Automated methods in the diagnosing of retinal images

MARTHINA JÖNSSON

Bachelor of Science Thesis in Medical Engineering
Stockholm 2012
This bachelor thesis project was performed in collaboration with
KTH, School of Technology and Health
Supervisor at KTH: Hamed Muhammed

Automated methods in the diagnosing of retinal images

Automatiska metoder för diagnosticering av ögonbottenbilder

MARTHINA JÖNSSON

Bachelor of Science Thesis in Medical Engineering
Basic level, 15 credits
Supervisor at KTH: Hamed Muhammed
Examiner Mats Nilsson
School of Technology and Health
TRITA-STH. EX 2012:5

Royal Institute of Technology
KTH STH
SE-141 86 Flemingsberg, Sweden
http://www.kth.se/sth
Abstract

This report contains a summation of a variety of articles that have been read and analysed. Each article describes different methods that can be used to detect lesions, optic disks, drusen and exudates in retinal images. I.e. diagnose e.g. Diabetic Retinopathy and Age-Related Macular Degeneration. A general approach is presented, which all methods more or less is based on.

Methods to locate the optic disk

- The PCA
- kNN Regression
- Hough Transform
- Fuzzy Convergence
- Vessel Direction Matched Filter
- Etc.

The best method based on result, reliability, number of images and publisher is kNN regression. The result of this method is remarkably good and that brings some doubt about its reliability. Though the method was published at IEEE and that gives the method a more trustful look. A next best method which also is very useful is Vessel Direction Matched Filter.

Methods to detect drusen – diagnose Age-Related Macular Degeneration

- PNN classifier
- Histogram approach
- Etc.

The best method based on result, reliability, number of images and publisher is the PNN classifier. The method had a sensitivity of 94% and a specificity of 95%. 300 images were used in the experiment which was published by the IEEE in 2011.

Methods to detect exudates – diagnose Diabetic Retinopathy

- Morphological techniques
- Luv colour space, Wiener filter an Canny edge detector.

The best method based on result, reliability, number of images and publisher is an experiment called “Feature Extraction”. The method includes the Luv colour space, Wiener filter (remove noise) and the Canny edge detector.

Keywords: retina, automatic, automated, analysis, image, optic disk, macula, age-related macular degeneration, drusen, diabetic retinopathy and diagnoses.
Sammanfattning

Den här rapporten innehåller en sammanfattning av ett flertal artiklar som har blivit studerade. Varje artikel har beskrivit en metod som kan användas för att upptäcka sjuka förändringar i ögonbottenbilder, det vill säga, åldersförändringar i gula fläcken och diabetisk retinopati.

Metoder för att lokalisera blinda fläcken

- PCA
- kNN regression
- Hough omvandling
- Suddig konvergens
- Filtrering beroende på kärlens riktning
- Mm.


Metoder för att diagnosticera åldersförändringar i gula fläcken

- PNN klassificeraren
- Histogram
- Mm.

Den bästa metoden baserat på resultat, pålitlighet, antal bilder och utgivare är PNN klassificeraren. Metoden hade en sensitivitet på 94 % och en specificitet på 95 %. 300 bilder användes i experimentet som publicerades av IEEE år 2011.

Metoder att diagnosticera diabetisk retinopati

- Morfologiska tekniker
- Luv colour space, Wiener filter and Canny edge detector.

Den bästa metoden baserat på resultat, pålitlighet, antal bilder och utgivare är ett experimentet som heter ”Feature Extraction”. Experimentet inkluderar Luv colour space, Wiener filter (brus borttagning) och Canny edge detector.

Sökord: retina, fundus, automatic, automated, analysis, image, optic disk, macula, age-related macular degeneration, drusen, diabetic retinopathy and diagnoses.
Table of Contents

1. Introduction ........................................................................................................................................... 1
   1.1 BACKGROUND.................................................................................................................................. 1
   1.2 SUBJECT ......................................................................................................................................... 1
   1.3 RESEARCH OBJECTIVES .............................................................................................................. 1
   1.4 LIMITATIONS ................................................................................................................................. 2
   1.5 METHODS ....................................................................................................................................... 2
   1.6 ASSUMPTIONS ............................................................................................................................... 2

2. Status report ......................................................................................................................................... 3

3. Basic theory ........................................................................................................................................... 5

4. Facts .................................................................................................................................................... 7
   4.1 GENERAL ......................................................................................................................................... 7
   4.2 DEFINITION OF TERMS .................................................................................................................. 7
   4.3 METHOD BASED ON THE GENERAL APPROACH ......................................................................... 10
      4.3.1 Example 1 - General Approach ............................................................................................... 10
   4.4 DETECTION OF THE OPTIC DISK ............................................................................................... 12
      4.4.1 Example 2 – The PCA Method ................................................................................................. 12
      4.4.2 Example 3 – kNN Regression .................................................................................................. 14
      4.4.3 Example 4 – Vessels Direction Matched Filter ....................................................................... 17
      4.4.4 Example 5 – Fuzzy Convergence .............................................................................................. 21
      4.4.5 Example 6 – Image Processing, Detect Optic Disk ................................................................ 24
      4.4.6 Example 7 – Geometrical Model of Vessel Structure ................................................................. 25
      4.4.7 Example 8 – Geometric Active Contours .................................................................................. 28
      4.4.8 Example 9 – Hough Transform .................................................................................................. 32
   4.5 AGE-RELATED MACULAR DEGENERATION .................................................................................. 34
      4.5.1 Example 10 – Drusen Detection ................................................................................................. 34
      4.5.2 Example 11 – Histogram Approach ........................................................................................... 37
      4.5.3 Example 12 – PNN classifier ..................................................................................................... 40
   4.6 DIABETIC RETINOPATHY ................................................................................................................ 42
      4.6.1 Example 13 – Image Processing, Detection of Exudates ............................................................ 42
      4.6.2 Example 14 – Classification of DR Lesions ............................................................................... 44
1. Introduction

1.1 Background

The retina is responsible for the conversion of incoming light to neural signals that is further processed in the visual cortex of the brain. The retina is a tissue that covers the interior of the eye which contains the scotoma (the blind spot, where nerve fibres enters the eye), the macula (that contain the fovea where the photoreceptors lie) and many blood vessels. These anatomical structures make the retina very important when it comes to diagnosis of many eye diseases and systemic diseases. The most frequent causes of blindness are often caused by glaucoma, diabetic retinopathy and the age-related macular degeneration. The best way to reach a diagnosis is to analyse images of the retina. The images can be taken noninvasively and be automatically processed [1] [2] [3].

There are an enormous amount of articles that describe everything from the first methods of diagnosing a retinal image to the automatic methods. In this study the fully automated methods are of most interest.

1.2 Subject

This assignment is to study automated methods in the diagnosing of retinal images and determine which methods that are presenting the best result. This is done by reading scientific articles that describes these methods. Each article describes different methods. After the articles have been read, they shall be analysed and compared to each other. In the end, the best methods shall be presented.

1.3 Research objectives

This study shall bring more knowledge about the latest methods for diagnosing retinal images. It is written for researchers, engineers or equal with knowledge in medical technology. The goal is to provide a distinct compilation of the methods used to analyse retinal images and knowledge of which methods that are presenting the best result.
1.4 Limitations

The articles that are being studied, treats the fully-automated methods for diagnosis of retinal images, i.e. articles published in the late 90’s and newer. For simplicity, only English written articles are read. The images represent the retina and diseases that present symptoms in the area of the blind spot (optic disk) and the macula, e.g. lesions. The focusing lies on articles that are published by the IEEE but other may be optional if the reliability is acceptable.

1.5 Method

The problem is solved by studying published scientific articles. These articles are to be found through Google Scholar. The methods shall be compiled in a systematic approach according to the limitations. This report comes together by the systematic compilation and interpretively of these articles. Internet searches leads to increased understanding of the topic and the correct definition for various terms. Those articles which seem reliable are categorised for a clearer presentation. Then the results from each article is analysed and conclusions are made. From these conclusions, the best methods are presented.

1.6 Assumptions

The oldest articles should contain more or less semi-automatic methods. As the time goes by the technique has developed and therefore should the fully-automated methods dominate. With a critical aspect of technique, the semi-automatic methods should be more reliable. A computer analyses and solves problems using algorithms which are created by general templates. I.e. nothing is exactly like the other. The technique has given us great possibilities in many medical fields and why not in this one.
2. Status report

KTH School of Technology and Health is a centre of education and research within the field area of Medical Technology. This report is the main part of a bachelor degree thesis of 15 ECTS in Medical Technology, which is made in cooperation between a student in Medical Technology and a PhD/researcher as a supervisor. The supervisor, who also suggested this topic, has medical imaging as research and lecture field.

The degree thesis is a literature study and therefore a theoretic work that includes this report and an oral presentation.
3. Basic theory

This degree thesis is based on earlier education years at KTH. That includes courses in Mathematics, Engineering and Information skills, Medical Imaging Systems, Anatomy, Physiology and Pathology. To fully understand this report you need to understand basic anatomical and physiological aspects of the eye. Some terms is described in chapter 4.2. This report is based on the knowledge that a student manages to receive and understand during a 15 ECTS course. To understand and analyse these methods, the following is required:

- A mathematical foundation on academic level.
- Understanding of medical imaging (including e.g. the importance of noise reduction).
- Understanding of basic pathology and physiology of e.g. vessels, blood and nerves.
- A foundation in English and medical terms.

There are so many published articles in this subject that it is impossible to address them all. The most important is to use those which seem most reliable, i.e. those which are mostly cited and that have reasonable results. There has not been published any similar summations of these methods before, but there have been small comparisons between results of different articles that share similar methods.
4. Facts

4.1 General

A general approach to an automated diagnose of a retina image is presented as follows:

- Pre-processing – A stage where the image quality is enhanced and noise is removed
- Locate anatomic structure – The optic nerve and blood vessels are detected
- Extract features – Defines e.g. texture of an image
- Screening process – Classification of retinal disease

4.2 Definition of Terms

[17]

Each example that follows from chapter 4.3.1 is a summation of an article. In the end of each headline, a [ ] can be read. The number inside the [ ] represents the reference for a specific article. All cites in one example are taken from that specific article. Other references may appear inside the examples, but not inside a cite. Those references are to explain certain terms that may need a further introduction. In the appendix, there are also some cites. Those are also taken from the specific article that is mentioned in the methods, respectively.

Contrast

The difference in visual properties that makes an object or its representation in an image distinguishable from other objects and the background.

Histogram

Histograms show the tonal distribution in a digital image. It plots the number of pixels for each tonal value. A histogram equalisation helps by distributing the intensity over the image. That leads to an under- or overexposed image to be more detailed.

Tresholding

Individual pixels are marked if their value is greater than a threshold value. The marked pixels are coloured black while the unmarked pixels are coloured white. The image is finally a binary image in black and white.

kNN and PNN

Two common ways to classify an image is using kNN and PNN. The kNN stands for “k-nearest neighbour algorithm” and is based on “close” training in the feature environment.
object is classified as a majority of its neighbours. If k=1 (K must be an integer), the object is assigned the same class as its closest neighbour. The PNN is based on probabilistic, e.g. it is very likely that the target value of an item is in the nearby neighbours.

**ASM**

ASM stands for Active Shape Model and is a statistical model of the shape of an object that is iterative deformed to suit another object. The good position is searched in the image, for example the existence of a sharp contour. Then the parameters of the model is updated to fit the new position.

**Lesions**

A lesion is a structural change that deviates from a healthy structure. They can be divided into background- and proliferative lesions. Background lesions are e.g. micro aneurysms, haemorrhages (bleedings – dot, blot and flame) [4], hard exudates, cotton wool spots (puffy white patches caused by damaged nerve fibres) [5], intraretinal microvascular anomalies and retinal oedema. Proliferative lesions are e.g. new vessels, preretinal haemorrhage, fibrous tissue and retinal detachment [6] [7] [8].

**Drusen**

"Drusens are deposits of cellular waste that accumulates beneath the retina". They are caused by Age-related Macular Degeneration (AMD) and they contribute to late-age blindness. They are seen as yellowish blobs in an image and they are classified as hard or soft. The hard drusen tend to be smaller and sharper but are generally less harmful than the soft ones. Soft drusen often lead to edema in macula and new vessel formations. The soft ones have very fuzzy boundaries and are therefore hard to discover [9]. Drusen detection is inhibited by number of difficulties, e.g. illumination that cause monotony is the images, object recognition and the alignment of images during the capture [10].

**Exudates**

“Hard exudates are yellowish intraretinal deposits, which are usually located in the posterior pole of the fundus”. It contains serum lipoproteins that have been leaked out from abnormal permeable blood vessels, especially from leaking micro aneurysms. They are a main hallmark of diabetic macular oedema. The exudates can also be soft and both of them are distinguished by their colour and sharpness of borders. “The grey level variation due to vessels is often as high as the one caused by the exudates”. The small hard exudates can often be confused with drusen [8].

**Hypertensive retinopathy**

The signs are vascular changes, like arteriolar narrowing and haemorrhages (bleeding). But it also includes patients with hypertension, e.g. micro aneurysms, retinal haemorrhage, or cotton wool spot [12].
The optic disc, OD

The optic disc a.k.a. the blind spot is where the optic nerve and the vessels enter the retina. It appears as a bright white/yellowish circular region. It can be used as a landmark that helps to reduce the search space and therefore simplify the analysis. The disc helps to determine the location of the macula, where lesions causing vision loss may appear. The optic disk is essential to be able to analyse the fovea. Its dimension is also used to measure abnormal feature due to glaucoma [11] [13].

Fovea

Central fovea of retina is a pit in the macula where the cones are located [14].

Age-Related Macular Degeneration, AMD

AMD is the main cause of the elderly blindness. Drusen is considered as the characteristic of AMD. It can be discovered during an inspection of the macula. AMD is classified as either wet/neovascular (WMD) or dry/non-neovascular (DMD). A WMD is more severe but fortunately less common. It is recognised by an extensive growth of blood vessels under the retina. A DMD is recognised by an attenuation of the retina and deposits that are being formed inside the retina [15].

Diabetic retinopathy, DR

Diabetic retinopathy is characterised with vascular changes in the retina. The signs can be divided into non-proliferative and proliferative retinopathy. There are many risks that come with diabetic, e.g. hyperglycaemia (high level of glucose in the blood), hyperlipidemia (high level of lipids in the blood) and hypertension (high blood pressure). These symptoms can reduce or increase the risk of retinopathy [12].

Specificity/Sensitivity

Specificity = \( \frac{B}{B+D} \) % healthy that received a negative test result [16].

Sensitivity = \( \frac{A}{A+C} \) % unhealthy that received a positive test result [16].

<table>
<thead>
<tr>
<th></th>
<th>Disease</th>
<th>Healthy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive test</td>
<td>A) True positive (TP)</td>
<td>B) False positive (FP)</td>
</tr>
<tr>
<td>Negative test</td>
<td>C) False negative (FN)</td>
<td>D) True negative (TN)</td>
</tr>
</tbody>
</table>

Figure 4.2.1 – Table explaining specificity and sensitivity.
4.3 Method Based on the General Approach

4.3.1 Example 1 – Automatic diagnosis of retinal diseases [17]

Proposed method
The system is based on the general approach with a special neural network technique. Signs of Drusen and Diabetic Retinopathy will be discovered.

Method

1. Pre-processing: Remove the noise from the retinal and therefore enable a reliable extraction of the features.
   a) The images are divided to non-overlapping blocks.
   b) A grey-scale conversion is made by extracting the RGB components.
   c) Enhancing of the contrast and image quality using histogram equalisation.
   d) A large median filter removes the noise from the image.

2. Locate anatomic structures and detecting lesions: Detect the optic nerve and vessels. The structures are bright features that are defined by linearity, connectivity, width and by Gaussian profile. The algorithm is based on the following:
   a) “Noise Reduction”.
   b) “Linear pattern with Gaussian-like profile improvement”.
   c) “Cross curvature evaluation”.
   d) “Linear filtering”.

This produces a binary image of the vasculature.

![Vascular segmentation with AMD (drusen)](image)

Algorithm 1 – Detection of Blood vessels boundaries using statistics:
Statistics like mean and standard deviation is used to detect boundaries. “The algorithm called DBDED, which stands for decision-based directional edge detector uses image statistics”. Each point in the image that passes a local threshold is
analysed. For the algorithm 1 – Detection of Blood vessels boundaries using statistics (See appendix 1).

Algorithm 2 – Extraction of blood vessel boundaries using deformable models
A recent method of contour detection and deformable models is snake. A snake is an active contour model that is manually initiated near to the contour of interest. This contour model deforms according to some criteria and image features to finally stay to the actual contour(s) in the image. An energy function is formulated to obtain an estimate of the quality of the mode in terms of its internal shape, and external forces e.g. underlying image forces and user constraint-forces. “The energy function integrates the weighted linear combination of the internal and external forces of the contour”.

Extraction of the core area of the blood vessel tree by tracing vessel centers:
“Algorithm: Extraction of blood vessel tree using the morphological reconstruction. Morphological reconstruction is to reconstruct an object in an image, called the marker image, containing at least one point belonging to that object from an image, called the mask image, containing that object and other objects and noise. An efficient implementation of morphological reconstruction can be described as follows:”

- Assign each connected component of the mask with a unique number.
- Determine which connected component that contains at least a pixel of the marker image.
- If the connected components are not of the previous ones, they can be removed.

3. Feature extraction
This can be done in two steps:
   a) Features detecting optic nerve.
   b) Features detecting diseases.

(See appendix 2).

4. Classification - Artificial Neural Networks (ANN)
ANN is a well-used method in medicine and medically related fields. A neural network is a signal-flow-model of self-learning algorithms for data processing, which tries to emulate the function in biological neurons. “Auto Associative Neural Network AANN is a network having the same number of neurons in input and output layers, and the less in the hidden layers”. The networks are being fed with the same input vector for it to learn the desired output value. This type of training leads to organised networks between the image layers. Each of the networks is independently trained for each class using the feature vector of the class. “The squared error between an input and the output is generally minimized by the network of the class to which the input pattern belongs. This enables a classification of an unknown input pattern. “The
unknown pattern is fed to all the networks, and is classified to the class with minimum squared error”.

