2008

A novel automated approach of multi-modality retinal image registration and fusion

Hua Cao
Louisiana State University and Agricultural and Mechanical College, hcao2@lsu.edu

A NOVEL AUTOMATED APPROACH OF MULTI-MODALITY RETINAL IMAGE REGISTRATION AND FUSION

A Dissertation
Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
in
The Department of Computer Science

by
Hua Cao
B.S., Southwest University of Finance & Economics, P.R. China, 2000
M.S., Louisiana State University, 2003
May 2008

I dedicate this work to my beloved husband, Dashun.

Acknowledgements

Firstly I wish to express my utmost gratitude to my advisor Dr. Iyengar and co-chair Dr. Brener for their excellent academic guidance and open-minded research support. They have generously contributed time and patience advising me during my study and research at LSU. Working with them for six years was fruitful, from which my future will benefit. I also extend my thanks for Dr. Ye-Sho Chen and Dr. Catallou for their willingness to serve on my committee and their valuable comments to the dissertation. I particularly want to thank Dr. Kannan for his detailed technical / critical comments which enabled to improve the quality of the dissertation. I am indebted to Dr. Thompson and Dr. Khoobehi for their efficient cooperation and advising during my dissertation research. I also wish to acknowledge all other excellent professors who had taught me various fields of computer science.

Secondly, I also want to thank my fellow students Jinfeng Ning and Gaurav Khnaduja for their help and support in my dissertation research. I also express my thanks to my friends: Dr. Mengxia Zhu, Dr. Qishi Wu, Dr. Tingzhong Guo, and Lixia Li who have been helping me and giving me encouragement.

Finally I would like to give my appreciation to my parents and my grandparents for their help and understanding in my life. Everyone in my family has been really supportive of my research and dissertation writing efforts.

Table of Contents

List of Tables

List of Figures

List of Diagrams

Abstract

Biomedical image registration and fusion are usually scene dependent, and require intensive computational effort. A novel automated approach of feature-based control point detection and area-based registration and fusion of retinal images has been successfully designed and developed. The new algorithm, which is reliable and time-efficient, has an automatic adaptation from frame to frame with few tunable threshold parameters. The reference and the to-be-registered images are from two different modalities, i.e. angiogram grayscale images and fundus color images. The relative study of retinal images enhances the information on the fundus image by superimposing information contained in the angiogram image.

Through the thesis research, two new contributions have been made to the biomedical image registration and fusion area. The first contribution is the automatic control point detection at the global direction change pixels using adaptive exploratory algorithm. Shape similarity criteria are employed to match the control points. The second contribution is the heuristic optimization algorithm that maximizes Mutual-Pixel-Count (MPC) objective function. The initially selected control points are adjusted during the optimization at the sub-pixel level. A global maxima equivalent result is achieved by calculating MPC local maxima with an efficient computation cost. The iteration stops either when MPC reaches the maximum value, or when the maximum allowable loop count is reached. To our knowledge, it is the first time that the MPC concept has been introduced into biomedical image fusion area as the measurement criteria for fusion accuracy. The fusion image is generated based on the current control point coordinates when the iteration stops. The comparative study of the presented automatic registration and fusion scheme

against Centerline Control Point Detection Algorithm, Genetic Algorithm, RMSE objective function, and other existing data fusion approaches has shown the advantage of the new approach in terms of accuracy, efficiency, and novelty.

Chapter 1

Introduction

1.1 Biomedical Image Registration and Fusion Background

An urgent requirement of clinical practice for the patients' optimal studies has, more than ever before, necessitated the development and implementation of the computer-based biomedical image registration and fusion techniques. Appropriate integration / alignment of useful data obtained from separate images are frequently desired. Fusing medical images captured by different sensors (multimodal analysis), such as Computed Topography (CT), Magnetic Resonance Imaging (MRI), Ultrasonography (US), X-rays, Single Photon Emission Computed Tomography (SPECT), Magnetic Resonance Spectroscopy (MRS), and Positron Emission Topography (PET) is able to provide physicians and medical research scientists with the complementary information, and thus help them more thoroughly understand both of the functional and structural information. The fused anatomical and functional information is important for the planning of the successful surgery for the patients as well. Recent advances in instrumentation and applications of computer imaging technologies provide the infrastructure which supports combination, integration, and registration of multimodality image data. The image fusion capability is also incorporated in some radiotherapy planning systems allowing the fused data to be used in preparation, monitoring, and optimization of the radiotherapy treatment plan.

There are four major steps involved in image registration and fusion. The first step will bring the images from single or different modalities into spatial alignment, a procedure referred as

"registration" [17]. The second step "optimization" is iteratively implemented in order to achieve the optimal fusion result. The third step "fusion" is required for the integrated display of the involved images. The final step "experiments" is a painstaking work of conducting many experiments in order to verify the fusion accuracy and validate the fusion scheme.

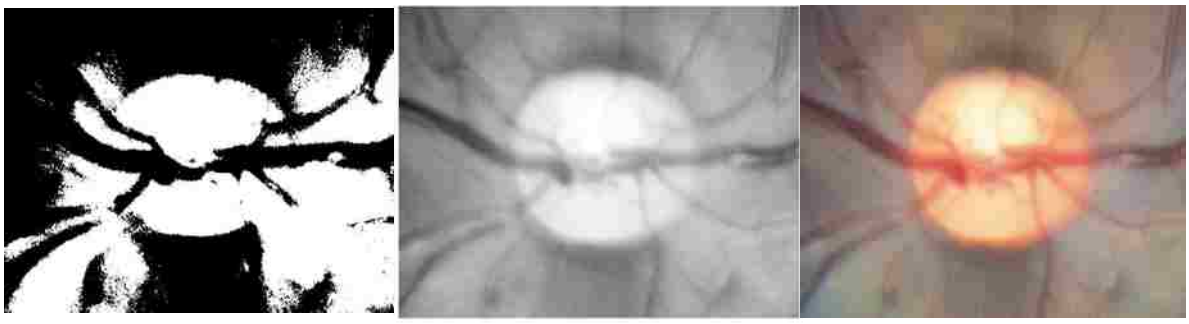
1.2 Brief Description of the Contribution

Two contributions are made to the biomedical image registration and fusion area in terms of novelty, efficiency, and accuracy. The first contribution is that this study has developed a new automated control point detection algorithm for multi-sensor retinal images. The new method is feature-based which employs retina vasculature and bifurcation features. The algorithm extracts retinal vasculature using Canny Edge Detector and identifies control points using adaptive exploratory algorithm. By locating control points at the global direction change pixels, local direction changes are efficiently avoided. Hence, this method made the number of control points as close to the minimum requirement of 2D affine transformation as possible.

The second contribution is the heuristic optimization algorithm for generating the optimal fusion image. In order to maximize Mutual-Pixel-Count (MPC) of the comparing regions, the iteration algorithm adjusts the initially selected control points at the sub-pixel level with the step size less than 1 pixel. A refinement of the parameter set is obtained at the end of each loop, and finally an optimal fused image is generated at the end of the iteration. To our knowledge, it is the first time that Mutual-Pixel-Count concept has been introduced into biomedical image fusion area. A global maxima equivalent fusion result is achieved by calculating local maxima with an efficient computation cost. The unique method presented in this dissertation is a promising step towards useful clinical tools for retinopathy diagnosis, and thus form a good foundation for further development.

1.3 Image Processing Background

There are three basic image types that image processing deals with, i.e. binary image, grayscale image, and color image.



Binary Image (1 bit) Grayscale image (8 bits) Color image (24 bits)

Figure 1.1: Three basic image types

1. **Binary Image**: each pixel is stored as a single bit with the value of 0 or 1. 0 represents black while 1 represents white. No other shades of gray colors exist.

2. **Gray Scale Image**: each pixel is stored as an 8-bit byte with the value ranging from 0 to 255. 0 represents black while 255 represents white. Other various shades of gray exist in the image with the value ranged from 1 to 254.

3. **Color Image**: each pixel is composed of three colors, i.e. red, green and blue (RGB value) in a certain proportions. Each color is stored as 8 bits. Therefore, each pixel is stored as 24 bits. A color image can be converted into a gray scale image using 0.2990R + 0.5870G + 0.1140B formula. These three weight parameters are the default values that have been widely used in today's image processing software, such as Adobe Photoshop® and Macromedia Fireworks®. In

order to achieve a particular purpose when converting the color image into grayscale, one can adjust those three parameters till a satisfactory grayscale image is obtained.

1.4 Retina and Retinopathy

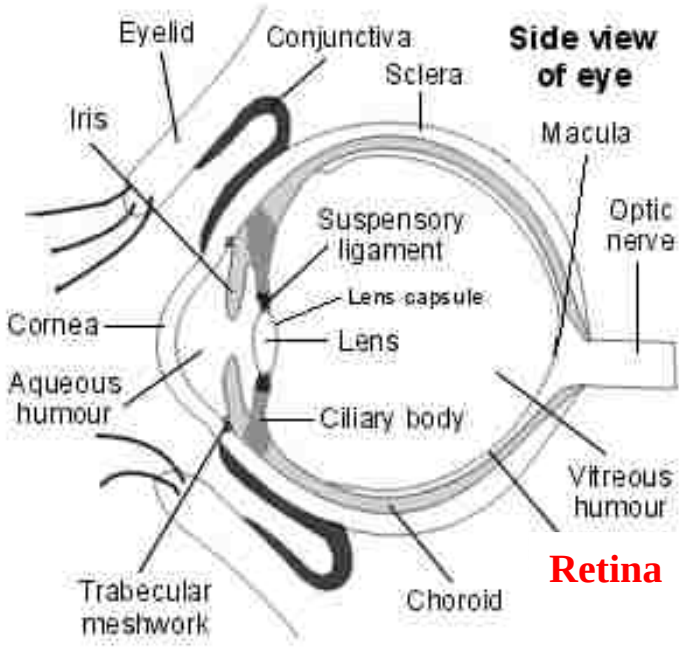


Figure 1.2: Eye Side View [67]

At the back of the eye, there is a nerve layer (Figure 1.2) that can sense the light passing through the lens, and send it to the optic nerve, and then onto the brain. This nerve layer is called *retina*. It is made up from special cells called rods and cones. There are many blood veins and vessels next to the retina. Those vessels take oxygen and nutrients to the cells of the retina. *Retinopathy* covers various diseases and disorders of the retina. Retinopathy is usually due to damage to the blood vessels next to the retina. Diabetes is a major cause to Retinopathy. Some other reasons such as very high blood pressure, age-related macular degeneration, and ocular inflammatory disease can also cause Retinopathy.

It is critical to treat eye diseases or injuries that affect the retina as early as possible. Otherwise, it is possible to lose some vision (blurred vision), or even get completely blindness. For example, Glaucoma, one of the retinal diseases, is the most common cause of blindness [63]. Early detection can allow timely treatment to prevent further vision loss, and to prolong effective years of usable vision. The novel automatic retinal image registration and fusion method presented in this dissertation provides a convenient way for early detection of the retina abnormalities.

1.5 Experimental Materials, Subjects, and Methods

1.5.1 Two Ophthalmologic Modalities

In practical clinical applications, the comparison of angiogram grayscale images with fundus true color images is often required in order to identify dynamic aspects of the circulation and evaluate various retinal vascular disorders. Figure 1.3 and Figure 1.4 are two ophthalmologic image modalities. They are hyperspectral fundus image and Intravenous Fluorescence Angiogram (IVFA) image, respectively. IVFA images are used to detect leaks and other artery abnormalities, and fundus images are useful to detect exudates and hemorrhages. By presenting the combined vessels captured from different sensors to ophthalmologists, blood vessels are displayed at one single view. Detecting minor disorders and damages of those blood vessels, such as ruptured, blocked or new blood vessels become relatively easier. Such fused imagery would allow the clinicians and ophthalmologists, in a convenient way, to monitor oxygen saturation following therapies at the optic nerve head for glaucomatous damages [63], proliferative diabetic retinopathy (DR), photodynamic therapy for age-related macular degeneration (AMD) (Figure 1.7), as well as the viability of the retina before and after laser treatment [37].

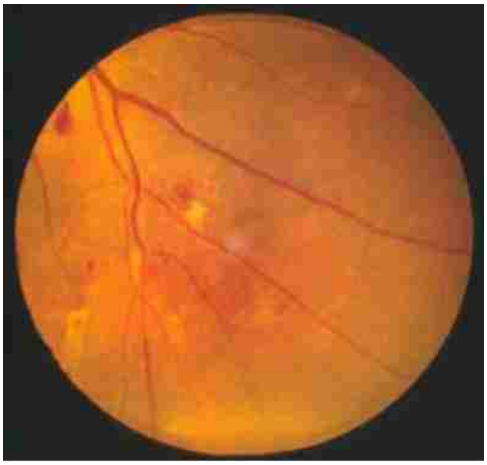


Figure 1.3: Hyperspectral fundus image
Spots: ruptured blood vessels.
Red spots: new blood.
Yellow spots: old blood.

Figure 1.4: Intravenous Fluorescence Angiogram (IVFA) Video Image.
Bright Areas: ruptured or blocked blood vessels.

1.5.2 Acquisition of Retinal Images

Retinal images presented in this dissertation were taken by a modified Topcon TRC-50EX fundus camera (Figure 1.5). The modification enables monochromatic fundus illumination and records over 400-720 nm range of wavelengths. The CCD camera was positioned in the Polaroid mounting part. A LCTF 7nm bandwidth filter was placed in the path of the illumination [37]. An internal flash tungsten halogen lamp provides the source of light. An Apogee U47 16-bit digital camera was positioned inside the fundus camera, which can provide low readout noise images.

1.5.3 Experimental Subjects and Materials

The subjects of the retinal images were Cynomolgus monkeys (Figure 1.6) of 4 to 4.5 years of age and 2.5 to 3 kg body weight with normal eyes [36]. The use of animals for taking retinal images was approved by LSU Health Sciences Center Institutional Animal Care and Use Committee [37]. The monkeys were housed in an air conditioned room with normal temperature and humidity with a 12 hour light-dark diurnal cycle.

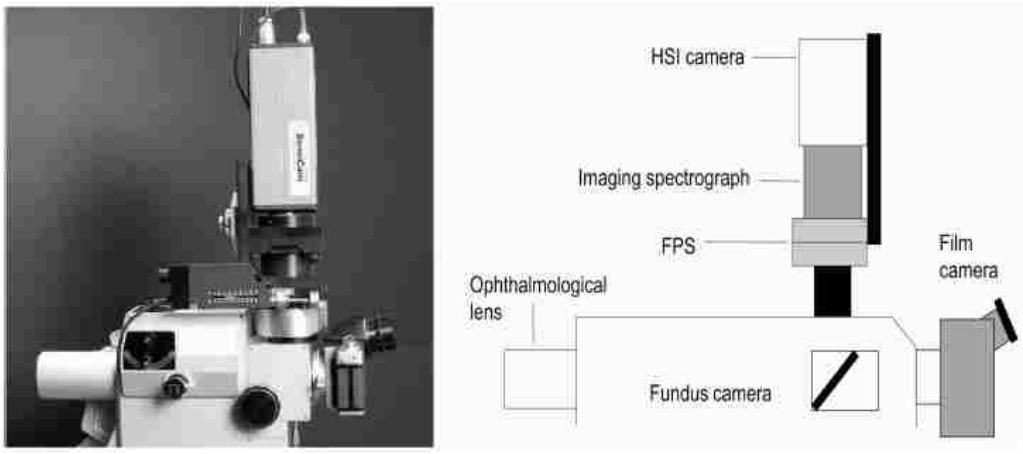


Figure 1.5: Hyperspectral imaging system in relation to the fundus camera [30]

Figure 1.6: Cynomolgus monkey [58]

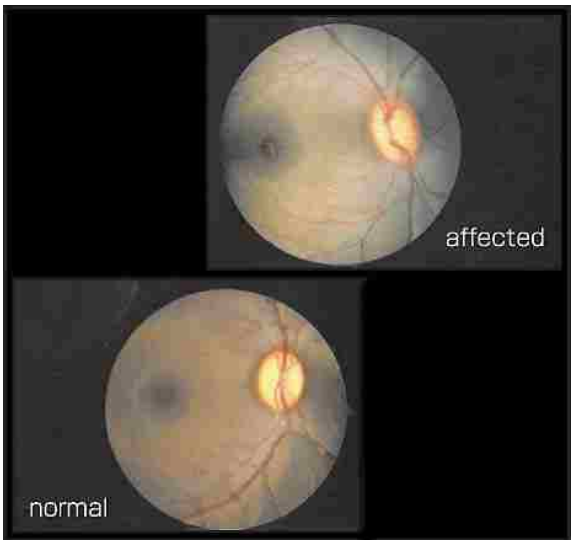


Figure 1.7: Monkey eye's macula degeneration (affected eye (upper) v.s. normal eye (below)) [59]

1.5.4 Experimental Methods

The experimental monkeys were anesthetized with intramuscular ketamine (7-10 mg/kg), xylazine(0.6-1 mg/kg), and intravenous pentobarbital (25-30 mg/kg). Administration of the

anesthetics was repeated alternately every 30 minutes as required to maintain the animals in deep, stage IV anesthesia [36].

In order to enhance the blood vessel appearance in angiograms, a sodium fluorescein dye was injected into the monkey's arm. Angiograms were acquired after 1-2 minutes after the intravenous injection. When the dye circulates through the retinal arteries, capillaries, and veins, it is progressively eliminated from the vasculature, while staining in the optic nerve head and lesions [4].

1.5.5 Contribution of Animal Models to Human Diseases

Establishing animal models is an essential prerequisite of the research and development of new therapeutic interventions on human diseases. Monkey species provide appropriate preclinical models that can closely reflect human's physical and physiological characteristics because of their close phylogenetic relationship to humans [59]. For example, human-similar macular degeneration happens to Cynomolgus monkeys (Figure 1.7).

1.6 Image Registration

1.6.1 Image Registration Definition

Prior to fusion, images need to be registered into the same reference coordinates so that the objects can be aligned pixel by pixel. For the multi-modality registration, aligning all of the data from different sensors to the same reference coordinate is obligatory because sensors provide different local reference frames. Registration is not only for integration purpose, it also corrects the inconsistence data in the images.

Image registration is a mapping function T between the coordinates of two images $I_{(1)}$ and $I_{(2)}$ such that image $I_{(1)}$ and transformed image $I_{(2)}(T)$ have similar structures/features at the same position [19]. Such mapping function is able to compensate for the differences in translation,

scaling, and rotation. The alignment between two images refers to the correspondence of pixel coordinates. The objective of registration is to achieve the image geometrical alignment during a certain time interval. For most medical image registration applications, a transformation T is a mapping function from 2D space to 2D space, from 3D space to 3D space or from 3D to 2D space.

Suppose $I_{(1)}$ is the reference image and $I_{(2)}$ is the input image to be registered to $I_{(1)}$, then the mapping function is:

$$I_{(1)}(x, y) = I_{(2)} \{ T(u, v) \}$$

where T is a 2D transformation function mapping of coordinates (u, v) to (x, y);

or,

$$I_{(1)}(x, y, z) = I_{(2)} \{ T(u, v, w) \}$$

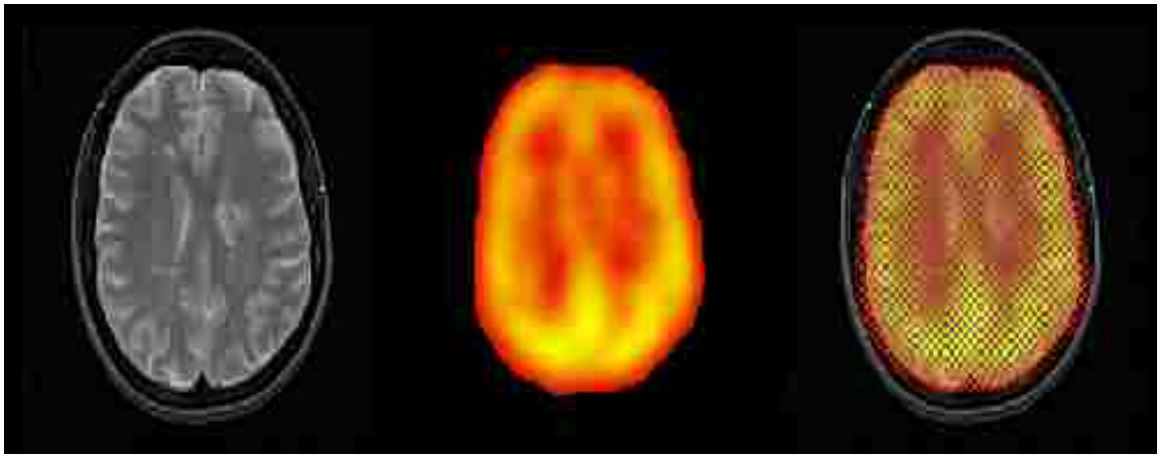
where T is a 3D transformation function mapping of coordinates (u, v, w) to (x, y, z).

1.6.2 Important Factors in Registration

There are a few important factors that most registration methods have to concern. They are image dimension, registration transformation models, intra-modality, and multi-modality models.

1.6.2.1 Image Dimension

2D methods are mainly applied to two-dimensional (x, y) image data such as digital angiography. Such method is also applied to registration of projection images or tomography slices of data which is naturally three-dimensional, such as SPECT, CT and MRI [21] [45] (Figure 1.8). Assume that the image slices to be registered lie on the same plane and that the subject has no significant motion during the involved modalities' work. 3D methods consider three dimensions (x, y, z) image data. The input volume can be constructed from a series of 2D slices.



MRI image **SPECT image** **Fused image**

Figure 1.8: Example of a Brain Image Registration of MRI with SPECT [24]

1.6.2.2 Registration Transformation Models

Transformation is a mapping function *T* between the coordinates of two images such that one original image and the transformed image have similar structures/features at the same position. Transformation can be local or global. Local transformation deals with subsets of the original image, according to the extracted features at the local regions. Local minima/maxima will be generated. Global transformation considers the entire image and uses a single uniform transformation for all regions, and thus can achieve global minima/maxima.

1.6.2.3 Intra Modality V.S. Multi-Modality

Registration dealing with images obtained from same modality at different time or location is called *intra-modality*. Generally, these intra-modality images are characterized by a high degree of similarity [21], i.e. certain level of pixel intensity values in one image must correspond to the same level of intensity values in the other image. Medical intra-modality image registration is widely used in serial scans of the same patient for staging a disease, a baseline and follow-up

exams to monitor the patient's condition [21]. *Multi-modality* registration covers images obtained from different modalities at same or different time, such as CT or PET modalities (Figure 1.9). The similarity assumption in intra-modality image registration is not required in multi-modality registration. Even though different multi-sensor data fusion methods have their own procedure, features, and models, they all share the basic data processing flow patterns – first place the measurements from different sensors into a common coordinate system, then extract feature vectors from the measurements, and finally combine the feature vectors to achieve the detection, estimation, and classification at a certain accuracy level [47]. One of the most desirable aspects of multi-modality image acquisition is that the sensors are collecting different types of information from the same scene, which makes the problem of multi-modality registration inevitably more complicated than intra-modality.

1.7 Image Fusion

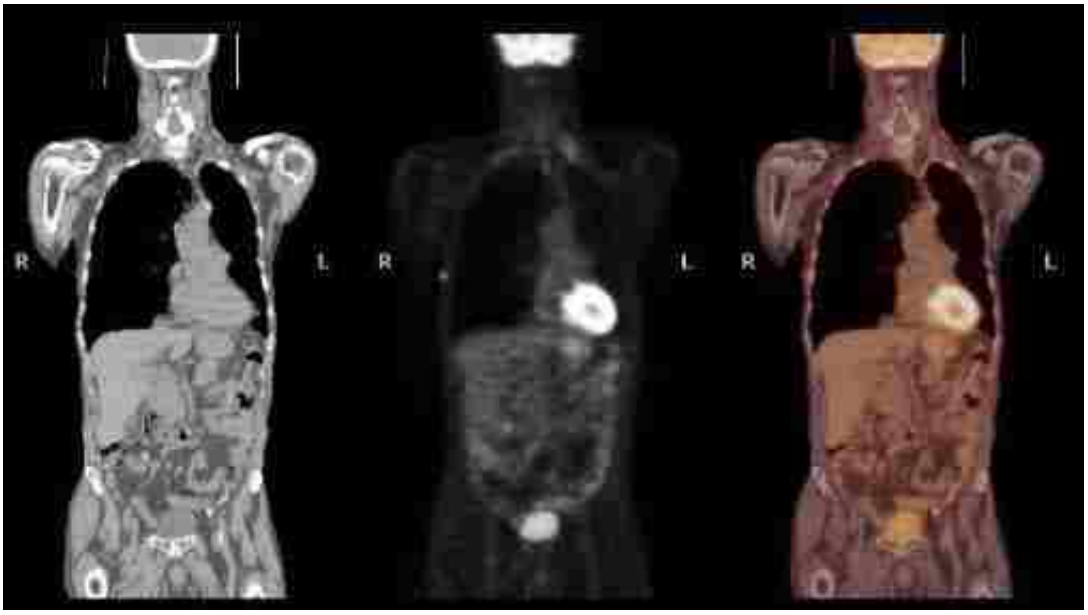
1.7.1 Definition

Image fusion produces a single image from a set of input images. The fused image should have more complete information which is more useful for human or machine perception [20]. Fusing multi-modality data can improve reliability by using mutual information and improve capability by providing complementary information. For example, optical images from Land-sat provide information on chemical composition, vegetation, and biological properties of the surface; while radar images from Sea-sat provide data on surface roughness and electrical properties [31].

1.7.2 Objectives of Image Fusion Schemes

The objectives of image fusion schemes are extracting all of the useful information from the source images without introducing artifacts or inconsistencies that will distract human observers

or the following process. Such objectives could be achieved through creating combined images that are more suitable for human perception and computerized image processing such as augmentation, segmentation, feature extraction, and object recognition [45]. The fused data should provide more complete information than the separate data set does and increase reliability, accuracy, and robustness to imperfection [20]. The accuracy of the fusion will directly affect further analysis of the combined data. Multi-modality image fusion can be described as an organized scheme for the use of different information from multiple sensors to assist in the overall understanding of a phenomenon, events measurement, and decision combination [29]. Medical fusion techniques are important for the purpose of enhancing information, and thus to simplify the comprehension of the image datasets in the clinical diagnosis process [45].



Anatomical (CT) **Function (PET)** **Fused Image**

Figure 1.9: Fusion of the body anatomical (CT) with the function (PET) images [25]

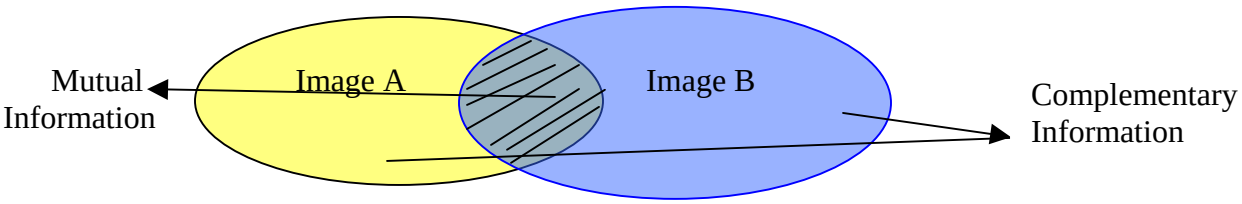


Figure 1.10: Mutual and Complementary Information

1.8 Application Areas of Image Fusion [20]

1.8.1 Intelligent Robots

- Motion control based on feedback from the environment from visual, tactile, force/torque, and other types of sensors
- Stereo camera fusion
- Intelligent viewing control
- Automatic target recognition and tracking

1.8.2 Medical Image

- Real-time image restoration applications [66]
- Same patient's serial studies with the images acquired at different times for staging of the diseases or of the therapy
- Compensation of patient motion during acquisition of the images
- Image guided surgery
- Biomedical sample characterization [43]
- Intelligent control on biometric identification and verification [44]

1.8.3 Manufacturing

- Electronic circuit and component inspection
- Product surface measurement and inspection

- Non-destructive material inspection
- Manufacture process monitoring
- Complex machine/device diagnostics
- Intelligent robots on assembly lines

1.8.4. Military and Law Enforcement

- Detection, tracking, identification of ocean (air, ground) target/event
- Concealed weapon detection
- Battle-field monitoring
- Night pilot guidance
- Biometric identification through retina patterns [60]

1.8.5. Remote Sensing

- Using various parts of the electro-magnetic spectrum
- Error resilient video trans-coding arising from video sensor networks or wireless video communication [46].
- Fusion techniques are classified into photographic method and numerical method
- Near-infrared and infrared imaging [42]

1.9 Organization of This Dissertation

The image registration and fusion methods presented in this dissertation may be applied, but not limited to eye, body, or brain images. The rest of the dissertation is organized as the following:

Chapter 2 provides a literature review and classification of the existing methods for image registration and fusion, as well as the commonly used mathematical models for coordinate transformation.

