OCULAR FUNDUS IMAGE SEGMENTATION FOR EARLY DETECTION OF DIABETIC RETINOPATHY

Dhanesh Ramachandram, Mandava Rajeswari,

Lee Seng Soon and Bakiah Shaharudin

UNIVERSITI SAINS MALAYSIA

OCULAR FUNDUS IMAGE SEGMENTATION FOR EARLY DETECTION OF DIABETIC RETINOPATHY

by

Dhanesh Ramachandram, Mandava Rajeswari, Lee Seng Soon and Bakiah Shaharudin

USM Short Term Research Grant Final Report

May 2006

ACKNOWLEDGEMENTS

This research has been made possible through the Universiti Sains Malaysia Short Term Research Grant. We would like to thank our consultant ophthalmologist, Dr Bakiah, from the Health Campus of University Sains Malaysia, for unselfishly rendering her expertise and valuable time for this research. We would also like to thank Dr Yunus and Dr.Elias Hussein for providing the colour fundus images for our analysis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

USM Universiti Sains Malaysia

WHO World Health Organisation

DR Diabetic Retinopathy

IRM Interetinal Microvascular Abnormalities

NVD Neovascularization at the disc

NVE Neovascularization elsewhere

GMRF Gaussian Markov Random Field

MRF Matched Filter Response

GVF Gradient Vector Field

OCULAR FUNDUS IMAGE SEGMENTATION FOR EARLY DETECTION OF DIABETIC RETINOPATHY

ABSTRACT

This work presents an automated fundus image analysis system for the detection of optic disc. The analysis proceeds in three distinct stages, optic disc localisation stage, a pre-processing stage to reduce noise and blood vessel structures, and finally an optic disc segmentation stage. The optic disc localisation results in a localised point that represents the centroid of optic disc region, whereas optic disc segmentation results in a complete contour of optic disc. In the optic disc localisation stage, a feature vector approach that employs four salient characteristics of the optic disc is implemented. These features include colour, convergence of blood vessels, the size of the optic disc and its shape. In order to obtain accurate optic disc localisation, an unsupervised colour thresholding approach is first applied to the image. The resulting clusters are then weighted according to its similarity to known optic disc characteristics. This knowledge about typical characteristics is obtained by manual observation of available fundus images, consultation from ophthalmologist and past research papers. Since there may be regions within a fundus image that closely resemble an optic disc, the true optic disc is chosen based on some weighting criteria and subsequent selection of the region with the highest weight. This indicates highest similarity to the known optic disc characteristics. The centroid of the highest-weighted region is then marked as the optic disc localisation point.

In order to validate the approach, a set of 81 colour fundus images were used to compare the obtained localisation point with the ground truth, a point which is manually annotated by qualified ophthalmologists. The proposed approach resulted in a success rate of 97.53% as opposed to 88.88% using a previously reported fuzzy convergence method.

Prior to optic disc segmentation, a non-linear selective filtering technique, known as adaptive anisotropic diffusion is implemented to reduce noise (unwanted data). The adaptive anisotropic diffusion utilises information at multiple scales to assign the appropriate amount of smoothing to edges at different scales so that strong as well as weak edges are well preserved. A comparative analysis of the performance of adaptive diffusion, anisotropic diffusion and coherence enhancing diffusion is made with respect to various levels of noise corruption. In addition to reducing noise in the fundus image, the adaptive anisotropic diffusion is also combined with a morphological operator to remove blood vessels that lie within the optic disc region. This creates a homogeneous region within the optic disc that facilitates the convergence of deformable contour to actual optic disc boundary.

The second stage in the analysis is the optic disc contour segmentation. In this research, an automatic initialisation method for deformable contours is introduced. Specifically, cubic spline interpolation is employed to deal with outlier edge pixels. The cubic spline interpolation re-positions these outliers such that the repositioned outlier pixels conform to the shape of the optic disc. Finally, the retained and re-positioned edges are used to construct a coarse contour that approximates the actual optic disc. The constructed contour acts as initial configuration for a deformable contour. This deformable contour is then deformed to match the actual boundary of the optic disc using fewer iterations and lower possibility of wrong convergence.

CHAPTER 1

INTRODUCTION

1.1 Introduction

Diabetic Retinopathy or simply referred as Retinopathy is a diabetic-related disease of retina of human eye. It is the complication of diabetes, which in turn is caused by low levels of insulin secreted by the pancreas or the consequence of a unhealthy diet consisting of high amounts of sugar. Retinopathy is caused by damaged blood vessels (arteries, veins and cappilaries) that have leaked, blocked or developed abnormally within the retina. Diabetic retinopathy may cause temporary or permanent visual impairment and irreversible blindness if it reaches a severe stage.

Based on the estimation by the World Health Organization (WHO), retinopathy is ranked as sixth significant (4.8%) cause of visual disability (*World Health Organization, Magnitude and Causes of Visual Impairment, Fact Sheet of World Health Organization*, 2004). In the United Kingdom alone, 12% of diabetes patients have been diagnosed with retinopathy (Kanski, 1997). In the United States, retinopathy has affected over 5.3 million citizens aged between 18 and older or 2.5% of the entire Americal population (*Vision Problems in U.S: Prevalence of Adult Vision Impairment and Age Related Eye Diseases in America: Prevent Blindness in America*, 2002). Various studies on frequency of retinopathy among diabetic patients have been conducted. For instance, one study on the frequency of prevalence of retinopathy at a diabetes center in Southern India shows that the prevalence of retinopathy is 34.1% or equivalent to 2319 patients out of 6792 (Kayani et al., 2003).

An immediate and effective strategy to counter the negative effects of Diabetic Retinopathy is to perform screening at an early stage. Screening is usually offered to people of selected population. They are asked questions or offered simple tests in order to identify the subjects who are more likely

to receive further test or treatment to reduce the risk of developing severe complication *UK National Screening Committee,http://www.nsc.nhs.uk/whatscreening/whatscreen_ind.htm* (2005). In other words, screening is a quick and simple test applies to a large number of people. It is the anticipation identification of subjects who might benefit from further testing and appropriate treatment, but it does not make the final diagnosis. Such a screening programme has been successfully implemented and it has became the basic model taken in the UK to develop national strategies for diabetic retinopathy screening (Wormald,1999).

1.2 Screening Approaches

There are two common screening approaches that have been applied to public screening programmes. They are the ophthalmoscope and fundus photograph. Both modalities are discussed in the following.

An *ophthalmoscope* is a lighted instrument that projects the light rays into the eye to examine the anatomical and pathological structure of the eye. Direct ophthalmoscope refers to the hand held instrument with a battery powered light source. It contains a series of lens to obtain a sharp view of the centre of the retina. Indirect ophthalmoscope is mounted on the head of doctor. It allows the doctor to view the entire retina by adjusting a set of magnifying glasses. Direct ophthalmoscope and indirect ophthalmoscope are illustrated in Figure 1.1. In order to obtain a good view of both the periphery and centre of the retina, the subject is placed in a dark room in order to dilate the pupils. This may turn to be an inconvenience to subjects and requires cooperative subjects. Ophthalmoscopy is less sensitive and less specific than fundus photography. There is a high probability of missing significant signs of retinopathy and greater risk of misdiagnosis of retinopathy. Opthalmoscope also does not enable the doctor to store the retinal record. The quality assurances and later assessments are even difficult to perform (Osareh, 2004).

The fundus photograph is the picture of inner part of eye, captured using fundus camera in the format of either digital or film-based photograph. The view of eye is captured directly via pupil as both the entrance point for camera's illuminating and exit point for imaging light rays. The

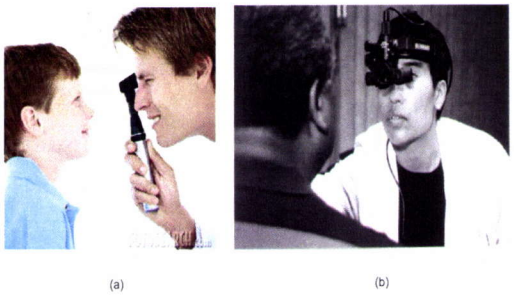


Figure 1.1: (a) Direct Ophthalmoscope (b) Indirect Ophthalmoscope

Figure 1.2: The Fundus Camera

patients sit in front of the camera with the chin rest on the platform and the forehead against the bar. After the lens of the camera are focused and aligned appropriately, a flash fired and a colour fundus photograph is taken. The illustration of this scenario is given in Figure 1.2 and the colour fundus image is illustrated in Figure 1.3.

Another format of fundus photograph is fluorescent angiograms photograph. In this method, a dye called Sodium Fluorescent liquid is injected into the patient's eye before the fundus angiogram photograph is taken using fundus camera. The dye that passes through the vessels allows blood vessels and microaneurysms to appear brighter than other parts of the retina. Here, haemorrhages

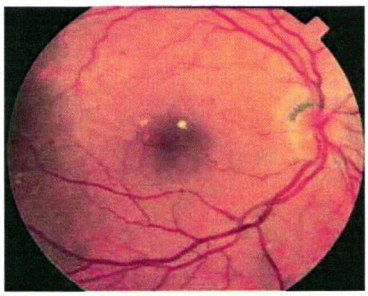


Figure 1.3: Colour Fundus Photograph

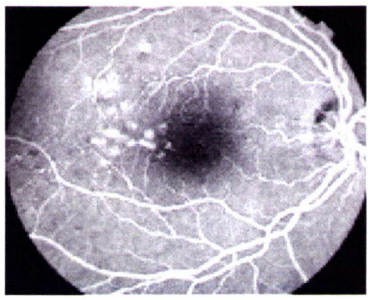


Figure 1.4: Fluorescent Angiogram Photograph

appear as darker regions. An illustration of fluorescent angiogram photograph is given in Figure 1.4

With the rapid development of computer technology, various novel image processing and computer vision approaches for screening with digital photograph are becoming available. Digital fundus camera are easily available, cheap to operate and portable - this translates into cost saving if these cameras are used as the screening instrument in a community screening programme. Digital fundus images can be compressed and saved into databases for later analysis or assessment. The efficient and effective assessment of fundus images is possible by combining the technology of telemedicine and computer vision methods. The captured fundus images can be analyzed instantly and automatically by the computer at the site of acquisition or sent to centralized databases via electronic mail for later assessment by opthalomologsits. If the fundus images are processed by the computer, this would eliminate the need of primary assessments by ophthalmologists. Only the patients who are true positives are required to meet an ophthalmologist for a more comprehensive examination. And at the same time, mass screening of large digital fundus image databases can be automated for faster and more consistent preliminary screening using a computerized screening programme. Considering the various advantages provided by the analysis of the digital fundus image, the work reported here is geared towards achieving a reliable and repeatable automated screening of fundus images for the early diagnosis of Diabetic Retinopathy.

1.2.1 Typical Computer Aided Detection and Diagnosis System

Generally, the input for a computer-aided detection and diagnosis system is the colour fundus image. The digital image is usually pre-processed to improve the image contrast, image quality and to remove unwanted data. One of the anatomical features to be detected first is the optic disc. With the location of optic disc as priori knowledge, the boundary of optic disc is then segmented. The priori knowledge of optic disc is also used to search for other important anatomical features. Following that, the existence of pathologies features are searched as well in fundus image. The number, the size and the spatial distribution of each particular pathologies features with respect to anatomical features in fundus image are taken as important information for accurate analysis and reliable assessment of diabetic retinopathy. The major steps in the analysis of fundus image are:

- Locate the arbitrary location of optic disc. The optic disc may be found at the center or at the side of the fundus image. Employ segmentation methods at the vicinity of optic disc to extract out the complete contour of optic disc. The optic disc feature is marked with red circle outline, as shown in Figure 1.5.
- Locate macula and fovea. Both macula and fovea may be found at the right or left side of optic disc. It is marked with blue circle outline.
- Track the blood vessels structures and then extract out the complete vasculature of blood vessels in fundus image. Identify the bifurcation and crossover point of blood vessels and the curvature of blood vessels. This feature is shown by yellow circle in Figure 1.5.
- Locate the location, the number, the size and spatial distribution of pathological features in fundus image. The pathological feature is circled with green outline, as in Figure 1.5.

The optic disc segmentation is the primary step in fundus image analysis. It serves as a landmark for subsequent detection and analysis of the anatomical and pathological features. With the location of optic disc as reference point, the macula and fovea is detected approximately two optic disc-diameters away from the reference point. The detection of both fovea and macula is

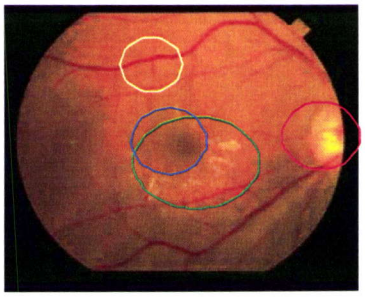


Figure 1.5: The Fundus Image

important as they contain a high concentration of photoreceptors which are responsible for daytime vision. The existence of any pathological features within the fovea and macula regions are taken as the criterion to assess the severity of retinopathy. The blood vessels structures are also one of the significant features for evaluation of retinopathy. The curvature of blood vessels, especially the major blood vessels and the growth of tiny and thin vessels are the symptoms of retinopathy. Therefore, the accurate detection of blood vessels and tortuosity of the blood vessels are necessary for computer-aided retinaopathy diagnosis. The spatial distribution of pathological features within the anatomical features can be used as a criterion for a more accurate assessment of the various stages of retinopathy.

This research work focuses on the optic disc detection and segmentation. As mentioned earlier, the optic disc detection and segmentation is the very first task to be performed in most of the fundus image analysis procedures. The accuracy of subsequent analysis relies heavily on the accuracy of the optic disc segmentation. The following briefly describes the importance of optic disc segmentation.

- Macula Localisation Diameter of optic disc is used as priori knowledge for locating the macula region (Osareh,2004) and fovea accurately (Lowell et al.,2004). The existence of lesions at the vicinity of macula helps to establish a measurement metric regarding the severity of retinopathy.

- Exudates Detection Owing to similar yellow colour appearance of optic disc and the exudates, misclassification of optic disc as exudates can be avoided by detecting the optic disc first and masking it out from the rest of yellow regions which may then be classified as exudates

(Osareh,2004).

- Starting Point For Blood Vessels Tracking Optic disc region also serves as the seed or starting point for blood vessels tracking (Tolias & Panas,1997) and (Osareh,2004). This is owing to the nature characteristics of blood vessels as most of the major blood vessels originated from the optic disc region. The tracking of blood vessels structures is of importance in retinal image registration, change detection and mosaic image synthesis (K.Fritzsche et al.,2002).

1.3 Motivation

Owing to the rapid development of image processing and computer vision research, a wide range of methods are becoming available for computer-aided screening system of retinopathy. It has become a common practice to pre-process and to segment the object of interest prior to further processing and analysis on the images. In a computer-aided screening system, the captured image can be processed and analyzed automatically with fairly consistent performance in terms of accuracy.

In a computer-aided screening system, *optic disc* is of critical importance in fundus image analysis and is a pre-requisite for any subsequent analysis tasks. In detecting the optic disc, an important aspect to be considered is the robustness of locating the center point of optic disc region considering the diverse manner in which the optic disc may appear in fundus images. Many published works on optic disc detection utilise single or multiple optic disc characteristics, such as the convergence point of blood vessels, region intensity, intensity fluctuation, region contrast, shape and colour (T. Teng & Claremont, 2001). However, these methods often fall short of its reliability since there is high variability in the location, size, shape and colour of the optic disc in a typical collection of fundus images.

This research proposes a novel method of utilizing multiple characteristics of the disc to detect the actual boundary of the optic disc using an ensemble of clustering technique, multiple scale edge informaton, shape analysis, colour measure, thresholding, morphological operations and non-linear filtering. The accuracy and complete boundary of optic disc is improved using a deformable contour method. The proposed algorithm is very robust considering the varying appearances of

the optic disc in real fundus image databases.

1.4 Research Scopes and Objectives

As mentioned earlier, this research proposes a robust optic disc detection system. The objectives of the research include:

- To propose and develop an accurate and reliable optic disc algorithm to locate the optic disc.
- To develop an algorithm to extract the accurate contour of optic disc in colour fundus image.
- To evaluate the proposed method through a set of fundus images.
- To compare the performance of proposed method with existing methods.

1.5 Overview of Approach

The input for the optic disc detection system, which is colour fundus image, is analysed automatically and the optic disc region is localised by locating the center point of the optic disc. This facilitates the subsequent optic disc segmentation process. Once the optic disc segmentation is complete, further analysis such as exudates detection and recognition, macula localisation and blood vessel tracking may be performed. The overall scheme of the approach presented in this research is shown in Figure 4.1. This process is described in the following.

Prior to selecting the most probable optic disc region, a number of candidate regions are identified in a colour fundus image. The determination of the candidate regions utilises an unsupervised colour thresholding technique. Following this, a feature vector approach is implemented to detect the optic disc. The feature vector approach uses multiple features instead of a single feature and thus increasing the robustness of optic disc detection. The features employed are the four salient characteristics of the optic disc: *colour, size, shape* and *convergence point of major blood* as the criteria for selecting the most probable optic disc region.

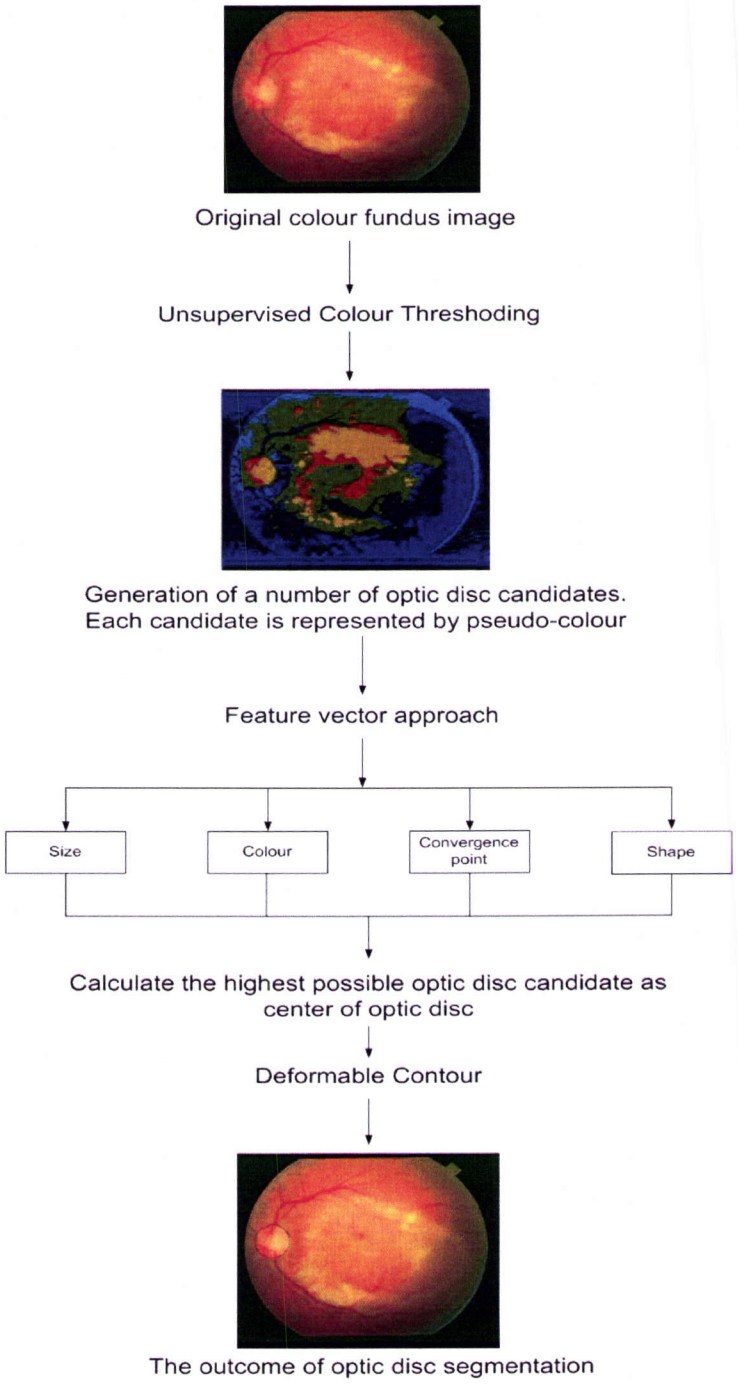


Figure 1.6: The overview of the proposed optic disc detection system

Each candidate region is then further processed to compute the selected features. These features are compared against the reference characteristics of the typical optic disc and weighted in accordance to their similarity to this reference optic disc. A single region with the highest weight is selected. Its centroid is calculated as the center of optic disc.

The next step is to segment the optic disc. Localizing the optic disc aids in initializing the deformable contour at the vicinity of the optic disc. The closed contour of the optic disc can be obtained using a Gradient Vector Flow (GVF) snake technique.

1.6 Contribution of the Research

This research proposes a novel method to accurately extract the actual boundary optic disc. This achieved through the use of an ensemble of techniques: multiple scale edge informaton, shape analysis, colour measure, thresholding, morphological operations, non-linear filtering, spline interpolation and deformable contour. The main contribution of this research is the capability to detect a wide range of appearances of optic disc at arbitrary location in colour fundus image.

1.7 Report Outline

The report is structured according to the tasks involved in the proposed automated optic disc detection approach.

Chapter 2 gives the necessary medical background which covers the anatomical structures of human eye and retinopathy description. This chapter addresses the complication of retinopathy to visual impairment, prevalence, implications of retinopathy, the importance of screening, the issues regarding optic disc detection, which covers the importances of optic disc, complication that difficult the task of optic disc detection and dominant characteristics of optic disc.

Chapter 3 presents a concise review of the pertinent literature in the areas of optic disc localisation and optic disc contour segmentation. We critically discuss various implementations

that utilise several optic disc characteristics for optic disc localisation. This chapter also continues with detailed description of the relevant work in optic disc contour segmentation.

Chapter 4 describes the theoretical concepts and algorithms of unsupervised colour thresholding, anisotropic diffusion and cubic spline interpolation. These methods are incorporated into optic disc detection algorithms.

Chapter 5 introduces the feature vector-based optic disc localisation algorithm. The initial section of this chapter is devoted to an overview of the proposed algorithm. A description of current algorithm covering the optic disc candidates generation, image pre-processing and measurement of respective optic disc's characteristics with corresponding method follows. In addition, this chapter also introduces a quantitative measurement to measure the performance of optic disc localisation semantically. The performance is measured in term of accuracy of obtained resulting localise point to actual centroid of optic disc by taking account of wide range of optic disc's sizes and shapes.

Chapter 6 presents the automatic positioning of deformable contour at the vicinity of optic disc region. This chapter describes the methodology for the selection of optic disc's edges, initiated at the given resulting center point. An experiment is carried out to validate the performance of obtained optic disc's contour against the groundtruth contour for a database of optic disc images.

The conclusion of the research in terms of advantages of proposed scheme, practical achievement and recommendations for future work in presented in **Chapter 7**.

