Fundus Image Mosaic Generation For Large Field of View

by

Tanveer Ahmad

2011-NUST-MS PhD-ComE-33

MS-69

Submitted to Department of Computer Engineering

in fulfilment of the requirements for the degree of

Masters of Science

in

Computer Engineering

Thesis Supervisor

Dr. Usman Akram

College of Electrical and Mechanical Engineering

National University of Science and Technology

December 2014

This page is intentionally left blank

ACKNOWLEDGEMENTS

"*Those who do not thank people, they do not thank Allah*." *(Al-Tirmidhi 1878).*

There is an old saying ‘It takes a village to raise a child’ and I am not an exception. It is with the grace of Almighty that I was led into the company of the people, whose generosity, enthusiasm, and good shepherding sustained me in producing this work: My friends always backed me up very strongly. Therefore, I thank every single one of them for their support.

This research is funded by National ICT R&D fund. I am also grateful to AFIO for their support and data provision.

To my Parents, Advisors and colleagues.

ABSTRACT

Considering the importance of retina in the eye as many ocular and systemic diseases are manifested in the retina, the need for detailed study of 2-D retinal images, 3-D retinal images and panoramic retinal images is increasing rapidly. Many structures in human body are responsible for the process of vision but our works is strictly related to the retina as most diseases of eye manifest themselves in retina of the human eye.

Digital fundus images are commonly used for computer aided diagnosis of different eye diseases such as diabetic retinopathy, glaucoma and age related macular degeneration. A very important problem with fundus cameras is that they provide fundus image only for a small field of view (FOV). Our thesis work presents a novel method to increase the FOV by stitching different fundus images of a particular patient. The proposed system uses weber local descriptor (WLD) based descriptors and generates a blended image by combining all available images. This work also compares the proposed system with Harris corner detector, SURF, SIFT and ASIFT based descriptors for the same application. A local dataset of 15 patients with different number of fundus images is gathered for proper validation of the proposed system.

Keywords – *SURF, SIFT, ASIFT, Harris Corner Detector, WLD, Image Stitching*

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1
INTRODUCTION

The Visual System is one of the most important systems in the human body. There are many causes of the malfunction of human visual system which in severe cases lead to blindness. Most of the diseases which affect the human visual system and some systemic diseases manifest themselves in the retina which makes the study of human Retina very important.

The advancement of digital photography in recent years has opened new dimensions in the area of medical diagnosis using images. Diseases are diagnosed nowadays with the help of digital images. Fundus Images are acquired by the fundus camera that take images of the interior portion of the eye and allow Eye-Specialists and Doctors to diagnose the eye diseases like lesions of diabetes, hypertension, maculopathy, eye ground arteriosclerosis and retinopathy, effectively and efficiently. The advancement in the technology and the development of new algorithms has introduced the automation of the detection of eye diseases using complex algorithms which are able to diagnose the diseases like diabetic retinopathy (DR) and Glaucoma with a greater accuracy than before. However, these algorithms are executed on more than one image to obtain accurate results, which is computationally an expensive task, so in order to obtain accurate results with less computation a panorama of the images is built which provides the complete structure of the inner eye.

1.1 Background and Motivation:

The images acquired from the digital fundus camera usually have a limited field of view, normally ranging from 30 to 45 degrees which restricts the visualization of the images to a limited area. Hence the analysis of the patient's disease is also restricted. To have a detail diagnosis of the disease a patient is suffering from, a set of images is acquired to enhance the view of field. To manifest the diseases all images should be taken under consideration, which is time consuming and can sometimes result in flawed diagnosis. So in order to manifest the eye diseases accurately with less computational time, the images covering different field of views of the eye should be stitched together to form a panorama and the algorithms should be applied to that single image for the diagnosis of the disease. Figure 1.1

shows the fundus image of a patient which does not show the entire eye but only a portion of eye while figure 1.2 shows the digital fundus image of eye from different fields of view.

The algorithms for image aligning, and stitching them to generate a mosaic are widely used in remote sensing, medical image, pattern recognition and computer vision [1]. There are basically two types of aligning techniques used in the photography. The first one is known as direct method. In direct method error metric has to be created and a suitable search technique must be devised to search for alienation. The main disadvantage of direct method is that it's too slow [2]. The second class of algorithms use the feature set extraction and the feature matching to match the images [3]. This method is more advantageous as it is more robust and potentially faster than the first method.

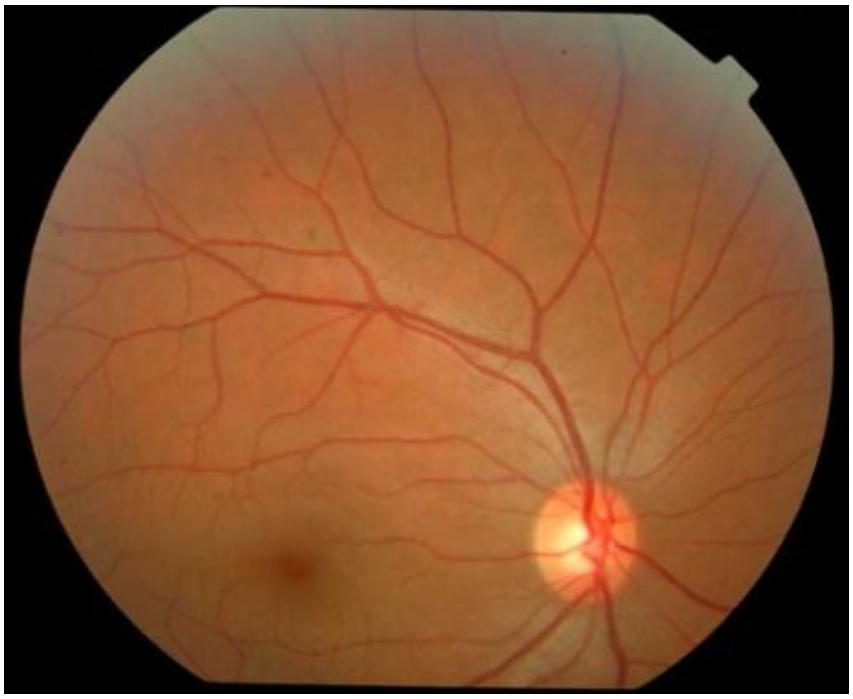


Figure 1.1: A Single Fundus Image of the patient. This image does not show the entire eye and the field of view is restricted

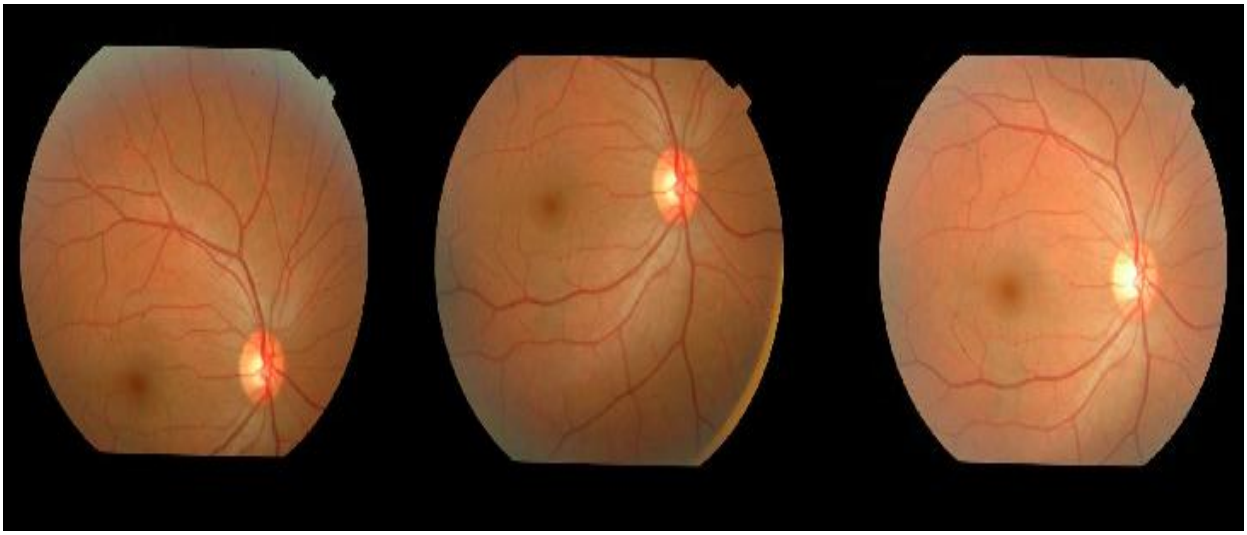


Figure 1.2: Three fundus images of the same patient

1.2 Overall Aim:

In our thesis, solution to the problem mentioned above is being proposed. WLD based automatic fundus image mosaic is presented. The features are extracted using Weber local descriptor, which is used because it is a dense descriptor and it is robust to noise and illumination [24]. The features are matched and the outliers are removed by using RANSAC (Random sample consensus) algorithm. After that the homograph of the most similar images is created. The last step is image registration and image mosaic implementation.

1.3 Outline:

The organization the thesis is as follow:

Chapter 1 contains the introduction to the thesis. Chapter 2 discusses the different techniques used in image stitching. In chapter 3, medical aspects and advantages of the thesis work. The design and implementation is presented in detail in Chapter 4. Chapter 5 presents the results of our work and its comparison with the state of art methods. And finally with chapter 6 we conclude our thesis and propose an outlook of possible extension as future work.

1.4 Summary:

Digital Fundus images have a limited field of view of the interior of human eye, so fundus image cannot show the entire eye. In order to view all the structures inside the eye more than one image should be taken of the interior part of the eye and hence we need more than one image and to diagnose eye diseases various algorithms should be applied on each of these images. But this process is computationally very expensive because the algorithms are applied on all the images, so in order to make it one image having a larger field of view image stitching is done. This was the real motivation behind this thesis.

Chapter 2

MEDICAL ASPECT OF IMAGE STITCHING

As we know that Retina is an important part of the human eye and it helps in the manifestation of many eye diseases as well as systemic diseases which can be cured if diagnosed at an early stage. This makes the study of retina very interesting. So in order to know about the importance of the retina we should first discuss about the structure of the retina.

2.1 Retinal Structure:

In order to study the structure of retina it is also important to study the structure of the human eye. Fig 2.1 shows the structure of the human eye.

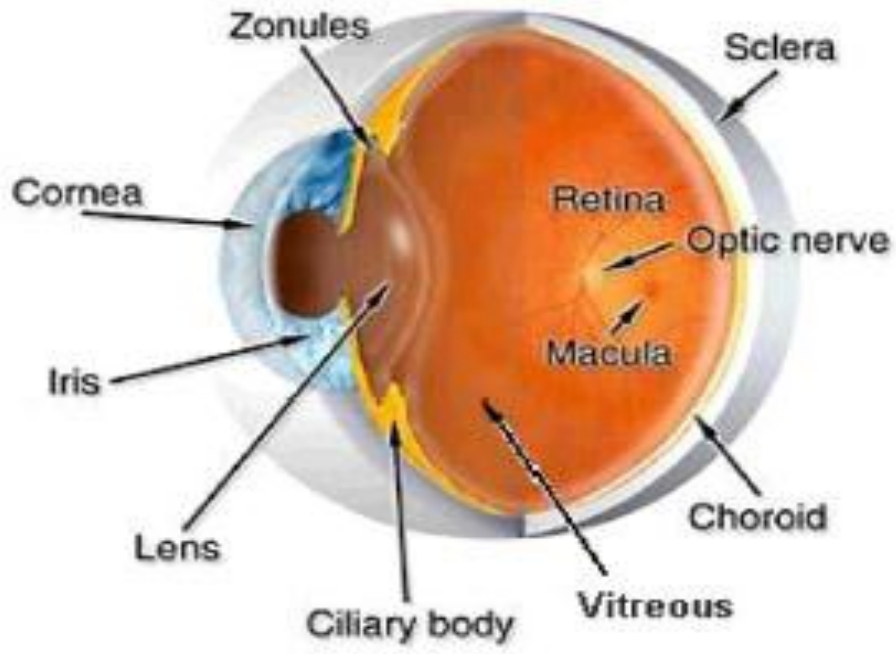


Figure 2.1: Anatomy of the Eye [20]

Retina has two main parts namely macula and the peripheral retina. Macula is the centre and the main portion of the human retina and it is responsible for the central vision of the human eye while the peripheral retina is responsible for the peripheral vision. Peripheral vision is the vision responsible for seeing objects out of the corner of the human eye. In other words with central vision, the finest details about the object which is in front of our eyes can be seen while with peripheral vision we can see objects that are not in front of our eyes but we see them with the help of the corner of the eye.

Figure 2.2 shows the structure of the retina. Fovea is shown in the figure which is the central part of the macula. The lines shown in the figure are the blood vessels and they bring nutrition and oxygen to the retina, while optic nerve is shown in the figure at the right side which is the brightest spot. All blood vessels rise from the optic nerve [13].

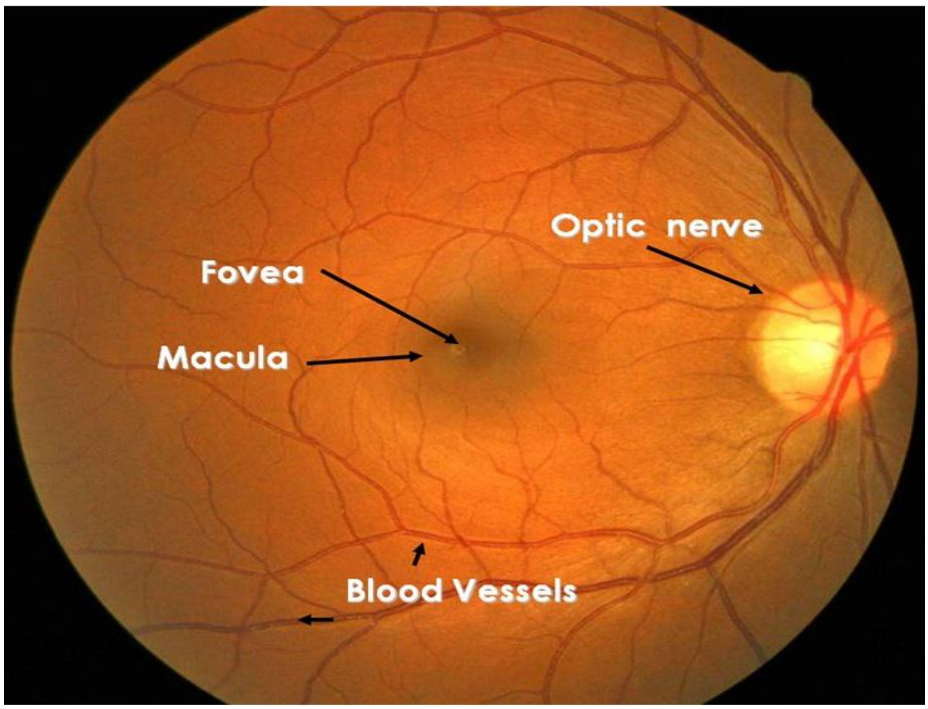


Figure 2.2: Structure of Retina and its main parts

.

2.2 Acquiring Retinal Images:

Retinal images are acquired by the Digital Fundus camera. Digital Fundus camera is low power microscope with a camera attached to it for the

purpose of acquiring the image of the retina. These fundus images are then used by the Eye specialists to diagnose ocular disease. Figure 2.3 shows a digital fundus camera which is used for acquiring retinal images.

Figure 2.3: Fundus Camera [21]

2.3 Ocular Diseases:

There are many eye diseases that may occur due to the disorder in the structure inside the retina [15]. Sometimes the disease effect the macula as shown in figure 2.4 which can be identified with the help of fundus images, because of this disease central vision is lost. Sometimes multiple new vessels appear on the surface of the optic nerve disc which causes a change in the structure of the retina as shown in figure 2.5, while other times new vessels appear on the other areas of the retina except the optic nerve disc as shown in figure 2.6.