Chapter 3 introduces the 2D area-based intra-modality retinal image registration method, which estimates the translation and the rotation of the input slices. Fourier transformation is employed to estimate the rotation angle and the translation. This method is used for alignment of slightly moving objects (subject's motion) where translation and rotation are involved.

Chapter 4 covers the novel automated feature-based control point detection method of multi-modality retinal images. The new method is based on extracting retina vasculature features using Canny Edge Detector. Control points are obtained at the global direction change pixels using adaptive exploratory algorithm. Shape similarity criteria are used to match the control points. The comparison with the Centerline Control Point Selection Algorithm shows the advantage of the new Exploratory Adaptive Control Point Detection Algorithm.

Chapter 5 describes the automated heuristic optimization algorithm as well as the experimental implementation results. Mutual-Pixel-Count (MPC) is utilized as the criteria for fusion accuracy measurement. Iteration stops either when the MPC reaches the maximum value or the maximum allowable loop count is reached. The fusion image is generated based on the current control point coordinates when the iteration stops. The evaluation study of the Root-Mean-Square-Error Minimization (RMSE) objective function proves the accuracy of the Mutual-Pixel-Count Maximization objective function. The Genetic Algorithm is compared with the heuristic optimization algorithm and the results shows that the heuristic algorithm has the advantage in terms of efficiency and accuracy. This chapter has also compared the new method with a few other existing data fusion approaches in terms of accuracy, efficiency, and automation. Finally, the challenges when applying the proposed algorithm to human retinal images have been discussed.

Chapter 6 is the conclusion about this dissertation and the future research direction toward the potential clinical application. The Appendix has the list of the tunable threshold parameters and their associated values specified in the program.

Chapter 2

Literature Review

2.1 Image Registration and Fusion Classification

A great number of different methods of biomedical image registration and fusion have been proposed or reported according to Dialogweb.com. These methods include feature extraction, segmentation, classification, pattern recognition, statistical measurements, texture analysis, and a wide variety of other approaches [45]. In order to achieve the satisfying results, scientists and medical research technicians have developed quite a few approaches, such as point-based method, deformation models [8], and voxel-based approach for 3D registration [9]. In [9], optimization is achieved by iteratively optimizing the similarity measurement calculated using voxel (3D) intensity values of both reference and to-be-registered images. Many methods are still considered barred from meaningful clinical application by the fact that they are as yet improperly validated [17].

Unfortunately, there is no method proposed so far can guarantee 100% registration or fusion success rate. In addition, all of them are limited to specific type of modality and organ images. There is no uniform image registration algorithm that can be applied for all different cases. The major reason is because of the difficulty in combining and aligning multi-modality images due to the non-uniform scale factors, different viewing angels, and different acquisition planes. These factors could cause vast content changes and non-uniform intensity distributions. As a consequence, the registration variables have to be corrected in many situations in order to achieve a perfect match.

2.1.1 Registration Classification by Level of Automation

Image registration process can be divided into three categories depending on how much the computer is involved. They are (1) expert manual registration, (2) semi-automated registration, and (3) automated registration.

2.1.1.1 Expert Manual Registration

Expert manual registration is a traditional approach that requires human experts' full interaction to match the image visually. In clinical practice, the manual positioning and alignment by physicians is still the common way that has been used for image fusion from single modality to multi-modality. Physicians have to determine the variables when doing registration without using computer-aided programs.

In the multi-modality scenario, the expert identifies and localizes structures and features in one modality, and then relates them to the corresponding areas on the other modality. The expert has to manually adjust variables of the images until the optimal registration result is accomplished. Although the fusion result is always roughly right as long as the images do not have overwhelming differences, this manual method is very subjective and time consuming. Furthermore, its reproducibility is very low.

2.1.1.2 Semi-automated Registration

In the semi-automated approaches, computer programs determine the mathematical transformation model and perform the transformation. Computers can also fulfill some pre-processing work, such as image shifting, re-orienting and zooming/scaling [21], noise removal, enhancement of contours, and outlining of certain anatomical. Programs have to depend on continuous human's operation and control during the registration process. Human is involved in adjusting the input parameters all the time during the registration process. Pelizzari and his

colleagues [22] have developed widespread three-dimensional semi-automated registration of CT, PET, and MR brain images. Another semi-automated approach is the iterative surface fitting algorithm, which requires human operators' guidance to avoid local minima. This method has been undergoing a few modifications and customizations for improvements in speed and accuracy [23].

2.1.1.3 Automated Registration

Automated registration does not require human's interaction at the control point selection, feature extraction, and the optimization procedures. A certain number of tunable thresholds are allowed. Therefore, it makes the process objective, reproducible and fast. However, fully automated method is not as reliable and robust as the one with human's interaction. It might generate erroneous fusion results in extreme cases. Quality control software might have to be involved to ensure the success rate of the registration. The algorithm this study has developed belongs to the automated registration category.

2.1.2 Fusion Classification by Level of Representation

Data fusion area is very broad and it can be grouped into four categories by levels of representation [47] [31]. They are signal-level, pixel-level, symbol-level, and feature-level fusions. This is only a rough classification, and in no way to capture every aspect of numerous fusion applications.

2.1.2.1 Signal-level Fusion

Signal-level fusion refers to the combination of sensors with the objective of producing a single signal that contains the same form as the original signals, but of better quality. A typical example is the generation of a high resolution image from a set of relatively low resolution input images. A common technique for signal-level fusion is Kalman filtering fusion [48] [49] [51].

First order and second-order Kalman filtering are implemented for the specific target tracking and scalable sensor fusion purpose in the sensor networks.

2.1.2.2 Pixel-level Fusion

Pixel-level fusion is through a method to enhance the useful information / features, so that the performance of the future image processing can be efficiently improved. The fused data can be generated either via pixel-by-pixel correlation or via the fusion of the local neighborhood. The useful features are combined by performing pixel-by-pixel superimpose from one image to the other. Many biomedical multi-sensor fusion methods can be categorized into this class. The new multi-modality fusion method presented in this dissertation belongs to the pixel-level fusion category as well.

2.1.2.3 Symbol-level Fusion

Symbol-level fusion is a way allowing the information from multiple sensors to be effectively used at the higher level of abstraction. It is often employed in the applications where sensors are deployed at different regions of environments. Simultaneous detection of lane and pavement boundaries is the widely used application of symbol-level fusion. The dominant technique of symbol-level fusion is Bayesian estimation [53].

2.1.2.4 Feature-level Fusion

Feature-level fusion is performed via the detection or extraction of the useful features with a certain level of confidence. It is suitable when different type of sensors measure different phenomena or features. The features in one modality are matched into the corresponding features of another modality. Generally, feature-level fusion does not require as much strict registration as the signal-level or the pixel-level fusion. Gauss-Markov random fields are a commonly used method to minimize total energy consumption during the feature-level data fusion [50] [52].

2.2 Registration Transformation Models

Translation model, affine model, and quadratic model are the three most frequently used transformation models. This study calls the image that remains same as “reference” and the one that needs to be registered as “input “image. In the case of automatically registering multi-modality retinal images, the goal is to achieve accurate registration with less human interaction. Therefore, the absence of clear reference anatomical vessel regions makes the angiogram image a better candidate to be a reference than the color fundus image.

2.2.1. Translation Model (Similarity)

Degree of freedom (DOF) = 2

$$\begin{pmatrix} U \\ V \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \qquad U(x, y) = x + a \qquad (2.1)$$

$$V(x, y) = y + b \qquad (2.2)$$

In the translation model, (x, y) are the coordinates of the reference image; (U, V) are the corresponding coordinates of the input image. Its advantage is that only two control points are needed in order to calculate the parameters a and b, which is suitable for some cases without sufficient control points. Its drawback is that the translation model gives rises to a significantly larger registration error than the higher order mathematical models. Translation model does not work well for cases of poor image quality and high geometric distortion.

2.2.2 Affine Model

Degree of freedom (DOF) = 6

$$\begin{pmatrix} U \\ V \\ 1 \end{pmatrix} = \begin{pmatrix} a_1 & a_2 & b_1 \\ a_3 & a_4 & b_2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \qquad U(x, y) = a_1x + a_2y + b_1 \qquad (2.3)$$

$$V(x, y) = a_3x + a_4y + b_2 \qquad (2.4)$$

Affine model's advantage lies in that it can measure lost information such as skew, translation, rotation, shearing and scaling that maps finite points to finite points and parallel lines to parallel lines. Its drawback lies in the strict requirement that at least six pairs of control points are needed.

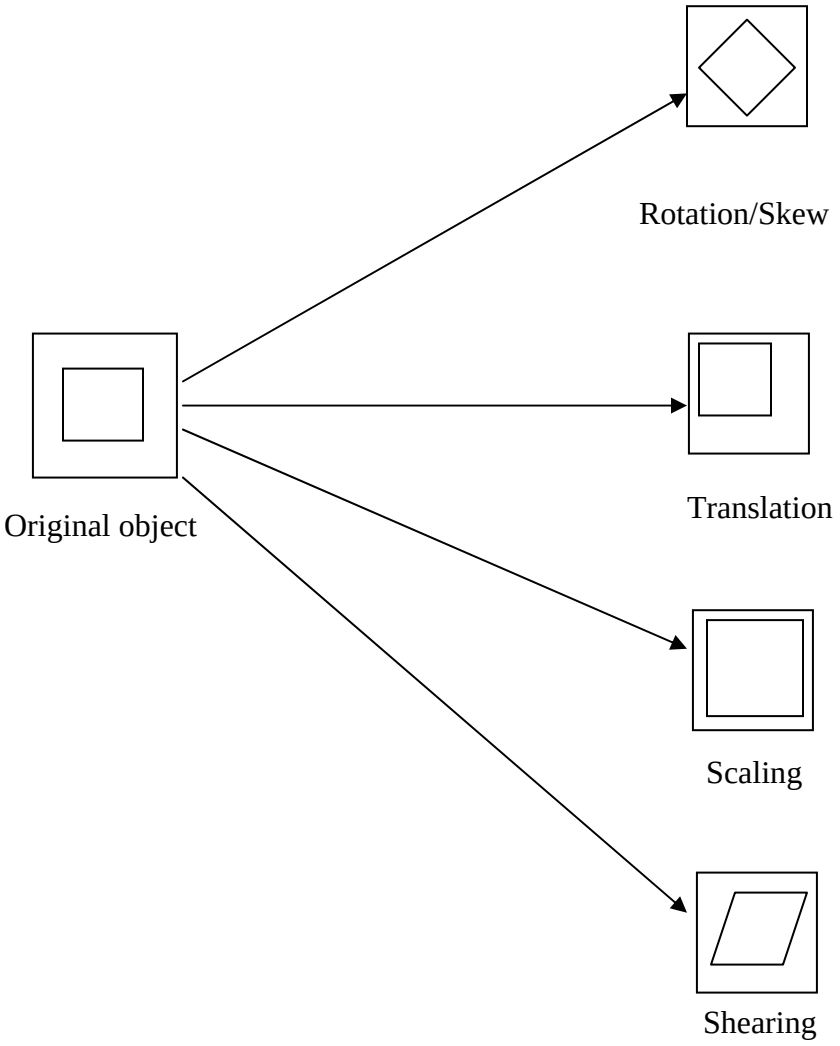


Figure 2.1: Transformation Graphs

Table 2.1: Transformation Descriptions

Transformation	**Description**
Rotation/Skew	Points are rotated by an angle θ.
Translation	A linear shift in the position of the vertical and horizontal coordinates of the image in one plane to another set in the same spatial domain.
Scaling	A transformation of the horizontal and vertical coordinate points characterized by a certain scale factor.
Shearing	A transformation in which all points along a line remain fixed while other points are shifted parallel to the line by a certain distance proportional to their perpendicular distance from the line [18].

2.2.3 Rigid 2D Transformation Model – Special Case of 2D Affine Model

When only transition and rotation are allowed in the input images, one can apply rigid 2D transformation model. Objects remain their relative size and shape.

Transformation: $T = D \times R$

$$D = \begin{pmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{pmatrix} \tag{2.5}$$

$$R_x = \begin{pmatrix} \cos\phi x & \sin\phi x & 0 \\ -\sin\phi x & \cos\phi x & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{2.6}$$

$$R_y = \begin{pmatrix} \cos\phi y & \sin\phi y & 0 \\ 1 & 0 & 0 \\ -\sin\phi y & \cos\phi y & 0 \end{pmatrix} \tag{2.7}$$

where R is the rotation matrix and D is the transition matrix. t_x and t_y are the transition parameters. ϕx *and* ϕy are the two angles on x and y direction respectively. Therefore, the transformation model can be written as:

$$U(x, y) = S\,(x \cos \phi x - y \sin \phi x) + t_x \tag{2.8}$$

$$V(x, y) = S\,(x \sin \phi y - y \cos \phi y) + t_y \tag{2.9}$$

where S is the translation scale.

2.2.4 Quadratic (2nd Order Polynomial) Model

Degree of freedom (DOF) = 12

$$\begin{pmatrix} U \\ V \end{pmatrix} = \begin{pmatrix} a_1 & a_2 & b_1 & c_1 & c_2 & d_1 \\ a_3 & a_4 & b_2 & c_3 & c_4 & d_2 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \\ x^2 \\ y^2 \\ xy \end{pmatrix} \tag{2.10}$$

$$U(x, y) = a_1x + a_2y + b_1 + c_1x^2 + c_2y^2 + d_1xy \tag{2.11}$$

$$V(x, y) = a_3x + a_4y + b_2 + c_3x^2 + c_4y^2 + d_2xy \tag{2.12}$$

The 2nd order polynomial model's advantage is that it can greatly reduce the registration error [2]. Its drawback is to locating at least 12 pairs of landmark control points, which can seldom be reached.

According to F. Laliberte and L. Gagnon [13]'s experiment, translation model achieved 61% fusion success rate; affine model achieved 65% while quadratic model achieved 66%. There is a tradeoff between the registration error and the computation complexity.

In the clinic practical implementations, the higher order models are rarely used due to the computational complexity and fusion distortion, regardless the relatively lower registration error.

The quantitative results of retinal images registration from different modalities and different times showed an advantageous performance of the affine transformation model, in terms of transformation selection and fusion of the registered data [15]. Based on the previous research and experimental result, this study has chosen 2D affine transformation as the registration model.

2.3 Area-Based Approach

2.3.1 Area-Based Registration

Area-based method is usually used for intra-modality image registration. A small window of points in the input image is statistically compared with windows of the same size in the reference image. Normalized cross-correlation is a commonly used measurement of the alignment.

2.3.2 Area-Based Fusion

Mutual Information (MI) [1] is the core optimization measurement in the area-based fusion. MI is originally a basic concept derived from information theory. It measures the statistical dependence between two random variables. In the image fusion area, Mutual Information describes the amount of information that one image contains about the other. In this dissertation, the Mutual Information concept is simplified into Mutual-Pixel-Count (MPC) which measures the overlapping pixels of the retinal vasculature.

If the images are perfectly geometrically aligned, MPC represents the maximal pixel correspondence. The searching area using a certain transformation model is usually large. Therefore, the great geometric distortion between the reference and input images will make the searching and computing very complicated and time-consuming. In other words, if there is no uniform intensity/contrast distribution in the original image acquisition or in the large homogeneous spatial domain, the performance of area-based techniques will be dramatically degraded.

2.3.2.1 Mutual Information Criterion [1] [7] [6]

Given two random variables *R* and *I*, $P_R(r)$ and $P_I(i)$ are their marginal probability distributions respectively. $P_{RI}(r, i)$ is the joint probability distribution. *MI(R, I)* is the measurement of the dependence of *R* and *I* by the distance between $P_{RI}(r, i)$ and $P_R(r) \times P_I(i)$. *MI(R, I)* is related to entropy by the following equation:

$$MI(R, I) = H(R) + H(I) - H(R, I) \tag{2.13}$$

Where *H(R) is* the entropy of *R*, a measure of the uncertainty or dispersion of the probabilities of event R; *H(I)*is the entropy of *I,* a measure of the uncertainty or dispersion of the probabilities of event I; and *H(R, I)* is the joint entropy of *R* and *I*, a measure of the reduction in the uncertainty of *R* by the knowledge of the other random variable *I*.

The global entropy [2] can be defined as a function of the state probability by the following equation.

$$H(R) = -\sum_{j=1}^{N} P_j(r) \log_2 P_j(r) \tag{2.14}$$

where $P_j(r)$ is the probability at the image's state *j*. In this dissertation, the global discrete entropy is used to verify the selection of RGB when converting the color image to grayscale image, which is described in details in chapter 4.

Maes and Collignon [1] proposed a registration algorithm based on Mutual Information. When MI is selected as the similarity measure, estimating the joint entropy is iteratively required. Frequent calculation of joint entropy is a time-consuming procedure. The difficulty with the MI algorithm is the possible existence of local maxima. Hence, coarse-to-fine multi resolution algorithm has been developed in [10] with the effort of achieving global maxima. In [11], the pyramid of the images could be obtained using the wavelet transformation. When using

image pyramids, they do not consider the preference for different regions of image pairs. Mutual Information was implemented by Klein and Staring to compare gradient descent, quasi-Newton, nonlinear conjugate gradient, Kiefer-Wolfowitz, simultaneous perturbation, Robbins-Monro, and evolution strategy [28]. Registration accuracy was assessed by computing the overlapping area of the segmented edges. Klein's conclusion is that the Robbins-Monro method is the best choice in most applications.

Although the area-based methods work well for intra-modality in various applications, they are not well-adaptive to multi-modality image registration because the gray/color characteristics of images to be matched are quite different [31].

2.3.3 Summary of Area-Based Approach

Most area-based registration and fusion method (Diagram 2.1) can be summarized in three steps:

Step1. Perform the initial processing of the original image pairs (usually obtained from same type sensors);

Step2. Conduct cross-correlation registration using Fourier Transformation;

Step3. Maximize MI to achieve the optimal fusion result and conduct feature-extraction at the end.

2.4 Feature-Based Method [4]

2.4.1 Feature-Based Registration

Feature-based registration method extracts and matches the common structures (features) from two images. It has been shown to receiving a higher successful rate at the multi-modality registration scenario [31] than the area-based method. Assume that the feature correspondences exist between the reference and input images, and the residual translation and rotation error are

small [31]. In addition, the images to be registered have to be coarsely aligned based on prior information.

Feature-based method requires feature extraction prior to registration. The feature means the salient structures, such as the central line of veins and the vessel bifurcation points in the retinal network. For example, Can and Shen have proposed an algorithm that can automatically trace the feature by using the direct exploratory method [38]. Their proposed algorithm is able to handle images with the poor condition, such as over / under-exposure and low contrast. Another possible solution is using *Matched Filter Technique* [55]. Banumathi and Devi have introduced and compared three approaches, i.e. Gaussian Matched Filter (GMF), Binary Matched Filtering (BMP), and Kirsch Template Matching Filter (KMF). According to their experimental results, GMF and BMF are suitable for grayscale angiogram images, while KMP can be applied to both of grayscale and color images. However, matched filter could be suffered from intensive computation cost and might not be suitable for real-time ophthalmic image processing.

Depends on the transformation model, feature-based method requires a certain number of control points available. The feature-based registration approach largely depends on sufficient and reliable pixel correspondences between image pairs. When there are no adequate control points in the image pairs, the feature-based method will fail.

2.4.2 Feature-Based Fusion

2.4.2.1 Registration Error Minimization

The matching process of feature-based fusion procedure detects the correspondences by minimizing registration error [62]. Zhou [5] proposed *SSD* minimization algorithm by defining *SSD* as the following equation:

$$\text{SSD} = \sum_{i \in OA} \left(I'(F(F(x;\Delta M);M)) - I(X) \right)^2 \tag{2.15}$$

where OA is the overlapping area of I_r and I_i; M is the parameter matrix; F is the transformation function. The transformation parameters are calculated using Newton's iterative optimization approach [16] where the iterative refinement equation is given by the following equation.

$$M_{n+1} = \mathrm{M_n} + (\mathrm{H_n})^{-1} \ \Delta \mathrm{M} \ \mathrm{F_n}. \tag{2.16}$$

where $\nabla M \ \mathrm{F_n}$ is the gradient vector of F in respect to M at $\mathrm{n^{th}}$ iteration. H_n is the Hessian Matrix of F in respect to M at $\mathrm{n^{th}}$ iteration.

$$H_n = \begin{bmatrix} \frac{d^2F}{dM_1^{\ 2}} & \frac{d^2F}{dM_1 dM_2} \cdots \frac{d^2F}{dM_1 dM_m} \\ \frac{d^2F}{dM_2 dM_1} & \frac{d^2F}{dM_2^{\ 2}} \cdots \frac{d^2F}{dM_2 dM_m} \\ \cdots & \cdots \quad \cdots \\ \frac{d^2F}{dM_m dM_1} & \frac{d^2F}{dM_m \ dM_2} \cdots \frac{d^2F}{dM_m^{\ 2}} \end{bmatrix} \tag{2.17}$$

2.4.2.2. Convergence Criteria

Convergence criteria are used to determine when the iteration should stop. The procedure is iteratively implemented until either of the following convergence criteria met [16].

(1). Predefined maximum number of loops has been reached.

(2). Updated parameters $(H_M^{\ n})^{-1} \nabla M \ F_n$ is smaller than ε, i.e. $|Mn_{+1} - M_n| < \varepsilon$. ε is a predefined threshold value.

The convergence criteria proposed by Averbuch and Keller [16] are adopted and modified in this study. The purpose is to make the optimization procedure suitable for the retinal image fusion.

2.4.3 Summary of Feature-Based Approach

Most feature-based registration and fusion method (Diagram 2.2) can be summarized in three steps:

Step1. Build a multi-resolution map for each image;

Step2. Conduct feature/control points selection at each featured area in the map;

Step3. Create a new image based on the fused multi-resolution map.

The final step would affect the fusion result most. Since the multi-resolution map is determined by some artificial selection rule, a "stability" problem arises when the inverse transform takes place when registration error minimization is applied as the fusion criteria [31].

2.5 Combined Approach

Chanwimaluang and Fan [2] proposed a combined method. In their proposal, vascular tree [38], which was the most prominent anatomical structure in the retina, was extracted. Vascular tree mapped the whole area of retina. In most cases, the map did not move. It contained adequate information for the localization of some anchors. There were five steps involved for extracting vascular tree: step1. A match filter was applied to enhance the prominence of vessels; step2. A local entropy-based threshold scheme was used to binarizing the images; step3. A length filter was placed to remove incorrectly binarized pixels; step4. The centerline of vessels were extracted by applying morphological thinning operation; step5. Bifurcation points were identified using a two-step window-based probing process. Registration process was done in three steps. Binary vascular trees were firstly extracted from retinal images using local entropy-based threshold method. 0^{th} order translation was then estimated by maximizing mutual information based on the binary image pair. Landmark points and sampling points were finally

selected and used for higher order transformation estimation. However, the proposed algorithm was not adaptive to large geometric distortion between the image pairs.

The presented new approach belongs to the combined method category because this study uses feature-based method in the control point detection and area-based method in the registration and fusion. In the feature-based part, this study made a contribution by developing an adaptive exploratory algorithm to identify the global direction change pixel at the Canny edges. By locating control points at the global direction change pixel, local direction changes are efficiently avoided. Therefore, this approach made the number of control points as close to the minimum requirement of 2D affine transformation as possible. This study made another contribution of the heuristic algorithm for the optimization area-based fusion. Mutual-Pixel-Count (MPC) objective function is a new and unique concept in the biomedical image fusion area as the fusion accuracy measurement. The fusion result is supposed to be optimal when MPC reaches the maximum value. A global maxima equivalent fusion result is obtained by calculating MPC local maxima with an efficient computation cost.

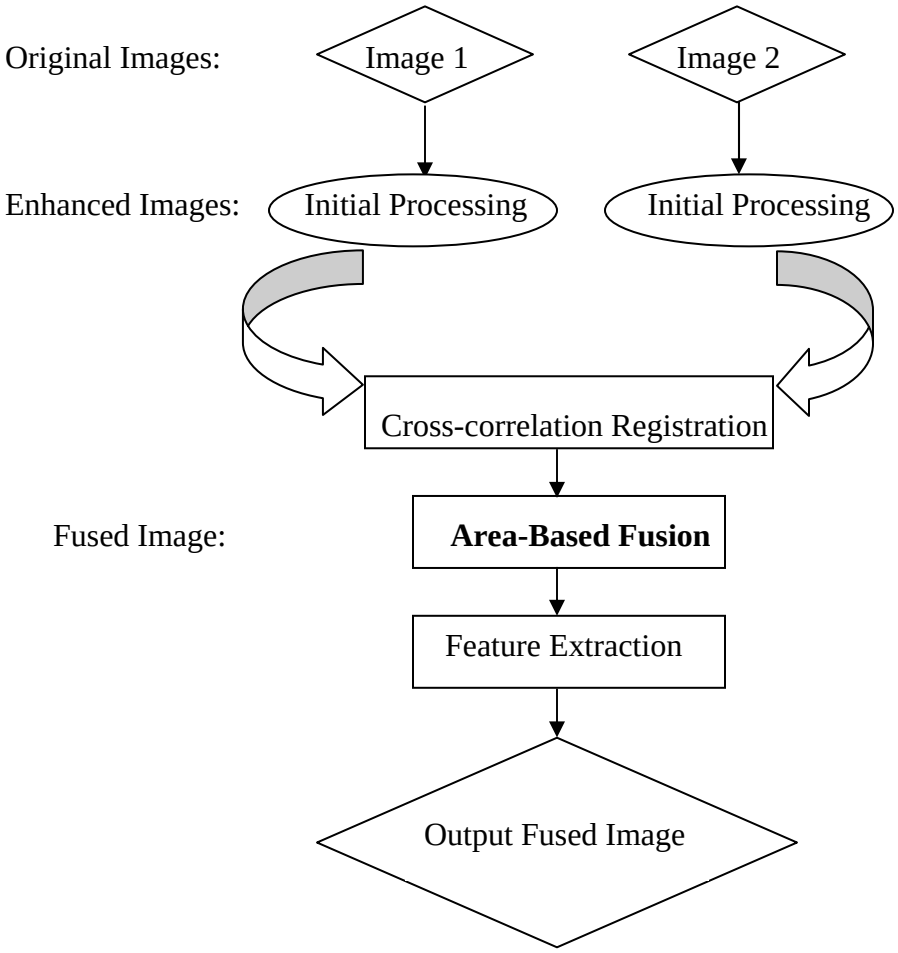


Diagram 2.1: Area-based Image Registration and Fusion Diagram

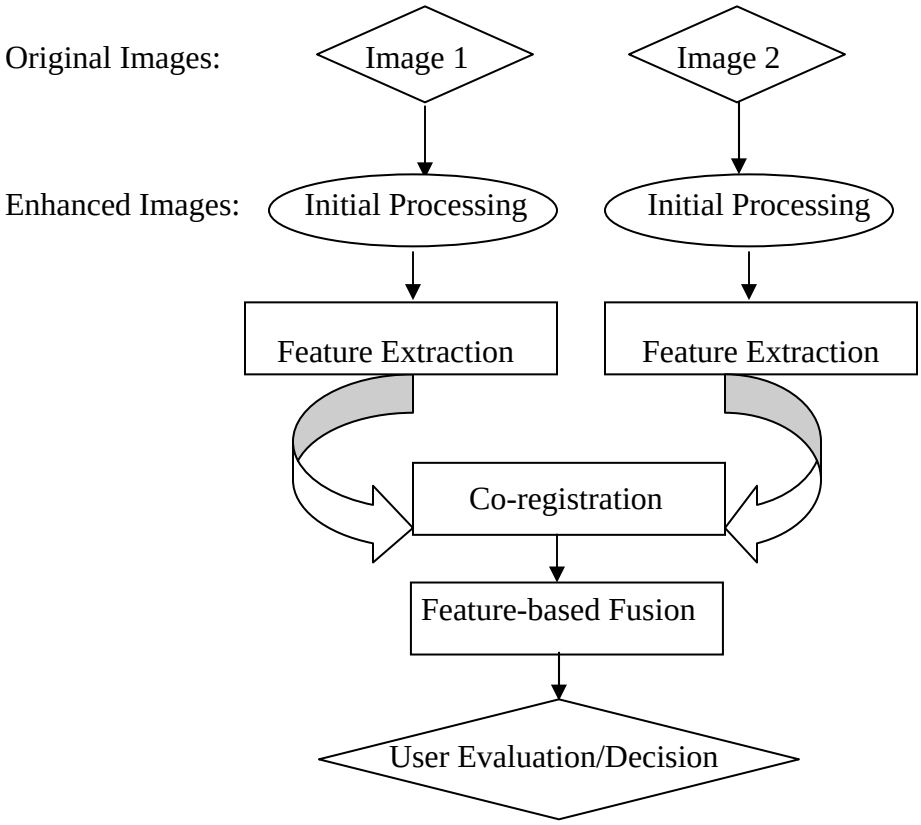


Diagram 2.2: Feature-Based Image Registration and Fusion

Chapter 3

Area-Based Intra-Modality Retinal Image Registration

3.1 Cross-Correlation Registration

With proper registration of the multi-spectral fundus images, clinicians are able to determine oxygen saturation in the primate retinal structures on a reasonable time interval [37]. Area-based two-dimensional cross-correlation image registration method has been implemented at LSU Eye Center, New Orleans, Louisiana to align slightly misaligned retinal images. Registration of region-of-interest (ROI) is the method for alignment of all 2D intra-modality slices in a stack to the first slice using cross-correlation [27]. The entire set of 2D slices is the 3D multi-spectral fundus images with the wavelength as the third dimension. Only translation, rotation, and scaling are allowed from the input images. A small window of points in the input image is statistically compared with windows of the same size in the reference image. The windows surrounding the control points are matched with each other in a hierarchical manner using area cross-correlation. It determines the registration parameters and transforms the entire image. The operator can select a ROI that is doing the motion that he/she wants to correct for in the large images, and thus can lower the computation cost.