CHAPTER 2

BACKGROUND

2.1 Introduction

This chapter presents the essentials of the human eye, a concise information on diabetic-related eye diseases and important issues pertaining to the optic disc. Important anatomical structures of human eye are introduced. Section 2.3 presents an overview of the effects of diabetes to human vision. Important a characteristics of the optic disc in most colour fundus images are descibed with relevant illustrations to aid the reader. Finally, challenges that are faced by the computer vision research community are discussed.

2.2 Anatomy of the Human Eye

Human eye is a complicated organ comprised of many anatomical structures working together as a vision instrument to produce images that can be interpreted by the brain. The anatomical structures of the human eye is illustrated in Figure 2.1. Cornea is the very first part of the eye to receive the incoming light rays reflected from the outside world and refracted onto the crystalline lens. In order to focus an image clearly onto the retina, ciliary muscles are used to alter the shape of the crystalline lens. The resulting light rays that fall onto the retina is upside down. Photoreceptor cells are then stimulated to produce electrical impulses which are then transmitted via the optic nerve, which is located at the back of the eyeball, to the brain. The brain then further translates and interprets the image in an upright position (Segal,1998). The whole process of refraction, focusing the light rays onto the retina and the subsequent transmission of the signal to the brain are similar to the way how a camera works. Generally, a film camera needs lens and film to capture the image. The lens are analogical to the human cornea and crystalline lens, while human retina represents the film. Figure 2.2 illustrates the analogy between a camera and the

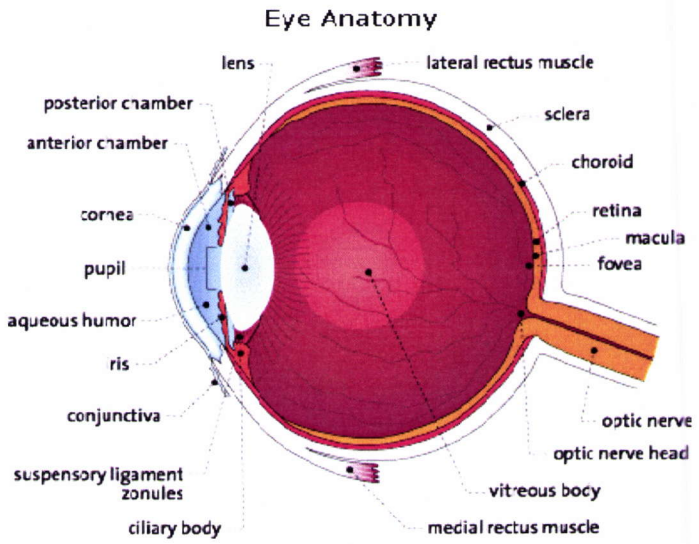


Figure 2.1: The cross Section of typical human eye in 2D

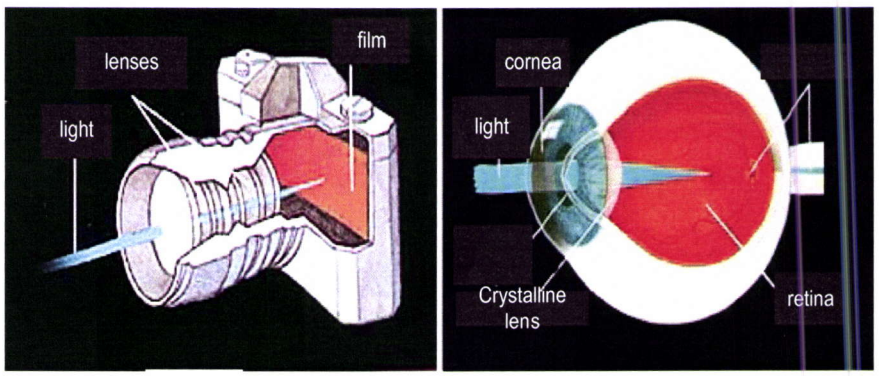


Figure 2.2: (a) the camera (b) the human eye

human eye.

The fundus image is the photograph of the inner layer of the human eye organ, captured using a fundus camera. The view of the eye is captured directly through the pupil. After the lens of the camera are focused and aligned appropriately, a flash is fired and a fundus image is captured. In a fundus image, there are several anatomical structures which are of particular importance to the current research. These prominent structures in the fundus image are described further in the following subsections.

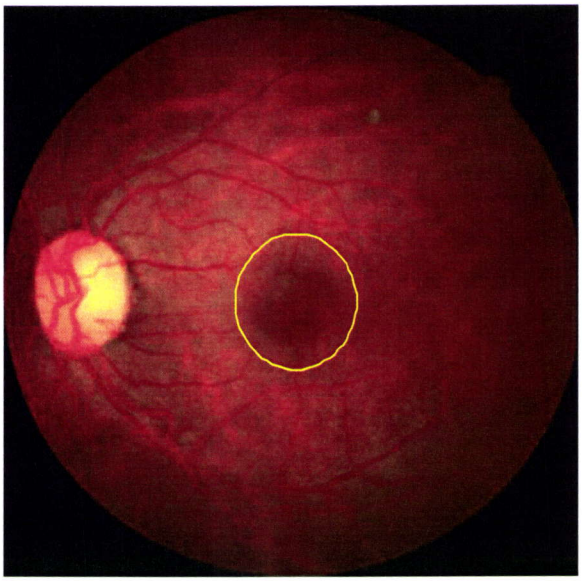


Figure 2.3: The highlighted location of macula in a typical colour fundus image

2.2.1 Vitreous Body

Vitreous body is the thick and transparent substances situated in between lens and retina. It is composed mostly of water and seen as clear jelly. It comprised $\frac{2}{3}$ of eye volume and responsible for maintaining the shape and form of eyeball (*Scott & Christie and Associates: Your EyeSight Is Our Vision*,n.d.). It is also known as vitreous gel and vitreous humor as well. The illustration of the vitreous body can be found in Figure 2.2(b).

2.2.2 Macula

The macula is located at the middle part of the retina and contains a high concentration of photoreceptors. It is responsible for central vision which allows small and detailed objects to be seen. Central vision is also responsible for tasks such as reading. The macula has a typical diameter of about 7mm or $\frac{1}{4}$ inch (*The Eye: Information About Vision Loss And Blindness*,2005). Physically, it appears as a darker circular patch at the optic disc axis and is situated about 2 disc diameter away temporally from the optic disc Goh et al. (2000). In a typical colour fundus image, the location of macula is highlighted as in Figure 2.3.

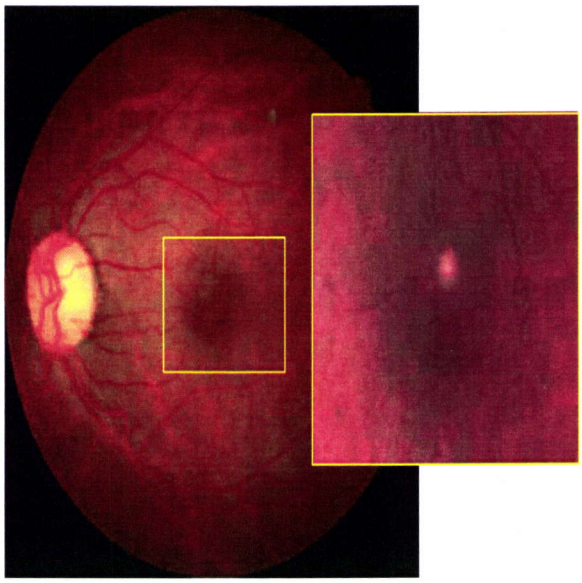


Figure 2.4: The fovea as highlighted in a typical colour fundus image

2.2.3 Fovea

The fovea is an indentation at the center of macula. It contains a high concentration of colour photoreceptors, known as cones which approximately number 6 million (*The Eye: Information About Vision Loss And Blindness*,2005). Cones are found to be sensitive to three different spectrums of wavelengths, i.e. red, green and blue respectively. These photoreceptors are responsible for daytime vision (photopic) and enable better colour vision and highest visual acuity. The typical size of the fovea about 1.5 mm (Osareh,2004)in diameter or approximately $\frac{1}{16}$ inch (*The Eye: Information About Vision Loss And Blindness*,2005). Figure 2.4 highlights the location of fovea in a typical colour fundus image.

2.2.4 Optic Disc

The optic disc is also known as *papilla* or the optic nerve head. The optic disc is the entrance and exit point for blood vessels and the optic nerve (Osareh,2004). Since there are virtually no photoreceptors in this region, any light rays that fall on the optic disc will be not be perceived. For this reason, it is also known as the *blind spot* of human vision. In a typical colour fundus image, the optic disc is observed as yellowish in colour (Walter et al.,2002). It appears relatively brighter

(a) (b)

Figure 2.5: The highlighted optic disc region in a typical colour fundus image. (a) At the posterior (b) At the anterior

than other regions due to the absence of the retina layer which allows much of the illumination to be received. At the posterior, the optic disc ranges from being round to oval in shape, with diameters ranging from 1.5 mm to 1.7 mm, while at the anterior, the optic disc appears as an ellipse (Goh et al.,2000). This is illustrated in Figures 2.5(a) and 2.5(b).

2.2.5 Retina

The retina is a thin sensory membrane which is found at the inner surface of the eye. The sensory cells, known as photoreceptors, that reside within retina consists of cones and rods. Rod cells provide dim and night vision and the ability to sense movement. It is mostly found at the vicinity of macula region. There are about 200 million of rods in the retina (*The Eye: Information About Vision Loss And Blindness*,2005). The underlying part of vision lies at the center part of retina, i.e macula and fovea. These two types of photoreceptors convert the light rays that fall on retina layer into electrical impulses, which are then sent via the optic nerve to the brain for image interpretation and understanding (Dan,2001). Figure 2.6 illustrates the region of retina.

2.2.6 Blood Vessels

Blood vessels are derived from capillaries arteries and veins. They act as an important transporter to carry in necessary oxygen and nutrition for healthy growth of retina tissues and carry out waste from retina. As observed in the colour fundus images, arteries appear as light red lines and the veins are appear as darker red lines. In general, the diameter of an artery is narrower than a

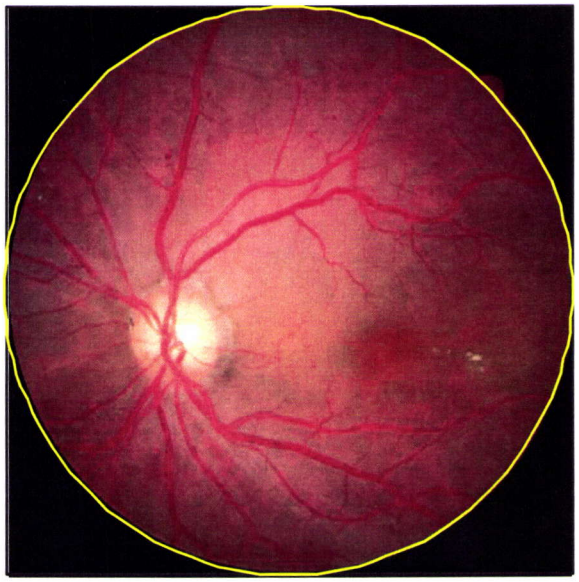


Figure 2.6: The region of retina is highlighted in typical colour fundus image

vein (L et al.,2003). Both arteries and veins appear as largely straight lines with gentle curves. They are more sharply outlined at the center than at the peripheral (Goh et al.,2000). Figure 2.7 highlights the region where arteries and veins are normally found within the eye.

2.2.7 Choroid

The choroid layer lies between sclera and retina layer. This layer or membrane contains blood vessels which transport nutrition and oxygen for the healthy growth of eye tissues (Jalali,2003). In a typical colour fundus images, it is always observed as being reddish in colour (Goh et al.,2000), as shown in Figure 2.3, Figure 2.4, Figure 2.5 and Figure 2.7

2.2.8 Optic Disc Characteristics

This Section briefly describes several dominant characteristics of the optic disc which are important for the detection of diabetic retinopathy. Their importance have been demonstrated by past research (Li & Opas,2001),(Hoover & Goldbaum,2003) and (Goh et al.,2000) as well as through consultations from ophthalmologists who were involved in this research project. The important characteristics include:

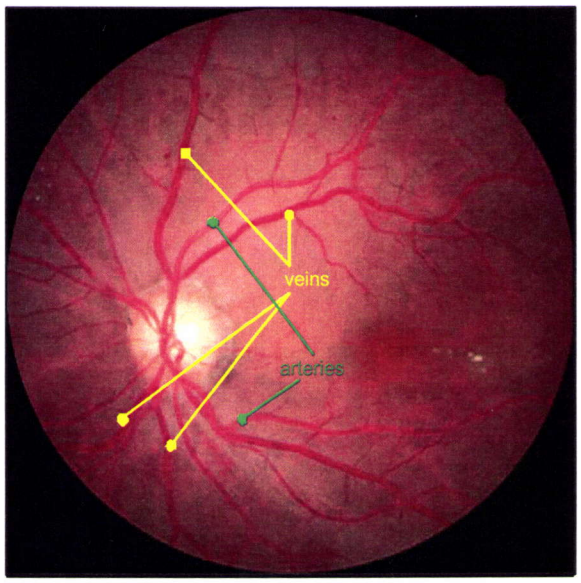


Figure 2.7: The arteries and veins are highlighted in a typical colour fundus image

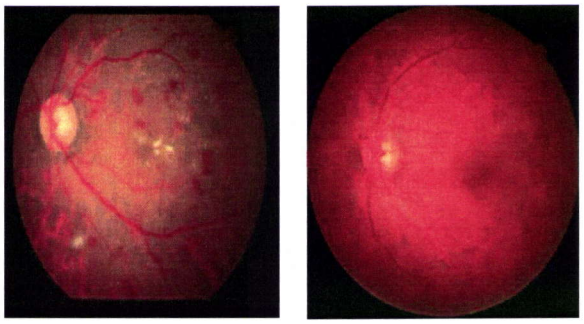


Figure 2.8: Typical colour of the optic disc. (a) Brighter yellowish region (b) dim orange

- Colour

  The optic disc appears as a yellow coloured region in most colour fundus images. At the temporal part, it is observed as a brighter yellowish region Walter et al. (2002). While at the nasal part, it is observed as dim or dull yellow in colour. In some colour fundus images, it may appear as dim orange or as a white circular patch due to uneven illumination. A typical optic disc is shown in Figure 2.8.

- High illumination

  The optic disc receives the most illumination due to the absence of a retinal layer. When observed in separate colour bands of a digital fundus image, the optic disc appears as a white

Figure 2.9: Appearance of the optic disc in different colour bands. (a) red band (b) green band (c) blue band

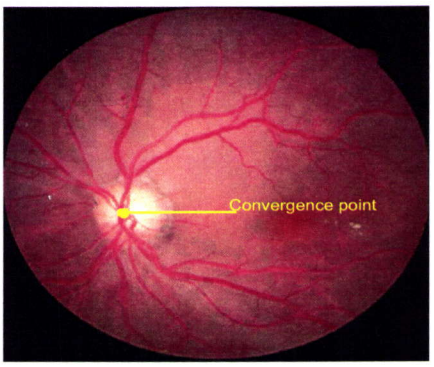


Figure 2.10: Typical optic disc with a convergence point

circular patch. In the green and blue colour bands, the optic disc is observed as a number of small fragmented white patches. Figure 2.9 illustrates the appearance of the optic disc in three different colour bands respectively.

- Obscuration of blood vessels

  The optic disc region is the only region that contains or traversed by a large number of major blood vessels, compared to other portions of fundus image. Owing to this nature, an examination of the optic disc exhibits a high variation of pixel intensity largely contributed by the low intensities of pixels belonging to the blood vessels and high intensities of pixels belonging to the optic disc.The entrance or exit point of a number of major blood vessels, known as the convergence point is also a vital feature of the optic disc. A typical convergence point of major blood vessels is illustrated in Figure 2.10.

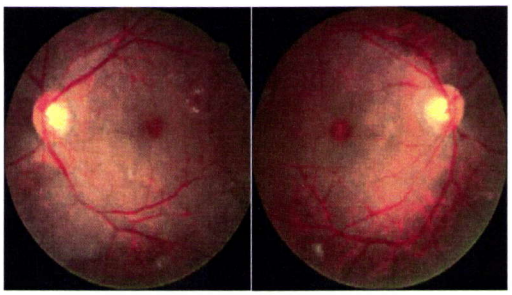


Figure 2.11: The size of the optic disc in typical fundus image

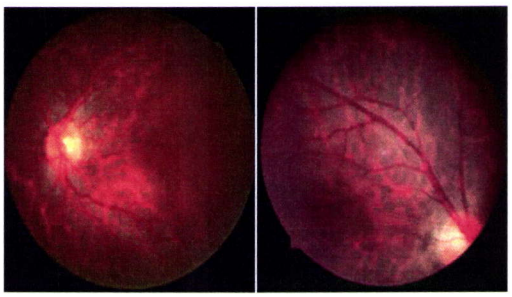


Figure 2.12: (a) An example of a circular optic disc at the anterior of the fundus image (b) An example of a ellipse optic disc at the posterior of the fundus image

- Size of the optic disc

  The size of the optic disc is smaller than or approximately equal to $\frac{1}{8}$ of size of correspond fundus image. The faction of image's size is based upon observations made from database of fundus images used in this research.

- Circularity

  Optic discs may exhibit a variety of shapes. For instance, the optic disc may be observed as circular at the anterior of the fundus image. While at the posterior or at the border of the fundus image, it may appear as ellipse in shape. In actual practice, it is impossible to observe the optic disc as a regular circle or ellipse since blood vessels traversing across the boundary of the optic disc may produce concave and convex boundaries.

2.3 Diabetic Retinopathy

To appreciate the contributions made in this research, it is important to understand the nature of diabetic retinopathy related complications. In general, vision impairments can be classified into

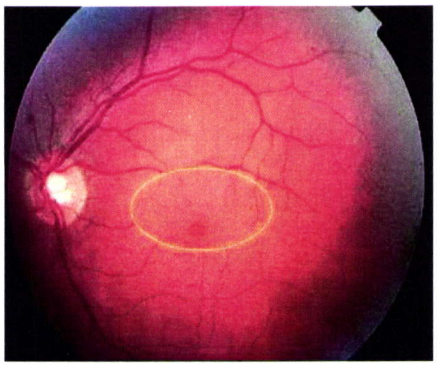


Figure 2.13: Background Retinopathy showing microaneurysms

several severity stages. The severity depends on the amount of damage caused by leakage, blockage and the growth of new blood vessels. There are four types of visual impairment associated with retinopathy; they are background retinopathy, macula oedema (maculapathy), pre-proliferative retinopathy and proliferative retinopathy. Each of these is briefly discussed in the following:

2.3.1 Background Retinopathy

Vision is progressively impaired in patients with background retinopathy. If no proper control of blood glucose is practised among diabetic patients, it may lead to more severe eye damage, i.e. macula oedema (maculapathy), pre-proliferative retinopathy and proliferative retinopathy. At this stage, several significant features can be observed in the fundus image, i.e. microaneurysms, hemorrhages and exudates. Microaneurysms are identified as isolated or a cluster of red spots in the retina due to the capillary damage, such as ballooned-out capillaries and dilated capillaries. However, microaneurysms do no significant harm to vision. The microaneurysms are marked with a yellow circle in Figure 2.13.

Haemorrhages are features that are spotted as red blots in the retina that are caused by the bleeding of damaged capillaries into the retina. This can be seen in Figure 2.14.

Exudates are observed as yellow or white patches around the choroid. It may exist in varying sizes, shapes and in arbitrary locations. The exudates are formed from the deposits of damaged blood vessels whose main components are lipids and proteins. Even though the existence of exu-

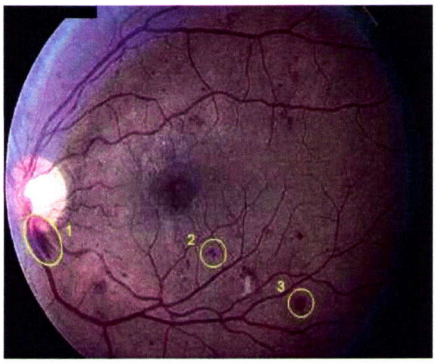


Figure 2.14: Background Retinopathy showing haemorrhages

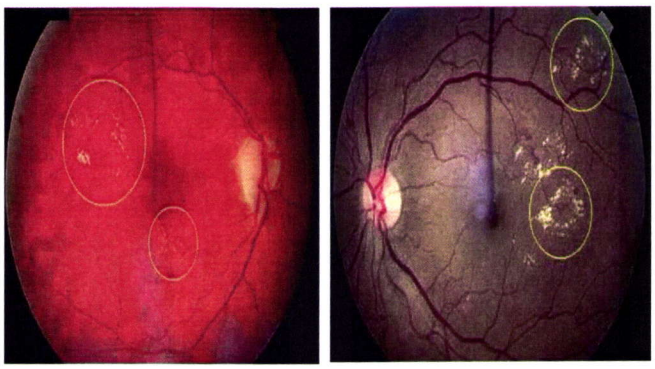


Figure 2.15: Background Retinopathy showing exudates

dates in the retina will not cause any significance harm to vision, but as the exudates accumulate and approach the central region of the macula, the vision is severely compromised. Figure 2.15 shows the common occurrence of exudates.

2.3.2 Maculopathy

Maculopathy occurs when the exudates of leaking blood vessels cover the important parts of the macula and causes it to swell. It is the most common cause of visual impairment. The visual impairments include the inability to focus well while reading and blurring at the side of central visual field. An example of maculopathy is shown in Figure 2.16(a) while the effect of maculapathy on vision is illustrated in Figure 2.16(b).

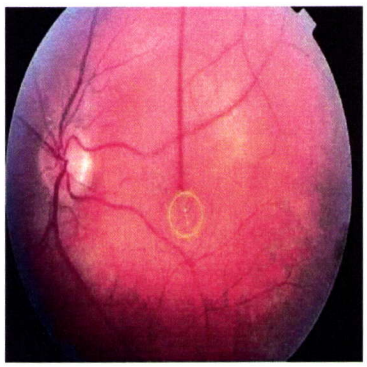


Figure 2.16: Exudates at the macular region

Figure 2.17: Pre proliferative Retinopathy showing IRMs

2.3.3 Pre-Proliferative Retinopathy

Pre-proliferative retinopathy is noted as an advanced stage of eye damage. It may go to the severe extent where vision is totally lost if the condition is not monitored regularly and treated at an early stage. Some of the tell-tale features i.e. interetinal microvascular abnormalities (IRMs) and cotton wool spots are seen at this stage of severity. Intraretinal microvascular abnormalities are areas with microvascular dilatation and intraretinal new vessel formation. They appear as a big patches around the vessels. Illustration of IRMs is shown in Figure 2.17. Cotton wool spots are observed as the white grayish cloudy or discolouration objects. It is the mark of damaged optic nerve. Because of uncontrollable blood flow through the damaged vessels, swelling occurs within the retina damages the surrounding optic nerve. One or two cotton wool spots have no clinical significance, but a cluster of cotton wool spots is regarded as pre proliferative retinopathy. The cotton wool spots are illustrated in Figure 2.18.

