise, the *glomerulus1* and *glomerulus2* data sets are very similar, and therefore only the results from the first data set have been included in this document. The rest of the Section 4 shows finals results of segmentation and registration of *nuclei* data set, *vasculature* data set and *ruben* data set. For the three stacks, outputs of two consecutive images are shown. A brief discussion of the registration and binarization technique is introduced here through the plots of horizontal and vertical movement and also the value of the threshold. The rest of the data sets of Table 4.1 are shown on Section B. Just one image of each stack is shown as a result, the complete stack of images in *png* format and also videos in *avi* format can be found on the following web page: https://redpill.ecn.purdue.edu/~kidney/thesisSerrano/ (See Section C).

Representative images showing the performance of our entire system for the *nuclei*, *vascular flow* and *ruben* data sets are given in Figure 4.2, Figure 4.3 and Figure 4.4, respectively. As it can be seen, nuclei with the greatest intensity are retained, while those with the lowest intensity are eliminated. Similarly, structures with the greatest intensity are retained in the *vascular flow* and *ruben* data sets, while those with the lowest intensity are eliminated. Note that in the *ruben* data set, the results images



(a) Image 29 of *nuclei* data Set



(b) Image 70 of *vascular flow* data Set

Fig. 4.1. Morphological segmentation results using initial approach

Name in Report	**Filename Prefix**	**Objective**	**Main Problem**
nuclei	BCHAN	Segment nuclei in sharp focus	Severe photon noise
vascular flow	rat1-infusion	Segment vasculature structures	Translational motion
hmang	hmang	Segment vasculature structures shown in red	Endosomes and tubular lumen have large red component
ruben	ruben34min	Segment vasculature structures shown in red	Registration requires non-linear distortion correction
unnamed	ruben62min	Segment vasculature structures shown in red	Registration requires non-linear distortion correction
ruben70	ruben70min	Segment vasculature structures shown in red	Registration requires non-linear distortion correction
unnamed	ruben120min	Segment vasculature structures shown in red	Registration requires non-linear distortion correction
glomerulus	glomerulus1	Segment vasculature structures	Artifacts due to respiration and acquisition
unnamed	glomerulus2	Segment vasculature structures	Artifacts due to respiration and acquisition
liver vasculature	vasculature-live	Segment vasculature structures shown in green	Significant green background fluorescence outside of blood vessels

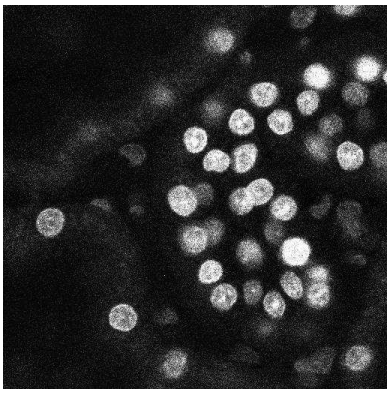
Table 4.1
List of Data Sets

are the original with the boundary of the segmentation instead of the output of the block diagram shown on Figure 3.1.

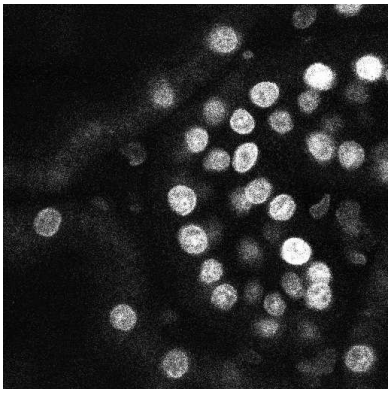
Registration results for the *vascular flow*, *nuclei* and *ruben* data sets are shown in Figure 4.5. Movements on *vascular flow* data set are almost constant values on positive direction for both, horizontal and vertical directions. Large spikes appear in the middle of the data set. It can be an indicator to possible distinguishes between multiple separate data sets. Playing the video for this data set, we observe a change of scene near image 103. The video also proves the constant value of the traslation in the same direction for the entire stack despite of some negative values. In a similar way, *nuclei* data set movements are also constant in the same direction in the first 60 frames. The last frames show a random movement. The reason of this movement is that the light intensity is too dark and the algorithm fails. For *ruben* data set we can appreciate a negative vertical movement but no horizontal movement is distinguished.

After each data set was denoised, filtered and registered, it was thresholded via K-means. Figure 4.6 shows the threshold value for the three stacks. *vascular flow* data set threshold is a constant value for all the images with a smaller value when a change of scene is present. In contrast, *nuclei* data set has a big variation of the threshold values. It's because of the fact of different light intensity on the stack having a low threshold values when the image is darker. Most clear is the case of *ruben* data set where there is a constant value until image 33 due to the light is constant in these images. From here to the end, there is a linear decreasing of the threshold due to the out of focus of the images.

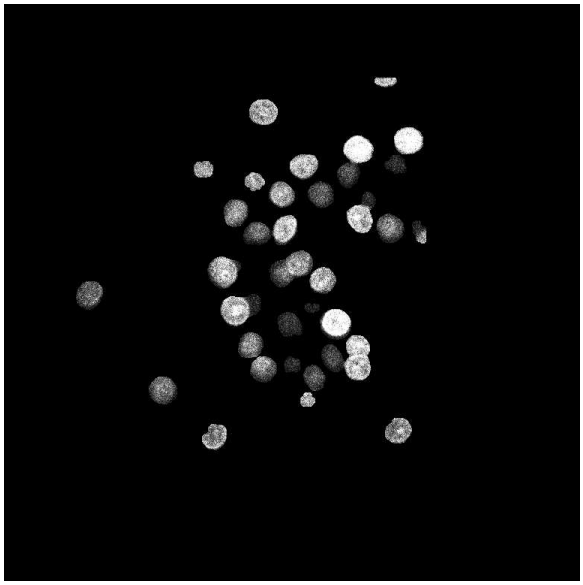
Finally, to demonstrate intermediate results from every block in Figure 3.1, the output of each block from the analysis of an example image from the *ruben* data set is shown in Figure 4.7.



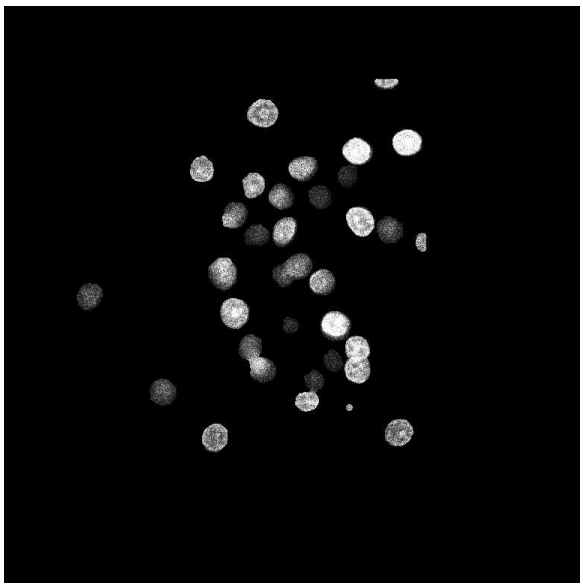
(a) Image 29



(b) Image 30

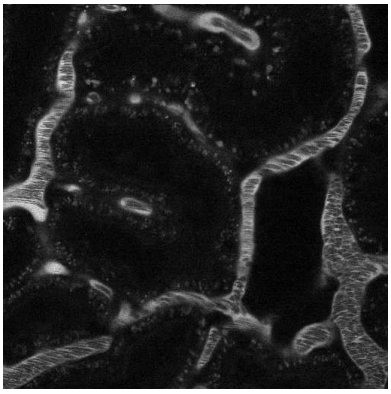


(c) Image 29

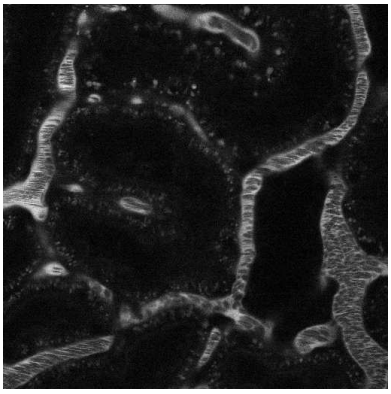


(d) Image 30

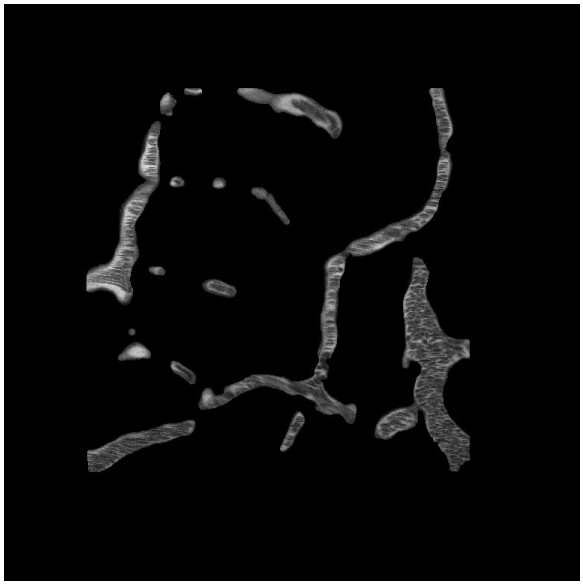
Fig. 4.2. Original Images (a, b) from *Nuclei* Data Set; Output Images (c, d)



