Running title:
Automatic Analysis of Retinal Images:
Retinopathy Detection and Grading

Enrico Grisan
Supervisor: Ing. Giuliano Barbaro, Nidek Technologies
Advisor: Prof. Alfredo Ruggeri, Università di Padova
Co-advisor: Dr. Peter Weller, City University, London
External Examiner: Dr. Anil Bharath, Imperial College, London

draft: 12th November 2004
Contents

Sommario xi
Summary xiii

1 Introduction 1
   1.1 Aim and Objectives ........................................ 2
   1.2 Outline of the Thesis ...................................... 3

2 Fundus Imaging and its Findings 5
   2.1 Fundus Oculi Examination ................................. 5
      2.1.1 Fundus Oculi Appearance .............................. 6
   2.2 Main Vascular Abnormalities ............................... 7
      2.2.1 Tortuosity .............................................. 7
      2.2.2 Generalized Arteriolar Narrowing .................... 7
      2.2.3 Focal Arteriolar Narrowing ......................... 8
      2.2.4 Bifurcations Abnormalities ......................... 8
      2.2.5 Crossing Abnormalities ............................... 8
   2.3 Main Non Vascular Findings ............................... 9
      2.3.1 Microaneurysms and Red Dots ....................... 9
      2.3.2 Hemorrhages .......................................... 10
      2.3.3 Hard Exudates ...................................... 10
      2.3.4 Cotton Wool Spots ................................ 11
      2.3.5 Drusen ................................................. 12
   2.4 Hypertensive Retinopathy Grading ....................... 12
   2.5 Diabetic Retinopathy .................................... 13

3 Image Acquisition Protocols and Clinical Evaluation 17
   3.1 Diabetic Retinopathy ..................................... 17
      3.1.1 ETDRS Protocol ..................................... 18
      3.1.2 EURODIAB Protocol ................................ 18
      3.1.3 Joslin Clinic Protocol ............................... 20
CONTENTS

3.1.4 Single Field Monochromatic .................................. 21
3.1.5 5-Fields Protocol .............................................. 21
3.1.6 Single-Field Colour Protocol .................................. 22
3.2 Hypertensive Retinopathy ........................................ 22
3.2.1 ARIC Protocol .................................................. 22
3.3 Used Protocol for Hypertensive Retinopathy ....................... 23
3.4 Conclusions ......................................................... 23

4 Clinical Evaluation .................................................. 29
4.1 Image Collection and Database Building .......................... 29

5 Luminosity and Contrast Normalization ............................ 33
5.1 Review of Available Methods ..................................... 34
5.1.1 Low-Pass Filtering ............................................. 34
5.1.2 Homomorphic Filtering ....................................... 35
5.1.3 Median Filtering ............................................... 35
5.1.4 Wallis Filter .................................................. 35
5.1.5 Non-linear filtering .......................................... 36
5.1.6 Local Contrast Enhancement .................................. 36
5.1.7 Vessel Sampling and Interpolation ............................ 36
5.2 A Simple Model of the Observed Image .......................... 37
5.3 Estimation of the Retinal Background ............................. 39
5.4 Estimation of Observed Local Luminosity and Contrast ........... 40
5.5 Estimation of the Illumination and Contrast Models Parameters 41
5.6 Luminosity and Contrast Normalization .......................... 42
5.7 Performance Evaluation ............................................ 42
5.7.1 Local Indexes .................................................. 43
5.7.2 Global Indexes ............................................... 44
5.7.3 Comparison with other correction techniques .................. 45
5.8 Conclusions ......................................................... 46

6 Vessel Tracking ....................................................... 57
6.1 Review of Available Methods ..................................... 58
6.1.1 Local vascular segmentation .................................. 58
6.1.2 Matched Filters .............................................. 59
6.1.3 Artificial Neural Networks .................................... 59
6.1.4 Vessel Tracking ............................................... 59
6.2 The Sparse Approach: Seed Points Extraction ................. 60
6.3 Bubble Analysis ..................................................... 62
6.4 Tracking ............................................................ 63
6.5 False Vessel Elimination .......................................... 67
## CONTENTS

8.7 Crossings Analysis: Gunn’s and Salus’ Signs .......................... 116  
8.8 Conclusions ............................................ 121  

9 Non Vascular Abnormalities Identification 123  
9.1 Review of Available Methods ................................. 123  
9.2 Bayesian Classification of Suspect Areas ......................... 125  
9.3 Blobs Analysis and Features Extraction ......................... 127  
9.4 Feature Selection ........................................... 129  
9.5 Linear Classifier ............................................ 133  
9.6 Blob Classification .......................................... 133  
9.7 Performance Evaluation ....................................... 134  
9.8 Conclusions .................................................. 135  

10 Hypertensive Retinopathy Grading 139  
10.1 From Sign Evaluation to Image Evaluation ..................... 139  
  10.1.1 Tortuosity ............................................ 139  
  10.1.2 Generalized Arteriolar Narrowing ....................... 141  
  10.1.3 Focal Arteriolar Narrowing ............................. 141  
  10.1.4 Gunn’s and Salus’ Signs .............................. 141  
  10.1.5 Non Vascular Lesions ................................. 141  
10.2 A Simple Explicit Network .................................... 144  
10.3 Performance Evaluation ..................................... 150  

11 Conclusions 153  
11.1 Achieving the Objectives ..................................... 153  
11.2 Contribution to Knowledge ................................... 153  
11.3 The way ahead .............................................. 153  
  11.3.1 Diabetic Retinopathy Grading ......................... 153  
  11.3.2 Possible Applications to other Retinopathies ............ 153  
  11.3.3 Applications to other Protocols ....................... 153

iv
List of Figures

2.1 An image of a normal fundus oculi. Papilla (a), fovea (b) and vessel networks are clearly visible ........................................ 6
2.2 ............................................................................................................................ 7
2.3 A definite focal narrowing .................................................................................. 8
2.4 ............................................................................................................................ 9
2.5 ........................................................................................................................... 10
2.6 .......................................................................................................................... 11
2.7 Different hard exudates .................................................................................... 12
2.8 Cotton wool spots ............................................................................................ 12
3.1 Standard ETDRS 7 fields .................................................................................. 19
3.2 Joslin Vision Network Non Mydriatic Fields (black circles), and standard ETDRS 7 fields (dotted blue) ......................... 20
3.3 Non Mydriatic Monochromatic Field proposed in [54] (black circle), and standard ETDRS 7 fields (dotted blue) .............. 21
4.1 ........................................................................................................................... 31
4.2 ........................................................................................................................... 32
5.1 ........................................................................................................................... 49
5.2 Reducing the difference ($\mu_a - \mu_b$) clearly increases the sum of the local contrasts, $\sigma_a^2$ and $\sigma_b^2$, if $\sigma_{a\cup b}$ is kept constant. ........................................ 50
5.3 Histograms of luminosity values in the 25 image blocks for the observed and normalized version of the same image .......... 51
5.4 Intensity profile on a sample row and intensity drift correction estimated by the proposed technique and by a low-pass filter. 52
5.5 Retinal image corrected with different methods .............................................. 53
5.6 Retinal image corrected with different methods .............................................. 54
5.7 Retinal image corrected with different methods .............................................. 55
5.8 Retinal image corrected with different methods .............................................. 56
LIST OF FIGURES

6.1 Bubble analysis at a bifurcation: correct vessel branches are identified ................................................. 64
6.2 Tracking sequence ................................................................. 72
6.3 Identification of internal and external regions of a vessel for evaluating its reliability ................................. 73

7.1 Fundus image with superimposed the main arcades model (blue double parabola). Vessel inside the parabola in the temporal region bend toward the medial axis more evidently than those in the nasal region ................................................................. 77
7.2 $c_1$ and $c_2$ (set equal in these figures) can accomodate different vessel bending behaviour. In blue is shown the double parabola and in red the model direction at sample points. ... 80
7.3 Different values of $c_1$ and $c_2$ can model the different vessel behaviour on the temporal and nasal part of the fundus image. In blue is shown the double parabola and in red the model direction at sample points. ................................. 81
7.4 Three slices of the pentadimensional objective function to be minimized ..................................................... 84
7.5 .......................................................................................... 85
7.6 .......................................................................................... 86

8.1 Inter-class variability: two arteries and one vein taken at different location of the same image. ......................... 89
8.2 Intra-class variability: three arteries taken at different location of the same image. ........................................ 90
8.3 Principal arcades of the retinal vessel network ................................................................. 91
8.4 Division of the fundus ........................................................................ 95
8.5 Selected vessels $\bigcup_{i=1}^{5} S_{Quad_i} \cap C$ in a sample image. In different colour vessels belonging to different quadrants set $S_{Quad_i}$. ................................. 95
8.6 Selected vessels $\bigcup_{i=1}^{5} S_{Quad_i} \cap C$ in a sample image ................................. 96
8.7 Features for the points from the vessels in $\bigcup_{i=1}^{5} S_{Quad_i} \cap C$. Circles are artery sample points, crosses are vein sample points. ................................. 96
8.8 Features for the points from the vessels in $S_{Quad_i} \cap C$. Circles are artery sample points, crosses are vein sample points. ................................. 97
8.9 Result of the A/V classification achieved by the proposed algorithm on a sample image; white vessels were classified as artery, black ones as vein. ................................................. 97
8.10 The first two curves (top and middle panel) have very different tortuosity but the same Length $L$ and Chord Length $\chi$. The second and last curves (middle and bottom panel) have the same average angular difference despite their different tortuosity. The curve in the bottom panel has a curvature integral of $\pi/2$ whereas the one in the middle panel has curvature integral $\pi$, even if that in the bottom panel is clearly perceived as more tortuous.

8.11 Curvature value along a vessel. The dashed lines indicate the hysteretic threshold for the convexity change.

8.12 Evaluation of Generalized Arteriolar Narrowing as computed in 8.29, $\beta = 0.1$ and $\gamma = 1$.

8.13 Focal Arteriolar Narrowing scheme.

8.14 Caliber analysis for the identification of focal arteriolar narrowing zones.

8.15 .

8.16 Bifurcation Scheme.

8.17 Bifurcation Scheme.

8.18 The typical hourglass appearance of a severe compression sign.

8.19 Gunn’s sign scheme.

8.20 Salus’ sign scheme.

9.1 Slices of the probability density functions of the three classes of interest.

9.2 .

9.3 .

9.4 Total area of the identified hemorrhages versus ophthalmologist grading.

9.5 Total area of the identified hard exudates versus ophthalmologist grading.

9.6 Total area of the identified cotton wool spots versus ophthalmologist grading.

10.1 Image tortuosity for the venous network on the DB60. The parameter used is $\beta = 0.52$. In black the least-squares line that interpolates the automatic measures with respect to the clinician grading.

10.2 Image tortuosity for the entire vessel network on the DB60. The parameter used is $\beta = 0.75$ and $\gamma = 0.73$. In black the least-squares line that interpolates the automatic measures with respect to the clinician grading.
LIST OF FIGURES

10.3 Example of the sigmoidal function used to normalize each vessel abnormality measure ................................. 145
10.4 Example of the combination of two sigmoidal functions used to normalize each non vessel lesion area .................. 146
10.5 Total estimated area covered by hemorrhages in the DB60 images versus the ophthalmologist grading .................. 147
10.6 Total estimated area covered by hard exudates in the DB60 images versus the ophthalmologist grading .................. 148
10.7 Total estimated area covered by cotton wool spots in the DB60 images versus the ophthalmologist grading .................. 149
10.8 Automatic grading versus the ophthalmologist grading. The interpolation line between the two gradings is also shown. ... 150
10.9 One of the images in which the system fails to identify non vascular lesions. The hard exudates in the peripapillary zone have a dull appearance and low contrast with the background 151
# List of Tables

2.1 Classification of hypertensive retinopathy as proposed in [45] and [76] .................. 13

2.2 Classification of diabetic retinopathy as proposed in [96] ... 15

2.3 Classification of diabetic macular oedema proposed in [96] ... 16

3.1 Early Treatment Diabetic Retinopathy Study Seven Standard Field Definitions, from [10, 22, 21] .................. 25

3.2 Joslin Vision Network Non-mydriatic Retinal Fields ......... 26

3.3 Results obtained in [74] for four screening protocols ......... 27

5.1 Indexes for local luminosity and contrast for observed ($\sigma_\mu, \mu_\sigma, \mu_c$) and normalized ($\sigma^N_\mu, \mu^N_\sigma, \mu^N_c$) images, and their percent differences. .................. 46

5.2 Indexes for local luminosity obtained with the proposed technique ($\sigma^N_\mu$), the low-pass correction ($\sigma^LP_\mu$), the Wallis filter correction [91] ($\sigma^W_\mu$), the correction proposed by Sinhanayothin [78] ($\sigma^S_\mu$), H. Wang [93] ($\sigma^{HW}_\mu$), and Y. Wang [95] ($\sigma^{YW}_\mu$) .... 47

5.3 Indexes for contrast obtained with the proposed technique ($\mu^N_\sigma$), the low-pass correction ($\mu^{LP}_\sigma$), the Wallis filter correction [91] ($\mu^W_\sigma$), the correction proposed by Sinhanayothin [78] ($\mu^S_\sigma$), H. Wang [93] ($\mu^{HW}_\sigma$), and Y. Wang [95] ($\mu^{YW}_\sigma$) .... 48

6.1 Fraction of recognized vessel on the images of the DB60 database. The fraction is computed as the ratio between the length of the correctly recognized vessels and the length of the manually identified vessels. .................. 70

6.2 .......................................................... 71

7.1 Optic disc identification performance on the STARE images 82

7.2 Optic disc identification performance on the DB60 images 83

8.1 Vessel classification performance .......................... 93
**LIST OF TABLES**

8.2 Correlation of the tortuosity measures ordering with the ophthalmologist tortuosity order ........................................ 106
8.3 Correlation of the measures CRAE, CRVE and A/V ratio with the ophthalmologist evaluation of generalized arteriolar narrowing ........................................ 108
8.4 Correlation of the measures obtained with the (8.29), $\beta = 0.1$ and $\gamma = 1$, with the ophthalmologist evaluation of generalized arteriolar narrowing ........................................ 109
8.5 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 113
8.6 Performance of the crossings recognition. The first column is the fraction of the correctly identified crossings with respect to the manually labelled. ........................................ 117
8.7 Performance of the bifurcation recognition. The first column is the fraction of the correctly identified bifurcations with respect to the manually labelled. ........................................ 118
8.8 Median of the mean junction exponent and of the mean branching angle of each image in the DB60. Between square brackets the first and third quartiles. ........................................ 119
9.1 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 127
9.2 Results for identification of images with non vascular lesions in DB60 ........................................ 134
9.3 Results for identification of images with non vascular lesions in DB200 ........................................ 134
9.4 Correlation between the estimated area covered by lesions and the ophthalmologist grading, on the images of the DB60 . . . 135
10.1 Tortuosity grading performance with respect to the ophthalmologist grading ........................................ 141
10.2 Grading performance of the automatic system with respect to the ophthalmologist grading ........................................ 150
Sommario

Questa tesi tratta dell’analisi automatica di immagini a colori del fondo dell’occhio. In particolare, essa si focalizza sull’applicazione di tale analisi alla valutazione quantitativa della retinopatia ipertensiva in primis, e della valutazione della retinopatia diabetica come naturale ricaduta dei metodi sviluppati.

Sia l’ipertensione che il diabete colpiscono, anche se con modalità e decorso temporale differente, il microcircolo sanguigno. La retinopatia è una delle conseguenze di tale danno circolatorio. La retina ed i vasi retinici sono assai sensibili a cambiamenti nella microcirculazione: è stato dimostrato che i singoli segni della retinopatia hanno un alto valore prognostico per infarto, sclerosi carotidea e danno coronarico.

Inoltre, la retinopatia è una malattia sociale, con una ricaduta economica (diretta ed indiretta) elevata: la perdita o la diminuzione della capacità visiva porta infatti ad una ridotta capacità lavorativa e di condurre una vita indipendente.

Anche altri organi sono sensibili ad alterazioni del microcircolo sanguigno, ma la retina ha il grande vantaggio di essere facilmente disponibile al controllo in maniera non invasiva. Questa caratteristica suggerisce un modo efficiente ed efficace per seguire il decorso di malattie sistemiche associate alla retinopatia.

Inoltre, con riguardo alla prevenzione della perdita della vista, il riconoscimento della retinopatia al suo insorgere è il punto più critico per evitare che degeneri in cecità. Ciò è particolarmente importante nella retinopatia diabetica in cui, allo stato attuale della farmacologia, i danni alla retina non recedono con il trattamento farmacologico o col controllo del diabete. Sfortunatamente, le fasi iniziali della retinopatia sono quasi asintomatiche. Un programma di screening potrebbe evitare alla maggior parte della popolazione a rischio lo sviluppo di retinopatie che minaccino la vista. Allo stesso tempo, nel mondo occidentale non ci sono abbastanza risorse, sia in termine di tempo che soprattutto in termine di disponibilità di oftalmologi esperti, per
organizzare uno screening di tal genere. É dunque presente un forte bisogno di uno strumento che valuti automaticamente la retina, per diagnosticare la presenza e la severità delle eventuali retinopatia.

Ogni metodo che voglia valutare automaticamente le retinopatie ipertensiva e diabetica deve seguire ad alcuni passi obbligati. In primo luogo deve identificare le strutture anatomiche principali della retina: i vasi sanguigni, il disco ottico e la macula. Quindi deve valutare in maniera quantitativa le anormalità presenti nelle strutture identificate. Infine deve essere in grado di identificare le eventuali altre lesioni presenti sulla retina.

In questa tesi verrà inizialmente descritto un algoritmo per la identificazione dell'albero vascolare. Quindi sarà presentato un metodo per identificare la posizione del disco ottico.

Per una identificazione precoce della retinopatia, anormalità dei vasi retinici devono essere identificate e valutate: saranno descritti degli algoritmi che ne misurino automaticamente alcune, utili per la caratterizzazione del quadro clinico della retina.

Sarà poi presentato un metodo per la identificazione di eventuali lesioni. Tra queste, i microaneurismi identificano gli stadi precoci della retinopatia diabetica, mentre altre lesioni sono essenziali nell'identificazione della retinopatia ipertensiva maligna e gli stadi più avanzati della retinopatia diabetica. L'identificazione di microaneurismi, emorragie, essudati duri e noduli cotonosi è dunque di fondamentale importanza nella valutazione della retinopatia, e sarà affrontata dall'algoritmo presentato.

Infine, sarà escritto un primo tentativo di valutare in maniera quantitativa la gravità della retinopatia ipertensiva. Questo è un passo necessario per poter seguire il decorso della retinopatia nel tempo, e per poter confrontare retine diverse in maniera consistente e quantitativa.

Gli algoritmi presentati in questa tesi rendono possibile pensare ad uno strumento utilizzato sia in ampi programmi di screening delle retinopatie ipertensiva e, con minimi cambiamenti, diabetica, sia per il controllo nel tempo del progredire della malattia.

La sua utilità è triplice. In primo luogo, potrà essere uno strumento diagnostico di aiuto alla pratica clinica. In secondo luogo, fornirà dettagli quantitativi sulle singole lesioni della retina, che possono essere utili nella ricerca medica e per meglio caratterizzare gli sviluppi della retinopatia. Infine, renderà possibile alla ricerca farmaceutica l'utilizzo di una misura quantitativa e riproducibile della evoluzione della retinopatia durante un trattamento farmacologico.
Summary

This thesis deals with the automatic analysis of colour fundus images, and with its applications in the evaluation of hypertensive retinopathy at first, and secondarily in the evaluation of diabetic retinopathy. Both hypertension and diabetes affect, although with different time courses, the microcirculation: retinopathy is one of the consequences of such a circulation damage. The retina and retinal vessels are very sensitive to changes in the microvascular circulation, and it has been demonstrated that single features of hypertensive retinopathy have a strong prognostic value for stroke, carotid stiffness, and coronary disease.

At the same time, retinopathy is a social burden, with heavy direct and indirect costs, since visual loss reduces the capacity of working and carrying an independent life.

Even if other organs are affected by microcirculation damage, the retina have the utter advantage over other them of being easily available for non-invasive examination, therefore suggesting a cost-effective way of monitoring the progression of the systemic disease associated with the retinopathy. Detection of retinopathy at its onset is the most critical issue to avoid blindness, particularly in diabetic retinopathy which does not recede with treatment, at the present state of pharmacology.

Unfortunately these first stages are almost asymptomatic. It has been demonstrated that a screening program could save most of the population at risk from developing sight-threatening retinopathy. In the western world there are not enough resources, in terms of time and available experts ophthalmologists, for carrying on an extensive screening. Thus, a reliable automatic tool for evaluating retinopathies is strongly needed.

Every automatic method for evaluating hypertensive and diabetic retinopathies must go through some well defined steps. First, it has to detect the main anatomical structures of the retina: blood vessels, optic disc and possibly the macula. Then it has to evaluate in a quantitative way the ab-
Summary

normalities present in the identified structures. Finally it has to identify any lesions present on the retina.
In this thesis a new algorithm to extract the vascular tree will be described. Then a method for identifying the position of the optic disc will be presented. For an early detection of retinopathy, retinal microvascular abnormalities must be identified and evaluated: algorithms for automatically measuring an extensive set of clinical features have been developed. A method to identify the presence of lesions will be described. Among the possible lesions, microaneurysms characterize the early stages of diabetic retinopathy, whereas other lesions are essential for identifying malignant hypertensive retinopathy and the more severe stages of diabetic retinopathy. Microaneurysms, hemorrhages, exudates and cotton wool spots identification is therefore a fundamental task in the evaluation of retinopathy and its severity. Finally, a first attempt to develop a grading of hypertensive retinopathy will be described. This is a fundamental step to monitor retinopathy progression in time and to compare different retinas in a consistent and quantitative way.

The algorithms presented in this thesis make possible to conceive a tool to be used both for mass-screening of hypertensive and, with small changes, diabetic retinopathies, and also for monitoring the progression of the diseases. Its usefulness will be threefold. Firstly, it shall provide a diagnostic tool to aid the clinical practice. Secondly, it will provide quantitative details of the single damages of the retina, thus providing a tool for clinical research. Finally, it will allow pharmaceutical research to obtain a quantitative and reproducible assessment of disease evolution during pharmacological treatment.
Introduction

Hypertensive retinopathy is associated with systemic arterial hypertension; retinal vascular changes that may occur are seen in both chronic and acute stages. It is likely that increasingly more patients will present hypertension, that is ranked as one of the top 10 risk factors for burden of diseases in developed countries by the World Health Organization. Ophthalmologists are in a unique position to detect the disease, as well as prevent visual loss from it; a patient with non diagnosed malignant hypertension will probably consult first an ophthalmologist with a complaint of visual loss, that is ultimately related to hypertension.

An even more dramatic situation characterizes diabetes related retinopathy. Diabetes is a growing epidemic in the world, due to population growth, aging, urbanization and increasing prevalence of obesity and physical inactivity: population with diabetes is estimated to grow by the 37 percent by 2030, and today it is already around 200 millions people worldwide. Following the trend of diabetes, diabetic retinopathy has an ever increasing importance as a cause of blindness: in the United States it is the first cause of blindness in people in working age, with all the consequent economic and social burdens. The timely diagnosis and referral for management of diabetic retinopathy can prevent 98% of visual loss. It is estimated that the underlying cause of blindness in the majority of diabetic patients is not diabetic retinopathy per se but the misdiagnosis of diabetic retinopathy. Currently, a periodic dilated direct ophthalmoscopic examination seems the best approach for a screening with near universal coverage of the population at risk, despite the proved low sensitivity of direct ophthalmoscopy [74, 86, 38]. However, the number of ophthalmologist available is the limiting factor in initiating an ophthalmologist based screening [90].
1. Introduction

With the increasing availability of digital fundus camera, there is a wide consensus opinion that automatic analysis of such digital images might relieve, at least partially, the burden of retinopathy screening from the ophthalmologists.

1.1 Aim and Objectives

The aim of the work presented in this thesis is to develop a set of tools for fully automatic analysis of retinal images. These tools would not only evaluate the general pathological state of the retina, but also grade the severity of single lesions and signs of retinopathy. In this way, a tool developed for, say, hypertensive retinopathy, would be easily customized to grade diabetic retinopathy by simply not considering some of the graded signs for the hypertensive, and by differently weighting the importance of those considered. With this in mind, abnormalities and lesions identification and evaluation will be useful for grading different retinopathies (say diabetic), even if for sake of clarity and brevity the thesis will focus on hypertensive retinopathy grading.

A number of abnormalities and lesions have to be identified and evaluated to obtain a grading of the hypertensive retinopathy, and the object of each module of the developed tool is the identification and evaluation of the different signs. These features can be roughly divided into two main groups. Those related to vascular morphological abnormalities, and those that are related to the alteration of the blood-retinal barrier or to more severe vascular damage, such as hemorrhages and ischemic events.

Therefore, in order to provide the retinopathy grading, a number of different steps must be performed, and the results of each step evaluated and validated.

First of all, as result of the acquisition procedure of the retinal image, the image often shows a marked illumination heterogeneity, accompanied by contrast variations within the same image. These illumination and contrast variabilities have to be corrected in order to provide a tool equally sensitive in every region of the image.

The second step deals with the identification of the retinal vessel network, and with its description in terms of position and vessel diameters. This is a prerequisite both for the analysis of the vascular tree morphology and its abnormalities, and also for the recognition of the optic disc and for the elimination of the vascular structures from the search of possible non-vascular lesions.
1. Introduction

The next step is the identification of the position of the optic disk, which provides the main landmark of the retinal coordinates. The area of the papilla has also to be excluded from the retinal regions searched for lesions.

A number of abnormalities of the retinal vessels have to be identified and graded to provide an evaluation of the mild and moderate hypertensive retinopathy. Most of these abnormalities take different forms, and have different time course if they are present on arteries or veins. The two vessel networks must thus be distinguished. Crossings and bifurcations have to be identified and analyzed, and the vessel tree morphology evaluated.

Then, other lesions have to be identified and classified in the part of the retina not covered by the previously identified anatomical structures: hemorrhages and microaneurysms, hard exudates, cotton wool spots and possibly drusen.

The last step addresses the grading of retinopathy severity. The quantitative evaluation of the single abnormalities and lesions have to be combined first in a image wide-index for each, and then together in a retinopathy severity grading.

1.2 Outline of the Thesis

Chap. 2-4 are introductory chapters describing retinal imaging, image acquisition protocols and the experimental setting of this thesis. The fundus camera examination, the appearance of the retina in a fundus image and the main findings of hypertensive and diabetic retinopathy will be described in Chap. 2.

In Chap. 3 the available imaging protocols for retinopathy evaluation will be reviewed, and the one used in acquiring the images used in this thesis described.

The images have been evaluated by experienced ophthalmologist both with regard to the gold standard classification grading for hypertensive retinopathy, and also with a more detailed scheme with grading for the single lesions and abnormalities. This will be described in Chap. 4.

The illumination correction procedure developed is presented in Chap. 5.

The extraction of the vessel structure is the object of Chap. 6, and the identification of the optic disc position of Chap. 7. In Chap. 8 a set of
1. Introduction

methodologies for the analysis of the identified vascular structure will be described, and in Chap. 9 the identification of non vascular lesion will be presented. Finally, in Chap. 10 a preliminary retinopathy grading scheme will be presented, and its changes to evaluate diabetic retinopathy. A brief discussion, summarizing the results presented in each chapter, with lines for further development will be contained in 11
2

Fundus Imaging and its Findings

In this chapter a brief review will be presented about what is seen in an image from a fundus camera examination and all the most relevant lesions to be found in the hypertensive and diabetic retinopathy.

2.1 Fundus Oculi Examination

The first instrument that made available to ophthalmologists the direct examination of the retina was the direct ophthalmoscope, which is still used today. It was first described by Helmholtz at the end of the XIX century, and since then it has not changed much. In its basic form is composed by a light source and a set of lenses. The light is projected through the dilated pupil onto the retina, and the lenses focus on so that the observer can look at the retina. Its use is widespread in the clinical practice, but it has been proved to provide poor sensitivity and results highly dependent on the observer experience.

In the middle of the XX century the first instrument able to acquire photographs of the retina appeared. This is a photographic 35mm back connected to an optic system that focuses on the fundus oculi, illuminated by a coaxial flash. This fundus camera enables the photography of different portions of the retina with different magnification, which ranges from $10^\circ$ to $60^\circ$.

Around 1990, the first digital fundus camera appeared. The optic system is not connected anymore to a traditional camera, but to a CCD, and the image is sent to a computer for visualization and storage.
2. Fundus Imaging and its Findings

2.1.1 Fundus Oculi Appearance

Figure 2.1: An image of a normal fundus oculi. Papilla (a), fovea (b) and vessel networks are clearly visible

Using a fundus camera, an image of the fundus oculi is acquired. The visible part of it is composed by the retina with its vascular network and the optic nerve head. The choroid is the structure below the retina and its usually obscured by it.

The retina is a multilayer structure, which is transparent except for the deepest layer, the pigmented epithelium. This gives to the retina its reddish colour. More superficially than the pigmented epithelium there is the sensorial retina, composed by the photoreceptor cells and by the gangliar cells. The axons of the gangliar cells runs to the papilla, or optic disc, or optic nerve head, which is the place where the bundle of nervous fibers forms the optic nerve, and leaves the optic bulb. From the center of the optic disc the ophthalmic artery enters into the optic bulb, and subsequently branches to provide vascularization to most of the retina. From the capillary network originates the venous vessels, which flow into the central retinal vein that exit the ocular bulb through the optic disc.