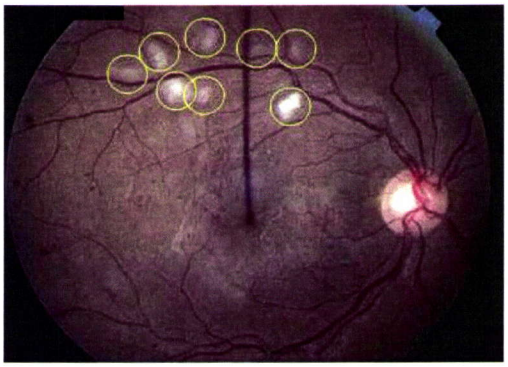


Figure 2.18: Pre proliferative Retinopathy showing cotton wool spots

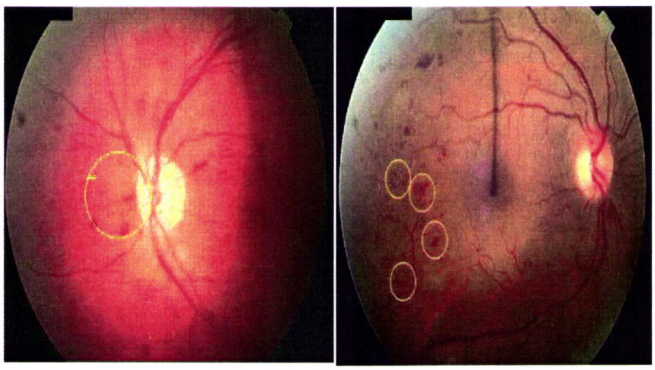


Figure 2.19: Proliferative Retinopathy showing neovascularization. (a) NVD (b) NVE

2.3.4 Proliferative Retinopathy

The main symptom of proliferative retinopathy is neovascularization. Neovascularization refers to the growth of new blood vessels due to shortage of oxygen supply in the retina. The new vessels grow in leaky, wavy and in uncontrolled directions. At the optic disc, sudden rise of blood pressure may rupture the wall of these new vessels easily and subsequently leads to bleeding. Outside the optic disc, the uncontrolled growth of new vessels leak their blood content and dissolve into the vitreous body, causing hazy vision. Two different areas of neovascularization are known as neovascularization at the disc (NVD) and neovascularization elsewhere (NVE). These two symptoms are illustrated in Figure 2.19 (a) and (b) respectively.

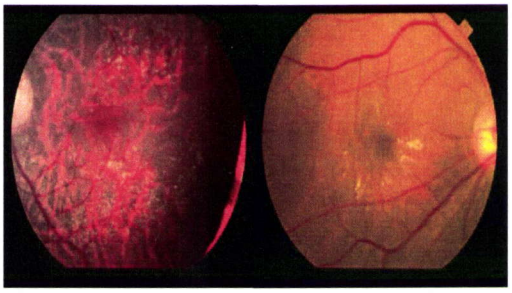


Figure 2.20: Shape variability at arbitrary location of the optic disc

2.4 Challenges in Optic Disc Localization and Segmentation

The computer-aided detection of the optic disc faces several major challenges, one of which is the high variability of colour fundus images. This variabilty is mainly caused by differing imaging conditions during the course of treatment, improper illumination, glare, fadeout and loss of focus. These problems may reduce the visibility of pathological and anatomical structures (K.Fritzsche et al.,2002). Figure 2.24 illustrates the variability of image's quality. The existence of artifacts such as reflection, refraction and dispersion during the image acquisition may mislead the process of the optic disc localisation (K.Fritzsche et al.,2002). The existence of white patches of high intensity at the border or center part of fundus image may also create problems. Figure 2.21(a) and (b). illustrate several examples.

Another factor that affects the quality of fundus image is the movement the patient's eye due to irritation induced by light, in addition to involuntary fixation the fovea on the light (K.Fritzsche et al.,2002). In some instances, the optic disc may not be sharply visible as a regular circular region, instead broken into a number of small regions by obscuration by traversed blood vessels.

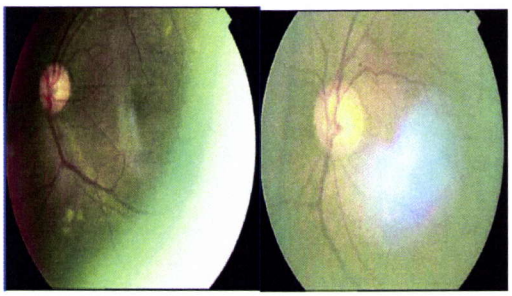


Figure 2.21: Existence of high illumination patch in typical fundus image

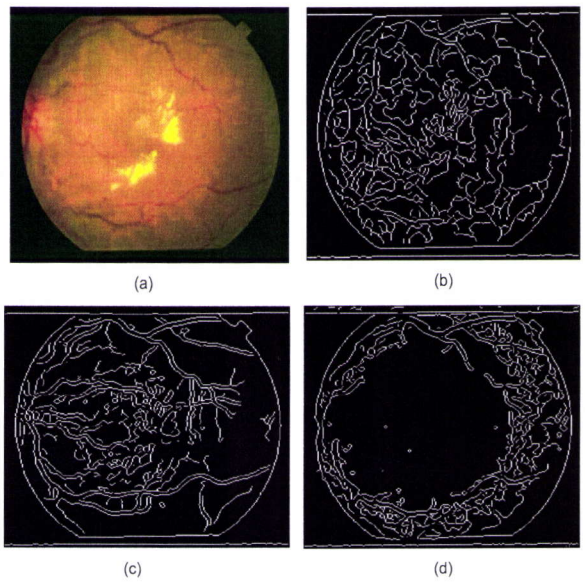


Figure 2.22: Fragmented edges of the optic disc. The contour of the optic disc is invisible

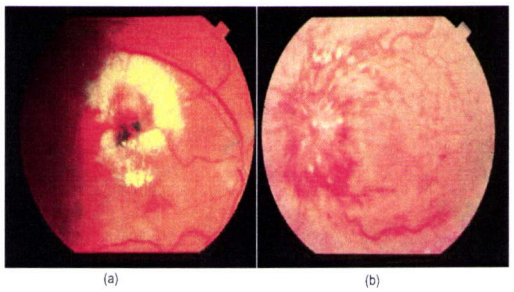
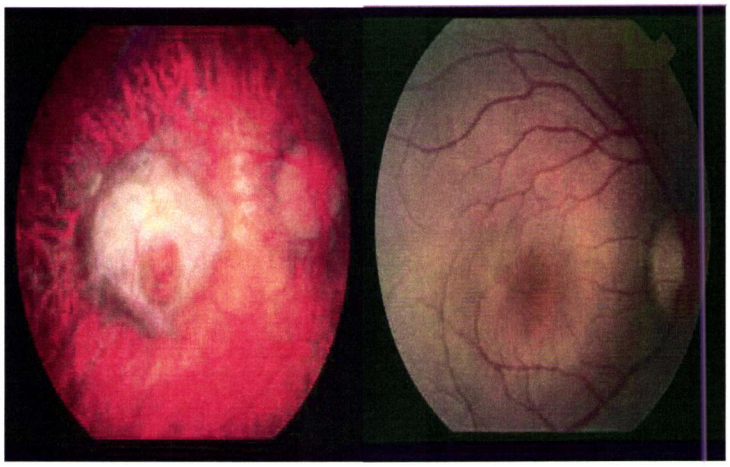


Figure 2.23: (a) low contrast image (b) hollow ring

This inhibits the task of obtaining a complete contour of the optic disc in order to compute the centroid of the optic disc region. Figure 2.22(b), (c) and (d) shows the fragmented regions of the optic disc in red, green and blue colour bands respectively.

In some other fundus images, low contrast and the existence of a hollow ring creates a large fluctuation in optic disc's boundary, which makes the optic disc detection more difficult. Figure 2.23 illustrates two examples of these problems.

Under the uneven illumination cases, the optic disc is observed as a region with lower intensity at the nasal part compare to temporal part of colour fundus image (Lowell et al.,2004). One of the example of uneven illumination is the vignetting. Vignetting is the result of an improper focusing of light which causes the brightness of image gradually decreases radially outward from the center



a b

Figure 2.24: Dim yellow colour of the optic disc produced by differing imaging condition of image (Hoover & Goldbaum,2003).

Besides that, the size and sharp of the optic disc vary significantly at different location of fundus image. In some fundus images, the optic disc is visible as ellipse shape, instead of circular at the the border part of fundus image, as shown in Figure 2.20.

Wide variability in colour appearances of colour fundus image owing to different skin pigmentation and iris colour from different patients may lower the robustness of the optic disc localisation. The variability makes the optic disc appear as bright yellow, dim yellow or orange in colour (A. Osareh & Markham,2002*b*). Figure 2.8 shows the colour variability of the optic disc. Apart from that, the existence of lesions, i.e. haemorrhages and exudates at the vicinity of the optic disc may taken wrongly as the edge of the optic disc and result in false detection of the optic disc contour. This may result in larger size of resulting optic disc contour compare to the actual optic disc contour.

2.5 Summary

In this chapter, the necessary basic knowledge of medical, i.e anatomical and pathological structures of fundus image have been reviewed. Appropriate background material concerning diabetic retinopathy, which includes its complication, prevalence, implications and screening have also been

covered. Understanding the importance of detecting the underlying feature and the common problems encountered by many published work, the following chapter presents a concise review of optic disc localisation as well as optic disc contour segmentation research work.

CHAPTER 3

LITERATURE REVIEW

3.1 Introduction

This chapter presents a review of the recent research in Retinopathy diagnosis using computer vision techniques. More specifically, this chapter highlights the strengths and drawbacks of published works on optic disc localisation and optic disc contour segmentation.

3.2 Computer-aided Fundus Image Analysis

The computer vision research community has in recent years devoted much effort in building an automated retinopathy detection system. A variety of image processing and analysis methods were used to maintain consistency and accuracy in detecting various anatomical and pathological structures. They include detection of optic disc, optic cup, blood vessels, fovea, macula, exudates and haemorrhages. Considering the high prevalence of retinopathy which is a serious health threat in many countries, a complete retinopathy detection system is very timely and necessary. Some of these systems have been already been developed and applied into mass screening programmes. One such screening system is the Telematics in Ophthalmology (OPHTEL) project. The OPHTEL project provides the telemedical services in the field of ophthalmology and diabetology (*OPHTEL. Telematics in Ophthalmology*,2006). The telemedical services includes teleconsultation and telescreening, monitoring and registration. In telescreening functionality, mass screening of retinopathy is allowed to perform and provides necessary information to medical expertise of non-ophthalmologist in primary care and any medical centers without ophthlomologic services. Image processing tools are employed to extract prominent data of fundus image, such as the number of location of pathological structures. Within the telescreening, a knowledge-based monitoring system,which acts as watch-dog is built to support monitoring of patients with chronic diseases or

high risk of getting vision loss. Any new findings in patients' fundus images or changes of patients' database will activate an alarm and on-line query directed to knowledge-based information system. Any new data collected is updated to registration functionality to monitor the progression of retinopathy for the prevention of blindness. Through the teleconsultation functionality, the ophthalmologists and patients are allowed to exchange data from different locations using ISDN-based video conference and Internet-based communication systems.

Another retinal screening system, known as the Digital Healthcare *Detecting And Monitoring Diabetic Retinopathy - Digital Healthcare Wins Prestigious Award For Innovative Healthcare Solution* (2006)was successfully developed by a company in the UK. The Digital Healthcare system is the leading software supplier of opthalmic and diabetic retinopathy screening solutions to over 600 private medical centers, hospitals and universities in US. The system is widely used to screen over 500,000 patients annually for the detection of diseases like glaucoma, retinopahty, cataracts and advance macular degeneration. The electronic administration system, developed by OptoMize IP technology is incorporated into retinal screening system to permit transfer screening result electronically across a network of computers at different locations.

We now move onto more specific aspects of computer-aided retinopathy detection, which may be considered as sub-components of a complete diagnosis system. First, we review optic disc detection methods, followed by optic disc localisation and optic disc contour segmentation methods that have been developed in recent years.

3.3 Optic Disc Localisation

By strict definition, optic disc localisation refers to the process of locating the optic disc by determining any point that lies within the optic disc region. In most of the optic disc localisation approaches, the initial step is to generate a number of possible optic disc candidates to reduce computation time and to provide higher accuracy. As mentioned in **Chapter 2**, there are several characteristics of the optic disc that are typically used to select the most probable optic disc center from a number of candidate regions. Current optic disc localisation algorithms utilise a

multitude of selection criteria and equally different approaches to measure the respective criteria. We review existing work on optic disc localisation from the perspective of the image segmentation approach that have been utilised. They include thresholding, intensity clustering, Haar-based pyramid decomposition and fuzzy c-means clustering.

3.3.1 Region Intensity

(Walter et al.,2002) employed region intensity as the sole criterion for optic disc localisation. Prior to selection of best possible optic disc candidate, a shade corrector operator is applied on fundus image in order to remove low-frequency intensity variation in the image background. The shade corrected operator suppresses any low intensity including those that close to bright region into lower intensity. As such, the contrast of exudate is reduced more than contrast of optic disc region. With suppression of low intensity in shade corrected image, a single thresholding is applied on shade corrected image to remove any cluster with low intensity. The cluster with intensity higher than threshold is taken as optic disc candidate.

The generation of optic disc candidates procedure proposed by (Li & Opas,2001) uses intensity clustering. Intensity is used as the sole criterion to initially determine a small number of possible optic disc candidates. Pixels with intensities higher than 90% of the highest intensity in the image are retained and clustered. These clusters correspond to regions that have higher possibility of being true optic disc region. The optic disc localisation procedure proposed by (Li & Opas,2001) encompasses two steps. First, a number of possible optic disc candidates are sought coarsely using intensity clustering. Intensity is used as the sole criterion to initially determine a small number of possible optic disc candidates. Pixels with intensities higher than 90% of highest intensity in the image are retained and clustered. These clusters correspond to regions that have higher possibility of being true optic disc region. Next, the true optic disc region is identified using principal component analysis. A traning set is first obtained by manually cropping subimages of optic disc. These subimages are transformed into disk space. In this context, disk space is a subspace defined by a set of eigenvectors of principal component analysis. Only the first six representative eigenvectors are chosen as the priori knowledge or model of optic disc. In the optic

disc recognition procedure, an euclidean distance between the optic disc candidate and a disk space defined by the model of optic disc is calculated to measure the similarity of the true optic disc. Thus, the pixel with the minimum Euclidean distance which corresponds to the most similar region to optic disc is located as the center of the optic disc.

3.3.2 Region Intensity and Circularity

M. Lalonde & Gagon (2001) utilised two dominant characteristics of the optic disc, which are intensity and circularity in selecting a highly plausible optic disc candidate. Both characteristics are used to determine a number of possible optic disc candidates. In this process, a multi-resolution decomposition using Haar-based discrete wavelet transform is first performed. Using the low-resolution decomposition, the image is searched for optic disc candidate regions which is typically of high intensity. A simple measured weight is computed for each optic disc candidate in order to measure how similar each region is to the intensity and circularity characteristic of a true optic disc. The measured weight is defined as the ratio between the average pixel intensity inside a circular region of fixed radius to the average intensity within its neighborhood. In this research, the radius of circular region is chosen approximately equal to the typical size of optic disc. The neighborhood is simply a rectangle which is larger than bounding box surrounding the circular region. Following this, a second characteristic of optic disc is used to further prune a number of optic disc candidates, retaining only those which are highly similar to the true optic disc. This procedure is performed using Hausdorff distance-based template matching. Hausdorff distance is used to measure a degree of similarity between a set of pixels of binary image and the template. Optic disc candidates with Hausdorff distances lower than a predefined threshold are retained. For the retained candidates, a second weight, defined as the proportion of template pixels overlapping edge pixels in binary image is calculated. The optic disc candidate with the highest weight is chosen as the center of the optic disc.

A. Osareh & Markham (2002*b*) proposed an optic disc localisation method using template matching. The coefficient of the template provides a *a priori* knowledge that carries region intensity and circularity characteristics of the optic disc. The coefficient of the template is assigned the

average intensity of 25 manually cropped subimages of the optic disc images. Following this, template matching is performed. The best match is the largest normalized correlation coefficient which corresponds to the center of the optic disc.

In a related work, Aliresa Osareh et. al. (Osareh,2004) proposed an alternative method, termed least square arc estimation, for optic disc localisation. The proposed method employs the size of optic disc as well as circularity as the criteria for optic disc localisation. The optic disc localisation starts with fuzzy c-means colour clustering. Clusters with more than 1000 pixels are retained for subsequent processing. With the remaining clusters, the circularity characteristics of each cluster is measured using a compactness criterion. Finally, the center of the most compact cluster is estimated using a fast iterative algorithm (Thomas & Y.T.Chan,1989) that corresponds to the center of true optic disc. The fast iterative algorithm is based upon the minimization of error between a set of pixels on the underlying image.

3.3.3 Obscuration of Blood Vessels

One of the dominant characteristics of optic disc, which is obscuration of blood vessels at optic disc region have been used in (Hoover & Goldbaum,2003). Hoover & Goldbaum (2003) proposed a voting method, known as fuzzy convergence. The fuzzy convergence method detects the convergence or origination point of obscuration of blood vessels at the vicinity of optic disc. Each vessel is modelled as line-like structures using fuzzy segment in binary segmentation of blood vessels. Fuzzy segment, F is simply a model to enumerate the number of pixels found in the predefined neighbourhood. For each pixel enumerated in F, a vote is cast. The procedure of casting vote is repeated for various neighbourhood on binary segmentation of blood vessels. The output of fuzzy segment is the convergence image which encompasses the number of votes that corresponds to the strength of convergence. Then, 1500 pixels with the highest vote are selected and clustered into regions using 8 neighbouring connectivity component labelling. Then, Fisher's linear discrimination is used to find the best separation among these obtained regions. The centroid of region that has best separation of other regions is taken as the center point of true optic disc.

Sinthanayothin et al.,1999 employed the obscuration of blood vessels as a dominant criterion for optic disc localisation. The obscuration of blood vessels is measured statistically by means of variance. In this method, a window of a predefined 80x80 pixels of a window is slid over each pixel on the intensity fundus image. The window with the highest average variance is considered as center of optic disc region.

Lam & Chutatape,2003 proposed an optic disc localisation method by searching the convergence point of major blood vessels which corresponds to obscuration of blood vessels characteristic of optic disc. Initially, segmentation of blood vessels is performed on intensity fundus image using amplitude modified second-order Gaussian filter in (L.Gang et al.,2002). The segmentation procedure produces a matched filter response (MRF) map for the subsequent tracking procedure. In the tracking procedure, the starting point or seed point for vessel tracking is determined by selecting the highest MRF value in both upper and lower section of MRF map respectively. These seed points are the points that are nearest to mid-section of MRF map. At the upper section, the tracking procedure is initiated towards the middle section. A similar procedure is performed at the lower section, where the tracking procedure is initiated up towards the middle section. Both tracking procedures at the upper and lower sections are performed alternately and independently until the stopping criteria are met. Thus, the convergence point is then defined as the middle point of both upper and lower end-tracked points that lie within a 120x120 pixel neighborhood.

3.3.4 Region Intensity and Obscuration of Blood Vessels

Baker,2003 employed the obscuration of blood vessels as the primary criterion and illumination as secondary criterion for optic disc localisation. These two characteristics are measured by average intensity and variance within a window. Initially, an intensity image is divided into a number of rectangular windows, which is about the size of optic disc. In each window, variance is computed. Only four windows with the most significant variance are remained. For the four windows, an average intensity is calculated. Only window with the maximum average intensity is taken as the center of the optic disc.

Lowell et al.,2004 employed a similar approach of template matching as proposed by (A. Osareh & Markham,2002*b*) in optic disc localisation. The template matching proposed by Lowell et al. (2004) consists of one vertical valley in the middle of Laplacian of Gaussian. The valley corresponds to the traversal of a major blood vessel across the optic disc region. The template is correlated with the underlying intensity fundus image, to account for both illumination and obscuration of blood vessels characteristics of optic disc. Pixel with the highest correlation is selected as center of true optic disc region.

3.4 Optic Disc Contour Segmentation

In general, the approaches involved in segmenting the optic disc region can be categorized into three classes: arc estimation, active contour and morphology. The subsequent subsections discusses these approaches.

3.4.1 Arc Estimation

In the arc-estimation class of approaches to optic disc segmentation, the boundary of the optic disc is estimated from the cracked edges of optic disc. This is because the optic disc is often fragmented into smaller regions due to crossing or traversed blood vessels. It is based upon the minimisation error between a set of edges and estimated contour. This type of approach is only adequate to provide an approximate contour of the optic disc. If the optic disc is located at the centre and is distinct from other features in the fundus image, the accurate boundary of optic disc may be obtained. However, if the optic disc is found at the border (posterior) of colour fundus image, the observed shape of the optic disc may not be circular or ellipse. This contributes to false boundary detection or the optic disc contour may not be detected at all.

In the automatic diabetic retinal image screening system (ADRIS) presented by Goh et al. (2000), a combination of various approaches including Sobel edge detector, ellipses fitting and neighborhood accumulation are used to localise the optic disc accurately. Initially, the Sobel edge detector is applied to extract edges in the intensity fundus image. The obtained edges alone are not

useful to locate the optic disc contour since the edges may come from blood vessels, macular and other lesions. In order to search for the location of the optic disc, a range of various ellipse shapes (with differing horizontal and vertical radius) are generated. These ellipses are then superimposed (fitting) on the extracted edges. The number of extracted edges that superimposes with the outline of ellipse indicates the likeliness to the optic disc. Hence, neighborhood accumulation is employed as the measure of accuracy of superimposition. It calculates the number of edges that coincide with the outline of ellipse. The ellipse that returns the maximum response of neighborhood accumulation and higher than a predefined minimum fit is selected as the optic disc and the outline of the ellipse is taken as the contour of optic disc. However, this series of algorithm fails to extract the significant features of optic disc in real fundus images as Sobel edge detector is noise sensitive. It may produce small over-segmented edges that influence the accuracy of responses for the neighborhood accumulation. Apart from that, it is hard for ellipse fitting to be superimposed on the contour of optic disc as it only partially visible at the periphery of the fundus image.

Z. Liu & Krishnan (1997) proposed an automatic optic disc contour segmentation via hough transform. The detection algorithm is applied on the red band of the colour fundus image. At first, optic disc candidate regions are found by searching for pixels having 2% highest intensities in the image within a 180x180 window. The Sobel edge detector is then applied on these candidate regions to extract edges that lie within each candidate regions. Next, the Hough transform is applied on extracted edges. The number of edges that coincides with the outline of Hough transform is counted. The Hough transform with the maximum count is taken as the optic disc's contour. However, this approach only manages to isolate the approximate contour of the optic disc, and not the actual contour.