The subject's eye motion causes serious problem in retinal oxygen analysis. People can never fully control the motion of eyes. Thus the motion compensation is necessary between the series of image slices. The algorithm assumes that the registration between any consecutive images can be well achieved by calculating rotation and translation. In this approach, every image acquired after the optic nerve head's movement was registered with the one acquired prior

to the movement. The interval of each slice's capturing time by the CCD Camera and transmission to the computer disk was one second. Different retinal features were reflected at different wavelengths. Seven monochromatic images were taken at 522, 542, 548, 560, 577, and 586 nm wavelengths [37]. 522 nm image was selected as the reference. Other six images were registered /aligned to this reference. Since those images were taken sequentially within a few seconds, they typically contained the same scene and displayed a similar feature. The classical template algorithm employed a template from the reference image and various moving windows in the input images to determine *x* and *y* translations and the rotation angle [37]. The algorithm is described as follows:

Step1. Read the reference image first and then each of the input images (Figure 3.4). The input image and the reference image differ from each other by a translation and a rotation. The optic nerve head is apparently moving toward east during the examination (Figure 3.1 and 3.2).

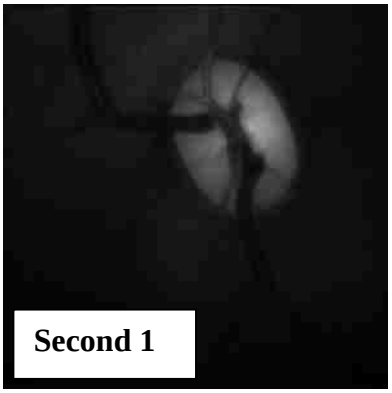


Figure 3.1: Reference image slice at the wavelength = 522

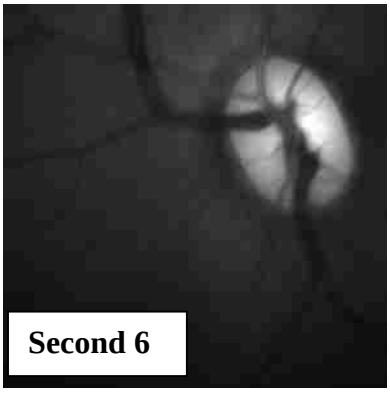


Figure 3.2: Input image slice at the wavelength = 586

Step2. Eliminate the translation and only allow rotation remains. Rotation of an image $I(x, y)$ in a spatial domain corresponds to a rotation of its Fourier transform $F(u, v)$ by the same angle [21].

Step3. Apply Fourier Transform based approach to the 2D cross-correlation. Find the maximum of the cross-correlation yielding the rotational difference. Suppose $I_{(1)}(x, y)$ is a spatial domain and $F_{(1)}(u, v)$ is the corresponding Fourier transform, it can get:

$$F_{(1)}(u, v) = \iint I_{(1)}(x, y) e^{-2\pi j(ux+vy)} dxdy ; \quad (3.1)$$

It can also get the inverse Fourier transform, which is:

$$I_{(1)}(x, y) = \mathrm{F}_{(1)}^{-1}(\mathrm{u}, \mathrm{v}) = \iint F_{(1)}(u, v) e^{2\pi j(ux+vy)} dudv \quad (3.2)$$

The cross-correlation function is defined as:

$$I_{(1)}(x, y) \otimes I_{(2)}(x, y) = \iint I_{(1)}^{*}(\alpha, \beta) I_{(2)}(x+\alpha, y+\beta) d\alpha d\beta \quad (3.3)$$

Two images' cross-correlation ($\otimes$) is equivalent to multiplication of the Fourier transform of one image with the complex conjugate of the Fourier transform of the other.

$$I_{(1)}(x, y) \otimes I_{(2)}(x, y) \Leftrightarrow F_{(1)}^{*}(u, v) \bullet F_{(2)}(u, v) \quad (3.4)$$

And the inverse transform is

$$I_{(1)}^{*}(x, y) \bullet I_{(2)}(x, y) \Leftrightarrow F_{(1)}(u, v) \otimes F_{(2)}(u, v) \quad (3.5)$$

Step4. Rotate the input image by the calculated angle.

Step5. A classic template in the reference image is placed in the spatial domain [37]. Various moving windows in the input images are manipulated in order to determine the *x* and *y* translation (Figure 3.3). The moving window location *(u, v)* has a normalized correlation (Eq. 3.1.6) with the reference images' *(x, y).* The similarity between the reference template *(X)* and the regions in the input image *(Y)* was determined using this correlation [39]. A normalized correlation was computed as Eq 3.7. Translation, which corresponds to a phase shift in the Fourier transform, is calculated (Figure 3.7 and 3.8) between the reference and the input images.

Then the peak of the correlation and the corresponding translation parameters are determined using the following equation.

$$I(x - x1, y - y1) \Leftrightarrow F(u,v)e^{-2\pi j(ux1+vy1)/N} \tag{3.6}$$

where I is an N × N image.

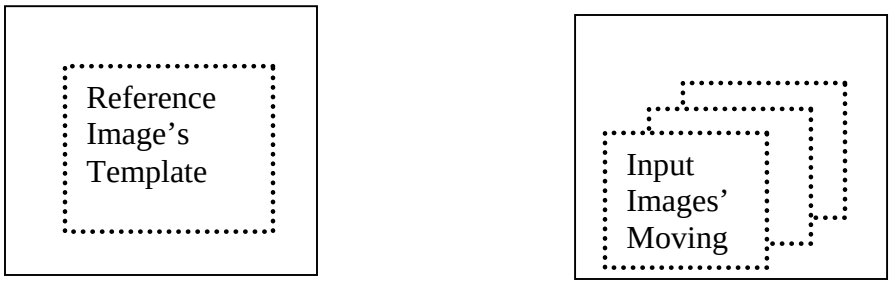


Figure 3.3: Using moving windows to determine *x* and *y* translation

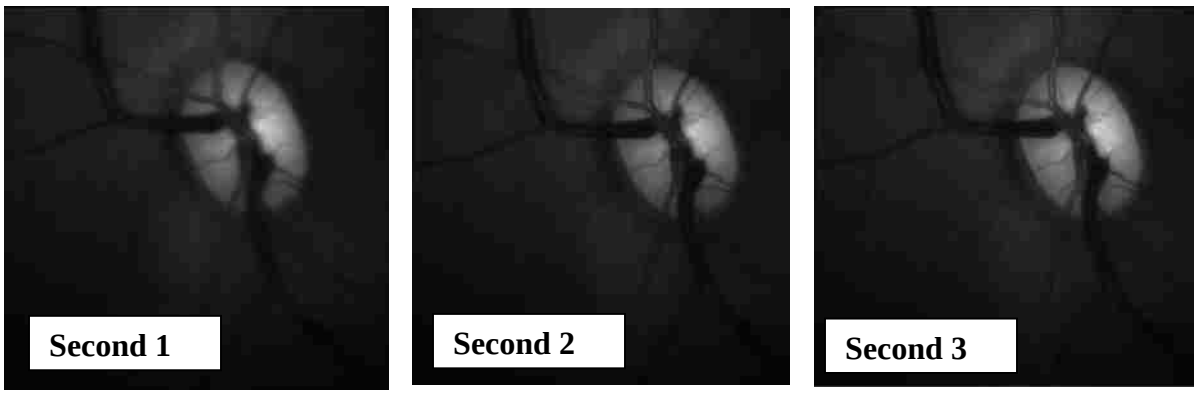
$$Correlation(X,Y) = \frac{\sum_{i=1}^{m}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{m}(X_i - \bar{X})^2 (Y_i - \bar{Y})^2}} \tag{3.7}$$

Step6. Apply 2D cross-correlation transformation on the reference image and the rotated one obtained at step2. Find the maximum of the cross-correlation, which presents the translational difference.

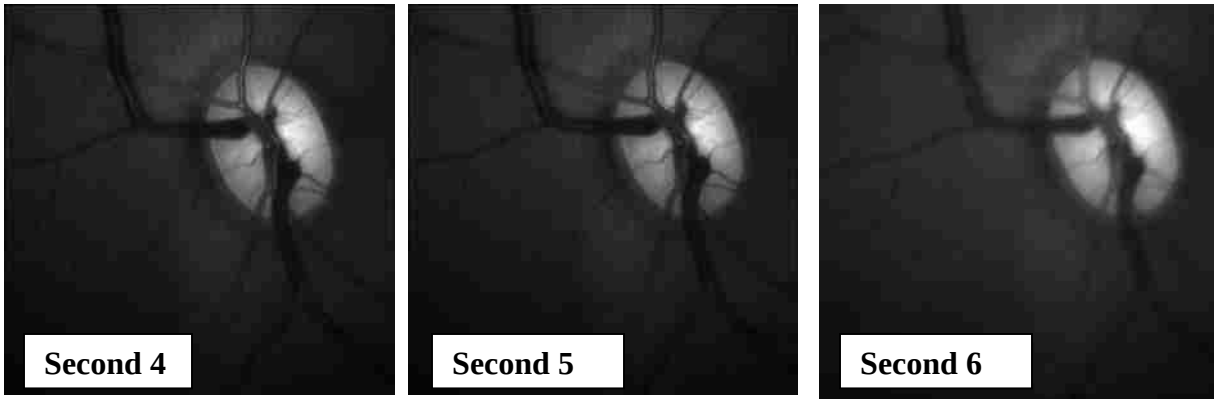
Step7. The input image is now completely registered into the reference image (Figure 3.5).

3.2 Discussion

The Fast Fourier Transform based algorithm usually requires large scene overlaps between the images to be registered. Although the area-based methods work well for intra-modality registration in various cases, they are not well-adaptive to multi-modality image registration. The reason is because the gray-level and RGB characteristics of images to be matched are quite different [31].



Slice 1, 2, 3 – input images during eye motion with the wavelength of 542, 548 and 560



Slice 4, 5, 6 – input images during eye motion with the wavelength of 569, 577 and 586

Figure 3.4: Eye motion causes the optic nerve head's movement

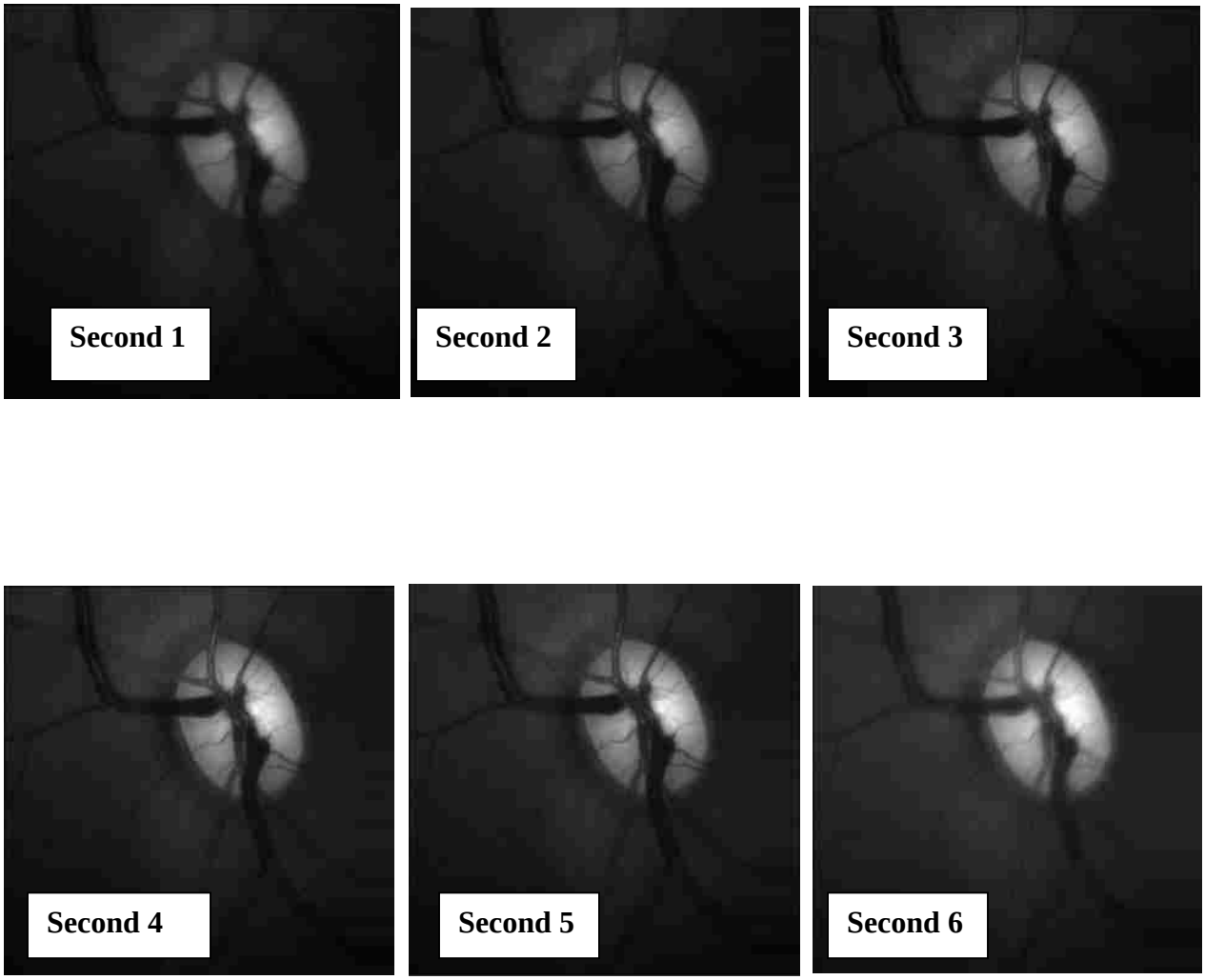


Figure 3.5: All input images were successfully registered.

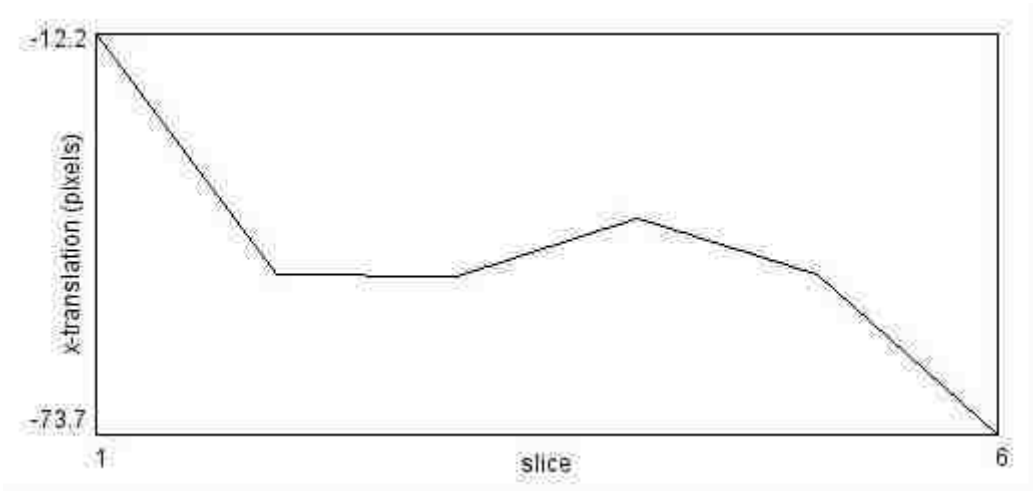


Figure 3.6: X-translation pixels for slice 1 – 6

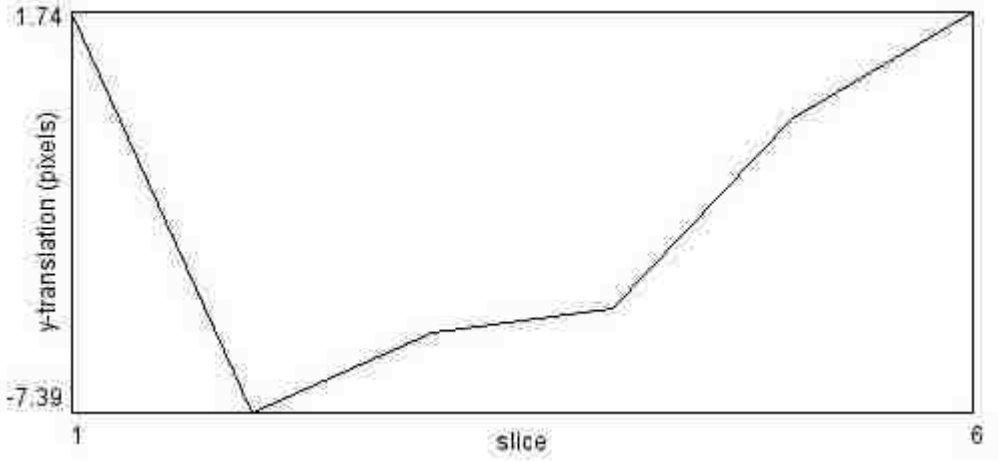


Figure 3.7: Y-translation pixels for slice 1-6

Chapter 4

Feature-Based Control Point Detection of Multi-Modality Retinal Images

4.1 Multi-Modality Image Registration Criteria

Multi-modality images taken by different sensors are difficult to register due to variations in color and intensity across image frames. A good multi-modality image registration algorithm should meet the following three criteria:

(1) A strong ability to process images with different gray-level and RGB color combination characteristics;

(2) Reliability and robustness;

(3) Efficient computational cost.

Compared to area-based registration, the feature-based approach is more appropriate to multi-modality image registration as discussed in Chapter 2. The registration method uses vessel boundaries and strong edges as the matching primitives. The strong edges are those with strong contrast from the background color. Weak edges will not be extracted because of their similar color intensities with the background of the retinal network. Retinal vasculature is extracted using Canny Edge Detector and the control points are identified using adaptive exploratory algorithm. The algorithm extracts the retinal vasculature using Canny Edge Detector and identifies control points at the global direction change pixel. Therefore, local direction changes are efficiently avoided. The presented approach made the number of control points as close to the minimum requirement of 2D affine transformation as possible. Shape similarity criteria are used to match the control points.

4.2 Problem Definition - Control Point Detection

4.2.1 Knowns

The following information is available for the adaptive exploratory control point detection algorithm.

(1) 2D images' resolutions and sizes (width: m pixels; height: n pixels).

(2) Corresponding binary m × n map with 1 standing for white and 0 standing for black.

(3) Edge detection algorithm is available for retinal vessel detection with 1 standing for non-edge pixels and 0 standing for edge pixels.

4.2.2 Assumptions

It is assumed that both of the reference and the input images do not have huge rotation, scaling or translation, and thus the same features on each image are close to each other.

4.2.3 Constrains

The parameters used in the control point detection algorithm have been adjusted for the monkey's retinal images. When applying the new algorithm on other type image's registration, some of the parameters might need to be tuned according to the specific image features.

4.2.4 Problem Formulation

Three control points will be identified at the reference images, and the three corresponding control points will be identified as well at the input images. Each pair of control points (one from the reference image and the other one from the input image) should represent the same feature at the retinal vessel's network.

$$(CP_1, CP_2, CP_3) \Leftrightarrow (CP_1', CP_2', CP_3') \tag{4.1}$$

where (CP_1, CP_2, CP_3) are from the reference image and each of them represent one feature of

the retinal vessels; (CP_1', CP_2', CP_3') are the corresponding control points from the input image and each represent same feature as (CP_1, CP_2, CP_3) respectively.

4.3 Automated Preprocessing of the Retinal Images

In the retinal network, the objects of interest are the retinal vessels, i.e. arteries and veins. The registration procedure will operates more efficiently if the retinal vessels have sharper contrast against the background than the fuzzy vessel boundaries. Therefore, image preprocessing is the pre-requisite step of the image registration. Because the retinal vessels and the optic nerve head background have comparable intensity structures and the vessel boundaries are not sharp on the raw images, there are two main processes involved. They are image binarization and retinal vasculature extraction. Currently, the preprocessing procedure is automated using Discriminant Analysis Method and Canny Edge Detector.

4.3.1 Image Binarization Using Discriminant Analysis Method [35]

Image binarization is the very first step in the registration method. The purpose to binarize the raw images is to provide the input data for Canny edge detector [33]. Different procedures have to be involved for handling different color properties of the fundus and IVFA images,

The IVFA image is a grayscale 8-bit image. In addition to black and white, it has other 254 possible gray level colors. This study employs a global adaptive threshold, which was developed by Otsu [32][35], to convert those 256 gray level colors to only black and white colors (Figure 4.3). The output binary BW image has values of 0 (black) for all pixels in the original image with luminance/intensity less than Otsu's threshold and 1 (white) for all other pixels. The threshold is a normalized intensity value that lies in the range [0, 1]. Otsu's method chooses the optimal threshold to minimize the intraclass variance of the black and white pixels and to maximize between-class variance in a grayscale image. The algorithm caculates the statistics of the image

itself to set the threshold and use the histogram (Figure 4.2) to choose its value at some percentile as a reference value of region strength. Therefore, Otsu's method is non-parametric and non-supervised. Consider a discrete image with N×M pixels, with the object (retinal vasculature) in one class Λ_1; and the background (optic nerve head) in another class Λ_2. Suppose *t* is the threshold between the two classes i.e., $f(m,n)_{\Lambda_2} \in [t, L]$, where $f(m,n)_{\Lambda 1}$ are the pixels in the class Λ_1, and $f(m,n)_{\Lambda 2}$ are the pixels in the class Λ_2, the probability of a intensity value *i* is estimated as:

$$P(i) = n_i / (N \times M) \tag{4.2}$$

where n_i is the number of pixels whose intensity value is equal to *i*. The class Λ_1 and Λ_2 's probabilities are defined as $p(\Lambda_1) = \sum_{i=0}^{t-1} pi$ and $p(\Lambda_2) = \sum_{i=t}^{L} pi$, respectively.

Suppose μ_t is the first order cumulative sample moment up to level *t*, and μ_{tot} is the total sample mean. Their definitions are given by $\mu_t = \sum_{i=0}^{t-1} ip_i$ and $\mu_{tot} = \sum_{i=t}^{L} ip_i$.

The sample class means are

$$\mu_1 = \sum_{i=0}^{t-1} ip(i \mid \Lambda_1) = \frac{\mu_t}{p(\Lambda_1)} \; ; \; \mu_2 = \sum_{i=t}^{L} ip(i \mid \Lambda_2) = \frac{\mu_{tot} - \mu_t}{1 - p(\Lambda_1)} \tag{4.3}$$

The between-class scatter sample variance σB^2, is defined as

$$\sigma B^2(t) = \frac{[\mu_{tot}\, p(\Lambda_1) - \mu_t]^2}{p(\Lambda_1)[1 - p(\Lambda_1)]} \tag{4.4}$$

The algorithm evaluates (Eq 4.4) for $0 \le t \le L$, when t^* that maximizes σB^2 is found.

More steps have to be involved when one tries to binarize a color image. A color image needs to be converted into an 8-bit gray scale prior to being converted into B/W image. The

weights of RGB are the three parameters of the conversion. 0.5236 for red, 0.1232 for green and 0.3256 for blue are the default weight parameters widely used in today's commercial image processing software, such as Adobe Photoshop® and Macromedia Fireworks®.

The presented method extracts blue color only when converting to gray scale image (Figure 4.4) because the blue intensity histogram is the closest one to a normal (bell-shaped) distribution (Figure 4.2), and the vessels on the retinal images contain stronger average blue information than red and green or any RGB combination (Figure 4.4). This clarification can be verified by calculating their global discrete entropy using the following equation (Table 4.1):

$$H(I) = -\sum_{j=1}^{N} P_j(I) \log_2 P_j(I) \qquad (4.5)$$

where $P_j(I)$ is the marginal probability distributions of image's RGB content at state j; I is the image intensity value.

Maximal entropy occurs when all potential outcomes are equal [41]. Simpler the image, lower the entropy it has. When the pixels in the image contribute more gray level intensities, the corresponding discrete entropy has higher value. The discrete entropy ranges from zero to $log_2 (n) = 8$ bits. Zero corresponds to the most uninformative case, and $log_2 (n)$ corresponds to the most informative one.

Among these retinal images, all quantities of the discrete entropy are within a range between 0 and 8. Higher the entropy, more information content the image's RGB contribute. When the color image has been converted into gray scale, the vessels become darker than the background. The darker the vessel is, the sharper the contrast will be, and thus making the vasculature extraction more accurate in the next step.

Table 4.1: Monkey1 and monkey2's color and gray images' global entropy

Global Entropy H(I)	Red (monkey1/monkey2)	Green (monkey1/monkey2)	Blue (monkey1/monkey2)
Color Image	4.0098 / 6.9310	5.4289 / 6.9670	6.0919/ 6.5670
Gray Scale Image	7.4186/ 6.7941	7.4186/ 6.7936	7.4183/ 6.8030

The color RGB intensities are inherently inconsistent among different retinal image pairs. One uniform threshold cannot be applied toward all image pairs. A modification on the threshold value (Diagram 4.1) has been made during this study when applying Otsu's method to the color image. Significant black areas are generated using Otsu's method (Figure 4.5a and c), which could not be removed during the vasculature extraction. The overwhelming black area adds noise to the measure of the objective function and causes the insufficient overlapping vessels at the image pair, and thus the noisy black pixels do not correspond to the match of the anatomical structures of the retinal network. Based on the fact that lower the threshold, less black area one will encounter, Otsu's threshold is adjusted by a negative AdjustedThresh. After the adjustment, if the black area has been decreased more than a Nonzero Percentage Threshold, the original Otsu's threshold is kept because the adjustment has high probability to eliminate some true vessels (Figure 4.5d). Otherwise, the adjusted threshold will be employed.