Topologically, the temporal vessel arcades delimit the posterior pole. At the center of the posterior pole there is the macula: its center is occupied by a small depression, the fovea, that is the region most densely packed with photoreceptor of the retina and is normally the center of vision. The macula is not fed by retinal vessels, but takes its nutrients from the choroid vessels below the retina.
2. Fundus Imaging and its Findings

Choroidal vessels are not usually visible in an image taken with a fundus camera, but if the pigmented epithelium is very lightly pigmented or in case of pathological depigmentation, the retina becomes almost transparent and the choroid becomes visible.

2.2 Main Vascular Abnormalities

2.2.1 Tortuosity

![Normal Vessel Course](image1.png) ![Tortuous Vessels](image2.png)

(a) Normal Vessel Course  (b) Tortuous Vessels

Figure 2.2:

In presence of high blood pressure, vessels may increase in length and vessel walls thicken, and as a result they become increasingly tortuous. This is at first seen in arteries, and only in more severe stages of retinopathy, also in veins.

2.2.2 Generalized Arteriolar Narrowing

The earliest fundus change due to hypertension is the thinning of the retinal arterioles. Narrowing of the arterioles is usually proportional to the degree of elevation of blood pressure. However, retinal arteriolar narrowing is imprecisely quantified from a clinical ophthalmoscopic examination, since the examiner should estimate the normal vessel width prior to the narrowing to evaluate severity of the latter.
2.2.3  Focal Arteriolar Narrowing

Figure 2.3: A definite focal narrowing

In severe hypertension states, irregularities in the caliber of blood vessels may appear. In arterioles, they are due to localized spasm and contraction of the wall. They appear as a focal thinning of the blood column: the narrowing may increase until the vessels become thread-like.

2.2.4  Bifurcations Abnormalities

Arterial diameters and topography at branch points are believed to conform to design principles that optimize circulatory efficiency and maintain constant shear stress across the network [65]. It has been suggested that arterial diameters at a bifurcation should conform to a power relationship, and arterial branches in various circulation have been shown to obey to this design. It has been shown that bifurcation angles are reduced with increasing hypertension, probably because the atheroma fibrosis of the central artery displace by contraction the arteries toward the disk. Although the mechanisms of bifurcation changes are not clear, both branching angles and also the value of the junction exponent seems to deviate from its optimal values with age [83]

2.2.5  Crossing Abnormalities

The abnormal changes in arteriovenous crossings result from the thickening of the wall of the arterioles due to hypertension and sclerosis, and associated changes in the veins at the crossings. The first appearance of crossing abnormalities is the compression of the vein by the artery, which may vary in
2. Fundus Imaging and its Findings

(a) Gunn’s sign (Vein compression)  (b) Salus’s sign (Vein deflection)

severity from a slight indentation to complete interruption of the vein where the artery crosses. When the sclerotic process in the artery extends to the adventitia of the vein, the blood column in the vein will be partially obscured and appear tapered on each side of the crossing.

Constriction and compression of the veins may impede the blood return, so that the veins become distended for some distance peripheral to the crossing: this is the so called Gunns sign.

The arterial sclerosis may cause deflection of the vein from its normal course at the point where the artery crosses. The vein may deflect both vertically (dipping under the artery or humping over it), or laterally. In this last case, instead of crossing the artery obliquely, the vein does so at right angles and appears as S-shaped at the bend, which has been referred as the Salus sign.

2.3 Main Non Vascular Findings

2.3.1 Microaneurysms and Red Dots

Retinal microaneurysms are the most characteristic lesion of diabetic retinopathy, but are present also in other pathologies that affect the microvessels. Micoraneurysms are a small dilation of a capillary wall. It is not clear if retinal microaneurysms are due to a vessel wall damage or to the beginning of a neovascularization. However, the result is the appearance of small saccular structures, of approximate dimension between 10µm and 100µm, that in the retinal fluorescein angiography appear as bright hyperfluorescent
2. Fundus Imaging and its Findings

Figure 2.5:

(a) Large microaneurysm with central reflex
(b) Microaneurysms

spots, whereas in colour fundus images appears as round, red spots. They are indistinguishable from small hemorrhages of the same dimension, since they both are small round regions, with a dark red colour. Therefore, both microaneurysms and hemorrhages smaller than the major vein caliber at the optic disc margin (usually 125 \( \mu m \)), are considered red dots, and evaluated as microaneurysms [21]. On the contrary, any red spot greater than that is considered an hemorrhage, unless features as round shape, smooth margins and a central light reflex suggest that it is probably a microaneurysm.

2.3.2 Hemorrhages

Retinal hemorrhages are blood deposits on the retina. Hemorrhages disappear as the blood is reabsorbed with time. They are due to the breaking of a vessel wall or of a microaneurysm, and the increase in their presence is a clear sign of diffuse retinal damage. They have very different shapes, going from the round red spot with sharp margins, to the blot hemorrhage, to the flame-shaped hemorrhage. As the blood is reabsorbed, hemorrhage margins fade and the characteristic red colour turns to a faint greyish-red before disappearing completely.

2.3.3 Hard Exudates

Hard exudates are small lipidic and proteinic deposits, which appear as white or yellowish-white areas with sharp margins. They may be arranged as individual dots, confluent patches or in partial or complete rings surrounding microaneurysms or zones of retinal edema. In the more severe cases of hypertensive retinopathy, they appear as a confluent ring around the macula (the macular star).
2. Fundus Imaging and its Findings

(a) Small hemorrhage
(b) Barely visible hemorrhage
(c) Flame hemorrhage
(d) Large hemorrhage
(e) Bright hemorrhage

Figure 2.6:

2.3.4 Cotton Wool Spots

Cotton wool spots are the consequence of retinal ischemic events, due to precapillary arterioles stenosis. This causes a swelling of the nerve fiber layer, with local deposit of cytoplasmatic material. They are round or oval in shape, white, pale yellow-white or greyish-white, with soft and feathery edges, that give their characteristic aspect and their name. They usually appear along the major vessel arcades, parallel to the nerve fibers, and are sometimes accompanied by the presence of microaneurysms.
2. Fundus Imaging and its Findings

2.3.5 Drusen

Drusen are deposits associated with thinning or hypopigmentation of the retinal pigment epithelium. They appear as deep, yellowish-white dots. To distinguish drusen from hard exudates, good stereoscopic view would be necessary, since drusen appear very deep while hard exudates are slightly more superficial. In the protocol used in this thesis the photographs are mono, therefore it is not easy to identify hard exudates from drusen. Several other features are used in distinguishing drusen from hard exudates. Drusen are usually scattered diffusely or scattered near the center of the macula. They are usually round in shape, while hard exudates are usually irregular in shape. Finally, drusen have often a faint border of pigment.

2.4 Hypertensive Retinopathy Grading

The classification of hypertensive changes in the retina in a severity scale was first proposed by Keith [?], in what is now currently known as the Keith-Wegener-Barker grading system. It was subsequently modified by Scheie [?] to better separate hypertensive from atherosclerotic abnormalities. In Tab. 2.1 the two classifications for hypertensive retinopathy are shown. It is
worth noting that recent literature challenges the prognostic significance of these classifications. The poor correlation with the severity of hypertension variation in the onset and progression of the clinical signs, has suggested the use of a classification of retinopathy into two grades: non-malignant and malignant [15]. This is further confirmed by the fact that density of perifoveal capillaries and capillary blood flow velocity analysed with an angiographic examination, correlate more with a two grade rather than with the classical four grade classification system. Nevertheless, the Keith-Wegener-Barker is still the standard *de facto* in the evaluation of hypertensive retinopathy.

<table>
<thead>
<tr>
<th>Keith-Wegener-Barker</th>
<th>Scheie</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Grade</strong></td>
<td><strong>Features</strong></td>
</tr>
<tr>
<td>I</td>
<td>Mild generalised retinal arteriolar narrowing. Increased arterial tortuosity</td>
</tr>
<tr>
<td>II</td>
<td>Definite focal narrowing and arteriovenous crossing abnormalities</td>
</tr>
<tr>
<td>III</td>
<td>The above and retinal hemorrhages, exudates and cotton wool spots</td>
</tr>
<tr>
<td>IV</td>
<td>Severe grade III plus papillar oedema</td>
</tr>
</tbody>
</table>

Table 2.1: Classification of hypertensive retinopathy as proposed in [45] and [76]

## 2.5 Diabetic Retinopathy

Two landmark clinical trials set the standard in grading diabetic retinopathy. They are the Diabetic Retinopathy Study (DRS) [88] and the Early Treatment Diabetic Retinopathy Study (ETDRS) [22]. The ETDRS severity scale was based on the Airlie House classification of diabetic retinopathy and is used to grade fundus photographs. It has been widely applied in research settings, publications and it has shown satisfactory reproducibility and validity. Although it is recognized as the *gold standard* for grading the severity of diabetic retinopathy in clinical trials, its use in everyday clinical practice has
2. Fundus Imaging and its Findings

not proven easy or practical. The first reason for this is that the photographic grading system has 90 levels, many more than what is necessary for clinical care. Given the number of levels to consider and the detailed specific definitions of the levels, and the requirement of comparison with standard photographs, it is not surprising that ETDRS grading procedure is difficult to remember and apply in a clinical setting.

Recently, simplified severity scales have been developed in an effort to improve both the screening of patient with diabetes and communication among caregivers. Yet, to overcome this proliferation of \textit{ad hoc} grading scales, it has been proposed in [96] a Diabetic Retinopathy Disease Severity Scale, in which separate scales were proposed to grade diabetic retinopathy (4 levels) and macular oedema (5 levels). The two scales are summarized in in Tab. 2.2 and Tab. 2.3.


## 2. Fundus Imaging and its Findings

<table>
<thead>
<tr>
<th>Disease Severity Level proposed in [96]</th>
<th>Findings Observable on Dilated Ophthalmoscopy</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Apparent Retinopathy</td>
<td>No abnormalities</td>
</tr>
<tr>
<td>Mild nonproliferative diabetic retinopathy</td>
<td>Microaneurysms only</td>
</tr>
<tr>
<td>Moderate nonproliferative diabetic retinopathy</td>
<td>More than just microaneurysms but less than severe nonproliferative diabetic retinopathy</td>
</tr>
<tr>
<td>Severe nonproliferative diabetic retinopathy</td>
<td>Any of the following: more than 20 intraretinal hemorrhages in each of 4 quadrants; definite venous beading in 2 or more quadrants; prominent intraretinal microvascular abnormalities in one or more quadrant and no signs of proliferative diabetic retinopathy</td>
</tr>
<tr>
<td>Proliferative diabetic retinopathy</td>
<td>One or more of the following: neovascularization, vitreous or preretinal hemorrhage</td>
</tr>
</tbody>
</table>

Table 2.2: Classification of diabetic retinopathy as proposed in [96]
2. Fundus Imaging and its Findings

<table>
<thead>
<tr>
<th>Disease Severity Level proposed in [96]</th>
<th>Findings Observable on Dilated Ophthalmoscopy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diabetic macular oedema apparently absent</td>
<td>No apparent retinal thickening or hard exudates in posterior pole</td>
</tr>
<tr>
<td>Mild diabetic macular oedema</td>
<td>Some retinal thickening or hard exudates in the posterior pole but distant from the center of the macula</td>
</tr>
<tr>
<td>Moderate diabetic macular oedema</td>
<td>Retinal thickening or hard exudates approaching the center of the macula but not involving the center</td>
</tr>
<tr>
<td>Severe diabetic macular oedema</td>
<td>Retinal thickening or hard exudates involving the center of the macula</td>
</tr>
</tbody>
</table>

Table 2.3: Classification of diabetic macular oedema proposed in [96]. Hard exudates are a sign of current or previous macular oedema. Diabetic macular oedema is defined as retinal thickening and requires a three-dimensional assessment.
Image Acquisition Protocols and Clinical Evaluation

Since the development of the fundus camera, the ophthalmoscopic examination was the standard procedure for evaluating the state of the retina. In the last 15 years, as the fundus camera took ground in the ophthalmologic practice, issues were raised about the sensitivity and specificity of fundus photographs, and about which acquisition protocols may assure the highest sensitivity and specificity for the early identification of sight-threatening diseases. The gold standard for fundus imaging is the ETDRS protocol. Its practical usefulness is reduced by the complexity of acquisition procedure for the camera technician, by the complexity of images evaluation for the ophthalmologist, and by the discomfort for the patient involved in the procedure. It is therefore not commonly used in the clinical practice, but mainly limited to large research studies. As fundus camera quality increases, a number of protocols simpler than ETDRS have been proposed and validated against that gold standard. Even if none has taken solid ground in the clinical practice, the future for the widespread utilization of fundus imaging will be wide angle, few fields photographs protocols. In this chapter a number of protocol proposed in the literature will be presented, and in the end the protocol for the acquisition of the photographs used in the current thesis.

3.1 Diabetic Retinopathy

The grading of Early Treatment Diabetic Retinopathy Study (ETDRS) seven standard field 35-mm stereoscopic color fundus photographs (ETDRS photos) using the modified Airlie House classification is the current gold standard for determining severity of diabetic retinopathy. The Airlie House Classifi-
3. Image Acquisition Protocols and Clinical Evaluation

cation of Diabetic Retinopathy provides the photographic basis for demonstrating the clinical characteristics and extent of clinically pertinent lesions of diabetic retinopathy. Early Treatment Diabetic Retinopathy Study 35-mm retinal photography and grading protocols provide an established and documented sensitivity for detecting and assessing severity of diabetic retinopathy. Compared with other retinal imaging methods, such as Polaroid photography or digital-video imaging, 35-mm slide retinal color photographs offer advantages of a large existing database, lower equipment cost, better resolution, and higher color fidelity. Disadvantages associated with 35-mm photography include the requirements for skilled photographers, the need for pupil dilation, uncomfortable examination sessions for the patients, higher costs for film and film handling, delays in film processing, and inefficient archiving with potential for loss or damage of slides. These disadvantages can impact the efficiency, convenience, and cost effectiveness of the procedure and can reduce compliance of patients with annual eye examinations for retinopathy assessment. A number of investigators have evaluated alternatives to ETDRS photos for retinal imaging and assessment of diabetic retinopathy severity.

3.1.1 ETDRS Protocol

The Early Treatment Diabetic Retinopathy Study is built over the Diabetic Retinopathy Study and uses the same protocol of image acquisition. Seven fields, mydriatic, stereoscopic photographs are taken on 35-mm film and subsequently evaluated on light-box. Both for image acquisition and retinopathy grading certified personnel is required. The seven fields are sketched in Fig. 3.1, and described in details in Tab. 3.1.

The grading procedure is based on a large set of reference photographs (the standard photographs) for each lesion or abnormality considered. In [22, 21] the complex procedure for classifying the features identified in the set of 7 stereo photographs is reported, together with the gradings to provide the level of retinopathy of the eye under examination.

3.1.2 EURODIAB Protocol

A wide angle retinal photography protocol was developed within the framework of the EURODIAB IDDM Complications Study [2], part of a European Community funded Concerted Action Programme into the epidemiology and prevention of diabetes (EURODIAB). Two 45° colour photographs of each eye are taken. One is centered on the macula, so that the exact centre of the optic disc lay at the nasal end of the horizontal meridian of the field of
view. The second is the nasal field, such that the optic disc was positioned one disc diameter from the temporal edge of the field, on the horizontal meridian of the field of view. Considering the partial overlap of the two fields, they provide a retinal view of approximately 80° horizontally by 45° vertically, that the authors judge as sufficient for detecting clinically significant or sight-threatening lesions of diabetic retinopathy. In order to test this, they compared the protocol with the recognised gold standard 7-field 30 degrees stereo photography (assessed using a modified Airlie House classification scheme). It was found that occasionally lesions occurred outside the field of view of either one or the other of the protocol [2]. Simple presence of retinal lesions was correctly detected by both systems in 43 of the 48 eyes, giving 100% agreement on detection. Both systems correctly identified the two known cases of confounding vein occlusion. In eyes with diabetic retinopathy (n = 41), when severity was expressed in three groups: mild background, moderate/severe background and proliferative/photocoagulated, at least one grader (out of five) using the new system matched the verified results in 38 out of 31 (93%) eyes and three or more graders matched in 31 (76%) eyes. In the view of the authors, the simplicity of application of the system compared to the ETDRS 7-field should prove useful especially in large clinical trials when consistently high quality ETDRS 7-field stereo images would be difficult to achieve.
3. Image Acquisition Protocols and Clinical Evaluation

3.1.3 Joslin Clinic Protocol

The Joslin Vision Network (JVN) is telemedicine platform designed to facilitate access of patients with diabetes into the chronic disease management program of diabetes within the Joslin Diabetes Eye Health Care Model. This system incorporates a commercially available nonmydriatic retinal fundus camera optimized for low-light level imaging of the retina. The protocol [10], developed with the aim of reducing patient discomfort and of providing an easy picture taking by non-certified photographers, is a three field nonmydriatic 45° photographs. These field are described in Tab. 3.2, and they are represented in Fig. 3.2 superimposed to the ETDRS protocol. This protocol was validated against the ETDRS 30° 7-field stereo. It was demonstrated that the determination of clinical level of diabetic retinopathy using the JVN stereoscopic nonmydriatic digital-video color retinal images from three distinct retinal regions obtained using a 45° nonmydriatic camera optimized for low-light level imaging (JVN images) is in substantial agreement (k =0.65) with dilated ETDRS seven standard field 35-mm stereoscopic 30° retinal photography (ETDRS photos).
3.1.4 Single Field Monochromatic

In [54], a single wide angle field, monochromatic was tested against ophthalmoscopy conducted by an experienced ophthalmologist and against the ETDRS 7-field stereo. The single field of the 45° photograph was centered on a point halfway between the temporal edge of the optic disk and the fovea, and included areas of the retina on either side of both structures (Fig. 3.3). Taking the ETDRS retinopathy level 35 as threshold for referral, digital imaging had a sensitivity of 78% and specificity of 86% when compared with standard seven-field color photography.

3.1.5 5-Fields Protocol

In [61] a non-mydriatic digital fundus photography without pupillary dilation, using a non-mydriatic 45 Topcon TRCNW6 fundus camera (Topcon Europe, Rotterdam, The Netherlands), was tested for sensitivity and specificity against the seven ETDRS mydriatic stereo standard fields. in the Early Treatment Diabetic Retinopathy Study (ETDRS) for DR screening. Five 45° non-stereoscopic images of five overlapping fields were taken for each eye: one image was centered on the macula, including the optic disc, and one each on the nasal, temporal, superior and inferior fields. The sensitivity of

Figure 3.3: Non Mydriatic Monochromatic Field proposed in [54] (black circle), and standard ETDRS 7 fields (dotted blue)
3. Image Acquisition Protocols and Clinical Evaluation

detection for moderately severe to severe forms of DR ranged from 92% to 100% and the specificity from 85% to 88%.

3.1.6 Single-Field Colour Protocol

In [74] four screening methods (an exam by an ophthalmologist through dilated pupils using direct and indirect ophthalmoscopy, an exam by a physicians assistant through dilated pupils using direct ophthalmoscopy, a single 45 degrees retinal photograph without pharmacological dilation, and a set of three dilated 45 degrees retinal photographs) were compared with a reference standard of stereoscopic 30 degrees retinal photographs of seven standard fields read by a central reading center. Sensitivity, specificity, and positive and negative likelihood ratios were calculated after dichotomizing the retinopathy levels into none and mild nonproliferative versus moderate to severe nonproliferative and proliferative. The sensitivities, specificities, and positive and negative likelihood ratios are summarized in Tab. 3.3

In [38], diabetic patients referred for screening were studied in a prospective fashion. A single 45 degrees fundus image was obtained using the non-mydriatic digital camera, and the validation was performed using as ground truth the diabetic retinopathy grading by a consultant ophthalmologist. The sensitivity for detection of any diabetic retinopathy was 38% and the specificity 95%.

3.2 Hypertensive Retinopathy

3.2.1 ARIC Protocol

The Atherosclerosis Risk in Communities (ARIC) Study is an epidemiological research study of the major factors contributing to the occurrence and trend of cardiovascular disease in middle-aged adults in the United States. It has two main objectives. Firstly, to investigate factors associated with both atherosclerosis and incidence of clinical cardiovascular disease, and then to measure coronary heart disease occurrence and trends and relate them to community levels of risk factors, medical care and atherosclerosis. Fundus photographs were used to evaluate changes in the retinal vasculature (presumed to be related to hypertension and/or arteriolar sclerosis) that may be prognostic for various cardiovascular outcomes, but other significant retinal conditions will be noted, such as diabetic retinopathy or vascular occlusions. Within the ARIC study, the protocol consists of one non-mydriatic
3. Image Acquisition Protocols and Clinical Evaluation

45° retinal photograph. The photographs are subsequently sent to the ARIC Retinal Reading Center for assessment of retinal status [3, 4]. To obtain consistent field specification even when non experienced technician take the photographs, the camera used in the ARIC study is provided with a mask on which to align the optic disc. These aligning masks are provided by the Retinal Reading Center and, when, properly attached to the monitor, they position the optic nerve centered from top to bottom and the nasal edge of the optic nerve falls between 1.5-2 optic disc diameters from the nasal edge of the photograph.

3.3 Used Protocol for Hypertensive Retinopathy

In the clinical routine, when checking the retina for retinopathies, it is common to find wide angle photographs, digital for the newest systems or film for the older, with a varying number of fields, depending on the whim of the taker, and if he/she judges that significant lesions are present in non-standard fields. Nevertheless, the most usual field used as starting point is the one centered on the macula: this is the most important area to check for sight-threatening lesions. Combining a macular field with wide angle photograph, it allows to evaluate most of the vascular arcades for abnormalities, and to have available for examination both the macula and the optic disc. Even if photographs are non-stereo, optic disc and macula appearances can suggest and sometimes clearly point out the presence of swellings and oedema.

The acquisition protocol used in this thesis is a single field photographs, non stereo, mydriatic, with a 50° angle and centered on the macula.

3.4 Conclusions

Several studies (see also [97] for an up-to-date bibliography) showed that, as a tool to detect vision-threatening retinopathy, single-field fundus photography interpreted by trained readers has sensitivity ranging from 61% to 90% and specificity ranging from 85% to 97% when compared with the gold standard reference of stereo-photographs of 7 standard fields. When compared with dilated ophthalmoscopy by an ophthalmologist, single-field fundus photography has sensitivity ranging from 38% to 100% and specificity ranging from 75% to 100%. Therefore, although single-field fundus photography is not a substitute for a comprehensive ophthalmic examination, it can serve as a screening tool retinopathies to identify patients for referral for ophthalmic
The advantages of single-field fundus photography interpreted by trained readers are ease of use (only one photograph is required), convenience, and ability to detect retinopathy. With this in mind, and with the consideration of the widespread utilization of wide angle retinal photography as examination procedure in the clinical practice, it was chosen as protocol for retinal imaging in this thesis a 50° single field centered on the macula.
3. Image Acquisition Protocols and Clinical Evaluation

<table>
<thead>
<tr>
<th>ETDRS Field</th>
<th>Field Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>F-1</td>
<td>Optic Disc: 30° field focused centrally on the optic disc</td>
</tr>
<tr>
<td>F-2</td>
<td>Macula: 30° field focused on the center of the macula</td>
</tr>
<tr>
<td>F-3</td>
<td>F-3 Temporal to Macula: 30° field focused so the nasal edge of the field crosses the center of the macula</td>
</tr>
<tr>
<td>F-4</td>
<td>F-4 Superior temporal: 30° field focused so the lower edge of the field is tangent to a horizontal line passing through the upper edge of the optic disc and the nasal edge of the field is tangent to a vertical line passing through the center of the disc</td>
</tr>
<tr>
<td>F-5</td>
<td>F-5 Inferior temporal: 30° field focused so the upper edge of the field is tangent to a horizontal line passing through the lower edge of the optic disc and the nasal edge of the field is tangent to a vertical line passing through the center of the disc</td>
</tr>
<tr>
<td>F-6</td>
<td>F-6 Superior nasal: 30° field focused so the lower edge of the field is tangent to a horizontal line passing through the upper edge of the optic disc, and the temporal edge of the field is tangent to a vertical line passing through the center of the disc</td>
</tr>
<tr>
<td>F-7</td>
<td>F-7 Inferior nasal: 30° field focused so the upper edge of the field is tangent to a horizontal line passing through the lower edge of the optic disc, and the temporal edge of the field is tangent to a vertical line passing through the center of the disc</td>
</tr>
<tr>
<td>F-8</td>
<td>30° field focused outside the seven standard fields</td>
</tr>
</tbody>
</table>

Table 3.1: Early Treatment Diabetic Retinopathy Study Seven Standard Field Definitions, from [10, 22, 21]
3. Image Acquisition Protocols and Clinical Evaluation

<table>
<thead>
<tr>
<th>JVN Field</th>
<th>Field Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>NM-1</td>
<td>45° field focused centrally between the temporal margin of optic disc and the center of the macula: Center the camera on the papil-lomacular bundle midway between the temporal margin of the optic disc and the center of the macula. The horizontal centerline of the image should pass directly through the center of the disc. A stereoscopic image is obtained by capturing one image through the left aspect of the pupil opening, shifting the camera laterally, and then capturing a second image through the right aspect of the pupil. A slight delay between the first and second image may be necessary to allow for adequate pupil mydriasis.</td>
</tr>
<tr>
<td>NM-2</td>
<td>45° field focused superior temporal to the optic disc: Center the camera laterally approximately one-half disc diameter temporal to the center of the macula. The lower edge of the field is tangent to a horizontal line passing through the upper edge of the optic disc. This image is taken temporal to the macula but includes more retina nasal and superior to the macula than standard Field 2.</td>
</tr>
<tr>
<td>NM-3</td>
<td>45° field focused nasal to the optic disc: This field is nasal to the optic disc and may include part of the optic disc. The horizontal centerline of the image should pass tangent to the lower edge of the optic disc.</td>
</tr>
<tr>
<td>NM-4</td>
<td>45° field focused temporal to the macula (obtained through dilated pupils only)</td>
</tr>
<tr>
<td>NM-5</td>
<td>45° any optional field focused beyond the definitions of NM fields 1-4</td>
</tr>
</tbody>
</table>

Table 3.2: Joslin Vision Network Non-mydriatic Retinal Fields
## 3. Image Acquisition Protocols and Clinical Evaluation

Table 3.3: Results obtained in [74] for four screening protocols

<table>
<thead>
<tr>
<th></th>
<th>Ophthalmologist</th>
<th>Physician Assistant</th>
<th>Single non-mydriatic photograph</th>
<th>Three mydriatic photographs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
<td>0.33</td>
<td>0.14</td>
<td>0.61</td>
<td>0.81</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.99</td>
<td>0.99</td>
<td>0.85</td>
<td>0.97</td>
</tr>
<tr>
<td>Positive Ratio</td>
<td>72</td>
<td>12</td>
<td>4.1</td>
<td>54</td>
</tr>
<tr>
<td>Negative</td>
<td>0.67</td>
<td>0.87</td>
<td>0.46</td>
<td>0.19</td>
</tr>
</tbody>
</table>
3. Image Acquisition Protocols and Clinical Evaluation
Clinical Evaluation

The images used in this thesis are acquired following the protocol described in Sec. 3.3. The photographs were acquired at the Ophthalmology Clinics of the Universities of Padova, Udine and Trieste by experienced technicians. The films were subsequently sent to the Department of Information Engineering of the University of Padova, where they were digitized at 1360 dpi, 24 bits per pixel, using a Canon scanner. Between May 2001 and January 2002, XXX retinal images of various quality and retinopathy level were digitized. An expert ophthalmologist chose 60 images as training set, creating what will be called the DB60. He chose other 200 images, the DB200 to be subsequently used as validation set.

4.1 Image Collection and Database Building

To provide a ground truth evaluation of the images chosen in the DB60 and in the DB200, a standard report form has been set up, and it is shown in Fig. ?? and Fig. ??.