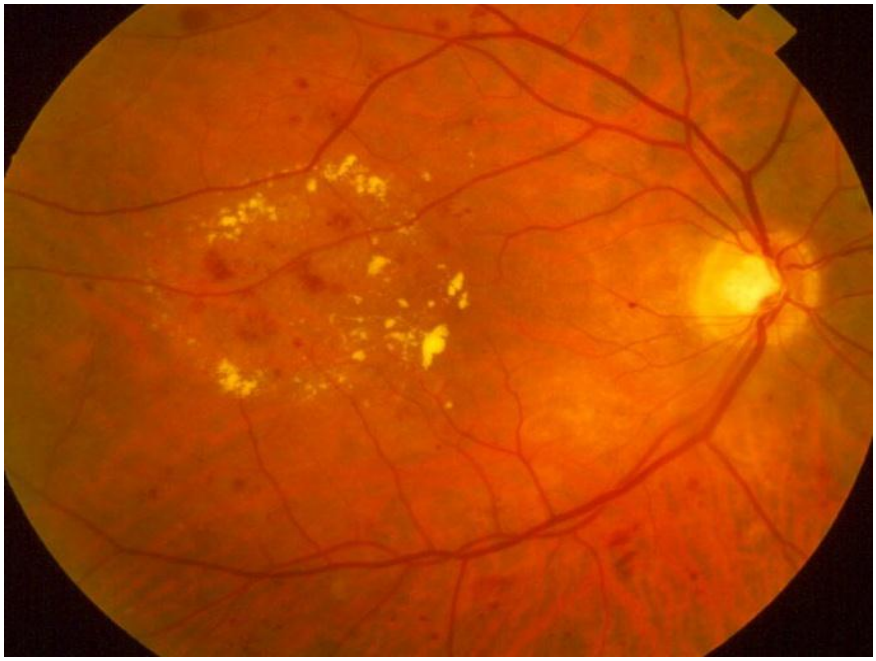


Figure 2.4: disorder of the Macula

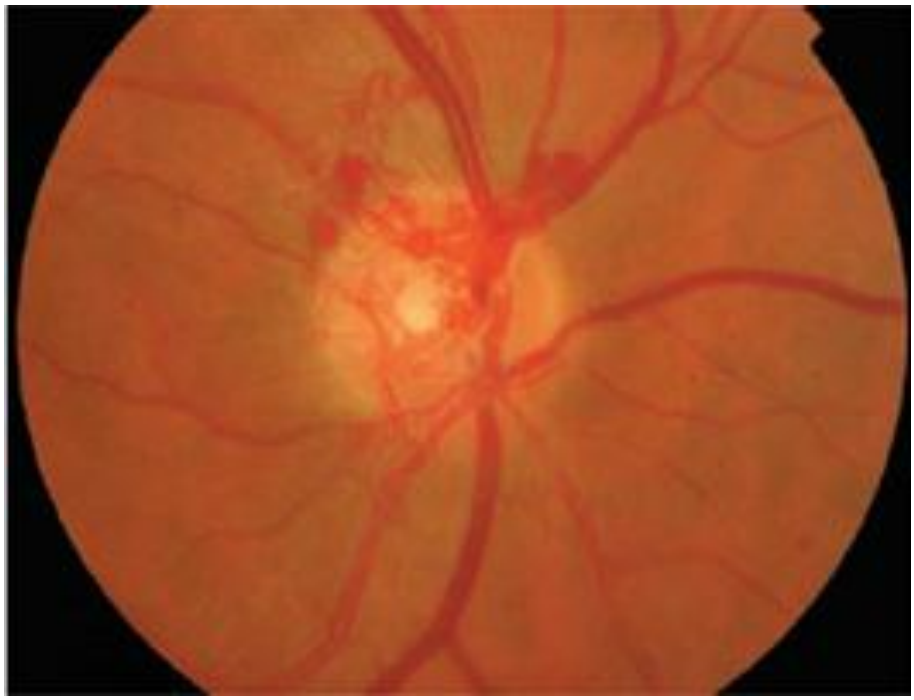


Figure 2.5: New Vessels on optic nerve disc

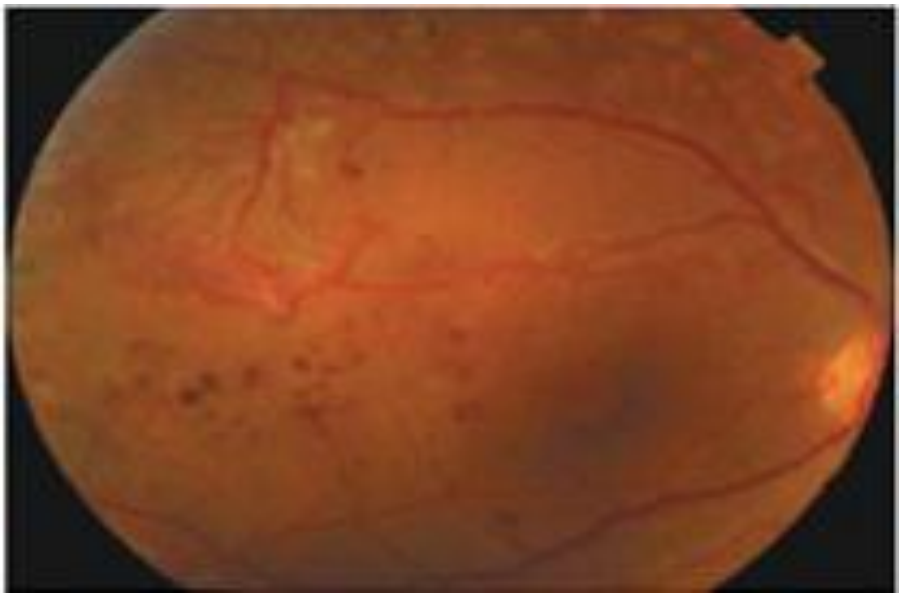


Figure 2.6: New Vessels on other areas except optic nerve disc

There are three main diseases which cause problems in the eye sight of the humans and in severe case these diseases may lead to blindness. During the early stage of these diseases these can be avoided and they can manifest themselves in retina during the early stages as these diseases cause a significant change in the structure of the retina.

The three main ocular diseases which may cause blindness are:

1. Diabetic Retinopathy:

   Diabetic Retinopathy is the severe form of Diabetes which develops normally after almost 10 years of Diabetes and it is the second most common disease that causes vision loss. It is caused when the blood vessels in retina are severely damaged. Severe damage of Retinal Vessels wall may lead to:

   (a) Ischemia:

   Ischemia is caused when blood supply to the tissues is restricted, which in turn reduces the percentage of oxygen and glucose in the tissue making it difficult for the tissue to keep alive. When the deficiency of oxygen and glucose occurs in retina it may sometime cause death to retinal cell and due to the death of the retinal cells the vision is lost.

(b) Diabetic Macular Edema (DME):

Diabetic Macular Edema is the swelling or thickening of macula (A yellow spot close to the centre of the retina of the eye). The main cause of the blindness caused due to Diabetes is Diabetic Macular Edema. During Diabetic Macular Edema leaking starts in capillaries [16], [17], which lead to the visual loss in severe case.

There are no symptoms or early indications of Diabetic Retinopathy. The patients of Diabetic Retinopathy live normal life during the early stages of the disease. They only way to manifest the disease in the early stages is through proper inspection of eyes regularly through fundus images of the Retina. Figure 2.7 shows the fundus image of a patient having diabetic macular Edema.

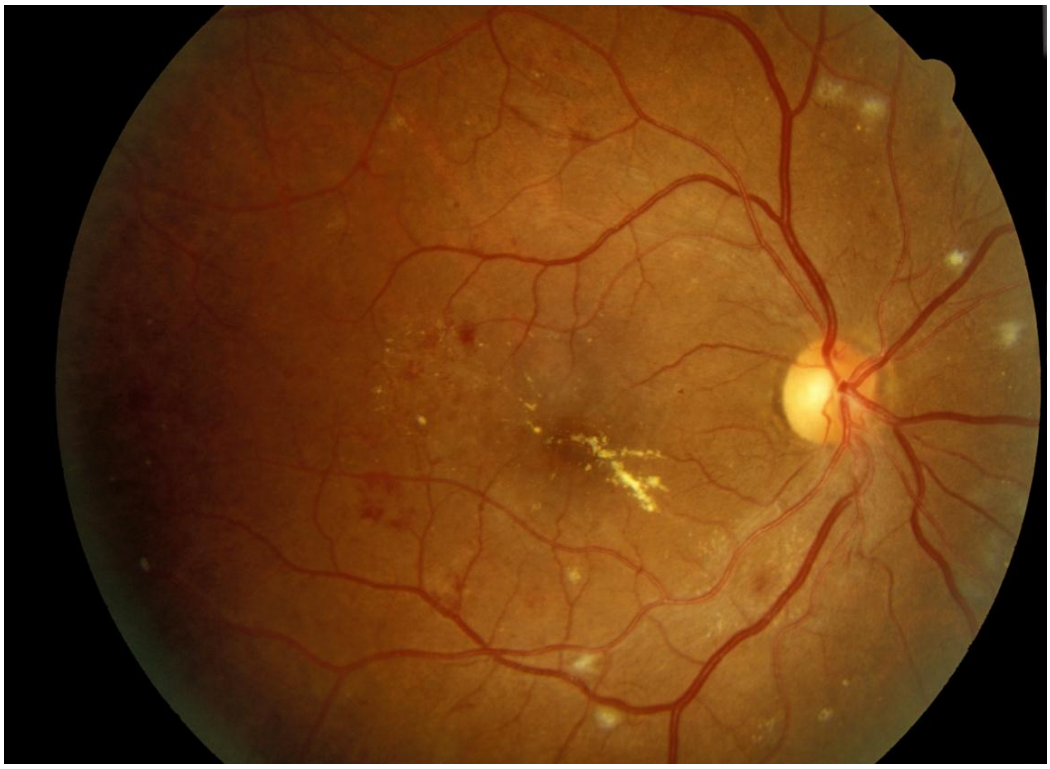


Figure 2.7: Fundus Image of a patient having DME

2. Age Related Macular Degeneration:

Age Related Macular Degeneration (AMD) is the leading cause of blindness for aged people (normally 45 plus). AMD is caused by the

damaged occurred to Macula which leads to central vision loss. There are two types of visions. One is central vision and the other is peripheral vision. Central vision is responsible for the things humans see directly in front of them while peripheral vision is responsible for the things humans see sideways. Macula is responsible for the central vision of human beings and due to age related macular degeneration macula is affected which cause in the loss of central vision. Figure 2.8 shows the fundus image of a patient having age related Macular degeneration.

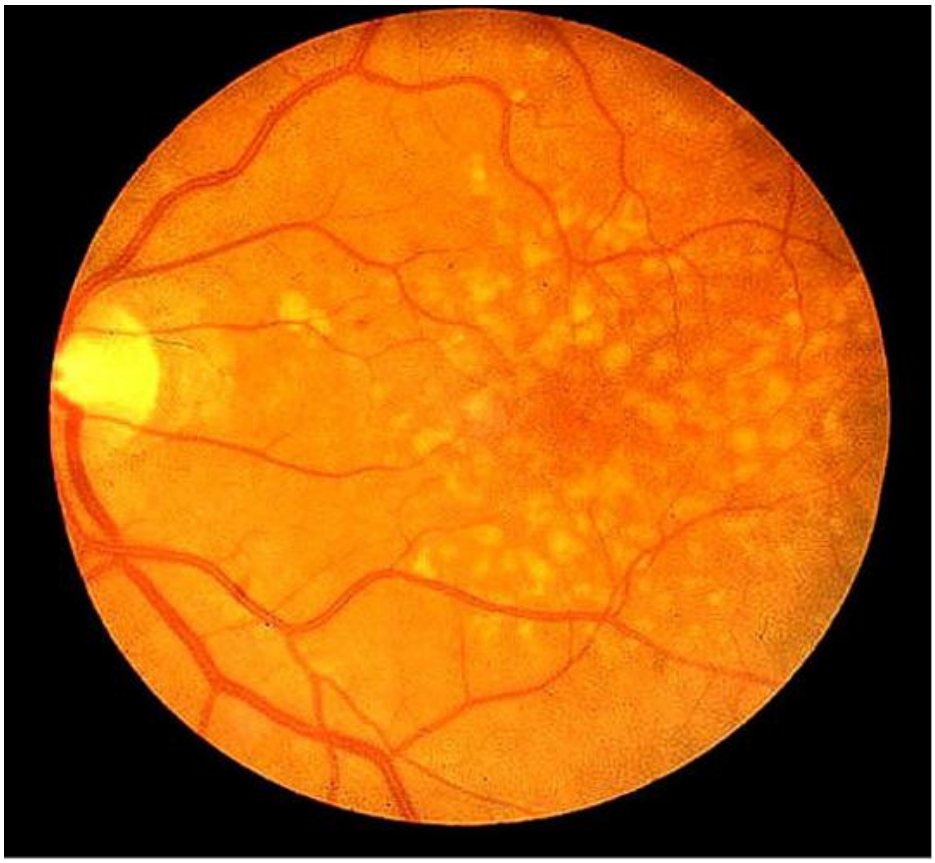


Figure 2.8: Fundus Image of a patient having Age related Macular degeneration

There are two types of Age Related macular degeneration which are discussed as follow:

(a) Dry Age Related Macular Degeneration:

Dry AMD starts with yellow pigments in the macula. It is the less threatening form of AMD in which the patients have a good vision in the early stages of the disease which leads to gradual loss in central vision.

Dry age related macular degeneration can be controlled through the use of particular diet [18].

(b) Wet Age Related Macular Degeneration:

Wet Age Related Macular Degeneration is caused by the extreme growth of blood vessels in macula which results in vision loss. The new abnormal blood vessels are very tenuous which leads to the leakage of blood inside macula which cause in the rapid loss of visions, that is the reason wet age related macular degeneration is considered as the most threatening form of macular degeneration.

During the early stages of the Macular degeneration it is very difficult to diagnose it as the patients normally have perfect vision. The only way to diagnose the disease is to have regular check-ups in which case fundus images and 3-D images of retina prove helpful.

3. Glaucoma:

Glaucoma is the third main cause of blindness in human beings. It is the disease which is occurred due to the damage in the eye's optic nerve. Optic nerve is damaged due to the increased pressure inside the eye. In severe cases Glaucoma results in resultant visual field loss [14]. Due to the early diagnosis the risk of the blindness caused by the Glaucoma can be minimized [19]. Figure 2.9 shows the fundus image of a patient suffering from advanced Glaucoma.

There are two types of Glaucoma which are discussed in brief below:

(a) Open Angle Glaucoma:

Open angle Glaucoma is the most common type of Glaucoma which occurs when there is some difficulty in the drainage of fluids from the eye. Open angle Glaucoma is a common disease which develops gradually over a period of time and it is not particularly painful.

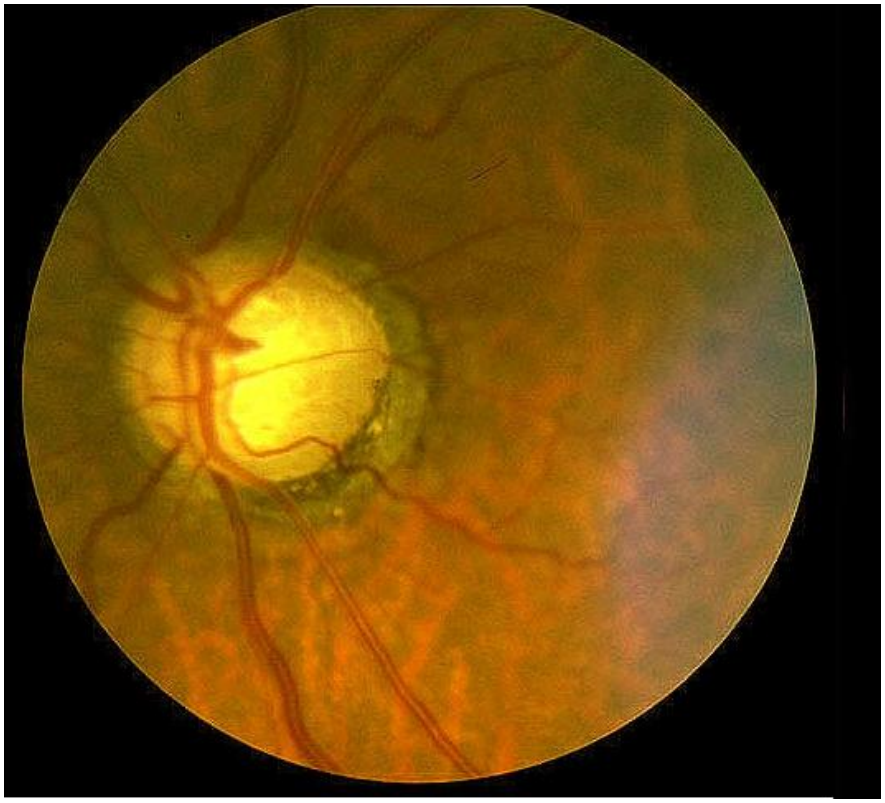


Figure 2.9: Fundus Image of a patient having advanced Glaucoma

(b) Close Angle Glaucoma:

Close angle Glaucoma is less common as compare to open angle Glaucoma. In this type of Glaucoma the angle between cornea and the iris is too narrow, so it makes difficult for fluids to flow. It develops all of a sudden and can be very painful.

It is very difficult to detect Glaucoma(specially open angle Glaucoma which is more common) in the early stages as the patients do not feel pain in the early stages of the open angle Glaucoma so it is very important to undergo regular check-ups so that it can be diagnosed with the help of fundus image in the early stage.

The motivation behind this thesis is the early manifestation of eye diseases which lead to visual loss. As discussed earlier that there are three types of diseases which are the main cause behind the visual loss in human eye and these diseases are very difficult to overcome in the final stages. So for this purposes early manifestation of the diseases is important to avoid visual loss. Fundus images and panoramic retinal images help in the manifestation of these diseases.