3.4.2 Active Contour

Snakes or active contour models are energy minimization splines. Snakes are a more general method of matching a deformable model to a desired feature of an image by means of energy minimization. The energy functional that are minimised are a combination of weighted internal and external forces. The internal forces compose of elasticity and bending forces which discourages stretching

and bending. They are emanating from the shape of snake itself. While the external forces consists of pressure and potential forces, which pull the snake towards the desired image features, i.e. image gradient. The contour of the snake is then defined to lie in the position in which the snake reaches a local energy minimum. Since optic disc region is usually fragmented into smaller cracked edges and traversed by blood vessels, conventional segmentation algorithms which are based solely on edge detection cannot accurately segment the optic disc contour. However, snakes which incorporate continuity, curvature and local edge strength properties possess an advantage over conventional edge segmentation. It has the ability to interpolate the missing edges of optic disc owing to obscuration of blood vessels. This advantage makes snakes highly appropriate approach for optic disc contour segmentation application. However, the initial contour of snakes must be close to the actual contour of optic disc in order to minimize the convergence of the snake onto undesired regions. Existing work on optic disc contour segmentation are detailed in the following.

In the work reported by A. Osareh & Markham,2002*b* the true contour of the optic disc is segmented by placing the Gradient Vector Field (GVF) Snake within optic disc region. GVF Snake's initial configuration is always set to 60 pixels radius and positioned at the center point which is obtained via the approach described in (A. Osareh & Markham,2002*a*). However, blood vessels within the optic disc region may negatively influence the convergence of GVF Snake. Thus, prior to convergence, blood vessels are removed using closing colour morphology. Closing colour morphology considers each pixel as a colour vector's component. The blood vessel removal procedure is carried out in *Lab* colour space to produces a more homogeneous region minus any interfering blood vessels. It preserves the optic disc edges well and leads to a more accurate segmentation of optic disc contour. The authors reported that the algorithm successfully segments the optic disc region in all 75 fundus images with a 90.32% accuracy.

F. Mendels & Thiran,1999 described a technique to identify the underlying contour of optic disc using active contour approach. The active contour approach is initialized within the optic disc region. Owing to the existence of blood vessels that traverses the optic disc region, the strong edges of blood vessels may influence the convergence of active contour into undesired contour. In order to overcome this problem, pixels belonging to blood vessels are removed and replaced

with the pixels corresponding to optic disc by two methods of image preprocessing. One is the minima detection. The blood vessels structures are detected by comparing the luminance changes in a local neighborhood. For each pixel, a set of pixels in the shape of crosses at 45^0 relative to each other are selected. The length of the cross is set to 1, 2,3 and 4 pixels long. For each set of pixels, if the luminance of center point of the cross is less than 90% compared to the extrema on all crosses, it is marked as vessel candidate. Next, morphological operation of dilation and erosion using a set of four simple 3x3 structuring elements are performed to eliminate connected structures of blood vessels. For any unconnected vessel, it is identified by searching any pixel which shares the common 4 connected neighbours. This results in a list of pixels which are classified as pixels belonging to blood vessels. The intensities of these pixels are then replaced by the average luminance of the optic disc region. Prior to minima detection, the unsharp mask is applied to enhance image contrast to improve the outcome.

3.4.3 Watershed Transformation

Watershed transformation is a mathematical morphology based segmentation. Watershed transformation has been used in many applications due to its simplicity, speed, intuitiveness and because they can be parallelized. Watershed operation produces a complete division of an image without contour joining Grau* et al. (2004). The concept of watershed transformation was motivated by the topographic view of an image. Assumes the analogue of a landscape and rain. The water gradually falling on the terrain (swiftest descent path) till the lake or sea formed. These lakes or seas are corresponding to region minima or catchments basin in which situated at the low gradient points. While watershed lines or catchment basin boundaries which separate catchment basins are situated at the high gradient points (M. Sonka & Boyle,1993). However, the watershed transform exhibits some pitfalls, such as oversegmentation and sensitive to noise (Grau* et al.,2004). In (Walter et al.,2002), after determining the center point of the optic disc, a subimage which includes the optic disc is obtained. Watershed transformation is then applied to this subimage to locate the contour of optic disc. However, the intensity variation within the optic disc region due to bright pixel and dark vessels' pixel may affect the accuracy of segmented contour. Therefore, the intensity variation is removed by applying the morphological closing operation, followed by

morphological reconstruction to facilitate watershed transformation. The authors tested the algorithm on 30 colour fundus images and found this method to show the exact contour of true optic disc in 27 images. However, the algorithm reveals its weakness if it is applied on images with low contrast, saturated regions and if the optic disc is only partially visible.

3.5 Conclusion

In Section 3.3, a brief yet concise literature review of optic disc detection methods have been presented. Among the published work in optic disc localisation, high illumination has been used as the primary criterion for computing the similarity between optic disc candidates and true optic disc. Obscuration of blood vessels and circularity are also incorporated as the selection criteria to search for the most plausible optic disc candidate. In most of existing algorithms, only a single criterion or a pair of criteria, at most, are employed to maintain simplicity and short computational time. This makes such approaches attractive for real time screening applications. However, there is no guarantee that their performance will be robust considering the highly variable presentation of actual fundus images. This work aims to develop feature vector approach, with an intention to incorporate four prominent characteristics of optic disc, namely obscuration of blood vessels, circularity, colour and size. With these selected characteristics of optic disc as the selection criteria, the feature vector approach exhibits robustness and high accuracy.

In Section 3.4, 3 general categories of optic disc contour segmentation approaches have been described, i.e. active contour, arc estimation and morphology. Of these 3 categories, active contour is the favoured approach as it incorporates continuity, curvature and local edge strength properties into the segmentation process. However, the work by (A. Osareh & Markham,2002*b*) and (F. Mendels & Thiran,1999) lack the automatic initialisation of the active contour. An existing method for automatic initialisation is to set the initial configuration of active contour within a 60 pixels radius from the detected center point of the optic disc. The disadvantage of this rigid approach to automatic active contour initialization is that wrong convergence to nearby lesions may occur if initial configuration of the active contour is positioned partially within the optic disc region. This is evident when the optic disc appears in uncommon shapes and sizes.

Therefore, this research proposes an approach to position the initial configuration of active contour close to the actual boundary of optic disc. With the proposed mechanism of automatic positioning of initial configuration of active contour, the complete boundary of optic disc can be obtained with fewer iterations and lower possibility of wrong convergence.

3.6 Summary

This chapter has presented the concise review for the work of optic disc localisation and optic disc contour segmentation. We have indicated that the use of a single or a pair of criteria to detect the optic disc as well as the lack of an efficient automatic active contour initialisation approach as major factors that are dealt with in this research work. The next chapter will present an overview of proposed optic disc localisation system and describes in detail in each aspect of the proposed system.

CHAPTER 4

THEORETICAL BACKGROUND

4.1 Introduction

This chapter presents the theoretical background necessary to understand the current research. The Section 4.3 reviews the formulation algorithm of unsupervised colour thresholding. This followed by Section 4.4 which presents the theory and fundamental algorithm of anisotropic diffusion, and Section 4.5 gives the fundamental theory of Spline Interpolation. The chapter concludes by discussing the theory of Gradient Vector Flow (GVF) Snake in Section 4.6.

4.2 Image Segmentation

Image segmentation is an essential step in feature extraction, measurement and analysis. It is a process of extracting homogeneous regions of an object of interest. We provide an overview of algorithms for image segmentation by categorizing them into thresholding-based segmentation, region-based segmentation, edge-based segmentation, clustering and deformable models.

Thresholding has became a simple but effective tool in many real industrial applications of image processing. Thresholding is used to separate objects from background as the pixel information in the histogram are distinguishable from the pixels belonging to background. The outcome of thresholding is usually in binary image, comprised of two states. One state represents the foreground objects, which is usually objects or regions of interest and the complementary state indicates the background. There are a number of strategies have been proposed in finding the optimal thresholding using various image properties. In (Rosenfeld & Torre,1983), the concavity point of histogram is applied as threshold. An alternative means is to define the valley point at the coarsest scale of histogram (Carlotto,1997). The actual location of threshold is successively refined

by backtracking from the lowest resolution to the corresponding extrema point in higher resolution. (Otsu,1979) searches the best threshold by minimizing the weighted sum of within-class variance of both background and foreground region. (Niblack,1986) proposed a approach that adjusted the threshold according to the mean and standard deviation within a certain window size. (Trier & Jain,1995) also employed the similar image properties by using a window size of 15 and bias setting of -0.2.

In the context of image processing, a region is defined as a connected set of pixels for which a homogeneity criterion is satisfied (Koschan & Skarbek,1994). The homogeneity criterion is generally based on the analysis of the chromatic characteristic, monochromatic characteristic and texture of a region. A homogeneous region within an image exhibits minimal standard deviation in any of the image properties. Two common segmentation approaches that utilise the concept of region homogeneity are the region growing and the split and merge. In a typical split and merge approach to image segmentation, the segmentation process starts from a an image that is split into four quadrants of equal dimensions. Each quadrant is then recursively tested against a homogeneity criterion. If each quadrant is found not to fulfill this homogeneity criterion, then the said region will be further sub-divided into four quadrants. This test and split procedure is repeated for every quadrant, until there exists no region that could be further subdivided. Following this, the region merging process commences. Adjacent regions that share similar image properties are merged. There are a number of strategies in implementing the split and merge. (Bow,19992) proposed quadtree representation which is a multiresolution scheme at the splitting phase, while region adjacent graph is employed to merge some of the fragmented regions at the merging phase. In (Panjwani & Healey,1995) work, the image properties of colour image are modelled by means of Gaussian Markov Random Field (GMRF). The image is initially split into a number of subregions until each of them contains the homogenous texture described by GMRF. The subregions are merged with agglomerative clustering. In another work of split and merge segmentation, (Shafarenko et al.,1997) applied Watershed transform on the $L^*u^*v^*$ colour model of gradient image and merges the subregions if the criterion homogeneous of colour contrast is met. The region growing process starts from a seed point. The region is 'grown' by comparing neighbouring pixels against certain criterion / criteria of homogeneity. If the neighbouring pixel satsifies this criterion /

criteria, then the pixel is merged with the seed. One of the major difficulties in this approach is the determination of the homogeneity criteria, to ensure no over-segmentation or undersegmentation occurs.

Edge-based segmentation depends on edge information in the image. The edge corresponds to image location of discontinuities in gray level, colour, texture and etc. There are a variety of image edge operators described in the literature. However, it is often insufficient to rely solely on edge-based segmentation as the final result. By incorporating the edge information into other segmentation approaches often produces more promising segmentation results. A simple method to improve the result of edge segmentation is by thresholding. Thresholding can be applied to filter out the non-significant edges (with small edge magnitudes) and retaining significant edges for segmentation. An alternative method of grouping the edges is to use heuristic graph searching. The graph searching algorithm starts from any single edge pixel. It groups the neighbouring pixels based on edge strength and its direction. A threshold value can be introduced to decide which pixel to be joined if more than one neighboring pixels are found. Another heuristic search takes into account of absolute value of image gradient. An arc is created which its gradient orthogonal to the edge direction if many pixels are joined. Pixels and adjacent pixels are considered as nodes and arcs respectively. The search for object boundary is made by weighting the nodes according to its gradient magnitude. For any pixel for which there is more than one path, a graph search is initiated to find the most weighted path.

Segmentation of the objects or structures of interest and reconstructing it into a compact geometric representation with high accuracy is often difficult. It is due to a number of factors such as variability and complexity of objects of interest, noise, artefacts, vague and disconnected of boundaries of region. Low-level image processing methods in the literature can make incorrect assumptions of edges and generate indistinct objects' boundaries. To alleviate this effect, deformable models, a promising model-based approach have been widely employed in computer-assisted image processing systems. This approach is more favourable as it combines geometry, physics and approximation theory (McInerney & Terzopoulos,1996) to provide a powerful image processing tool. The feasibility of deformable models have been widely recognized in segmentation, matching,

shape representation and tracking of structures of interest by exploiting constraints derived from the image itself and a priori knowledge about the location, shape and size of the these structures (McInerney & Terzopoulos,1996). Early works on deformable models include snake deformable model by (Kass et al.,1987), Non-Uniform Rational B-Spline by (K.M.Liang & B.E.Khoo,2002) and elastic template by (Bimbo & Pala,1989). The approach proposed in this research project implements a form of deformable model for segmenting the optic disc.

Fig 4.1 presents a graphical representation of the optic disc segmentation implemented in this work. We briefly describe our approach before discussing the theoretical aspects in the following Sections.

In our approach to optic disc localisation, a number of possible optic disc candidates are generated using unsupervised colour thresholding. This algorithm clusters the colour image into several clusters using a bit-code representation. The formulation for the algorithm of the unsupervised colour thresholding is presented in Section 4.3. From the pool of optic disc candidates, a single candidate that is most similar to a representative optic disc is chosen. In this research, a feature vector that utilises four salient characteristics of the true optic disc is employed to measure the similarity of a candidate region to the gold standard. Each of these characteristic is measured using unique approaches. The size of optic disc is a measure of the number of pixels within a cluster, and the colour is computed by calculating the Euclidean distance from a reference colour, derived from the HSI colour model. The third characteristic of shape of optic disc is simply measured by compactness criterion approach. The last characteristic requires the search for the convergence point or origination point of major blood vessels in fundus images. For faster computation, only major blood vessels are taken into consideration. Small blood vessels are removed prior to the search of convergence point. Other features, such as exudates, microaneurysms and haemorrhages are removed as well so that the convergence point of blood vessels is detected with high accuracy. In order to remove these mentioned features, anisotropic diffusion (Perona & Malik,1990) is chosen as it is capable of smoothing unwanted data while preserving the necessary data. Moreover, anisotropic diffusion is a more favourable filtering tool than Gaussian smoothing as it preserves the true location of object's edges (Acton,2000) and (Gerig et al.,1992). To provide clarity, the

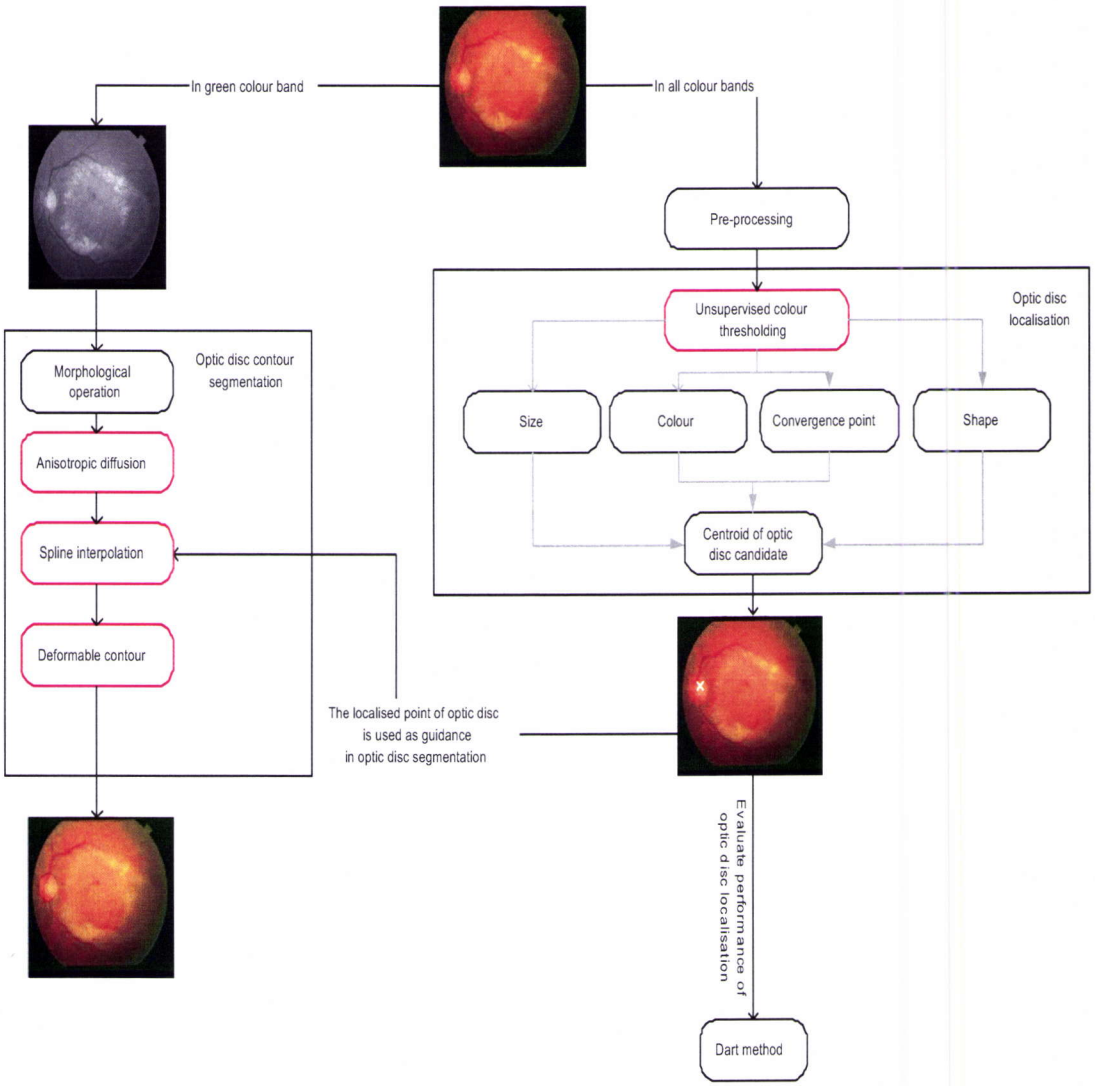


Figure 4.1: The overview of the proposed optic disc detection system.

theory and algorithm of anisotropic diffusion is given in Section 4.4.

In our optic disc contour segmentation approach, a combination of morphological operation and anisotropic diffusion eliminates the blood vessels structures that may obscure the boundary of the optic disc. With the image free of obscuring blood vessels, the edges belonging to the optic disc are obtained by calculating the high intensity gradients of the first order derivative of the intensity profile. Nevertheless, some of the obtained edges (outliers) may not belong to the true boundary of the optic disc. These outliers are repositioned using spline interpolation such that they comply with the shape of the true optic disc. The edges that comply with the optic disc's shape are then used as the initial configuration for the deformable contour. For a comprehensive understanding of the implementation of both spline interpolation and deformable contour, the mathematical elaboration of both these approaches are given in Section 4.5 and 4.6 respectively. The final outcome of this research is the accurate segmentation of the true contour of the optic disc.

4.3 Unsupervised Colour Thresholding

4.3.1 Otsu Thresholding

In the unsupervised colour thresholding approach used in this work, threshold values for each of the Red, Green and Blue colour bands are obtained using the Otsu method (Otsu,1979). Otsu thresholding (Otsu,1979) has been widely used to determine the optimal threshold for image segmentation and analysis. In the Otsu method, the probabilities of the foreground region, P_F^t and the background region, P_B^t of a gray level image are calculated using

$$p_B^t = \sum_{i=0}^{t} p_i \tag{4.1}$$

$$\begin{aligned} p_F^t &= 1 - p_B^t \\ &= \sum_{i=t+1}^{L-1} p_i \end{aligned} \tag{4.2}$$

where t is the selected threshold gray level value, and $\{p_i\}$ is the gray level histogram of each colour band. The mean and variance for the foreground and background regions are then computed as:

$$\begin{aligned} \mu_B^t &= \left(1/p_B^t\right) \sum_{i=0}^{t} i p_i \\ &\textit{and} \\ \mu_F^t &= \left(1/p_F^t\right) \sum_{i=t+1}^{L-1} i p_i \end{aligned} \tag{4.3}$$

where μ_B^t and μ_F^t are means associated with background and foreground, and the variance is computed as:

$$\begin{aligned} \mathrm{var}_B^t &= \left(1/p_B^t\right) \sum_{i=0}^{t} (i - \mu_B^t)^2 p_i \\ &\textit{and} \\ \mathrm{var}_F^t &= \left(1/p_F^t\right) \sum_{i=t+1}^{L} -1(i - \mu_F^t)^2 p_i \end{aligned} \tag{4.4}$$

where var_B^t and var_B^t are variances associated with background and foreground. Finally, the Otsu theshold value, t_{otsu} is the threshold, determined based upon two criteria: the maximization of the inter-class variance, $var_{between-class}^t$ and minimization of the intra-class variance, $var_{within-class}^t$. Otsu threshold, t_{otsu} is given by :

$$t_{otsu} = \arg\left\{ \max_{1 \le t \le L} \left(\mathrm{var}_{inter-class}^t\right) \right\} or \arg\left\{ \min_{1 \le t \le L} \left(\mathrm{var}_{intra-class}^t\right) \right\} \tag{4.5}$$

In Equation 4.5, the inter class variance is denoted as

$$\begin{aligned} \text{var}^t_{inter-class} &= p^t_B \left(\mu^t_B - \mu\right)^2 + p^t_F \left(u^t_F - \mu\right)^2 \\ &= p^t_B p^t_F \left(\mu^t_B - \mu^t_F\right)^2 \end{aligned} \tag{4.6}$$

where $\mu = \sum_{t=0}^{L-1} ip_i$ is the global mean of the image. Whereas, the intra class variance is given by:

$$\text{var}^t_{intra-class} = p^t_B \text{var}^t_B + p^t_F \text{var}^t_F \tag{4.7}$$

4.3.2 Colour Thresholding

Given that the threshold values denoted by t_r, t_g and t_b for red, green and blue colour bands respectively as in Eq 4.5. Each colour pixel is specified by $p_{i,j} = (r_{i,j}, g_{i,j}, b_{i,j})$ with (i, j) as the spatial coordinate of the corresponding pixel. A thresholding operation is applied on each colour pixel $\boldsymbol{p_{i,j}} = (\boldsymbol{r_{i,j}}, \boldsymbol{g_{i,j}}, \boldsymbol{b_{i,j}})$ according to

$$\boldsymbol{r_{i,j}} = \begin{cases} 1 & r_{i,j} > t_r \\ 0 & r_{i,j} \leq t_r \end{cases} , \boldsymbol{g_{i,j}} = \begin{cases} 1 & g_{i,j} > t_g \\ 0 & g_{i,j} \leq t_g \end{cases} , \boldsymbol{b_{i,j}} = \begin{cases} 1 & b_{i,j} > t_b \\ 0 & b_{i,j} \leq t_b \end{cases} \tag{4.8}$$

After the colour thresholding as described by Equation 4.8 is completed, each colour pixel is encoded with a 3-bit binary code, (c_1, c_2, c_3). The colour fundus image is then clustered into eight possible classes, $\{C_k\}_{k=1}^{8}$, resulting from the 3-bit encoding of the pixels. Following this, the mean intensity of each class is computed. Assuming that the k-th class, C_k, its mean is given by $\mu_k = (r_k, g_k, b_k)$. Each mean component is described by

$$r_k = \frac{\sum_{r_{i,j} \in C_k} r_{i,j}}{\sum_{r_{i,j} \in C_k} 1} \quad , \quad g_k = \frac{\sum_{r_{i,j} \in C_k} g_{i,j}}{\sum_{r_{i,j} \in C_k} 1} \quad , \quad b_k = \frac{\sum_{r_{i,j} \in C_k} b_{i,j}}{\sum_{r_{i,j} \in C_k} 1} \tag{4.9}$$

The definitions of intra-class variance and inter-class variance used in Otsu (Otsu,1979) are utilised as the criteria in an iterative clustering strategy. The intra-class variance is defined as follows:

$$\sigma_k = \frac{1}{N_k}\left\{\sum_{i,j\in C_k}\left[(r_{i,j}-r_k)^2+(g_{i,j}-g_k)^2+(b_{i,j}-b_k)^2\right]\right\} \tag{4.10}$$

where N_k is the number of pixels in class C_k. The second criterion, the inter-class distance, σ_{kj} between class C_k and class C_j with $k \neq j$ is defined as :

$$\sigma_{k,j} = \sqrt{(r_k-r_j)^2+(g_k-g_j)^2+(b_k-b_j)^2} \tag{4.11}$$

If the inter-class distance between class C_k and C_j is less than the intra-class distance for classes ($\sigma_k \geq \sigma_{kj}$ or $\sigma_j \geq \sigma_{kj}$), the two classes are merged into a single class. The clustering process for Eq 4.9, Eq 4.10 and Eq 4.11 is repeated until the intra-class distance of all classes are less than their corresponding inter-class distance. The details of numerical implementation of unsupervised colour thresholding are found in (Y. Du & Thouin,2004) and (Du & C.Chang 2003).