The first page of the report contains all the information relevant in the evaluation of the retinopathy. The second page require the ophthalmologist to sketch the approximate position of individual lesion (crossing abnormalities, focal arteriolar narrowing, non-vascular lesions), on the scheme. The standard clinical evaluation for hypertensive retinopathy is the Keith-Wegener-Barker grading scheme [45], or that proposed by Scheie [76]. The single abnormality is evaluated usually as none, mild, moderate and severe. These qualitative evaluations provide a very rough assessment of the pathology and of its constituting damages. In order to overcome this limitation, the ophthalmologist were required to fill in for every abnormality a grading in a...
4. Clinical Evaluation

percent scale, with percent intervals corresponding roughly to the none-mildmoderate-severe grading. The images in the DB60 were evaluated by two experienced ophthalmologists of the Ophthalmology Department of the University of Padova. The images in the DB200 were evaluated by one experienced ophthalmologist of the Ophthalmology Department of the University of Padova.
4. Clinical Evaluation

Univerisita’ degli Studi di Padova - Dipartimento di Elettronica e Informatica  
Biomedical Image Processing Group

<table>
<thead>
<tr>
<th>Feature</th>
<th>N</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arterial Tortuosity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Venous Tortuosity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Focal Narrowing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Generalized Narrowing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gunn/Salus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cotton Wool Spots</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hard Exudates</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hemorrhages</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Microaneurysms</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Venous Congestion</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drusen</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>K-W-B Grading</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image Grading (in %)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Contribution to image grading</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arterial Tortuosity</td>
</tr>
<tr>
<td>Venous Tortuosity</td>
</tr>
<tr>
<td>Focal Narrowing</td>
</tr>
<tr>
<td>Generalized Narrowing</td>
</tr>
<tr>
<td>Gunn/Salus</td>
</tr>
<tr>
<td>Cotton Wool Spots</td>
</tr>
<tr>
<td>Hard Exudates</td>
</tr>
<tr>
<td>Hemorrhages</td>
</tr>
<tr>
<td>Microaneurysms</td>
</tr>
</tbody>
</table>

Comments:
4. Clinical Evaluation

<table>
<thead>
<tr>
<th>VASCULAR FEATURES</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Vascular Features" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>NON VASCULAR FEATURES</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image2" alt="Non Vascular Features" /></td>
</tr>
</tbody>
</table>
Luminosity and Contrast Normalization

Retinal images are acquired with a fundus camera, which records, on film or digital sensors such as CCD, the illumination light reflected by the retinal surface. Very often these images are unevenly or non-uniformly illuminated and thus local luminosity and contrast variability are present. This problem may seriously affect the diagnostic process and its outcome, since lesions in some areas may become hardly visible to a human observer. Automatic computer-based methods are now often proposed to assist the eye specialist by deriving some parameters from the image, e.g. computing vessel tortuosity or detecting haemorrhages [23]. Images with large luminosity and contrast variability, both intra- and inter-image, are very difficult to analyze with such automatic systems and the obtained results may be of poor quality. In this case, the normalization step is a necessary pre-requisite, aimed at obtaining images with a common standardized value for luminosity and contrast. Several techniques have been used to improve non-uniform luminosity and contrast levels. The classic ones try to normalize image luminosity by disposing of low-frequency luminosity drifts by means of high-pass filtering [34]. More complex, space-variant filtering schemes have been proposed to enhance image appearance, with particular regard to locally adaptive contrast enhancement [91]. Other techniques have been proposed for the specific application to retinal images. Background normalization was applied by using a large median filter to extract slow variations of luminosity, which were then subtracted from the observed image [66]; a brightness adjustment procedure, based on a non-linear point transformation, was also proposed [93]; an adaptive non-linear contrast enhancement was described in [78]. Finally, a different approach was proposed in [95], which exploited the extraction of vessel pixels, the estimation of the illumination function drift in these pixels,
5. Luminosity and Contrast Normalization

and its subtraction from the observed image.
The main reason that calls for the development of a new and different normalization algorithm is the possible presence in the fundus of many features, e.g., optic disc or various types of lesions, whose visibility has to be preserved and hopefully improved by the normalization process. The approaches that estimate the correction from the whole image will fail in distinguishing luminosity variations due to the presence of these features from variations due to changes in illumination. The result will be a generalized smoothing of image luminosity variations. Locally adaptive non-linear filters, even if capable of producing a better local contrast, still decrease the global difference between bright and dark features and do not guarantee the reduction of luminosity variation throughout the image. The method proposed in [95] is a first attempt to overcome these problems, although the choice of vessels as descriptors of luminosity variation has some serious drawbacks. First, vessels are not evenly distributed throughout the image, and the macula region has no vessels at all: this will lead to a very sparse data set to be interpolated to obtain an estimate of the illumination drift. Secondly, there is a large variability in reflectance between arteries and veins, with vessels showing distinctive luminosity patterns [18, 51] that make very difficult to reliably estimate and interpolate luminosity drifts.
The objective of this chapter is to present a new method to normalize both luminosity and contrast in retinal images. The method is based on a model of the observed image, in which the background part of the image provides the information on the luminosity and contrast variability. Normalization is performed both within and between images.

5.1 Review of Available Methods

5.1.1 Low-Pass Filtering

In a typical fundus image, it can be safely assumed that the lower frequencies of its spectrum are related to the illumination variation, whereas the other frequencies carry all the information related to the anatomical structures and to possible presence of lesions.
The illumination variation can be therefore estimated by means of a lowpass filter, and then subtracted from the original image (that is equivalent to high-filtering the image).
5. Luminosity and Contrast Normalization

5.1.2 Homomorphic Filtering

Homomorphic filtering is a generalized technique for image enhancement and correction. It simultaneously normalizes the brightness across an image and increases contrast [34]. The idea behind the homomorphic filter is that an image can be represented as the product of illuminance and reflectance:

\[ f(x, y) = i(x, y) \cdot r(x, y) \]  

(5.1)

Taking the logarithm of the image \( f \) allows to separate the illuminance and reflectance contributions:

\[ \ln(f(x, y)) = \ln(i(x, y)) + \ln(r(x, y)) \]  

(5.2)

Since usually the illuminance part of the signal is characterized by slower variations than the reflectance contribution, a simple low pass filter would eliminate the illuminance component in the log-signal domain. Given a suitable low pass filter \( h(x, y) \), convolving it with the log-signal and exponentiating the result yields:

\[ s(x, y) = e^{h(x,y) \cdot \ln(f(x,y))} \cong e^{\ln(r(x,y))} = r(x, y) \]  

(5.3)

The drawbacks of this method is, as for all filtering scheme, that often, and especially in pathological images, low frequencies are associated with pathological signal of interest, which will be filtered out using these algorithms.

5.1.3 Median Filtering

5.1.4 Wallis Filter

The Wallis filter [91] is an adaptive, space-variant filter, which provides contrast enhancement. It is designed especially for images in which there are significant areas of high and low intensities. In this type of images a global contrast enhancement can not simultaneously produce good local contrast at both end of the image dynamic range. It shall in fact provide contrast enhancement in dark areas while saturating in the bright ones or vice versa. The Wallis filter operates a local histogram matching, setting local means and standard deviations to a user specified value. This produces good local contrast throughout the image, but it reduces the overall contrast between dark and bright areas.
5. Luminosity and Contrast Normalization

5.1.5 Non-linear filtering

In [93] a non-linear filter is proposed to correct the non-uniformity of illumination. A brightness adjustment procedure is described, which enhance only the darker regions. Regions regularly illuminated remain unchanged. This is achieved using a brightness transform function:

\[ y = \beta \cdot x^\alpha \]  

(5.4)

where \( x \) is the pixel intensity value of the original image, \( y \) is the pixel intensity value after the brightness adjustment, \( 0 \leq \alpha \leq 1 \) and \( \beta = x_{\text{max}}^{\alpha-1} \), with \( x_{\text{max}} \) is a saturating limit for the adjustment function.

5.1.6 Local Contrast Enhancement

To increase contrast in retinal images, a non-linear space-variant algorithm has been proposed in [78]. The main idea is to compute mean and standard deviation in a neighborhood of every image pixel, and then applying a local histogram stretch, which takes into account also the image grey-level range. Intensity maximum \( i_{\text{max}} \) and minimum \( i_{\text{min}} \) in the image are evaluated at first. Given a pixel \((x, y)\) with intensity \( i(x, y) \) and an \( M \times M \) neighborhood \( W \) of this pixel, intensity mean \( \mu_W \) and standard deviation \( \sigma_W \) of the pixels within \( W \) are computed. The adaptive contrast enhancement function is:

\[ i'(x, y) = \Psi_W(i(x, y)) - \Psi_W(i_{\text{min}}) \]  

\[ \Psi_W(i) = \frac{\mu_W - i}{\sigma_W} \]  

(5.5)

where \( \Psi() \) is a sigmoidal function:

\[ \Psi_W(i) = \frac{1}{1 + e^{-\frac{\mu_W - i}{\sigma_W}}} \]  

(5.6)

This technique should provide large contrast enhancement where there is an initially small \( \sigma \), and little contrast enhancement for an initially large \( \sigma \). This is done regardless of what is present in the image. It will therefore enhance contrast also in homogenous areas, where it should remain low.

5.1.7 Vessel Sampling and Interpolation

A first method to overcome the drawbacks of image filtering seen so far, was proposed in [95]. The authors identified blood vessels as the key feature,
and they argued that vessels are spread widely over the retina and that their gray level variation throughout the image is proportional to the illumination distribution of the image background. By choosing the vessel pixels as the sample seeds, and assuming that the reflectance of vessels is constant, illumination variation is estimated as the vessel gray level variation. Vessel points are spread on the image, thus the illumination has to be estimated by interpolating the available data points. This is done with a bicubic interpolation. Since vessel points are not evenly distributed on the image, a non-uniform interpolation algorithm is used.

This method suffer of two major drawbacks. Firstly, vessel reflectance is not constant but varies heavily: not only arteries have a remarkable difference in reflectance from veins, but atherosclerotic changes affect as well arterial reflectance. Secondly, vessels are not uniformly spread all over the retina, and there is no certainty of the amount of vessels present. This is highly variable and can lead to a very poor estimate of illumination variability.

### 5.2 A Simple Model of the Observed Image

An observed fundus image $I$ can be modelled as:

$$I = f (I^o) = f (I^o_b + I^o_f)$$

(5.7)

where $I^o$ is the original image, $I^o_b$ is the (original) background image, $I^o_f$ is the (original) foreground image, and function $f (\cdot)$ represents the acquisition transformation.

The background image $I^o_b$ is the ideal image of a retinal fundus free of any vascular structure, optic disc, or visible lesion. The vascular structures, the optic disc and any visible lesion are modelled as an additive term $I^o_f$ to the background image.

It is rather difficult to express properties of $I^o_f$, due to the wide variability of retinal features and lesions that can be found in a fundus image. The only assumption that has to be made regarding $I^o_f$ is that the set of pixels not covered by vascular structures, optic disc or lesions, called the background set $\mathcal{B}$, is not empty. This is a very reasonable assumption, since in all but the more severe pathologies (e.g. central retinal vein occlusion), the retinal background represents the main area of the fundus.

On the other hand, $I^o_b$ can be statistically modelled as:

$$I^o_b (x, y) \sim \mathcal{N} (\mu_b, \sigma_b)$$

(5.8)

i.e., as a white (independence between pixels is assumed) random field with
5. Luminosity and Contrast Normalization

mean value $\mu_b$, representing the ideally uniform luminosity value, and standard deviation $\sigma_b$, representing the natural variability of retinal fundus pigmentation. This model can be further simplified by imposing $\mu_b = 0$ and $\sigma_b = 1$; this latter assumption is acceptable as any bias or amplification can be arbitrarily lumped into the luminosity and contrast drifts introduced by the acquisition function.

The acquisition model $f(\cdot)$ describes the contrast and luminosity distortions introduced by the image observation process. Non-uniform contrast and luminosity within an image can be described as:

$$ I(x, y) = f(I_o(x, y)) = C(x, y)I_o(x, y) + L(x, y) \quad (5.9) $$

where $C(x, y)$ is the contrast drift factor and $L(x, y)$ is the luminosity drift term. Both contrast and luminosity drifts are space-dependent scalar functions and can therefore be considered as images themselves.

The recovery of an estimate $\hat{I}_o$ of original image $I_o$ is based on the estimation of $C$ and $L$ ($\hat{C}$ and $\hat{L}$), and the compensation of the observed image $I$ as:

$$ \hat{I}_o(x, y) = \frac{I(x, y) - \hat{L}(x, y)}{\hat{C}(x, y)} \quad (5.10) $$

Note that the acquisition model just described does not take into account any blurring or additive noise. The goal of this system is not, in fact, the restoration of the image; therefore it implicitly assumes that blurring and noise are already present in the original image $I_o$ it is trying to recover. The resulting normalized image can then be processed by any restoration algorithm.

$C(x, y)$ is assumed to be positive. Both $C(x, y)$ and $L(x, y)$ are assumed to have a spectral content concentrated in the low frequencies, which means that illumination irregularities do not present rapid changes. This is reasonable for regular fundus imaging techniques that adopt diffused light.

Estimation of drift images can be achieved by considering their effects on the background component of observed image. Combining (3) with (1) yields:

$$ I(x, y) = C(x, y)I_o(x, y) + L(x, y) \quad (5.11) $$

$$ = C(x, y)\left[I_o \circ (x, y) + I_f(x, y)\right] + L(x, y) $$

$$ = C(x, y)I_b(x, y) + C(x, y)I_f(x, y) + L(x, y) $$

Restricting the analysis to the background set $\mathcal{B}$, this expression simplifies
5. Luminosity and Contrast Normalization

to:

\[ I(x, y) = C(x, y) I^0_b(x, y) + L(x, y), \quad (x, y) \in \mathcal{B} \]  \hspace{1cm} (5.12)

since by definition \( I^0_f = 0 \) in \( \mathcal{B} \). Using the statistical model of \( I^0_b \) (5.8) and its further simplification, the statistical description of background pixels is:

\[ I(x, y) \sim \mathcal{N}(L(x, y), C(x, y)), \quad (x, y) \in \mathcal{B} \]  \hspace{1cm} (5.13)

In summary, the proposed method derives estimates \( \hat{L} \) and \( \hat{C} \) from the background component of the observed image \((I(x, y), (x, y) \in \mathcal{B})\) by estimating mean and standard deviation of (5.13), and uses them to recover an estimate \( \hat{I}^0 \) of the observed image \( I^0 \) by applying (5.10).

5.3 Estimation of the Retinal Background

The estimation of \( C \) and \( L \) requires the preliminary extraction of the background set \( \mathcal{B} \). To achieve this goal, the following assumptions have been made: for any pixel of the image, in a neighborhood \( N \) of appropriate size \( s \)

1. both \( L \) and \( C \) are constant;

2. at least 50% of the pixels are background pixels;

3. all background pixels have intensity values significantly different from those of foreground pixels.

The first assumption comes directly from the model hypothesis that the spectral content of \( L \) and \( C \) is concentrated in the low frequencies, whereas the second one indicates that a sufficient portion of background area must be present in each \( N \). The third assumption allows to determine whether pixels belong to background or not simply by examining their intensity, i.e., any two pixels in \( N \) having the same intensity both belong either to background or to foreground.

For each pixel \((x, y)\) in the image, mean \( \mu_N(x, y) \) and standard deviation \( \sigma_N(x, y) \) of the statistical distribution of intensities in \( N \) are estimated. As estimator \( \hat{\mu}_N \) for \( \mu_N(x, y) \) the sample mean is used; \( \hat{\sigma}_N \), estimator for \( \sigma_N(x, y) \), was the unbiased sample standard deviation. Pixel \((x, y)\) is considered to belong to the background set \( \mathcal{B} \) if its intensity is close to the mean intensity in \( N \). This is mathematically expressed by saying that \((x, y)\) belongs to \( \mathcal{B} \) if
5. Luminosity and Contrast Normalization

its Mahalanobis distance from $\hat{\mu}_N$, $d_M$, defined as

$$
\mathbf{d}_M = \left| \frac{I(x, y) - \hat{\mu}_N}{\hat{\sigma}_N} \right| \quad (5.14)
$$

is lower than a given threshold $t$.

The procedure for background pixel extraction could be implemented by means of two filters, evaluating $\hat{\mu}_N(x, y)$ and $\hat{\sigma}_N(x, y)$ respectively for each pixel $(x, y)$. The resulting images could then be combined to evaluate the Mahalanobis distance image, which could be segmented with threshold $t$ to identify the background pixels.

In order to reduce the computational burden, a different implementation has been chosen. The image was partitioned into a tessellation of squares $S_i$ of side $s$. For each $S_i$, $\hat{\mu}(S_i)$ and $\hat{\sigma}(S_i)$ were computed and represented as images. They are a sub-sampled version of the full images $\hat{\mu}_N(x, y)$ and $\hat{\sigma}_N(x, y)$ that would have been obtained with filtering. These full images were then approximated by means of a bicubic interpolation from the sub-sampled images $\hat{\mu}(S_i)$ and $\hat{\sigma}(S_i)$ respectively.

The choice of the value for square side $s$ is critical, since it must satisfy both the assumption on the low spatial dynamics of $C$ and $L$ and the one about the inclusion in $S_i$ of a sufficiently high number of background pixels. A value of $s = 200$ pixels has been empirically chosen, based on a visual inspection of the results. This correspond to an angle of

As regards threshold $t$, its value was set to 1, which means that, with normally distributed luminosity, about 68% of the square pixels are retained as background (note that Mahalanobis distance is by definition normalized for each pixel by its standard deviation $\hat{\sigma}_N(x, y)$). If all or most of the pixels in the square were background pixels, this choice might lead to some underestimation of background mean luminosity and contrast. This is however very rarely the case; in addition, the aim is the estimation of luminosity and contrast variations and thus, being the procedure the same for all squares, they are not significantly affected. An example of a background image is shown in Fig. 2 (top panel).

5.4 Estimation of Observed Local Luminosity and Contrast

Given the set of background pixels $B$, $\hat{L}(x, y)$ and $\hat{C}(x, y)$ were derived for each pixel. From (5.13) and under the first assumption (constant $L$ and $C$ in neighborhood $N$), background pixel intensities in each $N$ are independent,
identically distributed random variables. Thus, \( \hat{L}(x, y) \) and \( \hat{C}(x, y) \) could be derived for each pixel by estimating mean value and standard deviation of this distribution in \( N \).

This approach has to cope with the same computational problems mentioned in the previous section. Moreover, we are now dealing with a sparse set of pixels (background pixels are only a subset of all image pixels), which renders the application of filtering more difficult. A square-processing solution similar to the one presented in the previous section has been adopted. The image was divided into the same tessellation of squares \( S_i \), and from the set of background pixels \( B \) in \( S_i \), mean and standard deviation of intensity values were estimated by using sample mean and standard deviation estimators. Full \( \hat{L}(x, y) \) and \( \hat{C}(x, y) \) were then obtained by applying a bicubic interpolation on the sub-sampled images. Examples of computed \( \hat{L}(x, y) \) and \( \hat{C}(x, y) \) images are shown in Fig. 2 (middle and bottom panels).

The normalized image \( \hat{I}^o \), estimate of observed image \( I^o \), was eventually obtained by applying the point transformation (5.10) to each pixel of the image.

5.5 Estimation of the Illumination and Contrast Models Parameters

Using \( \hat{L}(x, y) \) and \( \hat{C}(x, y) \) directly to obtain the normalized image \( \hat{I}^o \), has proved to be an effective method compensate for illumination and contrast inhomogeneity in most of the images. When in the image there are hemorrhages or other lesions large with respect to the block size, their appearance can be completely changed by the procedure. If a block is mostly contained in a lesion, this latter will be considered as retinal background, and therefore in the normalized image the pixels in the lesion will have a distribution similar to the retinal background. This is a serious threat to the successive possibility of identifying this lesion.
5.6 Luminosity and Contrast Normalization

Normalization of RGB color images can be performed by independently normalizing each color channel with the above procedure:

\[
\begin{bmatrix}
\hat{R}_o (x, y) \\
\hat{G}_o (x, y) \\
\hat{B}_o (x, y)
\end{bmatrix}
= \begin{bmatrix}
\frac{1}{C_R(x,y)} & 0 & 0 \\
0 & \frac{1}{C_G(x,y)} & 0 \\
0 & 0 & \frac{1}{C_B(x,y)}
\end{bmatrix}
\begin{bmatrix}
R (x, y) \\
G (x, y) \\
B (x, y)
\end{bmatrix}
- \begin{bmatrix}
\hat{L}_R (x, y) \\
\hat{L}_G (x, y) \\
\hat{L}_B (x, y)
\end{bmatrix}
\]  

(5.15)

Being the two factors in this equation orthogonal, norm of differences between vectors are still meaningful (i.e., if the distance between vector \(a\) and \(b\) is different from zero, the distance between \(\hat{a}\) and \(\hat{b}\) will still be different from zero). Independent normalization of color components does not however maintain chromatic information. A recovery of the original chromatic distribution can be achieved by identifying an overall image chromatic statistical distribution in the observed image, given by the vectors of sample mean \([\mu_R \mu_G \mu_B]\) and sample standard deviation \([\sigma_R \sigma_G \sigma_B]\), and by forcing it on the normalized image:

\[
\begin{bmatrix}
\hat{R}_o^c (x, y) \\
\hat{G}_o^c (x, y) \\
\hat{B}_o^c (x, y)
\end{bmatrix}
= \begin{bmatrix}
\sigma_R & 0 & 0 \\
0 & \sigma_G & 0 \\
0 & 0 & \sigma_B
\end{bmatrix}
\begin{bmatrix}
\hat{R}_o^c (x, y) \\
\hat{G}_o^c (x, y) \\
\hat{B}_o^c (x, y)
\end{bmatrix}
+ \begin{bmatrix}
\mu_R \\
\mu_G \\
\mu_B
\end{bmatrix}
\]  

(5.16)

where \([\hat{R}_o^c \hat{G}_o^c \hat{B}_o^c]\) represents the vector of image components normalized with respect to luminosity, contrast and chromatic distribution.

It should be observed, however, that this last step, aimed at maintaining the image overall chromatic statistical distribution, was included to provide images that look more natural when observed by a human expert, but has no influence nor importance on the automatic analysis of images, which is performed on separate color channels.

5.7 Performance Evaluation

In order to better appreciate the results obtained with the proposed technique, results on both local and global indexes are reported. Using the green channel layer, each test image was partitioned into 25 blocks, 300x340 pixel each, and the results are presented first at block level (local) and then for the whole image (global).
5. Luminosity and Contrast Normalization

Comparing distribution of local luminosity and contrast between two images, e.g. before and after normalization, is meaningful only if the two images are normalized to the same global luminosity and contrast level. In fact, if we compare local distributions between two images having a systematic difference in global luminosity and contrast, the difference between the two local distributions would be affected both by the systematic difference and by the actual difference in local distribution. Let define $\mu(I)$ as the mean value (luminosity) and $\sigma(I)$ as the standard deviation (contrast) of intensity levels for image $I$. These two indicators can be arbitrarily modified by means of a linear point transformation on $I$, e.g., if the transformed image is $I^T = aI + b$ we obtain:

$$
\mu(I^T) = \mu(aI + b) = a\mu(I) + b \quad (5.17)
$$

$$
\sigma(I^T) = \sigma(aI + b) = a\sigma(I) \quad (5.18)
$$

Before computing performance indexes, the observed and normalized images have been transformed by 5.17 and 5.18 so as to have histograms with the same mean and standard deviation. This way, the improvement in luminosity and contrast are not simply due to histogram shift and stretching.

5.7.1 Local Indexes

In order to quantitatively assess the effect of the applied normalization, indexes of luminosity and contrast have been defined and evaluated for each block of the image. The analysis of the distribution of these local indicators over the image allows to express the spatial variation in luminosity and contrast.

For the $i^{th}$ block, luminosity was defined as the mean intensity level, $\mu_i$. Contrast has been quantified by means of two indexes. The first one is based on the standard deviation of intensity levels, $\sigma_i$. Since this quantity might be influenced by the presence of locally variable (non spatially stationary) noise, an additional, theoretically more robust index was defined, based on the distance between 10% and 90% quantile values, $c_i = rank_{90\%} - rank_{10\%}$. Both contrast indexes were evaluated only for blocks of the image containing a significant signal, corresponding to portions of the image containing a foreground part (vascular structures or lesions).

The normalization performed by the proposed algorithm provides a uniform luminosity and contrast throughout the image. It modifies local histograms, as can be seen in Fig. 3, where the histograms for the 25 image blocks are shown for the observed and normalized version of the same image. The applied normalization aligns the histogram mean value (luminosity)
5. Luminosity and Contrast Normalization

much closer to the center of the range, and spread them (contrast) over the full range of luminosity values. Normalizing image luminosity has thus a beneficial effect on the local contrast (inside single blocks), even if mean luminosity and contrast were forced to be the same in the normalized and in the observed images. This can be explained by considering that in the observed image a significant portion of the intensity range may be taken up by the luminosity drift, i.e., most of the contrast present in the observed image is due to the difference in mean intensity between different areas of the image. It can be shown that the distribution of intensity values in the union of the two regions, $a$ and $b$ (see Fig. 5.2), has the following standard deviation (overall contrast):

$$\sigma_{a\cup b} = \sqrt{\frac{1}{2} \left[ \sigma_a^2 + \sigma_b^2 + \frac{1}{2} (\mu_a - \mu_b)^2 \right]}$$

(5.19)

Reducing the difference $(\mu_a - \mu_b)$ clearly increases the sum of the local contrasts, $\sigma_a^2$ and $\sigma_b^2$, if $\sigma_{a\cup b}$ is kept constant.

5.7.2 Global Indexes

The following indexes were defined for the whole image:

- $\sigma_\mu = \sqrt{\frac{1}{N} \sum_i (\mu_i - \mu_\mu)^2}$, where $\mu_\mu = \frac{1}{N} \sum_i \mu_i$
- $\mu_\sigma = \frac{1}{N} \sum_i \sigma_i$
- $\mu_c = \frac{1}{N} \sum_i c_i$

where $N = 25$ is the number of blocks in the image.

Index $\sigma_\mu$ expresses the variability of the local luminosity throughout the image, and thus the lower this index, the more uniform the image luminosity. Indexes $\mu_\sigma$ and $\mu_c$ express the mean contrast level within the image.

The values of these indexes, evaluated on the set of 33 images, are reported in Table 1. The decrease in standard deviation of local luminosity ranges from -2% to -45%, with an average of -19%; in five images (out of 33), increments were obtained, albeit modest ones (range 1% - 12%). The mean contrast level $\mu_\sigma$ increases from a minimum of 2% up to 85%, with an average of 34%, and with only 2 instances of decrement (-5% and -6%); contrast index $\mu_c$ showed increments in the range 2% - 80% (average 31%), with 3 instances of increment (range -3% to -6%).
5.7.3 Comparison with other correction techniques

As already mentioned, luminosity and contrast normalization can be achieved also by means of other correction techniques. The low-pass correction, e.g., is obtained by subtracting a low-pass filtered version of the image from the observed image. The procedure proposed in [93] performs a pixel-wise nonlinear filtering, while the Wallis filter is based on a nonlinear and space-variant technique. At variance with these methods, the technique proposed in [95] works only over the vascular part to gather information about luminosity and contrast drift.

In order to allow a better appreciation of the performances of our technique with respect to the others, we have implemented all the above mentioned techniques and compared them with the proposed method. The low-pass filter was implemented by means of a convolution with a Gaussian kernel, whose $3\sigma$ value has been set to the square size $s$ of our normalization system. The Wallis filter was implemented as described in an on-line reference (http://www.microimages.com/fetupd/v55/wallis), while the other two techniques have been implemented according to the cited references. For the computation of the local mean and standard deviation, required by the Wallis filter, we have used the values obtained by our first stage (Estimation of background pixels), since this method is very similar to the partitioning-interpolating scheme proposed in the cited reference.