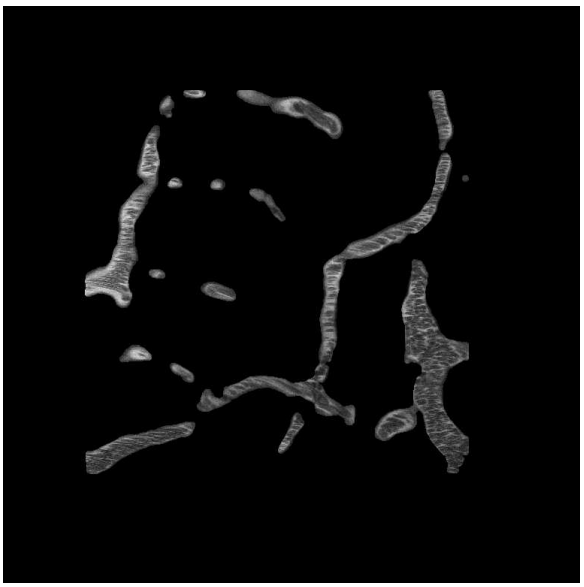
(a) Image 45



(b) Image 46

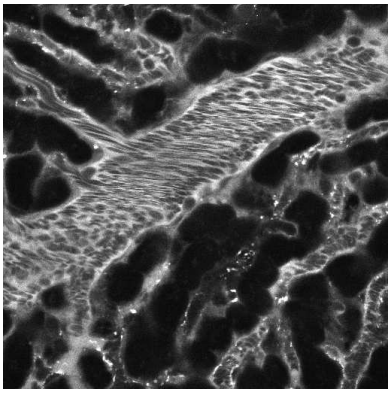


(c) Image 45

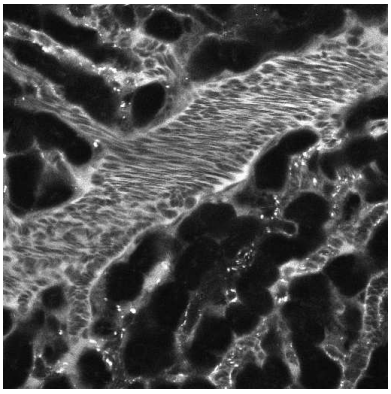


(d) Image 46

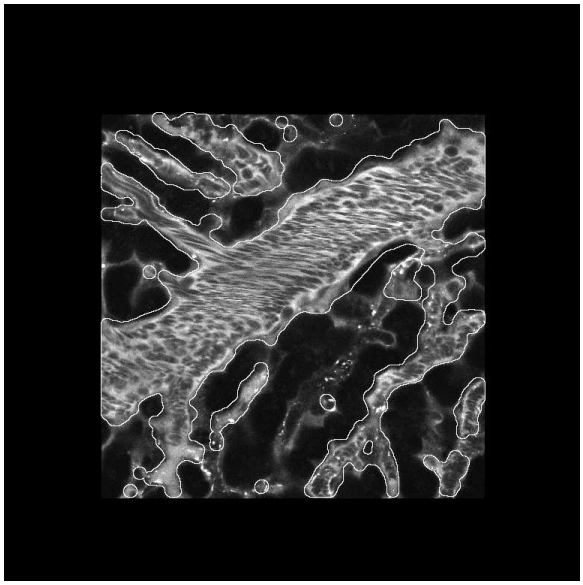
Fig. 4.3. Original Images (a, b) from *Vascular Flow* Data Set; Output Images (c, d)



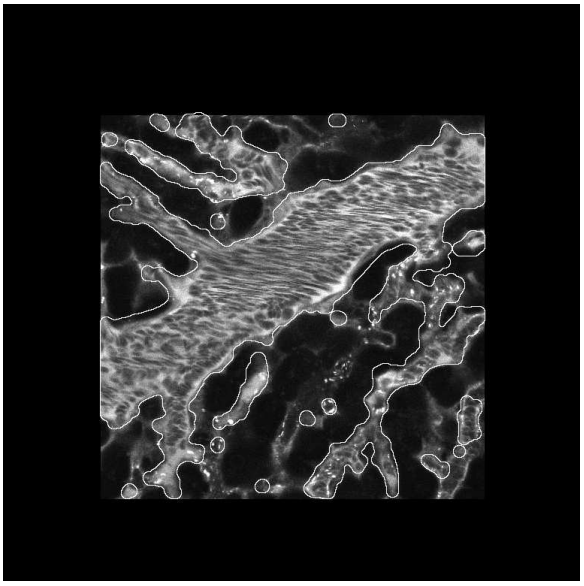
(a) Image 30



(b) Image 31

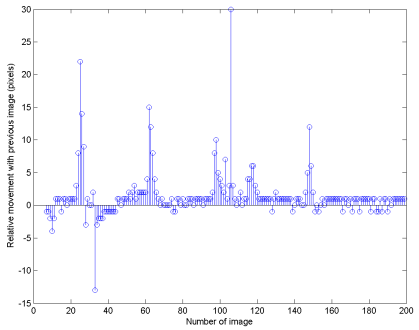


(c) Image 30

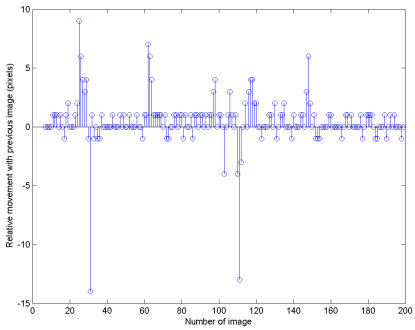


(d) Image 31

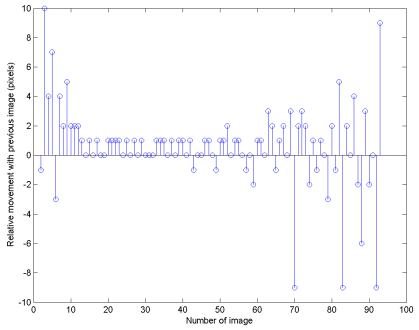
Fig. 4.4. Original Images (a, b) from *ruben* Data Set; Output Images (c, d)



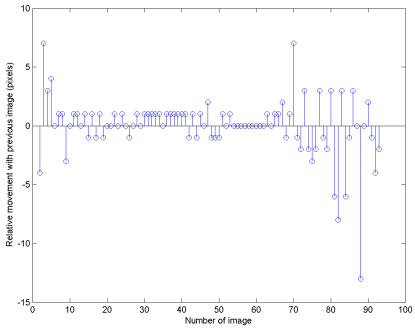
(a) Horizontal movement of *vascuar flow*



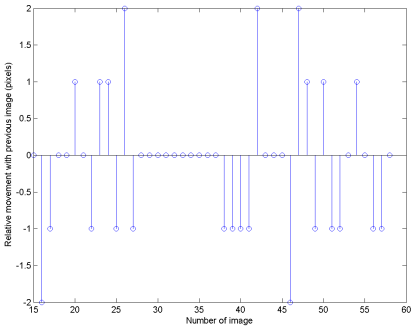
(b) Vertical movement of *vascuar flow*



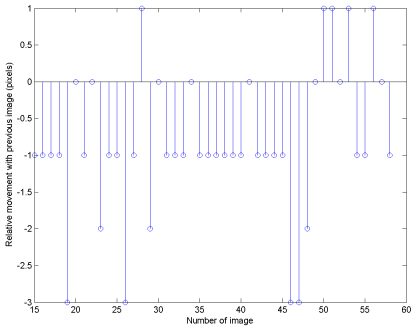
(c) Horizontal movement of *nuclei*



(d) Vertical movement of *nuclei*

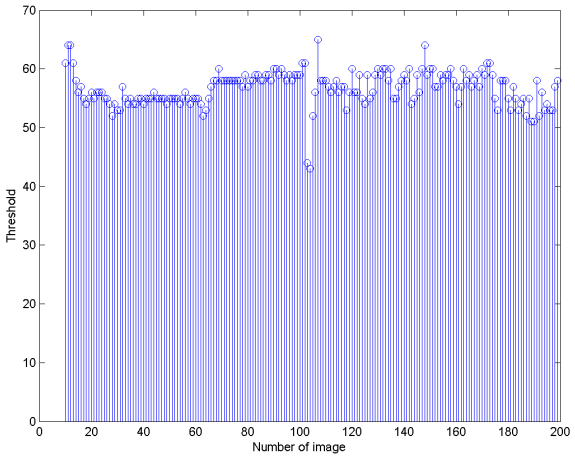


(e) Horizontal movement of *ruben*

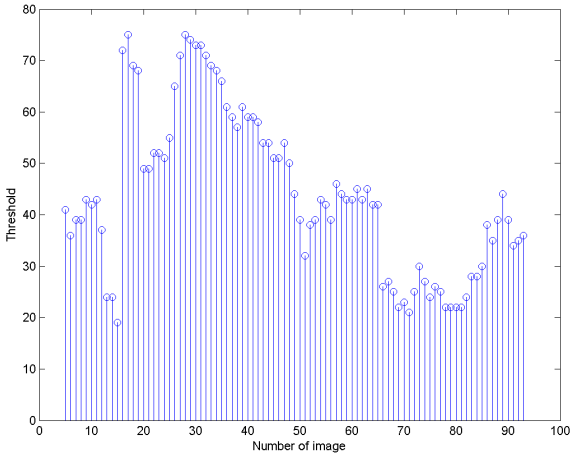


(f) Vertical movement of *ruben*

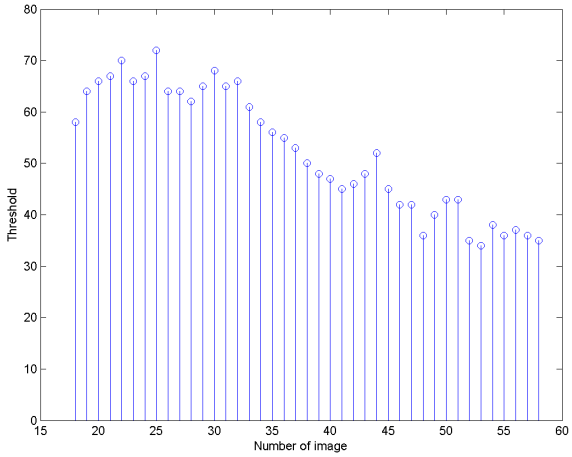
Fig. 4.5. Registration results



(a) *vascuar flow* data set

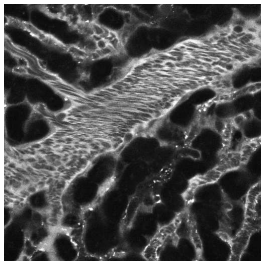


(b) *vascuar flow* data set

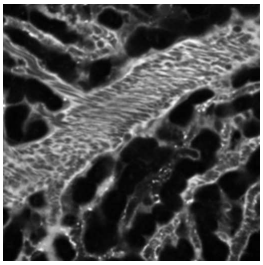


(c) *nuclei* data set

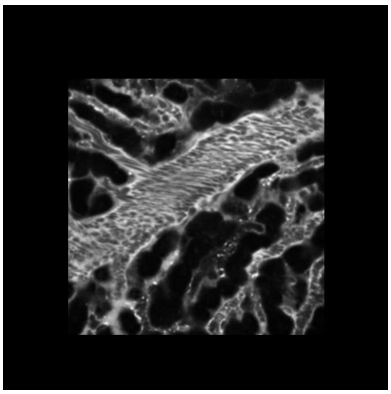
Fig. 4.6. K-means Threshold vs. Image Number