2.4 Summary:

In this chapter the structure of the eye and retina is discussed in brief and then the three basic ocular diseases are discussed which in severe case lead to vision loss. Early diagnosis of all these diseases is possible through regular check up of retina as the structure of the retina changes due to these diseases and it can be identified with the help of digital fundus images. The stitching of fundus images provides an easy way to manifest these diseases at an early stage as because of the stitching the field of view is larger.

Chapter 3

LITERATURE REVIEW

In this chapter, an overview of the existing state-of-the-art techniques used in the process of image stitching is given. We will discuss the benefits of using each of these methods and we will also discuss the disadvantages of some of these techniques. We will also compare various techniques used in the process the image stitching.

3.1 Overview:

Image Stitching is the process of combining multiple overlapped images to make it a single large image which is often known as panorama. The main focus of image stitching is to create a panorama in which the images are aligned seamlessly. In this section, we will discuss the state-of-the art techniques used in the process of image stitching and will point out its strengths and weaknesses.

3.2 Fundus Image Stitching System:

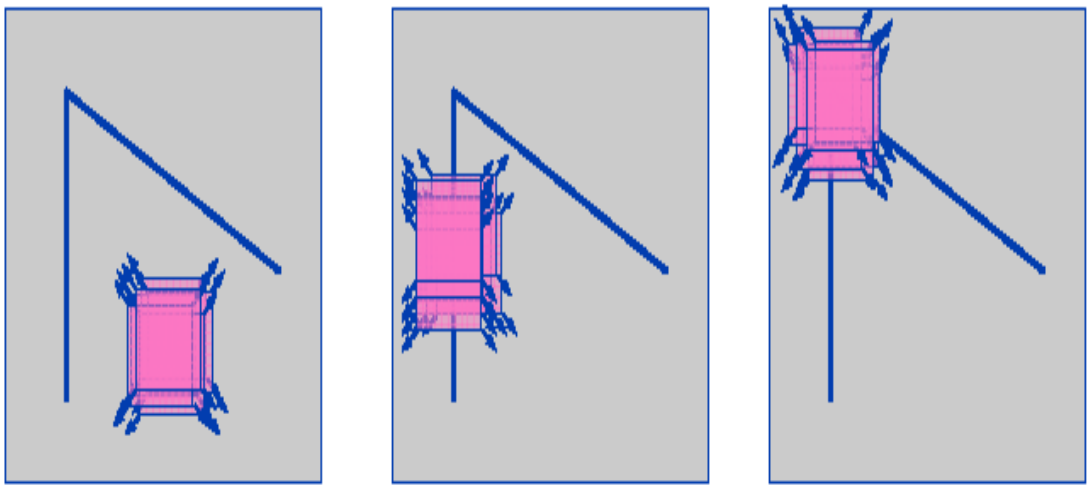
There are many techniques used in the process of image stitching. Some state-of-the art techniques used are as follow:

(A) Harris Corner Features:

Corner detection or the interest point detection is the process in which interest points are detected in an image and these points should be repeatable in another view of the same image. Interest points or corners are the expressive textures in an image. Interest points are the points where direction of the edge changes abruptly.

Finding the interest points or corners can be done by using a small window and shifting it (figure: 3.1). If there is no change occurred while shifting the window then that region is considered as flat region. When there is change in one direction if move the window then that

region is known as the edge but if there is significant change across both the directions while we move the window then that region is known as corner. That's the way to find the corner in an image.



a) Flat Region b) Edge c) Corner

Figure 3.1**:** Moving window used to find corner [23]

Harris Corner detection is an algorithm used to detect the corners. The process of computing the corners in the Harris Corner detection is as follow:

Harris corners are computed in following ways:

1. A window is swapped over gray scale image.
2. To view large variation intensity, window is expanded using Taylor's series.
3. Above expression is solved and represented in matrix form.
4. Score is calculated to check if a corner lies or not within it.

The equations 3.1, 3.2 and 3.3 for detection of corner are as follows [24]:

$$R = det(M) - k(trace(M))^2 \text{.....(3.1)}$$

$$det(M) = \lambda 1 \lambda 2 \text{.....(3.2)}$$

$$trace(M) = \lambda 1 + \lambda 2 \text{.....(3.3)}$$

The corners would only exist if R is greater than the threshold.

For time critical system the feature extraction and feature matching using SIFT (Scale Invariant Feature Transform) is quite slow as it takes more time in computation. A faster version than SIFT is already available which is known as SURF (Speeded up Robust Features). But Harris Corner detector is much faster than both SURF and SIFT. It computes features within 20 ms approximately. But the important point here is that Harris Corner Detector does not have its descriptor so either SIFT descriptor or a SURF descriptor can be used with it. So for corner detection Harris corner detector can be use because it is fast as compared to the state- of the art techniques and either SIFT or SURF descriptor should be used with it [12]. An advantage of Harris corner detector is that it is rotation invariant which means that if Harris points are computed in a image and then the image is rotated and then Harris detector is applied to it, the Harris points in the rotated and the un-rotated image will be in the same place. Harris Corner detector is not scale invariant. The feature extraction in the Harris corner detector is less efficient than other state-of the art techniques that's why the Harris corner detector cannot be used when the corners in an object are less prominent.

(B) SURF:

SIFT (Scale Invariant Fourier Transform) was introduced by David Lowe in 2004. It has been used for key point description and detection for a long time but the problem with SIFT is that it is slow and takes a lot of computational time, so a faster version was required. So for this purpose SURF (speeded up robust features) was proposed as it is faster than SIFT. SURF is scale and rotation invariant which means that if the images are rotated or scaled their interest points will remain the same. With respect to distinctiveness, robustness, repeatability SURF has outperformed previous techniques. Its advantage is that it is faster than other state of the art techniques [10].

In this approach, box filers are used for the approximation of the laplacian of Gaussian as shown in figure 3.2. A great advantage of using these box filters is that it makes the calculation in the convolution easy with the help of integral images, and the other positive thing about it is that it can be done for different scales at the same time.

In feature description phase it uses Haar-wavelet responses in the key points neighbourhood and its descriptor is a 64 dimensional feature vector hence it reduces the time for feature matching

and computation. In normal circumstances SURF works with same accuracy as other state-of-the-art methods in less time.

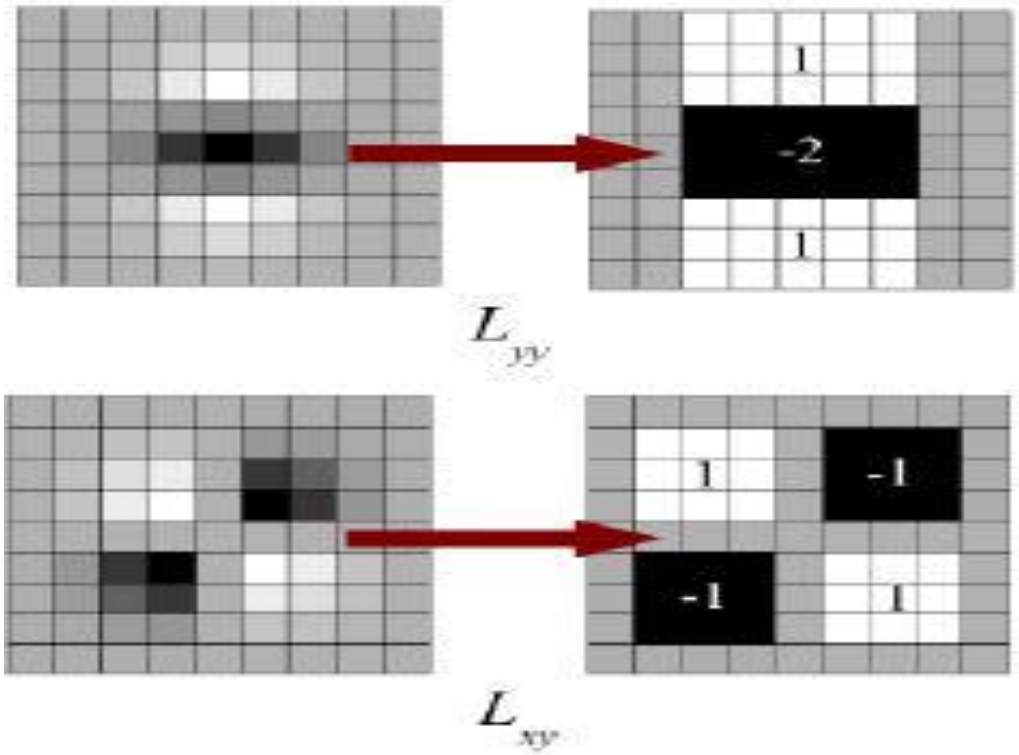


Figure 3.2: Box filters used in SURF [10]

The Procedure of SURF is as follow [10]:

SURF uses 2D Haar wavelet responses on integral images. The features are then extracted by summing the Haar wavelet responses around the points of interest. The use of integral images makes the algorithm robust. The key steps are:

1) SURF uses integral images.
2) Finds LOG using box filter.
3) Integral images can be computed in parallel.
4) Use of Hessian matrix for both scale and location determination.

(C) SIFT:

SIFT stands for Scale Invariant Feature Transform. It is a technique developed for the detection the local features of objects. It is important to note that features detected here are Scale invariant features which make the scale invariant Fourier Transform a good choice to compare the feature of two images with different scale and different camera orientation. The detailed insight into the algorithm can be found in [4], however the main steps of SIFT are as follow:

1) Scale Space Peak Selection:

To find the scale invariant features the scale change is applied by using Difference of Gaussian which is alternative of Laplace of Gaussian but is less costly as compared to Laplace of Gaussian. This process is done for different octaves of the Gaussian pyramid as shown in the figure 3.3.

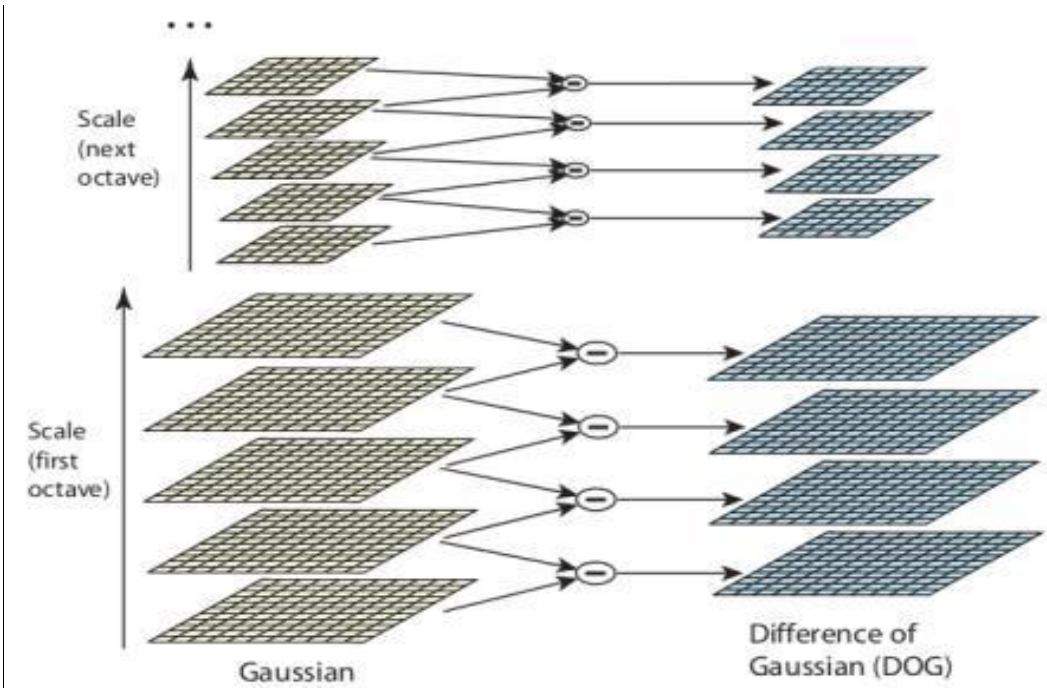


Figure 3.3: Approximation of LoG with Difference of Gaussian [4]

After the Difference of Gaussian is found, the local extrema is search in the images over different scales. A single pixel in an image is compared to all its neighbours and then it is compared with the neighbours in a scale above and a scale below it. It will be a potential interest point if it is an extrema otherwise the search will go on for the next pixel. This process is given in the figure 3.4. The number of octaves can vary but the most suitable

number of octaves is 4 and the number of scale levels per octave may vary based on the nature of image but various experiments show that the number of scales per active should be 3 as mentioned in [4].

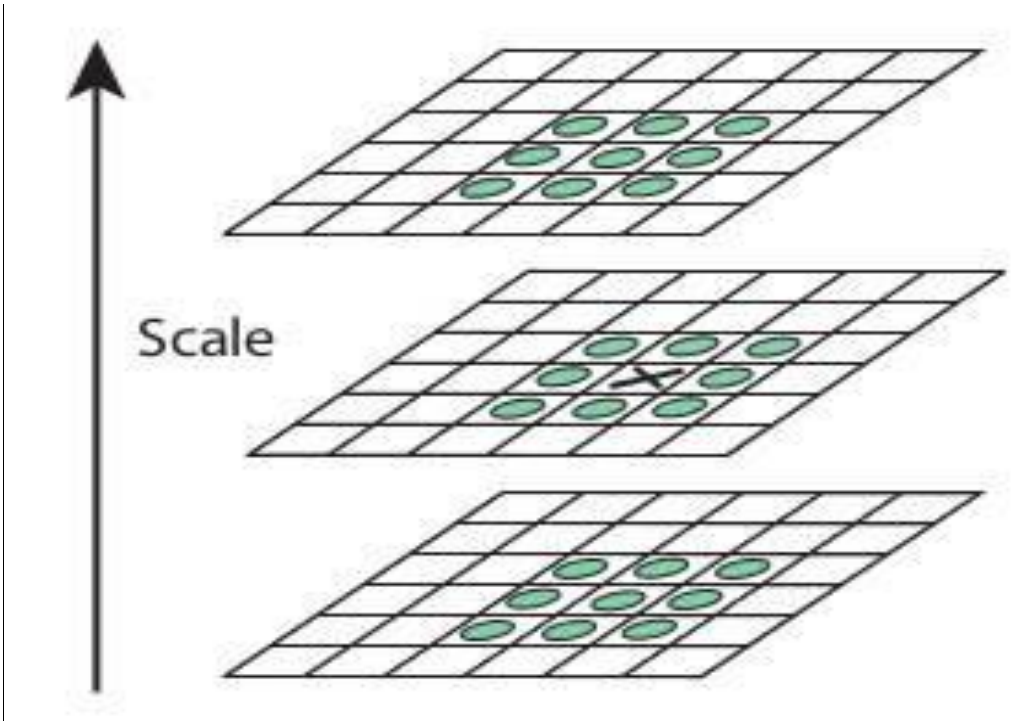


Figure 3.4: Finding potential key points across different scales [4]

2) Refining the Key points:

Once the extrema are located certain thresholds are used to reduce the noise present. Some of the extrema present in the image are not the actual interest points but they are presented as the interest points due to the noise. So to remove those extrema there are some thresholds used. The first threshold is the Contrast Threshold, which will remove those extrema which have low contrast and the threshold used in [4] is 0.03. So it will reject the extrema if its value is less than 0.03. The second threshold applied is known as edge Threshold where the corners are detected and the edges are discarded. It is done with the help of Hessian Matrix, if the ratio of the Eigen value is greater than 10 than the extrema is discarded as mentioned in [4]. These filters eliminate the noise on the low contrast areas and the edges of the image which were wrongly considered as interest points.

3) Key Point Orientation:

In this step the orientation of the interest points is calculated. For this purpose a weighted direction histogram of 36 bins is created, which means that the total 360 degrees is divided in 36 bins, so 1 bin will be of 10 degrees. Each pixel is assigned the weight and direction in the histogram. If the gradient magnitude of a pixel is high then it will count more as compared to the pixel will low magnitude. We will pick the one which has the highest peak and that will be the orientation of the interest point.

4) Descriptor:

SIFT descriptor is a 3-D spatial histogram of the image gradient. Samples are weighted and accumulated in 3-D histogram. Furthermore Gaussian weight is applied to give less weight to gradients away from key-point centre. Relative orientation and magnitude in a 16*16 neighbourhood is computed. The spatial coordinates are quantized into four bins whereas orientations are quantized into eight. SIFT Descriptor is a 128 dimensional vector.

5) Key point matching: A simple Euclidian distance error is calculated between the key point location and the nearest neighbour approach is used to find the closest match. However noise can be introduced from this method, to avoid this, a threshold is set. This threshold rejects any key points having distance error greater than the threshold.