The results of these comparisons are presented in Tab. 5.2, as regards local luminosity, and in Tab. 5.3, as regards contrast (results for the second contrast index, $\mu_c$, are not shown since they are very similar to the ones obtained with the first contrast index, $\mu_\sigma$). With respect to the proposed technique, image luminosity variability of all other techniques is on average higher (from 11% of the filter proposed by Y. Wang to 40% of low-pass filter), while achieved mean contrast is on average lower (from -3% of the Wallis filter to -28% of the filter proposed by Y. Wang). Figg. 5.5 to 5.8 report some examples of the results obtained by the different correction techniques on representative images. The algorithm described in [78] is that with results closest to the algorithm proposed in this chapter. By looking at images after the processing, it is clearly that it introduces an heavy shading on the periphery of the images, thus achieving a global low mean luminosity variation, while at the same time it performs a contrast enhancement on the center of the image, achieving a good contrast around larger vessel. Unfortunately this is hidden by the chosen performance indexes.
5. Luminosity and Contrast Normalization

<table>
<thead>
<tr>
<th></th>
<th>$\sigma_\mu$</th>
<th>$\sigma_{\mu}^N$</th>
<th>% difference</th>
<th>$\mu_\sigma$</th>
<th>$\mu_{\sigma}^N$</th>
<th>% difference</th>
<th>$\mu_c$</th>
<th>$\mu_{c}^N$</th>
<th>% difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.037</td>
<td>0.041</td>
<td>12</td>
<td>0.035</td>
<td>0.034</td>
<td>-5</td>
<td>0.085</td>
<td>0.079</td>
<td>-6</td>
</tr>
<tr>
<td>2</td>
<td>0.058</td>
<td>0.062</td>
<td>6</td>
<td>0.042</td>
<td>0.039</td>
<td>-6</td>
<td>0.105</td>
<td>0.098</td>
<td>-6</td>
</tr>
<tr>
<td>3</td>
<td>0.025</td>
<td>0.031</td>
<td>-21</td>
<td>0.033</td>
<td>0.034</td>
<td>4</td>
<td>0.085</td>
<td>0.087</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>0.051</td>
<td>0.053</td>
<td>-38</td>
<td>0.059</td>
<td>0.074</td>
<td>24</td>
<td>0.135</td>
<td>0.160</td>
<td>19</td>
</tr>
<tr>
<td>5</td>
<td>0.038</td>
<td>0.038</td>
<td>-1</td>
<td>0.036</td>
<td>0.043</td>
<td>17</td>
<td>0.095</td>
<td>0.110</td>
<td>16</td>
</tr>
<tr>
<td>6</td>
<td>0.058</td>
<td>0.043</td>
<td>-25</td>
<td>0.054</td>
<td>0.069</td>
<td>28</td>
<td>0.126</td>
<td>0.158</td>
<td>25</td>
</tr>
<tr>
<td>7</td>
<td>0.077</td>
<td>0.068</td>
<td>-11</td>
<td>0.039</td>
<td>0.057</td>
<td>45</td>
<td>0.094</td>
<td>0.134</td>
<td>42</td>
</tr>
<tr>
<td>8</td>
<td>0.074</td>
<td>0.060</td>
<td>-20</td>
<td>0.037</td>
<td>0.066</td>
<td>81</td>
<td>0.089</td>
<td>0.161</td>
<td>80</td>
</tr>
<tr>
<td>9</td>
<td>0.082</td>
<td>0.071</td>
<td>-14</td>
<td>0.042</td>
<td>0.060</td>
<td>45</td>
<td>0.103</td>
<td>0.149</td>
<td>45</td>
</tr>
<tr>
<td>10</td>
<td>0.182</td>
<td>0.197</td>
<td>8</td>
<td>0.070</td>
<td>0.125</td>
<td>80</td>
<td>0.176</td>
<td>0.314</td>
<td>78</td>
</tr>
<tr>
<td>11</td>
<td>0.029</td>
<td>0.026</td>
<td>-11</td>
<td>0.032</td>
<td>0.035</td>
<td>10</td>
<td>0.080</td>
<td>0.087</td>
<td>9</td>
</tr>
<tr>
<td>12</td>
<td>0.032</td>
<td>0.023</td>
<td>-28</td>
<td>0.030</td>
<td>0.034</td>
<td>13</td>
<td>0.077</td>
<td>0.086</td>
<td>12</td>
</tr>
<tr>
<td>13</td>
<td>0.043</td>
<td>0.031</td>
<td>-28</td>
<td>0.036</td>
<td>0.045</td>
<td>26</td>
<td>0.089</td>
<td>0.112</td>
<td>25</td>
</tr>
<tr>
<td>14</td>
<td>0.042</td>
<td>0.024</td>
<td>-43</td>
<td>0.035</td>
<td>0.043</td>
<td>22</td>
<td>0.092</td>
<td>0.111</td>
<td>21</td>
</tr>
<tr>
<td>15</td>
<td>0.041</td>
<td>0.035</td>
<td>-14</td>
<td>0.035</td>
<td>0.046</td>
<td>33</td>
<td>0.089</td>
<td>0.116</td>
<td>31</td>
</tr>
<tr>
<td>16</td>
<td>0.036</td>
<td>0.028</td>
<td>-21</td>
<td>0.034</td>
<td>0.039</td>
<td>15</td>
<td>0.088</td>
<td>0.100</td>
<td>14</td>
</tr>
<tr>
<td>17</td>
<td>0.040</td>
<td>0.029</td>
<td>-28</td>
<td>0.034</td>
<td>0.045</td>
<td>32</td>
<td>0.087</td>
<td>0.114</td>
<td>31</td>
</tr>
<tr>
<td>18</td>
<td>0.056</td>
<td>0.041</td>
<td>-28</td>
<td>0.043</td>
<td>0.056</td>
<td>32</td>
<td>0.108</td>
<td>0.141</td>
<td>31</td>
</tr>
<tr>
<td>19</td>
<td>0.096</td>
<td>0.053</td>
<td>-45</td>
<td>0.051</td>
<td>0.094</td>
<td>85</td>
<td>0.127</td>
<td>0.229</td>
<td>80</td>
</tr>
<tr>
<td>20</td>
<td>0.108</td>
<td>0.072</td>
<td>-34</td>
<td>0.055</td>
<td>0.094</td>
<td>70</td>
<td>0.134</td>
<td>0.219</td>
<td>64</td>
</tr>
<tr>
<td>21</td>
<td>0.089</td>
<td>0.067</td>
<td>-35</td>
<td>0.053</td>
<td>0.072</td>
<td>35</td>
<td>0.129</td>
<td>0.168</td>
<td>30</td>
</tr>
<tr>
<td>22</td>
<td>0.089</td>
<td>0.091</td>
<td>2</td>
<td>0.051</td>
<td>0.089</td>
<td>73</td>
<td>0.108</td>
<td>0.183</td>
<td>70</td>
</tr>
<tr>
<td>23</td>
<td>0.061</td>
<td>0.036</td>
<td>-41</td>
<td>0.052</td>
<td>0.075</td>
<td>44</td>
<td>0.118</td>
<td>0.165</td>
<td>40</td>
</tr>
<tr>
<td>24</td>
<td>0.078</td>
<td>0.058</td>
<td>-27</td>
<td>0.050</td>
<td>0.070</td>
<td>40</td>
<td>0.112</td>
<td>0.151</td>
<td>34</td>
</tr>
<tr>
<td>25</td>
<td>0.062</td>
<td>0.057</td>
<td>-9</td>
<td>0.045</td>
<td>0.057</td>
<td>26</td>
<td>0.095</td>
<td>0.112</td>
<td>18</td>
</tr>
<tr>
<td>26</td>
<td>0.090</td>
<td>0.091</td>
<td>1</td>
<td>0.034</td>
<td>0.048</td>
<td>42</td>
<td>0.080</td>
<td>0.111</td>
<td>40</td>
</tr>
<tr>
<td>27</td>
<td>0.066</td>
<td>0.048</td>
<td>-27</td>
<td>0.060</td>
<td>0.071</td>
<td>19</td>
<td>0.130</td>
<td>0.156</td>
<td>20</td>
</tr>
<tr>
<td>28</td>
<td>0.090</td>
<td>0.059</td>
<td>-34</td>
<td>0.065</td>
<td>0.094</td>
<td>44</td>
<td>0.148</td>
<td>0.202</td>
<td>37</td>
</tr>
<tr>
<td>29</td>
<td>0.066</td>
<td>0.052</td>
<td>-20</td>
<td>0.061</td>
<td>0.079</td>
<td>30</td>
<td>0.139</td>
<td>0.175</td>
<td>26</td>
</tr>
<tr>
<td>30</td>
<td>0.061</td>
<td>0.050</td>
<td>-17</td>
<td>0.042</td>
<td>0.050</td>
<td>20</td>
<td>0.088</td>
<td>0.103</td>
<td>17</td>
</tr>
<tr>
<td>31</td>
<td>0.087</td>
<td>0.060</td>
<td>-31</td>
<td>0.059</td>
<td>0.089</td>
<td>51</td>
<td>0.139</td>
<td>0.203</td>
<td>46</td>
</tr>
<tr>
<td>32</td>
<td>0.082</td>
<td>0.066</td>
<td>-20</td>
<td>0.036</td>
<td>0.056</td>
<td>53</td>
<td>0.086</td>
<td>0.128</td>
<td>48</td>
</tr>
<tr>
<td>33</td>
<td>0.072</td>
<td>0.071</td>
<td>-2</td>
<td>0.041</td>
<td>0.041</td>
<td>2</td>
<td>0.091</td>
<td>0.088</td>
<td>-3</td>
</tr>
</tbody>
</table>

Average: -19 34 31

Table 5.1: Indexes for local luminosity and contrast for observed ($\sigma_\mu$, $\mu_\sigma$, $\mu_c$) and normalized ($\sigma_{\mu}^N$, $\mu_{\sigma}^N$, $\mu_{c}^N$) images, and their percent differences.

5.8 Conclusions

In this chapter an algorithm for illumination and contrast correction in retinal images has been presented and compared with those proposed in the literature. It has been proved more robust and sound with respect to the computed indexes. In Figg. 5.5 to 5.8 are shown the results of various illumination correction procedures. The main drawback of the algorithm presented in this chapter is its dependency of the block size for estimating the retinal background. In fact in presence of large hemorrhages or exudates it might blur the information contents of these lesions as if they were retinal background: this happens when the kernel on which mean and standard deviation are estimated falls completely inside a lesion, or a non-background regions.
Table 5.2: Indexes for local luminosity obtained with the proposed technique ($\sigma_N^L$), the low-pass correction ($\sigma_L^P$), the Wallis filter correction [91] ($\sigma_W^S$), H. Wang [93] ($\sigma_H^W$), and Y. Wang [95] ($\sigma_Y^W$).

<table>
<thead>
<tr>
<th></th>
<th>$\sigma_N^L$</th>
<th>$\sigma_L^P$</th>
<th>% difference</th>
<th>$\sigma_W^S$</th>
<th>% difference</th>
<th>$\sigma_S$</th>
<th>% difference</th>
<th>$\sigma_H^W$</th>
<th>% difference</th>
<th>$\sigma_Y^W$</th>
<th>% difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.040</td>
<td>0.049</td>
<td>22</td>
<td>0.046</td>
<td>14</td>
<td>0.039</td>
<td>-3</td>
<td>0.036</td>
<td>-10</td>
<td>0.040</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0.058</td>
<td>0.067</td>
<td>91</td>
<td>0.064</td>
<td>12</td>
<td>0.055</td>
<td>-4</td>
<td>0.058</td>
<td>1</td>
<td>0.049</td>
<td>-15</td>
</tr>
<tr>
<td>3</td>
<td>0.019</td>
<td>0.037</td>
<td>107</td>
<td>0.040</td>
<td>107</td>
<td>0.028</td>
<td>47</td>
<td>0.025</td>
<td>27</td>
<td>0.029</td>
<td>51</td>
</tr>
<tr>
<td>4</td>
<td>0.032</td>
<td>0.061</td>
<td>38</td>
<td>0.044</td>
<td>38</td>
<td>0.054</td>
<td>66</td>
<td>0.048</td>
<td>49</td>
<td>0.047</td>
<td>47</td>
</tr>
<tr>
<td>5</td>
<td>0.039</td>
<td>0.054</td>
<td>15</td>
<td>0.047</td>
<td>22</td>
<td>0.043</td>
<td>12</td>
<td>0.038</td>
<td>-1</td>
<td>0.038</td>
<td>-3</td>
</tr>
<tr>
<td>6</td>
<td>0.043</td>
<td>0.069</td>
<td>5</td>
<td>0.058</td>
<td>34</td>
<td>0.056</td>
<td>30</td>
<td>0.057</td>
<td>31</td>
<td>0.045</td>
<td>5</td>
</tr>
<tr>
<td>7</td>
<td>0.068</td>
<td>0.068</td>
<td>0</td>
<td>0.065</td>
<td>-5</td>
<td>0.063</td>
<td>-8</td>
<td>0.079</td>
<td>16</td>
<td>0.047</td>
<td>-31</td>
</tr>
<tr>
<td>8</td>
<td>0.061</td>
<td>0.071</td>
<td>15</td>
<td>0.077</td>
<td>26</td>
<td>0.047</td>
<td>-22</td>
<td>0.077</td>
<td>25</td>
<td>0.052</td>
<td>-16</td>
</tr>
<tr>
<td>9</td>
<td>0.073</td>
<td>0.068</td>
<td>-7</td>
<td>0.069</td>
<td>-19</td>
<td>0.064</td>
<td>-13</td>
<td>0.085</td>
<td>16</td>
<td>0.056</td>
<td>-23</td>
</tr>
<tr>
<td>10</td>
<td>0.206</td>
<td>0.106</td>
<td>-49</td>
<td>0.144</td>
<td>-30</td>
<td>0.178</td>
<td>-14</td>
<td>0.196</td>
<td>-5</td>
<td>0.102</td>
<td>-51</td>
</tr>
<tr>
<td>11</td>
<td>0.026</td>
<td>0.043</td>
<td>62</td>
<td>0.033</td>
<td>24</td>
<td>0.035</td>
<td>32</td>
<td>0.028</td>
<td>8</td>
<td>0.032</td>
<td>23</td>
</tr>
<tr>
<td>12</td>
<td>0.023</td>
<td>0.041</td>
<td>80</td>
<td>0.037</td>
<td>62</td>
<td>0.032</td>
<td>40</td>
<td>0.032</td>
<td>38</td>
<td>0.028</td>
<td>24</td>
</tr>
<tr>
<td>13</td>
<td>0.032</td>
<td>0.055</td>
<td>70</td>
<td>0.042</td>
<td>30</td>
<td>0.043</td>
<td>35</td>
<td>0.043</td>
<td>35</td>
<td>0.045</td>
<td>41</td>
</tr>
<tr>
<td>14</td>
<td>0.024</td>
<td>0.054</td>
<td>122</td>
<td>0.046</td>
<td>89</td>
<td>0.043</td>
<td>76</td>
<td>0.042</td>
<td>72</td>
<td>0.037</td>
<td>53</td>
</tr>
<tr>
<td>15</td>
<td>0.034</td>
<td>0.055</td>
<td>62</td>
<td>0.052</td>
<td>51</td>
<td>0.044</td>
<td>30</td>
<td>0.041</td>
<td>20</td>
<td>0.043</td>
<td>26</td>
</tr>
<tr>
<td>16</td>
<td>0.028</td>
<td>0.047</td>
<td>68</td>
<td>0.044</td>
<td>56</td>
<td>0.037</td>
<td>33</td>
<td>0.036</td>
<td>28</td>
<td>0.033</td>
<td>18</td>
</tr>
<tr>
<td>17</td>
<td>0.029</td>
<td>0.053</td>
<td>83</td>
<td>0.050</td>
<td>72</td>
<td>0.042</td>
<td>44</td>
<td>0.041</td>
<td>41</td>
<td>0.037</td>
<td>29</td>
</tr>
<tr>
<td>18</td>
<td>0.042</td>
<td>0.063</td>
<td>50</td>
<td>0.047</td>
<td>11</td>
<td>0.052</td>
<td>23</td>
<td>0.056</td>
<td>32</td>
<td>0.048</td>
<td>14</td>
</tr>
<tr>
<td>19</td>
<td>0.059</td>
<td>0.094</td>
<td>59</td>
<td>0.096</td>
<td>63</td>
<td>0.088</td>
<td>51</td>
<td>0.097</td>
<td>66</td>
<td>0.087</td>
<td>48</td>
</tr>
<tr>
<td>20</td>
<td>0.072</td>
<td>0.094</td>
<td>30</td>
<td>0.102</td>
<td>41</td>
<td>0.084</td>
<td>16</td>
<td>0.110</td>
<td>52</td>
<td>0.079</td>
<td>9</td>
</tr>
<tr>
<td>21</td>
<td>0.070</td>
<td>0.084</td>
<td>20</td>
<td>0.074</td>
<td>6</td>
<td>0.076</td>
<td>9</td>
<td>0.081</td>
<td>17</td>
<td>0.103</td>
<td>48</td>
</tr>
<tr>
<td>22</td>
<td>0.149</td>
<td>0.083</td>
<td>-44</td>
<td>0.081</td>
<td>-46</td>
<td>0.074</td>
<td>-51</td>
<td>0.091</td>
<td>-39</td>
<td>0.075</td>
<td>-50</td>
</tr>
<tr>
<td>23</td>
<td>0.040</td>
<td>0.066</td>
<td>66</td>
<td>0.056</td>
<td>39</td>
<td>0.057</td>
<td>43</td>
<td>0.060</td>
<td>49</td>
<td>0.046</td>
<td>16</td>
</tr>
<tr>
<td>24</td>
<td>0.057</td>
<td>0.073</td>
<td>27</td>
<td>0.062</td>
<td>9</td>
<td>0.065</td>
<td>14</td>
<td>0.074</td>
<td>30</td>
<td>0.069</td>
<td>21</td>
</tr>
<tr>
<td>25</td>
<td>0.053</td>
<td>0.069</td>
<td>29</td>
<td>0.065</td>
<td>22</td>
<td>0.061</td>
<td>14</td>
<td>0.061</td>
<td>15</td>
<td>0.049</td>
<td>-7</td>
</tr>
<tr>
<td>26</td>
<td>0.074</td>
<td>0.087</td>
<td>18</td>
<td>0.089</td>
<td>21</td>
<td>0.076</td>
<td>3</td>
<td>0.089</td>
<td>20</td>
<td>0.065</td>
<td>-11</td>
</tr>
<tr>
<td>27</td>
<td>0.048</td>
<td>0.063</td>
<td>30</td>
<td>0.059</td>
<td>22</td>
<td>0.053</td>
<td>11</td>
<td>0.062</td>
<td>28</td>
<td>0.059</td>
<td>22</td>
</tr>
<tr>
<td>28</td>
<td>0.063</td>
<td>0.069</td>
<td>10</td>
<td>0.063</td>
<td>-1</td>
<td>0.063</td>
<td>-1</td>
<td>0.087</td>
<td>38</td>
<td>0.065</td>
<td>3</td>
</tr>
<tr>
<td>29</td>
<td>0.049</td>
<td>0.066</td>
<td>36</td>
<td>0.067</td>
<td>38</td>
<td>0.060</td>
<td>22</td>
<td>0.065</td>
<td>33</td>
<td>0.062</td>
<td>27</td>
</tr>
<tr>
<td>30</td>
<td>0.050</td>
<td>0.060</td>
<td>21</td>
<td>0.056</td>
<td>-2</td>
<td>0.056</td>
<td>1</td>
<td>0.058</td>
<td>15</td>
<td>0.046</td>
<td>-10</td>
</tr>
<tr>
<td>31</td>
<td>0.057</td>
<td>0.079</td>
<td>38</td>
<td>0.059</td>
<td>4</td>
<td>0.069</td>
<td>21</td>
<td>0.078</td>
<td>27</td>
<td>0.046</td>
<td>17</td>
</tr>
<tr>
<td>32</td>
<td>0.059</td>
<td>0.084</td>
<td>41</td>
<td>0.065</td>
<td>10</td>
<td>0.069</td>
<td>17</td>
<td>0.082</td>
<td>38</td>
<td>0.064</td>
<td>8</td>
</tr>
<tr>
<td>33</td>
<td>0.061</td>
<td>0.076</td>
<td>24</td>
<td>0.076</td>
<td>23</td>
<td>0.064</td>
<td>4</td>
<td>0.070</td>
<td>14</td>
<td>0.074</td>
<td>21</td>
</tr>
</tbody>
</table>

Average: 39, 26, 18, 25, 11
Table 5.3: Indexes for contrast obtained with the proposed technique (\(\mu^N_\sigma\)), the low-pass correction (\(\mu^{LP}_\sigma\)), the Wallis filter correction [91] (\(\mu^W_\sigma\)), the correction proposed by Sinthanayothin [78] (\(\mu^S_\sigma\)), H. Wang [93] (\(\mu^{HW}_\sigma\)), and Y. Wang [95] (\(\mu^{YW}_\sigma\)).
5. Luminosity and Contrast Normalization

(a) Estimated background image (white pixels)

(b) Estimated illumination drift $\hat{L}$

(c) Estimated contrast drift $\hat{C}$

Figure 5.1:
5. Luminosity and Contrast Normalization

Figure 5.2: Reducing the difference \((\mu_a - \mu_b)\) clearly increases the sum of the local contrasts, \(\sigma_a^2\) and \(\sigma_b^2\), if \(\sigma_{a,b}\) is kept constant.
5. Luminosity and Contrast Normalization

Figure 5.3: Histograms of luminosity values in the 25 image blocks for the observed and normalized version of the same image
Figure 5.4: Intensity profile on a sample row and intensity drift correction estimated by the proposed technique and by a low-pass filter.
5. Luminosity and Contrast Normalization

Figure 5.5: Retinal image corrected with different methods
5. Luminosity and Contrast Normalization

(a) Observed Image  (b) Low Pass Correction

(c) Homomorphic Filter  (d) Non-linear Transform

(e) Wallis Filter  (f) Adaptive Filter

(g) Vessel Sample and Interpolation  (h) Proposed Correction

Figure 5.6: Retinal image corrected with different methods
Figure 5.7: Retinal image corrected with different methods
5. Luminosity and Contrast Normalization

Figure 5.8: Retinal image corrected with different methods
Vessel Tracking

Most retinopathies, e.g. from hypertension or diabetes, could be early diagnosed and treated if an accurate and objective analysis of symptoms at their initial onset could be performed. The analysis should be accurate enough to detect minor pathological signs, and objective enough to be able to compare results with accepted clinical standards and with results obtained from the same patient at different times. This latter requirement is of paramount importance when assessing the effect of established therapeutic treatments and even more when evaluating in a quantitative way the efficacy of new drugs during their development.

Most of the early symptoms indicating the onset of retinopathies are related to morphological features of the retinal vascular tree [83]. When no major signs of retinal degeneration are present (such as cotton wool spots, haemorrhages, exudates), the clinical diagnostic procedure for retinopathy always starts with a careful evaluation of the main features of the network of retinal vessels, obtained from fundus camera images. The clinically most relevant signs taken into account by expert ophthalmologists are in general vessel tortuosity, vessel calibre and its distribution among different vessels, presence of vessel calibre irregularities along the same vessel, and the so-called Gunn and Salus signs, i.e., local calibre reduction or local deviation of vessel direction at the crossings between artery and vein [42].

In order to detect and quantitatively describe these diagnostic signs, the information to be extracted from the vascular network are the layout and the dimension of all the relevant vessels contained in the image. This task is relatively easy for an expert ophthalmologist if performed at a qualitative level, but rather cumbersome, highly subjective and error prone if a set of measurements is sought for [82, 24]. For this reason, a number of research projects have been carried out to develop automatic computerized systems for the extraction of retinal vascular structure.
6. Vessel Tracking

In order to track the entire vessel network, a modular algorithm has been
designed. The Seed Points Extraction procedure described in 6.2 identifies
on the image and stores in pool $S$ specific pixels, called seed points, which
will be the starting points of the tracking algorithm. The Sparse Tracking
procedure described in 6.4 considers each seed point and from them it tracks
a vessel step by step, recognizing its center and calibre at each step. It stops
when an exceptional condition (vessel bifurcation, crossing, or low contrast
area) is reached; the segment just recognized is added to the set of tracked
segments, $SEG$. A new seed point is then extracted from $S$ and from it Sparse
Tracking is started again. During these iterative steps, the procedure Bubble
Analysis described in 6.3 is called both when a seed point is first analyzed
to estimate initial vessel calibre and direction, and also at the end-points of
each tracked segment to generate new seed points, which are added to $S$.
The sparse tracking procedure ends when the pool of available seed points $S$
is empty.

All the tracked segments stored in $SEG$ are then examined, in order to discard
those that are not vessel. This is done by mean of a False Vessel Elimination
procedure described in 6.5, at the end of which all segments are smoothed
by means of a cubic smoothing spline. Finally, a Segments Connection pro-
cedure described in 6.6 will link the segments in order to reconstruct the
connected vessel structure.

6.1 Review of Available Methods

Automatic techniques for vessel identification have been proposed for general
angiography, e.g. [55, 64, 25], and also for more specialized areas such as
coronary angiography, e.g. [82, 24, 69, 84, 26, 81, 85, 12, 77], and retinal
angiography, e.g. [16, 100, 89, 11, 41]. In order to review the methods
proposed to identify vessels in retinal images, four classes of algorithms have
been considered [87]: local operators, matched filters, vessel tracking and
neural networks.

6.1.1 Local vascular segmentation

The simplest approaches to vessel segmentation try to exploit the fact that
usually retinal vessels are darker than their surroundings. Global or local
threshold and morphological operators on the image have been proposed to
identify the connected regions that represent vessels.

More recently local gradient of the image has been used [99, 94]to recognize
the edge pairs that identify the vessels. The second derivative has been used
to evaluate the local candidate vessel orientation [58], and then the pixels have been classified in vessel or non-vessels based on the gradient distribution.

6.1.2 Matched Filters

Matched filters are based upon a correlation measure between the expected shape looked for and the measured signal. In order to identify vessels, a set of kernels which mimic vessel shapes both in orientation and in scale needs to be convolved with the retinal image. Peak in the convolution result means a match between the portion of the image around the peak and the vessel-shaped kernel. A number of filter shapes have been investigated. In [16, 32, ?] extruded and rotated gaussian filters are used, but also filters based on lines [94] and partial Gaussian [11]. A number of strategies have also been proposed for identify true vessels from the filter response. A local entropy thresholding method has been proposed in [13]. In particular in [?] a strategy similar to a multiresolution approach was proposed, iteratively decrementing an initial threshold, and then checking the obtained connected component for vessel likeliness.

Matched filters are not guaranteed to work properly in presence of large variability in vessel calibre distribution, since the template has to be tuned to a fixed set of reference vessel calibres. This kind of algorithms are also very time-consuming, since they involve the processing of the entire image.

6.1.3 Artificial Neural Networks

Neural networks have been tried for detection and segmentation of the retinal vessels [29, 78]. The advantages of neural network is that they automatically tune the parameters of the net based on the example it has been given. The more complex and deep is the net structure, the more complex are the classification boundaries it can generate. As every supervised learning procedure, the choice of the training set needs utmost care, and the results can be quite disappointing in pathological images.

6.1.4 Vessel Tracking

The simplest The tracking techniques start from a point on a vessel and move along it as far as possible by analyzing consecutive local areas, e.g. drawing scan lines across the vessel. Typically, starting and ending points have to be
6. Vessel Tracking

provided by the user. In order to identify the vessel profile along the scan line, matched filters are still used [55, 64, 69, 100, 11, 41], together with simpler techniques like derivative analysis [85, 12], or more sophisticated ones such as morphological filters [25] or Fuzzy C-Mean classifiers [89].

6.2 The Sparse Approach: Seed Points Extraction

The goal of this procedure is to identify the set of seed points, i.e., points from which the segment tracking algorithm will start [11]. A set of equally-spaced rows and columns of $I$ is analyzed, on which candidate seed points are searched. A scan analysis is performed on each of these lines, in order to identify sequences of pixels corresponding to possible vessel profiles. Every line of pixels is treated as a one dimensional signal, which is at first processed by a 3-elements low-pass averaging filter. The signal is then numerically differentiated, to eliminate the effect of possible luminance drifts along the line.

On the derivative signal, a vessel profile would result in two spikes of opposite sign, corresponding to the transitions from background to vessel (first border) and then from vessel to background (second border). This is the pattern that is searched for in the derivative signal by the following finite-state machine, where $d$ is the derivative value and $th$ a suitable threshold:
Seed Points Extraction

```c
switch
    case No Vessel detected
        if d > th
    goto 2
    break
    case First Border Detected
        Save first vessel border coordinate
        if d < th
    goto 11
    break
    case Inside Vessel
        if d < −th
    goto 15
    break
    case Second Border Detected
        Save second vessel border coordinate
        if d > −th
    goto 2
    break
```

The $x, y$ coordinates of the two pixels identified as vessel borders are used to determine the center of the detected vessel, which is taken as the seed point we were looking for. Threshold $th$ is not known a-priori and is related to image contrast. The approach adopted to determine its value is an iterative one, in which decreasing threshold values are used until a fixed minimum number of seed points is extracted. This approach is based on the assumption that the number of seed points contained in the grid is fairly constant among images.

In order to reduce the total number of detected seed points, those positioned on non-significant or artifact areas are removed from $S$. Points on the optic disc, where vessels are not clinically significant, and on vessels with very large calibre (possible artifacts) or very small (negligible vessels) are searched and removed. Moreover, if we define a local intensity variability $\sigma_N$ as the standard deviation of pixel intensities in a 40x40 neighborhood around a seed point, seed points positioned on a vessel will have high values of $\sigma_N$, due to the presence in the neighborhood of both “vessel” (dark) and “non-vessel” (light) pixels. Mean $\mu$ and standard deviation $\sigma$ of the $\sigma_N$ of all the seed points are computed and seed points with a low value of the local intensity variability, e.g. lower than $\mu - k\sigma$, are also removed; $k$ is a user-selected constant.
6. Vessel Tracking

To further reduce the number of seed points, a clustering procedure is performed. It starts by clustering the two closest seed points whose distance is lower than a threshold $\delta$ and then proceeds by iteratively adding to the cluster all the seed points that have distances from all the seed points in the cluster lower than $\delta$. This iterative addition ends when no new seed points match the clustering condition. The whole procedure is iterated until all seed points have been examined and clustered. The seed points with the lowest intensity value (darkest) in each cluster are taken as the representatives of the cluster and are the ones returned by the clustering algorithm.

6.3 Bubble Analysis

Seeds points extracted in ?? have no information on the direction and caliber of their correspondent vessels. Before the tracking procedure can start, an initial estimate of this vessel direction has to be extracted. This is done by means of a “bubble” analysis technique, which will be further used to find possible vessel paths in critical situations, i.e. when the tracking is not able to move further along a vessel due to presence of bifurcations or crossings or to poor image contrast.

The idea of the “bubble analysis” is to look in concentric circular lines around a point, in an attempt to look “beyond” local information or critical points. The Bubble Analysis procedure consists of the following steps:

1. concentric circular scan lines are analyzed around each vessel end-point;
2. vessel profiles are extracted along these scan lines and vessel centers and calibers are determined;
3. center points are filtered by the Hough transform to eliminate spurious recognitions;
4. center points are transformed into polar space and clustered;
5. cluster centers are added to set $S$.

In the first two steps of the procedure, pixels of each circular line are clustered with a fuzzy c-mean algorithm [9], using their gray-level values as classification feature. After defuzzification, a two-class separation is obtained for the pixels along every circular line: the “vessel” and “non-vessel” classes. We can assume that in the small neighborhood around the end-point covered by bubbles, vessels that may be present can be considered as straight lines moving radially away from the end-point. This means that the center points of
6. Vessel Tracking

each vessel are aligned along a straight line. In order to remove spurious ves-

sel center recognitions, the whole set of center points is filtered by the Hough

transform. In the Hough transform space, each point is represented by a

sinusoid, and all the points aligned along a straight line have sinusoids with

a common intersection. Therefore, spurious center points are represented by

sinusoids that do not pass through these common intersections and can thus

be recognized and removed.

The remaining points are then transformed into a polar coordinates space,

having the vessel end-point as its origin. Under the assumption made above,

the polar representation allows a much better Euclidean separation of the

points into clusters than the Cartesian representation. A two-dimension fuzzy

c-mean technique is used for clustering, and the resulting clusters of center

points identifies the vessel segments that are present in the neighborhood.

Further fixing of these clusters is performed by merging close clusters and

removing very small clusters.