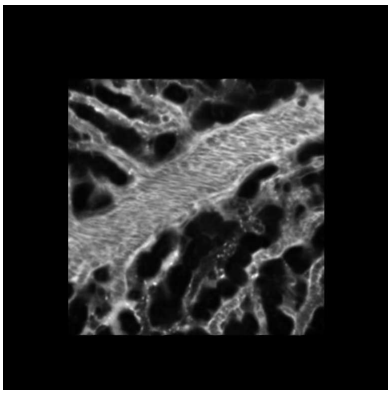
(a) Original image



(b) 2-D spatial low-pass filter



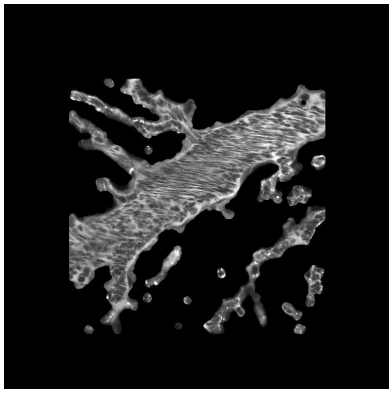
(c) Registration



(d) 1-D low-pass filtering

(e) K-means

(f) Morphological filtering



(g) Binary restoration

Fig. 4.7. Example of intermediate output images

5. USER ADJUSTABLE PARAMETERS IN PROPOSED APPROACH

The second approach has completely automated the image processing of the data sets. However, user interaction may be introduced to modify certain parameters of the analysis. Specifically, in the second approach, the following four parameters may be modified by the user: 2-D spatial low-pass filter, registration, 1-D low-pass filter, and morphological filter. These parameters may be adjusted depending on specific input data sets to produce more desired results. A diagram depicting these user adjustable parameters is shown in Figure 5.1.

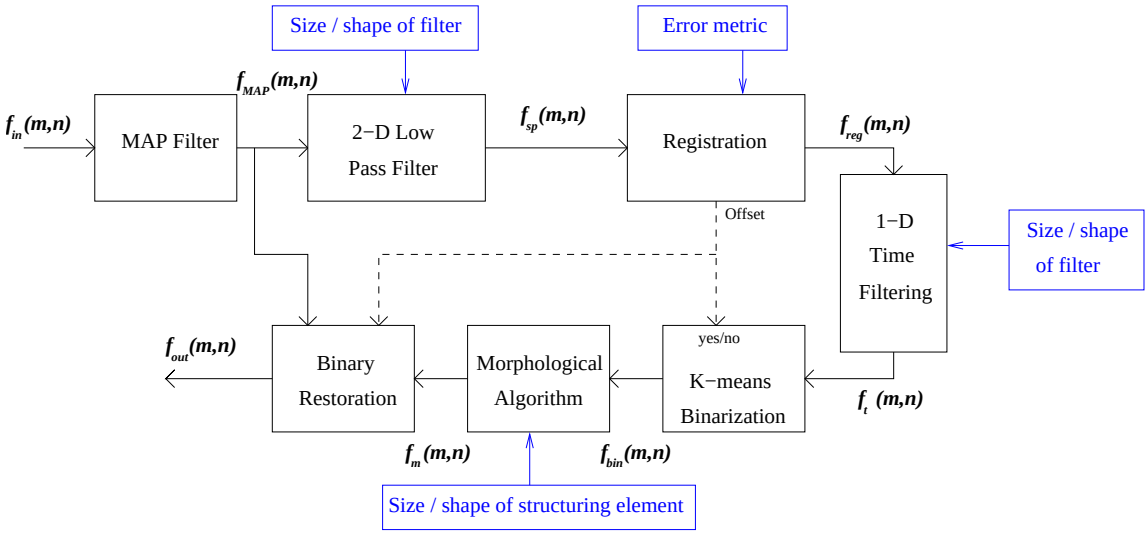


Fig. 5.1. User Adjustable Parameters in Block Diagram

5.1 2-D Spatial Low-Pass Filtering

The simplicity of the 2-D low-pass blurring filter lends itself well to this application. Other filters such as Gaussian smoothing [31] have a higher computation cost without necessarily producing improved results. An edge-preserving smoothing filter has also been studied. However, due to the characteristics of image noise, the edge preserving filter may perform poorly, as edges of objects are not well defined in the original data images.

Additionally, the form of the 2-D spatial low-pass filter introduced previously was selected experimentally. The coefficients and dimension of the filter may be modified to produce slightly different blurring effects. A filter with impulse response:

$$h = \begin{bmatrix} 0.03 & 0.03 & 0.03 & 0.03 & 0.03 \\ 0.03 & 0.05 & 0.05 & 0.05 & 0.03 \\ 0.03 & 0.05 & 0.12 & 0.05 & 0.03 \\ 0.03 & 0.05 & 0.05 & 0.05 & 0.03 \\ 0.03 & 0.03 & 0.03 & 0.03 & 0.03 \end{bmatrix} \tag{5.1}$$

with $h(0,0) = 0.12$ being the center element of the filter, has also shown to produce acceptable results.

5.2 Registration

In addition to the mean squared error metric used, the following metrics have been evaluated:

$$\begin{aligned} (u,v) &= \underset{(u,v)}{\operatorname{argmax}} \left\{ \sum_m \sum_n x_{i-1}(m,n) x_i(m-u, n-v) \right\} \\ (u,v) &= \underset{(u,v)}{\operatorname{argmin}} \left\{ \sum_m \sum_n |x_{i-1}(m,n) - x_i(m-u, n-v)| \right\} \end{aligned} \tag{5.2}$$

where $x_{i-1}(m,n)$ and $x_i(m,n)$ denote the pixels at location (m,n) within the reference image and target image, respectively. The registration was performed using

both metrics, where the results were visually identical. If the reference images are binary, computation will be greatly reduced using the first metric, where a logical `AND` operation is significantly faster than a subtraction operation. However, for grayscale reference images, the mean square error metric is a common metric with a large capture radius [9].

Additionally, the mutual information metric has been evaluated [32]. The mutual information metric consists of maximizing the following:

$$(u, v) = \underset{(u,v)}{\operatorname{argmax}} \{I(x_{i-1}(m, n), x_i(m - u, n - v))\} \tag{5.3}$$

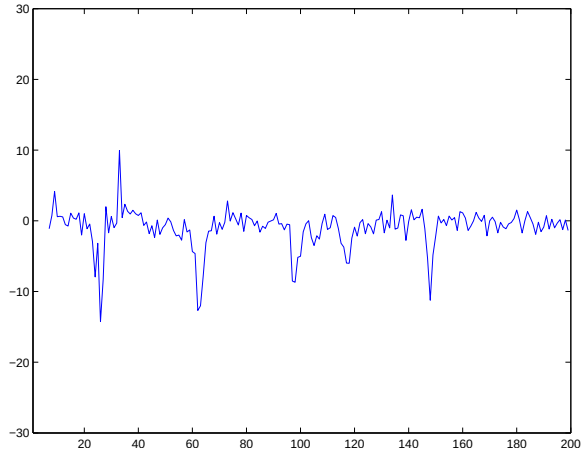
where

$$I(x(m,n), y(m-u, n-v)) = H(x(m,n)) + H(y(m-u, n-v)) - H(x(m,n), y(m-u, n-v)) \tag{5.4}$$

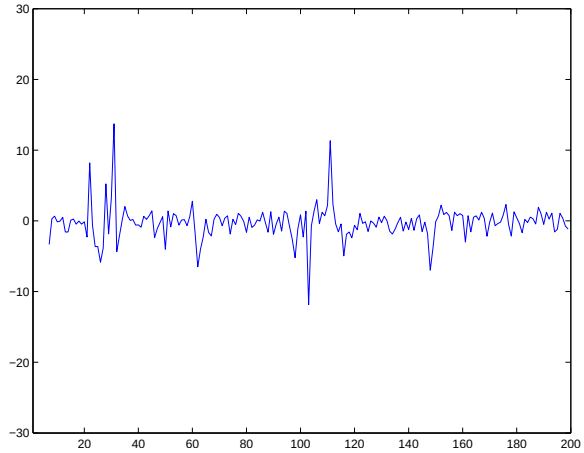
and the operator H denotes the entropy. The mutual information implementation from Mattes et al. [33] has been used on this approach. Figure 5.2 shows the registration results. Unlike the mean squared error metric, when there is significant movement between images, the mutual information metric does not produce large peaks in the metric value. Additionally, the mutual information metric values contains more variation than the mean squares metric values. More information about this technique can be found in [34, 35].

Surprisingly, worse results have been obtained using this metric. However, it is known that the mutual information metric produces optimal results when images of different modalities are involved. In such cases, metrics based on direct comparison of gray levels are not applicable. It has been extensively shown that metrics based on the evaluation of mutual information are well suited for overcoming the difficulties of multi-modality registration.