4.3.2 Vasculature Extraction

4.3.2.1 Chain Code Criteria

Chain code criteria are the well-known boundary representation technique [31]. A digital curve can be represented by an integer sequence based on the position of the current edge pixel to their eight neighbors at the 2D spatial domain**:**

$$N_i \in \{1, 2, 3, 4, 5, 6, 7, 8\} \quad (4.6)$$

where,

1 – South, corresponding to an angle of 270 ^{o};

2 – North, corresponding to an angle of 90 ^{o}

3 – East, corresponding to an angle of 0 ^{o}

4 – West, corresponding to an angle of 180 ^{o}

5 – Southeast, corresponding to an angle of 315 ^{o}

6 – Northwest, corresponding to an angle of 135 ^{o}

7 – Southwest, corresponding to an angle of 225 ^{o}

8 – Northeast, corresponding to an angle of 45 ^{o}

4.3.2.2 Vasculature Extraction Using Canny Edge Detector

Canny Edge Detector [33] [34] is used to extract the vasculature (Figure 4.6) of the B/W image obtained from the previous step. Canny's method detects edges at the zero-crossings of the second directional derivative of the image. It performs zero-crossings of

$$\frac{d^2(G \times I)}{dn^2} = \frac{d((\frac{dG}{dn}) \times I)}{dn} \quad (4.7)$$

where, n is the direction of the gradient of the image; G is the edge signal; I is the image intensity. The zero-crossings of Canny's method correspond to the first directional-derivative's maxima and minima in the direction of the gradient. Edges will be identified as the maxima in magnitude. Each pixel's edge gradient is computed and compared with the gradients of its neighbors along the gradient direction. If the central pixel is smaller, mark the current edge's intensity as 0; if largest among all neighbors, keep the original intensity. Based on the nine-pixel neighborhood, the normal to the edge direction has two u_x and u_y. In order to estimate the

gradient on the discrete sampling, two pixels closest to *u* are selected. A plane can be identified by the gradient magnitudes of three pixels. By using this plane, the gradient magnitude and the intensity at each pixel on the line can be locally estimated. The gradient magnitude at $P_{x+1, y+1}$ and $P_{x-1, y-1}$ (Figure 4.1) can be calculated as:

$$P_{x+1, y+1}: \quad G(P_{x+1,y+1}) = \frac{u}{u_y} G(x+1, y+1) + \frac{u_y - u_x}{u_y} G(x, y+1) \tag{4.8}$$

$$P_{x-1, y-1}: \quad G(P_{x-1,y-1}) = \frac{u}{u_y} G(x-1, y-1) + \frac{u_y - u_x}{u_y} G(x, y-1) \tag{4.9}$$

If the gradient at $P_{x,y}$ is greater than both of $G(P_{x+1,y+1})$ and $G(P_{x-1,y-1})$, $P_{x,y}$ will be identified as a maximum.

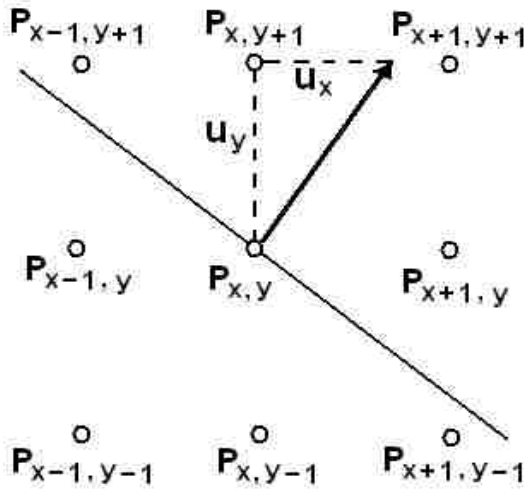


Figure 4.1: Canny Edge Detection – Localization of Maxima [34]

In order to make the localization of magnitude maxima accurate, Canny defined a filter by optimizing a performance index which enhances real positive and real negative. The filter is used to minimize the probability of non-detected edge points and false detection.

$$SNR = \frac{\left| \int_{-w}^{w} G(-x) f(x) dx \right|}{n_0 \sqrt{\int_{-w}^{w} f^2(x) dx}} \tag{4.10}$$

where, *SNR* stands for signal-to-noise ratios; *f* is the filter; denominator is the RMSE response to noise *n(x)*.

The identification of the real edge localization is defined as:

$$Localization = \frac{1}{\sqrt{E[x_0^2]}} = \frac{\left| \int_{-w}^{w} G'(-x) f'(x) dx \right|}{n_0 \sqrt{\int_{-w}^{w} f'^2(x) dx}} \tag{4.11}$$

The closer to the central of the real edge, the better of the localization is. The mean distance between two adjacent is derived as:

$$x_{zc}(f) = \pi \left(\frac{\int_{-\infty}^{\infty} f'^2(x) dx}{\int_{-\infty}^{\infty} f''^2(x) dx} \right)^{1/2} \tag{4.12}$$

Two adaptive threshold values are used in Canny's method. They are high threshold and low threshold. The high threshold is used to find the start point of strong edges. Any points that meet the high threshold will be selected as the edge point. These start points are growing into different directions as long as there is no edge strength falling below the low threshold.

4.4 Control Points Selection Using Adaptive Exploratory Algorithm

Control points selection is the essential step of the image registration. Good control point selection will ensure fused image generated at an efficient computational time. Bad control point selection will significantly increase the computation cost, or even cause the image fusion fail.

Vessels or some particular abnormalities make images not necessarily matching the retina structures. Even when structure and function correspond, the abnormality still happens sometimes if inconsistence exists between structural and functional changes. Further more, IVFA images usually have higher resolution and are rich in information, whereas fundus images have lower resolution and are indeed abstract with some details or even missing some small vessels. Practically, those situations are unavoidable and will create difficulties in extracting the control points because the delineation of the vein boundaries may not be precise.

Generally speaking, there are two broad groups of approaches for vasculature landmark control point analysis. The first group is called “pixel-processing approach” [38]. It uses matched filtering, segmentations, thinning, and bifurcation identification by processing every pixel and imposing numerous operations at each pixel [55] [56]. The “pixel-processing approach” scales poorly with large image size and can hardly meet short computation deadlines. The second group is called “exploratory algorithms” [57] [38]. The algorithm presented in this chapter can be classified into the second group. The vessel tracking/tracing is efficiently performed in the adaptive exploratory algorithm without traveling at every pixel. The algorithm traces the vasculature by locating an initial point and exploiting the local neighbors. Such method is suitable for real-time live ophthalmic image processing [66] with large image size. The following steps are involved for control point selection:

Step1. Split the entire image into two equal size blocks – West block and East block (Figure 4.7). This step is performed under the consideration of retinal image’s feature, i.e. the good control points always come out at the east side or west side of the optic nerve head. The bifurcation point at the north or south side will not be considered as control point candidates.

Based on this feature, this approach is able to split the entire image into two equal size blocks without losing any potential control points.

Step2. Initialize StepCount to 1. Starts at the West block's Northwest corner, read the edge pixels from West to East and from North to South.

Step3. As long as the edge pixels are heading toward East, no matter toward North East, East or South East, mark the current direction as "East".

Step4. Increment the StepCount by 1 as long as the pixel is moving by one pixel toward "East", while straight North or straight South does *not* count for steps (Figure 4.9). As long as the edge pixels are heading toward East, no matter toward North East, East or South East, mark the current direction as "East".

Step5. If direction starts to change, i.e. change from East to West, the StepCount needs to be compared with ROLLBACK threshold (ROLLBACK is a small integer value which is set to 3 in the program). If smaller, roll back the most recent change of StepCount; otherwise, keep the new step direction, and assign the StepCount to a parameter called previous StepCount, and then reset the StepCount to 1. ROLLBACK thresh is used to detect whether or not a direction change is local. When the rollback is triggered, it means the direction change is local, and thus it should not be considered as a real vessel bifurcation. The (X, Y) coordinates of the non-local changing pixel is a possible control point's coordinate that the algorithm will determine later.

Step6. Repeat step 3. If the StepCount is equal to MAXSTEP, a pre-defined threshold, and the previous direction's StepCount is greater than or equal to MAXSTEP, mark the most recent possible control point as a true control point.

Step7. At the East block, read the pre-detected edge pixels starting at the northeast corner and heading towards West. The rest of the procedures are same except the direction reading the

edge pixels is from East to West, and direction change sign is from heading westward to heading eastward.

4.5 Control Point Pair Matching Using Shape Similarity Criteria

The affine transformation model requires three pairs of control points $\{(x_i, x_i), (U_i, V_i)\}$ (i = 1, 2, 3). Firstly, the method takes the image having less number of control points as the grouping base. Suppose image I_1 has n control points, and image I_2 has m control points, and $m < n$, then m will be the group number that how many groups of control points it would have. Secondly, this method combines all control point in I_1 with each control point in I_2, and it totally gets $m \times n$ control point pairs. The method then calculates the distance $|d|$ between each control point pair within each group using $\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$. Thirdly, inside each group, the algorithm chooses the pair with minimum $|d|$. The assumption it uses the minimum distance as matching criteria of the control point pairs is based on the fact that the two images do not have huge rotation, shearing or translation, and thus the same features on each image are close to each other. This is called Shape Similarity Criteria [68]. If there are two or more control points in one image matching the same control point in the other image, the true match will always has smaller |d| than the false match based on the Shape Similarity Criteria assumption. Therefore, minimum |d| is the most qualified measurement to choose the best match of control points among other distance measurement. Finally, the method selects the three smallest distance $|d|$ control point pairs as the final *(x, y)* and *(u, v)* pairs for the transformation model (Table 4.2.4).

4.6 Evaluation Study – Compared with the Vessel Centerline Control Point Selection Algorithm

Vessel Centerline Detection algorithm was applied by Laliberte to identify control points [14]. Firstly, retinal vessel thin image was created with a thinning algorithm that guarantees a one-pixel width. Secondly, pixels with three or four neighbors were identified as the control

points. Thirdly, control points were matched by utilizing the following criteria: (1) control point pair located inside a distance thresh d; (2) with same number of neighbors; (3) distance between angle is less than a thresh σ; (4) eliminate the one without matches.

This study has tested Laliberte's Vessel Centerline Control Point Detection algorithm on the retinal images using both of the heuristic optimization algorithm and the genetic algorithm. The objective function's value is less than using the adaptive exploratory control point detection algorithm. And the fused image has poorer quality than the one created by the presented algorithm. Details will be discussed in the Chapter 5.

4.7 Discussions and Conclusion

In this chapter, an automatic approach is developed for multi-modality retinal image control point detection using feature-based method. Retinal image's feature determines that the ideal control point always come from the vessel bifurcations. The image's property determines the fuzzy vessel boundary, and therefore fuzzy vessel bifurcations. There is a set of control point pixels existing in the fuzzy bifurcation areas. The initial guess of the control point at the Canny edge's global direction change pixel is guaranteed falling into the fuzzy bifurcation areas, and thus close to the required solution. The automatic approach is efficient to handle multi-sensor retinal image registration, as long as the input image, compared with the reference image does not have large rotation or translation. The drawback, which common feature-based approaches have to deal with, is that if there is a huge rotation or translation of the input image, one has to implement area-based registration method to align the image beforehand. The reason is because, as previously discussed, the area-based registration is designed for alignment of single-sensor images with rotation, scaling or translation. Feature-based registration, however, is more suitable for multi-sensor registration that only allows small rotation, scaling or translation. In some rare

cases, image pairs have such bad quality regarding the visibility of the retinal network and the presence of major vessels, that one feels no retinal vascular-based registration techniques can work well on these cases. In such cases, image quality check is often a critical step for the development of a reliable computer-aided diagnostic system [14] [40]. Previous research work has shown that the task of automated multi-sensor image registration is very complex [31]. The feature-based multi-modality registration algorithm presented serves as fundamental step for the hybrid area-based and feature-based systems.

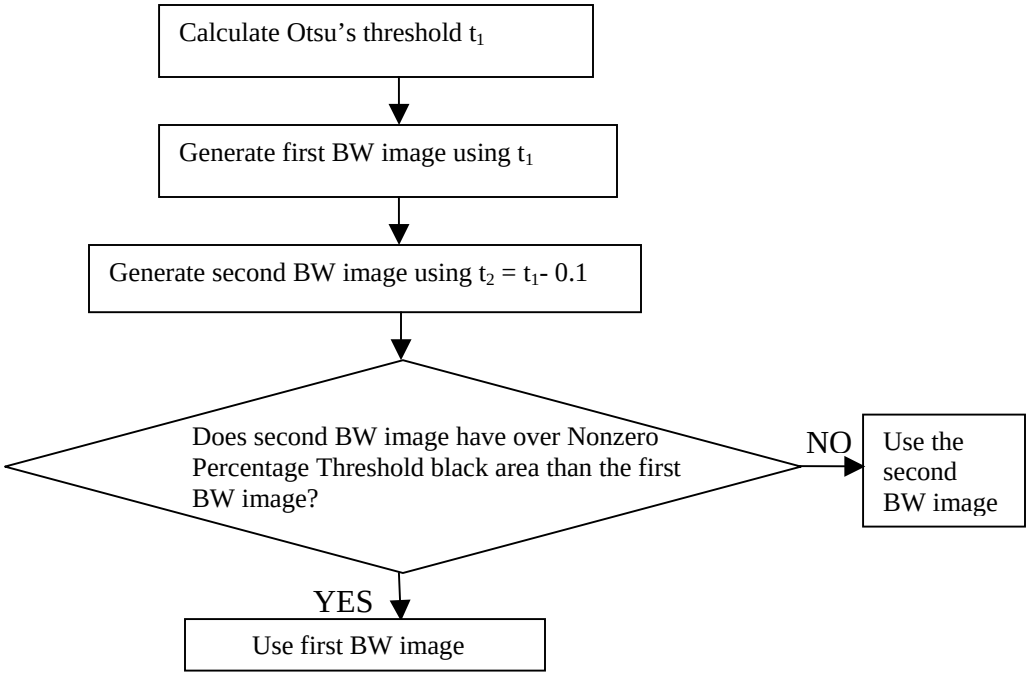
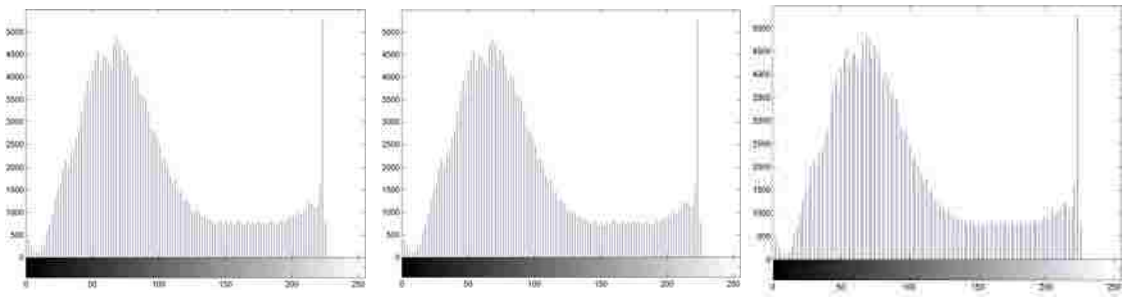
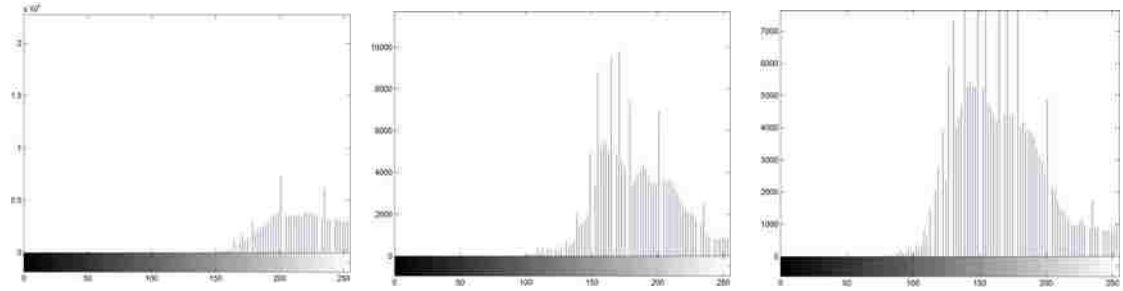


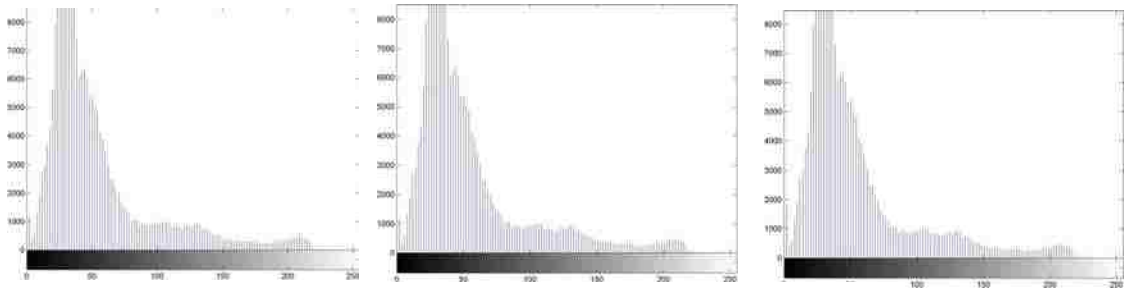
Diagram 4.1: Adjustment of Otsu's Threshold on Fundus Image



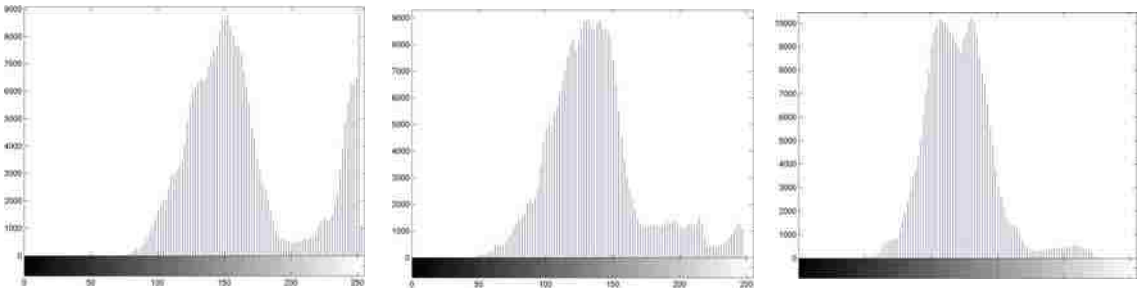
a. Monkey1' IVFA image RGB histogram



b. Monkey1' fundus image RGB histogram

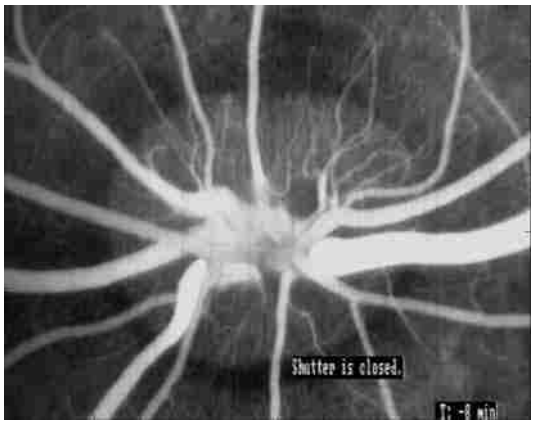


c. Monkey2' IVFA image RGB histogram



b. Monkey2' fundus image RGB histogram

Figure 4.2: Reference and input images' intensity distribution histogram (An image intensity distribution histogram is a chart that shows the distribution of intensities. The plot shows multiple equally spaced bins, each representing a range of color intensity. X-axis is the different color intensity values of the image. Y-axis is the number of pixels within each range.)



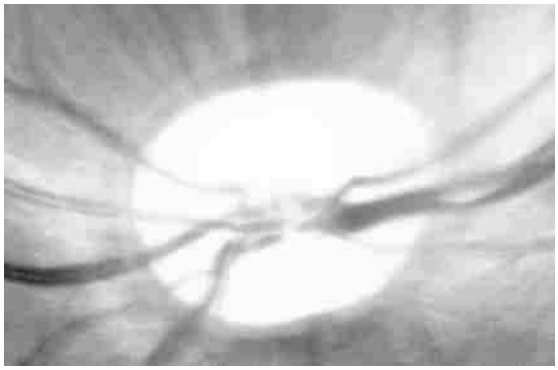
a: Monkey1' IVFA image

b: Monkey1 IVFA's BW image

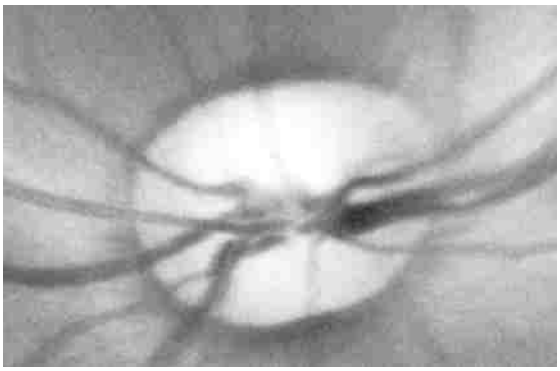
c: Monkey2' IVFA image

d: Monkey2 IVFA's BW image

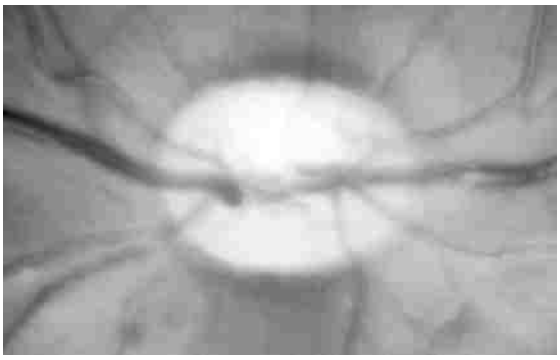
Figure 4.3: IVFA image and the corresponding B/W image using Otsu's threshold



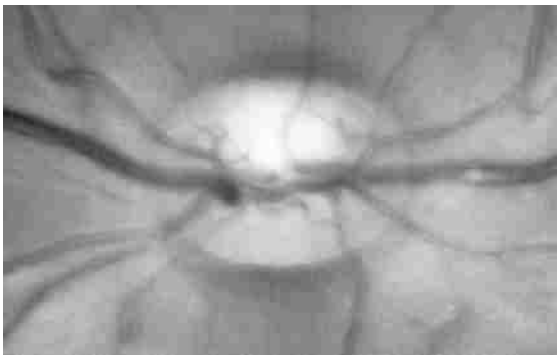
a: Monkey1's 8-bit gray scale image obtained by extracting RGB color using default weights.



b: Monkey1's 8-bit gray scale image obtained by extracting Blue color only.

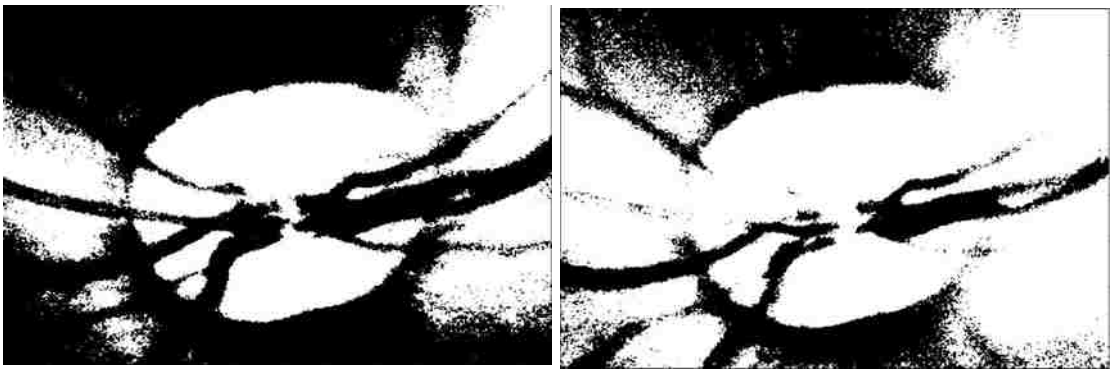


c: Monkey2's 8-bit gray scale image obtained by extracting RGB color using default weights.



d: Monkey2's 8-bit gray scale image obtained by only extracting blue color

Figure 4.4: 8-bit gray scale image obtained by extracting RGB



a: Monkey1’s fundus camera B/W image using Otsu's threshold

b: Monkey1’s fundus camera B/W image using Otsu's adjusted threshold

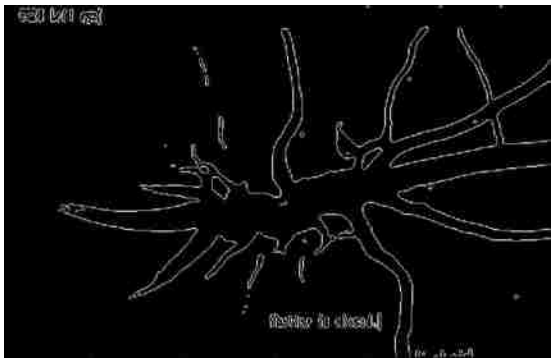
c: Monkey2’s fundus camera B/W image using Otsu’s threshold

d: Monkey2’s fundus camera B/W image using Otsu's adjusted threshold

Figure 4.5: Fundus camera image’s binarization

a. Monkey1's IVFA image contour

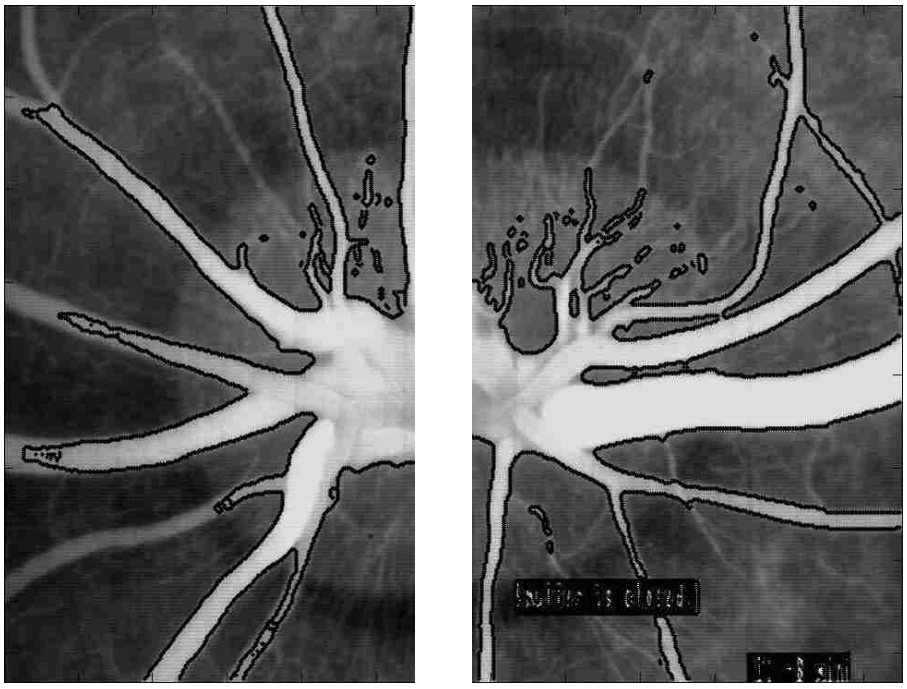
b. Monkey1's fundus camera Image contour



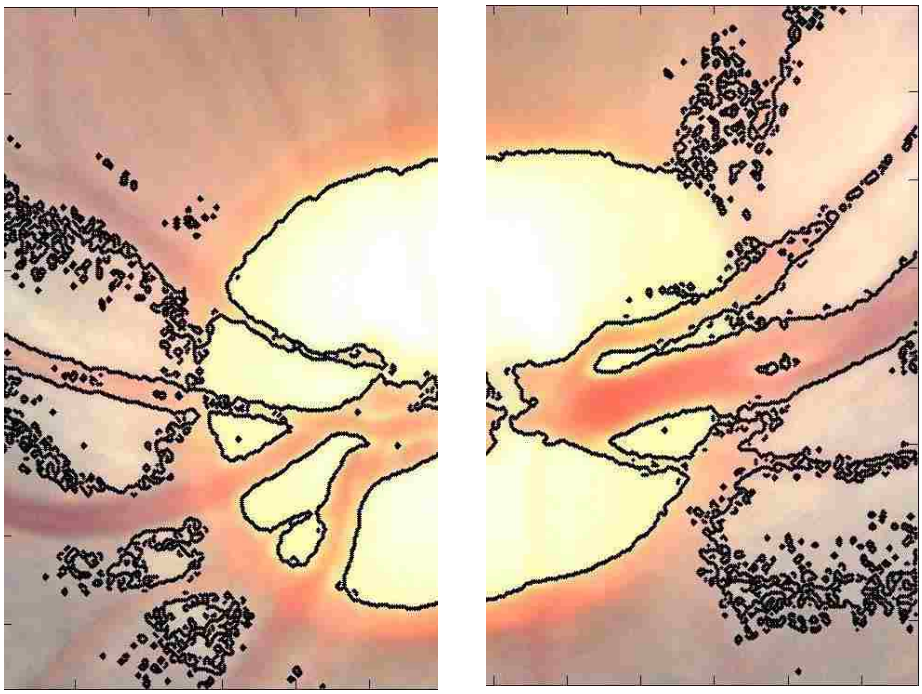
c. Monkey2's IVFA image contour

d. Monkey2's fundus camera image contour

Figure 4.6: Retinal image contour extractions using Canny Edge Detector (Because the Fundus image's B/W image is originally converted from a color image, the vessels are not as clear as those of IVFA images'.)