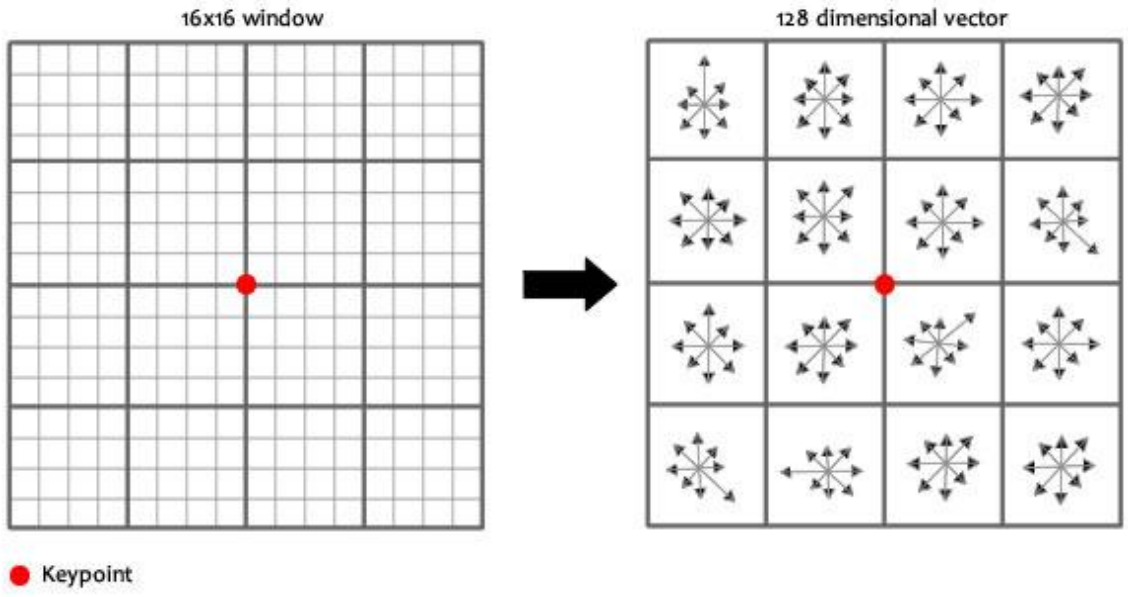


Figure 3.5: SIFT Descriptor [4]

The advantages of using SIFT are as follow:

(i) Local Features:

SIFT uses local features that's why it is robust to occlusion because some part will be occluded but other parts are not.

(ii) Distinctiveness:

These are distinct features and individual feature can be matched to a database of objects.

(iii) Quantity:

The number of interest points is high which is helpful in matching.

(iv) Efficiency:

SIFT gives close to real time efficiency.

(D) ASIFT:

In 2009 Jean Michael Morel and Guoshen Yu propose a method for fully affine scale invariant point detection and description and named it as ASIFT which is a variant of sift proposed by David Lowe in 2004. The SIFT detectors and descriptors are not fully affine and ASIFT is the affine extension of SIFT which means that the detector and descriptor of ASIFT is the same as that of SIFT but it gives the affine camera model and introduces another term "transition tilt", which is used for measuring the view point change from one view to another. ASIFT gives best results even if the there is distortion between images due to viewpoint change. ASIFT method could reliably identify features that have undergone transitional tilt of 36 degree or higher. The complexity of ASIFT algorithm is almost twice the complexity of the SIFT algorithm [11].

The difference between ASIFT and SIFT algorithm is that SIFT only simulates zoom outs of the images and normalizes the rotation and the translation which gives those features which are full scale invariant but are not affine invariant. ASIFT, however, simulates three parameters namely scale and camera (both the angles namely longitudinal angle and latitudinal angle), and the three other parameters namely rotation (both sides) and translation are normalized by it. The following steps are performed by the ASIFT algorithm to achieve fully affine invariant features.

1) Simulation of Affine:

Image is transformed to simulate all possible affine distortions caused by the change of longitude and latitude of the camera. The rotations are followed by tilts performed using t-subsampling. The steps of the sampling are short so that any important tilt is not missed [11]. The details of the sampling step can be found in paper [11].

2) Anti-aliasing: An anti-aliasing filter is applied which is typically a Gaussian with standard deviation of '$c\sqrt{t^2+1}$'where c is taken as 0.8 [3][11].

3) SIFT: The resulting images generated by tilt transformation are processed and compared with SIFT algorithm to extract the four parameters other parameters.

ASIFT has certain advantages over all the state-of the art techniques. ASIFT detects more interest points as compared to the other state-of the art techniques and give lesser number of mismatches as well. The other advantage of ASIFT is that it gives better results than the other state-of the art techniques in almost all type of scenarios.

(E) Homography:

Homography is used in computer vision to relate the pixel co-ordinates in two images. The homography works when the images are acquired from the same plane at the different angle.

The scenes captured by the cameras are 3D scenes each having three coordinates x, y and z respectively, but unfortunately we do not get the depth coordinate of pixel in the photographic image. However, the scene considered is planer which allows us to delete the last row [5]. The final coordinates in the images are 2D coordinates. The general equation (3.4) used is as follows:

$$\begin{bmatrix} x2 \\ y2 \\ z2 \end{bmatrix} = \begin{bmatrix} H11 & H12 & H13 \\ H21 & H22 & H23 \\ H31 & H32 & H33 \end{bmatrix} \begin{bmatrix} x1 \\ y1 \\ z1 \end{bmatrix} \qquad \text{……} \quad \textbf{(3.4)}$$

So x2 = H.x1

The coordinates are converted into 2D homogeneous coordinates. As the coordinates are homogeneous z=1 can be considered without the loss of generality. The equation is rearranged as follows:

$$\begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -x'_1x_1 & -x'_1y_1 & -x'_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -y'_1x_1 & -y'_1y_1 & -y'_1 \\ & & & & \vdots & & & & \\ x_n & y_n & 1 & 0 & 0 & 0 & -x'_nx_n & -x'_ny_n & -x'_n \\ 0 & 0 & 0 & x_n & y_n & 1 & -y'_nx_n & -y'_ny_n & -y'_n \end{bmatrix} \begin{bmatrix} h_{00} \\ h_{01} \\ h_{02} \\ h_{10} \\ h_{11} \\ h_{12} \\ h_{20} \\ h_{21} \\ h_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} \quad \text{…….. (3.5)}$$

The equation (3.5) could be solved using the sum squared error and the final form of the equation is as follows:

$$A^TAh = 0 \quad \text{……..…… (3.6)}$$

From the equation (3.6) h could be known by finding the Eigen vector of A^TA with the smallest Eigen value. This equation needs 4 or more points on the images to find the vector 'h'.

There are many applications of Homography, but the most prominent applications are the measurement of the camera motion, image registration and image rectification.

(F) RANSAC:

RANSAC (Random Sample and Consensus) is a method to estimate the parameters of a mathematical model containing outliers. It was first introduced by Fischer and Bolles in 1981 [8]. The whole data is divided into inliers and outliers. Inliers are the points which can be explained with the help of model parameters while outliers cannot be explained with the help of model parameters as they represent the noise or the data which not represent the model. RANSAC can handle outliners larger than 50% of the entire data set, this percentage is hard to achieve by other techniques used for parameter estimation. RANSAC is an iterative method which is composed of the following three major steps:

1) Hypothesize:

Minimal sample set (MSS) is selected from input dataset to compute the model parameters. The cardinality used to determine the model parameters is chosen to be the smallest sufficient.

2) Test:

In this step RANSAC checks for the consistency of the elements in the database with the parameters estimated in the first step. The selected elements are known as consensus set (CS).

3) Iterations:

Suppose y is the probability for a MSS to find accurate estimate of the model parameters. The probability to find at least 1 outline would then be 1-y. If we find the probability for large number of iterations than the expression will be $(1-y)^h$, where h represents the number of iterations. The value of h is selected so that the probability is smaller than equal to threshold. The equation (3.7) is used to find the limit of iterations.

$$\boldsymbol{T = \left[\frac{(\log \varepsilon)}{(\log(1-q))}\right]} \quad \boldsymbol{\ldots\ldots\ldots\ldots (3.7)}$$

RANSAC performs the above mentioned steps iteratively, it terminated when the probability of finding a better ranked CS drops below a certain threshold [9].

The main advantage of RANSAC is that it performs robust estimation of the parameters in a model. Even when there exist a high number of outliers in the dataset, RANSAC still produce good results by estimating the parameters with high accuracy.

Figure 3.6: A dataset which contains many outliers, RANSAC should be applied on it [22]

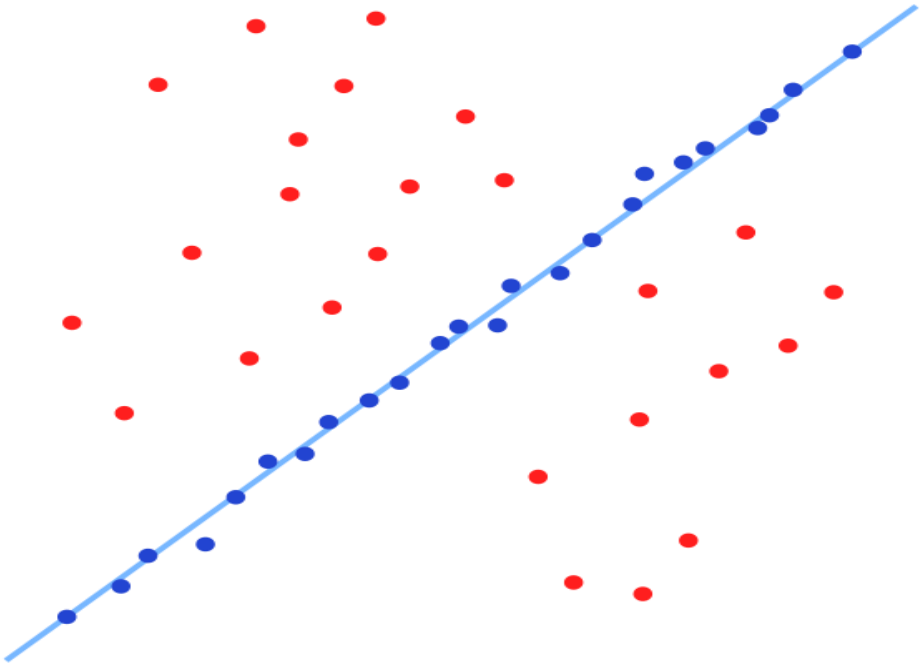


Figure 3.7: After applying RANSAC, the issue of outliers is resolved [22]

As shown in the above two figures, the data in the figure 3.6 contained both inliers and outliers. From the figure 3.6 it is clear that it is impossible to separate the inliers from the outliers without applying a technique on the data. But it is clear in the figure 3.7 that when we applied the RANSAC algorithm on the data then the outliers and inliers are separated from each other and the issue of the outliers is resolved.

3.3 Related Work on Retinal Image Mosaicing:

LI Jupeng *et al.* proposed an algorithm for automatic mosaic of the curved human color retinal images [25]. m-SIFT was used for feature detection as SIFT is not good at finding features when the surface area is flat. Second nearest neighboring strategy was used for the feature matching. After the feature matching the inliers identification between the two images was done. The technique of bilinear warping was then used to align the images. At the end multi-blending technique was used to create a panoramic image. The scope of their work was limited to the stitching of two images only.

LiFang Wei *et al.* presented a technique for retinal image mosaic [26]. In their proposed method feature detection was done with the help of PCA-SIFT (Principle Component Analysis- Scale Invariant Feature Transform). After the feature detection quadratic transformation model was used to simulate the anatomy of the eye. The feature matching was done with the help of k-nearest neighbors. The random sample consensus (RANSAC) was used to detect the inliers and to remove the outliers. At the end weighted mean was used to achieve seamless stitching.

Philippe C. Cattin *et al.* proposed an algorithm for retinal image mosaicing using local features [27]. Hessian matrix based detector was used for interest point detection and a 128 dimensional descriptor was use to describe the interest points. Feature matching was done with the help of k-nearest neighbors while RANSAC was used to match the inliers across the images. The outliers were also filtered out by using RANSAC. After using RANSAC, the anchor image selection and the mapping estimation was completed. The last step of their proposed algorithm was to blend the images by using Laplacian Pyramid.

Tae Eun Choe *et al.* propose a method of optimal global mosaic generation from retinal images [28]. In their proposed method Y-features were used for feature extraction and for feature

matching the windows enclosing Y-features were used. After the features matching RANSAC was used for inliers matching. Global pair-wise registration was used to find the optimal registration. The frame that gave the lowest selection error was selected as the reference frame. Other images were blended onto the reference frame based on shortest path.

3.4 Summary:

In this chapter various techniques namely Harris Corner detector, SURF, SIFT, ASIFT, Homography and RANSAC are discussed in brief. All these techniques are used in the process of image stitching. All the techniques have their own advantages and disadvantages which are discussed as well. At the end of the chapter the related work regarding retinal image mosaicing is discussed.

Chapter 4

Methodology:

4.1 Overview:

In this chapter the methodology which is used in our thesis work is discussed in detail. As discussed earlier, the main problem with digital fundus images is that the digital fundus images have a limited field of view which does not cover the entire interior of the eye. So to avoid this problem we are stitching the images having different limited fields of view to make single image having a field of view which can cover the entire interior part of the eye. The methodology we used for this purpose is discussed below:

4.2 Our Proposed Method:

In the proposed image stitching system for the fundus images we used some of the techniques which are discussed earlier. WLD (Weber Local Descriptor) algorithm is used for the feature generation and feature matching across the fundus images which we want to stitch to produce an image having a larger field of view. The key points obtained are then refined using the outlier removers. First of all the low contrast candidates and the edges which are wrongly considered as interest points are removed using the Taylor Series. After that Hessian Matrix is used to remove further outliers. In the process of feature matching the threshold is set to 0.8 to minimize the noise in the result.

After the process of feature matching RANSAC is used to find similarity between the two images. If the number of inliers generated by the RANSAC are greater than the specified threshold then the images are wrapped together. The data flow diagram of our proposed algorithm is shown in Figure 4.1:

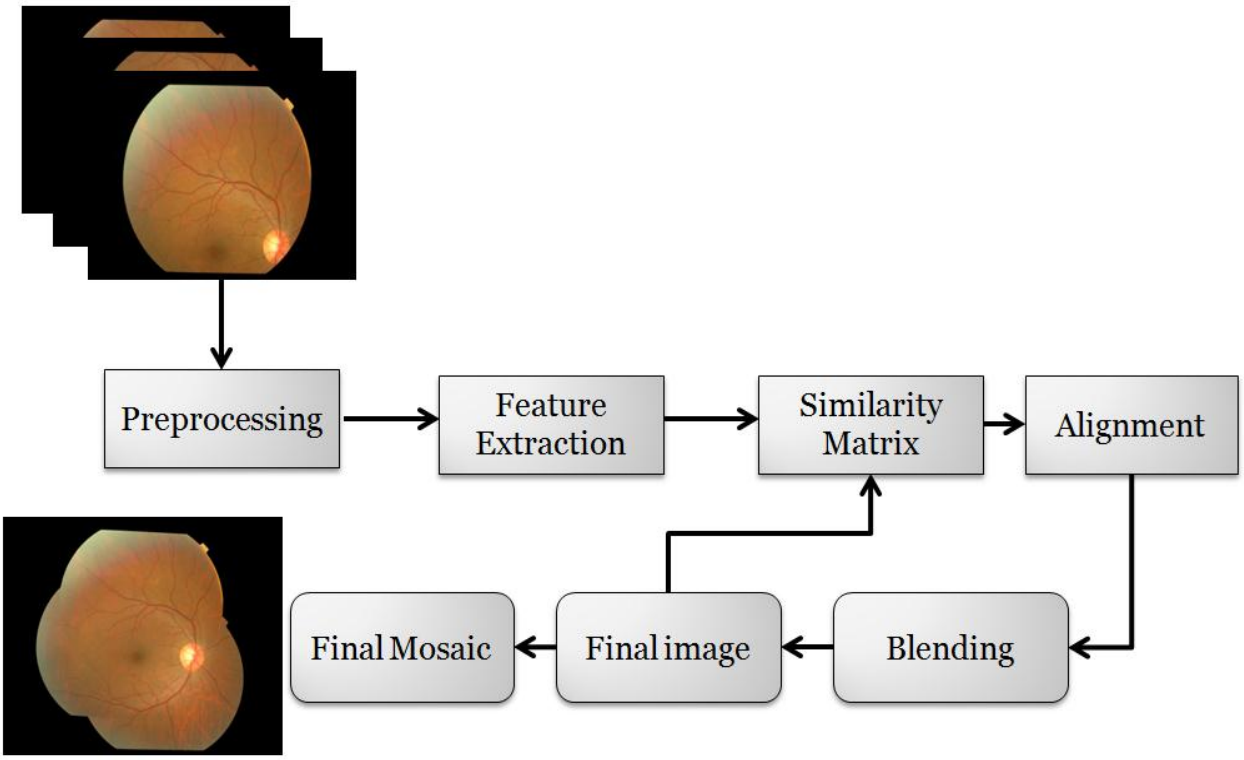


Figure 4.1: The data Flow Diagram of our proposed Algorithm

4.2.1 Weber Local Descriptor (WLD):

There are two types of descriptor namely sparse descriptors and dense descriptors. Sparse descriptors such as SIFT, SURF and ASIFT descriptors take the interest points like edges and corners which are located sparsely over the image while dense descriptors such as Local Binary Pattern and Weber Local descriptor extract local features pixel by pixel over the input image. Weber Local Descriptor is inspired from Weber's Law.