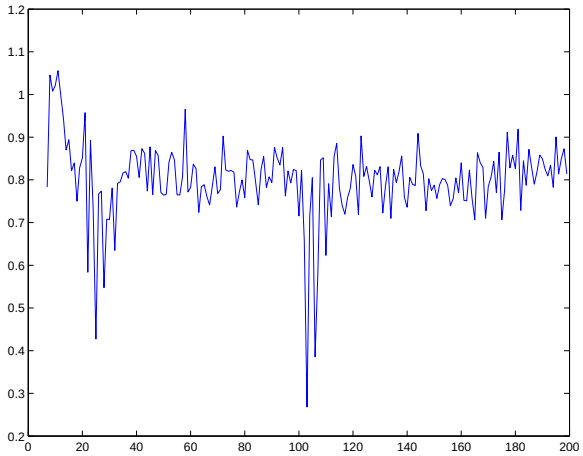
If the system is not required to be real-time where the entire stack of images is available, the system can be implemented using non-causal methods. Therefore, there is the advantage of being able to analyze the entire data set before creating



(a) Row (vert.) displacement



(b) Column (hor.) displacement



(c) Mutual information metric

Fig. 5.2. Registration Results from *Vascular Flow* Data Set Using Mutual Information Metric

any output images. In the second approach introduced previously, the output image size is unknown unless the maximum total displacement due to registration is known. However, this information is not available until the entire data set has been analyzed. For the system to be causal, an assumption about the output image size had to be made, and the placement of the first image in the output image had to be chosen. However, with a non-causal approach, the system is able to analyze the entire data set before producing any output images. After all images have been analyzed, the maximum displacement across all images can be determined to set the output image size so that no source images are truncated.

5.3 1-D Low-Pass Filtering

By continuing through with a non-causal approach, the 1-D vertical low-pass filter can be made non-causal as well. The filter can now be given by:

$$T_n = \frac{1}{10}I_{n-2} + \frac{2}{10}I_{n-1} + \frac{4}{10}I_n + \frac{2}{10}I_{n+1} + \frac{1}{10}I_{n-2} \tag{5.5}$$

where T_n denotes the n-th filtered image corresponding to the n-th input image I_n and its two preceding and two succeeding images I_{n-2}, I_{n-1}, I_{n+1}, and I_{n+2}, respectively. Implementing this filter requires analyzing all input images twice, once for the spatial low-pass filter, and a second time for the 1-D low-pass filter. For images at the beginning and end of the sequence, where not all of the four previous images are available, the n-th filtered image is obtained by:

$$\begin{aligned} T_n &= \frac{1}{3}I_{n-1} + \frac{1}{3}I_n + \frac{1}{3}I_{n+1} \\ T_n &= I_n \end{aligned} \tag{5.6}$$

Additionally, similar to the 2-D spatial low-pass filter, the form of the 1-D low-pass filter introduced previously was also selected experimentally. The coefficients and length of the filter, for both the causal and non-causal cases, may be modified to produce slightly different smoothing effects.

5.4 Morphological Filtering

The size and shape of the structuring elements used in the morphological filter were also selected experimentally. By altering the size and shape of the structuring elements for the opening and closing operations, the user can remove smaller or larger objects, and remove smaller or larger holes, respectively, from the binary mask.

6. CONCLUSIONS AND FUTURE WORK

This document described two techniques to segment microscopy image data sets. The second approach is deemed to be a more comprehensive method in addressing the issues and is chosen as the preferred method over the second approach. Experimental results indicated that this method was promising in segmenting objects in time series data sets as well as data sets comprised of images acquired at increasing tissue depths. Nevertheless, evaluation of results proves to be difficult without having any ground truth data available. This makes evaluation significantly more subjective and less objective. Future work will involve developing registration schemes that correct rotations and other non-linear distortions, especially those prevalent in the new data sets.

All methods have been implemented in both MATLAB and in C using ITK [1]. The computational complexity of the MATLAB implementation is extremely high, particularly for the registration component of the process. However, the ITK implementation has a speedup factor of 200 over the MATLAB implementation. All figures of analyzed results in Section 4 have been created with the ITK implementation. A complete guide of this image processing toolkit is presented in [22].

Lastly, it is becoming increasingly common in medical imaging to analyze high dimensional images. For this reason, future work can be based on the implementation of higher dimensional filters. The 2-D morphological filter may be replaced with a 3-D filter which takes into account the time/depth dimension. However, it is important to note that, registration must be performed on all the frames before applying any type of 3-D filter due to the specimen movements.

[1]National Library of Medicine Insight Segmentation and Registration Toolkit (http://www.itk.org)

Also, it's important to note that in future implementations; it can be useful, processing the different channels separately for obtaining a better performance in terms of segmentation and registration. For example, on *hmang* data set (see Figure B.1), red channel could be used for segmentation while green channel (Undesired objects) can be used for registration. Blood vessels (on red) are not useful for registration due to the movement of the blood inside them. Endosomes and tubular lumen, with a high ratio of green to red signal, however could be used for registration.

LIST OF REFERENCES

LIST OF REFERENCES

[1] R. L. Bacallao, W. Yua, K. W. Dunna, and C. L. Phillips, "Novel light microscopy imaging techniques in nephrology," *Current Opinion in Nephrology and Hypertension*, vol. 12, no. 1, pp. 455–481, 2003.

[2] B. Luck, K. Carlson, A. Bovik, and R. Richards-Kortum, "An image model and segmentation algorithm for reflectance confocal images of in vivo cervical tissue," vol. 14, pp. 1265–1276, Sept. 2005.

[3] A. Dufour, V. Shinin, S. Tajbakhsh, N. Guillen-Aghion, J.-C. Olivo-Marin, and C. Zimmer, "Segmenting and tracking fluorescent cells in dynamic 3-D microscopy with coupled active surfaces," vol. 14, pp. 1396–1410, Sept. 2005.

[4] M. Kyan, L. Guan, M. Arnison, and C. Cogswell, "Feature extraction of chromosomes from 3-D confocal microscope images," vol. 48, pp. 1306–1318, Nov 2001.

[5] K. Nandy, P. Gudla, and S. Lockett, "Automatic segmentation of cell nuclei in 2D using dynamic programming," in *Proceedings of 2nd Workshop on Microsopic Image Analysis with Applications in Biology, Piscataway, NJ, USA* (D. Metaxas, J. Rittscher, S. Lockett, and T. Sebastian, eds.), 2007.

[6] X. Chen, X. Zhou, and S. Wong, "Automated segmentation, classification, and tracking of cancer cell nuclei in time-lapse microscopy," vol. 53, pp. 762–766, April 2006.

[7] A. Dima, M. Scholz, and K. Obermayer, "Automatic segmentation and skeletonization of neurons from confocal microscopy images based on the 3-D wavelet transform," vol. 11, pp. 790–801, Jul 2002.

[8] S. Mallat and S. Zhong, "Characterization of signals from multiscale edges," vol. 14, pp. 710–732, Jul 1992.

[9] L. Brown, "A survey of image registration techniques," *ACM Computing Surveys*, vol. 24, no. 4, pp. 325–376, 1992.

[10] B. Zitová and J. Flusser, "Image registration methods: A survey," *Image and Vision Computing*, vol. 21, pp. 977–1000, October 2003.

[11] E. Meijering, I. Smal, O. Dzyubachyk, and J. Olivo-Marin, *Time-Lapse Imaging*, ch. 15, pp. 401–440. Elsevier Academic Press, Burlington, MA, USA, 2008.

[12] D. McCullough, P. Gudla, B. Harris, J. Collins, K. Meaburn, M.-A. Nakaya, T. Yamaguchi, T. Misteli, and S. Lockett, "Segmentation of whole cells and cell nuclei from 3-D optical microscope images using dynamic programming," vol. 27, pp. 723–734, May 2008.

[13] A. Bazen and S. Gerez, "Fingerprint matching by thin-plate spline modeling of elastic deformations," *Pattern Recognition*, vol. 36, no. 8, pp. 1859–1867, 2003.

[14] K. Rohr, H. Stiehl, R. Sprengel, T. Buzug, W. J., and M. Kuhn, "Landmark-based elastic registration using approximating thin-plate splines," vol. 20, p. 526534, June 2001.

[15] J. Pánek and J. Vohradský, "Point pattern matching in the analysis of two-dimensional gel electropherograms," *Electrophoresics*, vol. 20, no. 18, pp. 3483–3491, 1999.

[16] T. Akutsu, K. Kanaya, A. Ohyama, and A. Fujiyama, "Point matching under non-uniform distortions," *Discrete Applied Mathematics*, vol. 127, no. 1, pp. 5–21, 2003.

[17] D. Suter and F. Chen, "Left ventricular motion reconstruction based on elastic vector splines," vol. 19, pp. 295–305, April 2000.

[18] J. Kybic and M. Unser, "Fast parametric elastic image registration," vol. 12, pp. 1427–1442, November 2003.

[19] P. Salama, "A least squares approach to estimating the probability distribution of unobserved data in multiphoton microscopy," in *Proceedings of the SPIE Conference on Computational Imaging VI* (C. Bouman, E. Miller, and I. Pollak, eds.), vol. 6814, 2008.

[20] R. Haralick, S. Sternberg, and X. Zhuang, "Image analysis using mathematical morphology," vol. 9, pp. 532–550, July 1987.

[21] The MathWorks, Inc., "Morphological Reconstruction, Image Processing Toolbox."

[22] L. Ibáñez, W. Schroeder, L. Ng, and J. Cates, *The ITK Software Guide*. Kitware, Inc., 2005.

[23] L. Ibanez, L. Ng, J. Gee, and S. Aylward, "Registration patterns: the generic framework for image registration of the insight toolkit," pp. 345–348, 2002.