Conclusion

This example presents a framework for diagnosing human retina diseases and can be implemented using Matlab. “Each module can be tested individually with a test data of size 100. “The results can be classified into four phases: true positive, true negative, false positive, false negative”. The major goal of this example is to provide a foundation for researchers, i.e. it can be extended to other retinal diseases.

4.4 Detection of the optic disc

4.4.1 Example 2 – The PCA Method [13]

Proposed system

Potential regions where the optic disk maybe located are called candidate regions. These regions are being determined by clustering the brightest pixels in an intensity image. A special analysis called PCA, (Principal component analysis) is being used at these regions. “The minimum distance between the original retinal image and its projection onto “disk space” is located as the center of optic disk”.

One strategy to locate the optic disk is the bottom-up process and another is the top-down process. This method describes a combination of both.

Candidate regions

If a pixel has the highest 1 % grey level in intensity image, the pixel is selected. A clustering mechanism assembles the nearby pixels into clusters. The clusters that have specified radius are combined to one cluster. If a cluster has less than 100 pixels after combination, the cluster is abandoned. The remaining clusters are candidate regions. Each candidate region is defined as a square of 120 x 120 pixels in a retinal image of 512 x 512 pixels and the diameter of the optic disk is in the range of 65 ~ 100. Therefore the possibility to miss the optic disk decreases with clusters bigger than 100 pixels.

Figure 4.4.1 – Example of candidate regions
PCA method

An important aspect when trying to locate the optic disk is the information of blood vessels. This approach involves calculation of eigenvectors (vector $x$ satisfy $Ax = \lambda x$, $A$ = square matrix and $\lambda$ is a constant) [18]. The vectors are calculated from training images, “(...) projecting the new retinal image to the space specified by the eigenvectors and calculating the distance between the retinal image and its projection. The center optic disk is located at the point of the minimum distance”. (See more in appendix 3)

Result and discussion

The component analyse described in the previous section is applied to each pixel (with different scale 0.8 ~ 1.1) in the input retinal image. The result of the location and the distance map $E$ of the input retinal image are shown below.

![Figure 4.4.2 – The result is marked with +](image)

![Figure 4.4.3 – the Euclidian distance map $E$, where the result is the dark spot](image)

“The pixel with the minimum distance $E$ in all the candidate regions and among all the scales is located as the center of the optic disk”.

The following two pictures are illustrated as the PCA method. The left picture represents the left eye, diameter of optics disk by 80. The right picture represents the right eye, diameter of optic disk 100.

![Figure 4.4.4 – The result of the proposed algorithm](image)

The result of the proposed algorithm indicates that the experiment was a success and can give a sufficiently accurate location of the optic disk. This proposed method has an improved precision compared to older methods in detecting optic disks in colour retinal images. The result shows that the proposed algorithm is robust and good in the presence of large area of light lesions.
4.4.2  Example 3 – kNN Regression [19]

Introduction

1100 images were tested and 1000 of them came from a screening program. The images were 540 pixels in diameter and they were taken with three different cameras. This proposed method finds the relationship between the variable d, the distance to the optic center, and a feature vector measured around a circular template. Two independent observers marked out the center of the optic disks and the optic disks borders for all images in the test set. The first observer was chosen as the reference standard.

The Optic Disc Distance Regression Model

By using kNN regression, the middle of the optic disk is located. The relationship between the distance d (distance to the optic disk center) and a feature vector is measured around a circular (to represent the optic disk) template. The template is placed at different locations in the image and measures the features around the template. The distance between the center of the template and the center of the optic disc is estimated by feature vector. “The template itself is divided into 4 sections (...). The template radius r is a free parameter which should be set during training”. A number of features are measured. These features consist of:

1. Number of vessels.
2. Average vessel width.
3. Standard Deviation of the vessel width.
4. Average vessel orientation.
5. Standard Deviation of the vessel orientation.
6. Maximum vessel width.
7. Maximum width vessel orientation.
8. Average intensity (green plane) under the template.
9. Standard deviation of the intensity (green plane) under the template.
10. Average Vessel width under the template.
11. Number of vessel pixels under the template.

“Features 1-7 are measured for each template piece and features 8-11 are measured once for the entire template. The entire measured feature vector thus consists of 32 features”.

![Figure 4.4.5 – The circular template, divided into 4 quadrants. The dot on the border is the point where feature measurements start. The radius r of the template is a free parameter.](image-url)
Vessel segmentation and analysis

In order to measure some of the features around the template, a segmentation and analysis of the vasculature is made. To do this, a pixel classification based approach was used [20]. This approach, which was supervised, was trained using images from DRIVE database. All images vascularity had been segmented [21]. To determine whether a pixel is a vessel pixel, a posterior probability map is produced. This map indicates for each pixel in the field of view, the probability that a pixel is a vessel pixel. A binary vessel segmentation is produced by thresholding this map. The vessel centrelines are determined by thinning the binary segmentation. This in turn enables the orientation of the vasculature to be measured [22].

“Next, individual vessel segments were found by removing all vessel bifurcations and crossings. This was necessary as the vessel orientation and width is not well defined in these points. Bifurcations and vessel crossings were removed by removing all pixels from the thinned image which have more than two neighbors”.

The local orientation of each vessel segment was determined by selecting each centreline pixel and its 3 neighbours, and the applying PCA (described in previous example) on their coordinates. “The direction of the largest eigenvector of the covariance matrix is the local orientation”. The local orientation is used to measure the width of the vasculature for each centreline pixel. “Perpendicular to this orientation, the distance from the centerline pixel to the edge of the vessel in the posterior probability map as produced by the vessel segmentation was measured”.

Feature Measurement

Once the segmentation and analysis of the vasculature has been completed, the previously mentioned features can be measured everywhere in the image. This “To measure features 1-7 we traversed the circular template starting at the position marked with the dot in (…)” See figure 4.4.5 “(…) and proceed along the template border in a clockwise fashion. When a centerline pixel is encountered, which is not part of a vessel segment previously encountered, this is counted as one vessel found in a certain quadrant. The vessel width and orientation were recorded as well. After a quadrant was completed, the vessel widths and orientations were averaged and their standard deviation was calculated”. Then the width and orientation of the widest vessel for this quadrant was stored. That is done for all quadrants and then the average pixel intensity of the green image plane under the template and standard deviation of the pixel intensities was measured. “The average vessel width was determined by dividing the total number of vessel pixels under the template by the number of centerline pixels”.

Training the System

Vessel segmentation and analysis was applied to all images in the training set. In every 10th pixel in a rectangular grid centered on the optic disk center and of size 2r squared was sampled. This was done by centering the template on a certain pixel and extracting the feature. “Each feature vector was stored together with a value d, the distance to the optic disc center. After the entire grid was sampled a number (for the present system, 200) of randomly
selected different locations in the image are also sampled. The maximum distance $d$ to the optic disc center was limited to the value of $r$ as it is unlikely the system is able to estimate a distance much larger than $r$. The complete set of samples for all training images formed the training dataset”. Before the set can be used all features were normalized to unit standard deviation. “The training set can then be used in the final system to estimate the distance from any location in the image to the optic disc center. To estimate the distance $d$ to the optic disc center given a feature vector, the $k=11$ nearest neighbors in the feature space are found and the values $d$ of these nearest neighbors are averaged. This average value is returned as the estimate of $d$. The value of $k=11$ was empirically determined using the training set”.

Feature Selection

The complete set of features may not be the optimal set of features to give the best regression results. The system performance can be enhanced by applying a feature selection method. “A supervised feature selection method was applied, called Sequential Floating Feature Selection (…)”[22]. That algorithm adds features to an empty feature set and removes features if that enhances the overall performance. This enables clustered groups of good features to be found. Performance was measured as the minimum average regression distance. The training set was randomly divided into a feature selection training and test set.

Then the following features were selected: “The number of vessels, as well as the orientation of the widest vessel, were selected in every quadrant, resulting in 8 features”. For the whole template, the average and standard deviation of the image intensity under the template as well as the average vessel width and number of vessel pixels was selected. The total number of features used in the system was 12”.

Applying the System

All pixels that are part of a vessel as indicated by the vessel segmentation, were searched to find the optic disk centre. The distance $d$ to the optic disk was determined for every vessel pixel. “Because the highest value returned by the regression is $r$, the distances for all pixels, not part of a vessel, are set to $r$”. Then the image is blurred with a Gaussian kernel $\sigma = 15$. “This value was determined empirically, but does not have a large influence on the final result of the method. In the end, the pixel with the lowest value in the image was selected as the optic disk centre. See figure 4.4.6 and 4.4.7.”
Experiment and results

The approach was applied to 1000 images in the test set. “Parameter r was chosen as 50 pixels, which makes the template a little larger than the average optic disc diameter of around 80”. The system was able to find the correct position of the optic disc border as indicated by the reference standard, in 99.9% of all cases. The average distance of the found OD center to the real OD center was 9.75 pixels with a standard deviation of 16.14 pixels. The second human observer was able to find the OD center in 100% of all cases. “Average time was 1 minute, 30 seconds for the vessel segmentation and 30 seconds for the analysis”.

Conclusion

The results are good. In 999 out of 1000 retinal images, the optic disk claims to have been located successfully. 10% of the images were pathology. The failed image, failed because of the vessel segmentation due to low contrast. This approach is recommended for fast optic disk location in images that are obtained in early diagnosis and screening projects.

4.4.3 Example 4 – Vessels’ Direction Matched Filter [24]

Material

Two datasets was used, STARE and DRIVE. [25].

STH KTH, Flemingsberg
Proposed method

A) A binary mask is generated – The mask technique is to mark out those pixels that belong to “region of interest” (ROI), and therefore exclude the background of the image. Same method as [26]. Threshold, t is applied to the red band of the image (t=35). “And then the morphological operators (opening, closing and erosion) were applied respectively (to the result of the preceding step) using a 3 x 3 square kernel to give the final ROI mask (…)”.

![Figure 4.4.8](image1)

Figure 4.4.8 – The left image shows the ROI mask and the right shows the green-band image.

B) Illumination and contrast is equalised – The illumination of retinal images is nonuniform because of the variation of the retina response. Hoover and Goldbaum overcame this challenge and equalised each pixel using the equation [27].

\[ I_{eq}(r, c) = I(r, c) + m \]

![Figure 4.4.9](image2)

Figure 4.4.9 – The left image shows the image after the illumination equalisation and the right image shows the image after the adaptive histogram equalisation.

C) Adaptive Histogram Equalisation (AHE) – “AHE is applied to an illumination equalized ‘I_{eq}’, inverted green-band image (…)”[28], “where each pixel p is adapted using this equation:”

\[ I_{AHE}(p) = \left( \sum_{p' \in R(p)} \frac{s(I(p); I(p'))}{h^2} \right)^2 \cdot M \]
Where $M = 255$, $R(p)$ means the pixel $p$’s surrounding area (a square window with length $h$), $s(d) = 1$ if $d>0$, otherwise $s(d) = 0$. The value of $h$ and $r$ is 81 and 8, respectively.

**D)** The retinal vascularity is segmented and compared to a proposed filter which represents the expected vessel direction in optic disks – A standard edge fitting algorithm was used [29], “where the similarity between a predefined 2-D Gaussian template and the fundus image is maximized. Twelve “15 x 15” filters (template) were generated to model the retinal vasculature along all different orientations (0° to 165°) with an angular resolution of 15°, then applied to each pixel where only the maximum of their responses is kept. In order to generate a binary vessel/nonvessel image”, “the maximum responses are thresholded using the global threshold selection algorithm proposed by Otsu (...)” [30]. Instead of applying the 12 templates to an averaged green-band image as suggested by (...)” [29], “applying them to the adaptively histogram equalized image significantly improves the segmentation algorithm and increases the sensitivity and specificity of the detected vessels (...)” [31]. “A vessel direction map (VDM) can be obtained from the segmentation algorithm by recording the direction of the template that achieved the maximum response at each pixel. Then, for all the pixels labelled as nonvessel, the corresponding values in the VDM can be assigned to “ -1 ” or not-a-number (NAN) in order to exclude them from further processing”.

**E)** Vessels’ Direction Matched Filter – A matched filter is a special filter that describes the expected look of a desired signal, for purposes of comparative modelling. The optic disk is discovered with a simple vessels’ direction matched filter. The filter matches the direction of the vessels at the optic disk surrounding area.

```
135  120  105  105  90   75   75   60   45
150  135  120  105  90   75   60   45   30
165  150  135  120  90   60   45   30   15
165  165  150  135  90   45   30   15   15
   0   0   0   0   90   0   0   0   0
   0   15  30   45  90  135  150  165  165
   15  30   45  60  90  120  135  150  165
  30  45   60  75  90  105  120  135  150
  45   60   75   75  90  105  105  120  135
```

Figure 4.4.10 – Proposed “vessels’ direction at the OD vicinity (...)” matched filter.

“The 9 x 9 template is resized using bilinear interpolation to sizes 241 x 81, 361 x 121, 481 x 161, and 601 x 201 to match structure of the vessels at different scales. These sizes are specially tuned for the STARE and DRIVE, but they can be easily adjusted to other datasets. The difference between all four templates (in the single given direction) and a Vdm is calculated, and the pixel having the least accumulated difference is selected as the optic disk center (...”). Matched filters are applied to candidate pixels picked from the fundus image and therefore reduce the computational burden.

The binary image is thinned to reduce the amount of pixels labelled as vessels into the vessels’ centreline. “All remaining vessel-labeled pixels that are not within a 41 x 41
square centered on each of the highest 4 % intensity pixels in the illumination equalized image are relabeled as nonvessel pixels (...). The main goal of this step is to reduce the number of OD candidates. Altering the size of the square or the amount of highest intensity pixels has no significant effect. The remaining vessel-labeled pixels are potential OD center, thus selected as candidates for applying the four sizes of the matched filter”.

Figure 4.4.11 – To the left, the binary vessel/non-vessel image and to the right, thinned version of the preceding binary image.

Figure 4.4.12 – To the left, final OD-center candidates and to the right the OD detected successfully (marked with white cross).

Result

This proposed method had a success rate of 98.77 %. I.e. of the 81 images contrained in the STARE dataset, 80 OD was detected correctly. In the DRIVE dataset, all 40 images, the OD was detected successfully. The average distance between the estimated OD centers was 17 pixels.
4.4.4 Example 5 – Fuzzy Convergence [27]

Method

This method is based on an algorithm called fuzzy convergence. “This algorithm identifies the optic nerve as the focal point of the blood vessel network”. In the absence of a strong convergence, the optic nerve is detected as the brightest region in the image after an illumination equalisation.

![Diagram of method for optic nerve detection]

Figure 4.4.13 – Outline of this method for optic nerve detection

a) Fuzzy convergence
   In this method, each vessel is modelled with a line segment (of finite length). The proposed method can be studied in appendix 6.

b) Illumination equalisation
   “A retinal image is captured by viewing the inner rear surface of the eyeball through the pupil. The lens of the camera works in conjunction with the lens of the eyeball to form the image. Since the position of the eye relative to the camera varies from image to image, the exact properties of the vignetting also vary from image to image”.
   “Vignetting is the result of an improper focusing of light through an optical system. The result is that the brightness of the image generally decreases radially outward from near the center of the image. The uneven illumination hinders absolute interpretation of the intensities in the image. In a healthy retina, the optic nerve is usually the brightest feature. Applying a simple high threshold to an image of a healthy retina should yield pixels inside the optic nerve. However, a retinal image is often captured so that the fovea appears mid-image, with the nerve to one side. Because of the vignetting, the nerve may appear darker than areas central to the image (...).”

To undo the vignetting illumination equalization is applied to the image. “Each pixel $I(r,c)$ in the image is adjusted as follows”:

$$I_{eq}(r,c) = I(r,c) + m - A(r,c)$$
“where \( m \) is the desired intensity (128 in an 8-bit greyscale image) and \( A(r,c) \) is the local average intensity. The local average intensity is computed independently for each pixel as the average intensity of the pixels within an \( N \times N \) window. The window size \( N \) is variable, so that averages near the border of the image use the same number of pixels (between 30 and 50 in our experiments) as averages in the center of the image. The local average intensities are also smoothed using the same windowing, to prevent blurring the image features. This process can be implemented using a sliding window algorithm (…), so that the computations are fast”.

c) Hypothesis generation

The fuzzy convergence and illumination equalization algorithms indicate the presence of the optic nerve as the brightest pixels. “In order to generate a hypothesis of nerve location, either image is thresholded to identify the brightest 1500 pixels (or darkest 1500 pixels(…))”.

Figure 4.4.14 – Sparse scale (two out of six) and Denser scale (five out of six)

A total of 1500 was chosen based upon the observation. The observation is the average number of pixels that stand out in an optic nerve in our images (our imaging resolution is 605 x 700 pixels on a 35° field-of-view). The pixels are then grouped into regions using standard 8-connected component analysis. “Any regions within five pixels of each other are grouped using standard morphological operators. This last step groups areas that may be separated by blood vessels, as commonly occurs in the nerve”.

According to the size, the regions are sorted and separated into two classes. This process is made by using Fisher’s linear discriminant. “This statistical test finds the best separation of the regions into those that are “large” and those that are “small” or insignificant. It works as follows. The regions are sorted by size, and repeatedly partitioned into two sets, \( A \) and \( B \). The initial partition has only the largest element (region size) in set \( A \), and all other elements in set \( B \). Each partitioning moves the largest remaining element from set \( B \) into set \( A \), until there is only one element left in set \( B \). Thus, for \( n \) total elements (regions) there will be \( n – 1 \) total partitions. For each partition \( p \), the discriminant statistic \( F_p \) is computed as”

\[
F_p = \frac{(\bar{\mu}_A - \bar{\mu}_B)^2}{\sigma_A^2 + \sigma_B^2}
\]

“where \( \mu \) and \( \sigma \) are the mean and standard deviation of each set. The largest value of \( F_p \) indicates the best partition. If only one region passes this test (resides in set \( A \) for
the largest value of $F_P$), its centroid is deemed a hypothesis of the optic nerve location. If more than one region passes this test, then the result is deemed inconclusive, and no hypothesis is generated”.