Weber Law is a psychological law which states that the ratio of the increment threshold to the background intensity is constant. Increment threshold is the slightest difference in the intensity which can be perceived by the human eye. Weber law can be expressed in the equation form as follow:

$$\frac{\Delta I}{I} = k \qquad \text{..... (4.1)}$$

In the above equation ΔI represents the increment threshold, *I* represents the initial stimulus intensity and *k* represents constant which means that the left hand side of the equation remains the same even if there is change in the intensity. ΔI/I is also called the Weber fraction. The example of weber law is that one can hear even a whisper if there is complete silent but if there is noise around then in order to get heard one should scream.

The weber local descriptor consists of two main components: differential excitation (ξ) and orientation (θ). Both of these components are discussed as follow:

i) Differential Excitation:

Differential Excitation is the ratio between two terms. One is the difference between average intensity of the neighbors and the current intensity and the other is the intensity of current pixel.

Differential excitation is computed by using the weber's law as it is computed by the ratio change in the current pixel and the initial intensity of the current pixel. Change of the current pixel is calculated by the differences between its neighbors and a current pixel. The equation for it can be written as follow:

$$V_s^{00} = \sum_{i=0}^{p-1}(\Delta x_i) = \sum_{i=0}^{p-1}(x_i - x_c) \quad \text{.....} \quad \textbf{(4.2)}$$

In the above equation $\mathbf{V_s^{00}}$ stands for the change in the intensity, p represents the number of neighbors, x_c represents the current pixel while $\mathrm{X_i}$ (i=0,1,…p-1) denotes the i-th neighbors of x_c. As discussed above that differential excitation is the ratio of two terms so it can be expressed as follow:

$$G_{ratio}(x_c) = {v_s^{00}}\Big/{v_s^{01}} \quad \text{.....} \quad \textbf{(4.3)}$$

$\mathbf{V_s^{01}}$ represents the intensity of the current pixel.

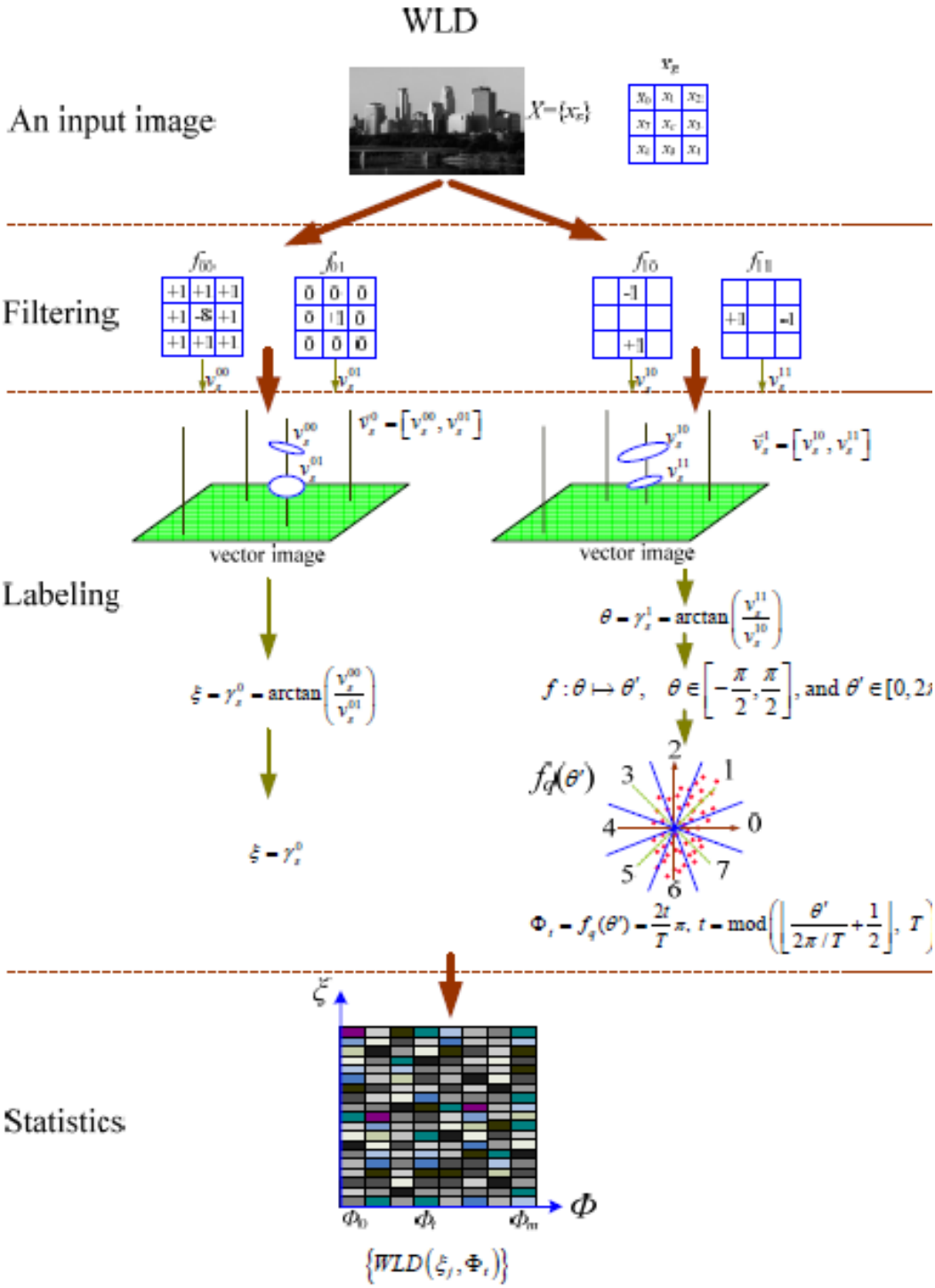


Figure 4.2: Illustration of the computation of weber local descriptor [24]

The problem with normal values is that it increases and decreases rapidly when the input is too large or too small, so for that purpose arctangent function is applied on it which limits the output. So after applying the arctangent function, the equation can be written as:

$$\xi(x_c) = arctan\left[\frac{v_s^{00}}{v_s^{01}}\right] = arctan\left[\sum_{i=0}^{p-1}\left(\frac{x_i - x_c}{x_c}\right)\right] \quad \ldots\ldots \textbf{(4.4)}$$

The above equation shows that weber local descriptor is robust to change in image contrast because the initial intensity of the current pixel is multiplied with the change in the contrast in each pixel and then it is divided by the intensity of the current pixel.

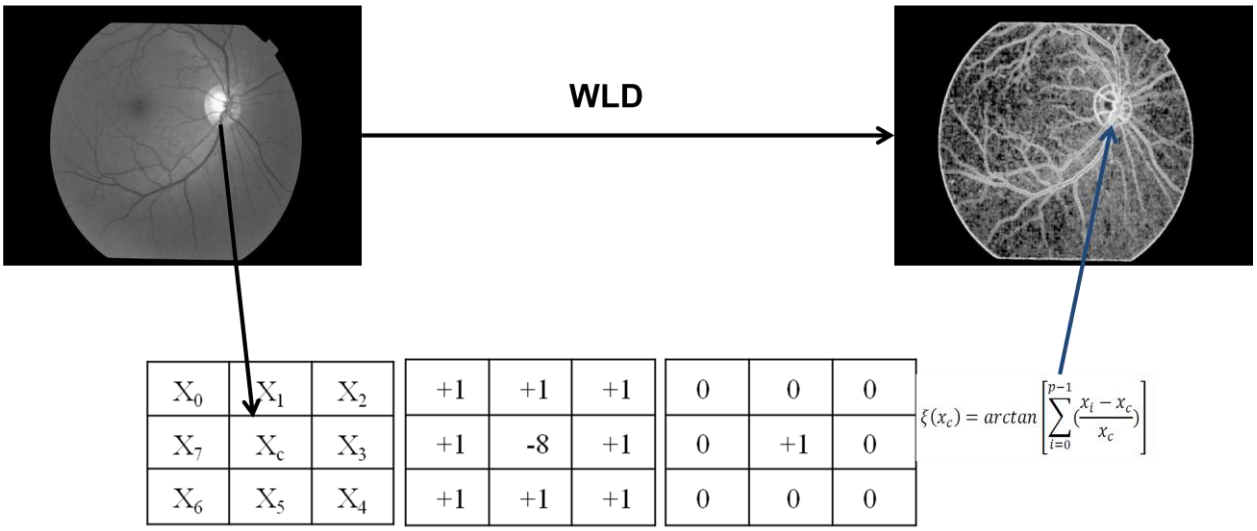


Figure 4.3: computing Differential Excitation for a single pixel in an image

The above figure explains that how the differential excitation is calculated for a single pixel in a digital fundus image. The two filters in the figure represent the difference of current pixel with the neighbors and the intensity of the current pixel respectively.

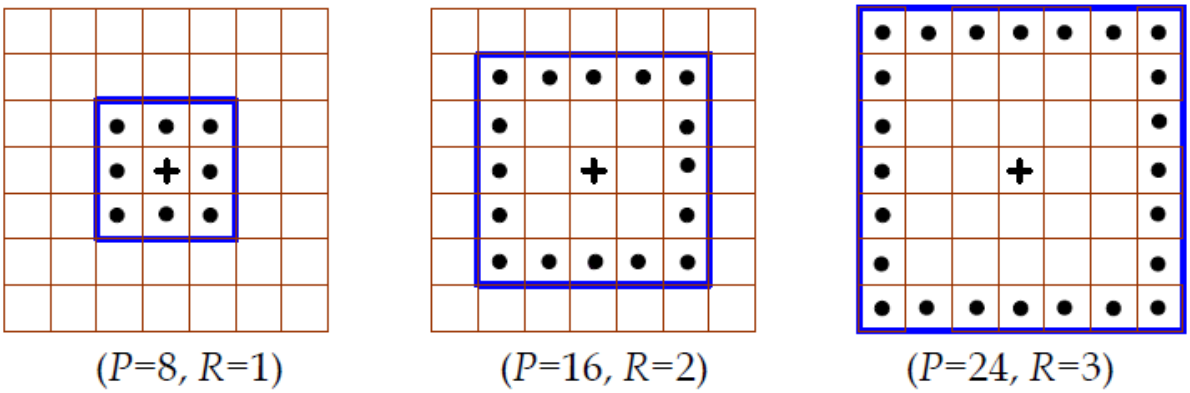


Figure 4.4: Square Symmetric Neighborhood of different (P, R) [24]

As it is already discussed that the differential excitation component of weber local descriptor depends upon the intensity of the neighoring pixels, so the above figures explains how the neighboring pixels are selected. The neighbor pixels are selected based on their distance from the current pixel. If we use R=1 that means that pixels which are at distance 1 will be selected as the neighbors. If we use R=2 then the pixels at distance 2 from the current pixel will be selected as the neighbors and so on. We use R=1 with 8 neighbors in our thesis work. As shown in the above figure that once the distance from the current pixel is increased the number of neighbors is also increased.

ii) Orientation:

The second component of weber local descriptor is the orientation which is the gradient orientation. It can be computed as:

$$\theta(x_c) = \gamma_s^1 = \arctan\left(\frac{v_s^{11}}{v_s^{10}}\right) \quad \text{.....} \quad \textbf{(4.5)}$$

In the above equation v_s^{10} and v_s^{11} are the outputs of the filters f_s^{10} and f_s^{11} which are the difference of the vertical and horizontal neighbors respectively.

$$\boldsymbol{v_s^{10} = x_5 - x_1} \text{ and}$$

$$\boldsymbol{v_s^{11} = x_7 - x_3}$$

x_0	x_1	x_2
x_7	x_c	x_3
x_6	x_5	x_4

Figure 4.5: The position of the current pixel and neighboring pixels

We have used only the differential excitation component of WLD for the extraction of features in our thesis work to minimize the computational complexity and differential excitation is good at detecting edges and it is robust to noise and illumination change.

After extracting the features, the next step in our thesis work is feature matching in the images and finding the inliers and outliers in the images. So when the feature matching is done, the inliers in both the images are matched.

Figure 4.6 which shows all the possible inline matches of the digital fundus images is given below:

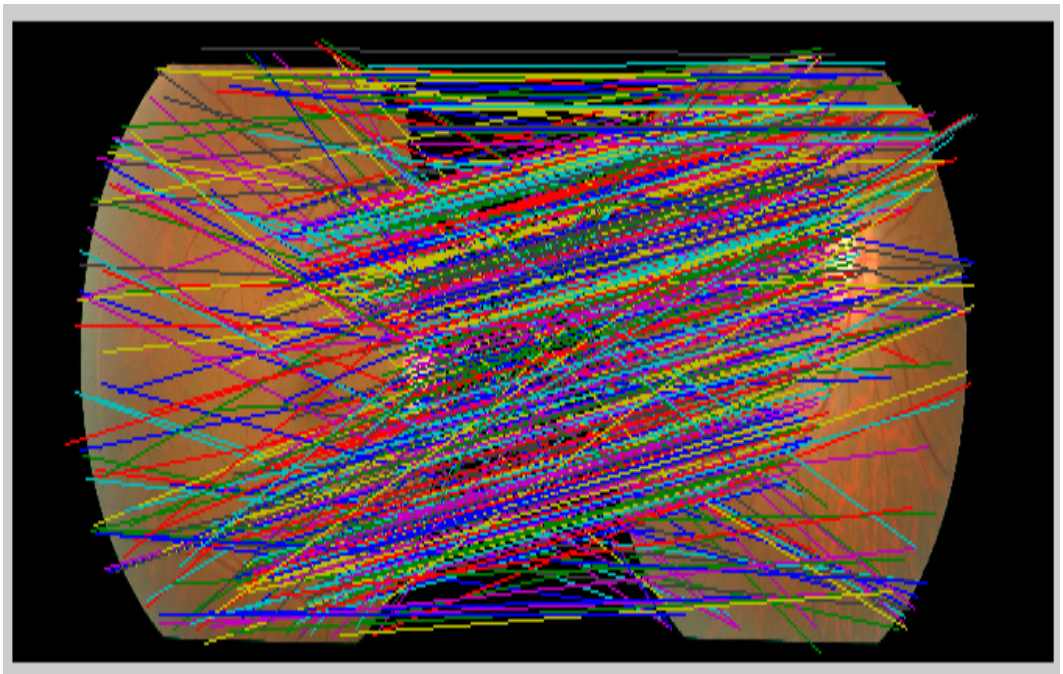


Figure 4.6: The all possible inliers matches

The above figure shows all the possible inliers combination between the two images. It is the first step of the RANSAC algorithm and as we know that RANAC is an iterative process so Figure 4.6 will have a high number of outliers as well which will be removed in the later stages. In the second step the combination with the maximum number of inliers is chosen.

Figure 4.7 shows the matching of refined inliers:

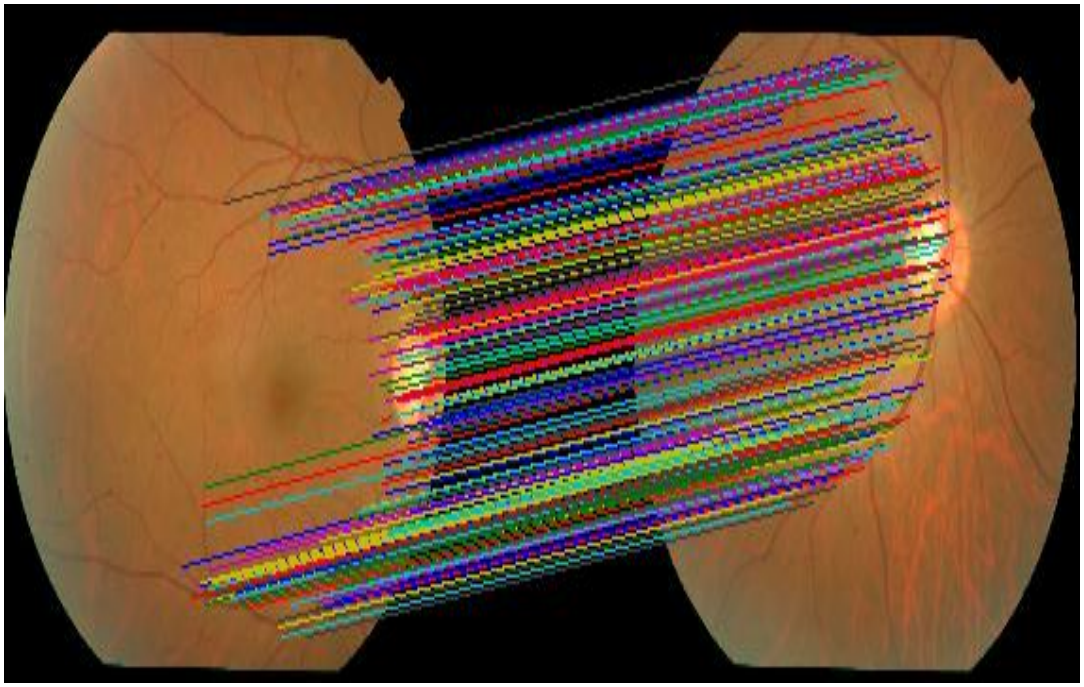


Figure 4.7: Matching of refined Inliers

The above figure shows the maximum number of inliers combination between the two images. In the next step these inliers are compared against the set threshold. If the inliers are greater than the set threshold, the images are wrapped together and mosaic is formed. So after the completion of this step a panorama using 2 images is built which means that half of the work will be done after the next step as we want to make a panorama of 3 images.