4.4 Anisotropic Diffusion

The diffusion equation originates from the works of Perona and Malik (Perona & Malik,1990). It is a non-linear or tunable filtering operation that can be used for smoothing, enhancement and detection of object boundaries. The diffusion process uses a simple partial differential technique and the incorporates adiabatic boundary condition which prefers intra-region smoothing and arrests inter-region smoothing. Mathematically, the diffusion equation is described as:

$$\frac{\partial}{\partial t}I\left(\bar{x},t\right) = \nabla\cdot\left(c\left(\bar{x},t\right)\nabla I\left(\bar{x},t\right)\right) \tag{4.12}$$

where $\bar{x} \in (x, y, z)$ of image axes and t refers to the number of iteration step. The conductance function is described in Equation 4.13.

$$c\left(\bar{x}, t\right) = f\left(\left|\nabla I\left(\bar{x}, t\right)\right|\right) \tag{4.13}$$

The conductance function is a non-negative and monotonically decreasing function of the image gradient magnitude. It allows the amount of diffusion to be varied locally, where a neighbourhood of image is selected either for smoothing or enhancement. The diffusion function determines whether the filter behaves as a smoothing filter or as an enhancing operator. An example diffusion model which was introduced in (Perona & Malik,1990) is given by Equation 4.14.

$$c\left(\bar{x}, t\right) = \frac{1}{1 + \left(\frac{\nabla I}{k}\right)^2} \tag{4.14}$$

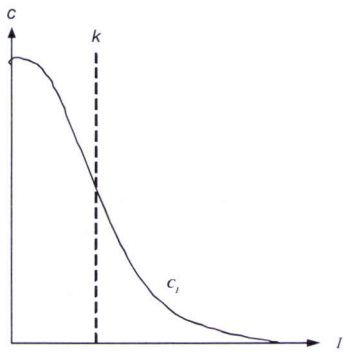


Figure 4.2: Non-negative and monotonically decreasing of diffusivity function

Figure 4.2 depicts the diffusivity function that inherits the two selective preserving and smoothing functionalities. It may be observed that the value of c decreases as the gradient magnitude is increased. The behaviour of the diffusivity function is directly dependent on the value of c. That is, for low values of c, smoothing is minimal and no smoothing takes place for c=0. This happens at high gradient magnitudes which are found typically at the object boundaries. In other words, if a large gradient magnitude is found, a discontinuity is assumed and diffusion process is arrested

(Gerig et al.,1992). In similar fashion, the maximum value of $c = 1$ indicates that smoothing is allowed to take place as the gradient magnitude decreases. This happens within the neighbourhoods of homogeneous regions. Referring to Equation 4.14, an additional parameter, k is introduced. The k parameter is the most important variable which controls the behaviour of the filter. It is known as the *diffusion constant* or flow constant. To understand how diffusivity function works on image, a flux, Φ is introduced in the subsequent paragraph.

The flux or flow function is defined as the rate of smoothing and is expressed mathematically as

$$\Phi\left(\bar{x},t\right) = c\left(\bar{x},t\right) \cdot \nabla I\left(\bar{x},t\right) \tag{4.15}$$

Flux is the product of the diffusivity function and the gradient magnitude. Figure 4.3 illustrates Φ as a function of gradient magnitude.

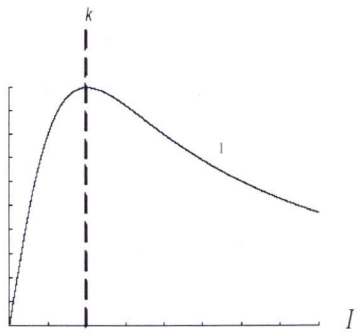


Figure 4.3: Non-negative and monotonically decreasing of flow function

As can be observed in Figure 4.3, the amount of flux (rate of smoothing) is zero as the gradient magnitude is nil, and reaches a maximum at k. After point k, the amount of flux decreases to a minimum. This behaviour implies that the diffusion process decreases the rate of smoothing on both sides of k. The gradient magnitude at $\nabla I \leq k$ indicates a homogeneous region or low fluctuation region. Hence, a minimal amount of smoothing is enough to smooth the neighbourhood. The

$\nabla I \geq k$ indicates a boundary region or a region with significant features. A small flow is generated across edges so that these features will not be over smoothed. The maximum flow or maximum rate of smoothing is only generated as the gradient magnitude becomes equal to the flow constant, k. With this in mind, in order to function as a smoothing operator, the k is chosen appropriately such that it corresponds to the gradient magnitude produced by noise. Whereas in order to preserve image features, the k is chosen to correspond less than the gradient magnitude of edges.

4.4.1 Discrete Implementation

In order to filter digital images using anisotropic diffusion, a 2D discrete implementation is used to map the continuous domain of Equation 4.12 to the discrete domain. Generally, a 2D discrete implementation is composed of two 1D discrete implementation. For clarity, a 1D discrete implementation is briefly discussed.

In the case of 1D discrete implementation, the intensity of one upper iteration is defined as the sum of the flow contributed by two neighbouring pixels intensities, $I(x + \Delta x)$ and $I(x - \Delta x)$. The diagram of 1D discrete case is illustrated as a network in Figure 4.4. The centre pixel represents the diffused node and it is linked by neighbouring arcs. The magnitude of the arc is determined by the magnitude of the flow function, as given in Equation 4.15.

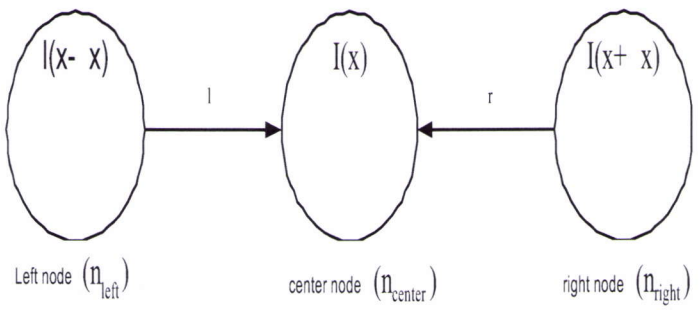


Figure 4.4: One dimensional discrete implementation network. The centre node, n_{center} is linked by both arcs, ϕ_{left} and ϕ_{right} from the neighbouring nodes

The 1-dimensional discrete implementation that derives from the continuous domain of Equa-

tion 4.12 is formulated by replacing the discrete approximation of the gradient into Equation 4.13 and is described in the following:

$$c\left(x,t\right) \approx f\left(\frac{1}{\Delta x}\left|I\left(x+\frac{\Delta}{x},t\right)-I\left(x-\frac{\Delta x}{2},t\right)\right|\right) \tag{4.16}$$

At any pixel, the conductance function, c in Equation 4.13 is composed of 2 neighbouring pixels. The c contributed from the right pixel, the c is defined as in Equation 4.17, while in Equation 4.18 shows the c from the left pixel. Both equations are stated as follows.

$$c\left(x+\frac{\Delta x}{2},t\right) \approx f\left(\frac{1}{\Delta x}\left|I\left(x+\Delta x,t\right)-I\left(x,t\right)\right|\right) \tag{4.17}$$

$$c\left(x-\frac{\Delta x}{2},t\right) \approx f\left(\frac{1}{\Delta x}\left|I\left(x,t\right)-I\left(x-\Delta x,t\right)\right|\right) \tag{4.18}$$

By substituting the discrete approximates into Equation 4.12, the diffusion equation can be described clearly in the form of flux, ϕ contributed from neighbouring pixels. It is formulated as:

$$\begin{aligned}
\tfrac{\partial}{\partial x} I\left(\bar{x},t\right) &= \nabla \cdot \left(c\left(\bar{x},t\right)\nabla I\left(\bar{x},t\right)\right) \\
&\approx \tfrac{1}{\Delta x} \cdot \left(c\left(x,t\right)\tfrac{1}{\Delta x} I\left(x,t\right)\right) \\
&\approx \tfrac{1}{\Delta x^2}\left(c\left(x+\tfrac{\Delta x}{2},t\right)\cdot\left(I\left(x+\Delta x,t\right)-I\left(x,t\right)\right)-c\left(x-\tfrac{\Delta x}{2},t\right)\cdot\left(I\left(x,t\right)-I\left(x-\Delta x,t\right)\right)\right) \\
&\approx \tfrac{1}{\Delta x^2}\left(\Phi_{right}-\Phi_{left}\right) \\
&\approx \Phi_{right}-\Phi_{left}|\Delta x = 1
\end{aligned} \tag{4.19}$$

where $\bar{x} \in x$ axis, t is the current iteration and $(t + \Delta t)$ is the subsequent iteration.

For each iteration, the value of the centre pixel (depicted in Figure 4.4) is changed according

to the flow contribution of its neighbouring pixels. An expression of one dimensional discrete implementation for a single pixel can be derived as:

$$\begin{aligned} I(x, t+\Delta t) &\approx I(x,t) + \Delta t \cdot \frac{\partial}{\partial t} I(x,t) \\ &\approx I(x,t) + \Delta t \cdot (\Phi_{right} - \Phi_{left}) \end{aligned} \tag{4.20}$$

4.4.2 2D Discrete Implementation

The 2D discrete implementation can be easily extended from the 1D discrete implementation's concept. It can be assumed as a 2 dimensional network structure composed of 1D horizontal and 1D vertical discrete implementation, as illustrated in Figure 4.5.

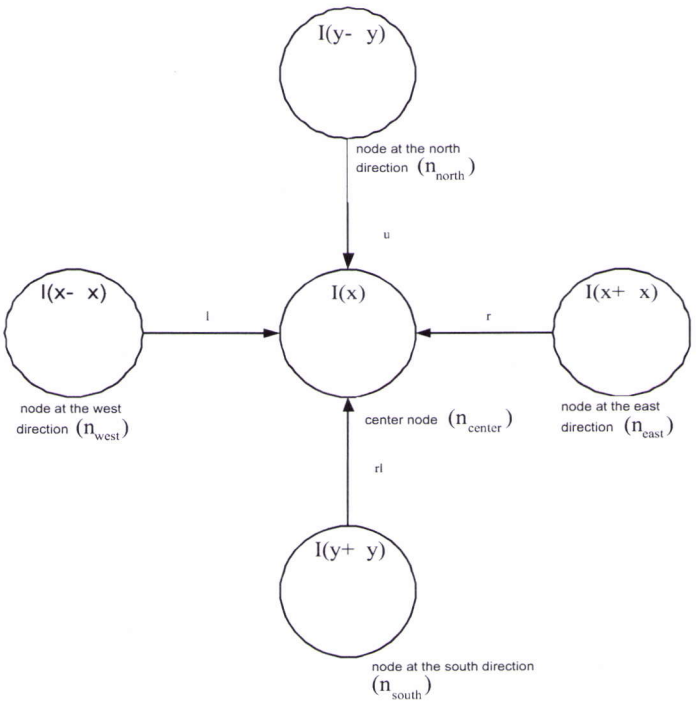


Figure 4.5: As with 2D case, the flow may contributed from 4 neighbouring pixels, which are represented by *north*, *south*, *east* and *west*

In the 2D context, Equation 4.12 can be rewritten as:

$$\frac{\partial}{\partial x} I(x,y,t) = \nabla_x \left[c(x,y,t) \cdot \nabla_x I(x,y,t) \right] + \nabla_y \left[c(x,y,t) \cdot \nabla_y I(x,y,t) \right] \tag{4.21}$$

Thus, the 2D diffusion equation can be described in terms of ϕ as follows:

$$\begin{aligned}
\tfrac{\partial}{\partial t} I(x,y,t) &\approx \tfrac{\partial}{\partial x} \left[c(x,y,t) \, \tfrac{\partial}{\partial x} I(x,y,t) \right] + \tfrac{\partial}{\partial y} \left[c(x,y,t) \, \tfrac{\partial}{\partial y} I(x,y,t) \right] \\
&\approx \tfrac{1}{(\Delta x)^2} \left[\begin{array}{l} c\left(x + \frac{\Delta x}{2}, y, t\right) \cdot \left(I(x + \Delta x, y, t) - I(x,y,t) \right) \\ -c\left(x - \frac{\Delta x}{2}, y, t\right) \cdot \left(I(x,y,t) - I(x - \Delta x, y, t) \right) \end{array} \right] \\
&+ \tfrac{1}{(\Delta y)^2} \left[\begin{array}{l} c\left(x, y + \frac{\Delta y}{2}, t\right) \cdot \left(I(x, y + \Delta y, t) - I(x,y,t) \right) \\ -c\left(x, y - \frac{\Delta y}{2}, t\right) \cdot \left(I(x,y,t) - I(x, y - \Delta y, t) \right) \end{array} \right] \\
&\approx \tfrac{1}{(\Delta x)^2} \left(\Phi_{east} - \Phi_{west} \right) + \tfrac{1}{(\Delta y)^2} \left(\Phi_{south} - \Phi_{north} \right) \\
&\approx \Phi_{east} - \Phi_{west} + \Phi_{south} - \Phi_{north} | \Delta x = \Delta y = 1
\end{aligned} \tag{4.22}$$

With the flow value calculated from each of its neighbouring pixels, the centre pixel can be upgraded as follows:

$$I(x,y,t + \Delta t) \approx I(x,y,t) + \Delta t \cdot \left(\Phi_{east} - \Phi_{west} + \Phi_{north} - \Phi_{south} \right) \tag{4.23}$$

Thus, based on Equation 4.23, the intensity of the center pixel at the subsequent iteration, $(t + \Delta t)$ can be obtained by the summation of intensity at the current iteration, t and the flow contribution from its neighbouring pixels.

4.5 Spline Interpolation

The basic idea for cubic spline interpolation is inspired from the drawing tool used by an engineer to draw a smooth curve through a number of defined points. From a numerical data analysis point of view, the idea of cubic spline interpolation was developed to correlate the real world numerical data effectively (de Boor Carl,1978). Spline consists of a series of unique cubic interpolation that

are fitted between each of the data points. In this case, these data points are the numerical data. In spline, a flexible strip with respective weight at interval, is bent across data points, resulting a continuous and smooth curve. In the mathematical spline context, the weights are the coefficients in the cubic spline interpolation. These coefficients bend and turn the strip (line) that passes through each data points without any erratic behaviour or breaks in continuity.

Cubic spline interpolation is a piecewise polynomial with points that are smoothly connected together. Assuming a one-dimensional spline for a set of $n-1$ points, $(y_1, ..., y_n)$, a polynomial of cubic spline interpolation is described as a function of order 3 as in Equation 4.24.

$$p_i(t_k) = a_i + b_i t_k + c_i t_k^2 + d_i t_k^3 \tag{4.24}$$

,where $i = 1, ..., n-1$ and t_k denoted as the length of k-th interval, $x_{k+1} - x_k$. It is a parameter of $[0, 1]$.

The cubic spline interpolation conforms to two criteria. They are:

1. The piecewise function, $P(t_k)$ interpolates all the data points.
2. $P(t_k)$ is a continuous function of its (n) arguments (de Boor Carl,1978). They are:
   (a) $P(t_k)$ is continuous on the interval $[x_1, x_n]$.
   (b) $P'(t_k)$ is continuous on the interval $[x_1, x_n]$.
   (c) $P''(t_k)$ is continuous on the interval $[x_1, x_n]$.

The zero derivative of Equation 4.24 with $t_k = 0$ $(x_{k+1} = x_k)$ is defined as

$$\begin{aligned} p_i(0) &= a_i \\ &= y_i \end{aligned} \tag{4.25}$$

With the similar equation, with $t_k = 1$ $(x_{k+1} \neq x_k)$, then

$$\begin{aligned} p_i(1) &= a_i + b_i + c_i + d_i \\ &= y_{i+1} \end{aligned} \tag{4.26}$$

Taking the first derivative of $p_i(x)$ for each interval,

$$p_i^t(t_k) = b_i + 2c_i t_k + 3d_1 t_k^2 \tag{4.27}$$

If $t_k = 0$, then

$$\begin{aligned} p_i^{'}(0) &= b_i \\ &= D_i \end{aligned} \tag{4.28}$$

and with $t_k = 1$,

$$\begin{aligned} p_i^{'}(1) &= b_i + 2c_i + 3d_i \\ &= D_{i+1} \end{aligned} \tag{4.29}$$

,where the D is the slope parameter (de Boor Carl,1978).

Thus, in order to determine the four weights a_i, b_i, c_i and d_i for the polynomial equation, the following equations are defined:

$$a_i = y_i \ (derived\ from\ equation\ \ 4.25) \tag{4.30}$$

$$b_i = D_i \ (derived\ from\ equation\ 4.28) \tag{4.31}$$

Solving Equation 4.29 - 3x(Equation 4.26), gives

$$c_i = 3\left(y_{i+1} - y_i\right) - 2D_i - D_{i+1} \tag{4.32}$$

Solving Equation 4.28 - 2x(Equation 4.26), gives d_i as

$$d_i = 2\left(y_i - y_{i+1}\right) + D_i + D_{i+1} \tag{4.33}$$

Taking the second derivative of $p_i\left(t_k\right)$ in each interval,

$$p_i^{''}\left(t_k\right) = 2c_i + 6d_1 t_k \tag{4.34}$$

If $x = 0$, then

$$p_i^{''}\left(0\right) = 2c_i \tag{4.35}$$

and with $t_k = 1$ is

$$\begin{aligned} p_i^{''}\left(1\right) &= 2c_i + 6d_i \\ &= 2c_{i+1} \end{aligned} \tag{4.36}$$

Following this, by rearranging Equation 4.36, 4.33 and 4.32 leads to the following condition:

$$D_i + 4D_{i+1} + D_{i+2} = 3r_{i+1} \tag{4.37}$$

where the r_k is defined as

$$r_i = t_i\delta_{i-1} + t_{i-1}\delta_i \quad where\, i \in [2, 3, ..., n-1] \tag{4.38}$$

with $\delta_k = \frac{y_{k+1}-y_k}{t_k}$ as the first divided difference.

Equation 4.37 is applied to each points, $(x_2, ..., x_{n-1})$ by rearranging them into a symmetric tri-diagonal system.

$$\begin{bmatrix} 1 & 4 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 1 & 4 & 1 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 4 & \cdots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 4 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 1 & 4 & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 4 & 1 \end{bmatrix}_{(n-2,n)} \begin{bmatrix} D_1 \\ D_2 \\ D_3 \\ \vdots \\ D_{n-2} \\ D_{n-1} \\ D_n \end{bmatrix}_{(n,1)} = 3 \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ \vdots \\ r_{n-2} \\ r_{n-1} \\ r_n \end{bmatrix}_{(n,1)} \tag{4.39}$$

The preceding matrix equation with $n-2$ rows and n columns only involves $n-2$ equations in solving the n unknowns D_i. Thus, the Equation 4.39 is under determined. The cubic spline interpolation employed in this research uses the "not-a-knot" strategy owing to unknown end points (x_1 *and* x_n) derivatives.

In the x_1 *and* x_n strategy, the $p'''(x)$ is required to be continuous across x_2 and x_{n-1}, which involves the first two subintervals $x_1 \leq x \leq x_3$, and last two subintervals $x_{n-2} \leq x \leq x_n$, therefore,

two conditions are imposed in solving the unknowns end points of x_1 *and* x_n:

$$\begin{aligned} & p_1 = p_2 \\ & and \\ & p_{n-1} = p_n \end{aligned} \tag{4.40}$$

Thus, Equation 4.41 (derived from $p_1 = p_2$) is added at the beginning of and Equation 4.42 is appended to the matrix equation in Equation 4.39. The details of cubic spline interpolation with x_1 *and* x_n) condition is described in (de Boor Carl,1978) and (Weisstein,2005).

$$D_1 t_2 + D_2 (x_3 - x_1) = \frac{t_1 + 2(x_3 - x_1) t_2 \delta_1 + t_1^2 \delta_2}{x_3 - x_1} \tag{4.41}$$

$$D_{n-1} (x_n - x_{n-2}) + D_n t_{n-2} = \frac{(t_{n-1})^2 \delta_{n-2} + (2(x_n - x_{n-2}) + t_{n-1}) t_{n-2} \delta_{n-1}}{x_n - x_{n-2}} \tag{4.42}$$

With these two *not a knot* conditions included, a n linear equations in n unknowns are obtained. This produces a tri-diagonal system with n equations to solve for n unknowns:

$$\begin{bmatrix} 1 & 2 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 1 & 4 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 1 & 4 & 1 & \cdots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & 4 & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 4 & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 2 & 1 \end{bmatrix}_{(n,n)} \begin{bmatrix} D_1 \\ D_2 \\ D_3 \\ \vdots \\ D_{n-2} \\ D_{n-1} \\ D_n \end{bmatrix}_{(n,1)} = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ \vdots \\ r_{n-2} \\ r_{n-1} \\ r_n \end{bmatrix}_{(n,1)} \tag{4.43}$$

4.6 Gradient Vector Flow Snake

Active Control models or Snakes (Kass et al.,1987) are a more general method of matching a deformable model to a desired feature of an image by means of energy minimization. They are elastic manifolds defined within an image domain move under the influence of internal forces and external forces. The internal forces compose of elasticity and bending forces which discourages stretching and bending, coming from from the shape of snake itself. The external forces consists of pressure and potential forces, which pull the snake toward the desired image features, i.e. image gradient. These depend on the image data. The contour of the snake is then defined to lie in the position in which the snake reaches a local energy minimum. The snake is parametrically defined as

$$v(s) = [x(s), y(s)], s \in [0, 1] \tag{4.44}$$

If *x(s)* are the parameterized manifolds where *(*s) is the parameter of the manifold, the final contour minimizes the representation of the energy of the manifold *(E)* consisting of internal energy ε_{int} and external energy ε_{ext} as follows:

$$E = \varepsilon_{\text{int}} + \varepsilon_{ext} \tag{4.45}$$

$$\varepsilon_{\text{int}} = \int_0^1 \left(\alpha \left| X^{'(s)} \right|^2 + \beta \left| X^{'(s)} \right|^2 \right) \tag{4.46}$$

where α and β is the weighting parameter of contour's tension and weighting parameter of contour's rigidity respectively. The $X^{'}$ is the first derivative of $X(s)$ with respect to s while $X^{''}$ is the second derivative of $X(s)$ with respect to s.