An initial calibre is estimated as the average of the calibers of the points

belonging to the cluster, while an initial vessel direction is extracted by a

Principal Components Analysis on the same set of points. In case of one

vessel present around the center of the circular lines, only one direction will

be found, and this will be the initial estimate for the tracking module. If one

or more relevant directions are found, the seed point is discarded. However,

when looking for new tracking paths beyond a critical point, all relevant di-

rections in the Hough transform are kept: seed points are placed at a suitable

distance from the critical point along each identified direction. (Fig. 6.1):

this ensures the possibility for the tracking algorithm to overcome bifurca-

tions and crossings.

6.4 Tracking

This module is based on a step by step analysis of consecutive linear scan

lines $SL$. A similar approach to segment tracking in retinal angiograms was

used by Sun [84], Zhou et al. [100] and later also by Tolias and Panas [89].

The position of the scan line in the image determines which image pixels

are under analysis. For the scan line $SL_i \subset \mathbb{Z}^2$, examined at step $i$, a scan

line analysis classifies the pixels and identifies all possible vessel profiles

$VP_i \subset \mathcal{C}_{2^2}$, where $\mathcal{C}_{2^2}$ is the class of all possible subsets of $\mathbb{Z}^2$. In general,

$n_i$ different vessel profiles may be identified on $SL_i$, since more than one

vessel or a “silver-wire” reflection (see below) or just random noise may be

present. Center $c_{ij}$ and calibre $d_{ij}$ of every $j^{th}$ ($j = 1 \ldots n_i$) vessel profile $vp_{ij}$
6. Vessel Tracking

Figure 6.1: Bubble analysis at a bifurcation: correct vessel branches are identified

are determined.

Among all the \( n_i \) vessel profiles \( vp_{ij} \) detected on \( SL_i \), a vessel profile selection function \( P \) selects \( vp_i \), the sole profile to be associated with \( SL_i \), i.e., the one belonging to the vessel being tracked, using information on centers and calibers detected on the previous scan line \( SL_{i-1} \): \( vp_i = P (\mathcal{VP}_i, c_{i-1}, d_{i-1}) \)

Scan line updating, i.e., the determination of the new scan line \( SL_{i+1} \) given the past history, is determined by a scan line updating function \( U \), which determines the position of \( SL_{i+1} \) using information on centers and calibers detected on the previous and current scan lines, \( SL_{i+1} = U (c_i, c_{i-1}, d_i, d_{i-1}) \).

Scan line analysis

The pixel classifier adopted here is based on a Fuzzy C-Means (FCM) clustering algorithm \([9, 89]\). The intensity values of the pixels \( p_k \) of a generic scan line \( SL \) are examined and classified into two classes: “vessel” \((v)\) and “non-vessel” \((\neg v)\). The classifier returns a vector \( \{P^v_k(SL)\} \), which contains the probability of each pixel \( p_k \) in \( SL \) of being “vessel” (the probability of being “non-vessel” can be evaluated by taking the complementary to the previous one, i.e., \( P^v_k(SL) = 1 - P^\neg v_k(SL) \)). The classification of pixel \( p_k \) is then done by applying a threshold to its probability \( P^v_k(SL) \).

The scan line is analyzed to determine all the sequences of consecutive pixels classified as “vessel”, leading to a first set of candidate vessel profiles \( \mathcal{VP}_i \), which is examined as follows:

- all vessel profiles \( vp_{ij} \in \mathcal{VP}_i \) that are smaller than \( \delta_0 \) are deleted from
6. Vessel Tracking

\[ \mathcal{VP}_i; \]

- all vessel profiles \((vp_{ij}, vp_{ik}) \in \mathcal{VP}_i\) that are closer than \(\delta_1\) are merged into a single profile;

- all vessel profiles with contrast \(\kappa\) lower than threshold \(\delta_2\) are deleted from \(\mathcal{VP}_i;\)

where \(\delta_0, \delta_1, \delta_2\) are user-selected threshold values and the contrast \(\kappa\) of a generic vessel profile \(vp\) on \(SL\) is defined as:

\[ \kappa (vp, I) = \frac{|\beta^v (vp, I) - \beta^{-v} (vp, I)|}{\max \{\beta^v (vp, I), \beta^{-v} (vp, I)\}} \quad (6.1) \]

where

\[ \beta^v (vp, I) = \frac{\sum_{k=1}^{m} P_k^v (SL) I(p_k)}{\sum_{k=1}^{m} P_k^v (SL)} \quad (6.2) \]

\[ \beta^{-v} (vp, I) = \frac{\sum_{k=1}^{m} P_k^{-v} (SL) I(p_k)}{\sum_{k=1}^{m} P_k^{-v} (SL)} \quad (6.3) \]

\(I(p_k)\) is the intensity of pixels \(p_k\) and \(m\) is the number of pixels in \(SL\).

The resulting set of vessel profiles \(\mathcal{VP}\) will not contain any profile with calibre lower than \(\delta_1\) or any pair of profiles separated by gaps smaller than \(\delta_0\). The first condition was meant to eliminate isolated or short sequences of noisy pixels classified as “vessel” by the FCM classifier, while the second condition was aimed at eliminating from the recognized profiles the isolated pixels incorrectly classified as “non-vessel”. Moreover, the second condition also takes care of the so-called “silver wire” effect, a light reflection often present at the center of large vessels; it results in a bright, center-line artefact on the vessel lumen, which would cause the erroneous recognition of two separate vessel profiles.

For each recognized vessel profile \(vp\), center and calibre are evaluated. Vessel center is defined as

\[ c (vp, I) = \frac{\sum_{k=1}^{m} P_k^v p_k}{\sum_{k=1}^{m} P_k^v} \quad (6.4) \]

and vessel calibre as

\[ d (v, I) = \max_{i,j} \|p_i - p_j\| \quad (6.5) \]
6. Vessel Tracking

Note that center evaluation does reach sub-pixel accuracy, while calibre is limited to pixel-accuracy, even though it considers pixel distances in Euclidean terms and not pixel-step terms.

Vessel profile selection

The correct vessel profile \( v_{pi} \), one for each scan line \( SL_i \), is provided by the profile selection function \( P \) defined as

\[
v_{pi} = P (\mathcal{V} \mathcal{P}_i) = \arg \min_{v_{p_{ij}} \in \mathcal{V} \mathcal{P}_i} s (v_{p_{ij}})
\]

where \( s \) is a score function evaluated for each of the vessel profiles \( v_{p_{ij}} (j = 1 \ldots n_i) \) identified on \( SL_i \). The proposed score function is a weighted combination of the absolute changes in calibre and direction with respect to reference calibre and direction signals \( d^*_i \) and \( \theta^*_i \), which are the 2-elements autoregressive filtered values of the calibre and direction series identified on previous scan lines. Score function \( s \) is defined as

\[
s (v_{p_{ij}}) = \frac{\left| \theta (v_{p_{ij}}) - \theta^*_{i-1} \right|}{w_\theta} + \frac{\left| d (v_{p_{ij}}) - d^*_{i-1} \right|}{w_d}
\]

where \( \theta (v_{p_{ij}}) (0 \leq \theta (v_{p_{ij}}) < 2\pi) \) is the direction defined by connecting the center of the current vessel profile, \( v_{p_{ij}} \), and the one selected on the previous scan line, \( c_{i-1} \); \( \theta^*_i \) is the reference direction; \( \theta_{i-1} \) is the direction defined by the centers of the vessel profiles selected on the \( (i - 2)^{th} \) and the \( (i - 1)^{th} \) scan lines, i.e., \( c_{i-2} \) and \( c_{i-1} \). Likewise, \( d (v_{p_{ij}}) \) is the calibre of the current vessel profile; \( d^*_i \) is the reference calibre; \( d_{i-1} \) is the calibre of the vessel profile selected on the \( (i - 1)^{th} \) scan line. The two parameters \( w_\theta \) and \( w_d \) are used to give more importance in the selection of the correct vessel profile either to constancy of direction \( (\theta) \) or to calibre regularity \( (d) \).

The reference values provided by the low-pass filtering have been introduced in order to force some regularity in vessel calibre and direction, preventing sharp variations to affect the subsequent steps in vessel profile detection.

A maximum value for the score function has been set and profiles for which this maximum value is exceeded are not considered for the selection. This was done in order to avoid tracking of artifacts when no vessel profile is actually present in the scan line. In this case, the detected vessel profiles...
might be either artifacts or neighboring vascular structures, which will differ markedly in terms of direction or calibre from the vessel being tracked and will therefore yield a high value of the score function.

**Scan line updating**

The new scan line $SL_{i+1}$ i.e., the one obtained moving $SL_i$ one step forward along the vessel, is defined as the set of pixels $p$

$$\{p \mid p(\lambda) = c_i + \rho (\cos \theta_i, \sin \theta_i) + \frac{\lambda}{2} (\cos (\theta_i + \pi/2), \sin (\theta_i + \pi/2)), -\lambda_{i+1} \leq \lambda \leq \lambda_{i+1}\}$$

(6.10)

where $\rho$ is the tracking step, i.e., the distance between $SL_i$ and $SL_{i+1}$, and $\lambda_{i+1}$ is the new scan line size. Scan line size $\lambda_{i+1}$ is dynamically set to twice the present reference caliber: $\lambda_{i+1} = 2d_i^*$, so that it can adapt to the dynamic variation of the vessel calibre size. This allows to always have scan lines containing both “vessel” and “non-vessel” pixels and to avoid to examine too many “non-vessel” pixels when the vessel calibre decreases. The tracking step $\rho$ is kept constant until the tracking algorithm reaches a termination point, i.e., when the profile selection function is not able to identify any profile. At this point, in order to overcome a stop that might be caused only by small local artifacts, the tracking step is iteratively decreased by 10% steps, until either tracking is restarted or the maximum number of attempts is reached and tracking ends.

In Fig. 6.2 a combination of the tracking procedure with the bubble analysis is shown: the tracking module follows a vessel until a critical point, in this case a crossing. Then it stops and the bubble analysis is ran, yielding new seed points with the correct estimated directions beyond the crossing point.

### 6.5 False Vessel Elimination

The procedure ?? may recognize vessels also in zone of very low image contrast, or where choroidal vessels appear beyond the retina, or if the system identifies retinal pigmentation changes as vessels. Removal of these false vessel has been taken care by a module that compares each extracted vessel with its neighborhood. Mean $\mu$ and variance $\sigma^2$ are evaluated both for the vessel and for its surrounding retinal background and their relative distances are
evaluated. The vessel is discarded as false if:

\[ d_\mu = \mu_{\text{vessel}} - \mu_{\text{background}} < 0.15 \cdot \mu_{\text{vessel}} \]  
\[ d_{\sigma^2} = \sigma^2_{\text{vessel}} - \sigma^2_{\text{background}} < 0.5 \cdot \sigma^2_{\text{vessel}} \]  

### 6.6 Connection

The tracking strategy may result in the splitting of a single vessel into two or more segments, because of the presence of critical areas where the tracking algorithm has stopped. These segments belonging to the same vessel are likely to have end-points which face each other and that are likely to be similar with respect to direction (modulus \( \pi \)), caliber and intensity. In order to identify those extrema candidates to be connected, we have devised a greedy algorithm that exploits four features: distance between two end-points, their respective caliber, their direction and the difference in their mean grey-level values, evaluated in a 5x5 neighborhood. Since the meaning of these features is very different, and it is difficult to robustly normalize and combine them, four score functions have been defined between every pair of end-points \( p_i \) and \( p_j \) in the vessel structure. Being \( \delta_{i,j} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \) the Euclidean distance between the two points, and denoting with \( \overline{g}_i \) the mean grey-level value in a 5x5 neighborhood around \( p_i \), and being \( \sigma_\delta, \sigma_\theta, \sigma_d \) scaling constants, the scores are:

\[ \Delta_{i,j} = e^{-\frac{\delta_{i,j}}{\sigma_\delta}} \]
\[ \Theta_{i,j} = e^{-\frac{\theta_i - \theta_j}{\sigma_\theta}} \]
\[ D_{i,j} = e^{-\frac{d_i - d_j}{\sigma_d}} \]
\[ G_{i,j} = e^{-|g_i - g_j|} \]  

Pairs of end-points having all four scores above minimum thresholds are considered for connection. The algorithm searches hierarchically for pairs with the greater \( \Delta \), and among them those with the greater \( \Theta \). The resulting pairs are connected, yielding to a unique vessel, then the matrices \( \Delta, \Theta, D \) are updated by removing the linked pairs. The procedure continues until there are no more end-point pairs having scores in (6.13) above the thresholds.
6.7 Performance Evaluation

The algorithm described in this chapter has been tested on the images of the DB60. Firstly, it has been evaluated the capacity of the algorithm to recognize, and in what measure, the vessels present in the retinal images. In Tab. 6.1 are reported the fraction of recognized vessel length, compared with a manually identified vessel network. The results show that most of the vessels are identified by the method. Secondarily, the ability of the algorithm to connect separate recognized segments belonging to the same vessel. It has been chosen to evaluate the connection capacity on the six major vessels, manually identified. Two indexes were computed, $i_{frac}$ and $i_{totconn}$. The first show how much the identified segments still fraction the vessels after the connection. This is computed as:

$$i_{frac} = \frac{n_{vessels}}{n_{segments}}$$  

(6.14)

where $n_{vessels}$ is here equal to 6, the number of major vessels considered, and $n_{segments}$ is the number of remaining segments after the connection belonging to the six vessels. The second index is the number of major vessels that are totally connected after the linking procedure, that is how many major vessels are described by a single segment.

6.8 Conclusions

In the chapter an algorithm for identifying the vessel network starting from a sparse set of seeds points has been described. The method is based on the bubble analysis for initialization and looking beyond critical point, whereas is based on a fuzzy clustering tracing for iteratively exploring the vessel network.

Performance evaluation shows that the algorithm is robust and reliable, even if no quantification of its specificity is available at the moment. Its main limit lies in that it has no ability in reconstructing the topography of the network, and it can not solve situation in which a blind clustering procedure is not able distinguish correctly between vessel pixels and non-vessel pixels. This is particularly evident when large silver wires run along an artery, or when an artery and a vein are intertwined along their course.
6. Vessel Tracking

<table>
<thead>
<tr>
<th>Image</th>
<th>Recognised Vessel Length Fraction</th>
<th>Image</th>
<th>Recognised Vessel Length Fraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.926</td>
<td>31</td>
<td>0.953</td>
</tr>
<tr>
<td>2</td>
<td>0.940</td>
<td>32</td>
<td>0.884</td>
</tr>
<tr>
<td>3</td>
<td>0.950</td>
<td>33</td>
<td>0.958</td>
</tr>
<tr>
<td>4</td>
<td>0.945</td>
<td>34</td>
<td>0.931</td>
</tr>
<tr>
<td>5</td>
<td>0.960</td>
<td>35</td>
<td>0.912</td>
</tr>
<tr>
<td>6</td>
<td>0.963</td>
<td>36</td>
<td>0.946</td>
</tr>
<tr>
<td>7</td>
<td>0.940</td>
<td>37</td>
<td>0.917</td>
</tr>
<tr>
<td>8</td>
<td>0.956</td>
<td>38</td>
<td>0.910</td>
</tr>
<tr>
<td>9</td>
<td>0.955</td>
<td>39</td>
<td>0.855</td>
</tr>
<tr>
<td>10</td>
<td>0.920</td>
<td>40</td>
<td>0.928</td>
</tr>
<tr>
<td>11</td>
<td>0.897</td>
<td>41</td>
<td>0.966</td>
</tr>
<tr>
<td>12</td>
<td>0.960</td>
<td>42</td>
<td>0.920</td>
</tr>
<tr>
<td>13</td>
<td>0.913</td>
<td>43</td>
<td>0.922</td>
</tr>
<tr>
<td>14</td>
<td>0.931</td>
<td>44</td>
<td>0.816</td>
</tr>
<tr>
<td>15</td>
<td>0.944</td>
<td>45</td>
<td>0.976</td>
</tr>
<tr>
<td>16</td>
<td>0.921</td>
<td>46</td>
<td>0.860</td>
</tr>
<tr>
<td>17</td>
<td>0.913</td>
<td>47</td>
<td>0.825</td>
</tr>
<tr>
<td>18</td>
<td>0.860</td>
<td>48</td>
<td>0.954</td>
</tr>
<tr>
<td>19</td>
<td>0.926</td>
<td>49</td>
<td>0.914</td>
</tr>
<tr>
<td>20</td>
<td>0.878</td>
<td>50</td>
<td>0.940</td>
</tr>
<tr>
<td>21</td>
<td>0.857</td>
<td>51</td>
<td>0.894</td>
</tr>
<tr>
<td>22</td>
<td>0.868</td>
<td>52</td>
<td>0.960</td>
</tr>
<tr>
<td>23</td>
<td>0.956</td>
<td>53</td>
<td>0.867</td>
</tr>
<tr>
<td>24</td>
<td>0.908</td>
<td>54</td>
<td>0.942</td>
</tr>
<tr>
<td>25</td>
<td>0.859</td>
<td>55</td>
<td>0.833</td>
</tr>
<tr>
<td>26</td>
<td>0.907</td>
<td>56</td>
<td>0.913</td>
</tr>
<tr>
<td>27</td>
<td>0.910</td>
<td>57</td>
<td>0.967</td>
</tr>
<tr>
<td>28</td>
<td>0.903</td>
<td>58</td>
<td>0.961</td>
</tr>
<tr>
<td>29</td>
<td>0.895</td>
<td>59</td>
<td>0.847</td>
</tr>
<tr>
<td>30</td>
<td>0.927</td>
<td>60</td>
<td>0.851</td>
</tr>
</tbody>
</table>

Table 6.1: Fraction of recognized vessel on the images of the DB60 database. The fraction is computed as the ratio between the length of the correctly recognized vessels and the length of the manually identified vessels.
## 6. Vessel Tracking

<table>
<thead>
<tr>
<th>Image</th>
<th>Vessel Connection Score</th>
<th>Major Vessels Totally Connected</th>
<th>Image</th>
<th>Vessel Connection Score</th>
<th>Major Vessels Totally Connected</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.750</td>
<td>4</td>
<td>31</td>
<td>0.857</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>0.750</td>
<td>4</td>
<td>32</td>
<td>0.857</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>0.545</td>
<td>4</td>
<td>33</td>
<td>0.857</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>0.857</td>
<td>5</td>
<td>34</td>
<td>0.750</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>0.667</td>
<td>4</td>
<td>35</td>
<td>0.857</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>0.857</td>
<td>5</td>
<td>36</td>
<td>0.545</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>0.666</td>
<td>4</td>
<td>37</td>
<td>0.666</td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td>0.857</td>
<td>5</td>
<td>38</td>
<td>0.666</td>
<td>4</td>
</tr>
<tr>
<td>9</td>
<td>0.666</td>
<td>4</td>
<td>39</td>
<td>0.666</td>
<td>4</td>
</tr>
<tr>
<td>10</td>
<td>0.666</td>
<td>3</td>
<td>40</td>
<td>0.666</td>
<td>4</td>
</tr>
<tr>
<td>11</td>
<td>0.600</td>
<td>3</td>
<td>41</td>
<td>0.857</td>
<td>5</td>
</tr>
<tr>
<td>12</td>
<td>0.666</td>
<td>4</td>
<td>42</td>
<td>0.545</td>
<td>2</td>
</tr>
<tr>
<td>13</td>
<td>0.545</td>
<td>4</td>
<td>43</td>
<td>0.587</td>
<td>5</td>
</tr>
<tr>
<td>14</td>
<td>0.857</td>
<td>5</td>
<td>44</td>
<td>0.857</td>
<td>5</td>
</tr>
<tr>
<td>15</td>
<td>0.750</td>
<td>4</td>
<td>45</td>
<td>0.857</td>
<td>5</td>
</tr>
<tr>
<td>16</td>
<td>0.857</td>
<td>5</td>
<td>46</td>
<td>0.545</td>
<td>2</td>
</tr>
<tr>
<td>17</td>
<td>0.600</td>
<td>4</td>
<td>47</td>
<td>0.600</td>
<td>3</td>
</tr>
<tr>
<td>18</td>
<td>0.750</td>
<td>4</td>
<td>48</td>
<td>1.000</td>
<td>6</td>
</tr>
<tr>
<td>19</td>
<td>0.857</td>
<td>4</td>
<td>49</td>
<td>0.750</td>
<td>4</td>
</tr>
<tr>
<td>20</td>
<td>0.857</td>
<td>3</td>
<td>50</td>
<td>0.600</td>
<td>4</td>
</tr>
<tr>
<td>21</td>
<td>0.666</td>
<td>5</td>
<td>51</td>
<td>0.666</td>
<td>4</td>
</tr>
<tr>
<td>22</td>
<td>0.857</td>
<td>5</td>
<td>52</td>
<td>0.857</td>
<td>5</td>
</tr>
<tr>
<td>23</td>
<td>0.600</td>
<td>4</td>
<td>53</td>
<td>0.857</td>
<td>5</td>
</tr>
<tr>
<td>24</td>
<td>0.600</td>
<td>3</td>
<td>54</td>
<td>0.600</td>
<td>3</td>
</tr>
<tr>
<td>25</td>
<td>0.666</td>
<td>4</td>
<td>55</td>
<td>0.666</td>
<td>3</td>
</tr>
<tr>
<td>26</td>
<td>0.545</td>
<td>4</td>
<td>56</td>
<td>0.600</td>
<td>3</td>
</tr>
<tr>
<td>27</td>
<td>0.750</td>
<td>5</td>
<td>57</td>
<td>0.857</td>
<td>5</td>
</tr>
<tr>
<td>28</td>
<td>0.666</td>
<td>3</td>
<td>58</td>
<td>0.750</td>
<td>4</td>
</tr>
<tr>
<td>29</td>
<td>0.600</td>
<td>3</td>
<td>59</td>
<td>0.750</td>
<td>4</td>
</tr>
<tr>
<td>30</td>
<td>0.666</td>
<td>3</td>
<td>60</td>
<td>0.750</td>
<td>4</td>
</tr>
</tbody>
</table>

|       | μ                       | 0.720                           |
|       | σ                       | 0.116                           |

Table 6.2:
6. Vessel Tracking

(a) Region of interest of a retinal image

(b) The tracking module identifies samples (cyan crosses) of a vessel until it founds a critical point (blue diamond)

(c) Bubble analysis run around the critical point: the black circles are the sampling lines, the yellow circles are the estimated vessel points

(d) New seed points (blue stars) with their directions (green lines) are estimated

(e) The tracking module identifies the vessel (cyan crosses) departing from the crossing from the identified seeds
6. Vessel Tracking

(a) Internal (blue) and external (green) region for vessel checking

(b) Identified region superimposed to a vessel

Figure 6.3: Identification of internal and external regions of a vessel for evaluating its reliability
6. Vessel Tracking
Optic Disc Identification

The optic disc (OD), which in fundus images usually appears as a round region brighter than the surrounding, is the image of the optic nerve. From it, the central retinal artery and vein emerge, to cover, with further branching, most of the retinal region. Locating the OD position in fundus images is quite important for many reasons. Many important retinal pathologies may affect the optic nerve. Since the OD may be easily confounded with large exudative lesions by image analysis techniques, its detection is also important to exclude it from the set of possible lesions. Moreover, OD detection is fundamental for establishing a frame of reference within the retinal image, and is thus important for any image analysis application. The detection of OD position is also a prerequisite for the computation of some important diagnostic indexes for hypertensive/sclerotic retinopathy based on vasculature, such as Central Retinal Artery Equivalent (CRAE) and Central Retinal Vein Equivalent (CRVE) [42]. Other techniques have been recently proposed, which try to exploit the information provided by the vessel structure, i.e., the fact that all retinal vessels originate from the OD. In Koozekanani et al. [47], an OD tracking technique was developed for OCT (Optical Coherent Tomography) images, using a tiered scheme based on the Hough transform, eigenimage analysis and geometrical analysis based on a vasculature model. Other techniques proposed on retinal fundus images are reviewed briefly in 7.1. Our proposed method is based on a model of the geometrical directional pattern of the retinal vascular system, which implicitly embeds the information on the optic disc position as the point of convergence of all vessels. However, the resulting method is not just based on the detection of the area of convergence of vessels (as in [40]), but rather on the fitting of a model with respect to the entire vascular structure.
7. Optic Disc Identification

7.1 Review of Available Methods

7.1.1 Methods based on Luminosity

Many techniques have been proposed to detect the OD, mainly based on its specific round shape and relatively high brightness, as compared to the rest of the fundus image (see e.g. [62, 79, 7, 50, 52, 48, 31, 68, 92]). These techniques, however, often fail on pathological images, where other regions of fundus may be characterized by round shape and/or elevated brightness, e.g. large exudative lesions, or where the optic disc may be obscured by hemorrhages or may have a dark pigmentation.

7.1.2 Method based on image decomposition

The approach proposed in [51] exploits the similarity among optic discs in retinal images acquired with the same magnification. It needs a set of images of optic disc, which are used as training set. On this image set, a set of eigenimages are calculated by a principal component analysis (PCA) method, which represents the basis of the disc space. Since the image that form the basis of this space have dimension similar to that of a typical optic disc, only regions of the retinal image of that dimension can be decomposed with this eigenimages. For every pixel of the fundus image, its neighborhood of the right dimension is considered and then projected on this disc space, and a distance between the original image and its projection is computed. The center of the optic disc is located at the point with the minimum distance.

7.1.3 Method based on the Hough transform

Since usually the optic disc appears as a roundish bright shape, lead some authors [7, 73] in exploiting this feature to identify the papilla in a retinal image. They use the Hough transform to detect pixels aligned in circles in an edge map of the fundus image. To reduce the computational complexity of the three dimensional search required (the coordinates of the circle center and its radius) by the Hough transform, they have to make some heavy assumption. In [7] a fixed radius for the optic disc is assumed, and no results of the algorithm are reported, whereas in [73] only a small region of the image is searched, identified with a priori knowledge on the expected location of the optic disc.
7. Optic Disc Identification

7.1.4 Method based on vessel crossing

In [40] is claimed that the only constant feature of the optic disc is that it is the convergent point of the blood vessel network. The method proposed there try to exploit this characteristic by finding the convergence of blood vessels using what the authors call fuzzy convergence. This is a voting based method, in which vessels are modelled with fuzzy segments, whose area contributes votes to its constituent pixels. Summing the votes coming from all the identified vessel segments, generates an image map in which every pixel has a value proportional to its convergence strength. After blurring and thresholding the obtained map, the highest convergence peak is retained as the optic disc center. In order to make the method more robust, the authors combine the fuzzy convergence approach with a supervised classification based on brightness. This method achieved 89% of correct identifications on an image data set developed within the STARE project and containing many pathological images (http://www.ces.clemson.edu/ahoover/stare).

7.2 A Model of the Vascular Tree

![Fundus image with superimposed the main arcades model (blue double parabola). Vessel inside the parabola in the temporal region bend toward the medial axis more evidently than those in the nasal region](image)

Figure 7.1: Fundus image with superimposed the main arcades model (blue double parabola). Vessel inside the parabola in the temporal region bend toward the medial axis more evidently than those in the nasal region.

Defining a directional model for retinal vessels requires the definition on
7. Optic Disc Identification

the whole image of a function:

\[ \vartheta^{\text{mod}}(x; y; p), \quad -\frac{\pi}{2} \leq \vartheta^{\text{mod}} \leq \frac{\pi}{2} \] (7.1)

which represents the preferential direction in any retinal image of a vessel present at point \((x, y)\). Vector \(p\) is the set of parameters defining the model and its positioning, and thus it will include the optic disc coordinates. By visual inspection of retinal fundus images (see a representative example in Fig. 2.1), 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 modelled 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 \(\Gamma\):

\[ \Gamma = \{(x, y) \mid ay^2 = x\} \] (7.2)

where \(a\) is the parameter governing the aperture of the parabolas (for sake of simplicity, let’s assume for the time being that the origin of the coordinate system is the vertex of the parabolas). 7.1 shows an example of one such locus overlapped to the retinal image. For a generic point \((x, y)\) belonging to locus \(\Gamma\), i.e., on the parabola, the directional model is expressed by the implicit equation:

\[ \vartheta^{\text{mod}}(x; y; p) = \arctan \left( \frac{\text{sgn}(x)\text{sgn}(y) \cdot 1}{2\sqrt{a \cdot |x|}} \right) , \quad (x, y) \in \Gamma, \quad x \neq 0 \] (7.3)

where function \(\text{sgn}(\cdot)\) returns the sign of its argument and vector \(p\) contains parameter \(a\). The above expression 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 parabolas) 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 towards the macula in the temporal region (left-hand side in Fig. 7.1), whereas in the nasal region this inward deflection happens at a much slower rate (right-hand side in Fig. 7.1). The tangent equation 7.3 was
7. Optic Disc Identification

thus extended to accommodate points outside by adding a correction term $d$:

$$d(x, y; p) = \frac{y - \text{sgn}(y)\sqrt{|x|}}{c(x)}$$  \hspace{1cm} (7.4)

$$x(x; p) = \frac{c_1}{1 + e^{-x}} + \frac{c_1}{1 + e^{-x}}, \quad c_1 > 0, c_2 < 0$$  \hspace{1cm} (7.5)

The numerator of 7.4 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 expression 7.5, which expresses the rate of divergence of the direction at any given $x$ coordinate. For increasing $|x|$, this rate tends towards the value of parameter $c_1$ for positive values of $x$ and $c_2$ for negative values of $x$. Values of $c_1$ and $c_2$ represent therefore the limit rates of convergence towards $\frac{\pi}{2}$ (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 the OD. Given a generic origin for the Cartesian coordinates 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 $(x_{OD}, y_{OD})$, as shown in Fig. 7.1, a translation transformation had to be applied to the model:

$$\begin{cases} 
x^* = x - x_{OD} \\
y^* = y - y_{OD}
\end{cases}$$  \hspace{1cm} (7.6)

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

$$\vartheta^{mod}(x, y; p) = \arctan \left( \frac{\text{sgn}(x^*)\text{sgn}(y^*)}{s\sqrt{a|x^*|}} + \frac{y^* - \text{sgn}(y^*)\sqrt{\frac{|x^*|^3}{a}}}{\frac{c_1}{1 + e^{-x^*}} + \frac{c_1}{1 + e^{-x^*}}} \right)$$  \hspace{1cm} (7.7)

Fig. ?? shows an example of one such model overlapped to the retinal image. For sake of clarity, directions $\vartheta^{mod}$ are shown only for some points of the image and optimal values for model parameters are used for this simulation.
7.3 Parameter Estimation Via Simulated Annealing

Figure 7.2: $c_1$ and $c_2$ (set equal in these figures) can accommodate different vessel bending behaviour. In blue is shown the double parabola and in red the model direction at sample points.