Result and discussion

A total of 81 images were tested, the nerve is visible in all 81 images (in 14 images, the nerve was located on the border so only a segment of the nerve was visible and in five images the nerve was obscured by hemorrhaging) and 50 of those with pathological signs. “All these images were acquired using a TopCon TRV-50 fundus camera at 35° field-of-view, and subsequently digitized at 605 x 700 pixels in resolution, 24 bits per pixel (standard RGB)”. In this method, the green band of the image was used (the red band inclines to be saturated, and the blue band inclines to be empty). “All our images and results may be viewed at (...)”[25].

Four variations were tested.

1) Equalized brightness—Illumination equalisation followed by hypothesis generation.

2) Fuzzy convergence, single scale—Fuzzy convergence to a single blood vessel segmentation, followed by hypothesis (each at the same scale).

3) Fuzzy convergence, multiple scales—“This variation applies fuzzy convergence to six segmentations of the blood vessels, each at a different scale (...). The scales correspond to the first six of the ten parameter values given in [10]. The distances between the hypotheses generated for each scale are measured relative to the radius of the average optic nerve (60 pixels in our images). If more than half of the hypotheses (three or more out of six) are all within this distance, the centroid of these locations is hypothesized as the optic nerve location. If there is no such consensus, then the result for the image is deemed inconclusive, and no hypothesis is generated”.

4) Fuzzy convergence and equalized brightness – This is a combination of variation 1 and 3. If the result from variation 3 is inconclusive, then the result from variation 1 is used. “The center point of the nerve in each image was recorded manually, for ground truth. A nerve is considered successfully detected if the hypothesis generated by the automated method is within the optic nerve, measured as within 60 pixels of the ground truth location”. If the method does not produce a hypothesis or if the hypothesised location is wrong then the nerve detection is unsuccessful.

The figure below shows the success rate of the variations on all 81 images. The “(...) brightness of the nerve is more easily confused in diseased retinas than in healthy retinas”. 

STH KTH, Flemingsberg
The multiple scales seem to work better than the single scale when it comes to finding true negative. The last variation shows a 100 % performance on the healthy retina test cases. The unsuccessful images were due to a convergence of blood vessels around a bright lesion. An additional operator that searched for circularity and bright regions may improve the result.

### 4.4.5 Example 6 – Image processing, detect optic disk [32]

**Outlines**

An approximate location of the macula [34] is determined because the diameter delivers a calibration of the measurements [33].

**Approach based on the watershed transformation**

1. *Colour Space*: The contours of the optic disk appear to be most continuous and most contrasted against the background in the red channel $f_R$ of the RGB colour space. In order to localise the optic disk, it is more reliable to work on the luminance channel $f_L$ of the HLS colour space. Because the red channel has a very small dynamic range and the optic disk belongs to the brightest parts of the colour image.

2. *Localising the Optic disk*: In this proposed method, a local grey level variation is used to find the locus of the optic disk. “*As the optic disc a bright pattern, and as the vessels appear dark, the grey level variation in the papillary region is higher than in any other part of the image*”. Assuming that there are no exudates on a dark background. See more in appendix 7.

3. *Finding the Countours of the Optic Disc Using the Watershed Transformation*: In the first step, the image $f_R$ in order to eliminate large grey level variations within the papillary region. First, the vessels are being filled, applied a simple closing with a hexagonal structuring element $s_1B$ bigger that the maximal width of vessels.

$$p_A = \emptyset^{(s_1B)}(f_R)$$

(See more in appendix 7).

---

**Figure 4.4.15 – Result in form of a table.**

<table>
<thead>
<tr>
<th></th>
<th>equalized brightness</th>
<th>fuzzy convergence, single scale</th>
<th>fuzzy convergence, multiple scales</th>
<th>fuzzy convergence and equalized brightness</th>
</tr>
</thead>
<tbody>
<tr>
<td>healthy retinas</td>
<td>77%</td>
<td>74%</td>
<td>87%</td>
<td>100%</td>
</tr>
<tr>
<td>diseased retinas</td>
<td>52%</td>
<td>74%</td>
<td>74%</td>
<td>82%</td>
</tr>
<tr>
<td>all retinas</td>
<td>62%</td>
<td>74%</td>
<td>79%</td>
<td>89%</td>
</tr>
</tbody>
</table>
Result

30 colour images were tested. They were 640 x 480 pixels. The optic disk was located in 29 of them. In 27 of 29 of the images, the exact contours were found. “However, in some of the images, there were small parts missing or small false positives. These shape irregularities in the segmentation result are due to the outgoing vessels or to low contrast”. The shape could be regularised by using standard morphological filter techniques. In two images, the contrast was too low or the red channel was too saturated.

4.4.6 Example 7 – Geometrical model of vessel structure [37]

Introduction

This method is based on the detection of the main retinal vessels. All vessels reach the retina through the optic disk and their path follows a parabolic course in all images. To describe this direction, a geometrical parametric model is proposed. Two of the model parameters are the coordinates of the OD center. “Using as experimental data samples of vessel centerline points and corresponding vessel directions, provided by any vessel identification procedure, model parameters were identified by means of a simulated annealing optimization technique”. Estimated values provide the coordinate of the center of OD.

Method

This method implicitly embeds the information on the OD position as the point of convergence of all vessels. “The result is not based on(...)” [27] “but on the fitting of a model with respect to the entire vascular structure”.

A. A geometrical model of retinal vessels direction – “Defining a directional model for retinal vessels requires the definition on the whole image of a function”

\[ \theta^{\text{mod}}(x, y; \mathbf{p}) = \pi \leq \theta^{\text{mod}} \leq \pi \]

“which represents the preferential direction in any retinal image of a vessel present at point \((x, y)\). Vector \(\mathbf{p}\) is the set of parameters defining the model and its positioning and, thus, it will include the OD coordinates. By visual inspection of retinal fundus images (...), it appears that a common vascular pattern is present among images: the main vessels originate from the OD and follow a specific course that can be geometrically modeled as two parabolas, with a common vertex inside the OD. The definition of the directional model can, therefore, be based on this assumption. If we assume a Cartesian coordinate system, these parabolas can be described as the geometrical locus”

\[ \Gamma = \{(x, y) : ay^2 = x\} \]

”where \(a\) is the parameter governing the aperture of the parabolas (for sake of simplicity, let us assume for the time being that the origin of the coordinate system is the vertex of the parabolas). (...). For a generic point \((x, y)\) belonging to locus \(\Gamma\), i.e., on the parabola. The directional model is expressed by the implicit equation”.
\[
\tan(\theta^{\text{mod}}(x, y; p)) = sgn(x)sgn(y) \frac{1}{2a} \sqrt{\frac{|x|}{a}}
\]

where \(x \neq 0\).

Vector \(p\) contains parameter \(a\). The function \(sgn()\) returns the sign of its argument. The equation above states that on the parabolas the preferential vessel direction is tangent to the parabolas themselves. "In order to completely define the model, it is necessary to express the preferential direction also outside of the parabolic geometrical locus \(\Gamma\) implicitly divides every quadrant in two areas: the internal area (with respect to the convexity of the parabola) and the external area. Anatomical knowledge indicates that vessels bifurcate when moving away from the OD, and branch vessels tend to diverge from the main vessel. In particular, vessels inside the parabolas quickly bend toward the macula in the temporal region (…), whereas in the nasal region this inward deflection happens at a much slower rate (…). The tangent equation above was, thus, extended to accommodate points outside \(\Gamma\) by adding a correction term \(d\)."

\[
d(x, y; p) = \frac{y - sgn(y)\sqrt{\frac{|x|}{a}}}{c(x)}
\]

\[
c(x; p) = \frac{c_1}{1 + e^{-x}} + \frac{c_2}{1 + e^x}
\]

\(c_1 > 0; c_2 < 0\).

"The numerator of (...)" equation \(d(x, y; p)\) "is zero for a point belonging to \(\Gamma\), whereas for a point outside \(\Gamma\) its absolute value increases in a way proportional to the vertical distance between the point and \(\Gamma\). This increment in tangent magnitude is modulated by (...)" equation \(c(x; p)\), "which expresses the rate of divergence of the direction at any given \(x\) coordinate". "For increasing \(x\), this rate tends toward the value of parameter \(c_1\) \((c_2)\) for positive (negative) values of \(x\). Values of \(c_1\) and \(c_2\) represent, therefore, the limit rates of convergence toward \(\pi / 2\) (vertical direction) of the vessel directions for positive and negative \(x\) values. These two rates are in principle different, to take care of the different degree of curvature of vessels in the nasal and temporal side of retina: the lower the absolute value of this constant, the higher the curvature of vessels as they move away from OD. Given a generic origin for the Cartesian coordinate system in use (e.g., upper-left corner in the image), in order for the parabolas to be centered at the coordinates of the OD center \((x_{OD}, y_{OD})\), (…), a translation transformation had to be applied to the model".

\[
x^{*} = x - x_{OD}
\]
\[
y^{*} = y - y_{OD}
\]

"The complete model for vessel direction \(\theta^{\text{mod}}\) at any point \((x, y)\) in the image is given by the following equation":

\[
\theta^{\text{mod}}(x, y; p) = \arctan\left\{\frac{sgn(x - x_{OD})sgn(y - y_{OD})}{2a}\sqrt{\frac{|x - x_{OD}|}{a}} + \frac{(y - y_{OD})sgn(y - y_{OD})}{\frac{c_1}{1 + e^{-(x - x_{OD})}} + \frac{c_2}{1 + e^{(x - x_{OD})}}}\right\}
\]
Directions $\theta^{\text{mod}}$ are shown for some points of the image and optimal values for model parameters are used for this simulation (see below).

B. Model Parameter – Identification of the optimal value for $p$ and, thus, for $(x_{OD}, y_{OD})$, can be identified for any image, given set of data, by using suitable model parameter identification techniques. “The data are the vessel directions $\theta_i$ measured at points $(x_i, y_i), i = 1, ..., N$, belonging to the vascular structure”. “Our choice for the identification of model parameters has been the minimization of the weighted residual sum of squares (RSS)”

$$RSS = \sum_i w_i [(\theta(x_i, y_i) - \theta^{\text{mod}}(x_i, y_i; p))^2].$$

“Minimization is performed with respect to model parameters and operator ‘.’ indicates a modulus $\pi$ difference between directions. Quantities $w_i$ are weights, used to modulate the importance of each term in the summation. Different options have been investigated to describe these weights and the best results were obtained with $w_i$ proportional to vessel caliber $c_i$. Optimized values of parameters $(x_{OD}, y_{OD})$ represent the best positioning of the OD according to the model fit on the available data $(x_i, y_i, c_i, \theta_i)$”.

“Minimization of RSS with classical gradient-based techniques is rather critical, since this function exhibits many local minima. Fig. 4 represents, e.g., a plot of RSS as a function of parameters $x_{OD}$ and $y_{OD}$ only. The absolute minimum is correctly found when $(x_{OD}, y_{OD})$ are inside the OD, but a gradient-based algorithm would be easily trapped in one of the many local minima. To overcome this problem, a simulated annealing (SA) optimization algorithm has been adopted. SA is a global stochastic optimization algorithm that theoretically guarantees the convergence toward global minimum (…)”[47]. “The working parameters of this procedure (e.g., number of data points, initial model parameters value, initial temperature, termination criterion, etc.) have been empirically tuned using are presentative subset of 20 images. The resulting set of values was then used for model parameter estimation in all 81 images of the entire data set. In order to overcome the stochastic nature of SA algorithm, several optimization runs were performed, starting the procedure from different points in the parameter space, and the final RSS values were compared to select the smallest one. A number of six runs for each image proved to be sufficient in our test set”.

Results

81 fundus images of the STARE project data set were used. 31 images were healthy and 50 were pathological (numerous types and severity). “In order to assess the robustness of the proposed procedure to detect OD position with respect to different vessel detection algorithms, we have used the measured vessel directions provided by two sets of vessel data. The first one (Track-1) was obtained by applying a binary segmentation procedure developed by Hoover (…)”[48] “and used to provide input data also to their own system for OD detection (…)”[32]. “For each image, 10 different sets of vessel structure data were provided, to produce vessel segmentations at different scales, from 0 to 9. At variance with
“where 6 segmentations for each image were used to detect vessel convergence and, thus, OD detection, we used only one segmentation (scale value of 4). The second set of vessel data (Track-2) was obtained by applying our own procedure, based on a sparse tracking algorithm (...)”[49]. “As proposed also in (...)”[32], “the OD position was considered correctly detected if the estimated coordinates were inside the contour of the OD, i.e., within 60 pixels of its center, as manually identified for ground truth”. This method found 79 out of 81 images.

<table>
<thead>
<tr>
<th></th>
<th>Track-1</th>
<th>Track-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>OD identified</td>
<td>79</td>
<td>79</td>
</tr>
<tr>
<td>OD not identified</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>81</td>
<td>81</td>
</tr>
</tbody>
</table>

Figure 4.4.16 – Result

Conclusion

The proposed method, based on a model of the vascular structure, “(...) are dependent on the availability of a good portion of this structure in the image, whereas are independent of the actual visibility (or even presence) of the OD”. The vascular structure that is spread all over the image, is less affected by the presence of obscuring or confounding pathological areas. The availability of a vessel extraction process is a necessary assumption for this technique and the performances directly affect the correct positioning of the optic disk. The remarkably good results were obtained from this described method.

4.4.8 Example 8 – Geometric active contours [38]

Introduction

This example proposes two efficient approaches for detection and boundary estimation of optic disk. The approach for optic disk detection uses the vessel branch with the most vessels. The algorithm for boundary detection, involves two steps. First, “(...)the color mathematical morphology in Lab space is used to have homogeneous optic disk region, then the boundary of the optic disk is estimated by using geometric active contour with new variational formulation”.

Method

1. **Localisation of Optic disk Based on the Branch with the Most Vessels** – OD is normally located by finding the largest cluster of brightest pixels, but it fails when it comes to find OD in images where the area of bright lesions is large or the OD is obscured by blood vessels. In this example the OD is found when the branch with the most vessels located. "The proposed method turns the vessel probability map into a network of vessels and branches. In this method network information is stored about the connections between vessels and branches. By means of this network, the branch
with most vessels connected to it can be selected. The selected branch is used to determine the optic disc”. The binary blood vessel skeleton map is obtained from [50].

“For each skeleton pixel the amount of neighbor skeleton pixels is determined. If the amount of neighbors is smaller than three, then the pixel is added to the vessel-image. Otherwise the pixel is added to the branch-image. A vessel is a collection of points, starting at a point with only one neighbor and ending at a point with only one neighbor. The vessels are detected as follows: If a begin point of a vessel is detected within the vessel-image, then the vessel is traced towards its endpoint. The begin point and the endpoint are marked to avoid tracing a vessel twice. The branches are detected by applying eight-connected component analysis on the branch-image. The constructed vessel-branch network can be used to find the optic disc in many ways. A very simple algorithm is the selection of the branch with most vessels. For each branch of the network the amount of vessels connected to it is stored. The optic disc contains the optic nerve from which a few main vessels split up into many smaller vessels which spread around the retina. Vessel segments in this area of the retina are often small and are therefore often combined into one large branch of the network with many vessel objects connected to it. An increasing amount of vessel connections of a branch also increases the probability of the branch being located in the optic disc area”.

The algorithm is performed as follows:
1. “Select the branch with the most vessel connections. If there are several branches with the highest number of vessel connections, then the branch with the most branch pixels is selected”.
2. “Take the bounding box of the branch with the most vessel connections”.
3. “Select the center of the bounding box as the OD-center”.

Figure 4.4.16 – (1) Localisation of the OD, (2) Binary vessel map, (3) Overlay of the vessel branch network and the original image, (4) The bounding box of the best branch and the determination of the OD center, (5) Detected OD center.

2. Boundary Detection of Optic Disk by Geometric Active Contour Model – “The accurate detection of the optic disk boundary can be used to assess the progress of eye disease and the treatment results. Some parts of the disk boundary are not well defined and some parts are partly obscured by the blood vessels in retinal images, which make the detection of disk shape complicated. A geometric active contour model is proposed to detect the disk boundary in retinal images. Firstly, the original color retinal image is preprocessed using color mathematical morphology in Lab space. This helps to remove blood vessels more cleanly and provides a more homogeneous optic disc region for the geometric active contour to lock onto. We performed dilation
first to remove the blood vessels in optic disk region and then an erosion to restore the boundaries to their former position. The morphology in Lab space is performed using the method described in (...)”[51]. “For each arbitrary point x in the color space, the definitions for dilation ($I_d$) and erosion ($I_e$) by structuring element K is defined as”:

$$I_d(x) = \{ I(y) : I(y) = \max[I(z)], \forall Z \in K_x \}$$

$$I_e(x) = \{ I(y) : I(y) = \min[I(z)], \forall Z \in K_x \}$$

A symmetrical disc structuring element of size 13 was used, since the blood vessels were determined to be not wider than 11 pixels. “The optic disk boundary is determined by fitting a geometric active contour model with variational formulation. The initial contour for a snake must be close to the desired boundary otherwise it can converge to the wrong resting place”. A method described for localizing the optic disc was used. This method enabled to automatically position an initial snake. “In general, a snake is a set of points initially placed near the contour of interest, which are gradually brought closer to the exact shape of the desired region in the image. This is carried out through iterative minimization of an energy function comprising an internal and an external term”:

$$\varepsilon(\phi) = \mu P(\phi) + \varepsilon_{g,\lambda,v}(\phi)$$

“where $P(\phi) = \int_\Omega \frac{1}{2} (|\nabla \phi| - 1)^2 \, dx \, dy$ a metric to characterize how close a function $\phi$ is to a signed distance function in $\Omega \subset \mathbb{R}^2$, $\mu > 0$ is a parameter controlling the effect of penalizing the deviation of $\phi$ from assigned distance function. $\varepsilon_{g,\lambda,v}(\phi)$ is the external energy for a function $\phi(x,y)(...)$” and it is defined as:

$$\varepsilon_{g,\lambda,v}(\phi) = \lambda L_g(\phi) + v A_g(\phi)$$

“where $\lambda > 0$ and $v$ are constants, $g$ is an edge indicator function defined for an image I as”

$$g = \frac{1}{1 + |g_{G_\sigma}|^2}$$

where $G_\sigma$ is the Gaussian kernel with standard deviation $\sigma$. The terms $L_g(\phi)$ and $A_g(\phi)$ are defined as”

$$L_g(\phi) = \int_\Omega g \delta(\phi) \, |\nabla \phi| \, dx \, dy \quad \text{and}$$

$$A_g(\phi) = \int_\Omega g H(\phi) \, |\nabla \phi| \, dx \, dy$$

“respectively, where $\delta$ is the univariate Dirac function, and $H$ is the Heaviside function. The external energy $\varepsilon_{g,\lambda,v}$ drives the zero level set toward the object boundaries, while the internal energy $\mu P(\phi)$ penalizes the deviation of $\phi$ from a signed distance function during its evolution. By calculus of variations, the Gateaux derive (first variation) of the functional $\varepsilon(\phi)$ can be written as”

$$\frac{\partial\varepsilon}{\partial\phi} = \mu \, (\nabla \phi - \text{div} \left( \frac{\nabla \phi}{|\nabla \phi|} \right)) + \lambda \delta(\phi) \text{div} \left( g \, \frac{\nabla \phi}{|\nabla \phi|} \right) + v g \delta(\phi)$$

“This gradient flow is the evolution equation of the level set function in the proposed method. The second and the third term in the right hand side of (...) correspond to the gradient flows of the energy functional $\lambda L_g(\phi)$ and $v A_g(\phi)$, respectively, and are responsible of driving the zero level curve towards the object boundaries”.
Results

“The proposed algorithms are tested and evaluated on four publicly available databases of color retinal images(...)”: STARE [32], DRIVE [52], DIARETDB0 [53], and DIARETDB1 [54] databases. “The DIARETDB0 database consists of 130 color retinal images of size 1500 × 1152. The DIARETDB1 database consists of 89 color retinal images of size 1500 × 1152. The DRIVE database contains 40 color images divided into 20 training and 20 test images. The downloaded images were of size 565×584. The STARE database consists of 81 slides which were digitized to 700 × 605 pixels, 8 bits per color channel”.