[24] T. Lehmann, C. Gonner, and K. Spitzer, "Survey: Interpolation methods in medical image," vol. 18, pp. 1049–1075, November 2001.

[25] G. Grevera and J. Udupa, "An objective comparison of 3-d image interpolation methods," vol. 17, pp. 642–652, August 1992.

[26] M. Arnold, S. Fink, D. Grove, M. Hind, and P. Sweeney, "A survey of adaptive optimization in virtual machines," vol. 93, pp. 449–466, February 2005.

[27] J. Swisher, P. Hyden, S. Jacobson, and L. Schruben, "A survey of simulation optimization techniques and procedures," vol. 1, pp. 119 – 128, December 2000.

[28] G. Penney, J. Weese, J. Little, P. Desmedt, D. Hill, and D. hawkes, "A comparison of similarity measures for use in 2-d-3-d medical image registration," vol. 17, pp. 586–595, August 1992.

[29] I. Kim, S. Yang, P. Le Baccon, E. Heard, Y.-C. Chen, D. Spector, C. Kappel, R. Eils, and K. Rohr, "Non-rigid temporal registration of 2D and 3D multi-channel microscopy image sequences of human cells," *Biomedical Imaging: From Nano to Macro, 2007. ISBI 2007. 4th IEEE International Symposium on*, pp. 1328–1331, April 2007.

[30] T. Ridler and S. Calvard, "Picture thresholding using an iterative selection method," vol. 8, pp. 630–632, Aug. 1978.

[31] C. Wählby, I.-M. Sintorn, F. Erlandsson, G. Borgefors, and E. Bergtsson, "Combining intensity, edge and shape information for 2D and 3D segmentation of cell nuclei in tissue sections," *Journal of Microscopy*, vol. 215, pp. 67–76, July 2004.

[32] J. Pluim, J. Maintz, and M. Viergever, "Mutual-information-based registration of medical images: a survey," vol. 22, pp. 986–1004, Aug. 2003.

[33] D. Mattes, D. Haynor, H. Vesselle, T. Lewellen, and W. Eubank, "Non-rigid multimodality image registration," *Medical imaging 2001: Image processing*, pp. 1609–1620, 2001.

[34] P. Viola and I. Wells, W.M., "Alignment by maximization of mutual information," *Computer Vision, 1995. Proceedings., Fifth International Conference on*, vol. 17, pp. 16–23, June 1995.

[35] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," vol. 16, pp. 187–198, April 1997.

[36] M. Minsky, "Microscopy apparatus. united states patent #3013467," 1957.

[37] J. Pawley, *Handbook of biological confocal microscopy*. New York: Plenum, 1995.

[38] T. Wilson, *Confocal microscopy: Basic principles and architectures.* New York: Wiley-Liss, 2002.

[39] W. Denk, J. Strickler, and W. Webb, "Two-photon laser scanning fluorescence microscopy," *Science*, pp. 73–76, April 1990.

APPENDIX

A. MICROSCOPY IMAGING

This chapter describes the techniques that are currently in use for acquiring the images that have been analyzed on that project. Microscopy involve the diffraction, reflection, or refraction of electromagnetic radiation interacting with the subject of study, and the subsequent collection of this scattered radiation in order to build up an image. An extended review of the novel light microscopy imaging techniques in nephrology can be found in [1]. Application of these new technologies to problems specific to areas of interest to the nephrology community presents both opportunities and challenges. The challenges arise from the unique anisotropy of the kidney for light in the visible spectrum. Although many tissues have a homogeneous refractive index that simplifies image acquisition, fluorescently labeled structures in the kidney are more difficult to image due to a heterogeneous network of intertwined or anastomosing tubules, vessels and nerves with intervening stroma [1].

Confocal microscopy This optical imaging technique was first presented by Minsky [36] in 1951. A confocal microscope uses point illumination and a pinhole in an optically conjugate plane in front of the detector to eliminate out-of-focus information. Only the light within the focal plane can be detected, so the image quality is much better than that of wide-field images. As only one point is illuminated at a time in confocal microscopy, 2D or 3D imaging requires scanning over a regular raster in the specimen. The thickness of the focal plane is defined mostly by the square of the numerical aperture of the objective lens, and also by the optical properties of the specimen and the ambient index of refraction. These microscopes also are able to see into the image by taking images at different depths. Figure A.1 shows a schematic diagram of the confocal principle described before. Several papers describe this technique in detail [37, 38]

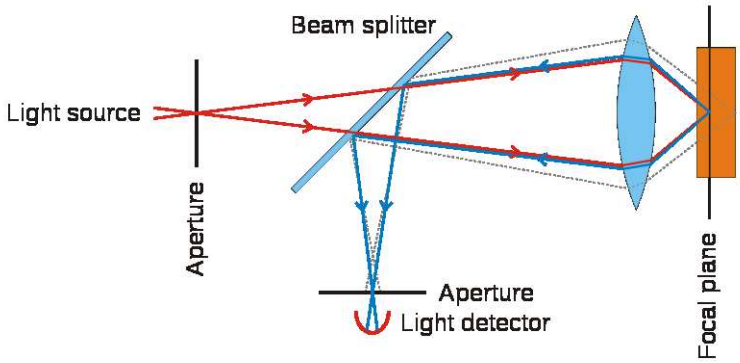


Fig. A.1. Laser scanning confocal microscope diagram

Multiphoton Fluorescence Microscope It uses pulsed long-wavelength laser light to excite fluorophores within the specimen being observed. The fluorophore absorbs the energy from two long-wavelength photons which must arrive simultaneously in order to excite an electron into a higher energy state, from which it can decay, emitting a fluorescence signal. It differs from traditional fluorescence microscopy in which the excitation wavelength is shorter than the emission wavelength, as the summed energies of two long-wavelength exciting photons will produce an emission wavelength shorter than the excitation wavelength. Multiphoton fluorescence microscopy has similarities to confocal laser scanning microscopy. Both use focused laser beams scanned in a raster pattern to generate images, and both have an optical sectioning effect. Unlike confocal microscopes, multiphoton microscopes do not contain pinhole apertures, which give confocal microscopes their optical sectioning quality. The two-photon excitation microscope [39] is a special variant of the multiphoton fluorescence microscope. The major advantage of this technique in front of confocal microscopy is the deeper tissue penetration. Figure A.2 shows the diagram of a two-photon microscope. There we can see the differences between Multiphoton and Confocal microscope mentioned above.

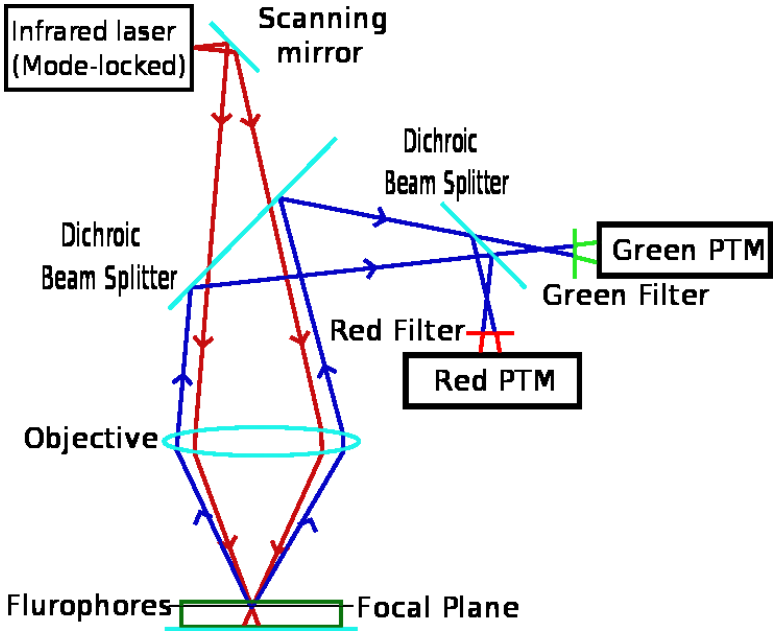


Fig. A.2. Two-photon microscope diagram

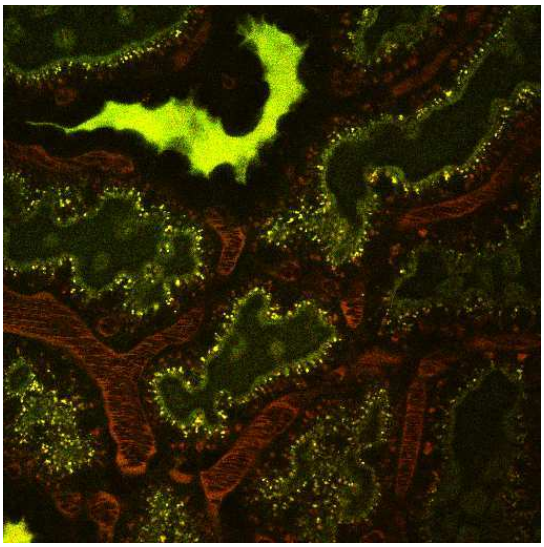
B. EXTENDED RESULTS

New data sets similar to the *vascular flow* data set are presented now, also consisting of liver structures of a rat across a series of time instances. These data sets will be referred to as the *hmang* data set, *ruben 70* data set, *glomerulus* data set, and *liver vasculature* data set. Example images from each data set are shown in Figure B.1. See Section C for the complete data set. The objective is to segment vasculature structures in all of these new data sets. Vascular structures are shown predominantly in red in the *hmang* and *ruben70* data sets, and predominantly in green in the *liver vasculature* data set. Therefore, regarding the *ruben70* and *liver vasculature* data sets in color, the respective color component is extracted. Then, only this color component is analyzed using the proposed approach, just as the *nuclei* and *vascular flow* data sets were analyzed. However, the *hmang* data set must undergo additional preprocessing due to the image set characteristics and desired objects to be segmented. This preprocessing is called *green component clamping*, and is explained in the following section.