Figure 4.8 shows the mosaic generated after stitching the first two fundus images.

Figure 4.8: The final mosaic of the two images

So after the completion of the above step the final mosaic of two images is created as shown in the figure 4.8 but the algorithm proposed in this thesis for creating mosaic is not just 2 fundus images but instead the algorithm proposed in the thesis is to create mosaic for multiple fundus images (in this case three images). It is an iterative method consisting of 2 main steps which are mentioned as follow:

1. A similarity matrix is created by finding the inliers between each pair of images available. The inliers for each pair are calculated only once to save the computation cost so that the algorithm may work faster. The pair of images having maximum inliers is found and is wrapped to generate the mosaic. After the completion of this step mosaic of two images is created.
2. After the mosaic is generated between two images, the next step is to take the wrapped mosaic of the first two images and then compare it with all the other images. The image with the maximum similarity is found and is then wrapped with the already existing mosaic to form a panorama of more than two images. So in this way the mosaic for three images is generated.

The second step is repeated again and again until all the images are wrapped to generate the final mosaic. The final mosaic created is then smoothed using morphological process of closing with diamond as structuring element. In our thesis we are generating a mosaic for three images which have a larger field of view covering the entire interior portion of the eye. So in our thesis, after the completion of the second step the final mosaic is smoothed using morphological process and with that the implementation of our work ends.

Figure 4.9 shows the final mosaic after stitching all three fundus images from the same patient, thus creating an image having a large field of view as compared to the original fundus images which had a limited field of view.

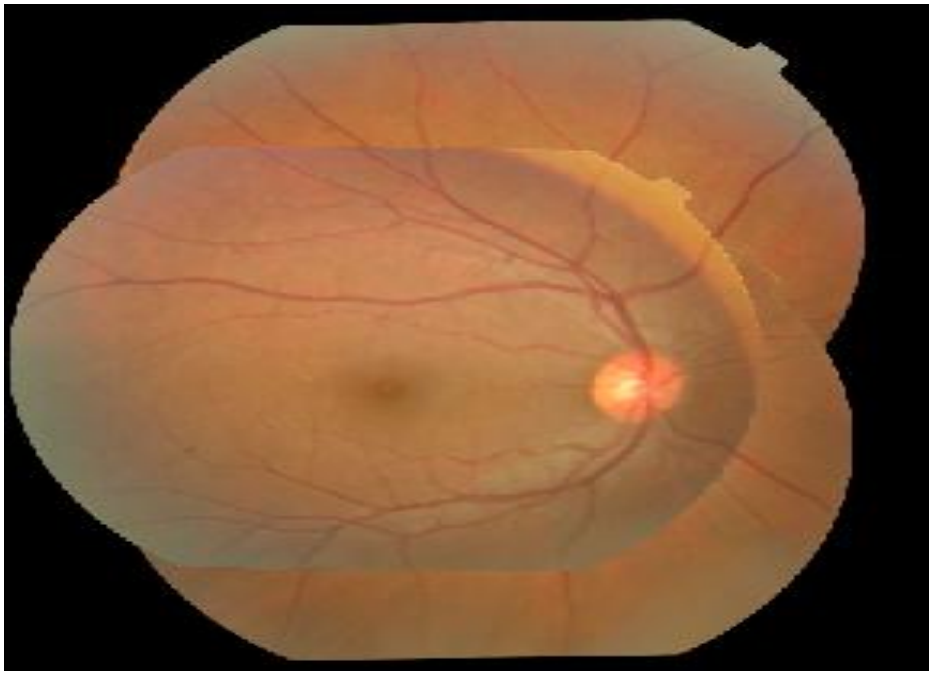


Figure 4.9: The final mosaic image after smoothing

4.3 Blending for Seamless Stitching:

A new algorithm has been proposed to achieve seamless stitching. Following are the steps for proposed algorithm for seamless stitching.

- Generation of **mask1** and **mask2**, the BG mask of both of the input images (image1, image2).
- Obtaining **fmask1**, by filtering the mask1 with Gaussian low pass filter in freq. domain.

- **Multiplication** of the image1 with fmask1.
- **Multiplication** of the image2 with inverted fmask2.
- The **result** after the summation of both the multiplications.
- Obtaining the **result1** by masking the result with the mask2.
- Obtaining the **result2** by masking the imag1 with a new mask obtain from the X-OR of mask1, mask2.
- The **final result** after summation of the result1 and result2.

The flow diagram of the proposed system used for seamless stitching is given below:

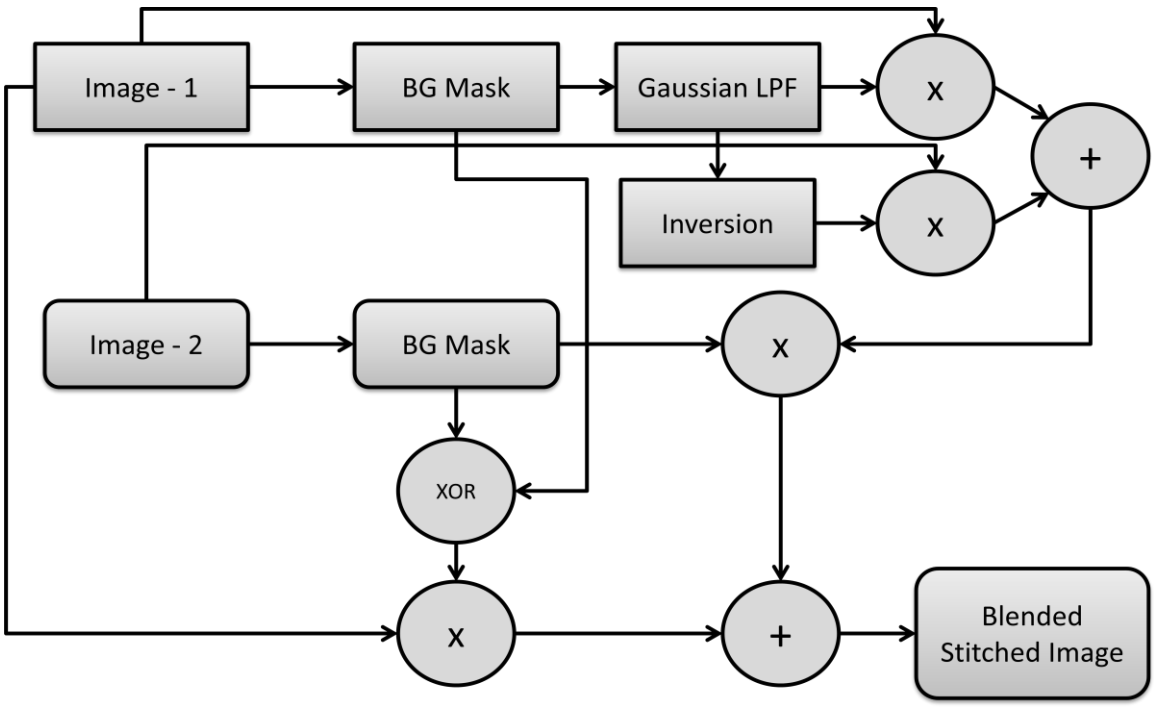


Figure 4.10: Flow diagram of proposed system to achieve seamless stitching

Figure 4.11 shows the comparison of applying seamless blending and normal stitching (without applying seamless blending algorithm)

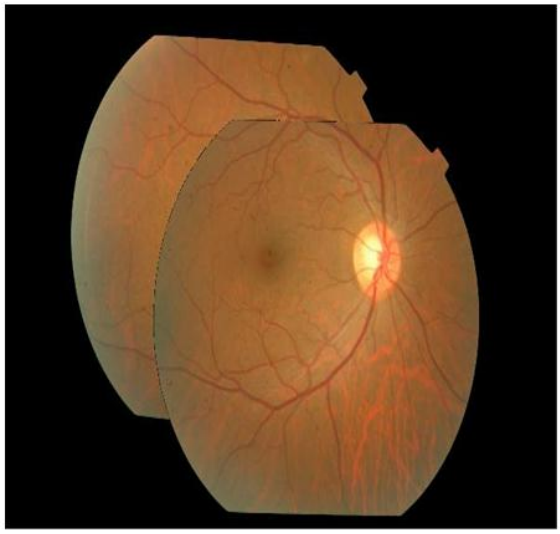


Figure 4.11: Stitching without Blending stitching after blending

4.4 Summary:

In this chapter the methodology used in our thesis work was discussed. For the purpose of feature extraction Weber local descriptor is used. For the removal of the outliers RANSAC is used. After removing the outliers, the images are wrapped together to form a mosaic for two images. To generate a mosaic for more than two images the already mosaic image is wrapped with a new fundus image and for that the whole algorithm is repeated.

Chapter 5

RESULTS

In this chapter we discuss the results produced after the implementation of our thesis work and after that the results produced by our thesis work are compared with other state-of the art techniques. We will discuss about the dataset of the images we have used in our thesis work as well.

The evaluation of the proposed system has been done on a locally gathered dataset. The dataset is collected with the help of Armed forces institute of ophthalmology (AFIO). The dataset is consisted of 56 images taken from Topcon non mydriatic camera NW 200. The images are from 15 different patients with varying number of images per patient. All the fifteen patients have at-least three images in the dataset and the maximum number of images of a patient in the dataset is six. The validity of proposed system has been tested using visual inspection.

5.1 Results:

The proposed system and other state-of the art techniques were applied on this dataset and the results showed that the performance of the proposed system was better than the other state-of the art techniques. The proposed method was applied to the whole dataset but here we will show our results on just three images of a patient and then we will derive our conclusion based on the results.

Figure 5.1 (a), Figure 5.1 (b) and figure 5.1 (c) show the three images of a patient which were used as an input for the proposed technique and the other state-of the art techniques as well.

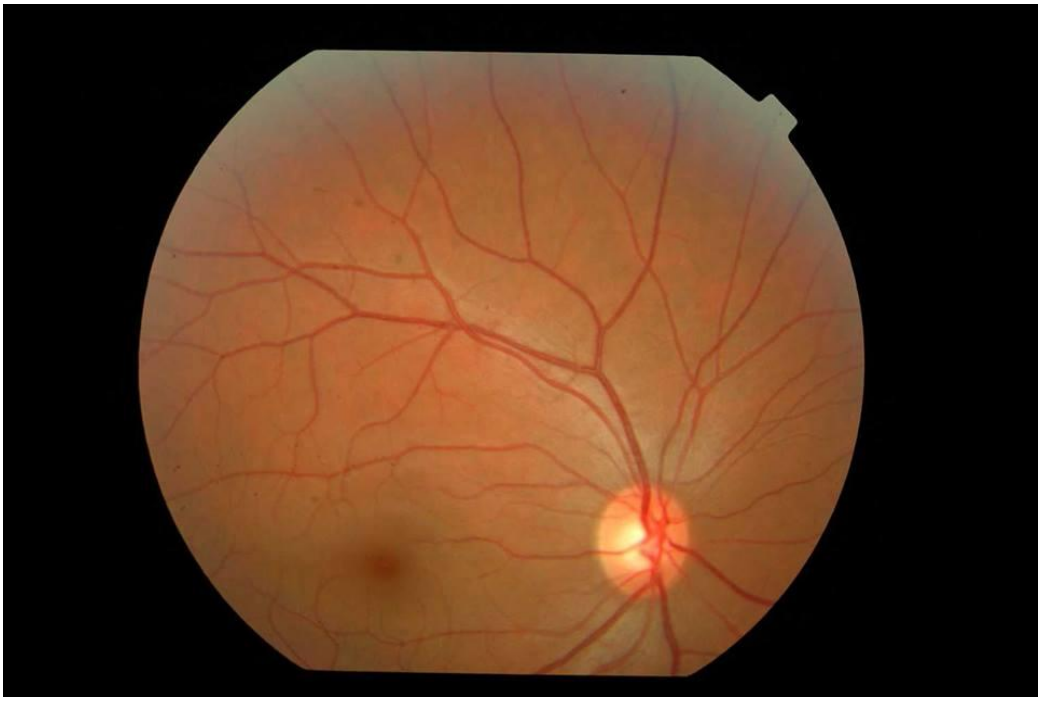


Fig 5.1 (a): first fundus image of the patient

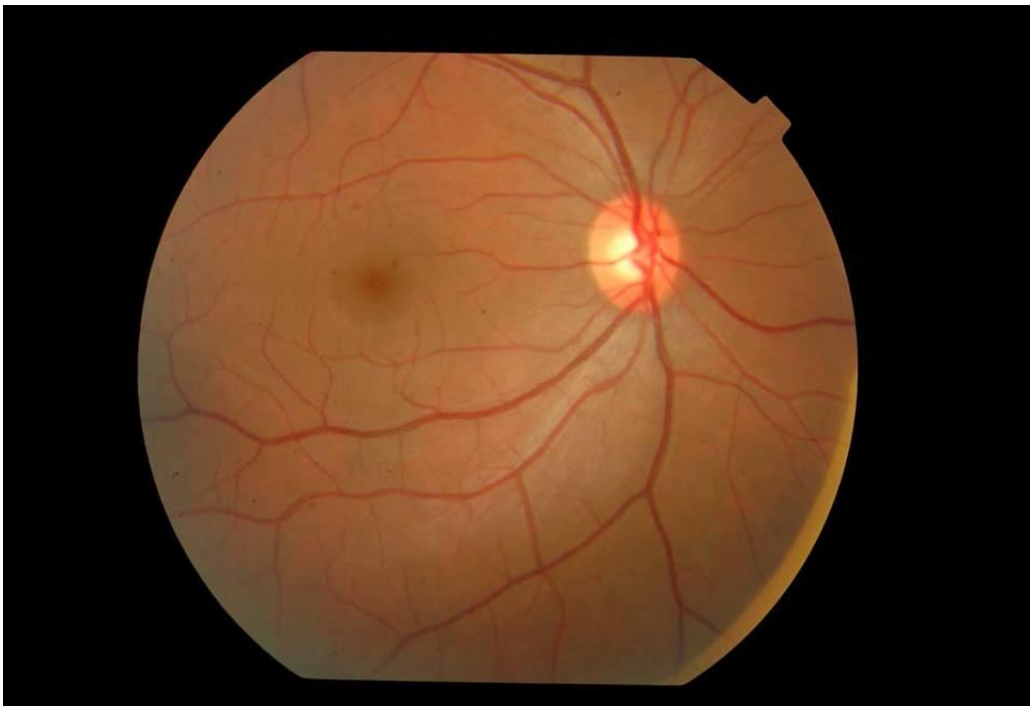


Figure 5.1 (b): 2nd Fundus image of the patient

Figure 5.1 (c): 3rd fundus image of the patient

To compare the results of the proposed method with the other state-of the art methods, first of all Harris corner detector was applied to the three input images. Based on the feature extraction of the Harris Corner detector the three images were stitched together. The results showed that the mosaic image was not up to the mark as the feature extraction of the Harris corner Detector for the given the dataset was not accurate. In most of the images the corners were either ignored or there were mismatched corners. The advantage of this technique was that it was very fast but the results were not accurate.

Figure 5.2 shows the mosaic image formed from the three input fundus images when Harris corner technique was applied on them.