By using suitable model parameter identification techniques, the optimal value for $p$, and thus for $(x_{OD}, y_{OD})$, can be identified for any image, given a set of data. The data are the vessel directions measured at points $(x_i, y_i)$, $i = 1, \ldots, N$, belonging to the vascular structure. Most of the algorithms developed for extracting the vascular structure (see e.g. [89, 41, 28]), provide, in addition to parameters such as vessel center-point position and caliber, also vessel direction at the center-point. This latter parameter can however be easily recovered from the identified vessel, e.g. by means of simple Principal Component Analysis on a set of vessel center-points. The detected vascular tree can therefore be represented by a set of quadruplets $(x_i, y_i, c_i, \vartheta_i)$, whose elements represent respectively the coordinates of vessel center-point $(x_i, y_i)$, the vessel caliber $c_i$ and vessel direction $\vartheta_i$ at that point. Our choice for the identification of model parameters has been the minimization of the weighted
7. Optic Disc Identification

![Figure 7.3: Different values of $c_1$ and $c_2$ can model the different vessel behaviour on the temporal and nasal part of the fundus image. In blue is shown the double parabola and in red the model direction at sample points.](image)

residual sum of squares RSS:

$$RSS = \sum_i w_i \left( \vartheta_i - \vartheta^{mod}(x_i, y_i; p) \right)^2$$  \hspace{1cm} (7.8)

Minimization is performed with respect to model parameters $p$ and operator $\mod$ 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 $(\hat{x}_{OD}, \hat{y}_{OD})$ represent the best positioning of the OD according to the model fit on the available data $(x_i, y_i, c_i, \vartheta_i)$. Minimization of $RSS$ with classical gradient-based techniques is rather critical, since this function exhibits many local minima. Fig. 7.4 represents 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 over-
7. Optic Disc Identification

Table 7.1: Optic disc identification performance on the STARE images

<table>
<thead>
<tr>
<th></th>
<th>Vessel Identified as in [41]</th>
<th>Vessel Identified as in Chap. 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct</td>
<td>79</td>
<td>79</td>
</tr>
<tr>
<td>Correct Fraction</td>
<td>0.97</td>
<td>0.97</td>
</tr>
</tbody>
</table>

come this problem, a Simulated Annealing (SA) optimization algorithm has been adopted. Simulated Annealing is a global stochastic optimization algorithm that theoretically guarantees the convergence towards global minimum [?]. 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.

7.4 Performance Evaluation

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 a representative subset of 20 images of the STARE data set. The resulting set of values was then used for model parameters estimation in all other images, both from the DB60 and the STARE data base. The optic disc is considered correctly identified if the estimated center is within the visible borders of the papilla.

7.4.1 The STARE Images and Vessel Data

The STARE project makes its images and results available on the web (http://www.ces.clemson.edu/ahoover/stare) upon request. In order to have comparison with the results reported in [40], the images and the vessel structures identified with the methods described in [41] were obtained courtesy of A. Hoover. The algorithm proposed here is able to correctly estimate the centre of the optic disc in 79 images on 81. As further assessment of the reliability and robustness of this method, the tracking algorithm described in Chap. 6 was run on these images, and the vessel structures used for estimating the optic disc centre. Even in this case the estimated centre is within the optic disc in 79 image, with the two missed that are different from those using the STARE vessel data. In Fig. ?? are shown the four images (two per tracking method) with the missed centre.
7. Optic Disc Identification

<table>
<thead>
<tr>
<th>Correct</th>
<th>DB60</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct Fraction</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Table 7.2: Optic disc identification performance on the DB60 images

7.4.2 DB60

The tracking algorithm described in Chap. 6 was run on the images of the DB60, and the vessel structures used for estimating the optic disc centre. In all images the estimated centre is within the optic disc margins.

7.5 Conclusions

A novel method for estimating the optic disc centre has been presented. It is based on a model of the vascular structure, which links the model vessel direction in every point on the image with the measured direction of the vessel samples.

It proved robust both with respect to the vessel extraction procedure used and to the image type. It estimated the optic disc centre within the human-assessed papilla borders in 89% of the images of the STARE database and in 100% of the images of the DB60 image set.

It is worth to be noted that the model does not take into account the fact that the main vessels are supposed to lie close to the locus $\Gamma$, and that the optic disc center is at the convergence of the largest vessels of the retina. This can lead to some convergence problem, when only the nasal or temporal network is visible and the vessel network directions are symmetric around the macula (vessels structure resembles more an ellipse than a parabola). This yields to a very prominent minimum of the objective function at a point symmetric to the optic disc with respect to the fovea.

The last thing to note is that the method does not evaluate the optic disc shape or dimension. This has been a topic of many works [6, 48, 52, 56, 62], but no methods has still been proven robust to pathological images such those in the STARE database. The choice adopted in the thesis was to approximate the optic disc diameter as ten times the largest vein in the image. This assumption comes form the widespread assumption that the main vain at the optic disc margin has a calibre of approximately $125\mu m$, and that a normal optic disc diameter is $2500\mu m$, even if it has been suggested that higher value should be used [21], and that is significantly varies with gender and race [63].
7. Optic Disc Identification

(a) RSS value as a function of $x_{OD}$ and $y_{OD}$ only, the other parameters fixed at their estimated optimal value

(b) RSS value as a function of $c_1$ and $c_2$ only, the other parameters fixed at their estimated optimal value
7. Optic Disc Identification

(a) Vessel samples with their direction

(b) \((x_{\text{od}}, y_{\text{od}})\) slice of the 4D objective function superimposed to the retinal image

(c) Estimated model directions on an arbitrary grid

(d) Estimated model direction on the vessel samples

Figure 7.5:
7. Optic Disc Identification

(a) Vessel samples with their direction

(b) \((x_{od}, y_{od})\) slice of the 4D objective function superimposed to the retinal image

(c) Estimated model directions on an arbitrary grid

(d) Estimated model direction on the vessel samples

Figure 7.6:
Vessel Structure Analysis

With hypertension, the retinal vasculature is initially protected from increases in blood pressure by metabolic and myogenic mechanisms. As the disease progresses, the endothelium, the muscularis and finally the entire vessel wall are damaged, causing an impairment of the contractile motility and an alteration of the blood-barrier function. Generalized narrowing of retinal arterioles is usually proportional to the degree of elevation of blood pressure [76], and it may or may not be present with mild hypertension, whereas is usually pronounced in severe hypertension. The increased hydrostatic pressure inside the retinal vasculature, also accounts for the tortuosity of arterioles and venules [30]. In more severe hypertensive states, irregularities in the caliber of blood vessels may be superimposed upon generalized narrowing. These are due to localized spasm and contraction of the wall of the arterioles, and appears as a focal narrowing of the vessel blood column. In chronic stages of hypertension, crosses abnormalities appear. These are caused by the enlargement of the retinal arteriolar wall by atherosclerosis, which compresses the crossing vein at their common adventitial sheath [76, 57]. Arterial branches have been shown to have such an architecture as to minimize sheer stress variation along the vascular network. This optimality is supposed to break with hypertension and atherosclerosis [83, 14].

To provide a sensible and robust diagnostic and prognostic tool to evaluate the retinopathy, all the aforementioned vascular changes have to be identified and evaluated: this is the aim of the methods described in the present chapter.
8. Vessel Structure Analysis

8.1 Arteries and Veins Separation

Changes in vessel structure can affect very differently arteries and veins. Tortuosity has a very different course in arteries and veins while retinopathy progresses, and some lesions are peculiar of only one type of vessels, e.g. focal narrowing for the arteries, beading for the veins. Moreover, one of the early signs of retinopathy is the so called *generalized arteriolar narrowing*, in which a ratio between arteries and veins diameters decreases.

In order to realize an automatic tool for the diagnosis and grading of retinopathy, it is mandatory to automatically distinguish between arterial and venous vessels (A/V classification). The presence of inter- and intra-image contrast, luminosity and color variability is a first problem, whereas the fading of the differences between the two types of vessels in the periphery of the retina is a second, but of no lesser impact, problem. Even after image contrast and luminosity normalization, A/V may be recognized reasonably well only in an area around the optic disc: inside the optic disc, vessels become so intertwined that it is very difficult even for an expert to track a vessel, whereas in the periphery of the image (far from the optic disc) they become thinner and thinner and almost undistinguishable. Moreover, even around the optic disc only vessels close to each other can be reliably recognized as arteries or veins, whereas vessels far apart from each other can be easily misclassified without any further knowledge other than their image features (see 8.2 and 8.1).

The retinal vessel network has a specific structure: the main vessels emerge at the optic disc and then follow a double-parabolic path by branching and thinning (see e.g. Fig. 7.1). This structure applies to both arteries and veins and we can assume that, at a small distance from the optic disc border, both types of vessels are randomly and homogeneously distributed around the optic disc.

These observations led us in our strategy to develop a reliable A/V classification technique. To begin with, we decided to classify the vessels only in a well defined concentric zone around the optic disc. Then, by using the vessel structure reconstructed by tracking techniques, we would propagate this classification outside this zone, where little or no information is available to discriminate arteries from veins. Second, the A/V classification algorithm would not be designed considering all together the vessels in the zone, but rather partitioning the zone in four quadrants and working separately and locally on each of them.
8. Vessel Structure Analysis

Figure 8.1: Inter-class variability: two arteries and one vein taken at different location of the same image.

8.1.1 Image preprocessing

Retinal images suffer quite often from inhomogeneity in luminosity and contrast, both within the same image (intra-image variability) as well as between images (inter-image variability). In order to obtain meaningful brightness information, it is necessary to compensate for this variability. To this end, we employed a previously developed algorithm [35], which analyzes the retinal background area to detect changes of luminosity and contrast, and through an estimation of their local statistical properties derives a compensation for their drifts.

The first task in fundus image analysis is to extract the vessel network, often through a vessel tracking procedure. It provides a set of vessel segments, usually described by a set $S$ of $n$ points $p$, representing the $x$ and $y$ Cartesian coordinates of samples of the vessel center line and the vessel diameter $d$ at
8. Vessel Structure Analysis

Figure 8.2: Intra-class variability: three arteries taken at different location of the same image.

these locations:

\[ p_i = (x_i, y_i, d_i) \quad (8.1) \]
\[ S = \{p_i : i = 1 \ldots n\} \quad (8.2) \]

In our study, the vessel network has been automatically extracted by a sparse tracking algorithm previously developed [27].

8.1.2 Divide

In order to exploit the local nature of the A/V classification procedure and the symmetry of the vessel network layout, we partitioned the retina in regions, in which it is more likely to have a similar number of veins and arteries, and in which the two types of vessels hopefully have different features. A concentric zone around the optic disc was identified and then partitioned into four regions, each containing one of the main arcs of the A/V network: superior-temporal (ST), inferior-temporal (IT), superior-nasal (SN) and inferior-nasal (IN), e.g. in Fig. 8.3.

To perform this partitioning, we first needed to identify the position of the optic disc and its approximate diameter; this can be done manually or
8. Vessel Structure Analysis

Figure 8.3: Principal arcades of the retinal vessel network

automatically, e.g. as in [41] or [75]. Then we identified the cardinal axes that divide the retina into the four quadrants \( \text{Quad}_i \), \( i = 1 \ldots 4 \) that contain the above mentioned main vessel arcs. Given the center of the optic disc \( c_{OD} \) and its radius \( r_{OD} \), the concentric zone \( C \) delimiting the region of analysis was centered in \( c_{OD} \) and had inner radius \( w_i \cdot r_{OD} \) and outer radius \( w_o \cdot r_{OD} \), with \( w_i = 1 \) and \( w_o = 4 \) (see Fig. 8.4).

For each quadrant \( \text{Quad}_i \), the 5 largest tracked vessels \( S_1, \ldots, S_5 \), i.e. the vessels that had the largest mean diameter, were automatically selected by the algorithm. Therefore, for each quadrant we had a set \( S_{\text{Quad}_i} = \{ S_1, \ldots, S_5 \} \) of vessels under analysis. An example is given in Fig. ??

This selection was performed to restrict the analysis to main vessels and avoid the possibly confusing information coming from small arterioles and venules. The balanced presence of veins and arteries in each of the four quadrants holds only if main vessels and their branches only are considered.

8.1.3 Data Extraction

For each quadrant, the set \( \mathcal{P} \) to be classified is represented by the sample points describing the vessels in \( S_{\text{Quad}_i} \) in \( C \).

\[
\mathcal{P}_{\text{Quad}_i} = \{ p_{i,j} = [x_{i,j}, y_{i,j}, d_{i,j}] \mid p \in S_{\text{Quad}_i} \cap C \} \tag{8.3}
\]
8. Vessel Structure Analysis

For each selected sample point $p_{i,j}$ in $\mathcal{P}_{\text{Quad}_{i}}$, a circular region $CR_{i,j}$ of diameter $d_{CR} = 0.8 \times d_{i,j}$ is identified:

$$CR_{i,j} = \{ p = [x, y] \mid \sqrt{(x - x_{i,j})^2 + (y - y_{i,j})^2} \leq d_{CR} \}$$ (8.4)

The triplet of values representing the RGB intensities of all pixels in $CR_{i,j}$ can also be expressed in another color space that more closely relates to a perceptual difference, such as the hue (H), saturation (S) and luminance (L) space. A non-linear mapping was used for this transformation [34]. A great deal of features can be extracted from the RGB and HSL values of the set $CR_{i,j}$. We performed an extensive statistical analysis to identify the features most discriminant for the A/V classification of $p_{i,j}$. The result was that the variance of red values $\text{Var}[R(CR_{i,j})] = R_{i,j}$ and the mean of hue values $E[H(CR_{i,j})] = H_{i,j}$ emerged as the best features to classify $p_{i,j}$ as belonging to an artery or to a vein. The fact that the classes of veins and of arteries are distinguished by looking at their average hue and homogeneity of their red component is also in agreement with clinical experience: when two vessels close to each other are compared for classification, the one with the darker red is classified as vein; if this difference is not significant enough, the one with the highest degree of uniformity is classified as vein.

8.1.4 Impera

After features extraction, we have four sets of data $F_{i}$, each containing a $2 \times N_{i}$ matrix with variance of red intensities and mean of hue intensities for the $N_{i}$ points in $CR_{i,j}$. The $j^{th}$ position in $F_{i}$ corresponds to the point $p_{i,j}$ and is:

$$F_{i}(j) = [R_{i,j}, H_{i,j}]$$ (8.5)

For each set $F_{i}$, the centers of the two classes are identified via a fuzzy clustering algorithm [9]):

$$c_{1} = [R_{1}, H_{1}] \text{ for arteries}$$ (8.6)

$$c_{2} = [R_{2}, H_{2}] \text{ for veins}$$ (8.7)

with $R_{1} > R_{2}$.

The Euclidean distance of each point from the centers is the criterion used for the classification:

$$p_{i,j} \in \text{Artery} \iff \| F_{i}(j) - c_{1} \| \leq \| F_{i}(j) - c_{1} \|$$ (8.8)
After all points are assigned to either of the two classes, an empirical probability $P$ for each vessel to be an artery (or vein) can be determined. Being $p_v$, the number of points $p_{i,j}$ of a vessel $S$ assigned to the vein class and $p_a$ the number of points assigned to the artery class, these probabilities are:

$$P[S \epsilon \text{ Artery} \mid \{p_i, i = 1 \ldots n\}] = \frac{p_a}{p_v + p_a}$$ (8.9)

$$P[S \epsilon \text{ Vein} \mid \{p_i, i = 1 \ldots n\}] = 1 - \frac{p_a}{p_v + p_a}$$ (8.10)

These values are then used to classify the vessel, according to the class having the higher probability.

### 8.1.5 Results and Discussion

In Fig. 8.7 features for all the points $p_{i,j}$ inside the concentric zone are plotted; in Fig. 8.8 features for the points of one quadrant only are plotted. It is quite evident that the two clusters of Fig. 8.8 are much more well defined than those of Fig. 8.7 and thus the strategy of partitioning the retina in four quadrants greatly improves the classification.

The A/V classification results obtained by the proposed algorithm on the whole set of 35 images were compared to those provided by a manual classification on the same vessels performed by an expert. The classification performances are shown in Table 8.1, where they are reported both for the whole data set and separately for the main vessels, i.e. those belonging to the main vascular arcades, and the secondary vessels. The origin of the latter ones is that it is possible that in some quadrant fewer than five segments belonging to the main vessels are present, in which case small arterioles and venules are also taken into account. Since they are quite similar to each other and their discrimination is not very significant from a clinical point of view, we can accept a higher error rate for them. A correct classification of the main vessels, on the contrary, has a far greater impact on the further diagnostic analysis of the retinal fundus.

<table>
<thead>
<tr>
<th></th>
<th>Number of vessels</th>
<th>Misclassifications</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>642</td>
<td>85</td>
<td>13.2 %</td>
</tr>
<tr>
<td>Main Vessels</td>
<td>384</td>
<td>27</td>
<td>7.0 %</td>
</tr>
<tr>
<td>Secondary Vessels</td>
<td>258</td>
<td>58</td>
<td>22.4 %</td>
</tr>
</tbody>
</table>

Table 8.1: Vessel classification performance

It is worth noting that going through the classification of single points,
8. Vessel Structure Analysis

and then of whole vessels, provides a more robust classification strategy, since small local variations in chromaticity/luminosity or vessel tracking errors have a reduced impact on the overall classification. An example of resulting A/V classification is shown in Fig. 8.9
Figure 8.4: Division of the fundus

Figure 8.5: Selected vessels $\bigcup_{i=1}^{5} S_{Quad_i} \cap C$ in a sample image. In different colour vessels belonging to different quadrants set $S_{Quad_i}$. 

95
8. Vessel Structure Analysis

Figure 8.6: Selected vessels $\bigcup_{i=1}^{5} S_{Quad_i} \cap C$ in a sample image

Figure 8.7: Features for the points from the vessels in $\bigcup_{i=1}^{5} S_{Quad_i} \cap C$. Circles are artery sample points, crosses are vein sample points.
8. Vessel Structure Analysis

Figure 8.8: Features for the points from the vessels in $S_{Quad_1} \cap C$. Circles are artery sample points, crosses are vein sample points.

Figure 8.9: Result of the A/V classification achieved by the proposed algorithm on a sample image; white vessels were classified as artery, black ones as vein.
8. Vessel Structure Analysis

8.2 Vessel Tortuosity

One of the first changes to occur in the vascular network of the retina as consequence of many diseases is the increasing in vessel tortuosity. In order to evaluate clinical significance of tortuosity changes with time, or to compare different levels of retinopathy, there is a strong need to develop a tortuosity measure that matches the clinical perception of ophthalmologists. At first, the main factors that influence the classification of a vascular structure as tortuous or non-tortuous will be outlined. This is most peculiar in analyzing retinal images, in which not only straight vessels but also long vessels presenting a smooth semi-circular shape are considered as non-tortuous. The need to develop a numerical index incorporating these factors arose from the observation that previously proposed methods (see [36] for a review) failed in differentiating the tortuosity of structures that visually appeared to be very different in tortuosity, as it will be shown in Sec. 8.2.2.

8.2.1 Tortuosity Properties

Vessel tortuosity measures do not have a formal clinical definition, but common practice has outlined an evaluation of the tortuosity with some well-defined properties. In order to obtain a clinically meaningful vessel tortuosity measure, i.e. a measure that is able to match the ophthalmologist’s evaluation, it is necessary to make these properties explicit. Then, using a proper mathematical formulation of the tortuosity functional, the proposed measure must be shown to satisfy these properties.

Affine Transformations

A sound approach to extract properties of a not formally defined index is to perform a preliminary study on the invariance properties of such index with respect to the most relevant transformation in the domain space. In particular, we will consider affine transformations of a vessel: translation, rotations and scaling.

Translation and rotation transformation are not supposed to influence the perception of tortuosity. In fact, these transformation are related to the geographical position and orientation of vessels in the retina, and do not alter in any way the clinical perception of tortuosity. Scaling of the single vessels does not seem to affect the clinical perception of tortuosity, but a warning flag should be raised since different clinicians may have different opinions about this issue. The scaling is particularly controversial when considering
also the vessel calibre, but for the purpose of evaluating single vessel tortuosity, its perception can be safely considered invariant to scaling.

Composition

Composition properties deal with the behavior of tortuosity perception when two vessel curves are merged into a single curve, or when various segments of the same vessel, in general with different tortuosity measures, build up to give the total vessel tortuosity.

Given two adjacent continuous curves \( s_1 \) and \( s_2 \), we define the combination of the two as:

\[
\mathcal{L} = s_1 \oplus s_2
\]

Since the two composing curves are supposed to belong to the same vessel, we can assume the continuity of \( \mathcal{L} \) without loss of generality.

In [36], an intuitive empirical principle was proposed:

\[
\tau(s_1) \leq \tau(s_2) \Rightarrow \tau(s_1) \leq \tau(s_1 \oplus s_2) \leq \tau(s_2)
\]

(8.11)

which means that, when composing two curves, the resulting tortuosity is between those of the composing curves. The counterexample shown in Fig. 8.10, top panel, clearly show that this statement can not be accepted in conjunction with the principle of invariance with respect to rotation and scale: two curves (e.g., \( L_1 \) and \( L_2 \)) perceived by themselves as almost non tortuous, when connected form an undoubtedly tortuous curve.

Therefore we propose a new composition property, such that a vessel \( s \), combination of various segments \( s_i \), will not have tortuosity measure less than any of its composing parts:

\[
\tau(s_i) \leq \tau(s_1 \oplus s_2 \oplus \ldots \oplus s_n)
\]

\( \forall i = 1 \ldots n | s_i \subseteq s \)

(8.12)

Modulation

It is useful now to express a monotonic relationships with respect to two other properties, which we will call frequency modulation at constant amplitude and amplitude modulation at constant frequency. It may be assumed that the greater the number of changes in the curvature sign (twist) is, the more tortuous the vessel can be considered. Similarly, the greater the amplitude (maximum distance of the curve from the underlying chord) of a twist is, the greater is the tortuosity associated with it.
8. Vessel Structure Analysis

For two vessels having twists with the same amplitude, the difference in tortuosity varies with the number of twists $\varphi$:

$$\varphi(s_1) \leq \varphi(s_2) \Rightarrow \tau(s_1) \leq \tau(s_2) \quad (8.13)$$

and, conversely, for two vessels with the same number of twists (with the same frequency), the difference in tortuosity depends on the difference in amplitude $\alpha$ of the twists:

$$\alpha(s_1) \leq \alpha(s_2) \Rightarrow \tau(s_1) \leq \tau(s_2) \quad (8.14)$$

8.2.2 Available tortuosity measures

Various tortuosity measures have been proposed in the literature, but all fails in certain aspects. In this section we will review the available methods for evaluating retinal vessel tortuosity and will present some counterexamples in which these methods provide results that do not match with clinical perception.

**Arc Length over Length ratio**

The simpler and most widely used measure of a vessel tortuosity is the ratio between its length and the length of the underlying chord [36] [37] [59]. The idea of using this ratio is that the greater the value of the ratio, the more distant the vessel is from a straight line, i.e., tortuous. Unfortunately, being the surface of the retina close to a semi-sphere, the non tortuous paradigm should be the circle arc. In fact, every vessel that has a constant and small curvature, regardless of the amplitude of the arc it describes (as for the main retinal vessels), will be regarded by an ophthalmologist as of negligible tortuosity. Moreover, it is shown in Fig. 8.10 that two vessels with very different tortuosity have the same arc length over length ratio measure.

**Measure involving curvature**

Hart [36] presented a number of tortuosity measures that involve the use of the integral of the absolute curvature or of the squared curvature. For a curve $s(l) = [x(l), y(l)] : D \subset \mathbb{R} \to \mathbb{R}^2, s \in C^1(\mathbb{R})$, with $l$ the curvilinear coordinate on it, the curvature $C_s(l)$ is:

$$C_s(l) = \frac{\dot{x}(l)\ddot{y}(l) - \ddot{x}(l)\dot{y}(l)}{(\dot{x}^2(l) + \dot{y}^2(l))^{3/2}} \quad (8.15)$$
Figure 8.10: The first two curves (top and middle panel) have very different tortuosity but the same Length $L$ and Chord Length $\chi$. The second and last curves (middle and bottom panel) have the same average angular difference despite their different tortuosity. The curve in the bottom panel has a curvature integral of $\pi/2$ whereas the one in the middle panel has curvature integral $\pi$, even if that in the bottom panel is clearly perceived as more tortuous.

The idea behind this is that this integral should be a measure of the variability of vessel direction. However, the example presented in Fig. 8.10 shows that a smaller curvature integral may correspond to a greater perceived tortuosity.

In our opinion, there are three main reasons for this result. The first is that curvature is non zero only along arcs, while it is negligible along straight or almost straight segments; but straight segments together with arcs dramatically change the tortuosity appearance. The second is that changes in convexity (curvature sign) of the curve are not taken into account. On the contrary, this is the main feature used by expert graders to assess tortuosity. Finally, integrating along domain possibly different in dimension yields to
8. Vessel Structure Analysis

measures depending on the aforesaid dimension. The ratios between the absolute curvature integral (or the squared curvature integral) and chord length (or vessel length) have been proposed to circumvent the latter point, but still the other two problems remain open with this definition of tortuosity.

Mean direction angle change

More recently, a measure of tortuosity based on a local directional changes of the vessel has been proposed [12]. It computes the average of the angles between sample points describing the vessel. For each point \(i\) of the vessel, it computes the versors from the previous \(d_{i-step}\) point and to the subsequent point \(d_{i+step}\):

\[
ad_s = \frac{1}{L_c - 2 \times \text{step}} \sum_{i=1}^{n} \arccos (d_{i-step} \cdot d_{i+step})
\] (8.16)

In addition to the high sensitivity to noise, both from the digital quantization and from the vessel extraction technique, it suffers of the problems mentioned above. Even the simple example of Fig. 8.10 shows that a vessel with constant curvature, such as a semi-circumference, and a vessel formed by the juxtaposition of two arcs of circumference have the same average angle variation, despite their difference in tortuosity. Moreover, also here vessel segments with no variation in direction do not weight in the tortuosity.

8.2.3 Available Data

The theoretical mathematical description of the vessel is a curve in a two dimension space:

\[
s(l) : D \subset \mathbb{R} \rightarrow \mathbb{R}^2
\] (8.17)

The available description of the vessel is a sampled, quantized and noise corrupted version of the theoretical curve:

\[
s(k) : s(l_k)
\] (8.18)

where \(l_k\) is a sequence of curvilinear coordinates that represent the sampling on the original vessel.

From the DB60, 30 major arteries and 30 major veins were identified and manually sampled, in order to eliminate tracking artifacts and therefore confounding the real tortuosity measure performances.
8. Vessel Structure Analysis

8.2.4 Vessel Representation

Sampling of a vessel may lead to a very sparse vessel description. Sparse description of the vessel further leads to a poor description of its dynamics, which may eventually fail to provide useful information on vessel direction and its derivatives. Sampling can be taken care of by means of an interpolation function, which describes the vessel between sampling points, where no data are present. Physiologically, vessels are structures continuous at least with their first derivative and this condition leads to the use of a cubic spline interpolation. Since the data are noisy, splines also offer the opportunity of filtering by using cubic smoothing splines.

The resulting spline curve is obtained from the available data by regularization:

\[ \hat{s} = \arg \min_{\hat{s}} \left[ (1 - \gamma) E_i(\hat{s}) + \gamma \| \hat{s} - s_q \| \right] \]  (8.19)

where \( E_i(s) \) is what is called internal energy term, and represents the smoothness of the curve; \( \gamma \) is a weighting parameter, varying between 0 and 1, that sets the compromise between following the available data \( s_q \) and having a 'smooth' behavior. For \( \gamma = 1 \) \( \hat{s} \) is the least square linear fit of the data, whereas for \( \gamma = 0 \) it is the variational cubic spline interpolant of the data. In this study we used \( \gamma = 0.005 \).

8.2.5 Tortuosity Calculation

When evaluating tortuosity, ophthalmologists integrate information about how many times a vessel twists (changes in convexity, or curvature sign), and how large is the amplitude of each of the recognized twist. We would therefore decompose any curve into a set of consecutive segments of constant curvature sign.