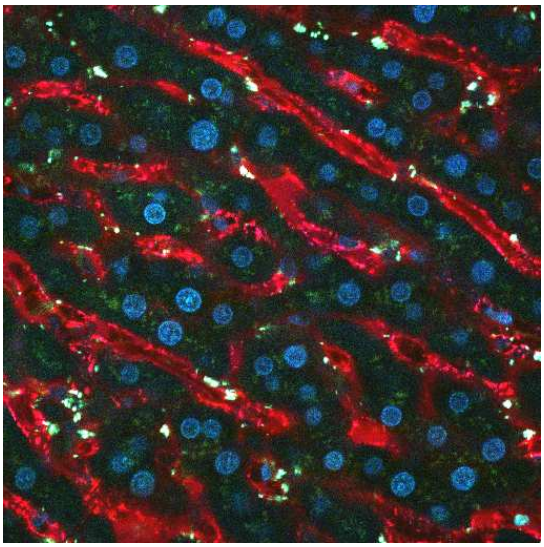
B.1 Green Component Clamping

The images in the *hmang* data set contain many yellow structures (endosomes and a tubular lumen have a combination of red and green fluorescence), which are undesired to be segmented. Therefore, simply using the red component of these images will cause segmentation of unwanted objects in addition to the vasculature. Therefore, to reduce the visibility of yellow structures, a new image is constructed by identifying unique ratio of red to green signal as follows:

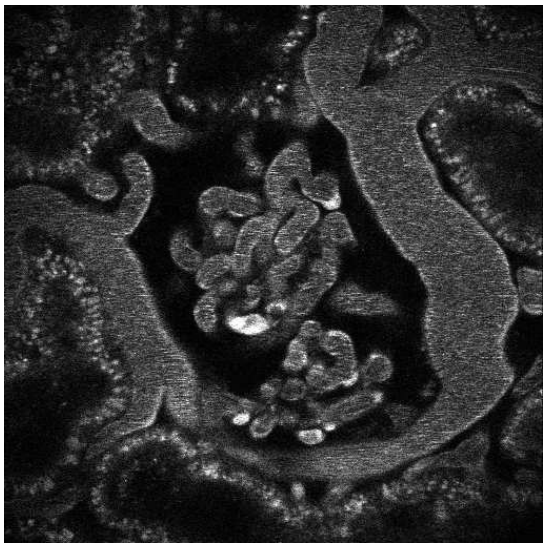
$$\tilde{f}_r(m,n) = \begin{cases} 0 & \text{if } \frac{f_r(m,n)}{f_g(m,n)} < \frac{1}{\alpha} \\ f_r(m,n) & \text{else} \end{cases} \tag{B.1}$$



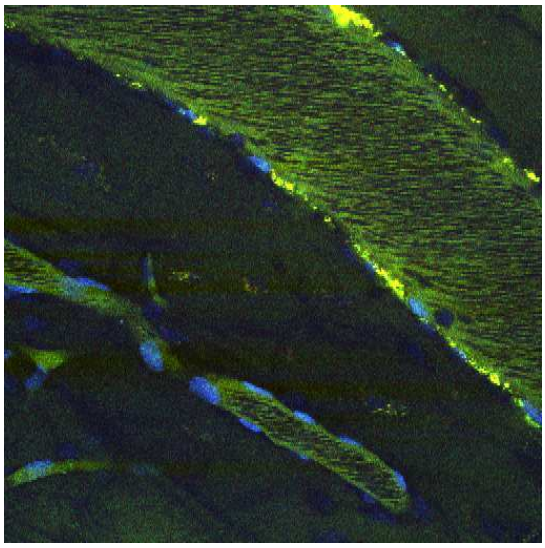
(a) Image 13 of *hmang* data set



(b) Image 26 of *ruben70* data set



(c) Image 20 of *glomerulus* data set



(d) Image 32 of *liver vasculature* data set

Fig. B.1. Example of New Images Used in Our Study

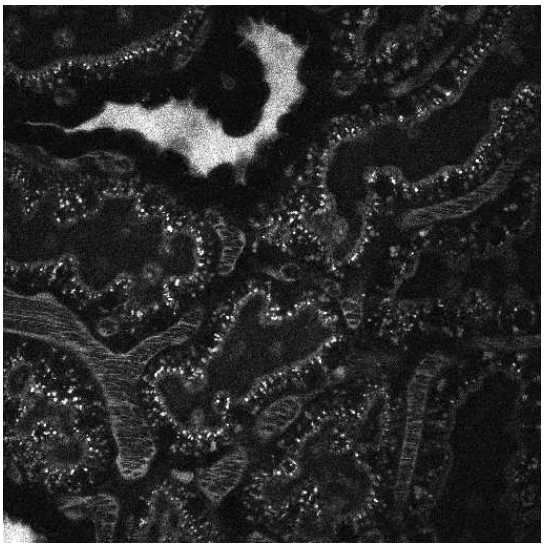
where f_r and f_g are the red and green components, respectively, of the current image in the stack being analyzed, and α is a constant greater than one. Selecting $\alpha = 3$ has been experimentally shown to produce sufficient results. Then, $\tilde{f}_r$ is analyzed using the proposed approach, but only with regard to the *hmang* data set due to its unique characteristics. The effects of this preprocessing can be seen in Figure B.2. Segmentation and registration results for *hmang* data set are shown on Figure B.3. Figures of the registration and thresholding are shown in Figure B.4. There, we can see that the horizontal and vertical movements are almost constant and equal to 1 pixel in the positive and negative axes respectively. The value of the threshold decreases on the second half of the stack due to the low light intensity on the last images.

B.2 Other results

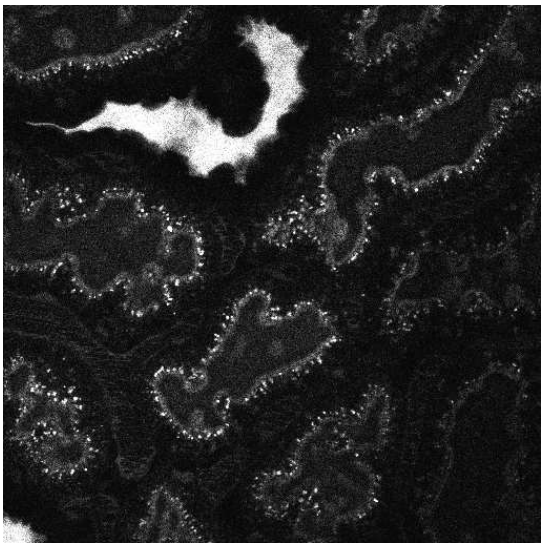
Results of *ruben70* data set are shown on Figure B.5. Red channel has been used for processing this stack of images. As it can be shown on the video (See Section C) there is some vibration on that image, so this is the nature of the horizontal and vertical movements shown in Figure B.6. Both movements can be represented by a sinusoid of a 3-4 images of period.

Glomerulus data set results are presented on Figure B.7. There we can appreciate that some non desirable structures have been segmented (See endosomes on the bottom left of the images). Also, as can be shown in the complete stack (See Section C, we can see that as the light intensity is decreasing (we are out of focus), it's even more difficult segmenting some parts of the image. For example, see the center of the image 19. Regarding to the movements on Figure B.8, we cannot appreciate a arranged movement in some direction as it was clear on the *ruben70* data set. The value of the threshold is almost constant for the entire stack.

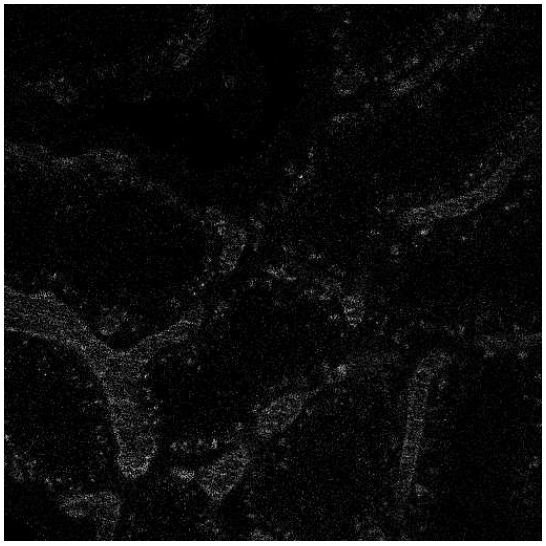
The last stack of images we are showing is the *liver vasculature* data set. Image results are shown on Figure B.9. Due to the noise present on the images and the low