1. Localization of Optic Disk

<table>
<thead>
<tr>
<th>Database</th>
<th>Number of Images</th>
<th>Maximum local variation method</th>
<th>Proposed Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>STARE</td>
<td>81</td>
<td>79</td>
<td>99</td>
</tr>
<tr>
<td>DRIVE</td>
<td>40</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>DIARETDB0</td>
<td>130</td>
<td>89</td>
<td>98</td>
</tr>
<tr>
<td>DIARETDB1</td>
<td>89</td>
<td>84</td>
<td>100</td>
</tr>
</tbody>
</table>

The input images are claimed to be in a low-contrast condition.

2. Boundary Detection of Optic Disk

“In order to evaluate the performance of our algorithm for detecting optic disk boundary, we compare results of the proposed algorithm with the state-of-the-art results obtained from GVF snake method (...)” [55], “2D Circular Hough Transform method and hand-labeled ground-truth segmentations. In GVF-snake, the images are preprocessed by morphological operation; and the parameters in the energy functions were carefully set to make a balance between the smoothness and the accuracy on the resulted boundary. In 2D Circular Hough Transform method, the dimensions of the normal circular Hough Transform histogram are reduced from 3 to 2 dimensions by assuming that the approximate OD radius is known. Only the first few circles are evaluated by using the maximum point from Hough space. The disk boundary manually marked by the experienced ophthalmologist is set to be the ground-truth. Then we use a simple and effective overlap measure to evaluate the accuracy of the detected boundary”.

\[ M = \frac{n(R \cap T)}{n(R)} \] where \( R \) and \( T \) correspond to the ground-truth and the detected optic disk region respectively and \( n(.) \) is the number of pixels in a region. Specifically, we classify retinal images into three categories, normal retinal images, abnormal retinal images with ill-defined optic disk, and retinal images with fuzzy elliptic optic disk. GVF snake give the successful results; Hough transform gives the failed result. The measured accuracies for the GVF snake, Hough transform and the proposed method are 99.3 %, 78 % and 99.5 % respectively. The example given in the second
column is an optic disk with ill-defined boundary and noises from the surrounding tissue. The proposed method correctly located the disk boundary, while the GVF snake and Hough transform methods failed. The accuracies for the GVF snake and Hough transform are 72% and 74% respectively. For the same image the proposed method has an accuracy of 99.1%. One more example of fuzzy elliptic optic disk is illustrated in the third column. The accuracies for the GVF snake and Hough transform are 81% and 82% respectively. For the same image the proposed method has an accuracy of 98.2%.

<table>
<thead>
<tr>
<th>DATABASE</th>
<th>NUMBER OF IMAGES</th>
<th>AVERAGE ACCURACY IN %</th>
<th>GVF SNAKE</th>
<th>PROPOSED METHOD</th>
</tr>
</thead>
<tbody>
<tr>
<td>STARE</td>
<td>81</td>
<td>83.8</td>
<td>89.8</td>
<td>94.2</td>
</tr>
<tr>
<td>DRIVE</td>
<td>40</td>
<td>93.5</td>
<td>98.6</td>
<td>99.3</td>
</tr>
<tr>
<td>DIARETDB0</td>
<td>130</td>
<td>85.4</td>
<td>95.1</td>
<td>96.8</td>
</tr>
<tr>
<td>DIARETDB1</td>
<td>89</td>
<td>89.2</td>
<td>95.4</td>
<td>97.5</td>
</tr>
</tbody>
</table>

4.4.9 Example 9 – Hough Transform [39]

Introductions

This is a method which can be used to find the OD in retinal images. The proposed method is based on the properties of the OD, including edge detection using the Sobel or the Canny method and detection of circles using the Hough transform. “The Hough transform assists in the detection of the center and radius of a circle that approximates the margin of the OD. Based on the feature that the OD is one of the bright areas in a fundus image, potential circles detected by the Hough transform are analyzed using intensity. Forty images of the retina from the DRIVE database were used to evaluate the performance of the proposed method”.

Methods

A. Dataset of Retinal Images and Preprocessing – 40 images form the DRIVE database. “After normalizing each component (dividing by 255), the result was converted to the luminance component Y, computed as $Y = 0.299R + 0.587G + 0.114B$, where R, G, and B are the red, green, and blue components, respectively, of the color image. The effective region of the image was thresholded using the normalized threshold of 0.1. The artifacts present in the DRIVE images at the edges were removed by applying morphological erosion (...)” [56] “with a disc-shaped structuring element of diameter 10 pixels. In order to avoid edge artifacts, each image was extended beyond the limits of its effective region (…)” [57][58]. “First, a four-pixel neighborhood was used to identify the pixels at the outer edge of the effective region. For each of the pixels identified, the mean grey level was computed over all pixels in a 21 × 21 neighborhood that were also within the effective region, and assigned to the corresponding pixel location. The effective region was merged with the outer edge...”
pixels, forming an extended effective region. The procedure was repeated 50 times, extending the image by a ribbon of width 50 pixels. After preprocessing, a 5 x 5 median filter was applied to the Y channel, to remove outliers in the image. Then, the maximum intensity in the image was calculated to serve as a reference intensity for the selection of circles”.

B. Detection of Edges – “The Sobel operators (…)” [56] “for the horizontal and vertical gradients are shown in (…)” figure 4.4.17. “The horizontal and vertical components of the gradient, \( G_x(x, y) \) and \( G_y(x, y) \), respectively, were obtained by convolving the preprocessed image with the operators (…)” shown in figure 4.4.17. “The combined gradient magnitude was obtained as \( G(x, y) = \sqrt{G_x^2(x, y) + G_y^2(x, y)} \). A threshold was applied to the gradient magnitude image to obtain a binary edge map”. Canny [59] “(…) proposed an approach for edge detection based upon three criteria for good edge detection, multidirectional derivatives, multiscale analysis, and optimization procedures. The MATLAB (…) of the Canny operator was used to obtain a binary edge map for comparative analysis”.

![The Sobel operators](image)

C. The Hough Transform for the Detection of Circles - Hough transform is a method to detect straight lines in images. It has been extended to identify circles and other parameterized geometrical shapes. “The points lying on the circle \((x - a)^2 + (y - b)^2 = c^2\) are represented by a single point in the three-dimensional (3D) parameter space \((a, b, c)\) with accumulators of the form \(A(a, b, c)\), which is also known as the Hough space. Here, \((a, b)\) is the center, and \(c\) is the radius of the circle. The procedure to detect circles involves the following steps”:

1) Obtain a binary edge map of the image.
2) Set values for parameters \(a\) and \(b\).
3) Solve for the value of \(c\) that satisfies equation above.
4) Update the accumulator that corresponds to \((a, b, c)\).
5) Update values for parameters \(a\) and \(b\) within the range of interest and go back to Step 3.

D. Procedure for the Detection of the OD – “The Hough accumulator is a 3D matrix, each cell of which is incremented for each nonzero pixel of the edge map that meets the stated condition. For example, the value for the cell \((a, b, c)\) in the Hough accumulator is equal to the number of edge map pixels of a potential circle in the image with the center at \((a, b)\) and radius \(c\). In the case of the images in the DRIVE
database (…), the size of each image is $584 \times 565$ pixels. The spatial resolution of the images in the DRIVE database is about $20 \mu m$ per pixel. The physical diameter of the optic disc is about $1.5 \ mm$ on average (…). Assuming the range of the radius of a circular approximation to the OD to be $600 - 1000 \ \mu m$, the range for the radius $c$ was determined to be $31 - 50$ pixels. Hence, the size of the Hough accumulator was set to be $584 \times 565 \times 20$. The potential circles indicated by the Hough accumulator were ranked, and the top 30 were selected for further analysis. Because we know that the OD is one of the bright areas in the image, a threshold equal to $0.9$ times the reference intensity (…) was used to check the maximum intensity within a circular area with half of the radius of the potential circle. If the test failed, the circle was rejected, and the next circle was tested”.

Results

The edge images that were obtained by the Sobel operators were binarised using a threshold of $0.02$ (the threshold was chosen by analysing the look of the OD in an edge image). The success rate was $92.5\%$ (37 out of 40). The Canny operator (threshold, automatically chosen by Matlab) missed seven ODs. Two results were acceptable. The operator had a success rate of $80\%$. The method was also tested with STARE dataset. 82 images were used and 51 ODs was located (33 by Sobel (auto threshold) and 18 by Canny (thresholded at 0.17)).

4.5 Age-Related Macular Degeneration, AMD

4.5.1 Example 10 – Drusen detection [41]

Texture-based drusen detection

Texture of drusens can be marked as local energy. The local energy is the sum of squared responses of orthogonal pairs of filters. Gabor functions and log-Gabor functions are popular choices for such filters. A local energy model using multi-channel log-Gabor filters is presented in this example.

Gabor filters can easily be adjusted to different scale and orientations. Its bandwidth decides the scale of the features. In a Gabor filter, the maximum bandwidth is 1 octave, which limits the feature size. Larger bandwidths are allowed on the log-Gabor filters, which makes it more reliable and informative. On a linear scale, the transfer function of the filter is:

$$
\Phi(r_o, \theta_o) = \exp\left\{-\frac{(\log \frac{r}{r_o})^2}{2(\log \frac{r}{r_o})^2}\right\} \exp\left\{-\frac{(\theta - \theta_o)^2}{2\sigma_\theta^2}\right\}
$$

It is based on the central radial frequency $r_o$, the orientation $\theta_o$, $\sigma_\theta$ and $\sigma_r$, represent the angular and radial bandwidths of the filter.

“The oriented local energy $\sum_{r_o} r_o \phi$ at every point $(x,y)$ in the image defines an energy map. This is obtained as:”
Where both terms are responses of the odd and even symmetric log-Gabor filters. “Let $Z(r_0, \Theta_0)$ be the filtered output. The responses of even and odd symmetric log-Gabor filters can be found as:

$$O_{\theta_0}^{r_o,\text{even}} = \text{Re}(Z(r_0, \Theta_0)); \quad O_{\theta_0}^{r_o,\text{odd}} = \text{Im}(Z(r_0, \Theta_0))$$

Experiment result

Drusens are best represented in the green channel of the input image and it was therefore used in this experiment. Three different scales and varied orientation (0°, 30°, 60°, 90° and 150°) were used. The filter parameters were chosen to scope the frequency domain. A good localization of the segmented regions was obtained by having a large radial bandwidth in the frequency domain and a small bandwidth in the spatial domain. The parameters have been chosen to capture high frequency components, i.e. drusens. The high frequency content is also present in veins and should be suppressed. Veins appear as dark pixels in the green channel of the retinal image and therefore all the dark pixels are replaced with the local mean of their neighbourhood.

The algorithm was tested on different datasets. The results of the tests show that the energy is maximum at the locations of drusens. Therefore it is quite easy to segment the drusens using the local energy. The high local energy also reveals the optic disc and it should also be suppressed. One solution is to extract closed boundaries for the drusens from the energy map and each boundary can be counted as one drusen. It is not an easy task. There are two deficiencies with the local energy based drusen detection:

- Selection of tuning parameters
- Obtaining of a reliable drusen count

Model-based drusen detection

Togographic models have successfully been used for feature detection. By visualising a 2D image function as a surface in 3D space, a different perspective of drusen is presented. A way of doing so is by using a property like the curvature of the image surface. The curvature can be used to detect features where the surface bends sharply. “The curvature at a point is measured of the bend in the surface along a particular direction”.

“Let $y = f(x)$ be a 1D function. Let the tangent at a point $P : x$ on this function make an angle $\theta$ with the $x$–axis as shown in (...)” figure 4.5.1. “If $dl$ is the differential arc length at the point $P$, then the extrinsic curvature of the function $f(x)$ at this point is defined as”:

$$k(x) = \frac{d\theta}{dl} = \frac{\frac{d\theta}{dx}}{\sqrt{1 + \left(\frac{dy}{dx}\right)^2}}$$

“Since, $\theta$ is the angle made by the tangent with the $x$–axis, it can be computed as:”
\[ \theta = \tan^{-1}\left( \frac{dy}{dx} \right) \]

“Hence, the numerator term \( \frac{d\theta}{dx} \) can be computed as:”

\[ Y(x) = \frac{d\theta}{dx} = \frac{d}{dx} \left[ \tan^{-1}\left( \frac{dy}{dx} \right) \right] = \frac{\frac{d^2y}{dx^2}}{1 + (\frac{dy}{dx})^2} \]

“Substituting the above expression (...)” in equation k (x) above:

\[ k(x) = \frac{\frac{d^2y}{dx^2}}{\left(1 + (\frac{dy}{dx})^2\right)^\frac{3}{2}} \]

“which is the true curvature measure. As the point P moves on the curve \( y = f(x) \), the tangent angle \( \theta \) changes. This change over a given arc length \( dl \) is the true curvature measure \( k(x) \).”

“In equation the first \( k(x) \), the numerator term \( Y(x) \) represent the rate of change the tangent angle with respect to the projection of the arc length over the \( x \)-axis. Comparing the equation of \( Y(x) \) and the second \( k(x) \), the two expressions differ by the power denominator. Especially \( Y(x) \), which peaks sharply at the locations of medial points of ridge profiles, “where the first derivative of the profile function vanishes and the second derivate is a negative maximum. Thus, \( Y(x) \) is an alternative to the true curvature measure \( k(x) \) since it also provides information about the rate of change of the tangent angle as a point moves along a curve (...). In the case of 2D images, \( Y(x) \) corresponds to a derivate of the angle made by a surface tangent line with the image plane, in some direction. Accordingly, we distinguish it from the true curvature measure, by calling it as the Surface Tangent Derivative (STD). We will use the STD as an estimate of the curvature of image surfaces. Since drusen have hilly profile, the STD can be used to detect them by detecting hill-like features in the images which are characterised by high values of STD in all directions. Here, we have used a hill detection algorithm presented in (...)” [60]. "At every point in a given image, STD is computed in N different directions and maximum value of chosen among all N directions. A point is declared as a hill point if the value at that point is above a threshold value and maximum among its neighbours”.

Experiment details and discussion

The STD computation was done in four directions: -45°, 0°, 45° and 90°. STD based hill detection was used to extract hill points in the image. The detected hill points are often due to drusen or due to noise. If it is noise, the hill points can be removed and those points whose value lies above a threshold are retained. A low threshold value detects a lot of drusens but due to noise, some false negatives. High threshold value misses some drusens and hence, the threshold value has to be carefully chosen. Low thresholds are chosen when to detect soft drusens. The false negatives are later suppressed by using information about local context. In this experiment, the value was decided empirically and was kept fixed for all images.
In order to detect all drusens, a multiscale computation of STD is made. STD was computed at the scales of 3, 7, 11, 21 and 31. The result were summed (logical OR). The results were confirmed by a retina expert. It is possible that a large drusen could be characterised by more than one hill point and that could affect the count.

4.5.2 Example 11 – Histogram approach [42]

Introduction

Histogram can be used to represent colour distribution in images. It is a simple way of representing the characteristics of an image and an effective representation for identifying objects in images.

There are two major methods of generating histograms:

i) Fixed binning
   ii) Adaptive binning

With the fixed binning, the same numbers and characters of bins are applied to all images. The adaptive binning, adapts to the current distribution of colours in the images. The fixed binning can adopt a wide range of similarity metrics (e.g. Euclidean distance) while the adaptive binning are more constrained by the number of variation measures available to measure the similarity between histograms. “Each bin in fixed binning histograms can be regarded as a vector and this makes it possible to apply various machine learning and data mining algorithms to representation for (...)” e.g. clustering and classification. A fixed binning will be presented in this method.