Having defined the curvature of a curve in (8.15), we define a subsegment \( s_i \) of \( s \) as a turn curve if:

\[ [C_{s_i}(l) \geq 0, \forall l \in D] \lor [C_{s_i}(l) \leq 0, \forall l \in D] \]  (8.20)

In real images, it is common to find small oscillations (changes of convexity with very small amplitude) around the main vessel direction, due to the presence of noise. These oscillations might affect the correct decomposition of the vessel, since they would create a great number of artificial turn curves, which are not clinically significant. An hysteretic threshold on curvature was thus used to deal with these small variations, as shown in Fig. 8.11. To eval-
8. Vessel Structure Analysis

Figure 8.11: Curvature value along a vessel. The dashed lines indicate the hysteretic threshold for the convexity change.

To evaluate tortuosity, we also need the Chord Length $L_\chi$ of a curve, defined as the distance between the two extremes of the curve:

$$L_\chi = ||s(\max(D)) - s(\min(D))||$$  \hspace{1cm} (8.21)

and the Curve Length $L_c$, defined as:

$$L_c = \int_{\min(D)}^{\max(D)} \left\| \frac{\partial s}{\partial l} \right\| \, dl$$  \hspace{1cm} (8.22)

An hysteretic threshold of 0.03 has been set for arteries and one of 0.01 for veins. The two types of vessels have been kept separate, since ophthalmologists grade arteries and veins differently. Due to the possible presence of turn curves with zero curvature, the decomposition is not unique. Since eliminating these turn curves might lead to the situation in which an increase in amplitude with straight segments would not increase the tortuosity, we chose to split the straight segments into two halves, assigning one to the preceding and one to the following turn curve.

Once a curve $s(l)$ is divided into $n$ turn curves

$$s_i : s = s_1 \oplus s_2 \oplus \ldots \oplus s_n$$  \hspace{1cm} (8.23)

we propose a new definition of vessel tortuosity as:

$$\tau(s) = \frac{n-1}{L_c} \sum_{i=1}^{n} \left[ \frac{L_{c_{s_i}}}{L_{\chi_{s_i}}} - 1 \right]$$  \hspace{1cm} (8.24)
This tortuosity measure has a dimension of $1/\text{length}$ and thus may be interpreted as a *tortuosity density*, allowing its comparison on vessels of different length. It is worth noting that when $n$ is equal to 1 then $\tau$ is equal to 0 and thus vessels with a constant convexity have zero tortuosity.

The proposed definition of tortuosity meets all the properties described in Sec. 8.2.1. In particular, the composition property is satisfied via the summation in (8.24), the amplitude modulation via the ratio of the length over the chord length for every turn curve, and the frequency modulation is taken care of both implicitly by means of the curve splitting and explicitly through the multiplicative term $n - 1$ in (8.24).

### 8.2.6 Results and Discussion

The 30 selected arteries and 30 veins have then been proposed to an expert ophthalmologist, who ordered them with increasing tortuosity. The proposed method was then compared with those presented in Sec. 8.2.2, for which, calling $\kappa$ the curvature of a curve, we define as in [36] the total curvature and the total squared curvature as the integral of $|\kappa|$ and $\kappa^2$, respectively.

Computation of $\kappa$ involves the evaluation of first and second derivatives along the curve, which was done both by numerical differentiation and by the same spline representation of the vessel used for the computation of the proposed measure $\tau$.

\[
\begin{align*}
tc &= \int_{\min D}^{\max D} |\kappa(l)| \, dl \\
tsc &= \int_{\min D}^{\max D} |\kappa(l)^2| \, dl
\end{align*}
\] (8.25)

In Table 8.2 values of the correlation with the clinician perception for all the tortuosity indexes mentioned in this paper are shown. The correlation is calculated on the order of vessels. In this way, possible non uniformity in the distribution of vessel tortuosity do not influence the correlation value. These results show that the proposed measure has the best correlation as regards both artery and vein tortuosity.
8. Vessel Structure Analysis

<table>
<thead>
<tr>
<th>Tortuosity Measure</th>
<th>Arteries</th>
<th>Veins</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L_c/L_\chi$</td>
<td>0.850</td>
<td>0.564</td>
</tr>
<tr>
<td>tc</td>
<td>0.889</td>
<td>0.741</td>
</tr>
<tr>
<td>tsc</td>
<td>0.863</td>
<td>0.707</td>
</tr>
<tr>
<td>$tc/L_c$</td>
<td>0.860</td>
<td>0.721</td>
</tr>
<tr>
<td>$tsc/L_c$</td>
<td>0.840</td>
<td>0.690</td>
</tr>
<tr>
<td>$tc/L_\chi$</td>
<td>0.883</td>
<td>0.745</td>
</tr>
<tr>
<td>$tsc/L_\chi$</td>
<td>0.863</td>
<td>0.707</td>
</tr>
<tr>
<td>ad</td>
<td>0.860</td>
<td>0.733</td>
</tr>
<tr>
<td>$\tau$</td>
<td>0.922</td>
<td>0.818</td>
</tr>
</tbody>
</table>

Table 8.2: Correlation of the tortuosity measures ordering with the ophthalmologist tortuosity order
8.3 Generalized Arteriolar Narrowing

The importance of generalized arteriolar narrowing goes well beyond the evaluation of retinopathy. Several studies have demonstrated a strong association of generalized arteriolar narrowing with blood pressure increase [42, 49, 46], and with carotid heart stiffness [53] and have identified in generalized arteriolar narrowing a risk factor for coronary heart disease [20]. The difficulty of its reliable estimate has excluded it from evaluation in all major studies [21], until the ARIC study devised a computer-aided method for quantitatively measures arterioles and venules around the optic disc (see below).

8.3.1 Proposed Measures

Generalized arteriolar narrowing is evaluated by computing the ratio between arteriolar width and venular width (A/V ratio). Normal ratio should be 0.75, and as hypertension increases, arteriolar spasm becomes widespread on the vessel network, yielding a reduced A/V ratio. Its evaluation is not practical and highly subjective, since the examiner should estimate what the original arteriolar caliber might have been or compare the current diameter of specific arterioles with those of approximately matching venules.

In order to quantify generalized arteriolar narrowing, Parr [70] postulated that several factors should be taken into account. Firstly, first-order arteries should be measured some distance away from the disc, where these vessels become unambiguously arteriolar rather than arterial. Secondly, since the vessel carrying capacity depends on its cross-section, arterioles should be described by a square function of their diameters. Finally, measurement of arterioles should account for branching pattern, since the total cross-section of the arteriolar system increases with each bifurcations.

The method described in [71, 72] reduces vessels which lay in a concentric zone around the optic disc, into a Central Retinal Arteriolar Equivalent (CRAE) and a Central Retinal Venular Equivalent (CRVE): these are used to compute the A/V ratio. To apply the method, a mapping of the branching pattern must be available. Then, every branch, starting with the more distal, is paired with its originating vessel, and its equivalent diameter calculated. The procedure iterate until all vessels have been taken into account. The equivalent caliber for every pair is computed as:

\[
\text{Arterioles } W_c = \sqrt{0.87W_a^2 + 1.01W_b^2 - 0.22W_a \cdot W_b - 10.76} \quad (8.26)
\]

\[
\text{Venules } W_c = \sqrt{0.72W_a^2 + 0.91W_b^2 + 452.05} \quad (8.27)
\]

\[
W_c = \sqrt{0.87W_a^2 + 1.01W_b^2 - 0.22W_a \cdot W_b - 10.76} \quad (8.28)
\]
8. Vessel Structure Analysis

where \( W_c \) is the equivalent caliber of the pair, \( W_a \) is the caliber of the smaller vessel and \( W_b \) is the caliber of the larger.

The ARIC Study devised a simplified procedure for combining the measurements to yield the central equivalents [3, 42]. The procedure combines the largest vessel with the smallest, then the next largest with the next smallest, and so on until all vessels are accounted for. This is carried on with a semi automatic procedure, both to identify vessel calibers, and to distinguish arterioles from venules. The tracking algorithm described in Chap. 6 provides the vessel caliber information, and the method in Chap. 7 gives the position and approximate dimension of the optic disc. Applying the algorithm described in Sec. 8.1, the main arterioles and venules around the optic disc are available for the ARIC procedure to be applied.

8.3.2 Results and Discussion

The evaluation of the central equivalents and of the A/V ratio was made on the DB60 data set both for the whole circular region around the optic disc, and also quadrant-wise. In Tab. 8.3 the correlation coefficients for some of the measures are shown. Combining some of these measures yield a slightly better correlation:

\[
\text{gan} = \frac{p(CRAE;90) + \beta p(CRAE;90) + \gamma p(CRAE;90) p(A/V;90)}{1 + \beta + \gamma} \tag{8.29}
\]

in which \( p(S; n) \) is the \( n \)th percentile of the set \( S \) of the four measures of each quadrant, and \( \beta \) and \( \gamma \) are heuristic parameters. The correlation obtained is:

<table>
<thead>
<tr>
<th>Measure</th>
<th>Absolute Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>A/V ratio</td>
<td>0.30</td>
</tr>
<tr>
<td>Median of the quadrant A/V ratio</td>
<td>0.18</td>
</tr>
<tr>
<td>90th percentile of the quadrant A/V ratio</td>
<td>0.38</td>
</tr>
<tr>
<td>CRVE</td>
<td>0.13</td>
</tr>
<tr>
<td>Median of the quadrant CRVE</td>
<td>0.23</td>
</tr>
<tr>
<td>90th percentile of the quadrant CRVE</td>
<td>0.15</td>
</tr>
<tr>
<td>CRAE</td>
<td>0.24</td>
</tr>
<tr>
<td>Median of the quadrant CRAE</td>
<td>0.06</td>
</tr>
<tr>
<td>90th percentile of the quadrant CRAE</td>
<td>0.43</td>
</tr>
</tbody>
</table>

Table 8.3: Correlation of the measures CRAE, CRVE and A/V ratio with the ophthalmologist evaluation of generalized arteriolar narrowing
8. Vessel Structure Analysis

<table>
<thead>
<tr>
<th>Measure</th>
<th>Absolute Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>gan</td>
<td>0.48</td>
</tr>
</tbody>
</table>

Table 8.4: Correlation of the measures obtained with the (8.29), $\beta = 0.1$ and $\gamma = 1$, with the ophthalmologist evaluation of generalized arteriolar narrowing.

The obtained correlations are nevertheless very low. This is a probable outcome of two concurrent sources of error. The first is related to the tracking procedure and the artery-vein separation algorithm. Misclassification of one major artery or vein may change substantially the result of the CRAE and CRVE computation, even if no sensibility analysis has been performed. The second source of error is related to the subjective nature of generalized narrowing evaluation of the ophthalmologist. The ophthalmologist evaluation used in this thesis was in no way computer aided, and no arterioles and venules caliber measures were taken. This yields a to estimation error in the visual inspection, especially in presence of arterial or venous sheathing. In any case, it will be required a more careful evaluation of generalized arteriolar narrowing automatic performances, and a tested reference evaluation needed.

8.4 Focal Arteriolar Narrowing

One of the first arterial lesions that can be identified in mild hypertensive retinopathy is the focal arteriolar narrowing. This appears as a local constriction of the vessel. The only guideline which defines the evaluation of focal arteriolar narrowing is the ETDRD [21], in which this lesion is considered definite only when the caliber reduction is of at least one third with respect both to the proximal and distal vessel calibers (see Fig. 8.11). Since vessels naturally diminish their caliber in their course from the optic disc toward the periphery, this trend has to be accounted for, in order to provide a reliable algorithm for the detection of focal narrowing. After estimating the natural vessel tapering trend, a set of candidate narrowed region is extracted. This regions are then classified by a finite-state machine to obtain the estimated location of the focal narrowing sites. The severity of the lesion is then evaluated.
8. Vessel Structure Analysis

Figure 8.12: Evaluation of Generalized Arteriolar Narrowing as computed in 8.29, $\beta = 0.1$ and $\gamma = 1$

8.4.1 Available Data

The vessel network has been automatically extracted by the algorithm described in Chap. 6. A set of vessel segments, described by a set $S$ of $n$ points $p_i$, representing the $x$ and $y$ Cartesian coordinates of samples of the vessel center line and the vessel diameter $d$ at these locations has been obtained:

$$p_i = (x_i, y_i, d_i) \quad (8.30)$$

$$S = \{p_i : i = 1, \ldots, n\} \quad (8.31)$$

Due to the nature of the lesion to be evaluated, only a subset of the recognized vessel is suitable to be analyzed. Given the image resolution, vessel with mean caliber smaller than 8 pixel are discarded, because they do not provide enough signal to noise ratio to detect a change of one third of the caliber (caliber variation in the order of the pixel or less should be detected). Given the $\text{pixel/\mu m}$ factor of the images in the DB60, this results in discarding all vessels thinner than approximately 65 $\mu m$. This is not so far from the data
8. Vessel Structure Analysis

Arteriolar Focal Narrowing

Figure 8.13: Focal Arteriolar Narrowing scheme

analyzed in [46], where discarded, but the images are taken at 30°.

8.4.2 Focal Narrowing Identification and Evaluation

Vessels have a physiological tapering in their course form the optic disc toward the periphery. This could lead to incorrect estimation of caliber narrowing. This decreasing trend has to be eliminated before any meaningful analysis on caliber irregularities can be performed. Therefore the parameters \((a, b)\) of the linear interpolation of caliber data \(d = \{d_1 \ldots d_n\}\) are estimated. Given the vessel curvilinear abscissa \(l = \{l_1 \ldots l_n\}\), the interpolation is:

\[
\hat{d}_i = a \cdot l_i + b \quad (8.32)
\]

The onset of a abnormal narrowing is considered when the caliber \(d\) reduces below \(\frac{5}{6}\hat{d}\). This is justified by the fact that significant focal narrowing is considered only when a reduction in caliber of \(\frac{1}{3}\) is present, and its onset may be safely set at half of the transition zones between narrowing and not narrowing. This is estimated as the location in which the vessel caliber is
less than the onset threshold (see Fig. 8.13):

\[
\text{th}_\text{onset}(l) = (1 - \frac{1}{2\Delta}) \hat{d}(l) = \frac{5}{6} \hat{d}(l)
\]  

(8.33)

Each vessel region, composed by 1-connected vessel samples with caliber below \( \text{th}_\text{onset} \) is considered a candidate focal narrowing. To be classified as definite focal narrowing it has to conform to some ophthalmologists inspired features. First of all, the region of interest must maintain a caliber reduced below the \( \frac{2}{3} \hat{d}(l) \) threshold for a substantial length: this corresponds to the notion that focal narrowing is considered clinically significant only when a caliber reduction of at least one third is present (see Fig. 8.14). Then it must be longer than a certain minimum length threshold, in order to not be confounded with caliber variation caused by noise or tracking errors. Finally, transitions from estimated normal caliber to focal narrowing and back to estimated normal caliber should be sharp compared to the vessel mean caliber. Results of such methods to a vessel with definite focal narrowing is shown in Fig. 8.15.
8. Vessel Structure Analysis

Figure 8.15:

(a) Arteriole with focal narrowing
(b) Identified FAN

<table>
<thead>
<tr>
<th></th>
<th>Focal Narrowing</th>
</tr>
</thead>
<tbody>
<tr>
<td>True Positives</td>
<td>12</td>
</tr>
<tr>
<td>False Positives</td>
<td>8</td>
</tr>
<tr>
<td>False Negatives</td>
<td>6</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.67</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.81</td>
</tr>
</tbody>
</table>

Table 8.5:

8.4.3 Results and Discussion

The results of the algorithm on the images belonging to DB60 database are summarized in Tab. 8.5. Since the evaluation of sensitivity and specificity on single lesion requires a carefully identified ground truth, it has not been possible to perform such an evaluation. The performance evaluation is therefore on the ability of the algorithm to correctly identify whole images with definite focal arteriolar narrowing.

Arteriolar narrowing has been considered definite when the ophthalmologist grading was greater than 15 (on a scale 0-100).
8. Vessel Structure Analysis

8.5 Bifurcations and Crossings Identification

The identification of both bifurcations and crossings is based on the extracted vascular structure. While crossings can be identified by a direct search of overlapping segments, bifurcations require a specific methodology, since the vessel tracking procedure described in Chap. 6 usually stops at bifurcations, so that the three converging branches are separate vessel segments.

All possible triplets of vessel extrema are considered, and their compatibility with regard to colour, convergence, diameters and directions is evaluated. Given two extrema \( p_i \) and \( p_j \),

\[
\delta_{i,j} = \sqrt{p(x_i - x_j)^2 + (y_i - y_j)^2}
\]

is the Euclidean distance between two points, \( g_i \) is the mean grey-level value in a 5x5 neighborhood around \( p_i \), \( d_i \) is the vessel diameter of the extremum \( p_i \) and \( \vartheta_i \) is its vessel direction. For three extrema \( p_i \), \( p_j \), \( p_k \), the centroid \( c_{ijk} \) is defined as the mean of the intersections of the three lines passing through the extrema and with their respective directions. Finally, a caliber compatibility is calculated as the minimum of the difference between the three possible power low relationships among the three extrema diameters:

\[
\zeta_{ijk} = \min\{ (d_i^3 - d_j^3 - d_k^3), (d_j^3 - d_i^3 - d_k^3), (d_k^3 - d_j^3 - d_i^3) \} \tag{8.34}
\]

In this way, the father trunk can be identified. The score functions to evaluate the extrema compatibility are:

\[
\Delta_{i,j,k} = e^{-\delta(i,j)+\delta(i,k)+\delta(j,k)} \tag{8.35}
\]

\[
\Theta_{i,j,k} = e^{-(2\pi - \vartheta_i - \vartheta_j + \vartheta_k)} \tag{8.36}
\]

\[
Zeta_{i,j,k} = e^{-\zeta_{ijk}} \tag{8.37}
\]

\[
G_{i,j,k} = e^{-|\varpi - \varphi| + |\varpi - \varphi| + |\varpi - \varphi|} \tag{8.38}
\]

\[
|\varpi - \varphi| \tag{8.39}
\]

Among all the possible triplets, only those that have the centroid \( c_{ijk} \) that lies inside the triangle having the three extrema as vertices are considered (see Fig. 8.16).

Among these, the one with minimum \( \Delta_{i,j,k} \) is chosen, and if the compatibility scores are below a given threshold, the triplets is accepted as bifurcation. The process is iterated until either there are no more available triplets or a minimum condition on \( \Delta_{i,j,k} \) is no more satisfied.

Care has to be taken to identify branchings, a special kind of bifurcations in which the tracking procedure has connected one branch with another, thus making the triplets search useless (see Fig. 8.17). To identify these situations, every vessel extremum not already part of a bifurcation is considered. The closest vessel point to this extremum is searched in the vessel structure.
If the distance of this vessel from the extremum is smaller than a given threshold, and the extremum direction is incident on that vessel, a bifurcation is detected.

8.5.1 Results and Discussion

In Tab. 8.6 and Tab. 8.7 the performances of the proposed bifurcation and crossing recognition algorithm are reported. The first score shown is
8. Vessel Structure Analysis

the fraction of recognized features with respect to those manually identified, then true positives and false positives are reported.

8.6 Bifurcation Analysis

Bifurcations is a feature seldom analysed, and few available data that link bifurcation morphological changes to retinopathy [14, 60], hypertension or diabetes are available, even if it seems that changes occur with age [43, 83]. Two parameters have been considered in analysing the bifurcations, that are somewhat indicative of non-optimality in the vascular tree. The first is the junction exponent $\gamma$. It has been argued that for optimal shear stress distribution across the vascular network, the diameters of the main trunk $d_0$ should be linked to those of the two branches $d_1$ and $d_2$ by power relationship:

$$d_0^\gamma = d_1^\gamma + d_2^\gamma$$  \hspace{1cm} (8.40)

For the same reasons, an increase in the branching angle value is a sign of non-optimality. Given the branches directions $\vartheta_1$ and $\vartheta_2$, the branching angle is defined as:

$$\omega = \vartheta_1 + \vartheta_2$$  \hspace{1cm} (8.41)

No ophthalmologist evaluation was available for the bifurcation analysis, but the results presented in Tab. 8.8 were consistent with those presented in [59] with regard to the bifurcation angle, while no comparison with respect to the bifurcation junction has been made since in that work no results have been given. A total of 652 bifurcations were analysed, from the images in the DB60, thus on average 10.87 bifurcations per image were analysed.

8.7 Crossings Analysis: Gunn’s and Salus’ Signs

Gunn’s sign appears as a tapering of a vein at both sides of an arteriovenous crossing. In the most severe cases the vein aspect becomes similar to a hourglass (see Fig. 8.18). A simple scheme of the Gunn’s (compression) sign is shown in Fig. 8.19, together with the measures needed to evaluate the presence and the degree of severity of the abnormality.

The simpler technique able to detect such a crossing abnormality, is based on the analysis of vein calibers close to the crossing ($C_{ic}$ and $C_{oc}$) and farther
Table 8.6: Performance of the crossings recognition. The first column is the fraction of the correctly identified crossings with respect to the manually labelled.
Table 8.7: Performance of the bifurcation recognition. The first column is the fraction of the correctly identified bifurcations with respect to the manually labelled.
Table 8.8: Median of the mean junction exponent and of the mean branching angle of each image in the DB60. Between square brackets the first and third quartiles.

<table>
<thead>
<tr>
<th></th>
<th>Junction Exponent</th>
<th>Branching Angle [rad]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Values in [59]</td>
<td>-</td>
<td>1.48 [1.26-1.71]</td>
</tr>
<tr>
<td>Values Obtained</td>
<td>1.74 [1.54-1.90]</td>
<td>1.52 [1.28-1.77]</td>
</tr>
</tbody>
</table>

Figure 8.18: The typical hourglass appearance of a severe compression sign from it ($C_{ir}$ and $C_{or}$). If significant reduction of vein caliber is found around the crossing, probable compression sign is present.

The evaluation of the severity of the compression at the $i^{th}$ crossing is done by computing the fraction of caliber reduction:

$$g_i = \frac{C_{ic} + C_{oc}}{C_{ir} + C_{or}}$$  \hspace{1cm} (8.42)

Salus’s sign is the deflection of a vein from its normal course at a arteriovenous crossing. The vein appear as S-shaped, the more sharp the S bend the more severe the sign (see Fig. ??).

The evaluation of the presence and the grading of the severity of the Salus’ (deflection) sign is based on the analysis of the vein directions around the crossing. Fig. 8.20 is a scheme representing a vessel changing course at a crossing. In order to quantitatively evaluate this change, vessel directions are computed both proximally ($d_1$) and distally ($d_2$ and $d_3$) to the crossing.
Gunn measures and examples

Figure 8.19: Gunn’s sign scheme

location. The severity of the deflection present in the $i^{th}$ crossing, is measured by the angles $a_1$ and $a_3$:

$$s_i = 0.5 \cdot (a_1 + a_3)$$  \hspace{1cm} (8.43)

Crossing directions can be estimates reliably enough to allow evaluation of Salus’ sign. The same can not be said for vessel calibers around a crossing. In fact, what appears as a venous narrowing, is usually only a fading of the vein blood column, caused by the dipping of the vein under the artery. This results in a very difficult estimation of vein calibers at the crossing, and of the vein apparent narrowing. And the more the vain fades, the more severe the Gunn’s sign but the more difficult its evaluation based only on estimated calibers.

A crossing model which adapts to image data would therefore be preferable to the actual algorithms, that do not consider the peculiarity of the crossing aspect.
8. Vessel Structure Analysis

Salus measures

Figure 8.20: Salus’ sign scheme

8.8 Conclusions

In this chapter a set of algorithm to automatically analyse the vascular network have been proposed. First of all a method to distinguish between arteries and veins. Then a method to identify from the extracted vascular network bifurcations and crossings. Finally, measures of vascular abnormalities such as tortuosity, generalized arteriolar narrowing, focal narrowing, Gunn’s and Salus’ sign have been performed and compared with the ophthalmologist grading. Even if the results of the various algorithms proposed are varied, ranging from the good correlation of the tortuosity measure to the poor one of the generalized narrowing, it is worth noting that this is the first time such a detailed analysis is automatically performed on the retinal vascular structure.
8. Vessel Structure Analysis
9

Non Vascular Abnormalities Identification

When the vessel structures and the optic disc have been eliminated from a retinal image, what remains is the retinal background and possible non vascular abnormalities, i.e. lesions such as hemorrhages, hard exudates and cotton wool spots.

Using the illumination corrected retinal images described in Chap. 5, and imposing the first two moments to the image histograms, the result is that colour content is an effective feature for the determination of the abnormal regions of the retina. Both normal fundus and lesions can only be described statistically, since their colour content is affected by noise and by an intrinsic variability of different retinas, and of different stage of evolution of the lesions. In order to cope with this variability, a two stage approach has been devised. The first stage has the aim of segmenting the candidate lesions from the normal fundus, assigning to each pixel a probability of being lesion or retinal background. This stage should be very sensitive, relaxing the requirements of specificity. The second stage, corresponding to a higher level of abstraction, classifies the candidate regions, identified in the first stage, into the lesions of interest. This is achieved through a supervised classifier, which looks at region-wide features, instead of pixel-wise. A training set containing lesions and regions of retinal background has been generated to develop the classifier.

9.1 Review of Available Methods

In the available literature, several algorithms were proposed to identify microaneurysms from fluorescein angiographic image [17, 8, 5, 1, ?]. They rely on the fact that every region containing blood will be uniformly filled
9. Non Vascular Abnormalities Identification

with contrast agent. This results in an hyperfluorescence of vessels, microaneurysms and hemorrhages. Microaneurysms detection was performed with morphological operators, Hough transform, or adaptive thresholding. In some case a second step was performed to test the identified microaneurysms with geometric features, since they are known to have a round shape. A semi-automatic approach to evaluate not only microaneurysms, but also other retinal lesions in fluorescein angiography was proposed in [19], in order to overcome the low sensitivity and specificity of the available methods for automatic lesion detection.

One of the first work in microaneurysm detection in colour fundus images is [66]. The methods is based on a shade correction using a large median filter, and then a thresholding that ensures that a user defined percentage of image pixels are classified as positive. Then a morphological operator is used to identify all round shapes among the positive classified pixels. The reported performance on 22 images are a sensitivity of 0.7 and specificity of 0.6 in the number of microaneurysms.

A similar approach was followed by [33], in which a further logit classification step is performed, yielding a sensitivity for bright lesions of 0.89 and for dark lesions of 0.78.

In [39] the same basic algorithm was described, with a larger number of features evaluated, and the classification performed by heuristic rules. The reported sensitivity is 0.85 and specificity of 0.76 in detecting patients with any retinopathy, regardless of effective presence of microaneurysms.

Top-hat morphological filtering is the core of the algorithm proposed in [98], which is followed by thresholding of extracted features. The identification of at least one microaneurysm in diseased retina is reported to be 0.9, whereas specificity is 0.8.

Others focused on exudative lesions detection. In [93] and [31] a Bayesian classification of image pixels has been described, based on a spherical colour coordinate system. It achieved a sensitivity of 1 and a specificity of 0.7, but no information is given on the separation between the training and validation set of images, thus suggesting the same images have been used for training and validation.

An adaptive thresholding coupled with a region growing procedure is described in [44], in which a reported sensitivity to exudates is 1 and the specificity is 0.71.

A fuzzy clustering approach do detect candidate exudate and an artificial neural network to refine the classification was the solution described in [67], which achieve a sensitivity of 0.92 and specificity 0.82.

In [92] the same approach used in [98] to detect microaneurysms was proposed to exudate identification, achieving a 0.92 sensitivity, but with no reported
9. Non Vascular Abnormalities Identification

specificity.
In [80] both microaneurysms and exudates are searched in retinal colour images. The first are identified with an original operator refined by an artificial neural network classification, and the sensitivity and specificity reported are 0.77 and 0.88 respectively. Exudate detection used simple thresholding with region growing algorithm, and obtained a sensitivity of 0.88 and a specificity of 0.99.

No reported work has tried to identify a set of lesions larger than microaneurysms and hard exudates. The only mention to classification of hemorrhages, microaneurysms, exudates and cotton wool spots is in [23], where different classification strategies are compared on manually segmented lesions.

9.2 Bayesian Classification of Suspect Areas

This is the first stage of the lesion identification procedure. In order to provide a high sensitivity to the lesions of interest (classes), the fundus image can be divided into three types of areas. Those that appear darker than the normal fundus $\omega_1$, which correspond to possible hemorrhages and microaneurysms, those that appear brighter than the retinal background $\omega_3$, which correspond to possible hard exudates and cotton wool spots, and those belonging to the normal fundus $\omega_2$. The colour content of each pixel of the image is assumed to be sufficient to identify the three classes. Given the image $I$, the colour content of the pixel at position $(x,y)$ can be described by the vector $f$ containing the intensity values of the three colour channel:

$$f(x,y) = [I_{\text{red}}(x,y), I_{\text{green}}(x,y), I_{\text{blue}}(x,y)]$$ (9.1)

Given the a priori probability $P(\omega_i)$, $i = 1, 2, 3$, of the three types of regions, and the conditional probability densities $p(f|\omega_i)$, describing the statistical colour distribution for the three classes, the probability that a pixel $I(x,y)$ belongs to the class $\omega_i$ can be computed by means of the Bayes theorem:

$$P(\omega_i|f) = \frac{p(f|\omega_i) \cdot P(\omega_i)}{p(f)} \quad i = 1, 2, 3$$ (9.2)

$$p(f) = \sum_{i=1}^{3} (p(f|\omega_i) \cdot P(\omega_i))$$ (9.3)

This formulation allows the straightforward description of a space variant a priori probability. For each pixel position $(x,y)$, the probability will be $P(\omega_i, x, y)$. Previously identified structures, as vessels and optic disc, can then be included in (9.2) through the a priori probability, forcing the pix-
els belonging to them to be considered normal fundus, excluding them from possible abnormal areas.