(a) Red component

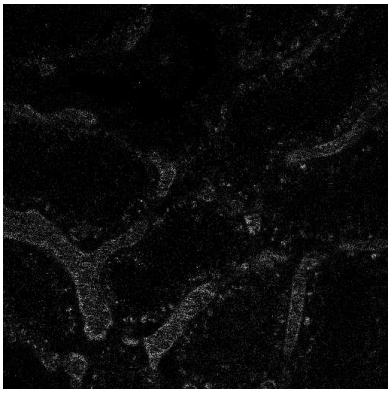


(b) Green component

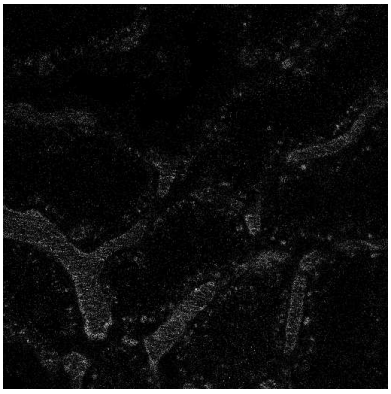


(c) New image

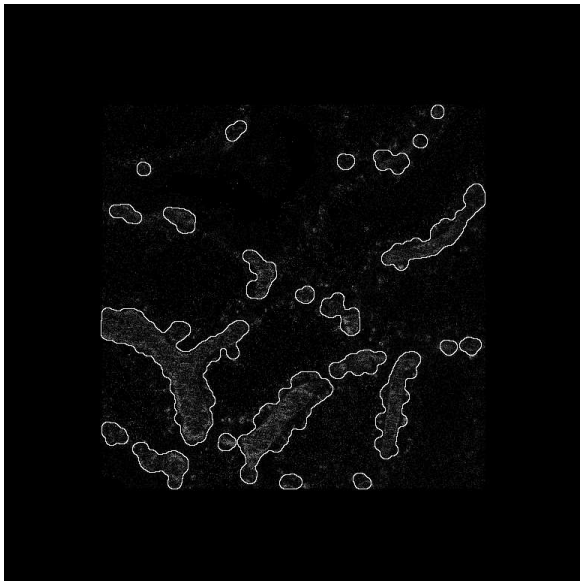
Fig. B.2. Image 10 of *hmang* data set



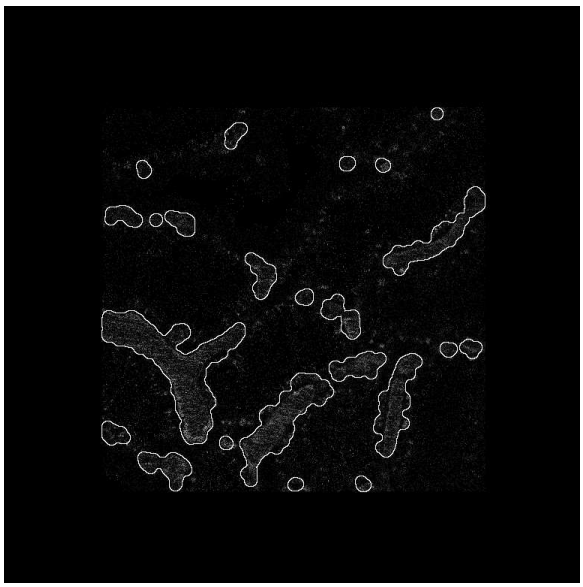
(a) Image 13



(b) Image 14

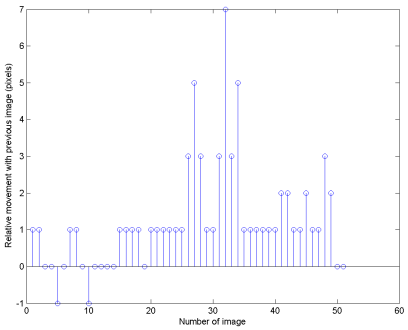


(c) Image 13

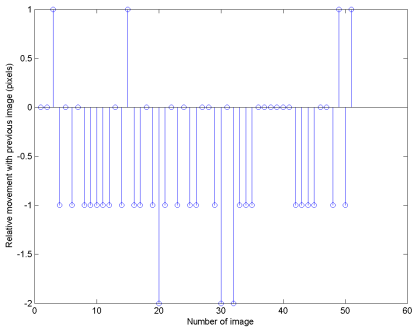


(d) Image 14

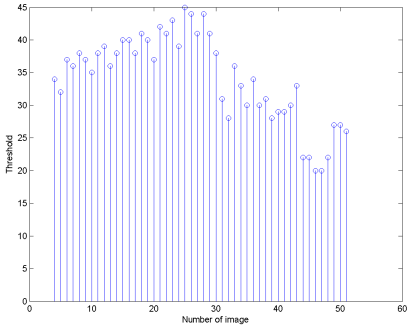
Fig. B.3. Original Images (a, b) from *hmang* Data Set; Output Images (c, d)



(a) Horizontal movement

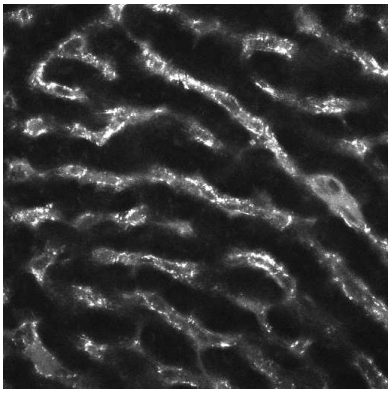


(b) Vertical movement

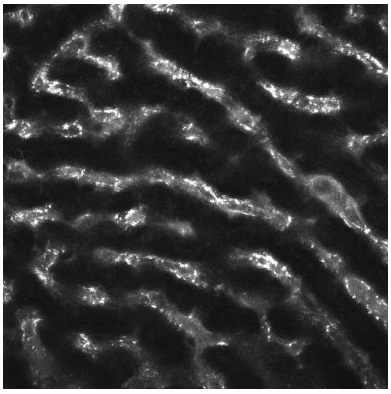


(c) K-means Threshold

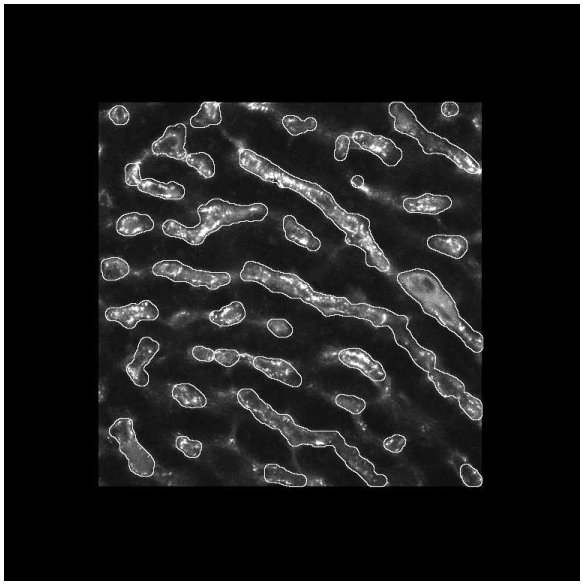
Fig. B.4. Graphs of *hmang* data set



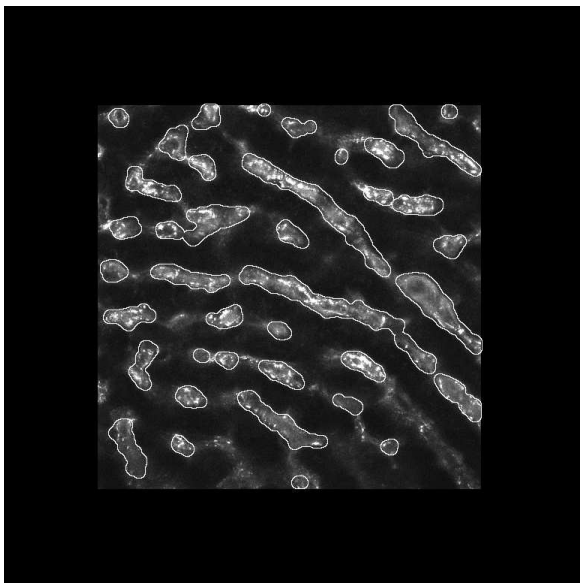
(a) Image 46



(b) Image 47

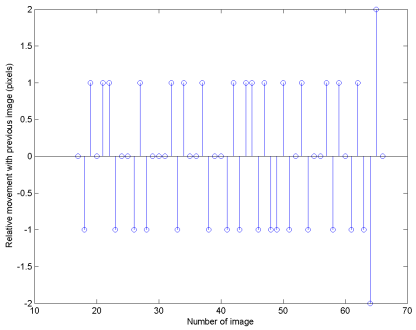


(c) Image 46

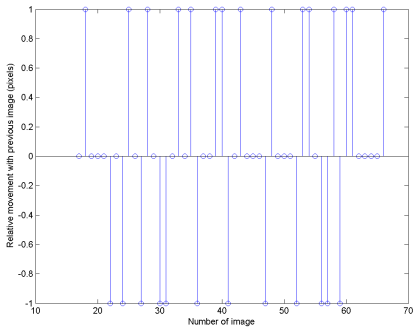


(d) Image 47

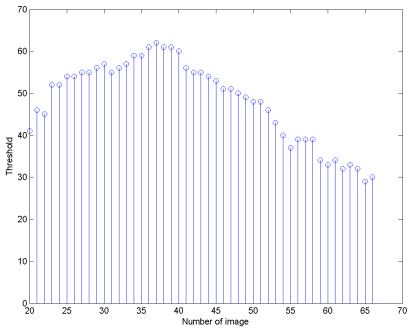
Fig. B.5. Original Images (a, b) from *ruben70* Data Set; Output Images (c, d)