Figure 4.5.2 – (1) Image with drusen, (2) Red channel image, (3) Green channel image, (4) Blue channel

STH KTH, Flemingsberg
Dynamic time warping

"Dynamic time warping (DTW) is a technique for measuring the similarity between two time series sequences." “Thus a histogram, of the form described above, can be interpreted as a time series. DTW uses a dynamic programming approach to align two time series and then generates a warping path that maps (align) the two sequences onto each other”. (See more in appendix 4)

Methodology

1) Image pre-processing – Any unnecessary information that are presented in the image, are being filtered out. E.g. the black area that surrounds the image. “Give I images, with rows r and columns C, R and C are fixed for each i∈I. The pixels information of each image is stored in a matrix X. Each element of X is referred as x_{r,c}(i)=γ, where 0≤r<R and 0≤c<C. γ is the colour value in hexadecimal format for pixel x_{r,c}(x∈X) of image i∈I”.

2) Histogram generation – the retinal images are translated into histogram representation, H. In this method, two suites of histograms are used. The first include histograms (per image) for each colour channel: Red, Green and Blue. The second one includes histograms describing the Hue, Saturation and Intensity components for each given image. All histogram are assumed to contain relevant information. As showed in figure above, the red and green channel give a better visual contrast of drusen as compared to the blue channel. Though the blue channel is also important.

“The length of each histogram is fixed to M bins, and represented as h_{i,b}(m)=β, where h∈H, b represents the RGB channel or HIS component, 0≤m<M, and X is the histogram value of m, normalized to the maximum value recorded by a m for a particular h_{i,b}. M is set to 256 (number of colour space cells) for each red, green and blue histogram, 360 (0° to 359°) for hue histogram and 101 (0 to 100) for both saturation and intensity histograms”.

3) Classification – “Using this histogram based approach a case base of pre-labelled (AMD positive or negative) histograms, P⊆H, where p=p1,p2,…, pa and a is a set of retinal images that have been hand classified by domain experts. The histograms for a new retinal image to be classified, N, where N∈H and N∈P are then plotted onto graphs to attain the curves of the histograms, as depicted in Figure 2, before DTW find pa∈P, a histogram that has the best warping path with N, of each colour channel or HIS component. Once identified, N will be classified into the same class as pa is”.
Experimental Setup

144 images were gathered of which 86 were AMD images. The images were separated into ten equally distributed datasets. Each dataset had approximately 9 AMD images and 6 control images. “Ten-fold Cross Validation was used to evaluate the performance of the proposed approach whereby the image set was divided into 10 subsets and ten evaluations runs conducted. For each run the case base was generated on a different nine tenths and the classification accuracy tested on the remaining tenth.” The goal of the experimental setup was:

- To analyse the effectiveness of using the colour channel histogram
- To investigate the effectiveness of the HIS histograms

“Three evaluation metrics were utilized to measure the classification performance: sensitivity, specificity and accuracy. Sensitivity, tells to measure the effectiveness of the classifier in identifying true positives (AMD images), and is formulated as sensitivity = TP/α, where TP (True-Positive) is the number of AMD images classified as AMD by the classifier, and α is the total number of AMD images in the test set. Specificity tries to measure the effectiveness of the classifier in distinguishing the normal control images by not falsely classifying the control image as AMD images. Specificity is defined as specificity = TN/δ, where TN is number of control images not misclassified as AMD images (True-Negative) and δ is the total number of control images in the test set. Accuracy will be used to measure the overall performance of the classifier in term of classifying retinal images correctly according to their class. Accuracy is defined as (TP+TN)/ε, where ε =α+δ”.

Result

The result looks promising according to the table below. The best specificity of 62 % was achieved with the green channel and a sensitivity of 83 % with the blue channel. Though the blue channel only had 69 % mean accuracy. The saturation component achieved the best specificity, sensitivity and accuracy, with 60 %, 82 % and 74 % of each.
Detection of the anatomic structure is the characterisation of the normal or disease state that exists in the retina.

Figure 4.5.4 – Table of result that shows a high specificity on the green channel and a high sensitivity on the blue channel.

The blue channel performed surprisingly well and demonstrates a high effectiveness, i.e. due to the high sensitivity of 83\%. “It is conjectured that the uneven distribution causes the DTW process to calculate an almost similar distance between points in unseen and knowledge base histograms, and consequently inaccurately classified the histogram. However, HSI did better in terms of classification accuracy, with a 74\% best accuracy recorded by the saturation component. This shows the ability of HSI in identifying patterns through the colours of the images”.

4.5.3 Example 12 – PNN classifier [35]

This method acquires a number of features such as area, mean, standard deviation etc. of the preprocessed image. These images are extracted to characterise the image content.

Then, after the extraction, a probabilistic neural network (PNN) classifier is used.

Pre-processing

Both normal and AMD images are being processed. Every image is 1280 x 1024 pixels. The green component is extracted from the images. Then histogram equalisation is used to enhance the contrast and quality. Finally an isotropic diffusion is applied to remove noise.

Locating anatomic structures and detecting lesions

Detection of the anatomic structure is the characterisation of the normal or disease state that exists in the retina.

- Discrete wavelet transform – an isotropic diffused image
  \[ [cA, cH, cV, cD] = \text{dwt2}(X, 'wname') \]  
  (1) 

  Calculates “…the approximation coefficients matrix \( cA \) and details coefficients matrices \( cH, cV \) and \( cD \) (Horizontal, vertical and diagonal, respectively), obtained by wavelet decomposition of the input matrix \( X \) where \( X \) is the given input eye image after applying an-isotropic diffusion. The stringname ‘\text{wname}’ denotes the wavelet name. The reversed biorthogonal wavelet is used”.

<table>
<thead>
<tr>
<th>Data</th>
<th>Specificity (%)</th>
<th>Sensitivity (%)</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R</td>
<td>G</td>
<td>B</td>
</tr>
<tr>
<td>1</td>
<td>100</td>
<td>80</td>
<td>40</td>
</tr>
<tr>
<td>2</td>
<td>33</td>
<td>100</td>
<td>50</td>
</tr>
<tr>
<td>3</td>
<td>50</td>
<td>67</td>
<td>67</td>
</tr>
<tr>
<td>4</td>
<td>33</td>
<td>33</td>
<td>33</td>
</tr>
<tr>
<td>5</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>6</td>
<td>33</td>
<td>83</td>
<td>50</td>
</tr>
<tr>
<td>7</td>
<td>67</td>
<td>83</td>
<td>33</td>
</tr>
<tr>
<td>8</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>9</td>
<td>33</td>
<td>50</td>
<td>33</td>
</tr>
<tr>
<td>10</td>
<td>33</td>
<td>50</td>
<td>33</td>
</tr>
<tr>
<td>Mean</td>
<td>47</td>
<td>62*</td>
<td>46</td>
</tr>
</tbody>
</table>
Kirsch template.

“The Kirsch operator is made up of a number of templates. Each template focuses on the edge strength in one direction. This edge detector performs convolution with 8 masks calculating gradients. The Kirsch edge detection algorithm uses a 3×3 table of pixels to store a pixel and its neighbors while calculating the derivatives. The 3×3 table of pixels is called a convolution table, because it moves across the image in a convolution-style algorithm. The Kirsch edge detection algorithm identifies both the presence of an edge and the direction of the edge and finally detects the blood vessels. For a convolution table, calculating the presence and direction of an edge is done in three major steps”:

1. Calculate the derivative for each of the eight directions.
2. Find the value and direction of the maximum derivative.
   \[
   \text{EdgeMax} = \text{Maximum of eight derivatives} \\
   \text{DirMax} = \text{Direction of EdgeMax}
   \]
3. Check if the maximum derivative is above the threshold

Feature extraction

In two steps:

- Detecting the optic nerve
- Detecting diseases.

(See appendix 5)

Classification using PNN

“PNN is a multilayered feed forward with four layers namely input layer, pattern layer, summation layer and output layer. The optimal procedure for this work is obtained by choosing the image features as inputs that is seven inputs and number of neurons in the pattern and summation layers as seven neurons and the output is obtained. The output corresponds to one of the three classes of the input images namely, normal, DMD and WMD separately. The model has been trained with 200 input images and tested with 100 images”.

Results

The result: Sensitivity 94 % and specificity 95 %.

“The Sensitivity and Specificity for the three classes of eye images using PNN classifier: Shows that 96 % is identified”.

<table>
<thead>
<tr>
<th>Input images</th>
<th>No. of trained images</th>
<th>No. of tested images</th>
<th>Classification %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal, DMD &amp; WMD</td>
<td>200</td>
<td>100</td>
<td>96.00</td>
</tr>
</tbody>
</table>
4.6 Diabetic retinopathy

4.6.1 Example 13 – Image processing, detection of exudates [32]

This method assumes that the optic disk has been located.

Approach based on morphological techniques

In the green channel $f_g$, the exudates appear most contrasted. This proposed algorithm can be divided into two parts:

- Finding of the candidate regions – The regions are characterized by high contrast and a high grey level. They can be confused with vessels if local contrast is applied. This can be avoided if the vessels are eliminated by a closing with $s_1$ such that $s_1B$ is larger than the maximal width of the vessels.

$$e_1 = \phi^{(s_1B)}(f_g)$$

(For further description see appendix 8)

- Finding the contours – To be able to find the contours and distinguish them from other well contrasted regions, some further calculation need to be made. (See appendix 8)

“This algorithm has three parameters: the size of the window (...), and the two thresholds $\alpha_1$ and $\alpha_2$. “The choice of the size of $W$ is not crucial, and we have found good results for a window size of 11 x 11. However, if the window size is chosen too large, small isolated exudates are not detected. We found that not very disturbing, because small isolated exudates do not play an important role for diagnostic purposes. The first threshold $\alpha_1$ determines the minimal variation value within the window that is suspected to be a result of the presence of exudates. If $\alpha_1$ is chosen too low, specificity decreases; if it is set too high, sensitivity decreases. The parameter $\alpha_2$ is a contrast parameter: It determines the minimal value a candidate must differ from its surrounding background to be classified as an exudate(...).”

Results

If a human grader does not agree with the algorithm, this can be due to an error of the human grader or due to an error of the algorithm. “One more problem arises, if (as for exudates) it has to be defined, when an exudates can be considered as having been detected and when not (pixelwise against objectwise comparison of segmentation results)”.

“If an algorithm can find all exudates, but not the borders in a correct manner, it will have good statistics but poor performance(...).” A variant of this method is a pixelwise comparison.

1) Comparison of the proposed method with human graders: The algorithm was tested on an image database of 30 images. “These images have not been used for the development of the algorithm”. 15 images did not contain exudates and 13 of these 15, no exudates were found by our algorithm. A few false positives (<20 pixels) were
found in two images. “In order to compare the results (for the 15 images containing exudates) obtained by the algorithm with the performance of a human grader, we asked a human specialist to mark the exudates on color images. In that way, we obtained a segmentation result \( t_1 \) that we considered to be correct. Then we applied the proposed algorithm and obtained a segmentation result \( t_2 \) that we considered to be correct. Let \( S(f) \) be the support of \( f \) defined as the set of pixels for which \( f(x) \neq 0 \), let \( \#S \) be the number of its elements and ‘\( \setminus \)’ the set difference. We define true positives, false negatives and false positives in the following way:

\[
TP = \#[S(r_1 \wedge r_2)] \\
FN = \#[S(r_1) \setminus S(r_1 \wedge \sigma^B r_2)] \\
FP = \#[S(r_2) \setminus S(\sigma^B r_1 \wedge r_2)]
\]

The result is shown below:

<table>
<thead>
<tr>
<th>Image</th>
<th>( TP )</th>
<th>( FP )</th>
<th>( FN )</th>
<th>sensitivity</th>
<th>pred. value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image 1</td>
<td>1318</td>
<td>137</td>
<td>39</td>
<td>97.1</td>
<td>90.6</td>
</tr>
<tr>
<td>Image 2</td>
<td>3792</td>
<td>136</td>
<td>404</td>
<td>90.3</td>
<td>96.5</td>
</tr>
<tr>
<td>Image 3</td>
<td>4188</td>
<td>82</td>
<td>311</td>
<td>93.0</td>
<td>98.1</td>
</tr>
<tr>
<td>Image 4</td>
<td>430</td>
<td>12</td>
<td>38</td>
<td>91.9</td>
<td>97.3</td>
</tr>
<tr>
<td>Image 5</td>
<td>531</td>
<td>40</td>
<td>31</td>
<td>91.9</td>
<td>89.8</td>
</tr>
<tr>
<td>Image 6</td>
<td>1204</td>
<td>106</td>
<td>71</td>
<td>94.4</td>
<td>91.9</td>
</tr>
<tr>
<td>Image 7</td>
<td>3758</td>
<td>190</td>
<td>193</td>
<td>95.1</td>
<td>95.2</td>
</tr>
<tr>
<td>Image 8</td>
<td>762</td>
<td>66</td>
<td>55</td>
<td>93.3</td>
<td>92.0</td>
</tr>
<tr>
<td>Image 9</td>
<td>975</td>
<td>42</td>
<td>103</td>
<td>90.4</td>
<td>95.9</td>
</tr>
<tr>
<td>Image 10</td>
<td>520</td>
<td>192</td>
<td>85</td>
<td>86.0</td>
<td>73.0</td>
</tr>
<tr>
<td>Image 11</td>
<td>1715</td>
<td>125</td>
<td>71</td>
<td>96.0</td>
<td>93.2</td>
</tr>
<tr>
<td>Image 12</td>
<td>2577</td>
<td>245</td>
<td>145</td>
<td>94.7</td>
<td>91.3</td>
</tr>
<tr>
<td>Image 13</td>
<td>430</td>
<td>8</td>
<td>71</td>
<td>85.8</td>
<td>98.2</td>
</tr>
<tr>
<td>Image 14</td>
<td>1311</td>
<td>163</td>
<td>47</td>
<td>96.5</td>
<td>88.9</td>
</tr>
<tr>
<td>Image 15</td>
<td>579</td>
<td>37</td>
<td>32</td>
<td>94.8</td>
<td>94.0</td>
</tr>
</tbody>
</table>

Figure 4.6.1 – Result of the evaluation method

2) Influence of the Parameters: “The robustness of an algorithm can be defined in respect to changes in the parameters or to image quality (influence of noise, low contrast, resolution (...). The behaviour of the algorithm concerning changes of parameters, were studied”. In order to study the influence of the parameters \( \alpha_1 \) and \( \alpha_2 \), we vary the two parameters within a wide range. “The algorithm with the following sets of parameters were tested”:

* \( \alpha_1 \in \{3, 6, 9\} \);

* \( \alpha_2 \in \{2, 3, 8, 10, 12, 14, 16, 20\} \)

The segmentation results get worse for \( \alpha_1 = 9 \), but a specificity of over 80% is obtained for a predictive value larger than 80% with a good choice of \( \alpha_2 \). The choice for the parameters is not independent: “The lower \( \alpha_1 \), the higher \( \alpha_2 \) has to be chosen in order to get acceptable results and vice versa”. The behaviour of the algorithm with respect to parameter changes is quite robust.
4.6.2 Example 14 – Classification of DR lesions [42]

This method angles at the classification of diabetic retinopathy lesions. It aims to distinguish drusen from hard exudates (Diabetic Retinopathy) and evaluate the severity of e.g. oedema.

Methods

Stereo camera geometry [61], "(...) where \( W \) is a world point that is imaged by a left and right camera. It can be shown that depth is related to disparity by":

\[
Z = \lambda - \frac{AB}{(x_2-x_1)}
\]

Where \( B \) is the baseline, \( \lambda \) is the focal length and \( Z \) is the true world coordinate for depth.

The problem of obtaining depth from images is a two part process:

- “features in the left and right images are matched to one other (correspondence process)"
- “the disparity in location of corresponding features is determined. From this disparity the location of the feature in three dimensional space is determined”.

“For corresponding points, the images have to be adjusted to satisfy the epipolar conditions (...). This implies that a point on the right image lies on the same horizontal line with the correspondent point on the left image. In order to achieve this process, one of the images of the stereo pair may need to be rotated, scaled and translated. A global registration technique based on Fourier Mellin Transform is used to determine the scaling and rotational factors (...). The result of this process is that the required scale change and rotation are reduced to translations. Once the images have been rotated and scaled, a correlation scheme is used to determine the vertical translation required to satisfy the epipolar condition”.

To obtain the highest contrast possible, the green channel of the image is used. Both left and right images undergo a radiometric correction which matches the radiometry of the images by making use of the variance, mean and dynamic range of intensities for each image. “The correspondence search is implemented as a window based correlation search (...)”
\[ \Psi(s, t) = \left( \frac{\sum_x \sum_y [f_d(x,y)] [w_d(x-s,y-t)]}{\sum_x \sum_y [f_d(x,y)]^2 \sum_x \sum_y [w_d(x-s,y-t)]^2} \right) \]

“where \( f_d(x,y) = f(x,y) - f_m(x,y) \)

\[ w_d(x-s, y-t) = f(x-s, y-t) - f_m(x-s, y-t) \]

and \( \Psi \) is the cross-correlation coefficient at point \( (s, t) \) over the neighbourhood defined by \( w \) \( (x, y) \). For an \( N \times N \) stereoscopic pair of images, \( x, y, s, t = 0, 1, 2, ..., N-1 \), and \( f_m, w_m \) are the mean values in the regions of the two images”.

“A 15 x 15 window is used as it gives good results in terms of least false matches. Both natural images as well as the segmented images are used. The natural images yield a dense depth map but matching is prone to errors in some regions. The segmented images yield a sparse but accurate depth map. Finally, the disparity map obtained from this process is smoothed using a 5 point box filter. Essentially, the segmented optic disc (...) “[62] “disparity is used as a reference”.

Results

“Regions of interest are selected in different retina images”. In figure 4.6.3 “(...) the region of interest includes hard exudates and in (...) “figure 4.6.4 “drusen are found. The measure of disparity is in pixels. The disparity maps obtained after segmentation and smoothing are shown in each figure as 3D perspective plots. In (...) “ figure 4.6.3, “the maximum disparity for hard exudates is approximately +14 pixels. For this same image, the optic disc shows a disparity of about 20 pixels. When viewed through a stereoscope the hard exudates do appear superficial”. In figure 4.6.4, (...) “the disparity for drusen is approximately -5 pixels. For this same image, the optic disc shows a disparity of about -4 pixels. When viewed through a stereoscope the drusen are located much deeper than the blood vessel network. In fact, the blood vessel network appears as if it were floating over the drusen”. 