And

$$\varepsilon_{ext} = E_{ext}\left(X\left(s\right)ds\right) \tag{4.47}$$

where E_{ext} is potential force field.

In order for curve $X\left(s\right)$ to minimize E, it must satisfy the Euler Equation (Kass et al.,1987).

$$\alpha X^{''}\left(s\right) - \beta X^{'''}\left(s\right) - \nabla E_{ext} = 0 \tag{4.48}$$

The Equation 4.48 can be expressed as a force balance equation:

$$F_{\text{int}} + F_{ext} = 0 \tag{4.49}$$

where $F_{\text{int}} = \alpha X''\left(s\right) - \beta X'''\left(s\right)$ and $F_{ext} = -\nabla E_{ext}$.

To solve Equation 4.48, the deformable contour associate X with t as follows:

To start designing a deformable contour, a new static external force known as *gradient vector flow* field is defined as in Equation 4.50 by substituting $-\nabla E_{ext}$ with $v\left(x,y\right)$. The substitution yields:

$$X_t\left(s,t\right) = \alpha X''\left(s,t\right) - \beta X'''\left(s,t\right) + v \tag{4.50}$$

The new external force can be computed from a number of image properties. In this Section, gradient of edge map $f\left(x,y\right)$ is chosen, as it is always pointing towards edge, which is desirable. It makes the deformable contour to converge to the stable configuration near the edges. Thus, $f\left(x,y\right) = -E_{ext}\left(x,y\right)$. Thus, gradient vector flow is the vector field $V\left(x,y\right) = \left[u\left(x,y\right), v\left(x,y\right)\right]$

that minimizes the energy functional.

$$\varepsilon = \int\int \mu \left(u_x^2 + u_y^2 + v_x^2 + v_y^2\right) + |\nabla f|^2 |V - \nabla f|^2 \, dxdy \tag{4.51}$$

The detail of numerical implementation of the GVF generation can be found in (Xu & Prince,1998) and (Xu & Prince,2000).

4.7 Summary

This chapter summarised the relevant theoretical and fundamental formulation of unsupervised colour thresholding, anisotropic diffusion and cubic spline interpolation for the development of the optic disc localisation and optic disc contour segmentation algorithms. The following chapter illustrates the algorithm in detail.

CHAPTER 5

OPTIC DISC LOCALISATION

5.1 Introduction

A novel algorithm for optic disc localisation proposed in this work searches and locates the center point of the optic disc in colour fundus images. Prior to the search, a number of optic disc candidates are generated. Each candidate consist of a set of connected pixels which satisfy the criteria of homogeneity in the red, green and blue colour bands. The search for the optic disc uses four dominant characteristics: its colour, the convergence point of major blood vessels, the size of the optic disc and its shape. The size of the optic disc is of critical importance. Observations made from available fundus image databases conclude that the size of optic disc is approximately $\frac{1}{8}$ of the size of a typical colour fundus image. This characteristic feature is used as the primary criteria in the proposed approach. In typical colour fundus images, the optic disc is always observed as a bright yellowish region. This is another representative characteristic of the optic disc and is used as the second criterion in the proposed algorithm. Apart from that, due to the fact that the optic disc is the only region that is traversed by a large number of major blood vessels, the convergence point of major blood vessels is incorporated as the third criterion to select the most probable optic disc region from a set of candidate regions. Finally, since the optic disc typically appears as a circular or ellipse region in most fundus images, the fourth criterion used is the shape of the optic disc. The four representative characteristics of optic disc is illustrated in Figure 5.1.

5.2 Generating Optic Disc Candidates

The localisation algorithm starts with a clustering process to generate a number of optic disc candidate regions. Each resulting cluster represents an optic disc candidate. Two algorithms are involved in generating the initial optic disc candidates regions: *Gaussian smoothing* and *unsuper-*

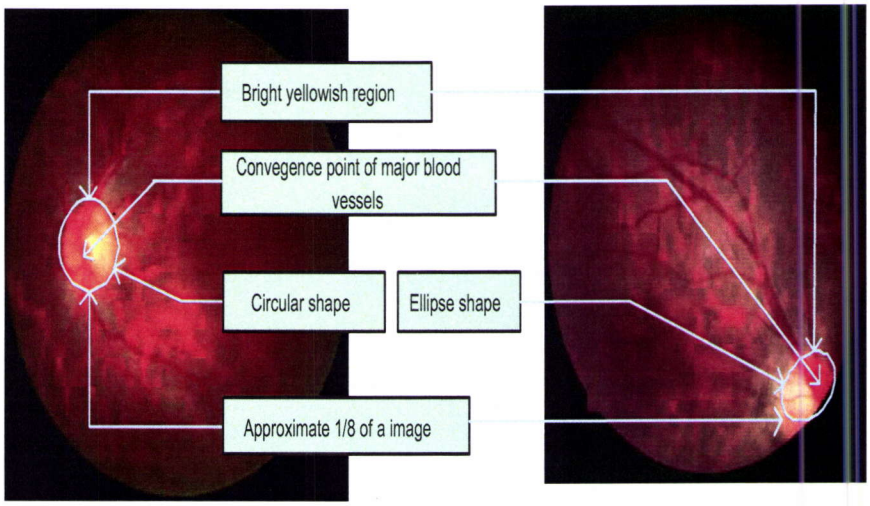


Figure 5.1: The four representative characteristics of optic disc are highlighted

vised colour thresholding. Both these approaches are described next.

5.2.1 Gaussian Smoothing

Gaussian smoothing is used on the colour fundus image to reduce noise and small features such as tiny blood vessels and microaneurysms. This produces more homogeneous regions in the image and aids the colour thresholding operation. Smoothing is performed by applying the normalised version of Gaussian smoothing operator on three primary colour bands respectively. The normalised version of Gaussian smoothing operator is given by:

$$G\left(x_i, y_i\right) = \frac{1}{2\Pi\sigma^2} e^{-\frac{x_i^2+y_i^2}{2\sigma^2}} \tag{5.1}$$

where x and y are the spatial information of image and subscript i denotes the colour band. σ is the standard deviation and is a spatial distribution that proportional to the size of neighbourhood of the Gaussian operator (M. Sonka & Boyle, 1993). Pixels that are more distant from the center of the operator exert smaller influence than those that are nearer to the center.

5.2.2 Unsupervised Colour Thresholding

Unsupervised colour thresholding (Y. Du & Thouin,2004) is a clustering-based thresholding method that is able to produce one or more representative clusters which are encoded in 3-bit binary code. The term "representative cluster" in this work refers to the cluster which lies within the optic disc region. It is the cluster that is either equal to or smaller than the size of the ground truth optic disc, and whose centroid is sufficiently close to the center of the ground truth optic disc.

Assume that the threshold values for the 3 primary colour bands, red, green and blue are denoted as t_r, t_g and t_b as per Equation 4.5. Each colour pixel is then specified by $p_{i,j} = (r_{i,j}, g_{i,j}, b_{i,j})$ with (i, j) as the spatial coordinate of the corresponding pixel. Each colour pixel is then thresholded into $\boldsymbol{p}_{i,j} = (\boldsymbol{r}_{i,j}, \boldsymbol{g}_{i,j}, \boldsymbol{b}_{i,j})$ according to Equation 4.8.

The binary thresholding operation results in each colour pixel being assigned a 3-bit binary value, denoted as (c_1, c_2, c_3). Using this 3-bit binary code, the colour fundus image is then clustered into eight possible regions, $\{C_k\}_{k=1}^{8}$ depending upon the 3-bit binary values that range from (0,0,0), (0,0,1)...(1,1,1). For each cluster, the cluster mean is computed. Refer to *Chapter* 4, *Section* 4.3 for more details on this algorithm.

Figure 5.2 illustrates two examples of colour fundus image to demonstrate the outcome of unsupervised colour thresholding. Figure 5.2 (a) shows the original colour fundus images and Figure 5.2 (b) shows the images in pseudo-colour after the colour clustering operation.

5.3 Optic Disc Selection

In this step, the task of selecting the best optic disc candidate is divided into four stages. Each stage measures one of the dominant characteristics of the optic disc. This process illustrated in the flow chart in Figure 5.3. Each stage is described in detail in subsequent sections.

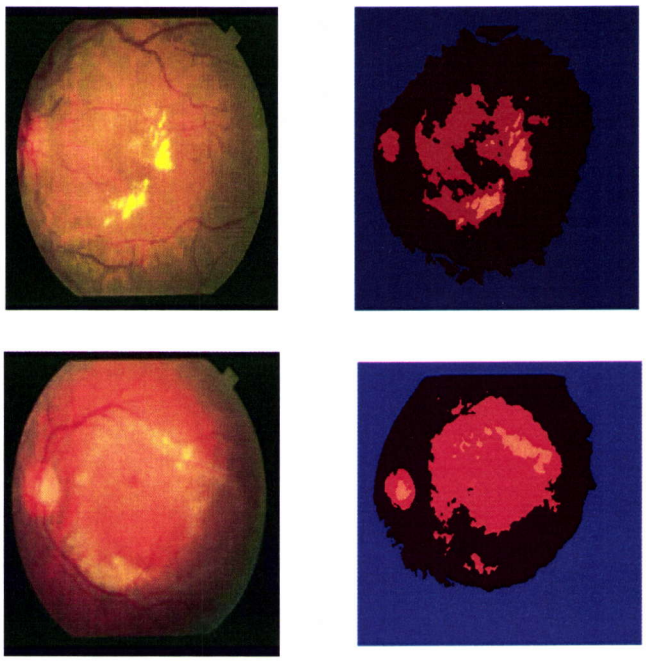


Figure 5.2: Each cluster is represented using its pseudo colour

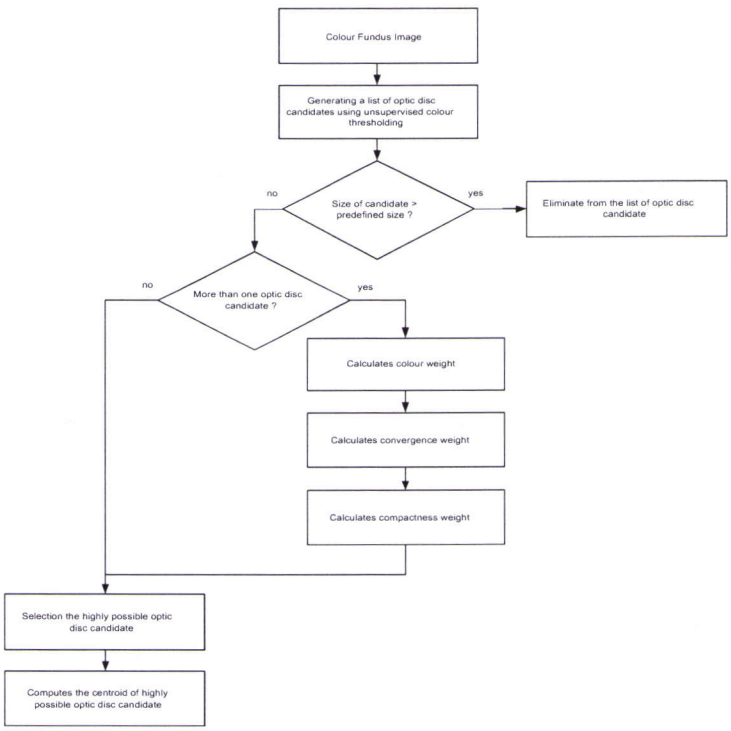


Figure 5.3: Overall flow of proposed optic disc localisation algorithm

5.3.1 Size

The size of each colour cluster should not exceed $\frac{1}{8}$ of the size of the fundus image. Clusters that do not satisfy this criterion are deleted. The formulation is shown in Equation 5.2:

$$C_k = \begin{cases} \text{remained if } C_k \geq T \\ \text{deleted if } C_k < T \end{cases} \tag{5.2}$$

where C_k is the k-th cluster represented by its respective bit code. The T is the $\frac{1}{8}$ of the size of the fundus image.

5.3.2 Colour Similarity

Each optic disc candidate region is then weighted based on the degree of similarity to a reference colour. The cluster that is most similar to the reference colour is assigned the highest value. The reference colour is determined by averaging the colour of healthy optic discs taken from a large database. The colour similarity measure is mathematically described as follows.

$$w_i^c = \sqrt{(r_i - r_r)^2 + (g_i - g_r)^2 + (b_i - b_r)^2} \tag{5.3}$$

where w_i^c is the colour weight of i-th colour cluster. r, g and b are the red, green and blue colour band from RGB colour model, and subscripted with r, g and b respectively. The subscript i refers to the reference colour.

Referring to Equation 5.3, the cluster that is most similar to the reference colour results in a smaller value, whereas the cluster that is least similar to the reference colour results in a larger value. The colour weight is then normalised into a range of between 0 and 1 and finally the cluster with the highest similarity to the reference colour is assigned the highest similarity value. This colour weight normalization is described in Equation 5.4.

$$W_i^c = 1 - \frac{w_i^c}{\sum w_i^c} \tag{5.4}$$

5.3.3 Convergence Point of Major Blood Vessels

In the green colour band of typical colour fundus images, the blood vessels are observed as long structures and consist of branches formed by minor vessels and small capillaries. The blood vessel structures are more visible in the green colour band compared to other colour bands, owing to better contrast. The red colour band is highly saturated blood vessels are only partially visible due to low contrast. The blue colour band often contains less important data, as well as other features such as haemorrhages and microaneurysms. Thus, only the green colour band is selected in the search for the convergence point of major blood vessels. Figure 5.4 illustrates the image contrast of the red, green and blue colour band for comparison, derived from the same colour fundus image.

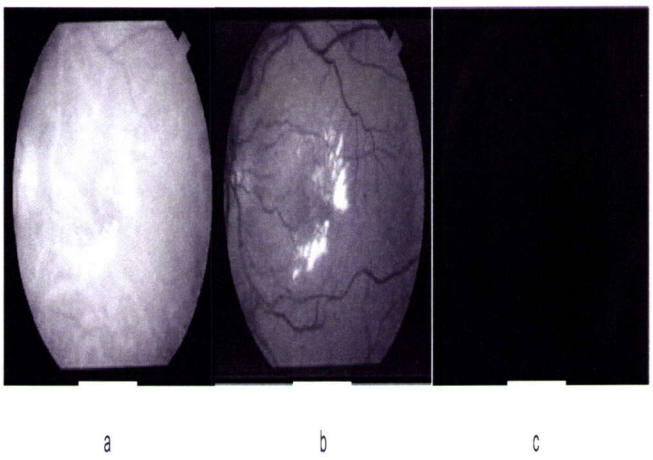


Figure 5.4: Image contrast for different colour bands (a) Red colour band (b) Green colour band (c) Blue colour band

Prior to locating the convergence point, a selective smoothing operation using anisotropic diffusion is applied on the green colour band of fundus image. Anisotropic diffusion is chosen for its ability to preserve major blood vessels while smoothing unnecessary features such as minor blood vessels, haemorrhages and microaneurysms. The detailed formulation of anisotropic diffusion is

given in Chapter 4 Section 4.4 Figure 5.5 illustrates a sample image before and after the anisotropic diffusion process.

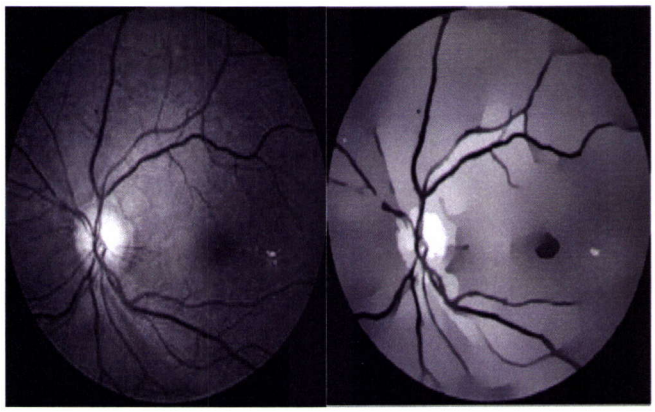


Figure 5.5: Sample image before and after the diffusion process. (a) Before the diffusion process (b) After the diffusion process

After preprocessing the fundus image using anisotropic diffusion, the intensities of fundus image are negated. After the image negation process, intensity line profile of blood vessels would appear as a Gaussian distribution. The mean of the Gaussian distribution corresponds to the centerline of the blood vessels. The location of the mean of the Gaussian distribution can be found by calculating the zero crossing of the first order derivative and the local maxima of the second order derivative of a one dimensional (1-D) window on the fundus image. As an example, refer to Figure 5.6. It can be noted that the center point of the blood vessel's intensity profile corresponds to the location of the zero crossing of the first order derivative and local maxima of the second order derivative.

Initially, points that lie at the zero crossing of first order derivative is considered as the centerline point candidates of the blood vessels structures. However, there is a possibility that these centerline point candidates (also known as point candidates) may not represent the actual centerline of blood vessels, but edges of exudates or marks from post laser treatment. Therefore, in order to increase the success rate of centerline structure detection, these point candidates are further verified using

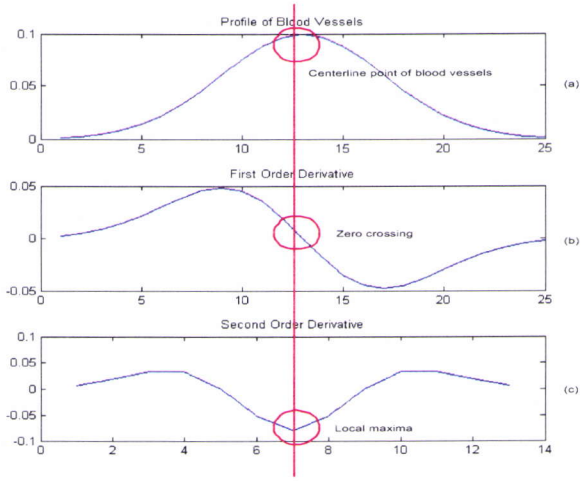


Figure 5.6: (a) Intensity Profile of Blood Vessel (b) First Order Derivative (c) Second Order Derivative

the local maxima of the second order derivative. The sliding window operation is repeated along four directions - horizontal, vertical, diagonal and counter-diagonal directions.

Although this approach only involves four principal directions, but it is found to be adequate to detect all the necessary centerlines of major blood vessels. In this report, the term *point candidates* is used to refer to the centerlines of major blood vessels.

Following this, the point candidates at arbitrary locations are grouped into set of connected centerline point candidates based on 8-neighbouring connectivity. These sets contain points belonging to major blood vessels as well as branches of blood vessels. However, the contribution of minor vessels (or branches) are not very important in determining the convergence point of major blood vessels. Therefore, these branches of vessels can be eliminated so that computation can be minimized. The following procedure illustrates the pruning process.

Assuming that the absolute magnitude of the second order derivative of point candidates for branch vessels is smaller than the major blood vessels, each set of connected point candidates is weighted as follows:

$$\nabla^2 I_{ij} = \sqrt{\left(\frac{\partial^2 I}{\partial x^2}\right)_{ij} + \left(\frac{\partial^2 I}{\partial y^2}\right)_{ij}} \tag{5.5}$$

$$w_i = n_i \cdot \sum_{j=1}^{n_i} \left|\nabla^2 I\right|_{ij} \tag{5.6}$$

$$s_i = \{r_1, r_2, ..., r_j\} \tag{5.7}$$

where w_i is the weight of i-th set of point candidates. n_i is the number of point candidates in the i-th set. I_{ij} is the intensity of j-th points that constitute i-th set, s_i. $\nabla^2 I_{ij}$ is the gradient magnitude of second order derivative of j-th point candidates in i set.

The value of w_i is then normalised to a range of 0 to 1 in Equation 5.8.

$$W_i = \frac{w_i}{\sum w_i} \tag{5.8}$$

In this procedure, only eight sets of point candidates, corresponding to eight major blood vessels are selected in the search for the convergence point of major blood vessels. Four point candidates with highest weights are selected from the upper and lower region of the fundus image respectively. The upper section is defined as starting from the top edge of the image to $\frac{7}{10}$ of the height of fundus image, and lower section is defiend as ranging from $\frac{3}{10}$ of the height of fundus image to the bottom edge of the image. Figure 5.7 illustrates the eight point candidates.

In order to separate the major blood vessels from its branching vessels, the end points are identified and extracted at first. This is performed by means of 3x3 window with the point candidate at the center of the window. The centered point candidate is considered as the end point if there exists only a single neighbouring point. The detected end points are marked with

Figure 5.7: Eight sets of valley points at the upper and lower section of fundus image

thick gray circles, as in Figure 5.8(a). With these detected end points, distances between any end points are calculated and stored. Next, with the similar procedure of end points detection, the branch point can be detected as well by searching the number of neighbouring points. The centered point underlying the 3x3 window is detected as branch point if two or more neighbouring points are detected. These branch points are marked with thin gray circles, as illustrated in Figure 5.8(a). Then, these branch points are removed and consequently the structure of blood vessels breaks into a number of wavy lines. Two types of wavy lines are formed, namely a line enclosed by one end point and one branch point (known as line α), and a line enclosed by two branch points (known as line β), as shown in Figure 5.8(b).

In the pruning process of the initial point set candidates, only α line types are taken into account. This procedure results in a long structure that corresponds to major blood vessels which traverse across the optic disc or grow towards the optic disc. In the pruning process, short α lines are prone to be eliminated. Only twoα lines that are the longest (having maximum distance between end points) are preserved. Finally, all the remaining lines, including both α and β lines are linked together by a simple dilation morphological operation. The similar procedure is repeated for other sets of point candidates.

In the subsequent step, the interpolation procedure is performed. The end point of each sets of connected point candidates are used as the seed point or starting point. The direction of the interpolation is determined by averaging the directions of the N-backtracked seed points. In the backtracking procedure, the number of pixels considered would be half the number of a

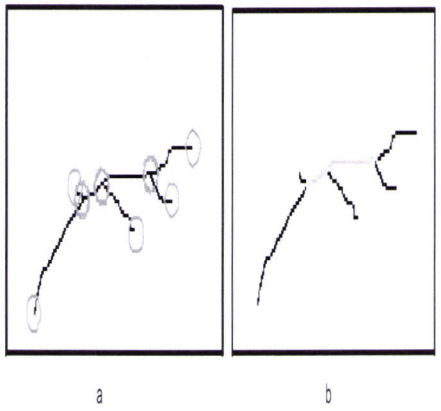


Figure 5.8: (a) End points and branch points are marked with thick gray circles and thin gray circles

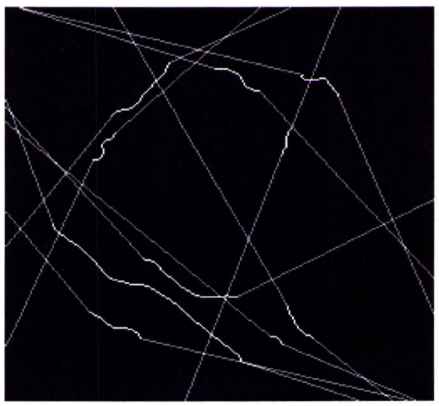


Figure 5.9: A number of crossed points, crossed by two interpolated lines are found

given point candidate. The backtracking and interpolation procedures are repeated for all the seed points. Using the interpolated lines, a number of intersection points can be obtained. The intersection point is the point crossed by any two interpolated lines, as illustrated in Figure **??**. Each intersection point is weighted by Eq 5.6 and is mathematically expressed as

$$W_{k,j} = W_{s_k} + W_{s_j} \tag{5.9}$$

where $W_{k,j}$ is the weight of convergence point crossed by k-th and j-th set of connected valley points. W_k and w_j are the weights of k-th and j-th set of connected valley points respectively.