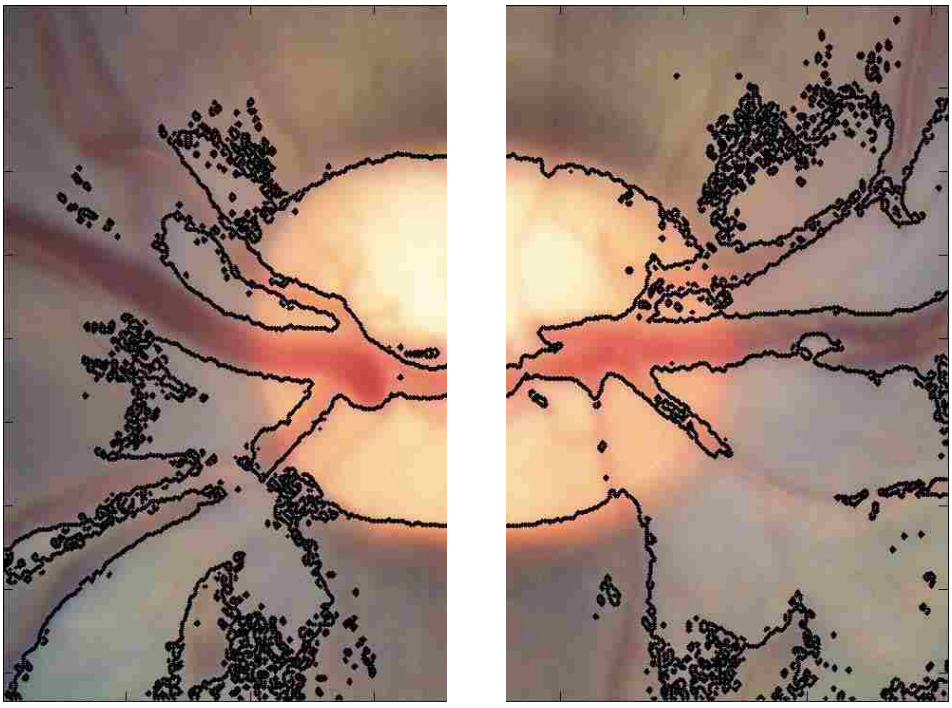
a. Monkey1's IVFA image's contour split into west and east blocks.



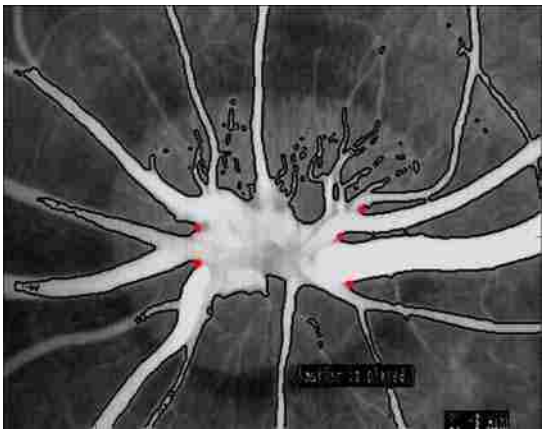
b. Monkey1's fundus camera image's contour split into west and east blocks.

Figure 4.7: West and east blocks of IVFA and Fundus images' contours (Figure Continued)

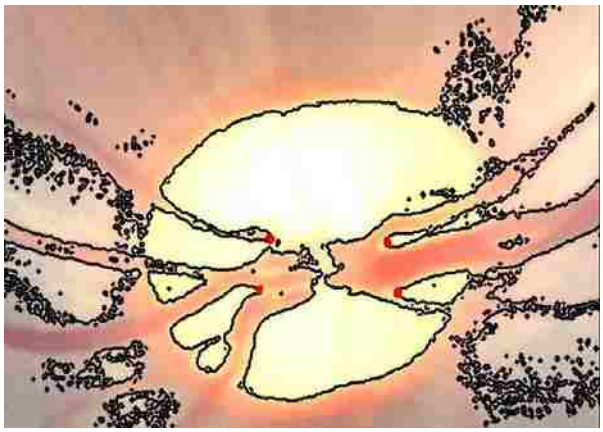
c. Monkey2's IVFA image's contour split into west and east blocks.



d. Monkey2's Fundus camera image's contour split into west and east blocks.

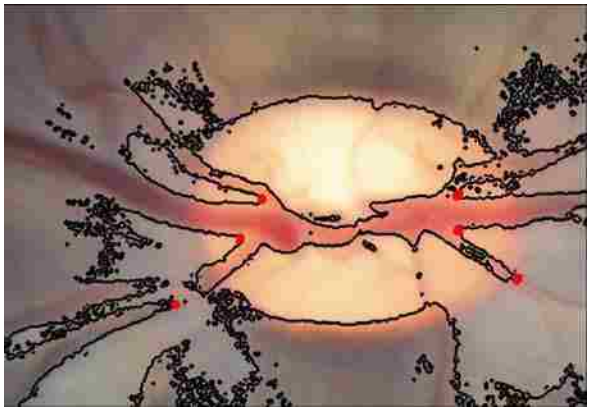


a. Monkey1's IVFA Image control point selection



b. Monkey1's fundus camera color image control point selection

c. Monkey2's IVFA image control point selection



d. Monkey2's fundus camera color image control point selection

Figure 4.8: Monkey1 and Monkey2's IVFA and fundus images' control point selection

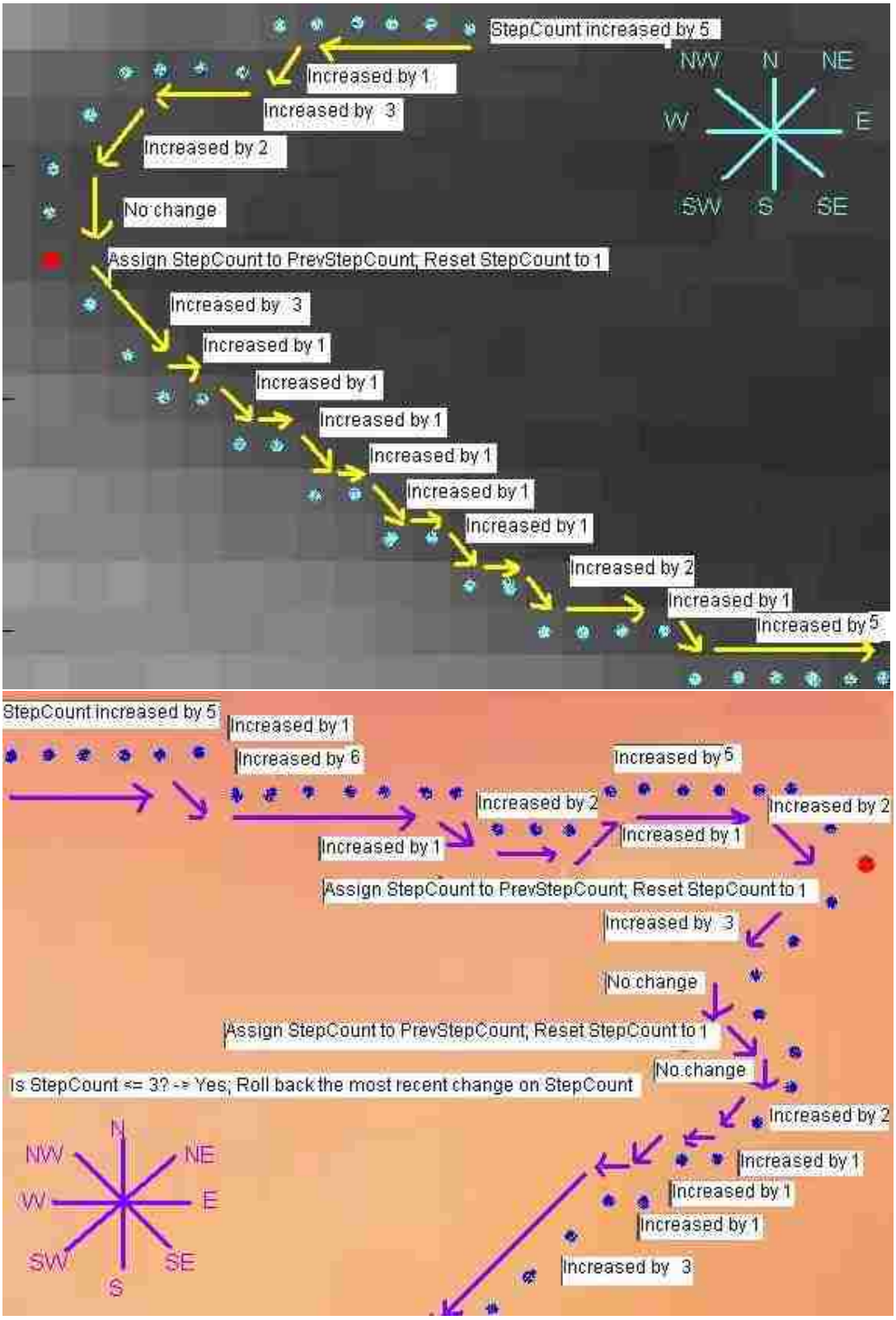
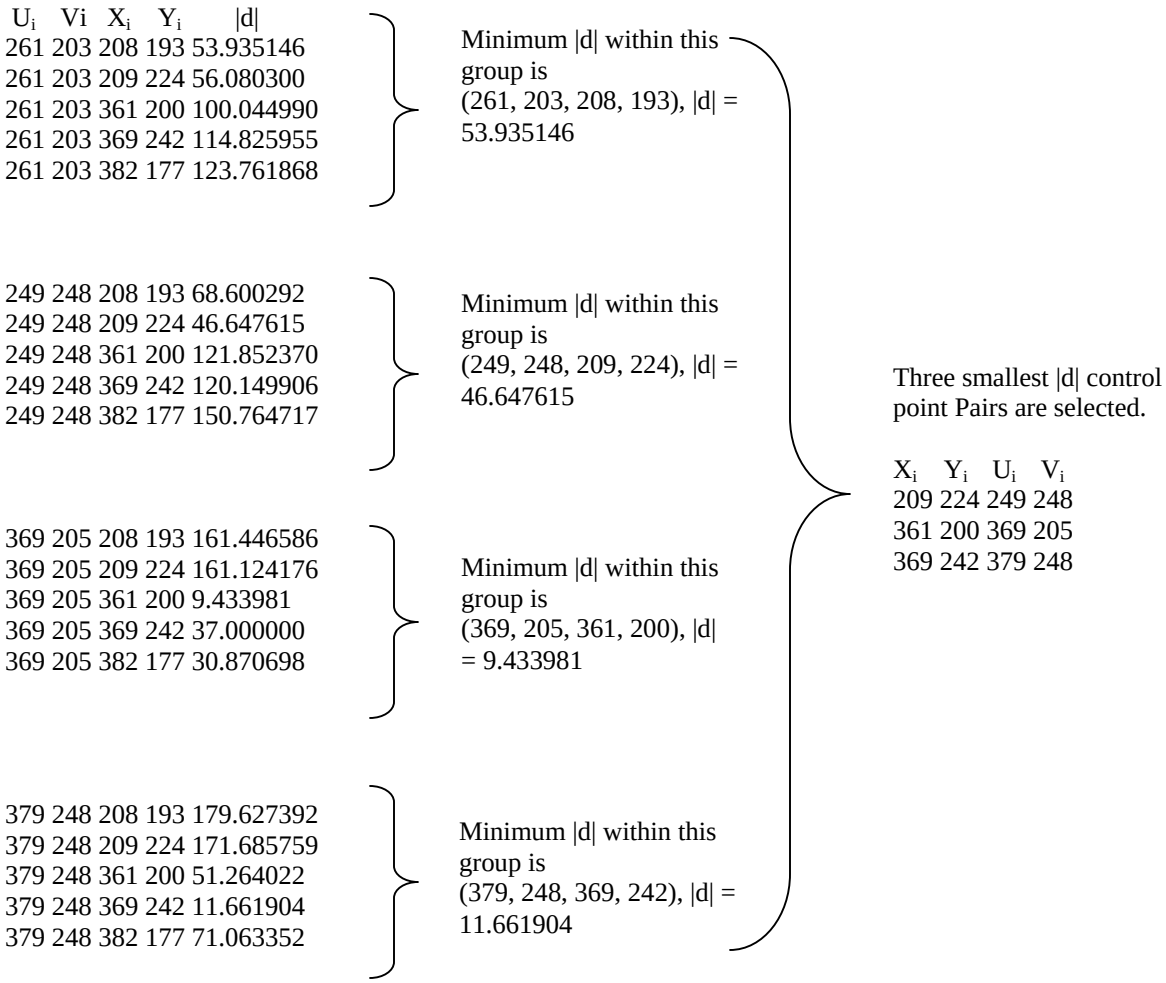


Figure 4.9: Logic to calculate the StepCount (upper- east block; lower- west block)

Table 4.2: Reference image and input image's control point pair matching and the final control point pair selection (Table Continued)



a. Monkey1's control point pair groups with their distance and the selection of minimum |d|.

X_i	Y_i	U_i	V_i	|d|	
216	251	180	293	55.317267	
216	251	250	227	41.617304	Minimum |d| within this group is (216, 251, 250, 227), |d| = 41.617304
216	251	271	190	82.134037	
216	251	475	187	266.790180	
216	251	474	221	259.738330	
216	251	538	269	322.502713	
438	212	180	293	270.416346	
438	212	250	227	188.597455	
438	212	271	190	168.442869	Minimum |d| within this group is (438, 212, 474, 221), |d| = 37.107951
438	212	475	187	44.654227	
438	212	474	221	37.107951	
438	212	538	269	115.104301	
429	179	180	293	273.855801	
429	179	250	227	185.324041	
429	179	271	190	158.382449	Minimum |d| within this group is (429, 179, 475, 187), |d| = 46.690470
429	179	475	187	46.690470	
429	179	474	221	61.554854	
429	179	538	269	141.354165	

b. Monkey2’s control point pair groups with their distance and the selection of minimum |d|.

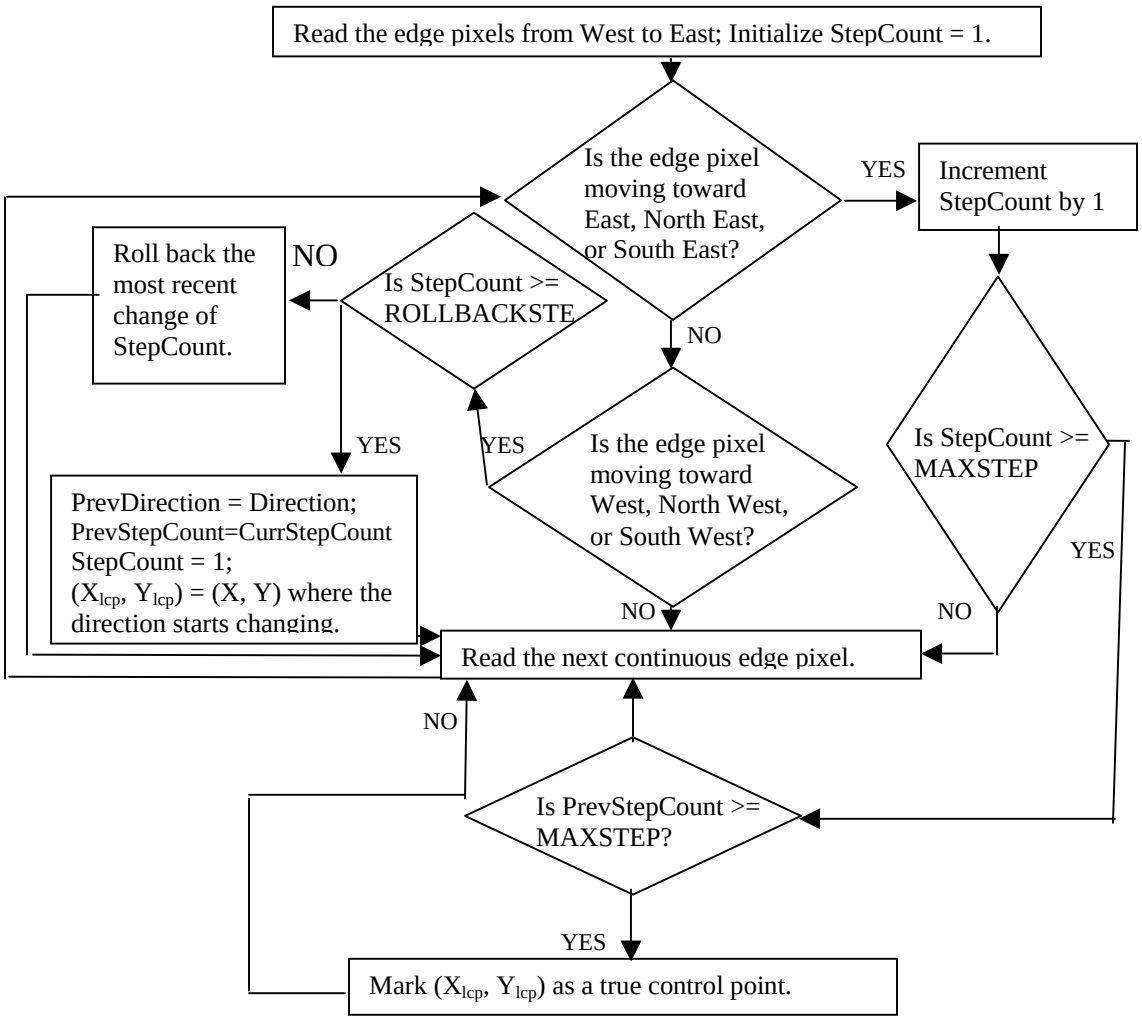


Diagram 4.2: Automated control point selection diagram at the west block (East block's diagram is same except the direction reading the edge pixels is from East to West, and direction change sign is from heading westward to heading eastward).

Chapter 5

Optimization Fusion Algorithm Based on Mutual-Pixel-Count Maximization

5.1 Problem Definition - Optimization

A successful fusion is the one with a good superposition of retinal blood vessels. The fused image created from the control points selected in the previous step does not meet such criteria (Figure 5.1). An optimization procedure is needed to adjust these control points in order to achieve an optimal result. The process of image iterative fusion can be formulated as a heuristic problem of optimizing an objective function that maximizes the Mutual-Pixel-Count match between the reference and the input images. An automatic heuristic optimization scheme based on maximizing Mutual-Pixel-Count has been developed. A refinement of the solution is obtained at the end of each loop (Figure 5.14 and 5.15), and finally a satisfying fused image is generated at the end of the iteration.

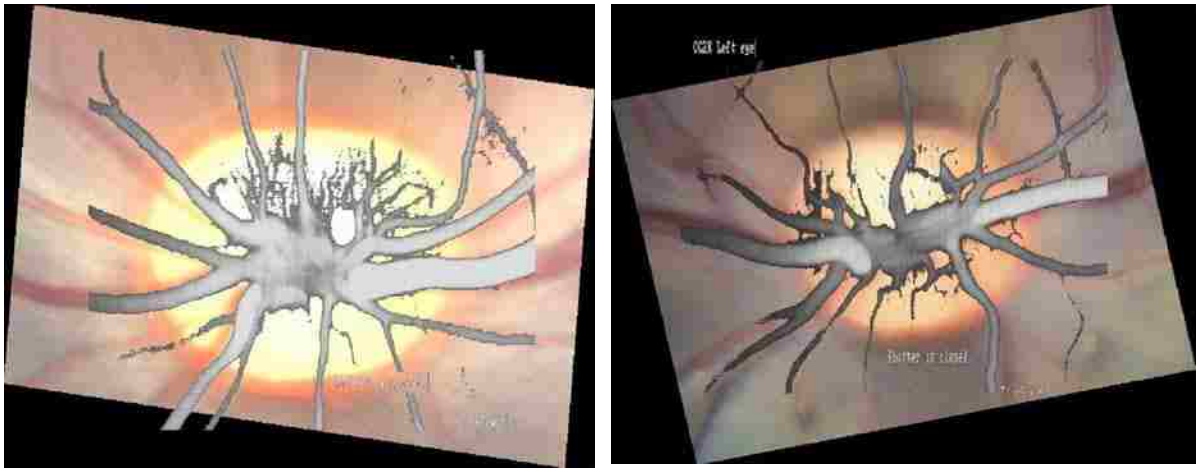


Figure 5.1: Fused images prior to optimization (left-monkey1, right-monkey2)

5.1.1 Knowns

The following information is available for the heuristic optimization algorithm.

(1) Initial guess of the control points

(2) 2D affine transformation model

(3) 2D Images' resolutions and sizes (width: m pixels; height: n pixels)

(4) Images' binary m × n maps with 1 standing for white and 0 standing for black

5.1.2 Assumptions

It is assumed that the retinal vessels are represented by 0 (black pixel) and background are represented by 1 (white pixels) in the binary 2D map.

5.1.3 Constrains

Due to the color image's properties, some of the background on the input image is represented by black pixels that do not contribute to the vessel overlapping calculation. The constrain is that such black background area is not overwhelming.

5.1.4 Problem Formulation

The heuristic optimization algorithm finds the optimal solution by refining the transformation parameters in an ordered way. The reference image's control points' x and y coordinates are fixed during the iteration. Only input image's control points' x and y coordinates are subject to adjustment. By maximizing the objective function, one image's retinal vessels are supposed to be well overlapped onto those of the other image. The fused image is assumed to be optimal when the objective function is maximized.

5.1.5 Objective Function: Mutual-Pixel-Count (MPC) Calculation

Mutual Information (MI) is originally a basic concept derived from information theory. In image fusion area, it describes the amount of information that one image contains about the

other. This study borrowed this idea and simplified it to Mutual-Pixel-Count (MPC) that measures the retinal vasculature overlap for corresponding pixels in both images. To our knowledge, it is the first time that MPC concept has been introduced into the biomedical image fusion. The problem is defined as determining the transformation parameters that are the best match of the input image onto the reference image. After the preprocessing, the binary images of the reference and input images are obtained, i.e. I_{ref} and I_{input}. Only black pixels from both images contribute to the value of the MPC. The most ideal case is that all black (nonzero) pixels of the input image should be mapped onto black pixels of the reference image. MPC is incremented by 1 whenever a mutual pixel is found on both images' retinal vasculature. For further explanation, when a vasculature (black) pixel's transformed *(u, v)* coordinates on the input image correspond to a vasculature pixel's *(x, y)* coordinates on the reference image, the MPC is incremented by 1. Otherwise, MPC remains same. The problem can be mathematically formulated as the maximization of the following objective function:

$$f_{mpc}(x,y,u,v) = \sum_{I_{ref}(x,y)=1 \text{ and } I_{input}(u,v)=1}^{u,v \in ROI} I_{input}(T_x(u,v), T_y(u,v)) \tag{5.1}$$

where f_{mpc} denotes the value of the Mutual-Pixel-Count. T_x and T_y are the transformations for u and v coordinates of the input image. ROI (Region-of-Interest) is the vasculature region where MPC is calculated. This calculation is formulated in pseudo codes as follows:

if B(x, y) = B(x_u, y_v)
then MPC = MPC + 1;
end if

where B is a binary 2D map. In this binary map, 1 denotes the vessel pixel and 0 denotes the background pixel. For all pixels *(u_i, v_i)*, where *(u_i, v_i)* $\in$ *ROI, (x_u, y_v) = T(u, v).*

Assume that the higher the MPC is, the better fusion result one will obtain. MPC is assumed to be maximal if the image pair is geometrically aligned by the transformation *T*. In the retinal image fusion, the high-intensity values of the vessels in the IVFA gray scale image are mapped on low-intensity values of the vessels in the fundus color image. This mapping pattern does not make such limitation on the MPC criteria that a certain intensity values in one image should be mapped onto a certain intensity values in the other image, because MPC criteria is highly independent of the modalities involved. Such property makes MPC a natural choice as the fusion accuracy measurement when multi-modality data are involved.

5.2 Objective Function Maximization Using Heuristic Optimization Algorithm

To solve the optimization problem, a global optimization scheme, e.g. brute force exhaustive search technique, guarantees successful achievement of the global maxima but with the tradeoff on prohibitively expensive computation cost. In the real scenario, the search domain range has to be narrowed down in order to accelerate the execution of the problem. A local optimization scheme is usually employed to reduce the computation cost. However, local optimization can be attracted to local maxima [65]. Heuristic optimization problem does not guarantee that the objective function always reaches the global maxima. The reason is not because the heuristic method is inefficiency, but the fact that the common features on the reference and input images are not identical in most of cases and contains noisy black pixels that could not be completely removed. Therefore, there is no transformation that can match every black pixel (retinal vessel pixel) of I_{input} onto the black pixels of I_{ref} [61]. In the retinal vasculature network, the optimal control points always come from the vessel bifurcations. However, the retinal image's internal properties determine the fuzzy vessel boundaries. Therefore, a set of the control points within the fuzzy bifurcation area are the candidates, where the best control points that correspond to the

global maxima of MPC exist. The initial guess of the control point at the Canny edge's global direction change pixel is guaranteed falling into the fuzzy set area. Ideally, the object function f_{MPC} appears containing a global maximum surrounded by numerous local maxima.

In order to achieve global maxima equivalent result with a lower computation cost, ROI is selected prior to calculating MPC. The entire BW image's retinal vasculature is split into three regions, i.e. west region, middle region, and east region. The middle region is the optic nerve head that has the highest value of intensity distribution among the entire retinal image. West and east regions appear to be darker than the optic nerve head because of their lower intensity value distribution. Only two regions among three are selected for MPC calculation. That which two comparing regions will be selected depends on the intensity property of the color image's corresponding BW image. The comparing regions will not be same if the color image property is inconsistent from different image pairs. The east region and the west region are the preferred ones. The middle region is the backup in case that either of the former regions is not qualified to be selected as the comparing region. When one judges how good a fused image is, the overlapped vessels at the west or east side of the optic never head is a good indicator. However, the vessels at the central optic nerve head are not because those central vessels always have better overlap than the side ones. The algorithm sets a BA percentage threshold to determine whether or not the west region should be selected as the comparing region. If the black area of the west region is greater than BA percentage threshold, it will not be selected. By only calculating partial retinal vasculature, an optimal result is obtained with an MPC which is very close to the global maxima at an efficient computation cost on local maxima calculation. The reason is because if the west region has over a certain percentage of black area, the retinal vessels cannot be well extracted from the black area. Hence, the overwhelming black area does

not make a good contribution to MPC even though one employs a certain global scheme to compare the black area.