(a) Horizontal movement

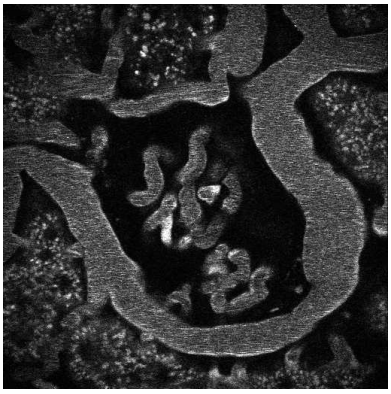


(b) Vertical movement

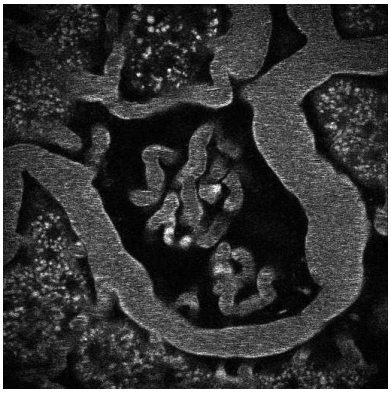


(c) K-means Threshold

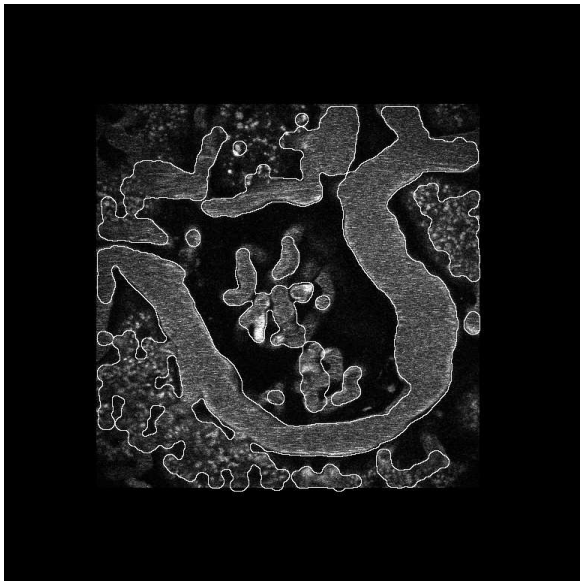
Fig. B.6. Graphs of *ruben70* data set



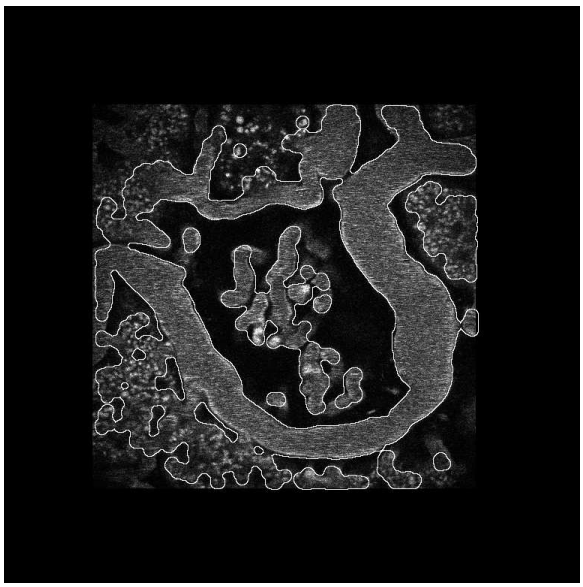
(a) Image 13



(b) Image 14

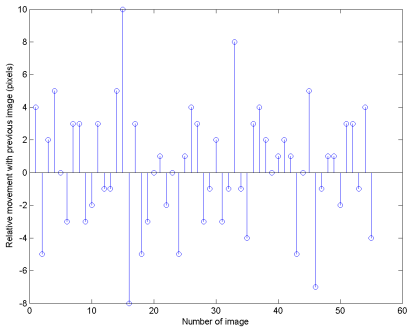


(c) Image 13

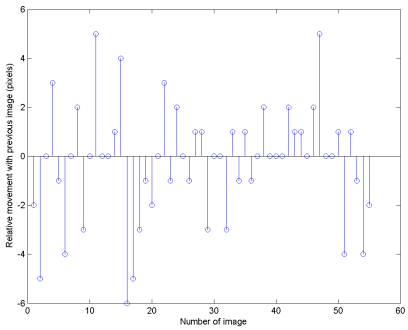


(d) Image 14

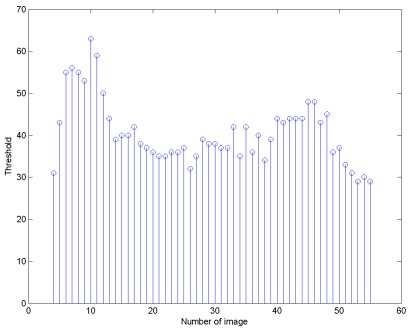
Fig. B.7. Original Images (a, b) from *glomerulus* Data Set; Output Images (c, d)



(a) Horizontal movement



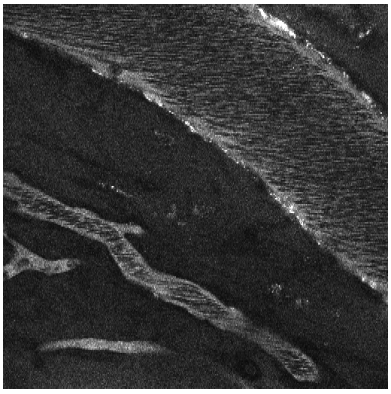
(b) Vertical movement



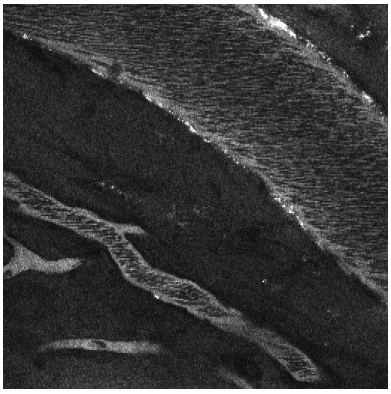
(c) K-means Threshold

Fig. B.8. Graphs of *glomerulus* data set

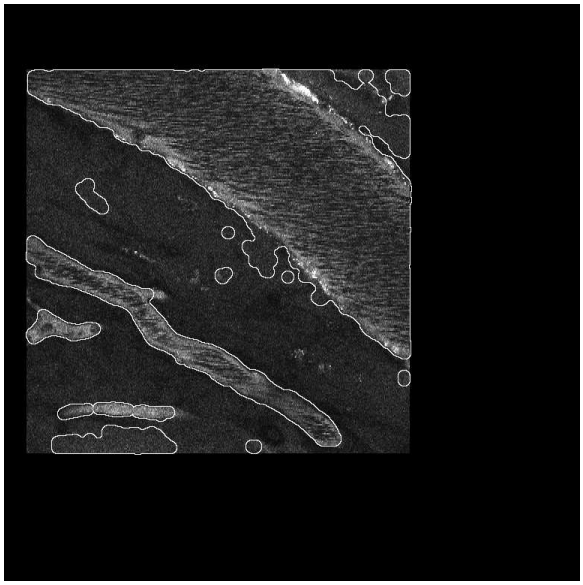
light intensity, after image 40, the method fails with the segmentation. Figure B.10 shows the horizontal and vertical movements and the threshold values after processing. There is a regular movement of one pixel in the positive value of both axes. Finally, the threshold try to adapt to the light intensity but as it's mentioned above, it fails.



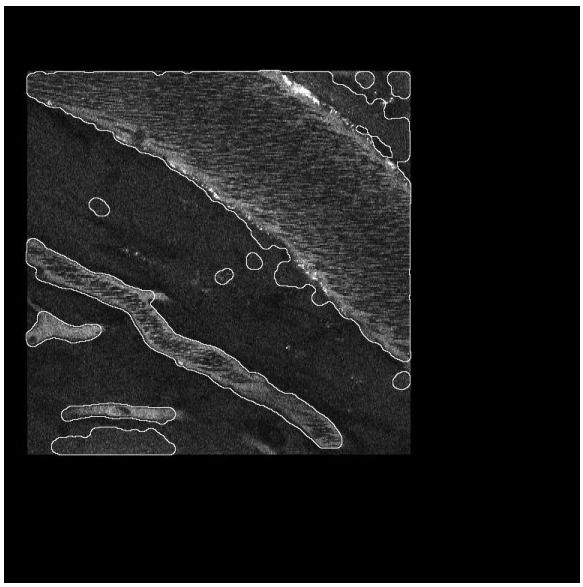
(a) Image 29



(b) Image 30

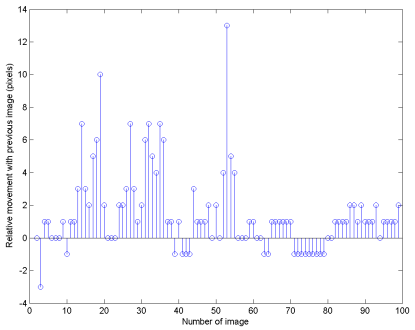


(c) Image 29

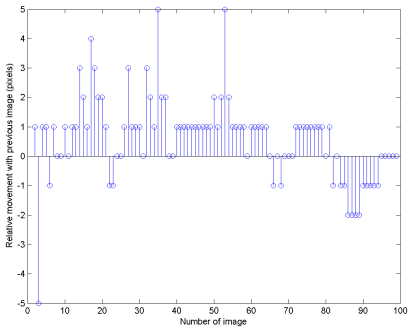


(d) Image 30

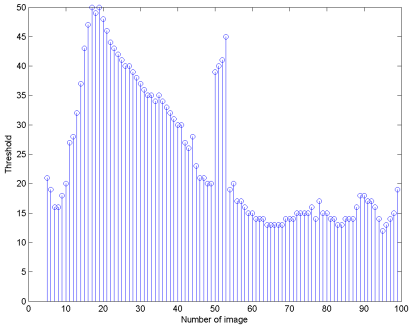
Fig. B.9. Original Images (a, b) from *liver vasculature* Data Set; Output Images (c, d)



(a) Horizontal movement



(b) Vertical movement



(c) K-means Threshold

Fig. B.10. Graphs of *vasculature* data set

C. WEB PAGE STRUCTURE

The segmentation effectiveness results can clearly be shown through the representation of some images. On the other hand, the graphics presented on section Section 4 and Section B, regarding to the registration process, show the direction and intensity of the movements but they don't show how the stack is lined up. For this reason, we refer to the following web page were, videos of the stacks can be easily played:
https://redpill.ecn.purdue.edu/~kidney/thesisSerrano/

Also, all the stacks presented on Table 4.1 have been processed and they are accessible from the web page. It can be seen there stacks of images similar to the stacks presented on this document.

The main page of the link divides the files into 2 groups: images and videos. Take in mind that just images have been processed with the method that this thesis proposes. The videos have been created by putting the images into a pile and then compression methods have been used. Therefore, just the registration process has to be observed through the videos. The problem of the segmentation is better seen through the image folder.

Each folder is then split into the original images or videos and the processed. When clicked each of this folder, we access to the files which can be downloaded or visualized on the web browser.

Finally, regarding to the segmentation process, for a clear understanding of the results, the images included are not the output of the process but the originals with the boundary of the segmentation mask. It allows the observer to objectively qualify the segmentation.