Consequently, the intersection point with the highest weight is chosen as the estimated convergence point. The convergence point, $\tilde{W}$ is defined as follows:

$$\tilde{W} = max\{W_{i,j}\} \tag{5.10}$$

where i and j are the two different sets of connected valley points.

5.3.4 Shape

The common shapes of the optic disc as observed in most colour fundus images are either circular, oval or ellipse. The ellipse form of optic disc appears at the border of fundus image where optic disc is partially visible. At the temporal and nasal part of fundus image, the optic disc appears as an incomplete circle as it obscured by major blood vessels. To quantitatively measure the shape of the optic disc, a method known as 'compactness' is used to measure the variability of its shape. Compactness is a shape descriptor that is independent of linear transformation, and is given by M. Sonka & Boyle (1993):

$$compactness = \frac{(region\_border\_length)^2}{area} \tag{5.11}$$

The most compact region is the circle. In our approach, pixel clusters are weighted according to the region compactness.

5.4 Optic Disc Center Point Detection

The center point of the chosen optic disc candidate is obtained using the center of gravity (COG) M. Sonka & Boyle (1993). COG, a measure calculated using geometric moments is given as :

$$x_c = \frac{M_{10}}{M_{00}} \quad , y_c = \frac{M_{01}}{M_{00}} \tag{5.12}$$

where x_c and y_c are co-ordinates of the center of gravity of the optic disc. M_{00}, M_{10}, and M_{01} are the zero order moment which represents the region area, first order moment about axis x and first order moment about y axis respectively.

5.5 Experimental Approach

An experiment with a set of 81 colour fundus images available in *STARE Project Website* (2003) are carried out to gauge the performance of proposed optic disc localisation algorithm. In the available fundus images database, optic discs can be distinguished as circular or oval-like in shape, and in arbitrary locations. In order to quantitatively measure the optic disc location, the center point of the groundtruth or gold standard optic disc is detected. The contour of the true optic disc in each fundus image is manually drawn or labelled by an ophthalmologist using a simple graphic package. Next, the groundtruth optic disc center point is obtained by computing the center of gravity (COG) as in Eq 5.12.The performance of the proposed algorithm is measured by comparing the resulting center point and the groundtruth center point. The proposed method, known as Dart Method provides more meaningfully data, compared to the existing means of measuring as found in Hoover & Goldbaum (2003). The concise explanation of the Dart Method will be given in subsequent section.

5.5.1 Dart Method

An optic disc is deemed successfully detected (also known as true detection) if the resulting center point lies within the true optic disc region. All the true detection results are further graded into three grades,i.e. excellent, good and poor. Three grades are determined based upon the proximity of the resulting center point to the groundtruth center point. The performance of optic disc localisation is classified as '*excellent*' if the location of resulting center point lies within 40% proximity of groundtruth center point. Similarly, '*good*' indicates that the resulting center point

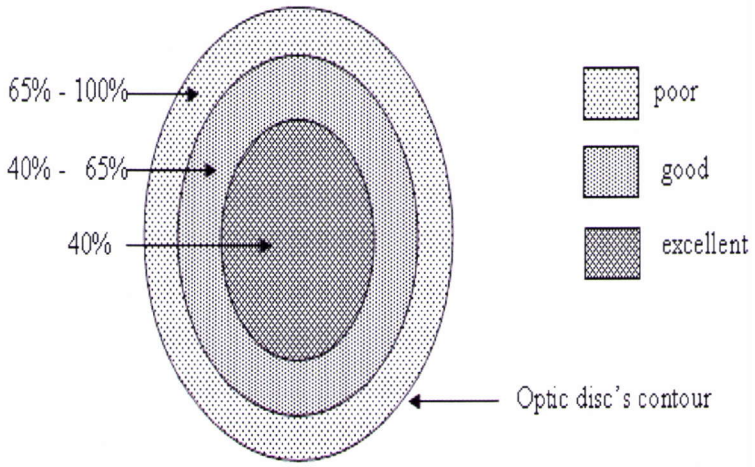


Figure 5.10: Dart Method: Quantitative measure of performance for optic disc localisation

Table 5.1: Results of fuzzy convergence and proposed algorithm

Method	True Detection	False Detection
fuzzy convergence	72	9
proposed	70	2

is located with 40% to 65% proximity of groundtruth center point, where '*poor*' performance indicates the resulting center point lies within 65% to 100% proximity of groundtruth center point. Figure 5.10 graphically illustrates the quantitative measure of the Dart Method.

5.6 Results and Discussion

The proposed optic disc localisation algorithm is tested on 81 colour fundus images that are freely available via *STARE Project Website* (2003). The provided fundus image database consists of 31 healthy fundus images and 50 diseased fundus images in 605x700 pixels resolution. The optic disc is visible in all the 81 colour fundus images, but there are 14 images in which the optic disc is only partially visible at the border of fundus image. In addition, there are 5 images in which optic disc is completely obscured by haemorrhages. The result of the proposed algorithm and the fuzzy convergence by Hoover & Goldbaum (2003) are tabulated in Table 5.1.

Out of 81 fundus images, the proposed optic disc algorithm outperforms the fuzzy convergence, proposed by Hoover & Goldbaum (2003), resulting in a 97.53% success rate compared to 88.89%

in Hoover & Goldbaum (2003). In addition, the fuzzy convergence Hoover & Goldbaum (2003) fails in 9 fundus images (including 2 images with no hypothesis of optic disc location), but for the same 9 images, our algorithm results in correct optic disc localisation. However, both approaches fail in 2 fundus images in the dataset. Figure 5.11 shows the comparative outcomes on 7 fundus images obtained by fuzzy convergence method and our proposed method. In these selected fundus images, it should be noted that the approach proposed in Hoover & Goldbaum (2003) fails to give the accurate resulting center point.

5.7 Summary

The experiment results indicate that the proposed feature vector approach provides superior results in optic disc localization as compared to recently reported approaches. The weighting factor for each measured characteristic plays a vital role to ensure that the optic disc localisation remains effective and robust.

The next chapter discusses the optic disc contour segmentation with combination of morphological operation, anisotropic diffusion, spline interpolation and Gradient Vector Field (GVF) Snake in detecting the accurate boundary of optic disc region.

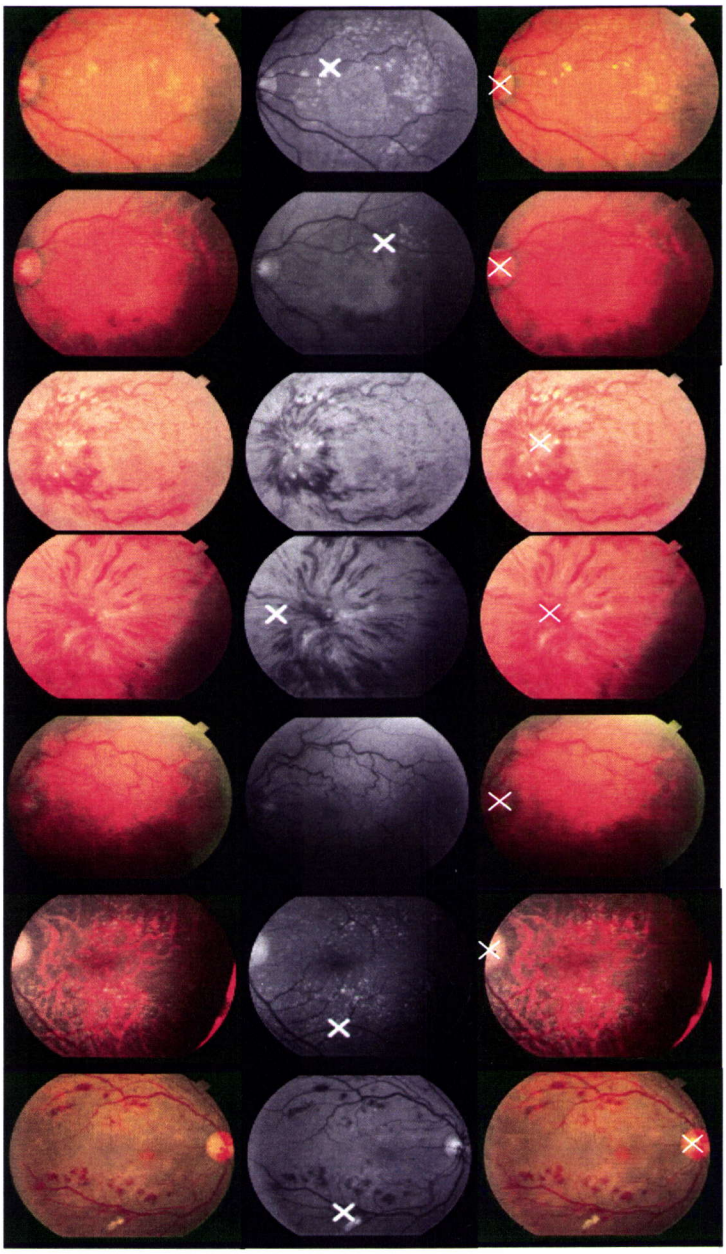


Figure 5.11: Comparison of results between fuzzy convergence and the proposed algorithm. Column (a) shows the original fundus images. Column (b) shows the localisation result by fuzzy convergence. Along that column, the 3^{rd} and 5^{th} images show the inconclusive results with no resulting center point, while the rest show the wrong localisation results. Column (c) shows the localisation results by proposed algorithm

CHAPTER 6

OPTIC DISC CONTOUR SEGMENTATION

6.1 Introduction

Optic disc contour segmentation is a prerequisite in the analysis of fundus images. However, as presented in **Chapter 2 Section 2.4**, there are varius challenges in optic disc contour segmentation. This chapter proposes a robust approach to generate the initial configuration for active contour which is ultimately used to segment the optic disc contour.

6.2 Overview

The proposed optic disc contour segmentation starts with the removal of blood vessels in the green band of a fundus image using a combination of morphological operations and anisotropic diffusion. The blood vessel structures are coarsely detected by applying a closing operation with a disk-like structuring element. Following this, the blood vessel structures in the red band of the fundus image are smoothed using the anisotropic diffusion method to create a homogeneous region within the optic disc. After the removal of blood vessels from the fundus image, the next step is to construct a coarse estimate of the optic disc contour. The method of constructing the coarse optic disc contour adopted in this work is inspired by Hiroyuki et al. (1998). The algorithm proceeds as follows:

Stage 1: Elimination of blood vessels structures

- Detection of blood vessels by means of morphological operation
- Elimination of blood vessels by means of anisotropic diffusion

Stage 2: Construction of coarse optic disc contour

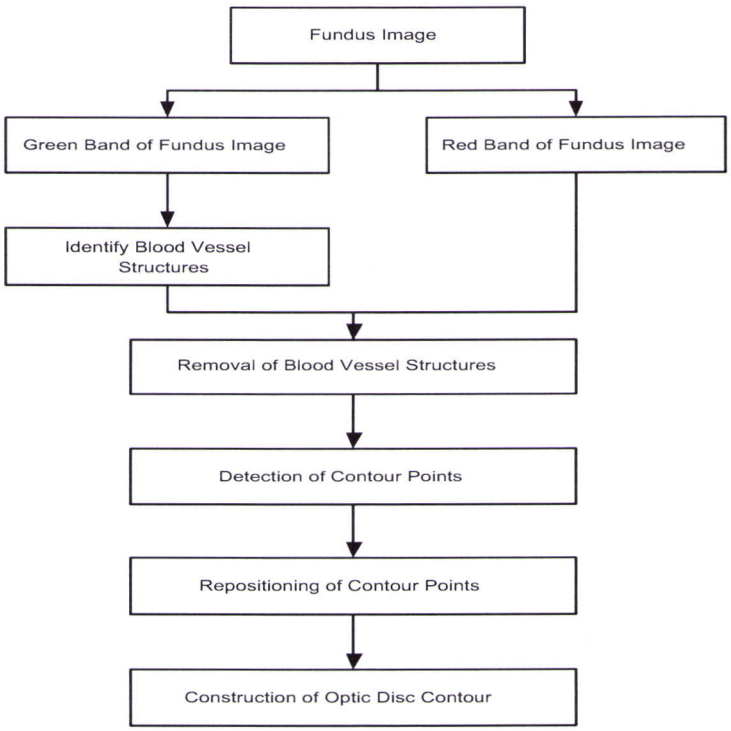


Figure 6.1: Overall flow of proposed optic disc contour segmentation algorithm

- Collection of contour point candidates
- Repositioning of outlier contour points by means of spline interpolation

The optic disc contour segmentation procedures is graphically illustrated in Figure 6.1.

6.2.1 Choosing the Appropriate Colour Band

The three primary colour bands of a fundus image provide different information about the anatomical and pathological structures of the eye. For instance, in the red colour band, the optic disc is observed as a high intensity region. However, the red colour band alone is not sufficient for optic disc contour segmentation as it is saturated and reveals other high intensity patches that do not correspond to the actual optic disc region. This is shown in Figure 6.2(a).

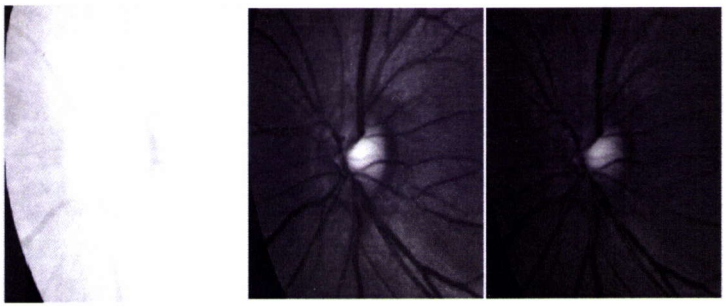


Figure 6.2: A typical fundus image shown in the three primary colour bands. (a) Red band (b) Green band (c) Blue band

6.2.2 Morphological Operation For Blood Vessel Detection

The traversing blood vessels near the optic disc may complicate the segmentation of optic disc contour and must be removed. Therefore, prior to the segmentation procedure, the blood vessels structures are first identified. A closing operation using a disk-like structuring element is applied on the green band of the fundus image. The closing operation consists of dilation followed by erosion morphological operation.

An absolute intensity difference image map is obtained by subtracting the morphologically processed image with the original image. With these absolute intensity differences, a cumulative histogram is constructed. The response variable whose number of occurrences exceed 90% of the maximum occurrence in cumulative histogram is taken as the threshold. Any absolute intensity differences that exceed this threshold is assumed to be pixels belonging to the blood vessel structures. Figure 6.3(a) and (b) illustrates the image prior to and after the morphological operation. Figure 6.3(c) shows the identified blood vessel structures.

6.2.3 Removal of Blood Vessel Structures

A selective filtering or non-linear smoothing method, called the anisotropic diffusion is applied on the red band of the fundus image. The anisotropic diffusion is applied in order to smooth unwanted data such as small and tiny blood vessels, haemorrhages, microaneurysms, while preserving important features such as optic disc contour and blood vessels. However, owing to the existence of blood vessels at the vicinity of optic disc, pixels belonging to blood vessels can be misdetected as

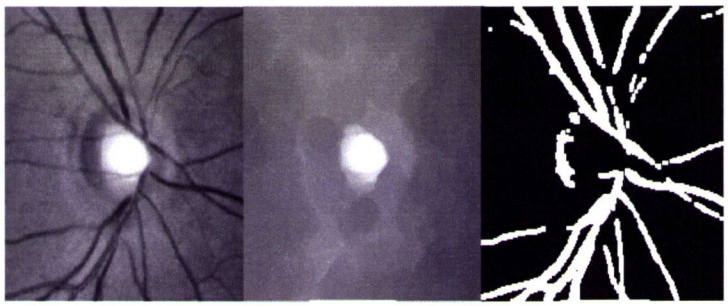


Figure 6.3: (a) The green band of a fundus image before morphological operation. (b) The green band of a fundus image after morphological operation. (c) The identified blood vessel structures

optic disc contour points. Therefore, the blood vessel structures are eliminated using anisotropic diffusion guided by morphological operation. The anisotropic diffusion is adapted in such a way that the diffusivity function of the pixels belonging to blood vessel structures are set to unity. This allows more smoothing within the blood vessel structures and at their edges for all iterations of the anisotropic diffusion process. The modified diffusivity function is shown in Equation 6.1.

$$(\nabla I)\begin{cases} 1 & \textit{within blood vessel structures} \\ \frac{1}{1+\frac{|\nabla I|^2}{k}} & \textit{not within blood vessel structures} \end{cases} \tag{6.1}$$

The details regarding the notation used in Equation 6.1 on anisotropic diffusion can be found in Perona & Malik (1990).

After several iterations of the diffusion process, small features and blood vessel structures are effectively removed, while optic disc contour is still well preserved. This facilitates the task of obtaining the true optic disc contour. Figure 6.4(a) and (b) show the optic disc region before and after removal of blood vessel structures using Equation 6.1.

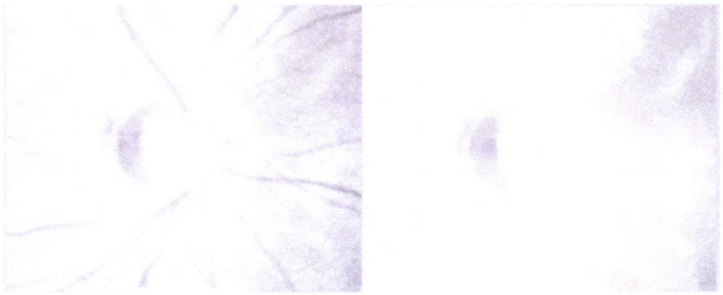


Figure 6.4: (a) Before being processed by anisotropic diffusion (b) After being processed by anisotropic diffusion

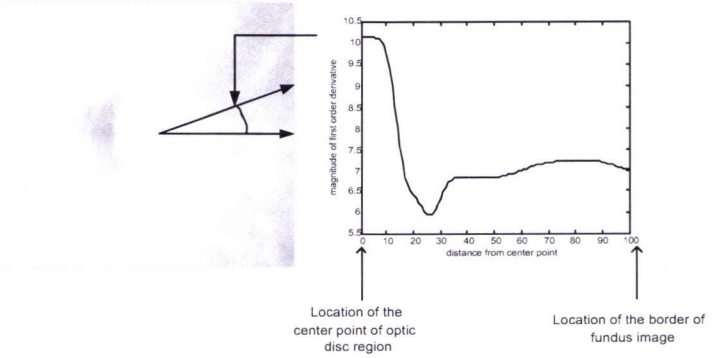


Figure 6.5: Detection of contour points by radial one-dimensional profile. (a) One-dimensional intensity profile is drawn by passing through the resulting center point. (b) The magnitude of first order derivative corresponds to the drawn profile in (a).

6.2.4 Optic Disc Contour Points Detection

A set of one-dimensional (1-D) intensity profiles that radially pass through the center point of the optic disc region at fixed intervals are obtained. The angular interval, $\nabla\theta$ between one intensity profile to the next intensity profile is set to one degree, 1^0. The first order derivative of the intensity profile is then calculated. From this first order derivative, local maxima points are identified. Subsequently, the local maxima point closest to the resulting center point is chosen as the contour points. This procedure is repeated for all the 360 intensity profiles. Figure 6.5 illustrates one of the intensity profile at the $\nabla\theta$ angle.

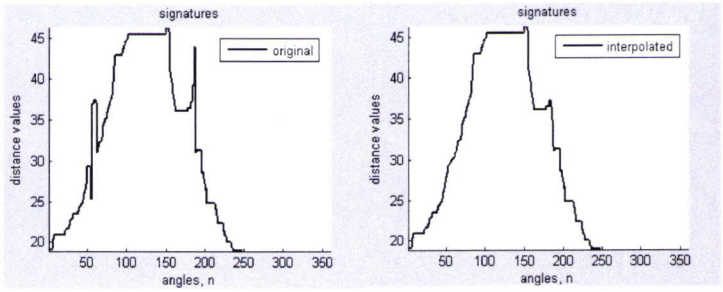


Figure 6.6: The signature representation for a detected contour points. (a) Original signature representation (b) Modified signature representation after repositioning by spline interpolation

6.2.5 Repositioning of Contour Points

During the contour point detection process, there is a possibility that some detected points do not represent the actual optic disc contour. These points are known as *outliers*. The detection of outliers is acquired from the signature profile. A signature is a one-dimensional representation of a boundary. A signature representation is drawn by plotting the distance between the resulting center point to all the obtained contour points as a function of its angle Gonzales & Woods (2002). From this signature representation, points that cause sudden or abrupt variations are considered outliers. The outliers are detected by searching for high gradient magnitudes of the first order derivative. Once identified, the positions of these outliers are adjusted to a new position (re-positioning) which better reflects the true position of the optic disc contour. This new position is estimated using spline interpolation de Boor Carl (1978). Figure 4.5(a) and (b) illustrate the signature representation of the detected contour points prior to and after the repositioning process using spline interpolation.

To provide more clarity, the obtained optic disc contours before and after the repositioning by spline interpolation are depicted as a two-dimensional image in Figure 6.7 (a) and (b) respectively.

6.2.6 Construction of Coarse Estimate Optic Disc Contour

A coarse estimate of the optic disc contour is constructed by connecting the re-positioned contour points using the *Snake* approach. The final outcome of proposed optic disc contour segmentation

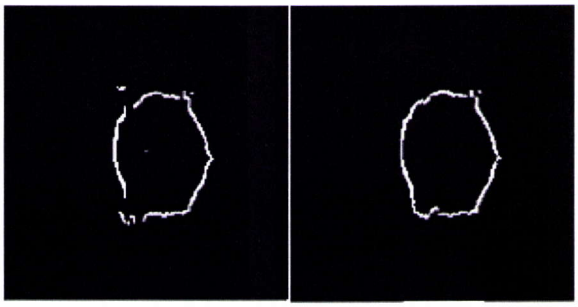


Figure 6.7: A two-dimensional representation of the detected contour points. (a) Before repositioning (b) After repositioning by spline interpolation

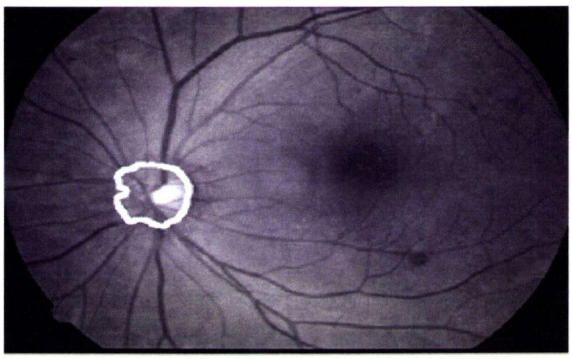


Figure 6.8: The coarse estimate of optic disc contour is highlighted

is depicted in Figure 6.8 as below.