Optimization algorithm finds the optimal solution by refining transformation parameters in an ordered way. During the iteration, the reference image's control points' x and y coordinates are fixed. Only input image's control points' x and y coordinates are subject to adjustment. The method first increments the 1st control points' x coordinate by step size *s*. If the MPC is increased because of this movement, the algorithm keeps increasing this coordinate by *s* till MPC stops increasing. Then *y* coordinate is adjusted by the same way. Step size *s* is a pre-defined value, which is set to less than 1 in order to ensure the registration at the sub-pixel accuracy level. Generally, lower the step size, higher registration accuracy one can receive, but higher computation cost one will suffer. Therefore, choosing a balanced step size can ensure good registration accuracy and also maintain a reasonable computation cost. The program selected 0.5 as the step size. Secondly, Decreasing the coordinates is applied by *s* at the same order. This procedure is repeated at each of the three control points till MPC is no long increased. Convergence criteria are used to determine when the iteration is finished. Adjustment of the coordinates is iteratively implemented until either of the following convergence criteria is met:

(1). Predefined maximum number of loops has been reached.

(2). Updated Mutual-Pixel-Count ∇MPC is smaller than ε, i.e.

$$|f_{MPC}{}^{n+1}(x, y, u, v) - f_{MPC}{}^{n}(x, y, u, v)| < \varepsilon \qquad (5.2)$$

where ε is a very small non-negative threshold value. Higher the ε, the faster the convergence happens. The smallest non-negative value is 0. The recommended range is [0, 10]. The fused image would have slightly difference when setting different ε values. The difference is such minor that human eyes cannot identify it. Therefore, the difference is negligible as long as ε

falls into the recommended range. ε is set to zero in the program. Inside each loop, a different set of transformation parameters is refined. This process is formulated in pseudo code as follows:

```
currentPixel = startPixel(u, v);
nextMPC = 0;
for all four neighbors of currentPixel
    loop do
        (u_i, v_i) = neighbor of currentPixel in an ordered way;
        if ( MPC(u_i, v_i) > nextMPC )
            nextMPC = f_MPC (u_i, v_i);
            currentPixel = (u_i, v_i);  // i: number of neighbors that have been visited.
        else
            break;
        end if
    end loop
```

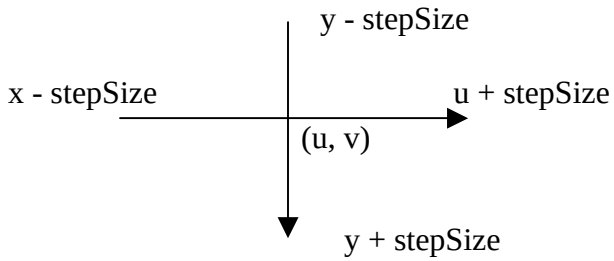


Figure 5.2: Definition of four-neighbor neighborhood

The iteration can also be summarized as the following steps (Diagram 5.1).

Step1. Initialize MPC using the control point pairs generated from the registration step. Initialize the loop counter to 0. The maximum allowable number of loops is set to L (L is set to be 7 in the program), with the purpose of preventing too much computation caused by the mismatched control points. If there is any mismatched control point pair from the previous registration step, the iteration will not stop after L loops. If the iteration loop number reaches L, it means the fusion failed because of the mismatched control point pairs (Figure 5.16). The average loop count of the successful fusion is significantly lower than L. The monkey1's fusion result reached the optimal result after 3 loops, while monkey2's is 2 loops.

Step2. Start with input image's 1st control point pair's x-coordinate u_1. Increase u_1 by step size s. Keep increasing u_1 till MPC stops increasing. Increase input image's 1st control point pair's y-coordinate v_1 by same step size s. Keep increasing v_1 till MPC stops increasing. Suppose ∇M is the changed volume of MPC, then:

$$\nabla M = MPC - MPC_{init} \Rightarrow \nabla M \begin{cases} > 0: & \text{Keep moving the coordinate toward the same direction} \\ \leq 0: & \text{Stop U or V coordinates' movement in that direction} \end{cases} \quad (5.3)$$

Step3. Decrease input image's 1st control point pair's x-coordinate u_1 by step size s. Keep decreasing u_1 till MPC stops increasing; Decrease input image's 1st control point pair's y-coordinate v_1 by same step size. Keep decreasing V_1 till MPC stops increasing.

Step4. Repeat step2 and step3 on the 2nd and 3rd control points of the input image till MPC stops increasing.

Step5. If L, which is the pre-defined maximum loop counter, is reached, stop the iteration and report fusion failed; if not reached, repeat step2 and step3 till MPC no long increases.

Step6. Generate the fused image using the current control points' coordinates through the transformation model T.

5.3 Evaluation Study - Compared with the Optimization Genetic Algorithm Using the Same Objective Function

As a well-known global optimization technique, Genetic Algorithm is often used to solve the optimization problem [61][69]. Genetic Algorithm has been implemented by Matsopoulos [61] to determine the optimal transformation parameters in the retinal image registration and fusion. The Genetic Algorithm Matsopoulos proposed has two operations, i.e. crossover and mutation (Figure 5.4 and 5.5).

In this comparative study, there are totally 3 groups of the populations, and each group stands for one control point of the input image. The population size remains same of each generation. The probability is 0.9 and 0.1 for crossover and mutation, respectively. Each individual is an unsigned char array (Figure 5.3). Array size (individual length) is a random number. The elements in the array are the random number s_n that ranges from 0 to 3, standing for 4 possible moving directions (East, West, North and South). The initial 3 individuals are the coordinates of the initial guess of the control points ((x_1, y_1), (x_2, y_2), (x_3, y_3)). Each individual's fitness is estimated by the MPC objective function. Half of the total individuals (parents + children) with higher MPC are selected. Offspring generation is iteratively implemented till the termination condition is true.

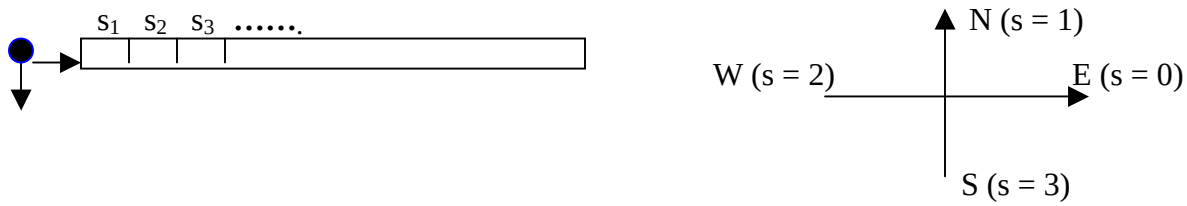


Figure 5.3: Individual's chromosome structure

5.3.1 Crossover

90% of the individuals will have the crossover operation. Two individuals are randomly chosen among one generation. One chromosome is randomly cut (several elements) from individual A. The same operation is applied on the second individual. Two chromosomes are switched. After crossover operation, two children are generated. Four individuals' fitness (two parents and two children) are calculated by evaluating the objective function f_{MPC}. Two individuals that have best fitness (higher MPC) as those in the new generation are kept and the worst two are discarded.

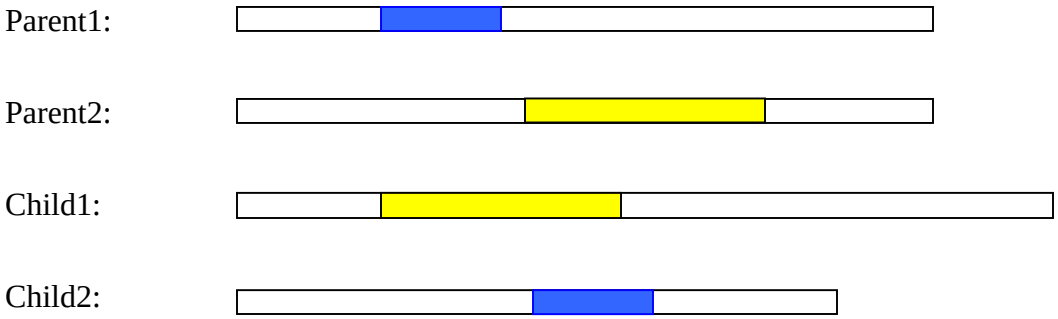


Figure 5.4: Parent individual and child individual by crossover operation

5.3.2 Mutation

10% of the individuals will have this mutation operation. One individual is randomly selected and one element of the chromosome is randomly selected as well. Randomly generate another s value and replace the original one with this newly generated s.

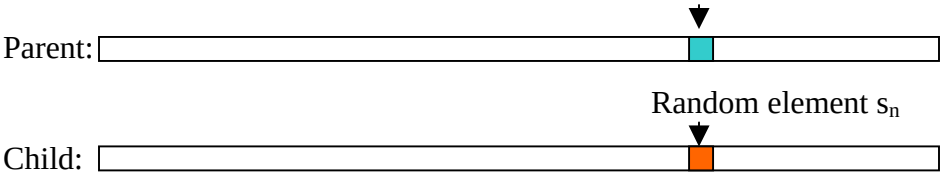


Figure 5.5: Parent individual and child individual by mutation operation

Calculate the new child's fitness and compare it with its parent. Put the better one to the new generation and discard the worse one.

5.3.3 Genetic Algorithm's Performance

The GA process is formulated in the pseudo code as follows:

Initialize the first generation composed of n individuals randomly;
while not meet generation number
- *calculate the objective function of the n individuals;*
- *apply crossover and mutation to generate offspring randomly;*
- *calculate the objective function of the offspring*
- *select n individuals from both of the parent and offspring generation with better fitness*
- *replace the current generation by the newly selected n individuals*

end while loop

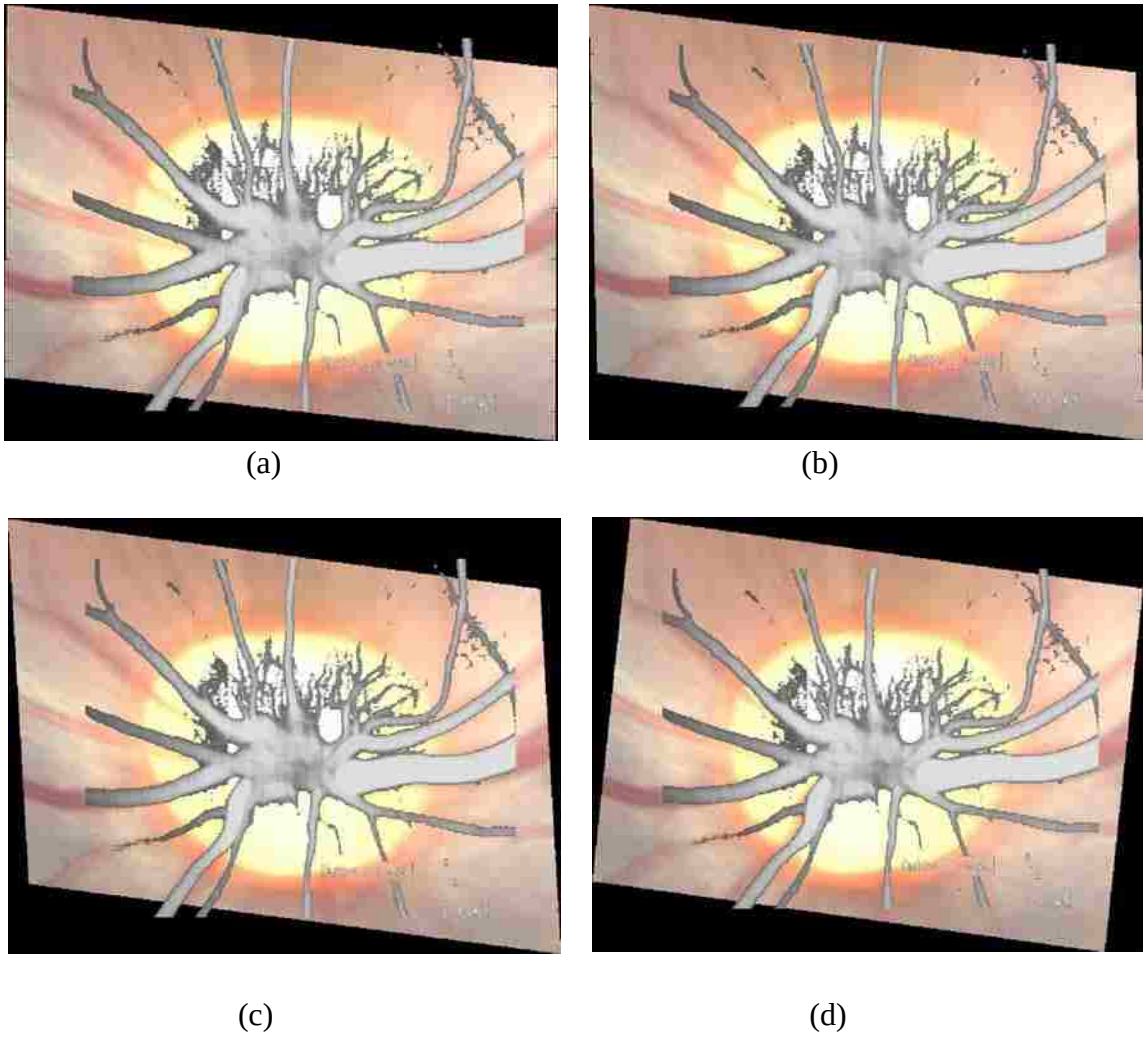
It can be observed that the heuristic algorithm achieved better optimization than GA at the less running time consumption (Figure 5.6 and 5.7). Probably the most important drawback of genetic algorithms is their strong dependence on a set of parameters (e.g., size of the population, number of generations, probabilities for applying the random operators, rate of generational reproduction, etc.) that have to be experimentally tuned for the optimization problem. Consequently, unless the user has experience in the resolution of the concrete optimization problem at hand by means of genetic algorithms, the choice of the most suitable values for all the parameters is converted itself into another optimization problem.

Table 5.1: GA's performance of 20 Runs

GA's MPC (objective function)	Monkey1	Running Time Monkey1	Monkey2	Running Time Monkey2
Run 1	6043	22 s	31082	7.5 m
Run 2	6095	21 s	31227	7.2 m
Run 3	6205	22 s	31397	7.7 m
Run 4	6210	22 s	31425	7.4 m
Run 5	6300	21 s	31487	7.2 m
Run 6	6421	21 s	31499	7.3 m
Run 7	6471	22 s	31569	7.6 m
Run 8	6650	21 s	31589	7.4 m
Run 9	6412	22 s	32007	7.4 m
Run 10	2062	22 s	32185	7.3 m
Run 11	6329	22 s	32189	7.3 m
Run 12	6062	23 s	32220	7.3 m
Run 13	5981	22 s	32247	7.3 m
Run 14	6239	22 s	31567	7.3 m
Run 15	6512	21 s	31128	7.2 m
Run 16	6551	22 s	32185	7.3 m
Run 17	6688	22 s	31395	7.3 m
Run 18	6513	22 s	32189	7.3 m
Run 19	6041	22 s	31403	7.3 m
Run 20	6342	22 s	29842	6.9 m

Table 5.2: Comparison of GA and the heuristic algorithm's average performance

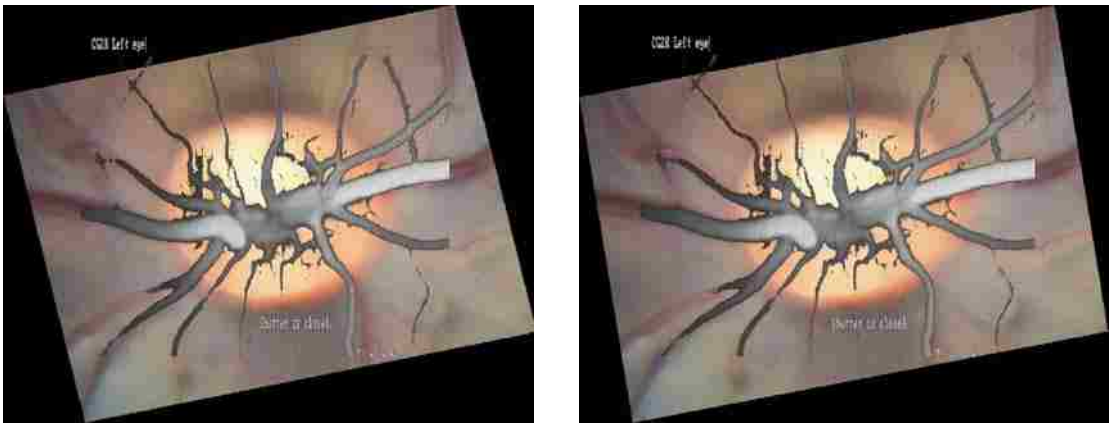
Average MPC (objective function)	Monkey1's Average MPC	Monkey1's Average Running Time	Monkey2's Average MPC	Monkey2's Average Running Time
GA	6186	22 seconds	31830	7.4 minutes
The heuristic algorithm	7732	13 seconds	32277	1 minute



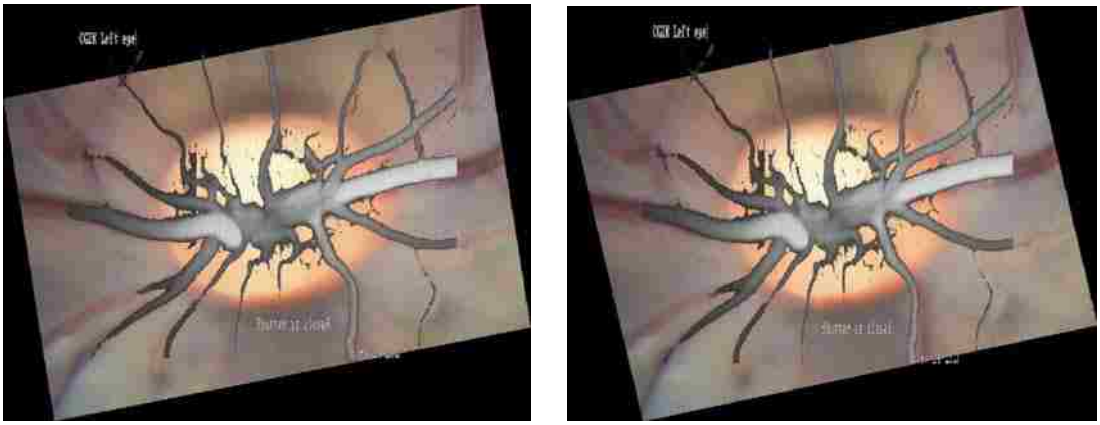
(a) (b)

(c) (d)

Figure 5.6: Monkey1's fused image generated by GA: (c) f_{MPC} = 5981; (d) f_{MPC} = 6095; (a) f_{MPC} = 6300; (b) f_{MPC} = 6471.



(a) (b)



(c) (d)

Figure 5.7: Monkey2's fused image generated by GA: (a) objective function f_{MPC} = 31227; (b) objective function f_{MPC} = 32007; (c) objective function f_{MPC} = 32185; (d) objective function f_{MPC} = 32220.

5.4 Evaluation Study – Comparison with the Centerline Control Point Detection Algorithm

This study has tested Laliberte's Vessel Centerline Control Point Detection algorithm which was described in chapter 4 on the retinal images using both of the heuristic optimization algorithm and the genetic algorithm. Based on the initial control points, monkey1 and 2's objective function f_{MPC} is 3686 and 7143, which are lower than the adaptive exploratory algorithm's result with 5144 and 31227, respectively. The final f_{MPC} after applying the heuristic

optimization algorithm is 6828 and 29720, which is apparently lower than the adaptive exploratory algorithm's final MPC value, i.e. 7732 and 32220 for monkey1 and monkey2, respectively. Therefore, the adaptive exploratory algorithm has better performance than the Centerline Algorithm in quantity in terms of the objective function and in human's observation (Figure 5.8).

5.5 Objective Function Comparison

5.5.1 RMSE Minimization Objective Function

Root Mean Square Error (RMSE) [11] [14] [15] between the coordinates of the reference image and the transformed image has been used for some type image fusion as the measurement for the fusion accuracy. RMSE is defined as:

$$\text{RMSE} = \sqrt{(\sum_{i=0}^{N}(x_{ture} - x_{reverse})^2 + (y_{ture} - y_{reverse})^2) / N} \qquad (5.4)$$

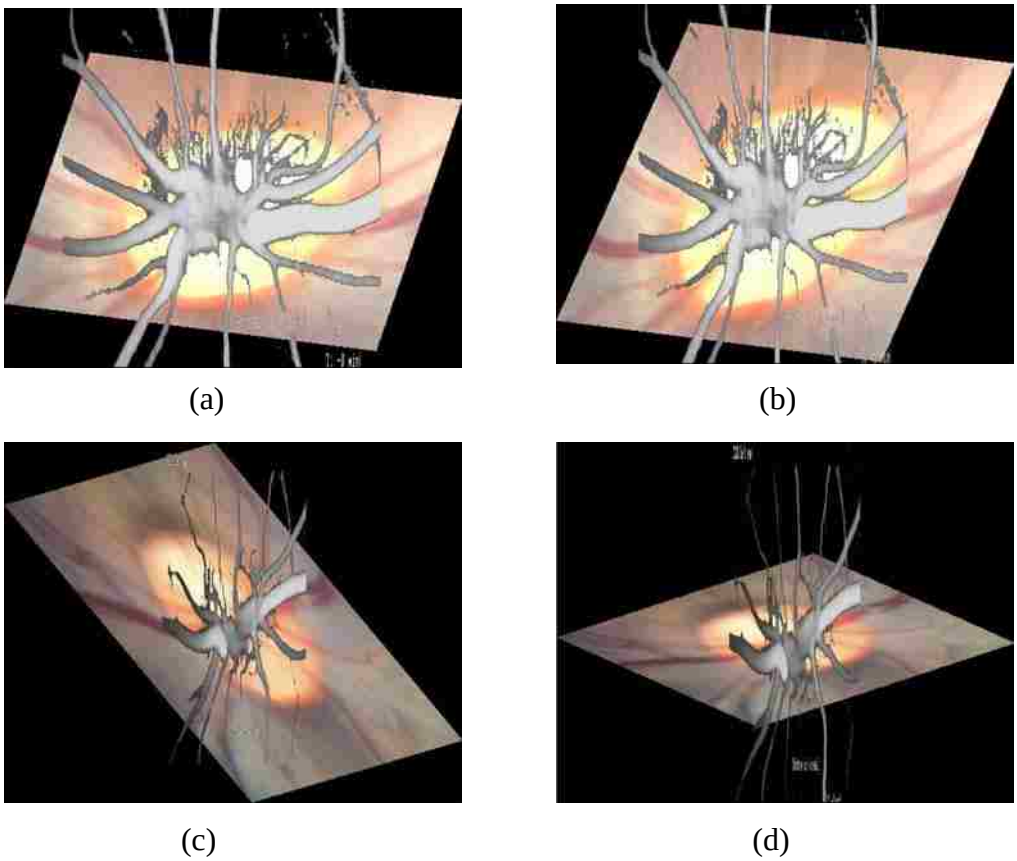
where N is the number of common pixels exist in both images. The x_{true} and y_{true} are the true coordinates from the reference image. The $x_{reverse}$ and $y_{reverse}$ are the reverse transformed coordinates by applying the reverse transformation of the known 2D affine polynomials.

The goal of the RMSE objective function is to minimize RMSE, in order to achieve the optimal fused image. This study has tested RMSE objective function on both monkeys (Figure 5.9 and 5.10). It can be observed that the fused image is not necessarily getting improved when RMSE is becoming smaller. The program has never been able to achieve same quality fused image as the f_{MPC} objective function does. RMSE is dropping from 120.25 to 17.98 for monkey1, and from 124.18 to 49.30 for monkey2 during the iteration. However, there is no clear relationship between the quality of the fused image and the value of RMSE that can follow a

certain pattern. Therefore, RMSE is not an ideal objective function when applying to monkey's retinal image fusion.

Table 5.3 – Comparison of Centerline Control Point Detection Algorithm with Adaptive Exploratory Control Point Detection Algorithm

Objective Function f_{MPC}	Centerline Algorithm (monkey 1)	The Adaptive Exploratory Algorithm (monkey 1)	Centerline Algorithm (monkey 1)	The Adaptive Exploratory Algorithm (monkey 2)
Initial f_{MPC}	3686	5144	7143	31227
GA final f_{MPC} (Average)	6268	6186	29337	31861
The heuristic final f_{MPC}	6828	7732	29720	32277



(a) (b)

(c) (d)

Figure 5.8: Monkey1 and Monkey2's fused images by applying Centerline algorithm.
(a): by heuristic optimization with f_{MPC} = 6828 for monkey1;
(b): by GA algorithm with f_{MPC} = 6268 for monkey1;
(c): by our heuristic optimization with f_{MPC} = 29720 for monkey2;
(d): by GA algorithm with f_{MPC} = 29337 for monkey2.