Figure 4.6.3 – Region of interest which includes hard exudates
Conclusions

The result shows that there is a relative difference between hard exudates and drusen. The difference can be used as an additional parameter in the classification process. This method shows that it is possible to use depth information from stereoscopic fundus image when you need to separate DR lesions such as hard exudates from non-DR lesions such as drusen. The severity of retinal thickening can be determined by automatic calculation of the thickened area and the maximum thickness within the area. This method can also be extended to photocoagulation scars. Further work to implement a robust stereo matching algorithm based on neural network technique, may enable more accurate disparity maps from the natural stereoscopic images in noisy environments.

4.7 Subprocesses/Others

4.7.1 Example 15 – Component extraction [43]

Introduction

Extracting the clear retinal image information is not an easy task. In order to do that with less data as assistant diagnosis, the histogram equalization preprocessing of images is researched based on its features. “The clariy is observed and analysed by comparing equalization images of colored retinal image’s gray image, red component image, green component image and blue component image”. The green component is the clearest one, followed by the red one and then the blue one.

Theoretical basis of equalization

“In the form of discrete \( r_k \) is used to represent discrete gray level, \( p(r_k) \) is used to represent \( p(r) \), then are set up under”:

\[
p(r_k) = \frac{n_k}{n} \quad 0 \leq r_k \leq 1, \quad k = 0, 1, \ldots, L - 1
\]
“There into \( n_k \) expresses the number of emerging pixels that is of \( r_k \) in image, \( n \) is the total number of pixels in the image \( n/n \) is the frequency in probability theory, and \( l \) is the total number of gray level. A given image’s gray level distribution within the range of \( 0 \leq r \leq 1 \) can be for any of \( r \) values within \([0, 1]\) to do the following transformation: \( s = T(r) \). The transformation function above should meet the following conditions: Firstly, if \( r \) is in the range of \([0, 1]\), then \( T(r) \)’s single-value increases; Secondly, if \( r \) meet the condition of \( 0 \leq r \leq 1 \), then \( T(r) \) is in the range of \([0, 1]\). The first condition ensures that the image gray level from white to black in the order unchanged. The second condition guarantees that the pixel gray values is to the extent permitted after mapping transformation of the dynamic range of consistency. In addition, anti-transformation should also meet the two conditions above. Anti-transformation from \( s \) to \( r \) can be expressed by the following formula”:

\[
r = T^{-1}(s)
\]

The transformation function \( T(r) \) can control the image grey level probability density function, therefore changing the image grey level, which is the basis of technical histogram amended. The equalization algorithm blow is based on this theory to achieve.

The equalisation algorithm are as follows:

- Take statistics to each grey level pixels of the original histogram, \( n_k \)
- List the original image grey level, \( r_k \)
- Calculate original histogram, \( p(r_k) \)
- Calculate cumulative histogram, \( s_k \)
- Ensure the transformation of the corresponding relationship between \( r_k \) and \( s_k \)
- Take statistics to each grey level pixels of the new histogram
- Calculate new histogram

The algorithm can achieve equalization of an image, in order to enhance the dynamic range of contrast.

**Equalization algorithm**

The idea of this approach is to transform the histogram of the original image into the form of homogeneous distribution [63], therefore increasing the value of the pixel dynamic range to achieve the complete enhanced image contrast effects. “Thereinto, cumulative function and mapping functions respectively are”:

\[
s_k = T(r_k) = \sum_{j=0}^{k} \frac{n_j}{n} = \sum_{j=0}^{k} p(r_j) \quad 0 \leq r_j \leq 1, \quad k = 0, 1, \ldots, L - 1
\]

\[
\hat{s}_k = INT\left[\frac{s_k - s_{\min}}{s_{\max} - s_{\min}} (L - 1) + 0.5\right]
\]

**Result analysis and conclusion**

This method uses the image as a pretreatment study. The images is firstly captured the patient’s retinal photo by CF-60U fundus imaging camera and then changed into digital.
photos by HP high-resolution photo scanner. The image is 714 x 927 pixels. “Gray image of the original image is obtained through the traditional hue and gray mapping formula (…) ”[64]; “three-component gray image from the color image are extracted (…) “[65]. “Subsequently, the gray image contrast enhancement is achievement through the gray distribution equalization algorithm which is described in previous section. Algorithm is achieved in Matlab 7.0 programming environment. The chart blow is a patient’s right eye fundus image, erythema in the chart for fundus bleeding, ant the realizing result is shown in Figure 2. Thereinto, the pictures from (a) to (c) are respectively the original image the images of the corresponding equalization and histogram; the picture from (d) to (f) are respectively the G component of the gray image extracted from the original image, the images of the corresponding equalization and histogram, in which the horizontal axis expresses different gray level, thus the vertical axis expresses the number of pixels of each gray level from gray image”;

“(g), (h) are respectively the images of the R component, B component of equalization from original image, in which vascular details are vague, containing more information interference. Use the method above to deal with the fundus vessels. Histogram equalization increases the dynamic range of gray image, and realizes the enhancement of the image, the comparison from (a), (b) and (d), (e) shows that the image after equalization is clearer than the original image. However, (e) is clearer than (b), which can be seen from the histogram (f) and (c), after dealing G component gray image into equalization image, the gray gap of the image in G component equalization image are much bigger than that of the original equalization image, which is reflected that the histogram is sparser in (f) than in (c), so it is clearer after dealing G component gray image into equalization image”.

![Figure 4.7.1 – Retinal images and the result after dealing](image)

A comparison of the results finds that, G component image not only keeps the basic characteristics of information with data compressed down to 1/3, but it also is clearer after equalization. Through the experiment, the method is applicable in retinal image and therefore to extract components from images. It has high clinical value and application.
4.7.2 Example 16 – Background and noise extraction [44]

Introduction

Retinal images are used for automated diagnosis of DR. This is a method for coloured retinal image preprocessing and enhancement. When diagnosing retinal images that may contain features for diabetic retinopathy, it is not necessary to process the background and noisy areas. It consumes a lot of time. By cutting/cropping out the interesting regions, valuable time can be saved. Binary mask are used to cut these regions out. 0 is for background and 1 is for true retinal image pixels.

The binary mask preprocesses the retinal images using morphological operations. The technique is tested on standard retinal to detection of features and abnormalities. Images from the databases Diaretdb0 and Diaretdb1 are used. “The validity of our techniques is checked against the experimental result”. The colour retinal images are normally of different qualities and need illumination equalization to enhance the image quality.

Method

**INPUT RETINAL IMAGE → BACKGROUND MASK + NOISE MASK → FINAL MASK → PREPROCESSED RETINAL IMAGE**

- **Background extraction mask** – Local mean and variance based method which creates a binary background mask.
  1. “Divide the input image $I(i, j)$ into nonoverlapping blocks with size $w \times w$. In this test, the value of $w = 8$”.
  2. Calculate the local mean value $M(I)$ for each block.

$$M(I) = \frac{1}{w^2} \sum_{i=-w/2}^{w/2} \sum_{j=-w/2}^{w/2} I(i, j)$$

  3. Use the local mean value computed in step 2 to calculate the local standard deviation value $std(I)$:

$$std(I) = \sqrt{\left(\frac{1}{w^2} \sum_{i=-w/2}^{w/2} \sum_{j=-w/2}^{w/2} (I(i,j) - M(I))^2\right)}$$

  4. “Select a threshold value empirical by working on different retinal images. If the $std(I)$ is greater than threshold value, the block is considered as original retinal image area otherwise it belongs to background”.

- **Noise removal mask** – Noise is normally due to noise pixels and pixels whose colour is distorted. This mask is applied on retinal image to ensure that the noisy area will not be processed in upcoming steps. The RGB retinal images are converted into HIS colour space. The HIS colour space is closer to the way a human experience colours and it makes noise removal easier.

  1. Divide the input retinal image exactly like step 1 in the “Background extraction mask”.
  2. “Use histogram equalization to enhance the contrast between background and foreground”.

STH KTH, Flemingsberg
3. “Use a 3 x 3 median filter to reduce the noise in background of the image (…)” [56].

4. Convert the equalized and filtered RGB retinal image into HIS color space using the following equations [56].

\[
H = \begin{cases} 
\theta & \text{if } B \leq G \\
360 & \text{if } B > G 
\end{cases} 
\]

Where \( \theta = \cos^{-1}\left(\frac{\frac{1}{2}(R-G) + (R-B)}{(R-G)^2 + (R-B)(G-B))^{1/2}}\right) \)

\[ R = \text{Red component}, \ G = \text{Green component}, \ B = \text{Blue component}. \]

\[ S = 1 - \frac{3}{(R+G+B)} \left(\min(R, G, B)\right) \]

\[ I_n = \frac{1}{3} \left( R + G + B \right) \]

5. Calculate noise factor, N, due to inadequate illumination using equation N (I):

\[ N \ (I) = \frac{H}{I_n} \]

6. “Select a threshold value empirically working on different retinal images. If the N (I) is less than threshold value, the block is considered as normal retinal image area otherwise it belongs to noisy area”.

- Final mask for preprocessing – The final mask is a combination of background mask and noise mask. Before the final mask is ready for use, it is applied with morphological operations i.e. morphological erosion and morphological dilation. The mask contains single pixel noise and edge pixels. A square structure element is used for erosion. The erosion removes all white single pixel noise but increases the black single noise. The same square structuring element is used for dilation. At the end, a noise free mask is applied for its preprocessing segmentation.

Results

The Diaretdb0 contains 130 retinal images and diaretdb1 contains 89 retinal images. These databases contain a total of 219 images of different qualities. Based on human eye observation, the decision for accurate processing and poor processing was made. The figures show that the technique gives good results for both high and low noisy areas.

<table>
<thead>
<tr>
<th>Masks</th>
<th>Accurately Processed (Numbers)</th>
<th>Accurately Processed (%)</th>
<th>Poorly Processed (Numbers)</th>
<th>Poorly Processed (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Background Preprocessing</td>
<td>217</td>
<td>99.08</td>
<td>2</td>
<td>0.92</td>
</tr>
<tr>
<td>Noise Preprocessing</td>
<td>202</td>
<td>92.23</td>
<td>17</td>
<td>7.77</td>
</tr>
<tr>
<td>Final Preprocessing</td>
<td>201</td>
<td>91.78</td>
<td>18</td>
<td>8.22</td>
</tr>
</tbody>
</table>

Figure 4.7.2 – Result from preprocessing
Example 17 – Feature extraction [45]

Introduction

This is a method to automatically extract the main features in colour retinal images. The OD is localised by the PCA and its shape is detected by a modified active shape model (ASM). Exudates are extracted and the fovea is located.

This is a novel method to detect optic disk, fovea, exudates and to set up fundus coordination system.

Method

- Disk localization by PCA – The method was described in [13]. The procedure is performed on an intensity image and contains two steps
  1. Determine the candidate regions
  2. The pixels with the highest 1% grey level are selected.

A single pass method [66] is used to cluster those pixels (from step 2). “A cluster is abandoned if the number of pixels in it is less than 0.004% of the total pixel number of the whole image”. This is because it is most likely caused by noise or bright lesions.

- Disk boundary detection by a modified ASM – The disk boundaries are not always well defined, which make the detection of disk shape complicated. A modified ASM is applied. ASM is proposed in [67][68]. The method is a searching procedure to fit the point distribution model, PDM, in a new image to find the modeled object. “A shape of optic disk is represented by the position of $n$ ($n = 48$) landmark points (...). The eight training shapes are aligned by a transformation that includes translation, rotation and scaling. The alignment is performed by minimizing the Euclidean distance between the shapes using a routine least square approach. PCA is next performed on the aligned training shapes. A shape model can be represented by”

$$x = \tilde{x} + \varphi b$$

“where $\tilde{x}$ is the mean shape of the aligned training set, $b = (b_1, b_2, K, b_i)^T$ is termed shape parameter vector, and $\varphi = (\varphi_1, \varphi_2, K, \varphi_r) \in \mathbb{R}^{2n \times t}$ is the set of eigenvectors corresponding to the largest $t$ eigenvalues of the covariance matrix of the training shapes. The first four eigenvectors are selected ($t = 4$) in our application”. The equation above is termed as PDM. “It is a statistical description of the disk shape and its variations of the training set”.

“The space defined by the input image is referred to as image space and the space described by the (...)” PDM equation “(...) is termed as shape space. The variables in the shape space and the image space are denoted by the lowercase and uppercase respectively in this paper. The transformation between the two spaces is defined by”

$$X = T(x) = \begin{pmatrix} s \cos \theta & -s \sin \theta \\ s \sin \theta & s \cos \theta \end{pmatrix} \begin{pmatrix} x_i \\ y_i \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \end{pmatrix}$$
“where $x$ and $X$ are the shape models in the shape space and image space respectively. $X_i$, $y_i$ denote the coordinates of $i$th landmark point in the shape space. $t_x$, $t_y$ represent the position of the model center in the image space. $\tau(s, \theta, t_x, t_y)$ determines the transformation between the two spaces, which is termed as pose parameter”.

“The first step in ASM is initialization. The disk localization ($L_x, L_y$) and the mean shape are utilized to initialize the shape model in the image space according to (...) equation $X$ above, (...) where $x = \bar{x}$, $i = 1$, $\theta = 0$, $t_x = L_x$, $t_y = L_y$. Matching point detection is the second task. The first derivative of the intensity distribution along the normal profile is employed to find the matching point for each landmark point. A blood vessel is identified by a negative pulse followed by a positive pulse within the width range of vessels, and there is a single negative pulse where disk edge appears. The last part of ASM is parameter update. The pose parameter $\tau(s, \theta, t_x, t_y)$ can be updated by minimizing the following expression”.

$$E_\tau = (Y - T(x))^T(Y - T(x))$$

“where $Y$ is the set of matching points in the image space. The inverse transformation is used to transform the matching points $Y$ in the image space back to $y$ in the shape space. The shape parameter $b$ is updated by projecting the matching points $y$ onto the shape space”.

$$b = \phi^T(Y - \bar{x})$$

“The constriction of $b_i \leq 3 \sqrt{\lambda_i}$ is applied to $b_i$ so that a new shape will be similar to those in the training set, where $\lambda_i$ is the $i$th largest eigenvalue. Finally the shape model is updated in the shape space and in the image space according to (...) PDM equation and equation for $X = T(x)$” respectively. The procedure of matching point detection and parameter update is iterated until the shape model $X$ is converged. The original ASM is improved in two aspects to eliminate the influence of the misplaced matching points: adding the self-adjusting weight and exclusion of the outlying points. These two modifications make the algorithm more favorable for the cases of weak edges. Weight factor is added to (...)” equation $E_\tau$ above.

$$E_\tau = \sum_{i=1}^{n} (Y_i - X_i)^T W_i (Y_i - X_i)$$

“where $Y_i$ and $X_i$ are the positions of the $i$th matching point and the $i$th model point in the image space respectively, and $W_i$ is the weight factor. The transform for alignment is performed twice in each iteration: once with the initialized weight $W_i$ and once with the adjusted $W_i$. The initialization of $W_i$ is expressed as the following”.

$$W_i = \begin{cases} 
1 & \text{if } Y_i \text{ is detected directly} \\
0.7 & \text{if } Y_i \text{ is estimated by nearby matching points} \\
0 & \text{if } Y_i \text{ is updated by } X_i 
\end{cases}$$

“$W_i$ is set to zero to eliminate the effect of $Y_i$ when $Y_i$ cannot be detected and the nearby matching points cannot be detected either. $W_i$ is adjusted as the following, which is a negative feedback”. 

W_i = \begin{cases} 
1 & E_i < 5 \\
5/E_i & 5 \leq E_i \\
1/E_i & E_i > 15
\end{cases}

“where E_i is the Euclidean distance between the matching point Y_i and the updated landmark points X_i. The pose parameter \( \tau \) in the iteration is finally obtained by minimizing the \( (...) \) “equation E_\tau \) above \( (...) \) with the adjusted weight factor. Another modification is excluding outlying points in the update of the shape parameter \( b \). In each iteration, the shape parameter \( b \) is obtained in the same way as the original ASM first. A matching point is considered to be an outlying point or misplaced matching point when \( E_i \) between the matching point \( Y_i \) and the updated landmark point \( X_i \) is larger than a constant value. Those outlying points will not be used in obtaining shape parameter \( b \)

\[
b = \hat{\phi}^T (\hat{y} - \hat{x})
\]

“where \( n_m \) is the number of the outlying points \( b \in R^t, \hat{\phi} \in R^{2(n-m)}, \hat{y} \in R^{2(n-m)}, \hat{x} \in R^{2(n-m)}, \hat{y}, \hat{\phi}, \hat{x} \) correspond to \( y, \phi \) and \( x \) in \( (...) \)” equation for shape parameter \( b \).