6.3 Results and Discussion

A set of 23 colour fundus images are tested to examine the performance of proposed algorithm. The results obtained by the proposed algorithm are compared to groundtruth images using the overlap measure A. Osareh & Markham (2002*b*). This simple and effective quantitative measure is given by Equation 6.2.

$$M = \frac{N(R \cap T)}{N(R \cup T)} \times 100\% \tag{6.2}$$

where **R** and **R** are the two regions to be compared, and $N(\cdot)$ is the amount of pixels in the

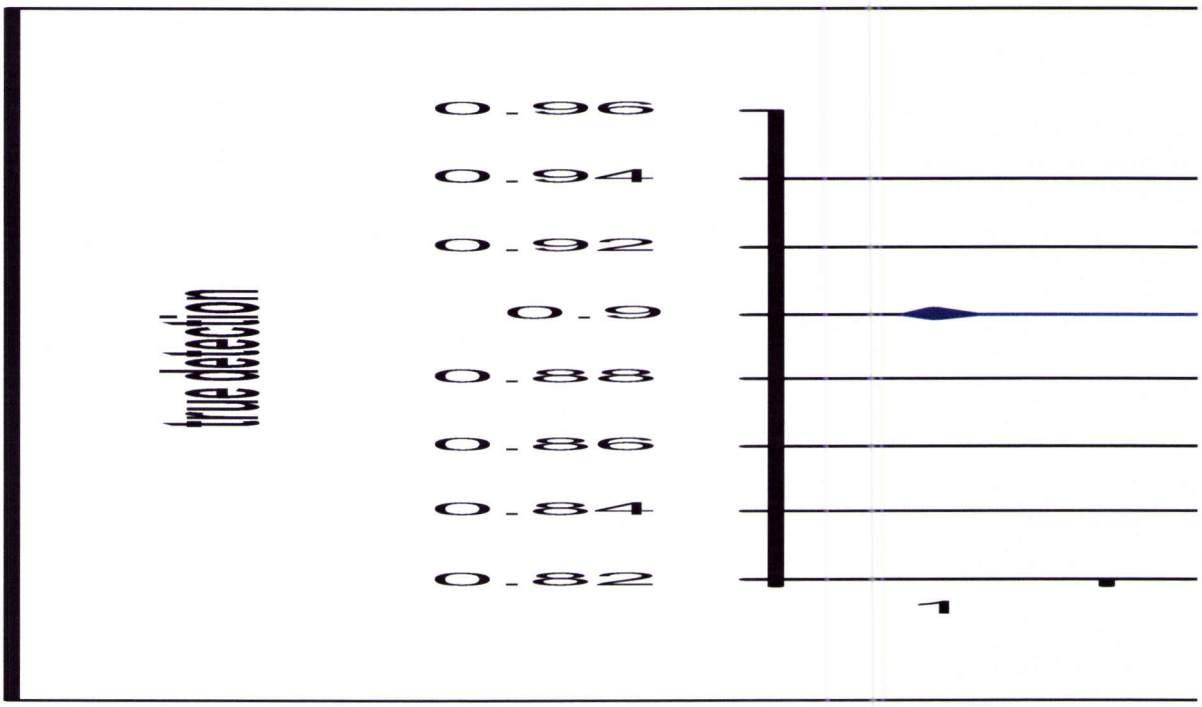


Figure 6.9: Performance of optic disc contour segmentation for each fundus image

set. The performance of the optic disc segmentation for each fundus image is illustrated in the line graph in Figure 6.9. The proposed algorithm achieves a minimum of 87% and a maximum of 95% accuracy, while the overall accuracy is 92% for this particular dataset.

One of the advantages of the proposed algorithm is that it is more reliable in detecting contour points for various shapes of optic discs, as compared to the method reported in Hiroyuki et al. (1998). In Hiroyuki et al. (1998), the outliers are identified reliably based upon the *Mahalanobis* distance. However, the weakness of this approach is that when an optic disc contour appears as an ellipse, some of the true contour points may fall outside the statistical range (computed as mean + standard deviation) and possibly be considered as outliers. The said approach only reliably eliminates outliers if the optic disc contour appears as a regular circle. However, in reality, most optic discs do not appear as regular circles. This is especially so in the fundus image database currently used for this work. Comparative results for outliers elimination contour points elimination between Hiroyuki et al. (1998) and the proposed method are illustrated in Figure 6.10.

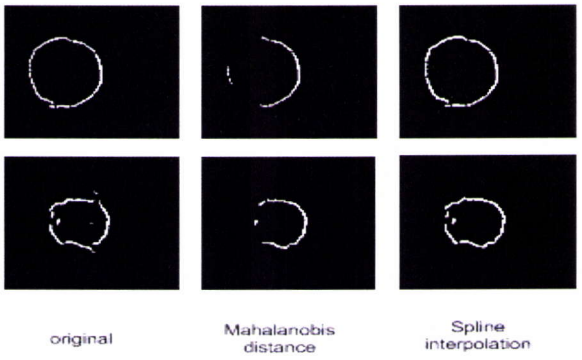


Figure 6.10: Performance of optic disc contour segmentation for each fundus image

As observed in Figure 6.10, some of the contour points are wrongly identified as outliers and therefore eliminated. This is the vital weakness of using the *Mahalanobis* distance metric. In the proposed approach, only contour points that produce sudden variation in distance in signature representation are considered as outliers, as illustrated in the last column of Figure 6.10.

The assessment of accuracy for optic disc contour segmentation is not an easy task. The hand labelled contour by human graders are not perfect and it may not correspond well to the actual optic disc contour. Hence, achieving an accuracy of 100% using the overlap measure is impossible since neither manually labelled optic disc contour nor the automatically segmented contour may correspond to the actual optic disc contour. In this case, a result of more than 90% is deemed to be fairly accurate.

6.4 Conclusions

The accuracy and effectiveness of proposed method in comparison to human graders have been evaluated using a set of 23 colour fundus images and the results are very promising. With the accuracy of 92%, it can be concluded that the estimated optic disc contour is very close to the actual optic disc contour. In subsequent work, this result may be further improved by using the estimated optic disc contour as an initial configuration for the Gradient Vector Flow (GVF) Snake to find a more accurate optic disc contour. The applicability of the proposed automatic optic disc contour segmentation for the practical use in a cost-effective mass screening system is investigated.

6.5 Summary

A novel algorithm of using combination methods of morphological, anisotropic diffusion and spline interpolation to generate a coarse or initial configuration of GVF Snake has been presented in this chapter. For most cases, the accuracy segmentation of up to 90% is very much acceptable.

CHAPTER 7

CONCLUSIONS

7.1 Summary of Research

In this research, an automated screening system that robustly detects the optic disc in digital fundus images is presented. The system is capable of detecting the optic disc through a wide range of uncertainties such as arbitrary locations, varying image conditions and colour. No elaborate parameter tuning or analytical modelling is required prior to deployment.

This research emphasises on optic disc localisation and optic disc contour segmentation. The significant contributions of this work is the introduction of novel algorithms for optic disc localisation and optic disc contour segmentation for realising an accurate segmentation of optic disc contour, and a recommendation for quantitative measurement of obtained resulting center point. The computational time is not considered in this work and the scope of the research is restricted to a set of colour fundus images, captured using CCD fundus camera.

In this research, a feature vector approach is proposed as the optic disc localisation algorithm. The proposed feature vector approach utilises the dominant characteristics of typical colour fundus images. They include colour, size, shape and obscuration of major blood vessels. Prior to selecting the optic disc that possess all or most of the dominant optic disc's characteristics, a number of optic disc candidates is generated. This is done using an unsupervised colour thresholding approach. Following this, each optic disc candidate is gauged and weighted according to the dominant optic disc's characteristics. The optic disc with the highest weight is most similar to a typical optic disc.

Past methods on optic disc localisation were assessed in terms of distance between resulting center point and groundtruth center point for small database of fundus images. The reported quantitative measurement of performance for optic disc localisation was often evaluated in terms

of either *true* detection or *false* detection. Most often, it was evaluated within 60 pixels distance basis. The *true* detection is only assigned if the resulting center point lies within 60 pixels from the groundtruth center point. If the resulting center point lies further than 60 pixels, it is then assigned as a *false* detection. However, this method of measurement is extremely rigid and does not account for the variation often observed in the optic disc's size. For instance, when a larger optic disc region is considered, even though the resulting center point lies further than the 60 pixels radius, it can still be considered as close to the groundtruth. Similarly, when the size of optic disc region is smaller than 60 pixels, the resulting center point that falls outside or at the boundary of optic disc region should not be assigned a *true* detection. In the proposed of semantic quantitative measure of optic disc localisation, a *Dart_Method* which takes into account the varying size of the optic disc region is considered.

Optic disc contour segmentation in colour fundus image analysis is important to distinguish the false positive optic disc region from the detected exudates candidates and macula localisation. A novel automatic initialisation of the *Snake* method is introduced. Using the center point obtained from the optic disc localisation, the proposed algorithm provides the initial configuration that approximates the actual optic disc contour. The obvious advantage of providing initial configuration using *Snake* is to prevent or reduce the probability of wrong convergence of the Snake to nearby lesions at the vicinity of optic disc region.

Snake method works well on gradient image. However, the Snake method become less effective when the boundary of the optic disc is fragmented into multiple cracked edges caused by traversing blood vessels. In order to remove the interfering major blood vessels that lie within optic disc region, a pre-processing operation is carried out. The pre-processing uses the combination of a morphological operation followed by anisotropic diffusion.

The result presented in this work is very promising. The system could be integrated into a mass screening system for real application of diabetic retinopathy and used by non-expert to diagnose patients who require referral to an ophthalmologist for further investigation and treatment. This contributes to cost savings since the number of fundus image to be manually reviewed by medical

professionals is dramatically lowered.

7.2 Practical Contributions of the Work

In essence, the practical contribution of this work is the successful integration and implementation of various approaches into optic disc detection system. Specific contributions of the work are as follows:

- the use of unsupervised colour thresholding to generate a list of optic disc candidates prior to selection of best candidate. This is contrary to the published method, i.e. fuzzy C-means clustering.
- the development of feature vector-based optic disc localisation approach to find the location of the optic disc in fundus images.
- the design and implementation of the *Dart* method to measure the performance of the optic disc localisation quantitatively and semantically for various appearances of the fundus image.
- A robust initial configuration of the *Snake* algorithm is developed. A morphological operation to detect the structure of blood vessels, followed by anisotropic diffusion for removing the interfering blood vessels while preserving or enhancing the boundary of optic disc. This approach demonstrates an improvement over previously published Snake initialisation.

7.3 Recommendations for Future Work

This research has shown promising results. However, there is still much possibility for improvement. As it usually happens in research, there are still many other approaches to be tried and developed given time and effort. Following are the recommendations proposed for improving the optic disc detection and mass screening for diabetic retinopathy.

- Develop a user interface program to ease ophthalmologist to define and draw the boundaries of anatomical and pathological structures for the purpose of collecting useful data.

- A series of algorithms to initially classify fundus image into either good quality and poor quality and detecting the region of poor quality are developed. Since poor quality of fundus image owing to noise and colour distorted may jeopardize the subsequent of optic disc localisation.

- It would be an advantageous to develop a colour-based active contour for optic disc contour segmentation. The gray level-based active contour can be extended to colour fundus image as the colour gradient information is obtained from the appropriate colour edge colour indication function. In this case, a suitable colour space for defining the colour edges should be carefully chosen.

- Optimization should be done in stand alone program to achieve an adequate processing rate for real application as the Matlab platform is relatively slow.

- In this research, groundtruth fundus image of one ophthalmologist is used to measure the performance of optic disc localisation as well as optic contour segmentation. Since ophthalmologist herself or himself may make mistakes, it would be worthwhile to use fundus image that have been commonly agreed or accepted by several ophthalmologists. It may be reasonable to measure variance among a number of ophthalmologists and variance between computed-generated result and groundtruth image.

- A system for detecting other lesions, i.e hard exudates, soft exudates, haemorrhages and microaneurysms for purpose of detection of early sign of retinopathy and retinopathy requiring imminent treatment.

- Detection of neovascularization at optic disc for screening of proliferative diabetic retinopathy. The severity of retinopathy can not be screened at the stage of background retinopathy and maculapathy.

- A method to compare and generate report on the changes in the number and distribution of lesions from one fundus image to another fundus image. This allows the progression of disease to guide a successful management and treatment of underlying diabetes.

BIBLIOGRAPHY

A. Osareh, M. Mirmehdi, B. T. & Markham, R. (2002*a*), Classification and localisation of diabetic related eye disease, In the Proceedings of 7th European Conference on Computer Vision, Springer LNCS 2353, pp. 502–516.

A. Osareh, M. Mirmehdi, B. T. & Markham, R. (2002*b*), Colour morphology and snakes for optic disc localisation, The 6-th Medical Image Understanding and Analysis Conference, BMVA Press, pp. 21–24.

Acton, S. (2000), 'Locally monotonic diffusion', *IEEE Transactions On Signal Processing* **48**(5), 1379–1389.

Baker, K. C. (2003), Locating retinal landmarks in fundus photography, Optics in the Southeast Poster Session, University of Central Florida, Orlando.

Bimbo, A. D. & Pala, P. (1989), 'Image retrieval by elastic matching of user sketches', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **19**(2), 121–132.

Bow, S. T. (19992), *Pattern Recognition and Image Processing*, Marcel Dekker, Inc New York.

Carlotto, M. J. (1997), 'Histogram analysis using a scale-space approach', *IEEE Transactions on Pattern Analysis and Machine Intelligence* (9), 121–129.

Dan, R. (2001), 'Anatomy of the eye'.
URL: *http://www.mdsupport.org/anatomy.html*

de Boor Carl (1978), *A Practical Guide to Splines*, Vol. 27 of *0-387-90356-9*, Springer-Verlag.

Detecting And Monitoring Diabetic Retinopathy - Digital Healthcare Wins Prestigious Award For Innovative Healthcare Solution (2006).
URL: *www.medicalnewstoday.com/medicalnews.php?newsid=37660*

Du, D. Y. & C.Chang (2003), An unsupervised approach to colour video thresholding, Vol. 3, Proceedings of IEEE International Conference on Multimedia and Expo (ICME), pp. 337–374.

F. Mendels, C. H. & Thiran, J. (1999), Identification of the optic disc boundary in retinal images using active contours, In the Proceedings of the Irish Machine Vision and Image Processing Conference, pp. 103–115.

Gerig, G., Kubler, O., Kinnis, R. & Jolesz, F. A. (1992), 'Non linear anisotropic diffusion of mri data', *IEEE Transactions on Image Processing* **11**(2).

Goh, K. G., Hsu, W., Lee, M. L. & Wang, H. (2000), *Medical Data Mining and Knowledge Discovery*, Springer-Verlag, chapter ADRIS: an Automatic Diabetic Retinal Image Screening system, pp. 181–210.

Gonzales, R. C. & Woods, R. E. (2002), *Digital Image Processing*, number 0-201-18075-8, second edn, Prentice Hall.

Grau*, V., Mewes, A. U. J., Alcaniz, M., Kikinis, R. & Warfield, S. K. (2004), 'Improved watershed transformation for medical image segmentation using prior information', *IEEE Transactions On Medical Imaging* **23**(4), 447–456.

Hiroyuki, Y., Bilgin, K., David, D. S., Abdullah, C., Omer, O. & Ahmad, S. (1998), Segmentation of liver tumors in ultrasound images based on scale-space analysis of the continuous wavelet transform, Vol. 2, in Proceedings of the Ultrasonics Symposium, pp. 1713–1716.

Hoover, A. & Goldbaum, M. (2003), 'Locating the optic disc in a retinal image using fuzzy convergence', *IEEE Transactions on Medical Imaging* **22**(8), 951–958.

Jalali, S. (2003), Retinal detachment, Vol. 16, Community Eye Health Journal, pp. 25–26.

Kanski, J. (1997), 'Diabetic retinopathy, clinical ophthalmology'.

Kass, M., Wiktin, A. & Terzopoulos, D. (1987), 'Snakes: Active contour models', **1**(4), 321–331.

Kayani, H., Rehan, N. & Ullah, N. (2003), 'Frequency of retinopathy among diabetics admitted in a teaching hospital lahore', *Journal Ayub Med Coll Abbottabad* **15**(4).

K.Fritzsche, A.Can, Shen, H., C.L.Tsai, J.Turner, Tanenbaum, H. L., Stewart, C. & Roysam, B. (2002), *Angiography and Plaque Imaging: Advanced Segmentation Techniques, CRC Press*, chapter Automated Model Based Segmentation, Tracing and Analysis of Retinal Vasculature From Digital Fundus Image.

K.M.Liang, R. & B.E.Khoo (2002), Free form shape representation using nurbs modeling, Proceedings of the Tenth International Conference in Central Europe on Computer Graphics, Visualisation and Computer Vision, Plenz, Chech Republic, pp. 67–73.

Koschan, A. & Skarbek, W. (1994), Colour image segmentation - a survey, Technical report.

L, H., Hsu, W., Lee, M. L. & Wang, H. (2003), A piecewise gaussian model for profiling and differentiating rretinal vessels, International Conference on Image Processing.

Lam, H. K. & Chutatape, O. (2003), Blood vessel tracking for optic nerve localisation for field 1-3 colour fundus image, ICICS-PCM, pp. 1437–1441.

L.Gang, Chutatape, O. & Krishnan, S. M. (2002), 'Detection and measurement of retinal vessels in fundus image using amplitude modified second order gaussian filter', *IEEE Transactions on Biomedical Engineering* **49**(2).

Li, H. & Opas, C. (2001), Location of optic disc in retinal images, International Conference of Image Processing, pp. 837–840.

Lowell, J., Hunter, A., Steel, D., Basu, A., Ryder, R., Fletcher, E. & Kennedy, L. (2004), 'Optic nerve head segmentation', **23**(2), 256–264.

M. Lalonde, M. B. & Gagon, L. (2001), 'Fast and robust optic disc detection using pyramidal decomposition and hausdorff based template matching', *IEEE Transaction on Medical Imaging* **20**(11), 1193–1200.

M. Sonka, V. H. & Boyle, R. (1993), *Image Processing, Analysis and Machine Vision*, International Thomson Publishing.

McInerney, T. & Terzopoulos, D. (1996), 'Derformable models in medical image analysis: A survey', *Medical Image Analysis* **1**(2), 91–108.

Niblack, W. (1986), *An Introduction to Image Processing*, Prentice-Hall, Englewood Cliffs, NJ.

OPHTEL. Telematics in Ophthalmology (2006).
URL: *http://ehto.org/ht_projects/html/dynamic/92.html*

Osareh, A. (2004), Automated identification of diabetic retinal exudates and the optic disc, phd thesis, Master's thesis.

Otsu, N. (1979), A threshold selection method from gray-level histograms, Vol. 9, IEEE transactions on Systems, Man and Cybernetics, pp. 62–66.

Panjwani, D. K. & Healey, G. (1995), 'Markov random field models for unsupervised segmentation of textured colour images', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **17**(10), 939–954.

Perona, P. & Malik, J. (1990), 'Scale space and edge detection using anisotropic diffusion', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **12**(7), 629–639.

Rosenfeld, A. & Torre, P. (1983), 'Histogram concavity analysis as an aid in threshold selection', *IEEE Transactions on System, Man and Cybernetics* pp. 231–235.

Scott & Christie and Associates: Your EyeSight Is Our Vision (n.d.).
URL: *http://www.scottandchristie.com/eye.cfm*

Segal, S. E. (1998), 'Pasadena eye associates'.
URL: *(How Does Human Eye Work?)http://www.pasadenaeye.com/faq/fraq15/fraq15_text/html*

Shafarenko, L., Petrou, M. & Kittler, J. (1997), 'Automatic wateshed segmentation of randomly textured colour images', *IEEE Transactions on Image Processing* **6**(11), 1530–1544.

Sinthanayothin, C., Boyce, J., Cook, H. & Williamson, T. (1999), 'Automated localisation of the optic disc, fovea and retinal blood vessels from digital colour fundus images', *British Journal of Ophthalmology* **83**, 902–910.

STARE Project Website (2003).
URL: *http://www.ces.clemson.edu/ahoover/stare*

T. Teng, M. L. & Claremont, D. (2001), 'Progress towards automated diabetic ocular screening: A review of image analysis and intellegent systems for diabetic retinopathy', *Medical and Biological Engineering and Computing* **40**, 2–12.

The Eye: Information About Vision Loss And Blindness (2005).
URL: *http://www.99main.com/ charlief/Blindness.html*

Thomas, S. & Y.T.Chan (1989), A simple approach for the estimation of circular arc center and its radius, Vol. 45, Computer Vision, Graphics and Image Processing, pp. 362–370.

Tolias, Y. A. & Panas, S. M. (1997), An unsupervised fuzzy vessel tracking algorithm for retinal images, Vol. 2, Proceedings of 6-th International Conference Fuzzy Systems (FUZZI-IEEE'97), pp. 325–330.

Trier, O. D. & Jain, A. K. (1995), 'Goal-directed evaluation of binarization methods', *IEEE Transactions on Pattern Analysis and Machine Intelligence* (17), 1191–1201.

UK National Screening Committee,http://www.nsc.nhs.uk/whatscreening/whatscreen_ind.htm (2005).
URL: *http://www.nsc.nhs.uk/whatscreening/whatscreen_ind.htm*

Vision Problems in U.S: Prevalence of Adult Vision Impairment and Age Related Eye Diseases in America: Prevent Blindness in America (2002), National Eye Institute and Prevent Blindness America.
URL: *http://www.nei.nih.gov/eyedata*

Walter, T., Klein, J., Massin, P. & A.Erginay (2002), 'A contribution of image processing to the diagnosis of diabetic retinopathy detection of exudates in colour fundus images of the human retina', *IEEE Transactions on Medical Imaging* **21**(10), 1236–1243.

Weisstein, E. W. (2005), 'Cubic spline from mathworld - a wolfram web resource'.
URL: *http://mathworld.wolfram.com/CubicSpline.html*

World Health Organization, Magnitude and Causes of Visual Impairment, Fact Sheet of World Health Organization (2004).
URL: *http://www.who.int/mediacentre/factsheets/fs282/en/index.html*

Wormald, R. (1999), 'Epidemiology in practice: Screening for eye disease, community and eye health', **12**(30), 29–30.

Xu, C. & Prince, J. L. (1998), 'Snakes, shapes and gradient vector flow', *IEEE Transactions on Image Processing* **7**(3), 359–370.

Xu, C. & Prince, J. L. (2000), Gradient vector flow deformable models, *in* I. Bankman, ed., 'Handbook of Medical Imaging', Academic Press.

Y. Du, C.-I. C. & Thouin, P. (2004), 'Unsupervised approach to colour video thresholding', *Optical Engineering* **43**(2), 282–289.

Z. Liu, C. O. & Krishnan, S. M. (1997), Automatic image analysis of fundus photograph, Proceedings of the 19-th International Conference IEEE/EMBS, Chicago, IL, USA, pp. 524–525.