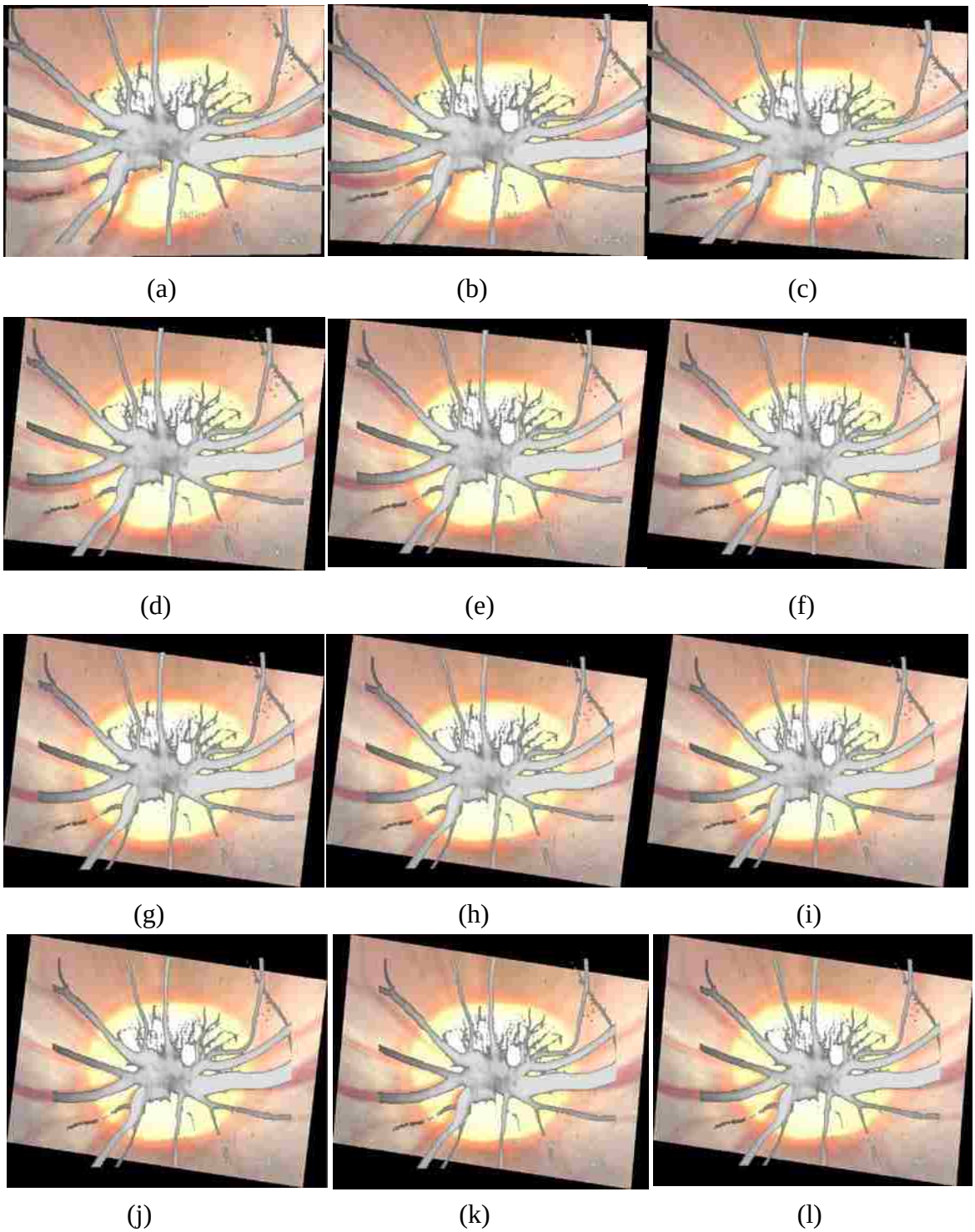


Figure 5.9: Monkey1's Fused images generated by using RMSE objective function during the iteration. RMSE (a) = 17.98; RMSE (b) = 28.68; RMSE (c) = 37.14; RMSE (d) = 49.21; RMSE (e) = 57.96; RMSE (f) = 68.87; RMSE (g) = 78.09; RMSE (h) = 89.33; RMSE (i) = 97.18; RMSE (j) = 109.15; RMSE (k) = 113.96; RMSE (l) = 120.25;

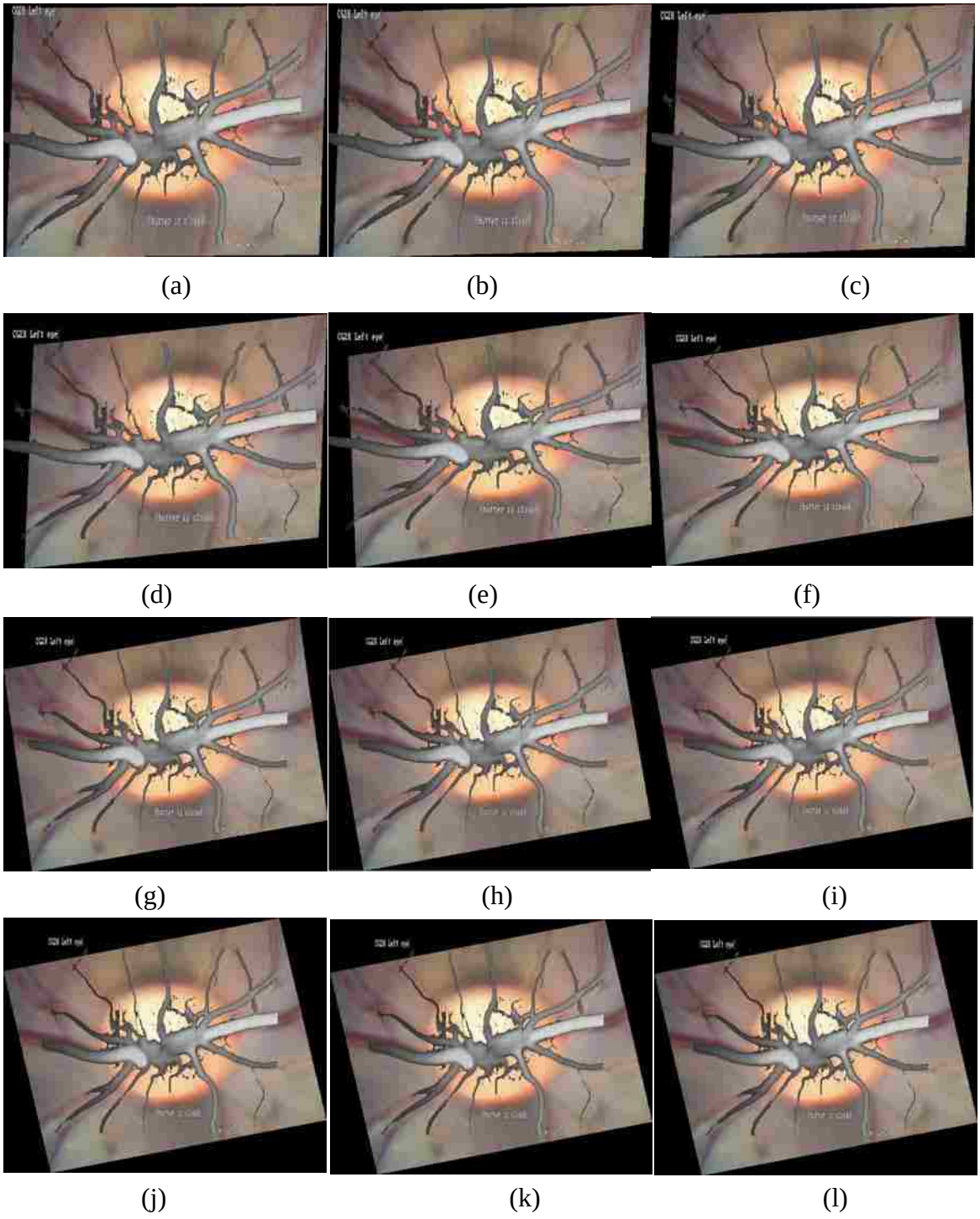


Figure 5.10: Monkey2's Fused images generated by using RMSE objective function during the iteration. RMSE (a) = 49.30; RMSE (b) = 57.01; RMSE (c) = 64.76; RMSE (d) = 71.34; RMSE (e) = 80.86; RMSE (f) = 90.05; RMSE (g) = 99.54; RMSE (h) = 106.61; RMSE (i) = 112.11; RMSE (j) = 119.51; RMSE (k) = 122.83; RMSE (l) = 124.18.

5.6 Parameter Extraction

Once three pairs control points' coordinates are available, one can solve the Gaussian matrix to get the parameters $P \in \{a_1, a_2, a_3, a_4, b_1, b_2\}$.

Gaussian Matrix

$$\begin{bmatrix} u_1 \\ v_1 \\ u_2 \\ v_2 \\ u_3 \\ v_3 \end{bmatrix} = \begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_1 & y_1 & 1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_2 & y_2 & 1 \\ x_3 & y_3 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_3 & y_3 & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ b_1 \\ a_3 \\ a_4 \\ b_2 \end{bmatrix}$$

$\Rightarrow$ Monkey1's P = {1.0488, -0.1067, -2.0575, 0.1211, 0.9429, -37.7558};

Monkey2's P = {0.9950, 0.1925, -79.1302, -0.1669, 0.8526, 98.7476}

5.7 Applying to Human's Retinal Images

During this study, humans' retinal image pairs have been tested using the presented control point detection, registration and fusion algorithm (Figure 5.17). The challenge is that the physicians have much more difficulty to control humans' eye motion than monkeys' during the examination. According to the information provided by LSU New Orleans Eye Center, the experimental monkey was anesthetized with intramuscular ketamine (7-10 mg/kg), xylazine(0.6-1 mg/kg), and intravenous pentobarbital (25-30 mg/kg). Administration of the anesthetics was repeated alternately every 30 minutes as required to maintain the animal in deep, stage IV anesthesia [36]. However, the anesthetics procedure cannot be simply repeated on humans just for retina examination purpose. Therefore, in many cases, the image quality of the retina optic nerve head does not meet the registration and fusion requirement. This is actually a common challenge that the real-time live ophthalmic image processing has to deal with [38] [54] [66]. The

problems include, but not limited to non-ideal imaging conditions, high image variability, and short computation deadlines. The control point selection is based on a good vasculature extraction assumption. However, retinal image appearance is highly variable in practice regarding to color, texture, structure, and lesions [14]. When applying to human's retinal images, human color pigmentation, age, and eye motion, etc. could impose some practical constraints that might cause the registration fail. Unlike the non-live imaging system where the conditions can be carefully controlled and adjusted in order to make the imaging optimal, retinal imaging are frequently subject to unavoidable body and eye movements. This research has applied the registration and fusion algorithm on two patients' retinal image pairs. Apparently, binarization of the fundus color images did not work very well. Therefore, the extraction of the vasculature did not ensure the control point selection's success. The plan is to address the issues with human eye motion and poor vessel extraction, as well as improve the robustness of the algorithm in the near future.

5.8 Computation Complexity

The code was written in MATLAB 6.5. The program was implemented on a Dell Latitude D620 Laptop with Intel Core Duo 1.66GHz processor, 2.0GB memory, 100GB hard drive, and Microsoft Windows XP Professional. The average running time is 13 seconds for the first monkey and 55 seconds for the second monkey. The difference is because the comparing regions where MPC was calculated were not same. Larger the comparing regions, higher the computation cost one would get. Monkey1's maximal MPC is 7,732, while monkey2's is 32,277 when the solution reaches the optimal result. The platform running the MATLAB code does not execute programs as fast as Desktop PC with similar configuration. The running time will be dropped by 20-30% if one runs the program on a Core Duo processor Desktop PC.

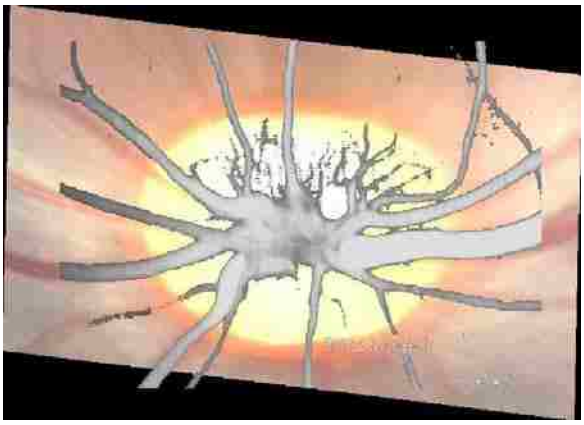
5.9 Comparison with Other Existing Fusion Approaches

5.9.1 Fusion Result Comparison – Accuracy

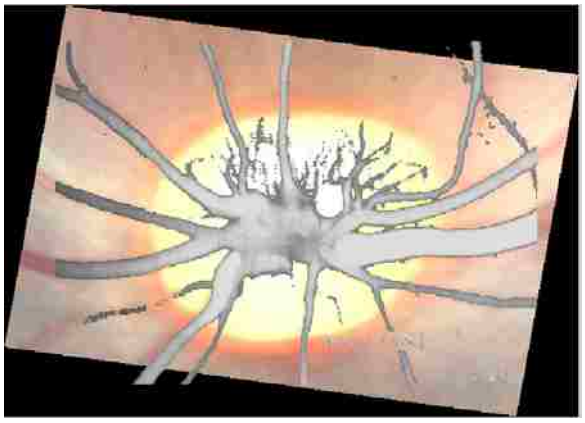
Manual approach is commonly used in clinical practice. The ophthalmologist identifies control points at vessel bifurcations, which are common to both images that are to be registered. The control points placed by the experts seemed appropriate. However, the fusion result might not be optimal at many cases (Figure 5.11a). The disadvantage of human-interactive approach includes, but not limited to inaccuracy in the placement of control points, inconsistency of the fusion results, and significantly increased interaction time during manual adjustment of the control points. Figure 5.11b shows an improved fused image after 30-minute manual adjustment. Figure 5.11c confirms the advantage of the automatic scheme in terms of accuracy against the manual approach.

5.9.2 Program Running Time Comparison -Efficiency

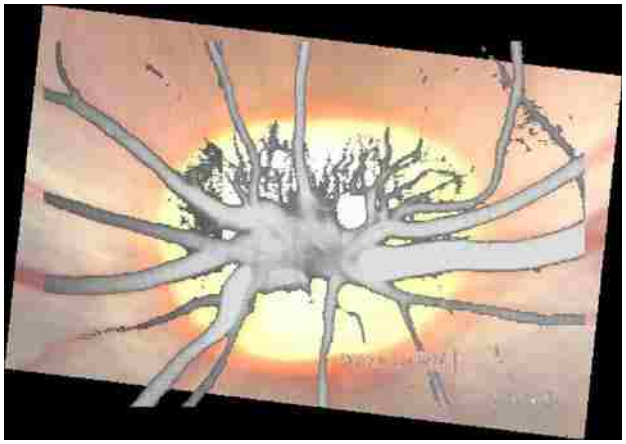
The average processing time for human manual registration and fusion is 35 minutes, including initial control point selection, manual adjustment of control points' coordinates, and evaluation of the fusion result after each adjustment. Three computer-aided strategies were proposed by B. Ma in [47]. In Ma's proposal, uniform spatial sub-sampling approach, vector quantization algorithm, and stratified sampling with centroid refinements were discussed and tested. Among these three approaches, stratified sampling with centroid refinements strategy achieved shortest average running time of 11 minutes for satisfactory fusion results. G. Matsopoulos and N. Mouravliansky proposed an automatic retinal image fusion scheme using global optimization techniques in [61]. They reported an average execution time of 4.5 minutes. With less than 1 minute running time, the new algorithm therefore, has significant advantage in terms of efficiency (Table 5.1).



a. Initial fused image by manual registration (Human interactive time is 2 minutes)



b. Improved fused image by manual adjustment (Human interactive time is 30 minutes)



b. Fused image by the automatic scheme (No human interaction with running time of 17 seconds)

Figure 5.11: Comparison between the automatic and the manual approach for monkey 1

5.9.3 Number of Threshold Parameters Comparison – Automation

Dependency of a few tunable thresholds is inevitable in various retinal image registration and fusion methods [2] [14] [61] [65]. More threshold parameters, more human's interaction required when the parameters have to be adjusted, and hence less automation level the program is. Laliberte's algorithm [14] has 10 threshold parameters of which 7 are dependent on the image resolution and 3 left main free. In the presented registration and fusion program, there are 9

adjustable threshold parameters. Among those thresholds, 5 are dependent on the image resolution/size and 4 are left main free.

5.10 Discussion

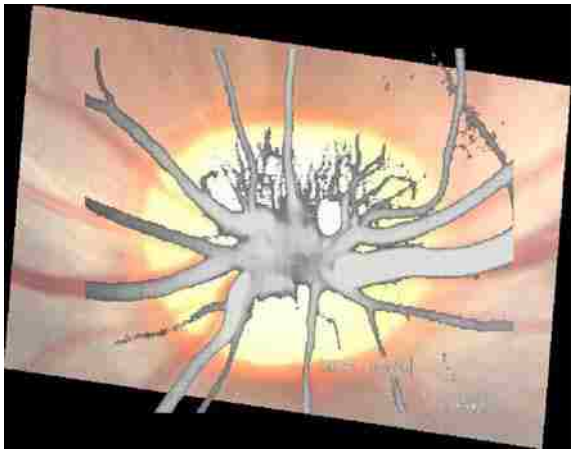
When generating the fused image, one always wants to display as much original information as possible. Therefore, the threshold extracting veins from the IVFA image is different from the threshold calculated from Otsu's method. Otsu' threshold was used for identifying the bifurcation points. As long as the bifurcation point is at the ideal location, the entire vessels are not required. However, displaying the entire vessel of the IVFA image would benefit the future analysis. Two thresholds are used to extract the IVFA's vessels because the darkness of gray levels at each side is not consistent.

Table 5.4: Running time comparison

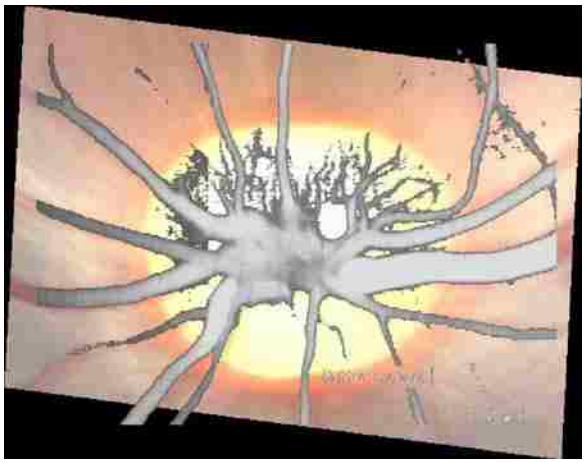
Methods	Average Running Time
Manual approach	35 minutes
Uniform spatial sub-sampling	24 minutes
Vector quantization algorithm	19 minutes
Stratified sampling	11 minutes
Matsopoulos's method	4.5 minutes
The new method	< 1 minute

Table 5.5: Threshold parameter number comparison

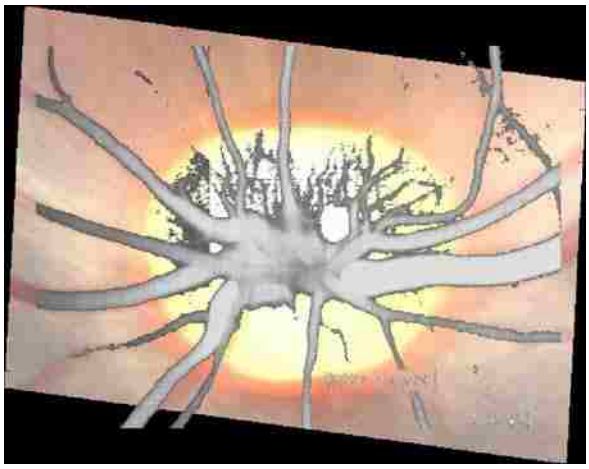
Methods	Total threshold Parameters	Image resolution dependent	Main free
Laliberte's method	10	7	3
The new method	9	5	4



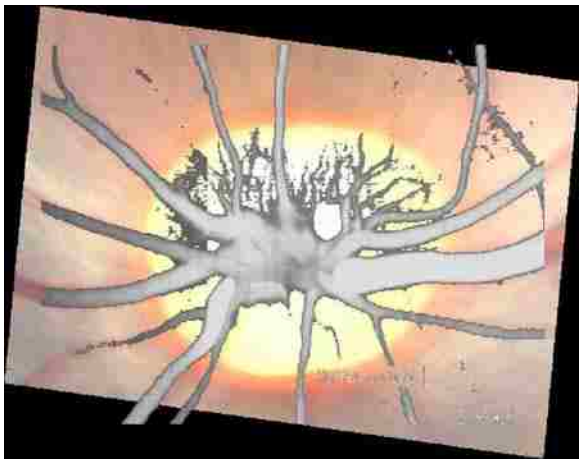
Fused image at MPC = 5144



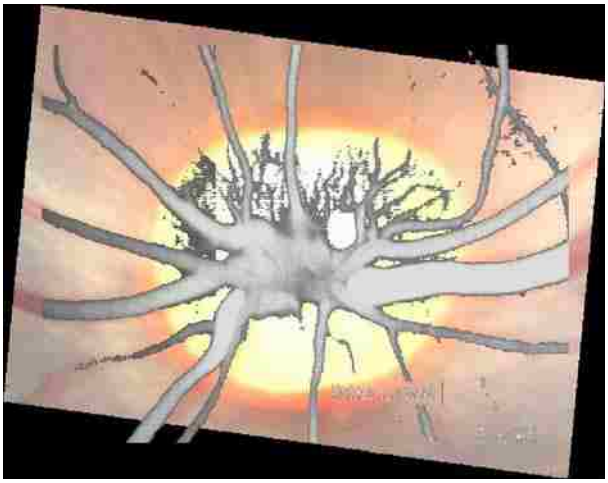
Fused image at MPC = 7396



Fused image at MPC = 7484

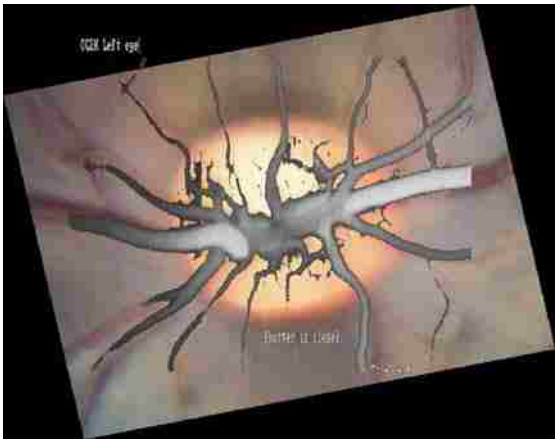


Fused image at MPC = 7681

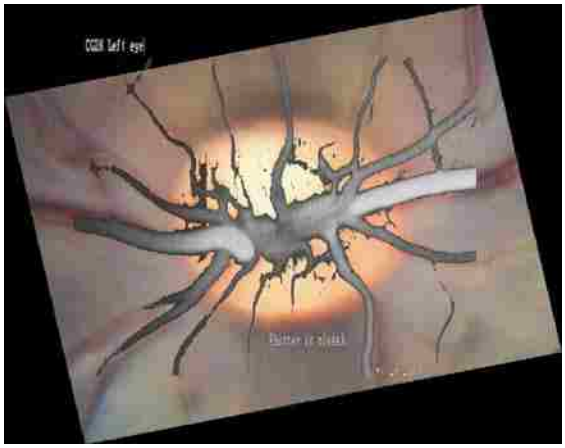


Fused image at MPC = 7732

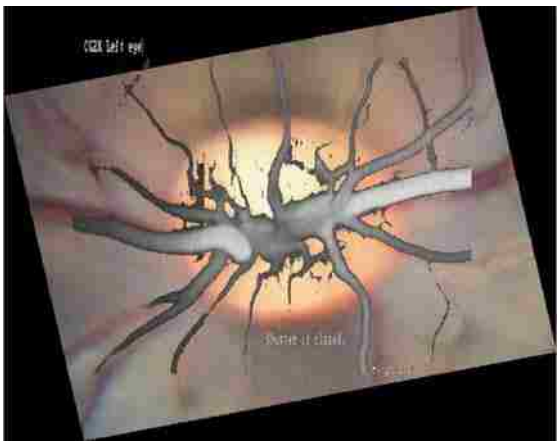
Figure 5.12: Monkey1's fused image improvement during the iteration



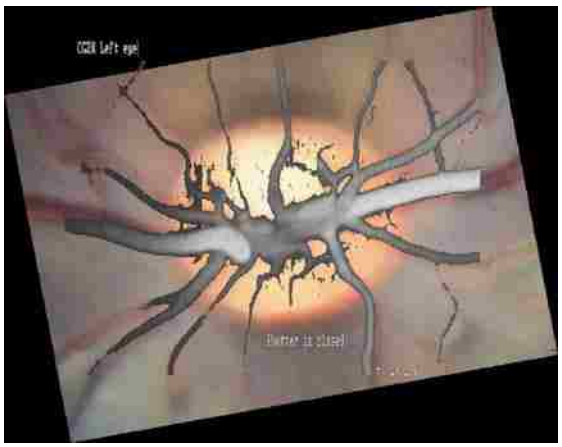
Fused image at MPC = 30732



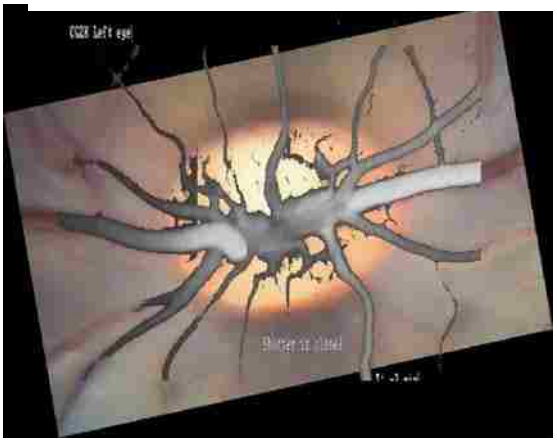
Fused image at MPC = 30888



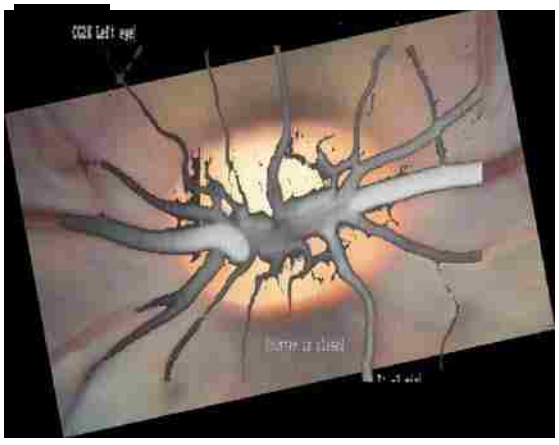
Fused image at MPC = 30914



Fused image at MPC = 31134



Fused image at MPC = 32254



Fused image at MPC = 32277

Figure 5.13: Improvement of Monkey2's fused image during the iteration.

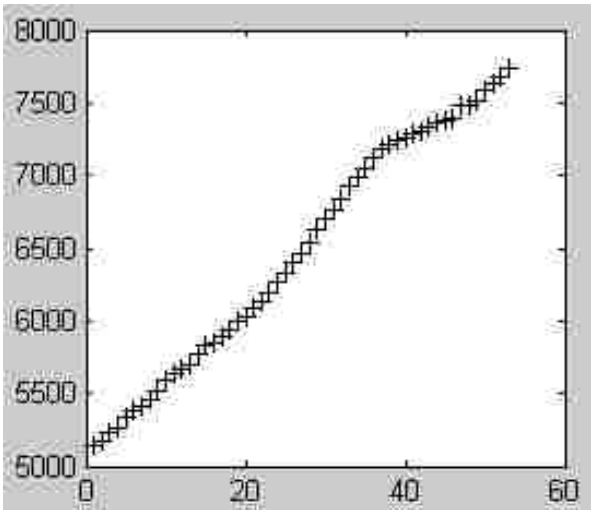


Figure 5.14: Monkey1's MPC increasing during the iteration. Y-axis is the MPC value; X-axis is the loop count.

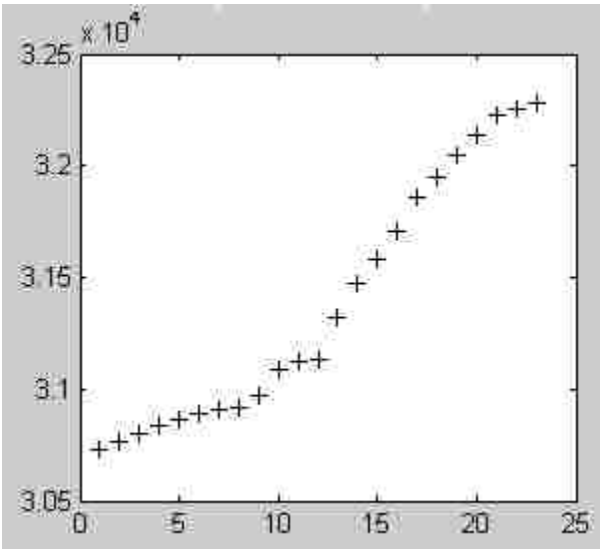


Figure 5.15: Monkey2's MPC increasing during the iteration. Y-axis is the MPC value; X-axis is the loop count.