“The final shape model is estimated from \( (...) \)” PDM equation “by reconstructing the shape model in 2n-D space with the same \( b \) obtained from \( (...) \)”

- **Foveal coordinate system establishment** – “A polar fundus coordinate system is established based on the fovea localization to describe the spatial locations of the features in fundus images. The fovea is the darkest part in most fundus images, while it is not obvious in some images due to high illumination or being covered by the lesions. It is situated about 2DD (DD = disk diameter) temporal to the optic disk in fundus images \( (...) \)” \[69\]. “The main courses of blood vessels are extracted by the modified ASM introduced in \( (...) \)” earlier. “They are represented by thirty landmark points \( (...) \). Eight landmark sets are utilized to derive PDM and the first four eigenvectors are chosen. Observing the main courses of blood vessels, its shape is roughly a parabolic curve. The extraction result is fitted to a parabola for the future localization of fovea. A generalized parabola can be described as”

\[
\begin{align*}
((x - x_c) \sin \theta + (y - y_c) \cos \theta)^2 &= 2p [(x - x_c) \cos \theta - (y - y_c) \sin \theta]
\end{align*}
\]

“where \( p/2 \) is the focal length, \((x_c, y_c)\) is the vertex, and \( \theta \) is the rotation of the directrix. Four parameters \((p, x_c, y_c, \theta)\) need to be estimated. The ideas of Hough transform and linear least square fitting are combined in the curve fitting. The rotation \( \theta \) is quantized to eliminate the nonlinear relationship between parameters. The vertex can be approximated at half optic disk radius nasal to optic disk, thus the parabolic fitting is simplified as estimating the only variable \( p \) by the least square fitting. The candidate region of fovea is defined as a circle area. Its center is located at 2DD from the disk center along the main axis of the fitted parabola and its radius is set as 1DD. The pixels with the lowest intensity in the candidate region are selected, the sum of which is the area of optic disk, because the fovea is about the same size as optic disk \( (...) \)” \[69\]. “These pixels are clustered by the single pass method \( (...) \)”
described earlier. “The mean intensity of each cluster is calculated and the lowest two are compared. As the fovea is not obvious in some images, the comparison is to avoid mistaking the peripheral area, where the illumination is relatively dark, as fovea. When the difference is obvious and the number of pixels in the cluster is greater than 1/6 disk area, the foveal center is located by the centroid of the cluster with the lowest mean intensity, otherwise it is estimated at the center of the candidate region. A polar coordinate system centered on the fovea is selected in our work according to the ETDRS report (...)” [70]. “A fundus image is divided into 10 sub-fields by three fovea-centered circles with the radii of 1 3DD, 1DD, and 2DD respectively. The 10 sub-fields are defined as: (1) central sub-field within the inner circle; (2) four inner sub-fields between the inner and middle circles; (3) four outer sub-fields between the middle and outer circles; (4) far temporal sub-field temporal to the outer circle”.

- Exudate detection – “Luv is selected as the suitable color space for exudate detection (...)” [71]. “As the illumination in fundus images is not homogeneous, a fundus image is divided into sixty-four sub-images. Exudate detection is performed in each subimage. The color difference of an object can be defined as:"

\[ D(i, j) = \sqrt{(L(i, j) - L_r)^2 + (u(i, j) - u_r)^2} \]

“where \( L(i, j) \) and \( u(i, j) \) are the colors of pixel \( (i, j) \) in the component \( L \) and \( u \) respectively. \( L_r \) and \( u_r \) are the reference colors of the object. The reference color is determined as the gravity center of the object (...)”[71]. ”Mean squared Wiener filter is performed to remove noise. A combined method of region growing and edge detection, which includes seed selection, edge detection and growing criteria, is employed here to detect the exudates”.

“It is noted that some local minima are from the retinal background since the retinal background is uneven. Local minima below a certain threshold are chosen as the seeds. The edges in a sub-image are detected by the Canny edge detector. As some weak edges still cannot be detected, other features are examined besides checking if the region has reached an edge. Three criteria are employed”: (a) “The gradient of the pixel is lower than a threshold \( T_1 \)”;(b) “The difference between the pixel value and the mean value of the region is lower than a threshold \( T_2 \)”;(c) “The difference between the pixel value and the value of the seed is lower than a threshold \( T_3 \)”.

Results

89 images were obtained. The disc localization algorithm was tested by all images. 35 images were provided by Singapore National Eye Center (SNEC) and was used as testing images for the other proposed algorithms, as verification from ophthalmologists is only available for this batch of images. The following success rate was achieved: 99 % disk localization, 94 % disk boundary detection and 100 % fovea localization. For exudates, the sensitivity was 100 % and the specificity was 71 %.
The details of the result:

- **Disk localization** – The algorithm failed in one of the testing images and that is probably because of the large area of lesions around the OD in that particular image. There are no such case in the training set and therefore it can’t recognize the pattern. “More constraints such as checking the convergence of the blood vessel network could be added to validate the localization of optic disk”.

- **Disk boundary detection** – “The modified ASM detected the disk boundary successfully in thirty-three images, while the original ASM failed in seven of them. In the twenty-six images where both methods succeeded, the modified ASM also achieved better or at least as good as the results of ASM. The modified ASM needs less iterations in all the images except two cases. The comparison shows that the modified ASM can give more robust result than the original ASM especially when there are several misplaced matching points and also converges faster”.

- **Foveal coordinate system establishment** – “The fovea is detected directly by the centroid of the darkest cluster in twentyone of the testing images. It is estimated in the other fourteen images. The localization of the fovea is within the region of fovea in all these thirty-five images. But the localization deviates slightly from the apparent center in three of the images when evaluated by the human eyes. The localization of fovea is estimated in all of these three images. The reason of the deviation is that the estimation may not be precise”.

- **Exudate detection** – “In the thirty-five testing images, seven images were identified to have no exudate by ophthalmologists. The presence of exudates was successfully detected in all the twenty-eight images. However, exudates were detected by our algorithm in two images in which no exudate is present”.

### 4.7.4 Other methods

The following methods are not presented in the report but they are based on previously mentioned methods. Their result can be studied in the table below.

<table>
<thead>
<tr>
<th>OD Detection methods</th>
<th>Local Dataset</th>
<th>STARE Dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tobin – Vasculature – related OD properties &amp; Bayesian classifier [46].</td>
<td>81 %</td>
<td></td>
</tr>
<tr>
<td>Lalonde et al.– Hausdorff-based template matching pyramidal decomposition…[47].</td>
<td>100 %</td>
<td>71.6 %</td>
</tr>
<tr>
<td>Walter and Klein – Largest brightest connected object [48].</td>
<td>100 %</td>
<td>58.0 %</td>
</tr>
</tbody>
</table>
The two latest shows a high success rate on their LOCAL dataset but a low success rate on STARE dataset.
5. Analysis

Finding the optic disk

Example 1 is a general approach and is not analysed.

Example 2 – The PCA method claims to be robust and good in the presence of large area of light lesions. It has been cited by other articles 72 times since it was published by the IEEE in 2001. The images show acceptable quality. The result is not presented in any numbers.

Example 3 - Use 1000 images and has a correction rate of 99.9 % which is remarkably good. It is almost too good to be true. The failed image claims to have had to low contrast. OD centers estimated to be 9.75 pixels. The article has been cited 17 times since it was published by the IEEE in 2006.

Example 4 – Use both STARE and DRIVE database. In STARE, ODs were found in 80 out of 81 images and in DRIVE all 40 image were a success. The OD centers were 17 pixel. The article has been cited 52 times since it was published by the IEEE in 2008.

Example 5 – 81 images were used and the nerve was visible in all of them. In 5 images, the nerve was obscured by haemorrhaging and in 14, a fraction of the nerve was visible at the border. The article has been cited 168 times since it was published by the IEEE in 2003.

<table>
<thead>
<tr>
<th></th>
<th>Equalised brightness</th>
<th>Fuzzy convergence, single scale</th>
<th>Fuzzy convergence, multiple scale</th>
<th>Fuzzy conv. + Equalised brightness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Healthy</td>
<td>77 %</td>
<td>74 %</td>
<td>87 %</td>
<td>100 %</td>
</tr>
<tr>
<td>Disease</td>
<td>52 %</td>
<td>74 %</td>
<td>74 %</td>
<td>82 %</td>
</tr>
<tr>
<td>All</td>
<td>62 %</td>
<td>74 %</td>
<td>79 %</td>
<td>89 %</td>
</tr>
</tbody>
</table>

Example 6 – 29 out of 30 ODs were found. 27 out of 29 borders were found. The method suffered from small false positives, small parts missing and shape irregularities. These problems were due to low contrast and/or outgoing vessels. The article has been cited 193 times since it was published by the IEEE in 2002.

Example 7 – 81 images from STARE database were used. 31 were healthy and 50 had various pathological signs. The article has been cited 122 times since it was published by the IEEE in 2004.

<table>
<thead>
<tr>
<th></th>
<th>TRACK 1</th>
<th>TRACK 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>OD identified</td>
<td>79</td>
<td>79</td>
</tr>
<tr>
<td>OD not identified</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>81</td>
<td>81</td>
</tr>
</tbody>
</table>

Example 8 – The active geometric contour. The article has been cited 4 times since it was published in India in 2009.
### Table 1

<table>
<thead>
<tr>
<th>FINDING OD</th>
<th>NR. Images</th>
<th>Max. local. Var. method</th>
<th>Proposed method</th>
</tr>
</thead>
<tbody>
<tr>
<td>STARE</td>
<td>81</td>
<td>79 %</td>
<td>99 %</td>
</tr>
<tr>
<td>DRIVE</td>
<td>40</td>
<td>100 %</td>
<td>100 %</td>
</tr>
<tr>
<td>DIARETDB0</td>
<td>130</td>
<td>89 %</td>
<td>98 %</td>
</tr>
<tr>
<td>DIARETDB1</td>
<td>89</td>
<td>84 %</td>
<td>100 %</td>
</tr>
</tbody>
</table>

The failure was due to low contrast.

<table>
<thead>
<tr>
<th>Detect boundary</th>
<th>Nr. Images</th>
<th>Hough</th>
<th>GVF snake</th>
<th>Proposed method</th>
</tr>
</thead>
<tbody>
<tr>
<td>STARE</td>
<td>81</td>
<td>83.8 %</td>
<td>89.8 %</td>
<td>94.2 %</td>
</tr>
<tr>
<td>DRIVE</td>
<td>40</td>
<td>93.5 %</td>
<td>98.6 %</td>
<td>99.3 %</td>
</tr>
<tr>
<td>DIARETDB0</td>
<td>130</td>
<td>85.4 %</td>
<td>95.1 %</td>
<td>96.8 %</td>
</tr>
<tr>
<td>DIARETDB1</td>
<td>89</td>
<td>89.2 %</td>
<td>95.4 %</td>
<td>97.5 %</td>
</tr>
</tbody>
</table>

Example 9 – The result was 37 out of 40, 92.50 % (good, acceptable) using Sobel operator. 7 ODs were missed. Two images was just acceptable and the rest were a success. That is a success rate of 80 %. This example also tested with STARE dataset. 33 ODs were found using Sobel operator and 18 using Canny edge detector. The article has been cited 8 times since it was published by the IEEE in 2008.

**Diagnosis of AMD**

Example 10 – Drusen detection. The article has been cited 10 times since it was published in India in 2007. The images seem to be great but there are no important numbers in the article. The result is only answered with words and not shown by numbers. The article has been cited 5 times since it was published by the IEEE in 2007.

Example 11 – The Histogram approach. 144 images were used. 86 of them had signs of AMD. They were separated into 10 equal datasets. Each dataset had 9 AMD images and 6 control images. Then a Ten-fold cross validation was made. The datasets were then divided into 10 subsets and 10 evaluation runs was conducted. Best specificity had the green channel of 62 % and the best sensitivity had the blue channel with 83 %. The best accuracy had the saturation channel with 74 %. The article has been cited 9 times since it was published in 2009. It comes from the Department of Computer Science at the University of Liverpool in UK.

Example 12 – This method had a sensitivity of 94 % and a specificity of 95 % when detecting AMD in images. 200 training images and 100 tested images were used and gave a total classification of 96 %. The article has not been cited since it was published by the IEEE in 2011. That is reasonable since the article is new.

**Diagnosis of Diabetic Retinopathy**

Example 13 – Detection of exudates. The method had an average sensitivity of 92.75 % and an average specificity of >80 %. The article has been cited 193 times since it was published by the IEEE in 2002.
Example 14 – Classification of DR lesions. The article has been cited 6 times since it was published by the IEEE in 1997. The article is showing that it is possible with the depth information from the stereoscopic fundus image, to separate DR lesions such as hard exudates from non-DR lesions such as drusen. The images are good.

**Subprocesses/Others**

Example 15 – Component extraction. The article has been cited 1 time since it was published by the IEEE in 2008. The green component image keeps the basic characteristics of information with data compressed to 1/3. The image is also much clearer after the proposed equalisation method.

Example 16 – Background and noise extraction. The background mask had a success of 99.08% while the noise mask had a success of 92.23%. The final preprocessing had a success rate of 91.78%. The article has been cited 1 time since it was published by the IEEE in 2009.

Example 17 – The following success rate was achieved: 99% disk localization, 94% disk boundary detection and 100% fovea localization. For exudates, the sensitivity was 100% and the specificity was 71%. The article has been cited 51 times since it was published by the IEEE in 2003.
6. Conclusion

After analysing the data collected, the following conclusion has been made:

Optic disk detection

Example 3 is the best method when it comes to locate the optic disk. The OD border was found in 99.9% of the cases. The experiment included 1000 images and the OD was successfully found in 999 out of those 1000. The number of images gives the experiment a solid ground. The remarkably good result can brings some doubts whether the results are reliable. Though the article was published by IEEE and that is a sort of trust stamp.

A less doubted method is example 4. Though that experiment only included images from the STARE and DRIVE dataset (81 and 40 images respectively). The OD was located in 80 from the STARE database and in all the DRIVE images. This approach is the next best and is very much useful.

Drusen detection – diagnosing Age-Related Macular Degeneration

Example 12 is the best method. It is newly published by IEEE. The experiment included 300 images, which is a good amount. The approach had a sensitivity of 94% and a specificity of 95%.

Exudates – diagnosing Diabetic Retinopathy

The highest sensitivity rate of all methods was 100% from example 17. The highest specificity rate came from example 13 with a little over 80%. Example 17 had a specificity rate of 71% and example 13 had a sensitivity rate of 92.75%. If example 17 is chosen as the best method, there will most likely be no false negatives. All affected patients will have be diagnosed and have chance to be treated. If example 13 is chosen, fewer patients may be wrongly diagnosed to have DR. Though with example 13, there is most likely to be some false negatives. From these facts I would say that example 17 is the best method. Maybe if this method could be combined with some other method to receive a higher specificity rate, as long as it doesn’t affect the sensitivity too much.
7. References


STH KTH, Flemingsberg


Appendix

1 Detection of Blood Vessels Boundaries using Statistics

The point (x,y) must fulfill the following equation to be called the one-dimension edge candidate. Where I(x,y) is the intensity in the point. $\eta$ is a constant threshold. $Av$ is the average deviation and Sd is the standard deviation.

$$I(x,y) > Av[2I(x+1,y),I(x+2,y),I(x+3,y)] + Sd[2I(x+1,y),I(x+2,y),I(x+3,y)] + \eta$$

The points are to be called 2-D edges and need to fulfill the following:

a) 2-7 points are a 1-D candidate
b) “If (x,y) is an edge candidate then, at least one of the immediate 8-neighboring points”: 
\{(x + \beta + \gamma) \beta \equiv (0, 1, -1), \gamma \equiv (0, 1, -1) \beta + \beta \gamma + \gamma \neq 0 \} “is also an edge candidate”.

2 Feature Extraction According to a General Approach

Features detection optic nerve:

I) Retina luminance $\ell (i, j)$ – A feature that measures the brightness in the retina. This can be used to help locate retinal lesions.

$$\ell (i,j) = s \frac{1}{M \times N} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I(i - m, j - m)$$

II) Vessel density, $\rho (i, j)$ – the number of vessels in a specific area. The vascularity that feeds the eye, enters in this region.

$$\rho(i,j) = \frac{1}{M \times N} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} b_1 (i - m, j - m)$$

III) Average vessel thickness, $t (i, j)$ – “the closer to the optic nerve, the thicker are the vessels”.

$$t(i,j) = \frac{\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} b(i - m, j - m)}{\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} b_1 (i - m, j - m)}$$

IV) Average vessel orientation, $\theta (i, j)$ – the vessels that are entering are pretty perpendicular to the horizontal raphe of the retina. “The result is an observation of vascular orientation being + - 90° relative to the horizontal raphe when entering the eye and becoming more parallel (i.e., 0°) as the distance from the optic nerve increases”.

$$\theta(i,j) = \frac{1}{M \times N} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} b_1 (i - m, j - m) \ast \cos \theta (i - m, j - m)$$

I-IV - “supports M x N region for every point (i, j) in the image”.

STH KTH, Flemingsberg
Features detecting diseases:

Each texture is described by a vector of properties. This vector represents a point in a multi-dimensional space.

i) “Mean - “The nth moment of about the mean is where m is the mean value of z (the average gray level)”
   \[ \mu_n (z) = \sum_{i=0}^{L-1} (z_i - m)^n \ast p (z_i) \quad (\mu_0 = 1 \text{ and } \mu_1 = 0) \]
   \[ m = \sum_{i=0}^{L-1} z_i \ast p (z_i) \]

ii) Variance – “The second moment (the variance \( \sigma^2 (z) = \mu_2 (z) \) is important for texture description”. This measures the grey-level contrast.
   
   \[ \mu_2 (z) = \sum_{i=0}^{L-1} (z_i - m)^2 p (z_i) \]

iii) Skewness – “The third moment
   \[ \mu_3 (z) = \sum_{i=0}^{L-1} (z_i - m)^3 p (z_i) \]
   is measure of the skewness of the histogram”.

iv) Entropy
   \[ H (z) = -\int p (z) \ln p (z) \, dz \]
   It is Shannon’s entropy of the image window z, p is the grey-level distribution. An approximation of the entropy:
   \[ H(z) \approx \frac{-1}{N_x} \sum_{z_j \in \Omega} \frac{1}{N_p} \sum_{z_j \in \omega_p} g_p(z_i - z_j) \]

v) Correlation distance – “The distances among all the images can be computed using correlation distance”
   \[ d_{rs} = 1 - \frac{(x_r - \bar{x}_r)(x_s - \bar{x}_s)}{\left[(x_r - \bar{x}_r)(x_r - \bar{x}_r)\right]^{1/2}\left[(x_s - \bar{x}_s)(x_s - \bar{x}_s)\right]^{1/2}} \]
   Where \( \bar{x}_r = \frac{1}{n} \sum_j x_{rj} \) and \( \bar{x}_s = \frac{1}{n} \sum_j x_{sj} \).

vi) Zernike moments – “the complex of Zernike moments of order n with repetition l are defined as”
   \[ A_{nl} = \frac{n+1}{\pi} \int_0^{2\pi} \int_0^{+\infty} [V_{nl}(r, \theta)]^* \ast f(r \cos \theta, r \sin \theta) \, r \, dr \, d\theta \]
   “where n=0,1,2,...., and l takes on positive and negative integer values subject to the conditions”
   \[ n - |l| = \text{even}, \quad |l| \leq n. \]
   “The symbol * denotes complex conjugate”.
   “The zernike polynomials
   \[ V_{nl}(x,y) = V_{nl}(r \cos \theta, r \sin \theta) = R_{nl}(r)e^{i\Omega} \]
   are a complete set of complex-valued functions orthogonal to the unit disk. The function f(x,y) can be explained in terms of Zernike polynomials over the unit disk as”.
   \[ f(x,y) = \sum_{n=0}^{\infty} \sum_{l=-\infty}^{\infty} A_{nl}V_{nl}(x,y) \quad n - |l| = \text{even}, \quad |l| \leq n. \]
3 PCA method

The training images are manually square cropped images from optic disk images. The intensity is adjusted linearly to eliminate any illumination difference between the images. “The image can also be considered as a vector of dimension $N^2$ where $N$ is set to 90 according to the size of optic disk. The training set of optic disk images is denoted as vector $T_1, T_2, ..., T_m$ where $m$ is the number of training images. Training images of right eyes are flipped horizontally to get their symmetrical images that can be calculated together with the images of the left eyes. The average image of the training set is defined by (...)” equation 1:

$$\psi = \frac{1}{M} \sum_{i=1}^{M} \Gamma_i$$

“The vector $\varphi_i$ denotes the difference between the training image and the average image“.