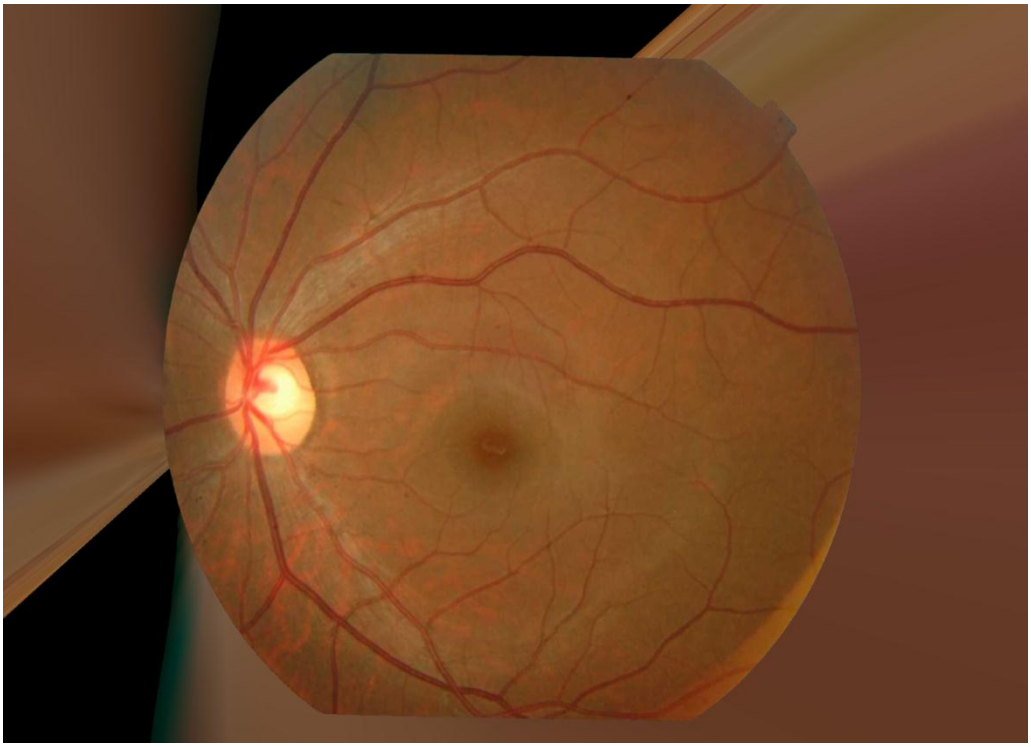


Figure 5.2: mosaic image by using Harris Corner Detector Technique

After using the Harris corner detector for the stitching of the input fundus images, the next technique we applied on these images was the SURF technique. So based on Surf technique the features were extracted and then based on these feature the mosaic of these fundus images was built by the same methodology that was discussed in chapter 4. The results showed that the mosaic of the images built by using SURF was not accurate. Most of the times the SURF feature extractor was not able to extract the interest points from the given fundus images. The advantage of SURF technique was that it was faster as compared to SIFT and ASIFT but its results were not accurate as compared to SIFT and ASIFT.

Figure 5.3 shows the result when mosaic of the input fundus images was formed by using the SURF technique.

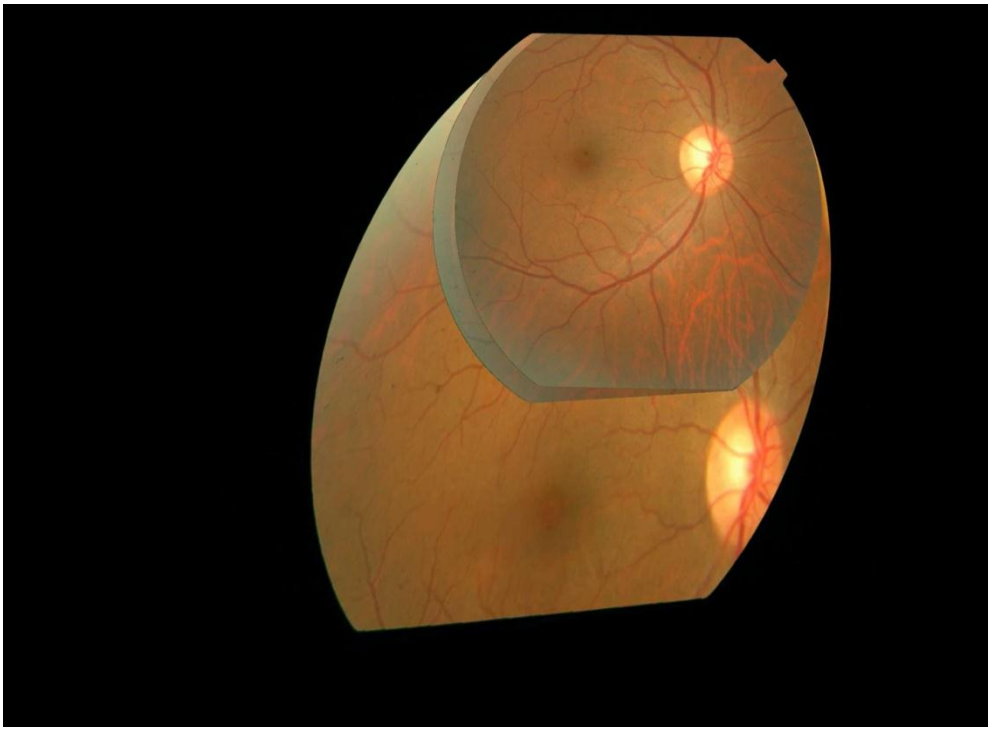


Figure 5.3: mosaic of the images when SURF was used for the feature extraction

After applying the Harris Corner Detector and SURF techniques on the given fundus images we applied the SIFT technique on the given input images. So SIFT was applied on the given images for the feature extraction which extracted the interest points better than the SURF and the Harris corner detector. Then the mosaic of the three fundus images was formed by the same method as discussed in the chapter 4. The results showed that mosaic of the images formed by the method of SIFT was inconsistent. Sometimes the results were too much below far and sometimes the results were close to being accurate. After applying the SIFT technique on the fundus images we found out that the SIFT technique was marginally slow as compared to SURF and Harris Corner detector but its results were better than that of the SURF and Harris corner Detector.

Figure 5.4 shows the mosaic image made from the three fundus images when the SIFT technique was used for feature extraction and feature matching.

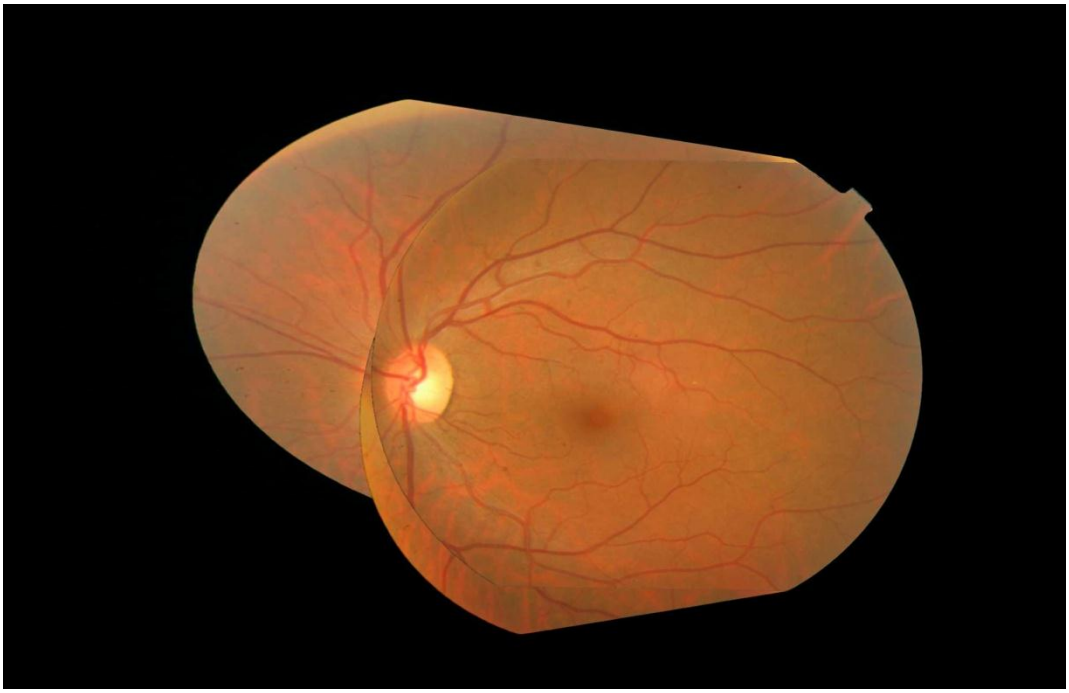


Figure 5.4: mosaic image using the SIFT technique

After applying Harris Corner detector, SURF and SIFT we applied ASIFT so that the distortion in the camera orientation can be simulated, then SIFT was applied on the fundus images so that the features can be extracted and matched. After that RANSAC algorithm is used to find the similarity between the images. After applying the RANSAC algorithm the images are wrapped together. Then this process was repeated to wrap the third image with the already built mosaic image as discussed earlier in the chapter 4.

The results show that Affine SIFT is better than the other state-of the art techniques. ASIFT is the best among all the sparse descriptors when it comes to feature extraction and feature matching. It is the best technique for extracting the interest points and it contains very few mismatches. But the main problem with this method is that it takes a lot of time as compared to the other techniques.

Figure 5.5 shows the result of ASIFT technique when it was applied to the input fundus images to build a mosaic image and it shows that the mosaic built is accurate as compared to the other techniques discussed above.

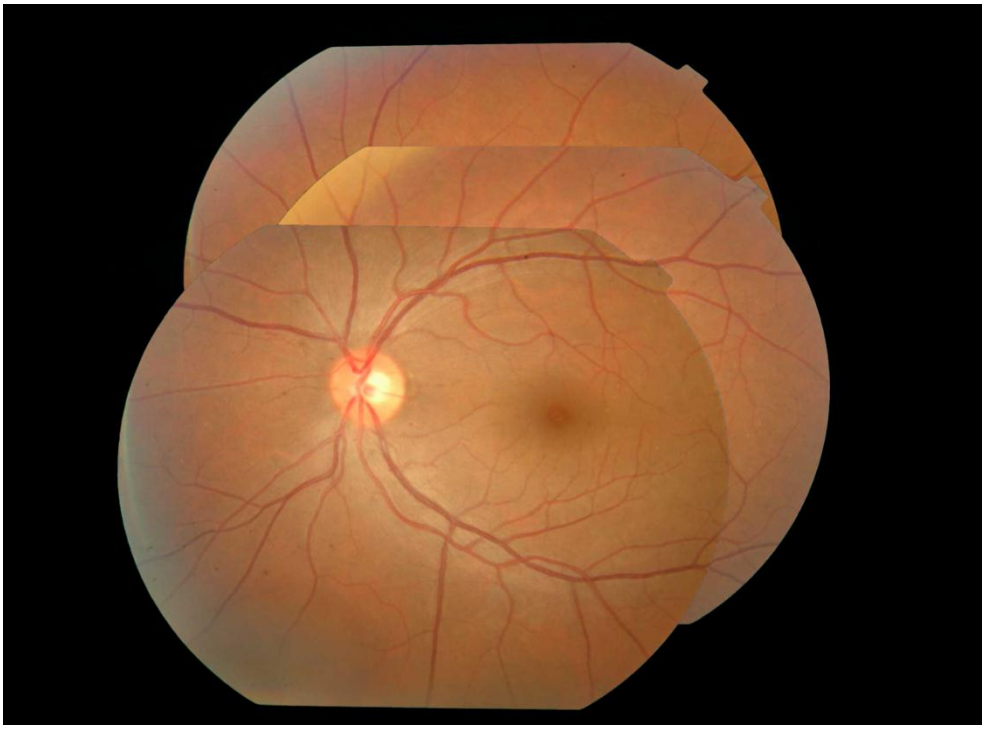


Figure 5.5: Image mosaic based on ASIFT technique.

So the above image shows that ASIFT is better than the other techniques we have used for the image stitching of digital fundus images.

After applying all these state-of-the art techniques on our dataset, we then applied our propose system on the given dataset of digital fundus images. In our proposed method we have used Weber local descriptor for the feature extraction, which works on the similarity of neighboring pixels. After the feature extraction, the feature matching is done and all the possible inliers are calculated in the images and the outliers are removed using RANSAC, then the homography of the images is calculated. One image is taken as a reference image while the other images are scaled and rotated in such a way that the images are aligned with the reference image. Once the images are aligned seamless blending algorithm is applied to the images to achieve seamless blending.

Figure 5.6 shows the stitching of our proposed method with the seamless stitching applied on 3 digital fundus images

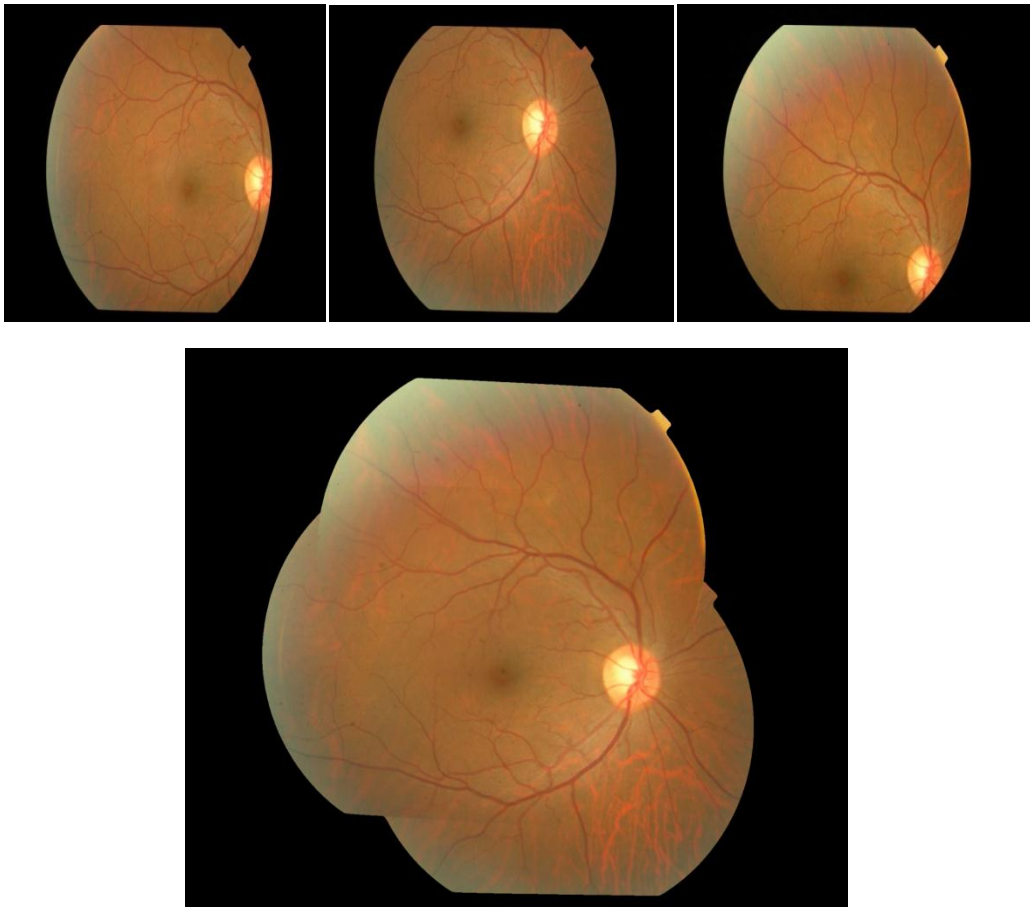


Figure 5.6: Mosaic of 3 fundus images.