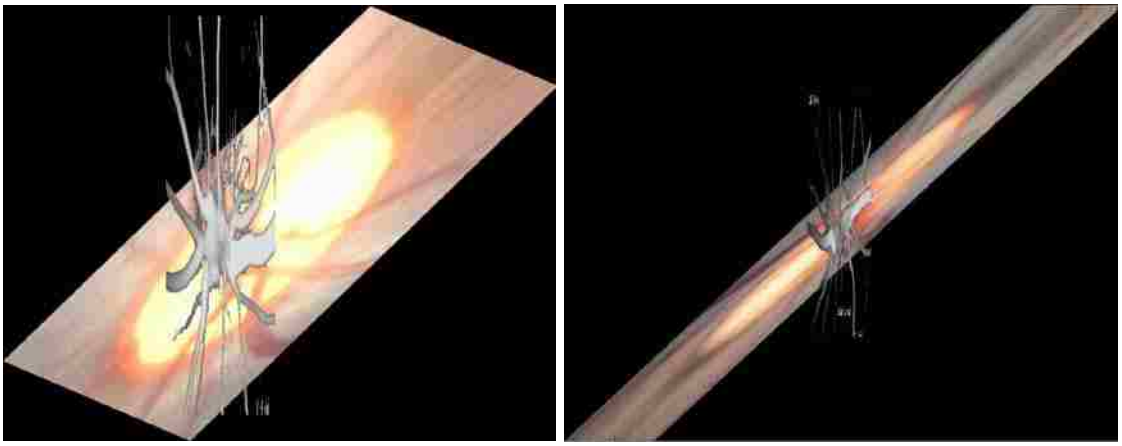
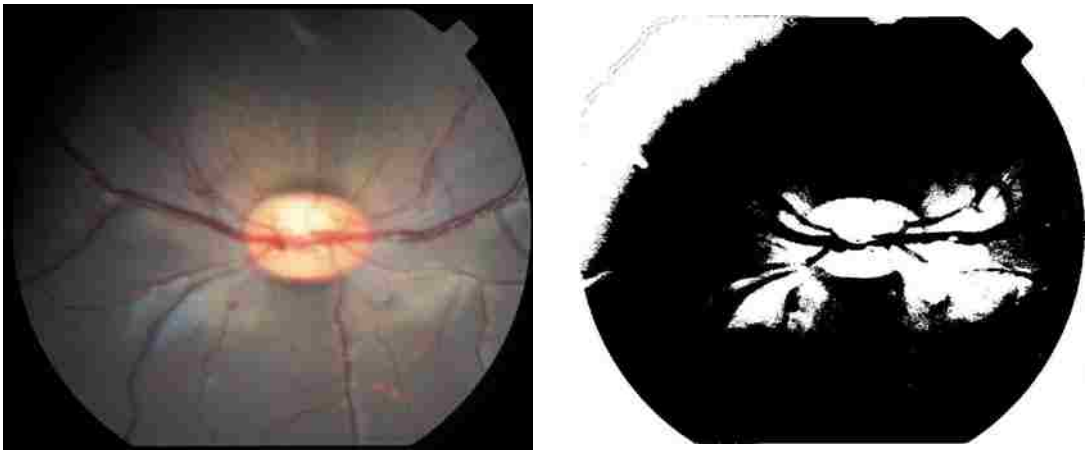
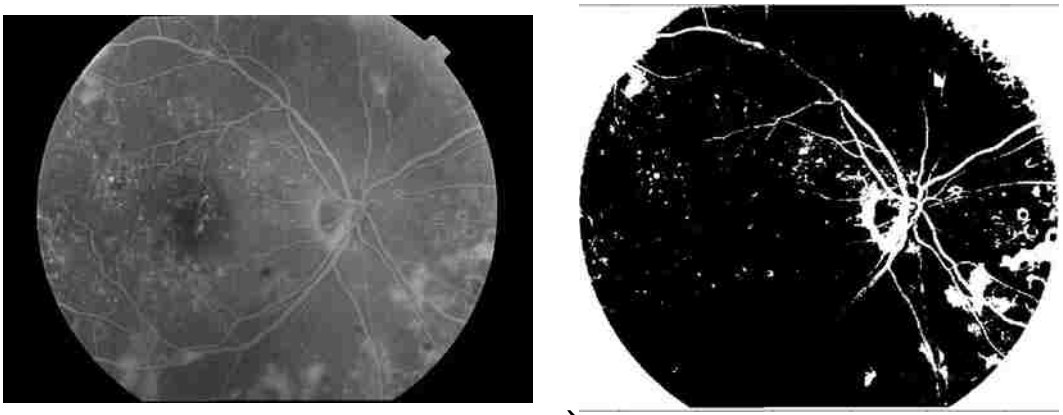


Figure 5.16: Mismatched control point pair causing fusion failed with loop number reached *L* (Left – monkey1; right – monkey2).



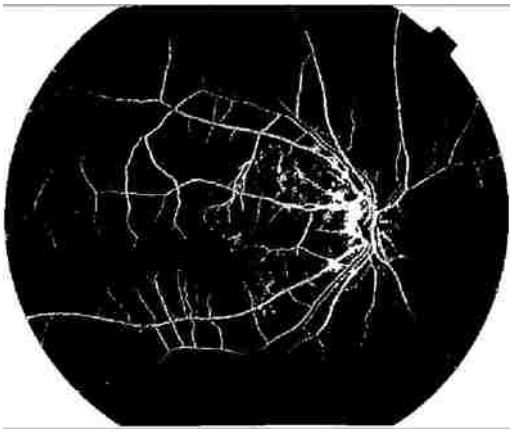
a. Monkey2's Fundus Image ⇒ Vasculature



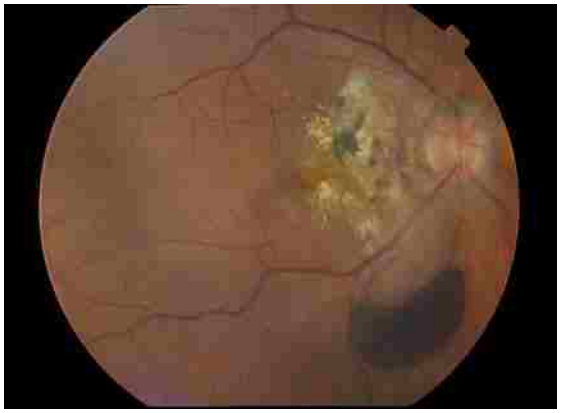
b. Patient 1's IVFA Image ⇒ Vasculature

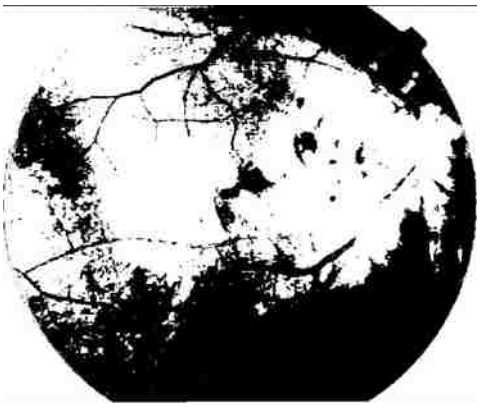
Figure 5.17: Monkey2's retinal image v.s. two human patients' retinal images (Figure Continued).

c. Patient 1’s fundus Image ⇒ Vasculature



d. Patient 2’s IVFA Image ⇒ Vasculature





e. Patient 2’s fundus Image ⇒ Vasculature

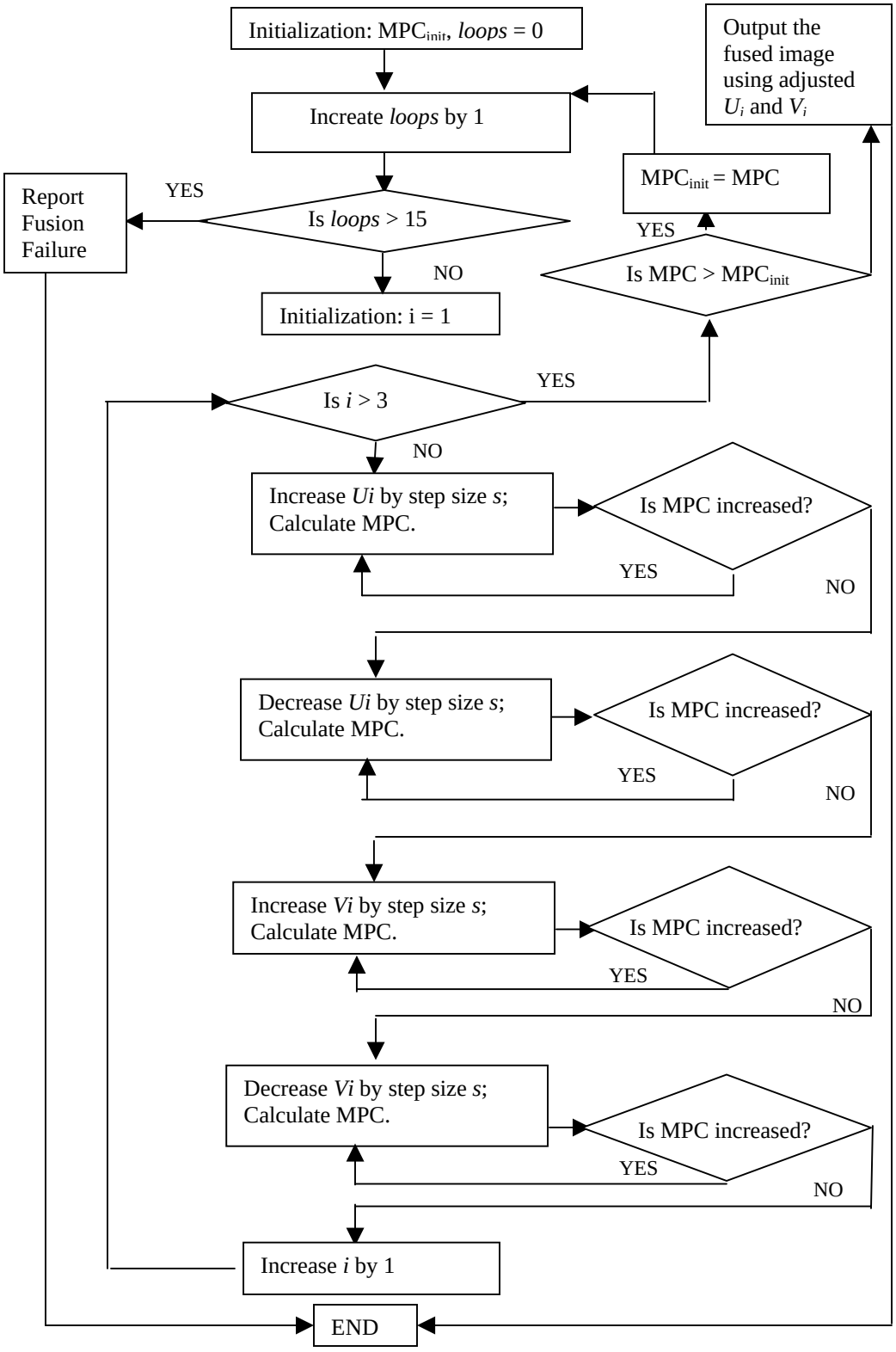


Diagram 5.1: MPC maximization based iteration scheme

Chapter 6

Conclusion and Future Research Direction

6.1 Conclusion

Multi-modality analysis has been emerging as a major trend in the remote sensing, computer visualization, and biomedical image fusion. Fusing biomedical images is a very challenging problem because of the possible vast content change and non-uniform distributed intensities of the involved images. Many methods are still considered barred from meaningful clinical application by the fact that they are as yet improperly validated [17].

This study has made two new and unique contributions to the biomedical image registration and fusion areas in terms of novelty, efficiency, and accuracy. The first contribution is the new automated control point detection algorithm for multi-sensor retinal images. The new method is feature-based which automatically extracts retina vasculature and identifies bifurcation points. 2D affine model is applied that incorporates transformations due to translation, rotation, scaling, and shearing of the subject images. The algorithm identifies control points using adaptive exploratory algorithm by locating the global direction change pixels. The evaluation study with the Centerline Control Point Selection Algorithm shows the advantage of the new Exploratory Adaptive Control Point Selection Algorithm.

The second contribution is the heuristic optimization algorithm for generating the optimal fusion image. In order to maximize MPC of the comparing regions, the iteration algorithm adjusts the initially selected control points at the sub-pixel level. A refinement of the parameter set is obtained at the end of each loop. The iteration stops either when MPC reaches the

maximum value, or when the maximum allowable loop count is reached. A global maxima equivalent fusion result is achieved with calculating local maxima of MPC at an efficient computation cost. The implementation has been successfully tested on two monkeys' retinal image pairs. The comparison with the Root-Mean-Square-Error Minimization (RMSE) objective function proves the accuracy of the Mutual-Pixel-Count maximization objective function. The genetic algorithm is compared with the new heuristic optimization algorithm and the program running results shows that the heuristic algorithm has the advantage in terms of efficiency and accuracy.

6.2 Future Work Direction

There are a couple of relevant issues that need to be addressed. The first task in the plan is to implement the methods in a larger experimental data of the control subjects. Significantly increased volume data could complicate the accuracy evaluation. At this moment, it is difficult to find large quantity of monkey retinal image materials. The number of Cynomolgus monkeys housed at LSU New Orleans Eye Center is very limited. When the physicians did the oxygen analysis experiment, only five monkeys were anesthetized and taken retinal images. Among them only two monkeys have both of the angiogram and fundus images available. Hopefully, this problem would be solved when more people and institutions are involved in the creation of the image materials.

Secondly, many biomedical registration and fusion methods are still primarily used for research activity [45]. Very few of them have been developed into the integrated user-friendly computer software. The eventual aim is to developing and distributing an advanced and easy-to-use application which is suitable for various clinical environments. This plan requires intensive user interface developing work, which allows users adjusting a few threshold parameters if

necessary. The user interface must be stable, simple and informative for every day clinical routine, because most of the end-users do not have much knowledge about algorithms and computer programs.

Working closely with clinicians is extremely important when applying the new methods to practical clinic applications. As computation speed has been dramatically increased, real-time live ophthalmic image processing [66] will be used to handle larger and larger volumes of data in short periods. The data transmission rate, image size, higher resolution pixels, and many other issues will inevitably stress such live imaging fusion systems. The algorithms presented in this thesis have potential ability to handle those challenges. The presented method is a promising step towards useful clinical tools for retinopathy diagnosis, and thus forms a good foundation for further development.

Bibliography

[1] F. Maes; A. Collignon; D. Vandermeulen; G. Marchal; P. Suetens; "Multimodality image registration by maximization of mutual information"; *IEEE Transactions on Medical Imaging*; 1997, V 16, N2, P 187-198.

[2] T. Chanwimaluang T, GL. Fan, SR. Fransen, "Hybrid retinal image registration", *IEEE Transactions on Information and Technology in Biomedcine*; 2006, V 10, N1, P 129-142.

[3] N. Ritter; R. Owens; J. Cooper; RH. Eikelboom; PP. Saarloos; "Registration of stereo and temporal images of the retina**";** *IEEE Transactions on Medical Imaging*; 1999, V 18, N5, P 404-418

[4] CV. Stewart; CL. Tsai; B. Roysam; "The dual-bootstrap iterative closest point algorithm with application to retinal image registration"; *IEEE Transactions on Medical Imaging*; 2003, V 22, N11, P 1379-1394

[5] Y. Zhou; H. Xue; M. Wan; "Inverse image alignment method for image mosaicing and video stabilization in fundus indocyanine green angiography under confocal scanning laser ophthalmoscope"; *Computerized Medical Imaging and Graphics*; 2003, V 27, N6, P 513-523.

[6] I. Vajda; "Theory of statistical inference and information"; Dordrecht. The Netherlands, Kluwer, 1989

[7] N. R. Pal; S. K. Pal; "Entropic thresholding"; *Signal Process*; V 16, P 97-108, 1989

[8] V. Noblet; C. Heinrich; F. Heitz; J. Armspach; "3D deformable image registration: a topology preservation scheme based on hierarchical deformation models and interval analysis optimization"; *IEEE Transactions on Image Processing*; 14 (5) MAY; P 553-566, 2005

[9] T.Netsch; P.Rosch; A. Muiswinkel; J. Weese; "Towards real-time multi-modality 3D medical image registration"; *Proceedings of the Eighth International Conference On Computer Vision*, P 718-725, 2001

[10] H.Chen; Parmod K.; Varshney. "A pyramid approcach for multi modality image registration based on mutual information"; *Proceedings of 3rd International Conference on Information Fusion*; P 9 –15, 2000

[11] T. Daubos; F. Murtagh; "Hight-Quality still images from video frame sequences"; *Proceedings of SPIE*; VOL. 4709, P 49-59, 2002

[12] M.Zibaeifard; M. Rahmati; “An improved multi-stage method for medical image registration based on mutual information”; *Department of Computer Engineering and IT; Amirkabir University of Technology, Tehran, Iran*.

[13] F. Laliberte; L. Gagnon; Y. Sheng; “Registration and fusion of retinal images: a comparative study”; *International Conference on Pattern Recognition* 2002, QUEBEC, AUGUST 11-15, P 715-718, 2002

[14] F. Laliberte; L. Gagnon; “Registration and fusion of retinal images - An evaluation study”**;** *IEEE Transactions on Medical Imaging*; 2003, V 22, N5, P 661-673

[15] GK. Matsopoulos; PA. Asvestas; NA. Mouravliansky; KK. Delibasis; “Multimodal registration of retinal images using self organizing maps”**;** *IEEE Transactions on Medical Imaging;* 2004, V 23, N12, P 1557-1563

[16] A. Averbuch; Y. Keller; “Robust Multi-Sensor Image Registration Using Pixel Migration”; *IEEE Sensor Array and Multichannel Signal Processing Workshop*, 2002, P 100-104

[17] J. B. Antoine Maintz; Max A. Viergever; “An Overview of Medical Image Registration Methods”; *Imaging Science Department, Imaging Center Utrecht*, 1996.

[18] Z. Millwala; “A dual-state approach to dental image registration”; Masters Thesis; *Lane Department of Computer Science and Electrical Engineering; West Virginia University*; 2004

[19] J. Zhang; “New information theoretic distance measures and algorithms for multimodality image registration”; Ph.D. Dissertation; *University of Florida*; 2005

[20] Electrical Engineering and Computer Science Department, Lehigh University, Bethlehem, PA 18015, http://www.ece.lehigh.edu/SPCRL/IF/image_fusion.htm

[21] M. F. Georgiou; “Intra- and multi-modality registration and fusion of body images in radiology”; Ph.D. Dissertation; *University of Miami*; 1997

[22] Pelizzari CA, Chen GTY, Spelbring DR,Weichselbaum RR, Chen CT; “Accurate three-dimensional registration of CT, PET, and/or MR images of the brain”; *J Comput Assist Tomogr*; V 13 P 20–26; 1989

[23] Turkington TG, Hoffman JM, Jaszczak RJ, MacFall JR, Harris CC, Kilts CD, Pelizzari CA, Coleman RE; “Accuracy of surface fit registration for PET and MR brain images using full and incomplete brain surfaces”; *J Comput Assist Tomogr*; V19 P 117–124; 1995

[24] Royal Perth Hospital Website: Division of Imaging Services – *Nuclear Medicine, 6 Selby Street Shenton Park, Western Australia*, 6008 http://www.rph.wa.gov.au.

[25] Duke University; Department of Medical Physics; Duke Advanced Imaging Labs 2424 Erwin Rd Suite 302, Durham, NC 27705. http://medicalphysics.duke.edu.

[26] Z. Lu, Y. Xu, X. Yang, S. Li, L. Traversoni; "2D Quaternion Fourier Transform: The Spectrum Properties and its Application in Color Image Registration"; *Multimedia and Expo, 2007 IEEE International Conference*; P 1715-1718; July 2007

[27] Michael Abràmoff, MD, PhD. Assistant Professor of Ophthalmology Department of Ophthalmology and Visual Sciences University of Iowa Hospitals and Clinics, Iowa USA. http://bij.isi.uu.nl/.

[28] S. Klein, M. Staring, J.P.W. Pluim, "Evaluation of Optimisation Methods for Nonrigid Medical Image Registration using Mutual Information and B-splines", *IEEE Transactions on Image Processing* (in press).

[29] Mohamed Ibrahim Elbakary; "Novel Pixel-Level And Subpixel-level Registration Algorithms for Multi-Modal Imagery Data", PhD Dissertation, *Department of Electrical and Engineering, The University of Arizona*, 2005.

[30] Khoobehi B, Beach J, Kawano H. "Hyperspectral imaging for measurement of oxygen saturation in the optic nerve head". *Invest Ophthalmol Vis Sci*. V 45 P1464–1472. 2004

[31] H, Li; "Multi-Sensor Image Registration and Fusion (Image Fusion)", PhD Dissertation; *University of California, Santa Barbara*; 1993

[32] N, Otsu; "A Threshold Selection Method from Gray-Level Histograms"; *IEEE Transactions on Systems, Man, and Cybernetics*; V9, P 62-66; 1979.

[33] J.F. Canny; "A computational approach to edge detection"; *IEEE Transactions on Pattern Analysis and Machine Intelligence;* V8, P 679-698; Nov. 1986.

[34] William E. Green; "Canny Edge Detector"; Department of Mechanical Engineering and Mechanics, *Drexel University, 3141 Chestnut Street, Room 2-115, Philadelphia, PA* 19104; 2002;

[35] Sahoo, P.K.; Farag, A.A.; Yeap, Y.-P; "Threshold Selection Based on Histogram Modeling"; *IEEE Transactions on Systems, Man and Cybernetics*; V1, P 351 – 356, Oct 1992

[36] J. Beach, J. Ning, B. Khoobehi; "Oxygen Saturation in Optic Neve Head Structures by Hyperspectral Image Analysis"; *Current Eye Research*; V32, P161-170, 2007

[37] B. Khoobehi, J. Ning, E. Puissegur; "Retinal Oxygen Saturation Evaluation by Multispectral Fundus Imaging"; Department of Ophthalmology, LSU Health Sciences Center; Department of Biomedical Engineering, Tulan University; Department of Computer of Computer Science, Louisiana State University; Jan, 2007; Submitted.

[38] Ali Can, Hong Shen, Turner, J.N. Tanenbaum, H.L. Roysam, B.;"Rapid automated tracing and feature extraction from retinal fundus images using direct exploratory algorithms"; *IEEE Transactions on Information Technology in Biomedicine*; V3, P125-138, June 1999

[39] Gonzalesz RC, Woods RE. “Digital Image Processing”; *Addison-Wesley Publishing Company*; 2002.

[40] Marc Lalondey, Langis Gagnony; "Non-recursive paired tracking for vessel extraction from retinal images"; *Proc. vision Interface*; P61-68, 2000

[41] Z. Ye, H. Mohamadian, Y. Ye; “Information Measures for Biometric Identification via 2D Discrete Wavelet Transform”; *Proceedings of the 2007 IEEE International Conference on Automation Science and Engineering*; Scottsdale, Arizona, USA; September 22-25, 2007

[42] Z. Ye, J. Luo, P. Bhattacharya; “Optimal Approach for Enhancement of Large and Small Scale Near-Infrared and Infrared Imaging”; *Proceedings of the 2005 IEEE International Conference on Industrial Electronics and Control Applications*; Nov.29 - Dec. 2, 2005, Quito, Ecuador

[43] Z. Ye; “Artificial Intelligence Approach for Biomedical Sample Characterization Using Raman Spectroscopy”; *IEEE Transactions on Automation Science and Engineering*; V2, P67-73; Jan. 2005

[44] Z. Ye, Y. Ye, H. Mohamadian; “Biometric identification via PCA and ICA based pattern recognition”; *IEEE International Conference on Control and Automation;* Guangzhou, China - May 30 to June 1, 2007

[45] R. Lundqvist; “Atlas-based fusion of medical brain images: methods and applications”; PhD dissertation; Uppsala University (Sweden); 2001

[46] M, Xia; “Issues in video/image communication”; PhD dissertation; *Princeton University*; 2004

[47] B. Ma; “Parametric and nonparametric approaches for multisensor data fusion”; PhD dissertation; *University of Michigan*; 2001

[48] B. Sadeghi, B. Moshiri; "Second-order EKF and Unscented Kalman Filter Fusion for Tracking Maneuvering Targets"; *IRI 2007 IEEE International Conference*; V13, P514- 519; Aug. 2007

[49] D. P. Spanos, R. Olfati-Saber, R. M. Murray; “Approximate distributed Kalman filtering in sensor networks with quantifiable performance”; *Fourth International Symposium on Information Processing in Sensor Networks*; P133–139; April 2005

[50] A. Anandkumar, L. Tong, and A. Swami; “Energy-efficient data fusion in Gauss-Markov random field,” *Tech. Rep. ACSP;* TR 01-07-02; Jan.2007

[51] K.C.Chou, A.S.Willsky, A.Benveniste; "Multiscale recursive estimation, data fusion, and regularization"; *IEEE Transactions on Automatic Control*; V39 P464 - 478; 1994

[52] H. Rue, L. Held; “Gaussian Markov Random Fields: Theory and Applications”; *Monographson Statistics and Applied Probability*; Chapman and Hall, London; V104; 2005

[53] B. Ma; S. Lakshmanan; A. O. Hero; "Simultaneous detection of lane and pavement boundaries using model-based multisensor fusion"; *IEEE Transactions on Intelligent Transportation Systems*;V01, P135-147; Sep. 2000

[54] D.E.Becker, A Can, J.N.Turner, H.L.Tanenbaum; "Image processing algorithms for retinal montage synthesis, mapping, and real-time location determination"; *IEEE Transactions on Biomedical Engineering*; V45, P105-118; Jan 1998

[55] A.Banumathi, R.K.Devi, Raju; "Performance analysis of matched filter techniques for automated detection of blood vessels in retinal images" *TENCON 2003. Conference on Convergent Technologies for Asia-Pacific Region* V2, P543-546; October 2003

[56] S.Chaudhuri, S.Chatterjee, N. Katz; "Detection of blood vessels in retinal images using two-dimensional matched filters"; *IEEE Transactions on Medical Imaging*; V8, P263 - 269; September 1989

[57] R. Poli, G. Valli; "An algorithm for real-time vessel enhancement and detection"; *Computer Methods and Programs in Biomedicine*; V52, P1-22; Nov.1996

[58] Image source: http://world.mongabay.com/romanian/travel/files/p6146p.html

[59] Tsukuba Primate Research Center, National Institute of Biomedical Innovation Hachimandai 1-1, Tsukuba-shi, Ibaraki, Japan http://tprc.nibio.go.jp/eng/index_e.html

[60] Z. Ye, H. Cao, S. Iyengar, H. Mohamadian; Chapter 6 - "Practical Approaches on Medical and Biometric System Identification for Pattern Recognition, Data Fusion and Information Measuring"; *Systems Engineering Approach to Medical Automation*; 2008

[61] GK. Matsopoulos; NA. Mouravliansky; KK. Delibasis; “Automatic retinal image registration scheme using global optimization techniques”; *IEEE Transactions on Information Technology in Biomedicine*; V3, P47 - 60; Mar 1999

[62] T. Choe, I.Cohen, L. Munwai; "Optimal Global Mosaic Generation from Retinal Images"; *18th International Conference on Pattern Recognition*; V3, P681 - 684; 2006

[63] M. Balasubramanian, S. Zabic, L. Zangwill1, H. Thompson, P. Wolenski, S. Iyengar; "A Framework for Detecting Glaucomatous Progression in the Optic 2 Nerve Head of an Eye using Proper Orthogonal Decomposition"; *IEEE Transactions on Information Technology in Biomedicine* (Submitted)

[64] Y.Zhu; "A Java program for stereo retinal image visualization"; *Computer Methods and Programs in Biomedicine*; V85, P214-219; 2007

[65] Y.Zhu; “Mutual information-based registration of temporal and stereo retinal images using constrained optimization”; *Computer Methods and Programs in Biomedicine*, V86, P210-215; June 2007

[66] M. Balasubramanian, et al.; “Real-time restoration of white-light confocal microscope optical sections,” *Electronic Imaging;* V 16, N3; 2007

[67] Howard Glicksman M. D. www.arn.org/docs/glicksman/eyw_041001.htm

[68] Sameer Antania; D.J. Leeb; L. Rodney Longa; George R. Thoma; "Evaluation of Shape Similarity Measurement Methods for Spine X-ray Images" *J. Vis. Commun. Image;* R. 15, P 285–302, 2004.

[69] D. Doldberg; “Genetic Algorithms in Optimization”; *Search and Machine Learning*; *Reading*, MA; Addison-Wesley, 1989.

Appendix: Tunable Threshold Parameters

Table A.1: Tunable threshold parameters

Parameter	Value
Comparing Edge	125
BW Thresh1	0.06
BW Thresh 2	0.14
Vein Extraction Thresh1	0.15
Vein Extraction Thresh2	0.14
Black Area Percentage	50 %
Max Steps Gray Image	60
Max Steps Color Image	30
Roll Back Steps	3

Vita

Hua Cao was born at Luzhou, China, in 1978. She is currently a doctoral student in the Computer Science Department at Louisiana State University. She received the Bachelor of Engineering degree in Information Systems and Management from the Southwest University of Finance & Economics, China, in 2000. Her master's degree is in systems science from the Computer Science Department at Louisiana State University in 2003. She has been working as a research assistant at Robotics Laboratory of Computer Science Department, LSU, since 2002. Her research interest includes biomedical imaging, data registration, image fusion, and large grid route planner.