Equation 2:

$$\Phi_j = \Gamma_j - \psi$$

“The vector $u_i$ and $\lambda_i$ are the eigenvector and corresponding eigenvalue of the covariance $N^2 \times N^2$ matrix $C$”:

$$C = \frac{1}{M} \sum_{i=1}^{M} \Phi_i \Phi_i^T = AA^T$$

“where $A = [\varphi_1 \varphi_2 ... \varphi_M]$. When $M$ is less than $N^2$, the eigenvectors $u_i$ of $C$ can be obtained by computing the eigenvectors of an $M \times M$ matrix $A^T A$. Assuming $v_i$ is the eigenvector of $A^T A$ such that”

$$A^T A v_i = \mu_i v_i$$

“$Av_i$ is the eigenvector of $C$, which can be seen by pre-multiplying both sides of (...)” the equation by $A$:

$$AA^T A v_i = \mu_i A v_i$$

“The eigenvector $u_i$ of $C$ can be calculated based on the eigenvector $v_i$ of $M \times M$ matrix $A^T A$. The subspaces defined by the eigenvector $u_i$ is called as “disk space” and eigenvector as “eigen disk”. The eigenvector $u_i$ is a linear combination of the original training – image vectors and arranged in descending order according to its associated eigenvalue. In the application, the $M'$ ($M' < M$) most significant eigenvectors as determined by their eigenvalues are enough to represent the training set of images (…)”.

“The $N \times N$ subimage $\Gamma$ is obtained by cropping an $N \times N$ square with the center pixel $(x, y)$ and its intensity adjusted linearly to the same output range of the training image. To project the subimage $\Gamma$ to “disk space”, the mean image $\Psi$ should be subtracted first as equation (2): $\phi = \Gamma - \Psi$. The subimage $\Gamma$ is projected onto the “disk space” by the following transformation”:

$$w_k \approx u_k^T (\Gamma - \psi), \quad k = 1,2, ..., M'$$
The vector of weights $\Omega^T = [w_1, w_2, ..., w_M]$ describes the contribution of each “eigen disk” while representing the newly input subimage by the “eigen disk”. The input subimage can be reconstructed as $\Gamma_p$, $\Gamma_p = \psi + \sum_{i=1}^{M'} w_i u_i$

“The distance between the original image and its projection (reconstruction) onto the “disk space” is calculated to measure the likeness of optic disk. The point with the acceptable small distance indicates the existence of optic disk. Denoting $\varphi_p$ as the projection of $\varphi$ onto “disk space”, the Euclidean distance $E$ at pixel $(x, y)$ is calculated as $(...)”$ equation:

$$E^2 = \|\Phi - \Phi_p\| = (\Phi - \Phi_p)^T (\Phi - \Phi_p) = \Phi^T \Phi - \Phi^T \Phi_p \Phi_p^T \Phi_p + \Phi_p^T \Phi_p = \Phi^T \Phi - \Phi_p^T \Phi_p + \Phi_p^T \Phi_p - \Phi_p^T \Phi_p = \Phi^T \Phi - \Phi_p^T \Phi_p - (\Phi - \Phi_p)^T \Phi_p - \Phi_p^T (\Phi - \Phi_p)$$

“As $\varphi_p$ is the projection of $\varphi$ onto the “disk space”, $\Phi_p [(\Phi - \Phi_p)]$, which means $(\varphi - \varphi_p)^T \varphi_p$ and $\varphi_p^T (\varphi - \varphi_p)$ are both equal to zero. $\Phi_p$ can be expressed as the linear combination of eigenvectors from equation (...)” $\Gamma_p$, “that is. $\Phi_p = \sum_{i=1}^{M'} w_i u_i$. Hence $\Phi_p^T \Phi_p = \sum_{i=1}^{M'} w_i^2$. The computation of Euclidian distance $E$ can be simplified as”

$$E^2 = \Phi^T \Phi - \sum_{i=1}^{M'} w_i^2$$

“The pixel with the smalles $E$ in the retinal image is located as the center of optic disk”.

4 Dynamic time warping

“To map two time series $T$ and $S$, of length $n$ and $m$ respectively, where $T=t_1,t_2,...,t_n$ and $S=s_1,s_2,...,s_m$ a n-by-m matrix will be formed, where the (i,j)th grid point corresponds to the alignment or distance between two points $t_i$ and $s_j$. The warping path, $W$, is then the set of matrix elements that defines a mapping between $T$ and $S$, defined as $W= w_1,w_2,...,w_k$, where $\max(m,n) <= K < m+n-1$. The distance $d(t_i,s_j)$ between two points $t_i$ and $s_j$ is used to identify potential warping paths. There are many distance measures that may be used, the most common one is the Euclidean distance, and this is the measure used in this paper. Thus, $d(t_i,s_j) = w_k = (t_i-s_j)^2$. The minimal warping path is selected by calculating the minimum cumulated distance between $T$ and $S$ as $DTW(T,S) = \min \left[ \sqrt{\sum_{k=1}^{K} w_k} \right]$. 

5 Feature extraction – section 4.5.3

Feature extraction – In two steps

- A) Features detecting the optic nerve – after preprocessing techniques we obtain a better contrast.
  1) Vessel density, $p(x,y)$
  “Vessel density is defined as the number of vessels existing in a unit area of the retina. Since the vasculature that feeds the retina enters the eye, the vessels tend to be most dense in this region”.

$$p(x,y) = b_i(x,y)*w_i(x,y)$$
“Where \( b(x,y) \) is the skeletonized image and \( w(x,y) \) is the convolution window combinedly as morphologically skeletonized window”.

2) Average vessel thickness, \( t(x,y) \)

“Vessels are also observed to be thickest near the optic nerve since most branching of both the arterial and venous structure does not take place until the tree is more distal from the optic nerve”.

\[
t(x, y) = \frac{b(x,y) \cdot w(x,y)}{bt(x,y) + w(x,y)}
\]

“Where \( b(x,y) \) is the binary image”.

- B) Feature detecting diseases – some of the features that are extracted for detecting diseases are given along with their formula:
  a) Area, \( A = \pi r^2 \)
  b) Radius, \( r = \sqrt{\frac{area}{\pi}} \)
  c) Perimeter, \( 2\pi r \)
  d) Mean, \( M = \frac{x}{y} \)
    “Where \( x = \) sum of the items, \( y = \) total number of items
  e) Standard deviation, \( \sigma = \sqrt{\frac{\Sigma x^2}{N}} \), where \( N \) = number of variables
  f) Variance, \( \nu = \sigma^2 \)
  g) Entropy – the “(...) statistical measure of randomness that can be used to characterize the textural feature of the input image, \( I: E = \text{entropy}(I) \)

6 Fuzzy convergence – section 4.4.4

“The proposed method runs in \( O(n) \) time, where \( n \) is the number of line-like shapes. It does not require any amount of inliers; instead, an absolute threshold for strength may be applied to determine if any area should be deemed convergent”.

1) “Fuzzy Segment Model: A line segment is defined by its two endpoints \((x_1,y_1)\) and \((x_2,y_2)\). In this section, a fuzzy segment model is proposed. The fuzzy segment, henceforward denoted as \( F \), is defined by a set of parametric line segments”.

\[
x(t) = x_1 + r\cos(\alpha + \theta) + (x_2 - x_1 - 2r\cos\theta\cos\alpha)t
\]
\[
y(t) = y_1 + r\sin(\alpha + \theta) + (y_2 - y_1 - 2r\cos\theta\sin\alpha)t
\]

Where \( 0 \leq t \leq 1, 0 \leq \theta \leq 2\pi, 0 \leq r \leq R \).

“The amount of “fuzziness” is controlled by the parameter \( R \), which, at zero reduces the fuzzy segment to the single line segment from \((x_1,y_1)\) to \((x_2,y_2)\). The parameter \( \alpha \) corresponds to the orientation of that single segment, and is computed as”

\[
\alpha = \frac{\pi}{2} - \arctan\frac{y_2 - y_1}{x_2 - x_1}
\]

“The fuzzy segment \( F \) defines a set of segments of
orientations and lengths ranging about a line segment. The motivation for the fuzzy segment is best demonstrated through its visualization”. The figure on the left illustrates the calculations for the subset of F where $r = R$. “It may be visualized as a moving line segment, whose initial endpoints are marked (...) at $\theta = 0$. (...) As theta moves from zero to $2\pi$, the endpoints trace the boundaries of circles, one in the clockwise direction, the other counterclockwise. By starting $\theta = 0$ at $a$, the shape of the fuzzy segment remains invariant to orientation”. “The fuzzy segment $F$ is proposed as a model for the area in which an observed line-like shape contributes to a sense of convergence. The model proposes that the contribution of the midpoint of the line-like shape is more compact than the endpoints. The width of the fuzzy segment at its midpoint is”

$$\sqrt{\frac{r^2}{(x_2-x_1)^2+(y_2-y_1)^2}}$$

“while the width at its endpoints is $2R$. The model also proposes that a line-like shape generally only contributes to a convergence in its “near” neighbourhood. The term “near” is subjectively applicable and implies the area within some perceptually relevant distance surrounding the line-like shape. The model allows for an interpretation of what is near via the parameter $R$”.

2) “Convergence Image: Given a binary input image, like the one depicted in (...) process for finding convergence works as follows”
   - “Thin the image (...)”.
   - “Erase (relabel as background) all branchpoints, breaking up the entire foreground into segments that contain two endpoints each. In a thinned image, endpoints may be discovered as any pixel for which a traverse of the eight bordering pixels in clockwise order yields only one foreground-to-background transition. Similarly, branchpoints may be discovered as any pixel for which the same traverse yields more than two transitions”.
   - “Extend each segment a distance of $R$ pixels in both directions. The extension is done along the vector made by the segment’s endpoints. For our experiments, we used $R = 15$ (the average distance between vessels in the nerve in our data set)”.
   - “Model each segment, via its two extended endpoints, with a fuzzy segment $F$. The image area covered by $F$ may be found by enumerating $F$ at $r = R$ with suitable discretizations of $\theta$ and $t$. For the experiments reported herein, $\theta$ was enumerated to produce unique pixel coordinates for endpoints. For each $\theta$, $t$ was also enumerated to produce unique pixel coordinates. Since the line segments at different $\theta$’s overlap, a second binary image is used to keep track of which pixels are found to lie in $F$. This image is cleared after each fuzzy segment enumeration and voting is completed”.
   - “For each pixel enumerated in $F$, a vote is cast. The image used to tally these votes is termed the convergence image. Several voting functions were explored, including weighting by segment size, and weighting by distance from midpoint.”
Interestingly, no function seemed to work better, overall, than the simplest: equal voting (+1) for all pixels in each $F$.

- "Smooth the convergence image to identify the center of the peak of convergences. The measures of convergence have been normalized so that the highest vote appears darkest. For our experiments, we used an 11 x 11 mean filter for smoothing”.

7 Watershed transformation – section 4.4.5

Localising the optic disk

A shade-correction operator is used to remove slow background variations, i.e. $L_1 = f_L - I(f_L) + k \text{ “(...)where k is a positive constant and } I(f_L) \text{ is an approximation of the slow variations of } f_L, \text{ that has been computed by a large median or mean filter, by homomorphic filtering or by alternating sequential filters. In order to avoid artifacts at the borders of bright regions, we use alternating sequential filters to calculate the background approximation with n sufficiently large to remove the optic disc”}.$

$\text{I}(f_1) = \text{ASF}(f_1) = \text{g}^{(\text{LB})}(\ldots \text{g}^{(B)} \left( \text{g}^{(B)}(f_1) \right) \ldots)$

“On the shade-corrected image we calculate the local variation for each pixel $x$ within a window centered at $x$. Let $W(x) \in D_f$ be the set of pixels within a window centered at $x$, $N$ the number of pixels in $W(x) \in D_f$ and let $\mu_{ll}(x)$ be the mean value of $l_1(\xi) \forall \xi \in W(x)$, then we can calculate $l_2$”.

$l_2(x) = \frac{1}{N-1} \sum_{\xi \in W(x)} \left( l_1(\xi) - \mu_{l_2}(x) \right)^2$

“If $k$ is chosen relatively high and the image $l_1$ is clipped, we reduce the contrast of exudates more than the contrast of the papillary region because the high contrast of the latter one is partially due to the vessels. The global maximum of $l_2$ is situated within the papilla or next to it. It allows one to work now on a subimage of the original that contains the optic disc but not large exudates. We note that this approach does not work with mean filters because they do not preserve the borders of image features. Therefore, pixels that do not belong to bright regions but that are close to borders of bright regions are darkened by the shade-correction operator and cause high grey level variation in the shade-corrected image”.

“Working on the subimage obtained in that way, we can apply a simple area threshold on it as we know approximately the size of the optic disc—in order to obtain a binary image $b$, that contains a part of the papilla. Its centroid $c \in D_b$, that can be calculated as the maximum of the discrete distance function of the biggest particle of $b$ (…) can be considered as an approximation for the locus of the optic disc”.
Finding the contour of the optic disk

“\(P_1\) is opened with a large structuring element, in order to remove large peaks. The morphological reconstruction is calculated as this alters the shape of the papillary region”.

\[ p_2 = R_{P_1}(\gamma^{(s_2B)})(P_1) \]

“Before the watershed transformation is applied to the morphological gradient, which may lead to oversegmentation of the image, internal and external markers are imposed”.

\[ \Delta p_2 = \delta^{(B)}p_2 - \varepsilon^{(B)}p_2 \]

“The center \( c \) that has been calculated in the section above are used as internal marker. As external marker, a circle \( C(c) \) with center at \( c \) with radius bigger that the diameter of the optic disk. The marker \( m \) of the image and the result \( P_{fin} \):”

\[
\begin{align*}
m(x) &= \begin{cases} 
\Delta p_2, & \text{if } x \in (c) \cup C(c) \\
t_{max}, & \text{if } x \notin (c) \cup C(c)
\end{cases} \\
p_{fin} &= WS[R_{\Delta p_2}^{*}(m)]
\end{align*}
\]

“with \( WS(f) \) being the watershed transformation of \( f \). This transformation assigns to each local minimum of \( f \) one catchment basin (one connected region), in a way that all \( x \in D_f \) belong to basins (the so-called watershed line). If we write \( P_{fin} = WS[R_{\Delta p_2}^{*}(m)] \), we mean that \( p_{fin}(x) = t_{max} \) for all pixels within the catchment basin that contains \( c \) and \( 0 \) elsewhere”.

8 Detection of exudates – section 4.6.3

Finding of the candidate regions:

After vessel elimination, calculations of the local variation for each pixel \( x \) within a window \( W(x) \) are implemented.

\[ e_2(x) = \frac{1}{N-1} \sum_{\xi \in W(x)} (e_1(\xi) - \mu_{e_1}(x))^2 \]

“Thresholding the image \( e_2 \) at grey level \( \alpha_1 \), we obtain all regions with a standard variation larger than or equal to \( \alpha_1 \), i.e., small bright objects and borders of large bright objects. In order to obtain the whole candidate regions rather than their borders, we fill the holes by reconstructing the image from its borders \( B_f \). We also dilate the candidate region in order to ensure that there are background pixels next to exudates that are included in the candidate regions; this is important for finding the contours”.

\[ e_3 = \delta^{(sB)}(T_{[\alpha_1,t_{max}]})(e_2) \]
\[ e_4 = R_{e_3}^{*}(b) \]

“with \( b =\begin{cases} 
0, & \text{if } x \in B_f \\
t_{max}, & \text{if } x \notin B_f
\end{cases} \).”

“The threshold \( \alpha_1 \) is chosen in a very tolerant manner, i.e., we get the regions containing some exudates, but we also get some false positives: The papillary region and some other areas that are characterized by a sufficiently high grey level variation due to illumination changes in the image. Finally, we have to remove the candidate region that results from the
optic disc. We remove a dilated version of the segmentation result of (...)” Section II in the article.” In that way, we obtain as candidate regions (...)

\[ e_5 = e_4 - e_4 \land \delta^{(SB)}(p_{fin}) \]

**Finding the contours:**

“We set all the candidate regions to 0 in the original image” “and we then calculate the morphological reconstruction by dilation of the resulting image under \( f_g \)”.

\[ e_6(x) = \begin{cases} 
0, & \text{if } e_5(x) \neq 0 \\
 f_g(x), & \text{if } e_5(x) = 0 
\end{cases} \]

“This operator propagates the values \( f_g(x) \) of pixels \( x \) next to the candidate regions into the candidate regions by successive geodesic dilation under the mask \( f_g \). As exudates are entirely comprised within the candidate region, they are completely removed, whereas regions that are not entirely comprised in the candidate regions are nearly entirely reconstructed. The final result (...) is obtained by applying a simple threshold operation to the difference between the original image \( f_g \) and the reconstructed image \( e_7 \)”.

\[ e_{fin} = T_{[\alpha_2, \alpha_{max}]}(f_g - e_7) \]