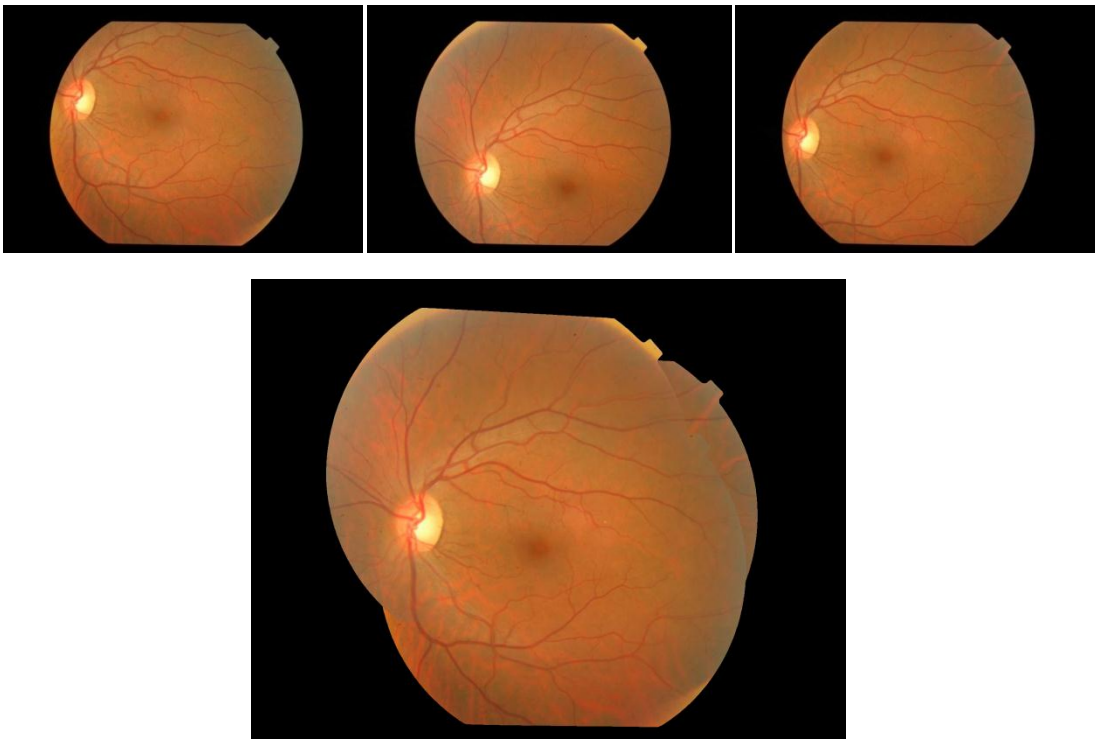


Figure 5.7: Mosaic Image of 3 Fundus Images

We have applied our proposed algorithm on 5 images as well; the result of it is given as follow:

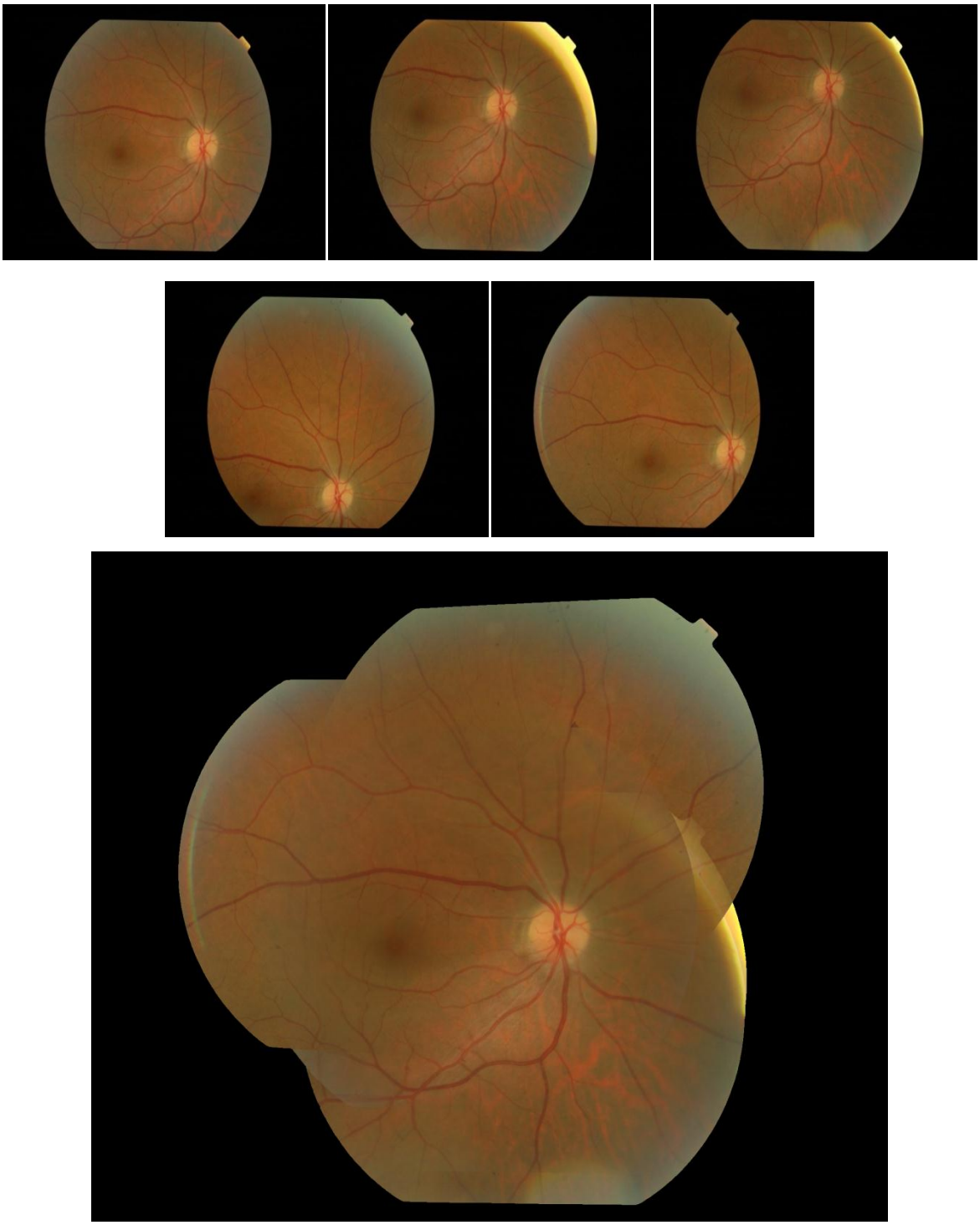


Figure 5.8: Result of the 5 images stitching

We have applied our proposed algorithm on 6 images as well; and it produced good results. The results are given as follow:

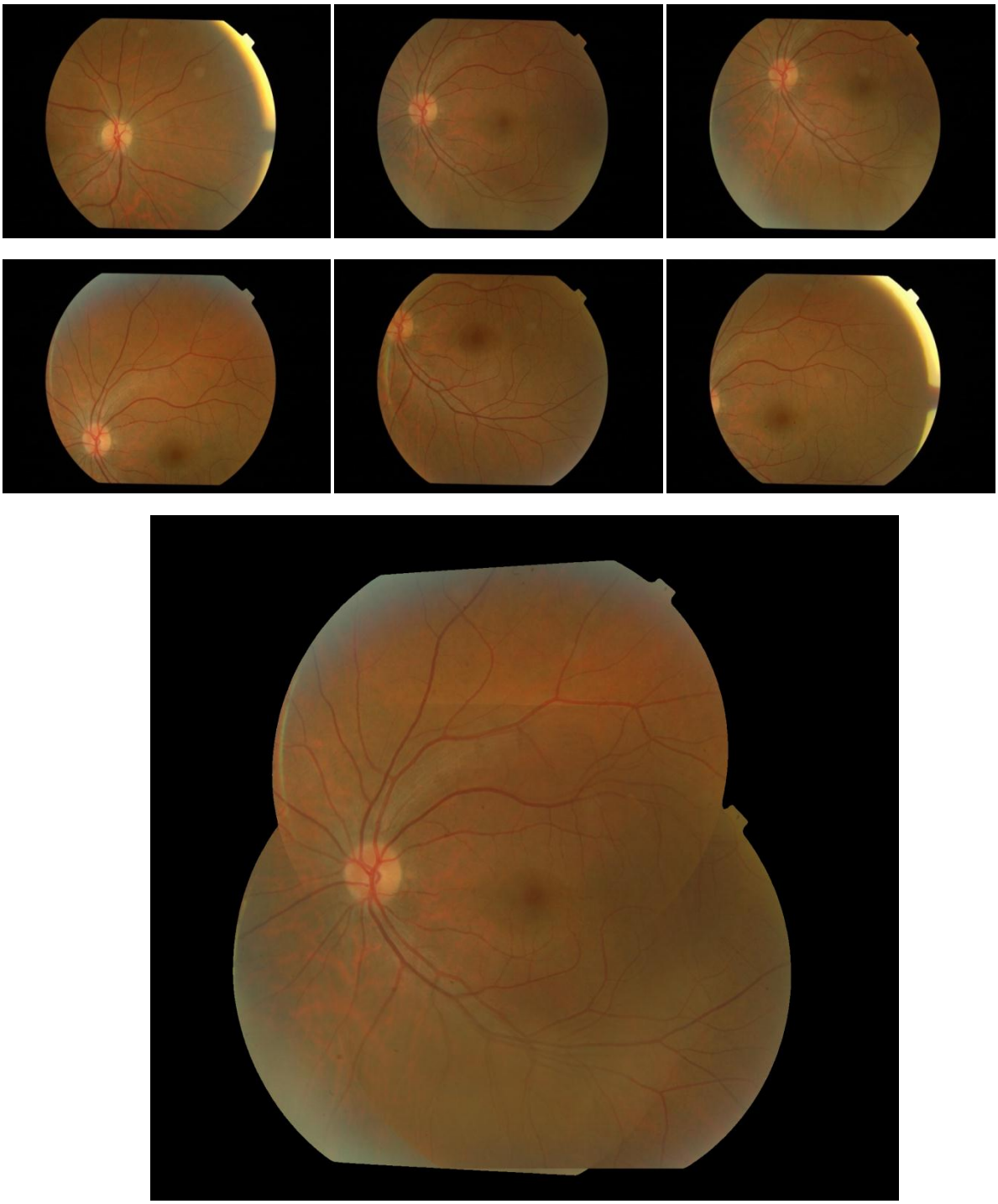


Figure 5.9: Result of the 6 images stitching

5.2 Performance:

We have evaluated the performance of our proposed algorithm by comparing it with the other state-of-the art techniques based on two criteria:

1) Accuracy and
2) Time

1) Accuracy:

We measure the accuracy of the results based on the continuity of the vessels in the images as perceived by the human. The accuracy of our proposed algorithm is compared with the other state-of the art techniques. The results are shown in the following tables.

Techniques Used	Total Patients	Correct	Accuracy %
SURF	15	11	73.33
SIFT	15	15	100
ASIFT	15	15	100
WLD	15	15	100

Table 5.1: Comparison table of two images stitching

It shows that SIFT, ASIFT and WLD performed the same over two images, while the results of SURF were poor among all the methods used.

Techniques Used	Total Patients	Correct	Accuracy %
SURF	15	10	66.6
SIFT	15	13	86.6
ASIFT	15	15	100
WLD	15	15	100

Table 5.2: Comparison table of three images stitching

Table 5.2 shows that ASIFT and WLD produced the same results while the results of SIFT and SURF were not up to the mark.

Techniques Used	Total Patients	Correct	Accuracy %
SURF	5	3	60
SIFT	5	4	80
ASIFT	5	5	100
WLD	5	5	100

Table 5.3: Comparison table of four images stitching

The image stitching on 4 images show that ASIFT and WLD produced the best results while the results of SIFT and SURF were not accurate.

Techniques Used	Total Patients	Correct	Accuracy %
SURF	2	0	0
SIFT	2	1	50
ASIFT	2	2	100
WLD	2	2	100

Table 5.4: Comparison table of 6 images stitching.

The mosaic generation for 6 images show that ASIFT and WLD produced the accurate results while the results of SIFT were inconsistent. The results of SURF were extremely poor.

2) Timing Table:

The average time taken by the various techniques is given in the table 05. We used on a Sony core i3 system with 1.90 GHz processor and 4 GB RAM

Techniques Used	2- Images	3- Images
SURF	13.78	68.13
SIFT	75.14	116.96
ASIFT	133.12	183.28
WLD	63.13	95.24

Table 5.5: Timing Table of various techniques on two and three images stitching

The above table shows that WLD produce faster results as compared to SIFT and ASIFT techniques.

5.3 Summary:

In this chapter the results of our proposed method and its comparison with the other state-of the art methods are shown. The accuracy and the timing table show that our proposed technique produced comparable results with the other state-of the art techniques.

Chapter 6

CONCLUSION AND FUTURE WORK

6.1 Conclusion:

The problem of limited Field of View (FOV) of fundus cameras for the diagnosis of the eye disease is solved in our thesis work. Single fundus image has a limited field of view so to extend the field of view image stitching is used. The algorithm proposed in our thesis work uses WLD descriptors to extract features from the fundus image. Features are then matched and RANSAC algorithm is applied for removing the outliers. The images are then stitched using the steps defined above to attain autonomous mosaic of more than two fundus images. The algorithm is applied to the dataset of 56 images from 15 patients. In the dataset each patient has different number of images but every patient has at least 3 images. The results of the proposed method are compared with other state-of the art techniques such as SURF, SIFT and ASIFT to judge the accuracy and the results show that our proposed method produced comparable results with the other state-of the art methods based on the accuracy and time.

6.2 Future Work:

Our proposed method gives comparable results to generate a mosaic from multiple digital fundus images. The images we have used in our thesis are of healthy human beings. In future work images of patients having eye diseases will be added and the results will be compared with other techniques such as Local binary pattern which is a dense descriptor. The other important part of the future work will be the smoothing of overall boundary of the stitched image.

References

1. Yuliang Wang, Jianxin Shen, Wenhe Liao and Lin Zho: Automatic Fundus Images Mosaic Based on SIFT Feature, (2010)
2. Richard Szeliski: Image Alignment and Stitching: A Tutorial, (2005)
3. Nice, France, Matthew Brown and David G. Lowe: "Recognising panoramas," International Conference on Computer Vision *(ICCV 2003)*, (2003)
4. David Lowe: "Object recognition from local scale-invariant features". Proceedings of the International Conference on Computer Vision, (1999)
5. Richard Szeliski: Computer Vision: Algorithms and Applications, (2010)
6. Homographies and RANSAC MIT CSAIL (2006)
7. David Kriegman: Homography Estimation, (2007)
8. M. A. Fischler and R. C. Bolles: Random sample consensus: A paradigm for model-fitting with applications to image analysis and automated cartography, Communications of the ACM 24. 381–395 (1981)
9. Marco Zuliani: RANSAC toolbox for Matlab™ & Octave
10. Herbert Bay, Andreas Ess, Tinne Tuytelaars and Luc Van Gool: SURF: Speeded Up Robust Features, Computer Vision and Image Understanding (CVIU), (2008)
11. Jean-Michel Morel and Guoshen Yu: ASIFT: A new framework for fully Affine Invariant image comparison, (2011)
12. Azad, P. Asfour, T.; Dillmann and R.,Oct: Combining Harris interest points and the SIFT descriptor for fast scale-invariant object recognition, (2009)
13. A.M.Mendonca and A. Campilho: Segmentation of retinal blood vessels by combining the detection of centerlines and morphological reconstruction. IEEE Trans. Med. Imag. 25(9), 1200-1213 (2006)

14. J. M. Tielsch, J. Katz, K. Singh, H. A. Quigley, J. D. Gottsch, J. Javitt, and A. Sommer: "A population-based evaluation of glaucoma screening: The Baltimore eye survey," Amer. J. Epidemiol., vol. 134, no. 10, pp. 1102–1110, (1991)
15. O. Kean: Pathogenetic Mechanisms of Retinal Detachment, in Retina, ed. Ryan, S.J., Elsevier Health Sciences, Philadelphia, PA, 2013-15, (2006)
16. Early Treatment Diabetic Retinopathy Study research group, "Early photocoagulation for diabetic retinopathy. ETDRS Rep. 9," *Ophthalmology*, vol. 98, no. 5 suppl, pp. 766–785, (1991)
17. T. A. Ciulla, A. G. Amador, and B. Zinman: Diabetic retinopathy and diabetic macular edema: Pathophysiology, screening, and novel therapies, *Diabetes Care*, vol. 26, no. 9, pp. 2653–2664, (2003)
18. Age-Related Eye Disease Study Research Group, "A randomized, placebo-controlled, clinical trial of high-dose supplementation with vitamins C and E, beta carotene, and zinc for age-related macular degeneration and vision loss: AREDS Rep. 8," *Arch. Ophthalmol.*, vol. 119, pp. 1417–1436, (2001)
19. A. Heijl, M. C. Leske, B. Bengtsson, L. Hyman, B. Bengtsson, and M. Hussein: Early Manifest Glaucoma Trial Group, "Reduction of intraocular pressure and glaucoma progression: Results from the early manifest glaucoma trial," *Arch. Ophthalmol.*, vol. 120, no. 10, pp.1268–1279, (2002)
20. A. Hoover, V. Kouznetsova, and M. Goldbaum: Locating blood vessels in retinal images by piecewise threshold probing of a matched _lter response. IEEE Trans. Med. Imag.19(3), 203-210 (2000)

21. X. Jiang and D. Mojon: Adaptive local thresholding by verification based multithreshold probing with application to vessel detection in retinal images. IEEE Trans. Pattern Anal.Mach. Intell 25(1), 131-137 Jan. (2003)
22. Marco Zuliani : RANSAC for dummies, July 4, (2014)
23. Chris Harris and Mike Stephens: A combined corner and edge detection, Plessey Research Roke Manor, United Kingdom, (1988)
24. Jie Chen, Shiguang Shan, Chu He, Guoying Zhao, Matti Pietikäinen, Xilin Chen and Wen Gao : WLD: A Robust Local Image Descriptor, IEEE Transactions on Pattern analysis and Machine intelligence, (2009)
25. LI Jupeng, CHEN Houjin, YAO Chang and ZHANG Xinyuan: A Robust Feature Based Method for Mosaic of the curved Human Color Retinal Images, International Conference on BioMedical Engineering and Informatics, (2008)
26. LiFang Wei, LinLin Huang, Lin Pan and Lun Yu: The Retinal Image Mosaic Based on Invariant Feature and Hierarchical Transformation Models, 2nd International conference on Image And Signal Processing, (2009)
27. Philippe C. Cattin, Hobert Bay, Luc Van Gool and Gabor Szekely: Retina Mosaicing Using Local Features, Medical Image Computing and Computer-Assisted Intervention, (2006)
28. Tae Eun Choe, Isaac Cohen, Munwai Lee and Gérard Medioni: Optimal Global Mosaic Generation from Retinal Images, Proceedings of the 18th International Conference on Pattern Recognition, ICPR, (2006)