An even more powerful utilization of this space variant probability can be envisaged. Since cotton wool spots are more likely to appear along the main vascular arcades, a probability image that gives a higher probability of being a bright lesions to pixels in the surroundings of those arcades might be computed. At the same time, since macular exudates are signs of a severe and sight-threatening retinopathy, a higher sensitivity to macular bright lesions could be set by modifying \( P(\omega, x, y) \) for pixels in the macula.

After the three conditional \textit{a posteriori} probabilities \( P(f|\omega_i) \) are evaluated, a pixel is assigned to one of the three classes with a maximum a posteriori (MAP) procedure. The maximum of the three probabilities leads the assignment of a pixel to a class, but a pixel is assigned with its probability.

A set of ten images, not belonging to the DB60 nor to the DB200, have been used to estimates the probabilities and the probability densities needed to evaluate the \textit{a posteriori} probabilities of (9.2). The lesions present in these images were manually segmented from the illumination corrected images described in Chap. 5. The intensity values of the three colour channels of the identified pixels have been pooled to estimate \( p(f|\omega_i) \). The colour probability density for each class has been modelled with a multivariate mixture of Gaussian. Normal fundus have shown a monomodal density, thus it is described by a single Gaussian, whereas both dark and bright classes shew a bimodal density, thus modelled with the combination of two Gaussian curves. Since significant correlation was found among the colour channels for the three classes, a full covariance matrix is used in the gaussian mixture model. The values of the model parameters, means \( \mu \) and covariance matrices \( \Sigma \), are reported in Tab. 9.1. The a priori probabilities \( P(\omega_i) \) have been empirically set. Since lesions are uncommon but in the more severe stages of retinopathy, the occurrence (relative frequency) of lesion pixels in the images would be excessively low, providing an unsatisfactory sensitivity:

\[
\begin{align*}
P(\omega_1) &= 0.01 \\
P(\omega_2) &= 0.98 \\
P(\omega_3) &= 0.01
\end{align*}
\]

In Fig. 9.1 three slices along the blue component plane of the probability density functions are shown, for increasing values of the blue component.
9. Non Vascular Abnormalities Identification

Dark Areas

<table>
<thead>
<tr>
<th></th>
<th>Normal Fundus</th>
<th>Bright Areas</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mode 1</td>
<td>Mode 2</td>
</tr>
<tr>
<td>$\mu_{red}$</td>
<td>0.54</td>
<td>0.55</td>
</tr>
<tr>
<td>$\mu_{green}$</td>
<td>0.38</td>
<td>0.49</td>
</tr>
<tr>
<td>$\mu_{blue}$</td>
<td>0.41</td>
<td>0.52</td>
</tr>
<tr>
<td>$\Sigma(1, 1)$</td>
<td>$17 \cdot 10^{-4}$</td>
<td>$92 \cdot 10^{-4}$</td>
</tr>
<tr>
<td>$\Sigma(2, 1)$</td>
<td>$8 \cdot 10^{-4}$</td>
<td>$6 \cdot 10^{-4}$</td>
</tr>
<tr>
<td>$\Sigma(2, 2)$</td>
<td>$23 \cdot 10^{-4}$</td>
<td>$28 \cdot 10^{-4}$</td>
</tr>
<tr>
<td>$\Sigma(3, 1)$</td>
<td>$3 \cdot 10^{-4}$</td>
<td>$-9 \cdot 10^{-4}$</td>
</tr>
<tr>
<td>$\Sigma(3, 2)$</td>
<td>$22 \cdot 10^{-4}$</td>
<td>$12 \cdot 10^{-4}$</td>
</tr>
<tr>
<td>$\Sigma(3, 3)$</td>
<td>$31 \cdot 10^{-4}$</td>
<td>$94 \cdot 10^{-4}$</td>
</tr>
</tbody>
</table>

Table 9.1:

9.3 Blobs Analysis and Features Extraction

In the previous section, a pixel-wise classification to perform a segmentation of candidate dark-lesions pixels and bright/lesions pixels from the normal retinal fundus has been described. The result of this procedure is a sparse set of pixels for the two classes of interest, that not convey any information on spatial relationship or global object features. Linking single pixels to the lesion they belong to would not be an issue if the proposed algorithm were 100% specific. However, since pixel-wise classification is not very specific, a refinement is needed: this is done by introducing a further level of abstraction. Every set of connected pixels with the same classification represent a separate object, that is a candidate lesion.

Small set of pixels classified as candidate lesions were excluded from the subsequent analysis. The ratio for doing this is that isolated pixels are more likely to be noise overlapped to normal fundus, and then misclassified, rather than tiny lesions. A pair of cascaded morphological operators well suit the task, providing both the removal of isolated areas of dimension comparable to the morphological kernel, and also the filling of small holes in connected regions. After the pixel classification has been cleaned by this filtering, all connected components are extracted separately, and each represent a candidate lesion.

In order to distinguish the different lesions, and the lesions from false positives, is necessary to provide a description of every identified region, to provide the most discriminating factors. This is not a trivial task, given the extreme variability in appearance of the same type of lesions (intra-class variability), and the close similarity of lesions to variations in pigmentation or to other lesions (inter-class variability).
9. Non Vascular Abnormalities Identification

The regions can be described roughly by means of four types of features. The first are chromatic features, which statistically describe the colour aspect of the region. The most important are those related to the first two moments of the vector of the colour intensities. To enhance the chromatic information, the HSV colour space in addition to the RGB available intensities has been used. This is supposed to be the most discriminant space for perceptually different colours [72]. This space is a non-linear transformation of the RGB space, which should eliminate the redundancy of this description, and provide a space in which the three channels are almost uncorrelated:

\[
V = \frac{R+G+B}{3} \\
H = \frac{1}{2\pi} \arccos \left( \frac{0.5 \cdot (2R-G-B)}{\sqrt{(R-G)^2 + (R-B)(G-B)}} \right) \\
S = 1 - \frac{3}{R+G+B} \cdot \min(R, G, B)
\]  

where red, green and blue value ranges is \([0, 1]\), and the same is for hue, saturation and value.

The second set of features provides a geometrical characterization of the region. It evaluates both perimeter and area features, in order to evaluate the compactness, but also aspect ratios, principal axes, eccentricity. Also the standard deviation of the radius is evaluated: given the region centroid, the variation of the distance between the region border and the centroid for every perimeter point, should give a measure of the smoothness of the region borders and of its roundness.

Then, on one hand the region should have a definite differentiation from the neighborhood, on the other the characteristics of features on the border may provide a hint on the particular lesion to be classified. Three areas are defined in each candidate lesion: the internal part \(B_i\), the internal border \(\partial B_i\) and the external border \(\partial B_o\), and they are shown in Fig. 9.2. They are easily defined using the morphological operations of dilation and erosion, with a kernel of suitable dimension. \(B_i\) is simply the eroded version of the original region \(B\), the internal border is what remains by subtracting form \(B\) the internal part \(B_i\), and the external border is what remains after the subtraction of \(B\) from its dilation.

The fourth set of features is composed by colour gradient of the regions. The gradients are computed by looking at intensity variation between growing region perimeters, starting from the centroid. Defining a sequence of boundaries by iteratively removing all perimeter points from the region, a sequence of vectors \(\{\mathbf{n}_1 \ldots \mathbf{n}_j\}\) normal to each boundary is computed. The defined gradient is therefore the intensity variation along the normal direction from the region centroid to its boundary. Sharp margins and soft margins
9. Non Vascular Abnormalities Identification

are likely to be identified via a combination of internal border and external border features.

9.4 Feature Selection

There are several reasons that suggest to keep as low as possible the number of features with which classify the candidate regions. The first is computational complexity. Another is that the increase in complexity in computing a larger number of features is not always matched by an increase in discriminatory power, because of the possible redundancy and correlation among features. The most important reasons lie however in the generalization power of a classifier. Since the number of features can be considered as the number of free parameters of the classifier, the smaller the ratio between this number and the cardinality of the training set, the best the performance on the training set but the less robust the classifier.

Given the large number of features that can be computed in the framework of the previous section, there is the need to select the most significant ones, so as to maintain the number of features used by the classifier small, and at the same time keep the discriminating power elevated.

In order to perform this feature selection, it is necessary to have a measure of the discrimination ability of a set of features. The most common measures involve the evaluation of the intra-class scatter matrix $S_w$ and the inter-class scatter matrix $S_b$. Given $M$ classes, a set of features vectors $\mathbf{x}$, each one belonging to one and only one class, the intra-class scatter matrix is:

$$ S_w = \sum_{i=1}^{M} P(\omega_i) \Sigma_i $$  \hspace{1cm} (9.9)

$P(\omega_i)$ is the a priori probability of the class $\omega_i$, and $\Sigma_i$ is its covariance matrix:

$$ \Sigma_i = E \left[ (\mathbf{x} - \mu_i)(\mathbf{x} - \mu_i)^T \right] $$  \hspace{1cm} (9.10)

The inter-class scatter matrix is:

$$ S_b = \sum_{i=1}^{M} P(\omega_i)(\mu_i - \mu_0)(\mu_i - \mu_0)^T $$  \hspace{1cm} (9.11)
9. Non Vascular Abnormalities Identification

Figure 9.1: Slices of the probability density functions of the three classes of interest
9. Non Vascular Abnormalities Identification

(a) A candidate region
(b) A region divided into its internal part $B_i$, internal border $\partial B_i$, and external border $\partial B_o$.

Figure 9.2:

(a) Original region (b) Dilated Region (c) Border Sequence

Figure 9.3:
9. Non Vascular Abnormalities Identification

with \( \mu_0 \) being the global mean:

\[
\mu_0 = \sum_{i=1}^{M} P(\omega_i) \mu_i
\]  

(9.12)

Various combination of these two matrices have been proposed to evaluate at the same time the closeness of samples coming from the same class, and the separation of the classes. Among the proposed criteria:

\[
J_1 = \text{trace}(S^{-1}wS_b)
\]  

(9.13)

\[
J_1 = \ln(\det(S^{-1}wS_b))
\]  

(9.14)

\[
J_3 = \frac{\text{trace}(S_b)}{\text{trace}(S_w)}
\]  

(9.15)

The main disadvantage of this criteria is that they do not have any relationship with the Bayesian classification error. On the contrary, the probability of error in classifying a point between two classes \( \omega_1 \) and \( \omega_2 \) is bounded by the Chernoff bound:

\[
\varepsilon \leq P(\omega_1)^{1-s}P(\omega_2)^s \int p(x|\omega_1)^{1-s}p(x|\omega_2)^s dx
\]  

(9.16)

\[
s \in [0, 1]
\]  

(9.17)

The optimum value of \( s \), which gives the lowest bound, satisfies:

\[
\frac{d\mu}{ds} = \ln \left( \frac{P(\omega_1)}{P(\omega_2)} \right)
\]  

(9.18)

\[
\mu(s) = -\ln \left( \int p(x|\omega_1)^{1-s}p(x|\omega_2)^s dx \right)
\]  

(9.19)

The optimum value of \( s \), which provides the minimum error, is not easy to find, so that the value \( s = 0.5 \) is often chosen to provide an upper bound, even if not the lowest. In this case, the value \( \mu(\frac{1}{2}) \) is called the Bhattacharayya distance, and it can be used as a criterion for class separability. For two normal distribution with mean \( \mu_i \) and covariance \( \Sigma_i, i = 1, 2 \), the distance can be evaluated in closed form:

\[
\mu(\frac{1}{2}) = \frac{1}{8}(\mu_1-\mu_2)^t \left( \frac{\Sigma_1 + \Sigma_2}{2} \right)^{-1} (\mu_1-\mu_2) + \frac{1}{2} \ln \left( \frac{0.5(|\text{Sigma}_1 + |\text{Sigma}_2|)|}{|\Sigma_1|^{1/2} + |\Sigma_2|^{1/2}} \right)
\]  

(9.20)

This distance is used to iteratively pick the most discriminating feature among in the set of available features, given the features already chosen.
In this way 29 features have been chosen to build the classifier for the bright classes and 30 for the classification of dark lesions.

9.5 Linear Classifier

The classifier that is build from the training set and the chosen features is based on the Linear Discriminant Analysis (LDA). Given a certain numbers of classes with supposedly different characteristics, LDA is a method for linearly mapping the high dimensional characteristics vector in a lower dimensional space, which maximize the separation between classes, supposing their distribution normal. LDA is based on the maximization of a function \( J \) that is an indicator of the class separation. Given \( N \) samples, \( M \) classes and a \( 1 \times m \) vector \( f \) of features, the function \( J \) considered is:

\[
J_1 = \text{trace}(S_w^{-1}S_b) \tag{9.21}
\]

with \( S_w \) and \( S_b \) defined by (9.11) and (9.9) respectively. The \( m \times m \) matrix \( S_w^{-1}S_b \) has rank \( M - 1 \). A linear transformation mapping the original \( m \)-dimensional features space into a new \((M1)\)-dimensional space, can therefore yields the same value for \( J \) while obtaining a lower dimensionality. This problem could be viewed as an eigenvalue problem: the matrix describing the linear transformation is in fact the matrix \( C \) having on its columns the \((M-1)\) non trivial eigenvectors of \( S_w^{-1}S_b \).

9.6 Blob Classification

In the case of lesion classification, both dark and bright lesions may belong to one of three classes to be distinguished. Bright lesions have to be separated in hard exudates, cotton wool spots and false positives. Dark lesions have to be distinguished in hemorrhages, vessels and false positives. Vessel have been incorporated in the classification of dark region separately from false positives, because of their features, separate from that of retinal background: consider them in a unique class together with the false positives would have resulted in poor classification performances.

Given the vector \( f_{ij} \) of the features belonging to blob \( j \) classified as class \( i \), the transformed features vector is obtained through the linear transformation:

\[
y_{ij} = Cf_{ij} \tag{9.22}
\]
9. Non Vascular Abnormalities Identification

<table>
<thead>
<tr>
<th></th>
<th>Hard Exudates</th>
<th>Cotton Wool Spots</th>
<th>Hemorrhages</th>
</tr>
</thead>
<tbody>
<tr>
<td>True Positives</td>
<td>5</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>False Positives</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>False Negatives</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.83</td>
<td>0.66</td>
<td>0.6</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
</tr>
</tbody>
</table>

Table 9.2: Results for identification of images with non vascular lesions in DB60

For each class $\omega_i$ of numerosity $N_i$, the sample mean and covariance matrix, $\mu_i$ and $\Sigma_i$, are evaluated from the transformed features of the training set:

$$\mu_i = \frac{1}{N_i} \sum_{j=1}^{N_i} y_{ij}$$

(9.23)

$$\Sigma_i = \frac{1}{N_i-1} \sum_{j=1}^{N_i} (y_{ij} - \mu_i)' (y_{ij} - \mu_i)$$

(9.24)

These means and covariances are used to evaluate the Mahalanobis distance from each class $\omega_i$ of the transformed features vector $\mathbf{y}$ of an unknown region:

$$d_i = (\mathbf{y} - \mu_i)' \Sigma_i^{-1} (\mathbf{y} - \mu_i)$$

(9.25)

The region has the highest probability of belonging to the class with the minimum distance, and the probability is proportional to the distance $d_i$

9.7 Performance Evaluation

<table>
<thead>
<tr>
<th></th>
<th>Hard Exudates</th>
<th>Cotton Wool Spots</th>
<th>Hemorrhages</th>
</tr>
</thead>
<tbody>
<tr>
<td>True Positives</td>
<td>22</td>
<td>15</td>
<td>22</td>
</tr>
<tr>
<td>False Positives</td>
<td>1</td>
<td>17</td>
<td>1</td>
</tr>
<tr>
<td>False Negatives</td>
<td>7</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.75</td>
<td>0.68</td>
<td>0.71</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.99</td>
<td>0.90</td>
<td>0.99</td>
</tr>
</tbody>
</table>

Table 9.3: Results for identification of images with non vascular lesions in DB200

The results of the algorithm on the images belonging to DB60 database are summarized in Tab. 9.2 and in Tab. 9.3 for those belonging to DB200. Since the evaluation of sensitivity and specificity on single lesion requires a carefully identified ground truth, it has not been possible to perform such an evaluation. The performance evaluation is therefore on the ability of the
9. Non Vascular Abnormalities Identification

<table>
<thead>
<tr>
<th></th>
<th>Hard Exudates</th>
<th>cotton Wool Spots</th>
<th>Hemorrhages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation</td>
<td>0.87</td>
<td>0.78</td>
<td>0.83</td>
</tr>
</tbody>
</table>

Table 9.4: Correlation between the estimated area covered by lesions and the ophthalmologist grading, on the images of the DB60

algorithm to correctly identify whole images with particular lesions. Nevertheless, the ability of the algorithm to estimate the severity of the non-vascular lesions can be evaluated by looking at the correlation between the estimated area of the retina covered by specific lesion, and the related ophthalmologist grading. This can be quantitatively assessed with the measure of correlation, reported in Tab. 9.4, and visually evaluated in Fig. 9.4, Fig. 9.6 and Fig. 9.5

Figure 9.4: Total area of the identified hemorrhages versus ophthalmologist grading

9.8 Conclusions

In this chapter a method for identifying and distinguish between retinopathy lesions has been presented. Its simplicity makes it appealing when very fast computation is required. Its low sensitivity for the identification of images
Figure 9.5: Total area of the identified hard exudates versus ophthalmologist grading

with hemorrhages poses some problems about its reliability in screening for retinopathy, especially for diabetic retinopathy, where microaneurysms are the first lesion to appear. The major weakness of the presented algorithm is the first Bayesian classification, which is not sensitive enough to segment from the normal fundus all hemorrhages and cotton wool. This is sometimes due to their faint appearance, so that the Bayesian classifier is not able to find a definite connected region of candidate lesion pixels, and then the sparse pixels are discarded.
Figure 9.6: Total area of the identified cotton wool spots versus ophthalmologist grading
9. Non Vascular Abnormalities Identification
10

Hypertensive Retinopathy Grading

In Chap. 8 and Chap. 9 a set of techniques able to quantify vascular changes and detect lesions as hemorrhages, microaneurysms, hard exudates and cotton wool spots has been described. The evaluation of the extent and severity of a single lesion or abnormality is only the first step to provide a comprehensive clinical profile of the retina and its pathology. The second step deals with gathering all the information from homogenous measurements (e.g. the tortuosity of the identified vessels in an image) into a measure for the whole retina. Finally, all the image-wide measurements of abnormalities have to be combined to yield a single retinopathy level. In this chapter an approach to the last two steps will be described, bearing in mind that is not trivial a task, since involves translating the implicit knowledge of the ophthalmologists judgement into an explicit knowledge.

10.1 From Sign Evaluation to Image Evaluation

10.1.1 Tortuosity

When several vessels with different tortuosities are present in a retinal image, it is not clear even to an experienced ophthalmologist how every single vessel appearance contributes to the overall perception of retinal tortuosity. It is worth noting two additional aspects of evaluating retinal tortuosity, as opposed to vessel tortuosity. The first is that it is commonly accepted that arterial an venous tortuosity should be evaluated separately. The second is that tortuosity has a different course with the disease for the two vessel
10. Hypertensive Retinopathy Grading

networks. As hypertension increases, arteries are the first vessel to become tortuous, followed only at the more severe stages of hypertension by veins. And more, as arteries become sclerotic, and the vessel wall is damaged by sustained hypertension, it happens that they become stiff and straight, while venous hypertension persists. At this stage, a match between the ophthalmologist evaluation of tortuosity and a image-wide automatic evaluation of tortuosity, separately for arteries and veins, will be proposed, regardless of the retinopathy level it is linked to. Given a set of vessel tortuosity measures $T = \{\tau_1 \ldots \tau_n\}$, the idea is to consider as representative measure its median. Since, given the same median, few vessels with high tortuosity make human perception bend toward a more severe evaluation of tortuosity, and the same happens when a high bottom-line tortuosity is present in a vessel network, the evaluation of the retinal tortuosity has to incorporate these deviation from the median. The proposed image arterial tortuosity index is thus:

$$\tau_{\text{network}} = \frac{p(T, 50) + \beta p(T, 75) + \gamma p(T, 25)}{1 + \beta + \gamma} \quad (10.1)$$

in which $p(S, n)$ is the $n^{th}$ percentile of the set $S$, and $\beta$ and $\gamma$ are heuristic parameters.

For veins, it seems that more weight is given on the largest veins. At the same time, the lowest value of vessel tortuosity set an flashlight for tortuosity severity grading. In order to do this, each vessel tortuosity $\tau_i$ has been multiplied by the correspondent vessel mean calibre $d_i$, normalized by the maximum vessel calibre of the retina $d_{\text{max}}$.

$$T = \left\{ \tau_1 \frac{d_1}{d_{\text{max}}} \ldots \tau_n \frac{d_n}{d_{\text{max}}} \right\} \quad (10.2)$$

The proposed image venous tortuosity index is thus:

$$\tau_{\text{network}} = \frac{p(T, 75)p(T, 25) + \beta p(T, 10)p(T, 90)}{1 + \beta} \quad (10.3)$$

When evaluating tortuosity of the whole vessel network, regardless of each vessel being an artery or a vein, the behaviour seems similar to that of the arterial network. The tortuosity measure for the whole network was computed with (10.1), with parameters set to $\beta = 0.75$ and $\gamma = 0.73$. In Fig. ?? is shown the plot of the automatic versus tortuosity computed
10. Hypertensive Retinopathy Grading

automatically and the mean of the arterial and venous tortuosity evaluated by an experienced ophthalmologist.

<table>
<thead>
<tr>
<th></th>
<th>Arterial Network</th>
<th>Venous Network</th>
<th>Mixed Network</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation</td>
<td>0.718</td>
<td>0.650</td>
<td>0.710</td>
</tr>
</tbody>
</table>

Table 10.1: Tortuosity grading performance with respect to the ophthalmologist grading

10.1.2 Generalized Arteriolar Narrowing

Generalized arteriolar narrowing is altogether a single score for every image, so that no further analysis is needed to obtain an image-wide severity index for this feature.

The severity is evaluated with one of the indexes described in Sec. 8.3.

10.1.3 Focal Arteriolar Narrowing

10.1.4 Gunn’s and Salus’ Signs

Crossing abnormalities evaluation poses the serious problem of mixing a present/not present type of grading with a continuous severity evaluation of the single abnormal event.

In fact the presence of even a single crossing abnormalities triggers an high ophthalmologist’s grading, but it is still not clear how multiple events concur to the final judgement.

The measure proposed in this thesis and used in evaluation image-wide the severity of crossing abnormalities (separately for compression and deflection signs), considers only single events measures above an empirical threshold, and then combines the measures weighting the mean value with the more severe event found:

\[
\gamma_{\text{gunn}} = \frac{0.5}{N} \sum_{i=1}^{N} (g_i) + 0.5 \cdot \max(g_i, i = 1 \ldots N) \\
\gamma_{\text{salus}} = \frac{0.5}{N} \sum_{i=1}^{N} (s_i) + 0.5 \cdot \max(s_i, i = 1 \ldots N)
\]

10.1.5 Non Vascular Lesions

Hemorrhages, hard exudates and cotton wool spots severity is evaluated based on the area of the retina covered by each lesion.
Figure 10.1: Image tortuosity for the venous network on the DB60. The parameter used is $\beta = 0.52$. In black the least-squares line that interpolates the automatic measures with respect to the clinician grading...
Figure 10.2: Image tortuosity for the entire vessel network on the DB60. The parameter used is $\beta = 0.75$ and $\gamma = 0.73$. In black the least-squares line that interpolates the automatic measures with respect to the clinician grading.
Even if weighting lesions by their distance from the macula would result in a measure more sensitive with respect to sight-threatening retinopathy, there are two problems that still need to be addressed to this regard. Firstly, there is no clear evidence that the severity of retinopathy is judged differently by the ophthalmologist if the same area of, e.g., hard exudates cover a region close to the macula, to the papilla, to the main arcades or elsewhere. Secondly, in order to quantify automatically the severity, there is the need to know where the macula is, and therefore if the eye is left or right, and no robust way has been devised to perform this task.

### 10.2 A Simple Explicit Network

The main characteristic of the hypertensive retinopathy grading according to the Keith-Wegener-Barker scale, is that vessel abnormalities define the severity of the retinopathy in the first two stages, corresponding to nonmalignant hypertensive retinopathy, whereas they are less and less used to modulate the retinopathy severity when non-vascular lesions are present. Thus, two separate indexes can be defined to globally evaluate a retina with regards to vascular abnormalities and to non-vascular lesions. They are a weighted mean of the severity measures of each lesion found:

\[
\gamma_{\text{vascular}} = \frac{1}{\sum_i w_i} \sum_i w_i \gamma_i \\
\gamma_{\text{non vascular}} = \frac{1}{w_1 + w_2 + w_3} (w_1 A_{\text{hem}} + w_2 A_{\text{he}} + w_3 A_{\text{cws}})
\]

(10.9)

where \( \gamma_i \) are the measures of arterial and venous tortuosity, focal arteriolar narrowing, generalized arteriolar narrowing, Gunn’s sign, and Salus’ sign, whereas \( A \) represents the area in \( \mu m \) of the subscripted lesion.

In order to have meaningful results, it is necessary that all measures are normalized into the same dynamic range. For the vascular abnormalities measures, this is done by means of a sigmoidal function (see Fig. 10.3) with dynamic range between 0 and 1. In this way it is possible to perform at the same time both the normalization and a rejection of the not significant measures, which are therefore regarded as zeros.

On the other hand, a slightly different approach has been used for non-vascular lesions. Since they need to drive the grading above the level II on a four grade scale, or above level 50 in a percentage grading scale, the sigmoidal function should have dynamic range \([0.5, 1]\). To avoid the discontinuity of the step between 0 and 0.5, a combination of two sigmoids have been used.
10. Hypertensive Retinopathy Grading

Figure 10.3: Example of the sigmoidal function used to normalize each vessel abnormality measure

One is very fast and provides a control on the onset of the non vascular lesions index, the second is smoother and modulate the severity level once significant lesions have been found. An example of the final normalization function for the nonvascular index is shown in Fig. 10.4.

Normalized severity measures for hemorrhages, hard exudates and cotton wool spots are shown in Fig. 10.5, Fig. 10.6 and Fig. 10.7. By a careful setting of the parameters of the two sigmoids, questionable lesions can be considered in the evaluation, but not weighting as much as definite lesions, therefore yielding normalized values below the 0.5 value that represent the onset of malignant retinopathy.

Once the two separate indexes (10.7) and (10.8) have been calculated, they are combined in a way that as long as the non vascular signs are absent, the grading is based on the vascular index alone, and when non vascular
10. Hypertensive Retinopathy Grading

Figure 10.4: Example of the combination of two sigmoidal functions used to normalize each non vessel lesion area

Lesions are present, grading is based mostly on the non vascular index:

\[ \gamma_{\text{retinopathy}} = (1 - \gamma_{\text{nonvascular}}) \cdot \gamma_{\text{vascular}} + \gamma_{\text{nonvascular}} \]  \hspace{1cm} (10.10)
Figure 10.5: Total estimated area covered by hemorrhages in the DB60 images versus the ophthalmologist grading
Figure 10.6: Total estimated area covered by hard exudates in the DB60 images versus the ophthalmologist grading
Figure 10.7: Total estimated area covered by cotton wool spots in the DB60 images versus the ophthalmologist grading.
10. Hypertensive Retinopathy Grading

10.3 Performance Evaluation

Using equation (10.10), the system was able to achieve a correlation of 0.75 with the ophthalmologist grading. In Fig 10.8 is shown the automatic grading versus the ophthalmologist grading, and is clear a remarkably good agreement between the two.

The main issue are in the few cases in which the system was not able to identify significant non vascular lesions, grading the image between 20 and

<table>
<thead>
<tr>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.75</td>
</tr>
</tbody>
</table>

Table 10.2: Grading performance of the automatic system with respect to the ophthalmologist grading

![Graph showing automatic grading versus ophthalmologist grading]

Figure 10.8: Automatic grading versus the ophthalmologist grading. The interpolation line between the two gradings is also shown.

identify significant non vascular lesions, grading the image between 20 and
10. Hypertensive Retinopathy Grading

40, instead of the 60-85 level graded by the ophthalmologist. One of the image is shown in Fig. 10.9, in which small hard exudates are present in the peripapillary region: since both the retinal background and the hard exudates appears yellowish and bright in that region, the system fails to identify the lesions.

Figure 10.9: One of the images in which the system fails to identify non vascular lesions. The hard exudates in the peripapillary zone have a dull appearance and low contrast with the background.
10. Hypertensive Retinopathy Grading
11

Conclusions

11.1 Achieving the Objectives

11.2 Contribution to Knowledge

11.3 The way ahead

11.3.1 Diabetic Retinopathy Grading

11.3.2 Possible Applications to other Retinopathies

11.3.3 Applications to other Protocols
11. Conclusions
Bibliography


[3] ARIC, ARIC protocol 14A retinal photography, ARIC Coordinating Center, Department of Biostatistics (CSCC), University of North Carolina, August 1995.

[4] _____, ARIC protocol 14B retinal reading protocol, ARIC Coordinating Center, Department of Biostatistics (CSCC), University of North Carolina, May 1996.


BIBLIOGRAPHY


[56] J. Lowell, A. Hunter, and D. Steel, *Quantitative measurement of retinal vessel widths from fundus images based on 2D modelling*, tech − report6,6,2, Department of Computer Science, University of Durham, South Road, Durham, DH1 3LE, UK, June 2002.
BIBLIOGRAPHY


BIBLIOGRAPHY


